

AD744204

*Preliminary Reports, Memoranda
and Technical Notes of the ARPA
Materials Summer Conference*

Woods Hole, Massachusetts

VOLUME I

July 1971

Sponsored by
Advanced Research Projects Agency
ARPA Order No. 0236



Department of Materials and Metallurgical Engineering

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



715

EC

DISCLAIMER NOTICE

THIS DOCUMENT IS THE BEST
QUALITY AVAILABLE.

COPY FURNISHED CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

PRELIMINARY REPORTS, MEMORANDA AND TECHNICAL NOTES
of the
ARPA MATERIALS SUMMER CONFERENCE
Woods Hole, Massachusetts
July 1971

ARPA Order Number: 0236
Program Code Number: 1D10
Contractor: The Regents of The University of Michigan
Effective Date of Contract: 1 May 71
Contract Expiration Date: 30 Apr 72
Amount of Contract: \$254,000
Contract Number: DAHC15-71-C-0253
Principal Investigator: Professor Edward E. Huckle
Department of Materials & Metallurgical
Engineering
The University of Michigan
Ann Arbor, Michigan 48104
(313) 764-3302

See AD-744206

7

VOLUME I

TABLE OF CONTENTS

- I. Foreword
- II. Steering Committee
- III. Participants
- IV. Consultants
- V. Preliminary Reports, Memoranda and Technical Notes

The following papers fall into two categories; (1) papers in a state ready for publication, and (2) reports and memoranda for limited distribution representing work in progress. The former category is available for general distribution and in some cases are in the process of publication in the appropriate technical journals. The limited distribution reports and memoranda represent initial ideas, problem suggestions, position papers, and status reports and are aimed primarily to stimulate discussion with the Council. However, they are available subject to the author's release by request to the Project Director. Titles marked with an asterisk are reports that are being published.

<u>TITLE</u>	<u>PAGE</u>
*Gradients in Polymeric Materials M. B. Bever and M. Shen.	1
Effect of Strain on the Fermi Energy R. Gomer and R. M. Thomson	15
Note on Density Determination of Amorphous Materials R. Gomer	29
Modified Null-Flux Magnetic Suspension and Propulsion System for High-Speed Transportation M. Tinkham and P. L. Richards.	33
*Magnetic Suspension and Propulsion Systems for High-Speed Transportation M. Tinkham and P. L. Richards.	73
Conservation Laws and Energy Release Rates B. Udiansky and J. Rice	123

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

II

<u>TITLE</u>	<u>PAGE</u>
Application of a Defect Model to Vitreous Solids R. A. Huggins.	135
Thermodynamic Properties of Liquid Metals J. L. Margrave	161
Stress Corrosion Cracking in Plastic Solids Including the Role of Hydrogen J. J. Gilman	181
A Unified Theory for the Free Energy of Inhomogeneous Systems L. A. Swanger, G. M. Pound and J. P. Hirth	205
Theory of Ionic Transport in Crystallographic Tunnels W. H. Flygare and R. A. Huggins.	235
Effect of Stress on Electrochemical Dissolution R. Gomer and R. M. Thomson	249
Amorphous Metallic Alloys P. E. Duwez.	257
Diffusion Through Anisotropic Polymer Systems J. D. Ferry.	277
Remarks on Montroll's Nonlinear Wave Equation G. H. Vineyard	285
Random Close Packing of Spheres G. H. Vineyard	291
Note on IR Windows N. Bloembergen	299
Line Tension on Kinks on Fracture Cracks J. J. Gilman, J. P. Hirth and R. M. Thomson.	311
States of Ease of Polymeric Entanglement Networks Crosslinked in Strained States J. D. Ferry.	327
Diffusion Through Composite Polymer Systems J. D. Ferry.	341
Some Problems in Bulk Polymeric Systems H. Reiss	349

<u>TITLE</u>	<u>PAGE</u>
A Note on the Ground State Energy of an Assembly of Interacting Electrons A. Ishihara and E. W. Montroll.	413
A Rigid-Plastic Model of Spall Fracture by Hole Growth F. A. McClintock	437
SD Effect D. C. Drucker.	455
Energy to Create New Surface During Crack Propagation D. C. Drucker.	459
On the Characteristics of Gradient Materials M. B. Bever.	463
Analysis of Stress Intensity Factors in a Plate with Any Given Distribution of Cracks: A Translation M. Ishida.	483
Edge Dislocation Arrays Around a Crack Under Tension F. A. McClintock	515
High-Temperature Stability of Silicon Nitride J. L. Wood, G. P. Adams and J. L. Margrave	531
*The Use of Levitation in Inorganic Synthesis J. L. Margrave, J. A. Treverton and P. W. Wilson	533
High-Temperature Properties of Nb and Zr D. W. Bonnell, J. L. Margrave and A. J. Valerga.	545
*Pyrolysis of Polymers and Simple Organic Molecules J. L. Margrave	557
Emissivities of Liquid Metals at Their Fusion Temperatures D. W. Bonnell, J. A. Treverton, A. J. Valerga and J. L. Margrave	559
Heats of Formation of Various Types of Carbon and Graphite by Combustion Calorimetry J. L. Wood and J. L. Margrave.	571
The Li/CFX Battery and Its Characteristics R. B. Badachhape, J. L. Wood, A. J. Valerga and J. L. Margrave	573

<u>TITLE</u>	<u>PAGE</u>
Stress Waves Due to a Short Duration Pressure Pulse on a Semi-Infinite Body of Layered Composite J. A. Krumhansl and E. H. Lee.	575
Determination of Stress Profiles for Waves in Periodic Composites L. Bevilacqua, W. Kohn, J. A. Krumhansl and E. H. Lee.	587
A Proposed Method for the Evaluation of the Thermo- dynamic Properties of the Glassy Carbon-Graphite Equilibrium E. E. Hucke and S. K. Das.	589
Flow Via Dislocations in Ideal Glasses J. J. Gilman	649
*Hardness - A Strength Microprobe J. J. Gilman	675

	<u>PAGE</u>
Abstract by J. P. Hirth and F. A. McClintock.	iv
I. Summary Report by H. H. Johnson, A. J. Sedriks and M. Cohen	1
II. Comments on Adsorption-Sensitive Cracking by A. R. C. Westwood and R. M. Latanision	62
III. Unit Processes in Stress-Corrosion Cracking by R. W. Staehle	104
IV. Defect Physics in Stress-Corrosion Cracking by R. Thomson and J. P. Hirth.	182
V. Fracture Mechanics for Stress-Corrosion Cracking by F. A. McClintock and J. R. Rice.	202
VI. Boundary Segregation Effects in Environmental Degradation by J. H. Westbrook	214
VII. Fractography of Stress-Corrosion Failures-- Mechanistic Aspects by E. N. Pugh.	218
VIII. Practical Problems of Stress-Corrosion Cracking in High Strength Metallic Materials by Markus O. Speidel	222
IX. Stress Corrosion in Nuclear Systems by S. H. Bush.	244
X. Some Thoughts on SCC Problems by B. F. Brown	285
XI. Notes on Electrochemistry by R. A. Oriani.	288
XII. A Mechanism of Stress-Corrosion Cracking in Plastic Solids by J. J. Gilman.	294
XIII. Environmentally Induced Brittle Delayed Failure: The Stress-Corrosion Problem by A. R. Troiano	296
XIV. Remarks on Stress-Corrosion Cracking by H. H. Uhlig.	305

Foreword

This collection of papers does not constitute a formal reporting of the activities of the ARPA Materials Research Council Summer Conference. Each report, memorandum or technical note is a draft of the author or authors and is their work alone. The Steering Committee, in conjunction with the authors, will decide how this material can best be presented as a formal report to ARPA.

Steering Committee

Professor Elliott W. Montroll
Secretary of the Steering Committee
Department of Physics & Astronomy
University of Rochester
Rochester, New York 14534

Professor Morris Cohen
Department of Metallurgy & Materials Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Professor John P. Hirth
Metallurgical Engineering Department
Ohio State University
Columbus, Ohio 43201

Professor John L. Margrave
Department of Chemistry
Rice University
Houston, Texas 77001

Professor Frank A. McClintock
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Professor Howard Reis
Department of Chemistry
University of California
Los Angeles, California 90024

Professor James Rice
Division of Engineering
Brown University
Providence, Rhode Island 02912

Professor Michael Tinkham
Department of Physics
Harvard University
Cambridge, Massachusetts 02138

Dr. George H. Vineyard
Brookhaven National Laboratory
Upton, Long Island, New York 11973

Project Director

Professor Edward E. Huckle
Department of Materials and Metallurgical
Engineering
The University of Michigan
Ann Arbor, Michigan 48104

Participants

Professor Michael B. Bever
Department of Metallurgy & Materials Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Professor Nico Bloembergen
Division of Engineering and Applied Physics
Pierce Hall
Harvard University
Cambridge, Massachusetts 02138

Professor Bernard Budiansky
Division of Engineering & Applied Science
Harvard University
Cambridge, Massachusetts 02138

Dean Daniel C. Drucker
Engineering College
University of Illinois
Urbana, Illinois 61801

Professor Pol E. Duwez
W. M. Keck Laboratory of Engineering
Materials
California Institute of Technology
Pasadena, California 91109

Professor John D. Ferry
Department of Chemistry
University of Wisconsin
Madison, Wisconsin 53706

Professor Willis H. Flygare
Noyes Chemical Laboratory
University of Illinois
Urbana, Illinois 61801

Dr. John J. Gilman, Director
Materials Research Center
Allied Chemical Corporation
Morristown, New Jersey 07960

Professor Robert Gomer
James Franck Institute
5640 Ellis Avenue
Chicago, Illinois 60637

Professor Robert A. Huggins
Center for Materials Research
Stanford University
Stanford, California 94305

Professor Walter Kohn
Department of Physics
University of California
La Jolla, California 92037

Professor James A. Krumhansl
Department of Physics
Clark Hall of Science
Cornell University
Ithaca, New York 14850

Professor Erastus H. Lee
Department of Applied Mechanics
Stanford University
Stanford, California 94305

Professor Donald J. Lyman
Materials Science & Engineering
The University of Utah
Salt Lake City, Utah 84112

Professor Paul L. Richards
Department of Physics
University of California
Berkeley, California 94720

Professor Albert J. Sievers
Laboratory of Atomic & Solid State Physics
Cornell University
Ithaca, New York 14850

Dr. Robb M. Thomson
National Bureau of Standards
Institute of Applied Technology
Washington, D.C. 20234

Guest Consultants

Dr. Joel A. Appelbaum
Bell Laboratories
600 Mountain Avenue
Murray Hill, New Jersey 07974
(201)582-2156

Dr. Francis Baratta
Army Materials and Mechanics
Research Center
Watertown, Massachusetts 02172
(617)926-1900

Dr. A. S. Barker, Jr.
Bell Laboratories
600 Mountain Avenue
Murray Hill, New Jersey 07974
(201)582-3000

Dr. Bernard Bendow
Air Force Cambridge Research
Laboratories
L. C. Hanscom Field
Bedford, Massachusetts 01730

Professor A. Bienenstock
Materials Science Department
Stanford University
Stanford, California 94305
(415)321-2300 Ext. 2617 or 2534

Dr. G. H. Bishop
Metals Division
Army Materials and Mechanics
Research Center
Watertown, Massachusetts 02172
(617)926-1900

Dr. Joseph I. Bluhm
Army Materials and Mechanics
Research Center
Watertown, Massachusetts 02172
(617)926-1900

Dr. J. C. Bokros
Medical Products Division
Gulf Energy & Environmental
Systems Company
P.O. Box 608
San Diego, California 92112

Professor R. H. Bragg
Department of Materials
Science
University of California
Berkeley, California 94720
(415)642-3815

Dr. S. H. Bush
Battelle-Northwest
P.O. Box 999
Richland, Washington 99352
(509)946-2223

Professor John M. Carpenter
Nuclear Engineering Department
300 Automotive Laboratory
University of Michigan
Ann Arbor, Michigan 48104
(313)764-4260

Dr. Fred Chomppf
Polymer Science Department
Scientific Research Staff
Ford Motor Company
P.O. Box 2053
Dearborn, Michigan 48121
(313)323-1423

Mr. B. Cohen
Aeronautical Systems Support
Branch
Materials Support Division
AF Materials Laboratory
Wright-Patterson AFB, Ohio 45433

Mr. W. J. Croft
Army Materials & Mechanics
Research Center
Watertown, Massachusetts 02172
(617)926-1900

Dr. Gordon T. Danby
Brookhaven National Laboratory
Upton, Long Island,
New York, 11973
(516)924-6262 Ext. 2471

Dr. Thomas Deutsch
Raytheon Research Division
28 Seyon Street
Waltham, Massachusetts 02154
(617)899-8400

Professor H. Ehrenreich
Pierce Hall
Harvard University
Cambridge, Massachusetts 02138
(617)495-3213

Dr. David Emin
Division 5134
Sandia Laboratories
Albuquerque, New Mexico 87115
(505)264-3431 or 5156

Dr. Sabri Ergun
U.S. Bureau of Mines
4800 Forbes Avenue
Pittsburgh, Pennsylvania 15213

Professor David Fischbach
Ceramic Engineering Division
Roberts Hall, FB-10
University of Washington
Seattle, Washington 98195
(206)543-8573

Professor Merton Flemings
Room 8-407
Department of Materials Science
Massachusetts Institute of
Technology
Cambridge, Massachusetts 02139
(617)864-6900 Ext. 3233

Dr. H. L. Gegel
Advanced Metallurgical Studies
Branch
Metals & Ceramics Division
AF Materials Laboratory
Wright-Patterson AFB, Ohio 45433

Dr. David Goldstein
Code 211
Naval Ordnance Laboratory
Silver Springs
White Oak, Maryland 20910

Dr. Alvin E. Gorum
Director, Army Materials
& Mechanics Research Center
Watertown, Massachusetts 02172
(617)926-1900 Ext. 275

Dr. Barry Granoff
Sandia Laboratories
Sandia Base
Albuquerque, New Mexico 87115
(505)264-5458

Dr. Donald E. Harrison
Manager, Materials Science
Research & Development Center
Westinghouse Electric Company
Pittsburgh, Pennsylvania 15235
(412)256-7000 Ext. 7336

Dr. J. D. Hoffman, Director
Institute of Materials Science
National Bureau of Standards
Gaithersburg, Maryland

Dr. Frank Horrigan
Raytheon Research Division
28 Seyon Street
Waltham, Massachusetts 02154
(617)899-8400

Dr. John R. Jasperse
Solid State Sciences Laboratory
Air Force Cambridge Research
Laboratories
L. C. Hanscom Field
Bedford, Massachusetts 01730
(617)232-5464

Professor H. H. Johnson, Head
Department of Materials Science
& Engineering
Cornell University
Ithaca, New York 14850
(607)256-4135

Dr. R. P. Kambour
Research & Development Center
General Electric Company
P.O. Box 8
Schenectady, New York 12301

Dr. R. Nathan Katz
Chief, Ceramics Research Division
Army Materials & Mechanics
Research Center
Watertown, Massachusetts 02172
(617)926-1900 Ext. 415

Mr. T. F. Kearns
AIR 320A-Research & Technology
Group
Naval Air Systems Command
Department of the Navy
Washington, D.C. 20360
(202)692-7416

Dr. R. Latanision
Research Institute for Advanced
Studies
1450 South Rolling Road
Baltimore, Maryland 21227
(301)247-0770

Dr. J. R. Low, Jr.
Department of Metallurgy &
Materials Science
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213
(412)621-2600

Mr. Arthur F. McLean
Manager, Turbine Research
Ford Motor Company
20000 Rotunda Drive
Dearborn, Michigan 48121
(313)322-3859

Professor W. G. McMillan
Department of Chemistry
University of California
Los Angeles, California 90024

Mr. David F. R. Mildner
Department of Nuclear Engineering
University of Michigan
Ann Arbor, Michigan 48104
(313)764-6220

Dr. Perry Miles
Raytheon Research Division
28 Seyon Street
Waltham, Massachusetts 02154
(617)899-8400

Dr. S. C. Moss
Energy Conversion Devices, Inc.
1675 West Maple Road
Troy, Michigan 48084
(313)549-7300

Dr. R. A. Oriani
E. C. Bain Laboratory
United States Steel Corporation
Research Center
Monroeville, Pennsylvania 15146

Mr. Antoni Paluszny
Supervisor, Turbine Research
Ford Motor Company
20000 Rotunda Drive
Dearborn, Michigan 48121
(313)337-5515

Dr. P. C. Paris
Del Research Corporation
427 Main Street
Hellertown, Pennsylvania 18055
(215)838-7069

Dr. H. W. Paxton, Head
Materials Science Department
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(412)621-2600

Dr. R. Pelloux
Room 8-305
Massachusetts Institute of
Technology
Cambridge, Massachusetts 02139

Professor G. M. Pound
Department of Materials Science
Stanford University
Stanford, California 94305
(415)321-2300 Ext. 4257

Dr. James Powell
T-318
Brookhaven National Laboratory
Upton, Long Island,
New York 11973
(516)924-6262 Ext. 7789

Dr. E. N. Pugh
Department of Metallurgy &
Mining Engineering
University of Illinois
Urbana, Illinois 61801
(217)333-4692

Mr. G. Quinn
Army Materials & Mechanics
Research Center
Watertown, Massachusetts 02172

Dean David V. Ragone
Thayer School of Engineering
Dartmouth College
Hanover, New Hampshire 03755
(603)646-2238

Dr. W. G. Ramke, Chief
Ceramics & Graphite Branch
Metals & Ceramics Division
Air Force Materials Laboratory
Wright-Patterson AFB
Ohio 45433

Mr. Leo F. Salzberg
Chief, Materials Physics Division
AF Materials Laboratory
Wright-Patterson AFB
Ohio 45433
(513)255-2433

Dr. Richard J. Sanford
U.S. Naval Ordnance Laboratory
Code 431
White Oak, Silver Spring
Maryland 20910
(301)495-8539

Dr. Wayne W. Scanlon
Applied Physics Department
U.S. Naval Ordnance Laboratory
White Oak, Silver Spring
Maryland 20910
(301)495-7773

Dr. R. J. Schaller
Research & Development Center
Westinghouse Electric Company
Pittsburgh, Pennsylvania 15235
(412)256-7000

Dr. A. J. Sedriks
Research Institute for Advanced
Studies
1450 South Rolling Road
Baltimore, Maryland 21227
(412)621-2600

Professor Mitchel Shen
Department of Chemistry
University of California
Berkeley, California 94720
(415)642-2111

Dr. M. J. Sinnott, Director
Materials Science Office
Advanced Research Projects
Agency
1400 Wilson Boulevard
Arlington, Virginia 22209
(202)694-3010

Dr. Morton C. Smith
Los Alamos Scientific Laboratory
Los Alamos, New Mexico

Dr. Richard A. Smith
TREE Section Head
U.S. Naval Ordnance Laboratory
Code 431
White Oak, Silver Spring
Maryland 20910
(301)495-8334

Dr. Marshall Sparks
The Rand Corporation
1700 Main Street
Santa Monica, California 60406
(213)393-0411

Dr. Marcus O. Spidel
Boeing Scientific Laboratory
P.O. Box 3981
Seattle, Washington 98124
(206)655-6279

Professor R. W. Staehle
Department of Metallurgical
Engineering
Ohio State University
116 West 19th Avenue
Columbus, Ohio 43210
(614)422-6255

Professor Eli Sternberg
Division of Engineering and
Applied Science
California Institute of Technology
Pasadena, California 91109
(213)795-6841 Ext. 1178

Dr. C. Martin Stickley
Deputy Director
Materials Sciences Office
Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209
(202)694-3010

Dr. Garth L. Tingey
Technical Leader
Ceramics & Graphite Section
Battelle-Northwest
P.O. Box 999
Richland, Washington 99352
(509)946-2419

Dr. A. R. Troiano
Department of Metallurgy and
Materials
Case Western Reserve University
Cleveland, Ohio 44106
(216)368-4234

Professor David Turnbull
Pierce Hall
Department of Applied Physics
Harvard University
Cambridge, Massachusetts 02138
(617)495-2038

Dr. H. H. Uhlig
Department of Metallurgy &
Materials Science
Massachusetts Institute of
Technology
Cambridge, Massachusetts 02139
(617)864-6900 Ext. 3313

Professor P. L. Walker, Jr.
Department of Materials Science
101 Minerals Industry Building
Pennsylvania State University
University Park, Pennsylvania 16802
(814)865-0497

Professor James O. Wilken
Department of Chemical
Engineering
University of Michigan
Ann Arbor, Michigan 48104
(313)764-2385

Dr. Mark L. Wilkins
L-24
Lawrence Radiation Laboratory
East Avenue
Livermore, California 94550
(415)447-1100 Ext. 8611

Dr. H. G. F. Wildorf, Chairman
Department of Materials Science
Thornton Hall
University of Virginia
Charlottesville, Virginia 22901
(703)924-3462

Mr. Harry V. Winsor
Air Force Weapons Laboratory/LRO
Kirtland Air Force Base
New Mexico 87117

Dr. Richard Weiss
Army Materials & Mechanics
Research Center
Watertown, Massachusetts 02172

Dr. J. Westbrook
Bldg. K-1, Rm. 3A44
General Electric Research &
Development Center
Schenectady, New York 12305

Dr. A. R. C. Westwood
Deputy Director, RIAS
1450 South Rolling Road
Baltimore, Maryland 21227
(301)247-0700

Professor A. Yariv
Department of Electrical
Engineering
California Institute of
Technology
Pasadena, California 91109
(213)795-6841 Ext. 1821

Professor G. S. Y. Yeh
Materials & Metallurgical
Engineering
University of Michigan
Ann Arbor, Michigan 48104
(313)764-9236

GRADIENTS IN POLYMERIC MATERIALS

M. Shen and M. B. Bever

Abstract

In this memorandum we consider the structure and properties of polymeric materials possessing spatial gradients. Potential applications of such materials are also discussed.

Gradients in the structure of polymeric systems may be generated by varying the chemical nature of the monomers, the molecular constitution of the polymers and the supra-molecular structure or morphology of the polymers. Gradients in each of these categories are possible for single-phase as well as heterophase systems. Such gradients are associated with gradients in properties.

The properties considered are chemical, mechanical, biomedical and transport properties. Structural gradients in the polymeric system may lead to a desired gradient in a single property, or to a combination of more than one property which may assume optimum values in different regions of the material. In the latter case, one of the properties is frequently related to mechanical integrity.

Possible applications of gradient polymeric systems include plastic gasoline tanks, biomedical implants, and damping materials for a wide frequency range.

GRADIENTS IN POLYMERIC MATERIALS

M. Shen and M. B. Bever

Introduction

The structure and properties of polymeric materials can be varied over wide ranges. Most of the pertinent efforts in this regard have been devoted to homogeneous and in some measure to laminated polymeric systems. Little attention has been directed toward gradient systems, that is, systems in which structure and properties vary continuously in space.

Gradient materials in general are beginning to attract some attention. Bever and Duwez¹ have considered gradients in composite materials. Ferry² has analyzed some aspects of gradients in the cross linking and swelling of polymers.

In this memorandum we shall systematically discuss possible gradients in polymeric materials. We shall be concerned primarily with the structure and properties of these materials rather than the techniques for their preparation. We shall mention various illustrative examples but shall not attempt to explore exhaustively the entire range of possible applications.

Structure

We shall consider the structural features of polymers on three levels of resolution:

- (i) the chemical nature of the monomers;
- (ii) the molecular constitution of the polymers; and
- (iii) the supramolecular structure or morphology of the polymers.

In each of these categories the possibility of gradients in space exists. We shall discuss these gradients first for single-phase systems and then for heterophase systems.

Single-Phase Systems

The nature of a polymer depends first of all on the constituent monomers. For instance, at room temperature the polymer of methyl methacrylate is a tough, hard plastic, commonly known as "Plexiglas"; by contrast the polymer of methyl acrylate is a soft rubber. These two monomers can be combined to form random copolymers of any composition ranging from pure poly(methyl methacrylate) to pure poly(methyl acrylate). A gradient polymer with a spatial variation in composition can therefore be prepared. Polymer chemistry offers many possibilities for creating similar compositional gradients by copolymerization techniques.

The degree of crosslinking is an important variable in many polymers. For example, the rubber in an elastic band is lightly crosslinked; an increase in the degree of crosslinking converts the same material into a hard plastic. Various techniques are available for controlling the degree of crosslinking, such as irradiation and the use of chemical crosslinking agents. These techniques can be adapted to producing

gradients in the degree of crosslinking. Some of these techniques have been considered by Ferry⁷ for gradient structures in hydrophilic polymers.

Differences in the degree of crosslinking are also associated with differences in the degree of equilibrium swelling (i.e., the amount of solvent imbibed in the polymer network). The degree of equilibrium swelling can also be controlled by changing the chemical nature of the monomers. For example, pure poly(hydroxyethyl methacrylate) can absorb an appreciable amount of water. The addition of even a small amount of a nonhydrophilic monomer, such as methyl methacrylate, to form a copolymer with hydroxyethyl methacrylate drastically reduces the degree of swelling due to the absorption of water. In general, the shear modulus of a polymer decreases with increasing degree of swelling.

The average molecular weight and the molecular weight distribution are other variables affecting single-phase polymers. Since these variables can in principle be manipulated (e.g., by high-shear flow), corresponding gradient structures may be achieved.

Oriented polymers possess certain unique features which distinguish them from unoriented ones. For instance, biaxially stretched polycarbonates offer superior resistance to crazing and uniaxially stretched nylon has improved yield resistance. Gradients in the degree and the direction of the orientation are of obvious interest.

Recent evidence suggests the possibility of the existence of supramolecular structures in single-phase, non-crystalline polymers. The morphology associated with such structures merits consideration. Their effects would probably be similar to those of the supramolecular structures encountered in heterophase polymers.

Heterophase Systems

The gradients discussed in the previous section for single-phase systems can also occur in heterophase systems. Other gradients are uniquely characteristic of heterophase systems; they will be discussed in this section.

Crystallinity - Crystalline polymers are always heterophase because total crystallinity can almost never be achieved, at least not in the bulk state. An amorphous fraction which is always present constitutes a second phase. The ratio of crystalline to non-crystalline material, however, can be varied. The degree of crystallinity may be changed by several methods. For a given material, this is possible by changing the rate of cooling from the molten state or by annealing of the solid polymer.

The ability of a polymer to crystallize is also determined by the regularity of its chain structure. An isotactic polystyrene, for example, being more regular in structure is semicrystalline, whereas atactic polystyrene is always amorphous because of the greater randomness of its structure. The degree of crystallinity is therefore affected by the degree

of tacticity. Similar effects can also be achieved by adjusting the relative ratio of cis and trans isomers of a block polymer. For instance, trans-polyisoprene is crystalline, while cis-polyisoprene, i.e., natural rubber, is amorphous.

A third method of affecting the crystallinity of a polymer is copolymerization. Polytetrafluoroethylene or "Teflon" is a highly crystalline polymer. The addition of an appropriate amount of perfluoropropylene as a random comonomer reduces the degree of crystallinity. In fact, a composition of 10% perfluoropropylene and 70% tetrafluoroethylene yields an excellent high-temperature resistant rubber, known as Teflon PEP, in contradistinction to Teflon, which is a hard plastic.

It is evident that each of the three methods of changing the degree of crystallinity described in the foregoing may be utilized in such a manner as to produce a gradient in the crystallinity of a polymer. Other aspects of crystallinity, such as the orientation and size of the spherulites, may also be varied. Their effects, however, are in general less marked than those of the degree of crystallinity.

Polymer Mixtures - Another important class of heterophase polymers can be prepared by mixing different polymers. Most polymers are insoluble in each other and thus show microphase separation. Most polymer mixtures, therefore, are heterophase. Three principal methods of mixing are available. The first consists of milling or solvent casting of two or more polymers, which produces the so-called "polyblends", the second

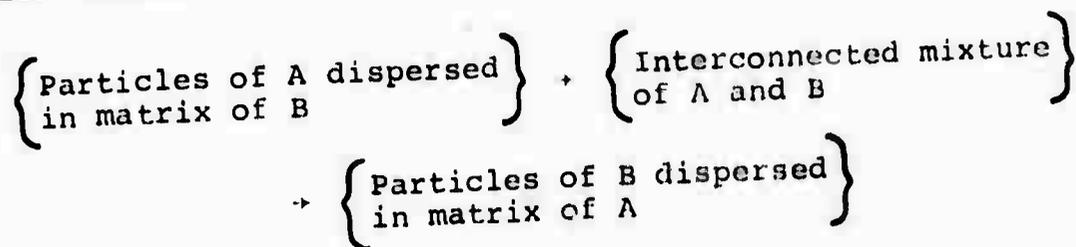
In the synthesis of block copolymers and the third the synthesis of graft copolymers. The preparation of polyblends by mixing is the simplest and most flexible method. The structures of block and graft copolymers are on a finer scale than those of the polyblends.

The morphology of heterophase mixtures results from the amount and configuration of the constituent phases. In one possible case, one phase is continuous, i.e., the matrix, and the particles of the other are dispersed in it. The variables are the amount of the dispersed phase and the sizes and shapes of the particles of this phase; for example, they may be spheres, rods or platelets. In another case, each of the two phases is interconnected with itself and the two phases are interpenetrating. In the third case, alternating layers of the two phases occur.

In polymers some of these different configurations can be generated by varying the relative amounts of the components or by the method of preparation. An example of the latter is the preparation of a polystyrene-continuous matrix of a styrene-butadiene-styrene block copolymer by casting from its solution in tetrahydrofuran/methyl ethyl ketone, whereas a polybutadiene-continuous matrix of the same block copolymer results if it is cast from a solution of benzene/heptane.

Gradients of the morphological features of heterophase systems may involve the amount, the particle size and particle shape of the dispersed phase. Typical examples of geometrical

possibilities, applied on the scale of composite materials, have been discussed by Bever and Duwez.¹ In the case of hetero-phase polymers, a gradient may be carried to the point where an inversion occurs according to the following scheme:



Gradients Involving Nonpolymeric Components

Reinforcing fillers, usually nonpolymeric in nature, cause marked changes in the properties of polymers, e.g., carbon black in natural rubber. The use of a concentration gradient of a filler in a polymer is an obvious possibility. The good bonding between carbon black and rubber suggests that glassy carbon may be used in conjunction with a polymeric phase.

A porosity gradient can be introduced into a polymer. For example, a foaming agent may be suitably incorporated to bring about this effect.

Another possible development is the joining of a polymer with an inorganic material, such as a metal or a ceramic, through a gradient. One type of such a joint requires a porosity gradient at the surface of the inorganic material and the impregnation of the pores by the polymer.

The doping of a polymer, such as the doping of polyethylene by iodine, changes the material from an insulator to a semiconductor. The use of a concentration gradient of the

doping agent constitutes a possible development.

Properties and Applications

The properties which make gradient polymers of interest are chemical, mechanical, biomedical and transport properties. In each of these categories, a variation in space of a given property can be achieved by a suitable gradient in the polymeric system. In addition, more than one property may be varied to satisfy multiple requirements of a potential application. Thus the coupling of several optimal properties is possible. In the following we shall consider the aforementioned classes of properties and suggest potential applications wherever possible.

Chemical Properties

Certain chemical properties of polymers are intimately related to the nature of the constituent monomers. For example, many polymers are hydrophobic. Hydrophilicity may be introduced by copolymerizing such hydrophobic monomers with hydrophilic monomers. One example is the copolymerization of the hydrophobic monomer of styrene with the hydrophilic monomer of styrene sulfonic acid. A gradient of hydrophilicity will result from continuously varying the ratio of these two monomers.

A possible application for materials of this type is the construction of gasoline tanks for aircraft or automobiles. A hydrophilic interior layer will prevent the gasoline from swelling the material, while a hydrophobic exterior layer will be inert to water in the environment. The strength may be

increased by blending the material with polyacrylonitrile and polybutadiene, like ABS resins. The use of a gradient in this case would be advantageous for maintaining structural integrity.

Other chemical properties that may be imparted to one surface layer of a gradient polymer may be increased resistance to oxidation, weathering, high-temperature degradation and irradiation. Resistance to specific chemical attack, such as by strong acids, may similarly be achieved.

Mechanical Properties

In general, rubbery polymers are characterized by a high elastic limit, low elastic moduli, low fracture stress, large fracture strain and high impact resistance. By contrast, glassy polymers have a low elastic limit, high elastic moduli, high fracture stress, small fracture strain and low impact resistance. The values of these properties of crystalline polymers approximate those of the glassy polymers. Most of the properties of one of these classes of materials can be modified by the addition of a material from another class. These modifications can be achieved by methods such as solvent swelling, degree of crosslinking, copolymerization, filler addition and blending. All of these methods have been shown in previous sections to be capable of producing gradients in composition and structure: corresponding gradients in mechanical properties can thus be achieved.

For the mechanical properties of some of the materials, geometrical configurations play an especially important role.

A good illustration of this is the case of mixtures of polystyrene and polybutadiene in nearly equal proportions. If polystyrene is the continuous phase and polybutadiene the dispersed phase, the composite is relatively rigid. If the same volume of polystyrene is present as dispersed particles in a matrix of polybutadiene, the material is rubbery. A transition between these two types of configuration requires the coexistence of domains of both types thus leading to a gradient in structure and properties.

Applications requiring a gradient in strength properties of polymeric materials can readily be envisioned. For example, if a relatively weak material is to be fastened to a structure, the region in which the fastener is to be applied should have a higher strength. This can be achieved by a suitable gradient material. Another example is the provision of a surface layer that is resistant to indentation and abrasion. The case of a soft surface layer on a hard substrate, which is of special interest for biomedical applications, will be mentioned in the next section.

The damping capacity of polymers, due to their visco-elastic nature, is highly frequency-dependent at a given temperature. The usefulness of a polymer as a damping material is confined to a limited frequency range. This range may be extended by using graded materials (prepared, for instance, by copolymerization) each of which has a damping maximum at a different frequency. The gradient in this application must not

be steep in order to achieve maximum effectiveness.

Biomedical Properties

Two essential requirements for biomaterials are compatibility with blood and tissues and inertness to body fluids. Both of these are surface properties, and hence can be achieved by the use of gradient polymers. Silicone rubber, for instance, is a versatile biomaterial. However, its lack of rigidity limits its range of applicability. By combining silicone rubber with a rigid base material in a continuous gradient, this limitation may be overcome. Ferry⁷ has analyzed mechanisms of achieving the same effect in hydrophilic polymers.

Glassy carbon is known to possess good biomedical properties. The possibility of forming a gradient material consisting of such polymers has been mentioned earlier. Such a material would be of obvious interest for some biomedical applications.

Transport Properties

The transport properties of interest here comprise the electrical conductivity, heat conductivity, and mass diffusivity. Each type of these transport properties depends on the characteristics of the polymeric system.

Generally polymers are good electrical insulators. However, by special modification some of them may be made into semiconductors. The example of iodine-doped polyethylene has been mentioned earlier. On the other hand, as insulators polymers vary in their dielectric strength. Polymeric systems

with spatially varying conductivities or dielectric strengths can be obtained by suitable gradients.

Since polymers are poor heat conductors, they are used as thermal insulating materials. They are particularly effective in the form of foams, e.g., "Styrofoam". A problem encountered in the use of polymeric foam as a thermal insulator for cryogenic fuel tanks in rockets has been delamination due to thermal shock. This difficulty can probably be overcome by the use of a gradient material, which combines in a continuous manner across its thickness the insulating capacity due to porosity and mechanical strength.

Many polymers permit the diffusion of liquids and gases of low molecular weight. In applications in which permeability is undesirable, such a polymer may be made impermeable by the introduction of an impermeable surface layer. The layer may confer greater mechanical strength by a gradient structure with the bulk. On the other hand, selective permeability for some substances may be useful. Examples are polymeric membranes used for sea water desalination or hemodialysis. In such instances a gradient structure may impart superior mechanical strength and rigidity without risk of delamination.

Combination of Properties

In the foregoing four sections, we have been primarily concerned with the spatial variations of a single property. However, several applications considered involve the combination of more than one property. An example of this is the use of

polymeric foam with variable porosity as a thermal insulator in which effects of both the thermal and mechanical property gradients are operative. Applications requiring combinations of mechanical properties with any of the other properties can certainly be envisioned. Coupling of gradients of various properties is generally feasible.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAH03-71-C-0253.

References

1. M. B. Bever and P. Duwez, On Gradient Composites, Preliminary Reports, Memoranda and Technical Notes of the ARPA Materials Summer Conference, July 1970, pp. 117-140.
2. J. D. Ferry, Control of Mechanical Properties of Swollen Hydrophilic Network Polymers; Layer and Gradient Structures, Preliminary Reports, Memoranda and Technical Notes of the ARPA Materials Summer Conference, July 1970, pp. 611-627.

EFFECT OF STRAIN ON THE FERMI ENERGY
OF SIMPLE METALS

R. Gomer and R. M. Thomson

Abstract

The effect of the strain on the chemical potential and work function of a metal is reinvestigated. The result is of some importance for electron emission from metal tips under high fields and at crack tips. The results show that the chemical potential change is more complex than that predicted in an earlier work by Schrieffer and Tiller, but usually of the same sign.

EFFECT OF STRAIN ON THE FERMI ENERGY OF SIMPLE METALS

R. Gomer and R. M. Thomson

The effect of elastic deformation on the Fermi energy of metals is of some relevance to stress corrosion cracking, since high stresses and consequently strains may occur at the crack tip. It is also of some interest in connection with field emission, and even more, field ionization and field desorption, since the very high electric fields required for the latter cause stresses of the order of 10^{11} dynes/cm² and thus strains of the order of $\delta = 0.02$ to 0.05 . The problem has been considered by Schrieffer and Tiller¹ in terms of the parameters of a free electron gas without, however, taking into consideration that the deformation changes the lowest energy level of the free electron gas, i.e., shifts the bottom of the box. It turns out that this effect is appreciable, and we present here a calculation which would be reasonably valid for free electron-like metals, i.e., those in the left hand portion of the periodic table. We emphasize at the outset that no attempt has been made to calculate the change in surface dipole moment caused by the deformation, so that our calculation applies strictly only to the inner potential.

For metals which can be treated by the Wigner-Seitz method the average energy per electron is given by²

$$E = \epsilon_0 + 2.21\alpha/r_e^2 + 1.2Z^{2/3}/r_e - .916/r_e + W(r_e) \text{ Rydberg} \quad (1)$$

where ϵ_0 is the energy of the lowest electron ie state, $\alpha = m/m^*$ with m^* an effective electron mass, Z the valence of the atoms of the metal, W a correlation energy, and r_e the radius of a sphere containing just one electron, in Bohr units:

$$\left(\frac{4}{3}\right)\pi a_H^3 r_e^3 = V/n_e \quad (2)$$

where $a_H = 0.531 \text{ \AA}$, V the volume of the system, and n_e the number of electrons in it.

The chemical potential μ is defined as

$$\mu = \left(\frac{\partial (n_e E)}{\partial n_e}\right)_V = \epsilon_0 + \left(\frac{5}{3}\right)E_k + \left(\frac{4}{3}\right)E_{cx} + W\left(1 + \frac{1}{3} \frac{r_e}{r_e + 7.8}\right) \quad (3)$$

where

$$E_k \equiv 2.21(m/m^*)/r_e^2 \quad (4a)$$

$$E_{cx} \equiv (1.2Z^{2/3} - .916)/r_e \quad (4b)$$

$$\text{and } W \equiv - \frac{.88}{r_e + 7.8} \quad (4c)$$

where we have used in (4c) the Wigner interpolation form of the correlation energy².

Let us assume that a piece of metal is compressed (dilated) uniformly and reversibly. The mechanical reversible work U done on the system is

$$U = V k' \delta^2 / 2 \quad (5)$$

where k' is an appropriate modulus. Thus, the energy increase per electron u is

$$vk'\delta^2/2 = \Delta c_0 + \Delta E_k + \Delta E_{cx} + \Delta W \quad (6)$$

where v is the volume per electron, $(4/3)\pi a_{II}^3 r_e^3$. Note that u is always positive regardless of the sign of δ , when a system is deformed from its equilibrium state. The Δ 's in Equation 6 refer to the changes in the various terms, defined in Eqs. 4a-c when the electron radius changes. In an ab initio calculation we could determine the elastic constants of the metal from Eq. 6 if the form of $c_0(r_e)$ were known. We content ourselves here with a much more modest approach; we will assume the moduli to be known and will thus solve for Δc_0 :

$$\Delta c_0 = u - \Delta E_k - \Delta E_{cx} - \Delta W \quad (7)$$

Since the signs of all but the first terms on the right side of Eq. 7 depend on the sign of the strain, Δc_0 can be positive or negative, of course. The change in μ is now found very simply from Eq. 3 and 7,

$$\begin{aligned} \Delta\mu &= \mu(r'_e) - \mu(r_e) = \Delta c_0 + (5/3)\Delta E_k + (4/3)\Delta E_{cx} + \Delta W + \frac{1}{3} \left(\frac{W(r'_e)r'_e}{r'_e + 7.8} - \frac{W(r_e)r_e}{r_e + 7.8} \right) \\ &= u + (2/3)\Delta E_k + \frac{1}{3} \Delta E_{cx} + \frac{1}{3} \left(\frac{W(r'_e)r'_e}{r'_e + 7.8} - \frac{W(r_e)r_e}{r_e + 7.8} \right) \end{aligned} \quad (8)$$

which yields in terms of strain $\delta = (r'_e - r_e)/r_e$

$$\Delta u = vk' \delta^2 / 2 - (2.94/r_0^2) \frac{\delta}{(1+\delta)} - \left((0.42^{2/3} - .3) / r_0 \right) \delta - \frac{.29 r_0}{(r_0 + 7.8)} \delta \quad (9)$$

In terms of Young's modulus k , and of Poisson's ratio ν ,

$$k' / 2 = \frac{9k}{6(1-2\nu)} \quad (10)$$

so that we have for the first term on the right of Eq. 9

$$u = 3.9 \cdot 10^{-11} \frac{k}{(1-2\nu)} r_0^3 \delta^2 \quad \text{Rydberg} \quad (11)$$

Numerical Results

Table I lists the results for potassium, aluminum, and tungsten, based on the parameters listed and a value of $\nu = 0.3$. The effect of the u term is considerable as can be seen from the differences in the absolute values of Δu for tensile (δ positive and compressive (δ negative strain). In the case of tungsten the values listed are based on an effective valence of 5, i.e., the d-electrons are included. It may be more correct to assume that the largest contribution comes from s-electrons, in which case the r_0 value for W would be ~ 3 . If this were the case, the u term would predominate to the extent of changing the sign of Δu for tensile strain. It must be added however that in the case of transition metals part of the work term (or alternately a fraction of Young's modulus) must be assigned to core-core interactions, so that the effective k should undoubtedly be smaller than the overall value. This would tend to deemphasize the effect of the u term.

Acknowledgment

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAKC15-71-C-0253 with The University of Michigan.

References

1. W. A. Tiller and R. Schrieffer, *Scripta Met.*, 4, 47 (1970).
2. S. Raimes, "The Wave Mechanics of Electrons in Metals", North-Holland Press (1967).

TABLE I

δ	$K r_e = 5$ $k = .04 \times 10^{12}$	$\Delta L r_e = 2.1$ $Z = 3 \quad k = 1 \times 10^{12}$	$W r_e = 2$ $Z = 5 \quad k = 5 \times 10^{12}$
	$\Delta\mu$	$\Delta\mu$	$\Delta\mu$
+ .1	-0.09 eV	-0.95 eV	-0.8 eV
- .1	0.2 eV	1.2	2.0 eV
+ .05	-0.06 eV	-0.5	-0.3 eV
- .05	0.1	0.6	0.6 eV
+ .01	-	-	-0.1 eV
- .01	-	-	0.1

(Positive $\Delta\mu$ means a decrease in work function)

Change in Work Function with Mechanical Stress

We consider at first only free-electron like metals to which the Wigner-Seitz method can be applied.

The average energy per electron is given (Raimes p. 276, Eq. 9.92) by

$$E_0(r_s) = \epsilon_0 + \alpha \frac{2.21Z^{2/3}}{r_s^2} + \frac{1.2Z}{r_s} - \frac{.916Z^{2/3}}{r_s} + W\left(Z^{-1/3}r_s\right) \quad (1)$$

where $\alpha = m/m^* \approx 1$; $W = \frac{-0.88}{r_e + 7.8}$

Z = valence of atom

$$\frac{4}{3} \pi r_s^3 a_H^3 = \left(\frac{n \text{ atom}}{V}\right)^{-1} \quad \text{i.e., } r_s = \text{atomic ratios in Bohn units, } a_H = .53\text{\AA}.$$

$$r_e = r_s/Z^{1/3}$$

Therefore, in terms of r_e

$$E(r_e) = \epsilon_0 + \alpha \frac{2.21}{r_e^2} + \frac{1.2Z^{2/3}}{r_e} - \frac{.916}{r_e} + W(r_e) \quad (2)$$

The chemical potential is defined as

$$\left(\frac{\partial E_{r \cdot r}}{\partial n_e}\right)_V = \left(\frac{\partial nE(r_e)}{\partial n_e}\right)_V = E(r_e) + n_e \left(\frac{\partial E(r)}{\partial r_e}\right)_V \left(\frac{\partial r_e}{\partial n_e}\right)_V$$

To find $\frac{\partial r_e}{\partial n_e}$:

$$\frac{4}{3} \pi a_H^3 r_e^3 = V/n_e$$

$$\frac{\partial r}{\partial n_e} = \frac{V}{n_e^2} - \frac{1}{4\pi a^3 r^2} = - \frac{V}{2n_e^2} \frac{r_e}{\frac{4}{3}\pi a^3 r_e^3} = - \frac{V}{3n_e^2} \frac{n}{V} = - \frac{r_e}{3n_e}$$

$$\begin{aligned} \therefore \left(\frac{\partial E_{r \cdot r}}{\partial n_e} \right)_v &= \epsilon_0 + \frac{.21}{r_e^2} + \frac{1.2Z^{2/3}}{r_e} - \frac{.916}{r_e} - \frac{.88}{r_e + 7.8} \\ &+ \frac{2}{3} \frac{2.21}{r_e^2} + \frac{1}{3} \frac{1.2Z^{2/3}}{r_e} - \frac{1}{3} \frac{.916}{r_e} - \frac{1}{3} \frac{.88r_e}{(r_e + 7.8)^2} \\ &= \epsilon_0 + \frac{5}{3} E_k + \frac{4}{3} E_{cx} + W \left(1 + \frac{.33r_e}{r_e + 7.8} \right) \end{aligned}$$

where $E_k = \frac{2.21}{r_e^2}$; $E_{cx} = \frac{1.2Z^{2/3}}{r_e} - \frac{.916}{r_e}$; $W = - \frac{.88}{r_e + 7.8}$

Now a strain, caused by mechanical work increases $nE(r_e) = E_{r \cdot r}$ by $\Delta U = V_k \delta^2 / 2$ ($\delta =$ strain; $k =$ modulus). Thus if $v =$ volume per electron

$$\Delta U = v k \delta^2 / 2 = \Delta \epsilon_0 + \Delta E_k + \Delta E_{cx} + \Delta W$$

where $\Delta = f(r_e') - f(r_e)$ with r_e' referring to effect in the electron radius for the strained system. Thus,

$$\Delta \epsilon_0 = \Delta U - \Delta E_k - \Delta E_{cx} - \Delta W$$

Whence

$$\mu' \cdot \mu = \Delta \epsilon_0 + \frac{5}{3} \Delta E_k + \frac{4}{3} \Delta E_{cx} + \Delta W + \frac{.33W'r'_e}{r'_e + 7.8} - \frac{.33Wr_e}{r_e + 7.8}$$

$$= \Delta U - \Delta E_k - \Delta E_{cx} - \Delta W + \dots$$

$$\Delta \mu = \Delta U + \frac{2}{3} \Delta E_k + \frac{1}{3} \Delta E_{cx} + \frac{.33W'r'_e}{r'_e + 7.8} - \frac{.33Wr_e}{r_e + 7.8}$$

$$= k\delta^2/2 + .66 \times 2.21 \frac{r^2 - r'^2}{r^2 r'^2} + .4Z^{2/3} \left(\frac{r-r'}{r \cdot r'} \right) - .305 \left(\frac{r-r'}{r \cdot r'} \right)$$

$$\left(\frac{.293}{F+7.8} \left(\frac{r'-r}{F+7.8} \right) = - \frac{.293 \Delta r}{(F+7.8)^2} \right)$$

$$\Delta\mu^2 = k\delta^2/2 + 1.47 - \frac{\Delta r \cdot 2}{F^3} - \frac{.4Z^{2/3} \Delta r}{F^2} + \frac{.305 \Delta r}{F^2} - \frac{.293 \Delta r}{(F+7.8)^2}$$

where we have approximated

$$\frac{r^2 - r'^2}{r^2 r'^2} \approx - \frac{2r \Delta r}{r^2 r'^2} \approx - \frac{2\Delta r}{F^3}$$

would be better as $-\frac{\Delta(r^2)}{F^4}$

$$\Delta\mu^2 \approx v_e k\delta^2/2 - 1.47 \frac{\Delta(r^2)}{F^4 r^2} + \frac{\Delta r}{F^2} \left(.305 - .4Z^{2/3} \right) + \frac{.293 \Delta r}{(F+7.8)^2}$$

$$\Delta\mu = k\delta^2/2 - 1.47 \frac{\Delta(r^2)}{r^2 \cdot r'^2} - \frac{\Delta r}{F^2} (.4Z^{2/3} - .3) - \frac{.29 \Delta r}{(F+7.8)^2}$$

We need now a relation between ΔU and Δr

First $v_e = \frac{4}{3} \pi a_H^3 r_e^3$

$$\therefore \Delta v_e \approx 4 \pi a_H^3 r_e^3$$

and $\frac{\Delta v_e}{v_e} \approx 3 \frac{\Delta r_e}{r_e}$ also, $\frac{\Delta v}{v} = 3\delta \therefore \frac{\Delta r_e}{r_e} = \delta$

ΔU for compressive stress is $\frac{1}{3(1-\nu)} \left(\frac{\Delta v}{v} \right)^2$ per unit volume,

where k = Young's modulus, and ν = Poisson's ratio $\approx .3$

$\therefore \Delta U/\text{electron}$ is

$$\begin{aligned}
& v_e \frac{k}{\delta(1-2\nu)} \left(3 \frac{\Delta r_e}{r_e} \right)^2 \\
&= \frac{4}{3} \pi a_H^3 r_e^3 k' (9) \frac{(\Delta r)^2}{r_e} \\
&= \frac{4}{3} \pi a_H^3 k' (9) r_e \Delta r_e^2 \quad \text{in ergs}
\end{aligned}$$

To convert to Rydberg

$$\begin{aligned}
\Delta U &= \frac{4 \times 3.14 \times (.53)^3 \times 10^{-24} \times 9}{3 \times 9 \times 2.18 \times 10^{-11}} k' \text{ exponents } r_e (\Delta r)^2 \\
&= 2.6 \times 10^{-13} k' r_e (\Delta r_e)^2 = 2.6 \times 10^{-13} k' r_e^3 \delta^2 \text{ Rydberg}
\end{aligned}$$

or

$$\begin{aligned}
\Delta U &= v_e k' \left(\frac{\Delta v}{v} \right)^2 & \frac{\Delta v}{v} &= .36 \\
&= v_e k' 9\delta^2 \\
&= v_e \frac{3k}{2(1-2\nu)} \delta^2 \quad \text{ergs} \\
&= \frac{v_e \times 10^{-11}}{2.8} \frac{3k}{2(1-2\nu)} \delta^2 \quad \text{Rydberg}
\end{aligned}$$

$$v_e = \frac{\rho_M}{6 \times 10^{23} Z} = \quad \rho_M^2 = \text{metal density}$$

$$\therefore \Delta U = \frac{(\rho_M/Z) k' 9\delta^2}{6.02 \times 10^{23} \times 2.8 \times 10^{-11}} \text{ Rydberg}$$

$$= (\rho_M/Z) k' \delta^2 \times 5.3 \times 10^{-13}$$

$$\Delta U = \frac{2.6 \times 10^{-13} k r_e^3}{\delta(1-2\nu)} \delta^2$$

let $v = .3$

$$\Delta U = \frac{2.6 \times 10^{-11}}{6 \times .4} k r_{\oplus}^3 \delta^2$$

$$= 1.1 \times 10^{-11} k r_{\oplus}^3 \delta^2 = 1.1 (k/10^{11}) r_{\oplus}^3 \delta^2$$

Now

$$\Delta \mu = .11 (k/10^{11}) r_{\oplus}^3 \delta^2 - \frac{1.47 \times 2}{(1+\delta)^2 r_{\oplus}^2} \delta - \frac{(.42^2 / .3 - .3)}{r_{\oplus}} \delta$$

$$+ \frac{.29 r_{\oplus}}{(r_{\oplus} + 7.8)^2} \delta$$

$r_{\oplus} = 2$ $z = 1$

Let $\delta = +.1$ $\frac{k}{10^{11}} = 2$

$$\Delta \mu = 1.1 \times 8 \times 2 \times .01 - \frac{1.47 \times .1}{2} - \frac{.1 \times .1}{2} - \frac{.59 \times .1}{(9.8)^2}$$

$$.16 \quad - .074 \quad - .005$$

$$= \underline{+.08 \text{ Rydberg}} = \underline{+1.08 \text{ eV}}$$

$\delta = +.05$

$$\Delta \mu = .16/4 \quad - .037 \quad - .0025 = \underline{.0005 \text{ Rydberg}}$$

$$= \underline{.006 \text{ eV}}$$

For Al $\frac{k}{10^{11}} = .14$

$r_{\oplus} 5.0$	K	39	$k = .011$	$45 \rho_M = .86$
2.1	Al	27	$k = 1$	$10/3 \rho_M = 2.7$
2.02	W	184	$k = 5$	$(9.9/5) \rho = 19.3$

$$V_{\infty} A^0 = .74 \times 10^{-9} (V/Z)^{1/2}$$

$$V_{\infty} / a_H = \frac{.74 \times 10^{-9}}{.53 \times 10^{-8}} (V/Z)^{1/2}$$

$$= \underline{1.4 (V/Z)^{1/2}}$$

Numerical Results

$$K \quad k \neq 10^{12} = .04 \quad r_0 = 5$$

$$\Delta\mu = .0044 \times 125 \delta^2 - \frac{2.94}{25} \frac{\delta}{(1+\delta)^2} - \frac{.1}{5} \delta + \frac{1.5}{152} \delta$$

$$= .55 \delta^2 - .118 \frac{\delta}{(1+\delta)^2} - .02\delta + .0092\delta$$

for $\delta = +.1$ (dilation)

$$\Delta\mu = .0055 - .0098 - .002 - .0009$$

$$= \underline{-.0063 \text{ Rydberg}} = \underline{-.085\text{eV}}$$

for $\delta = -.1$ (compression)

$$.0153 \text{ Rydberg}$$

$$= \underline{.2\text{eV}} + \underline{.2\text{eV}}$$

for $\delta = +.05$

$$\Delta\mu = .0013 - .0005 - .001 + .0005$$

$$= \underline{.0050 \text{ Rydberg}} = \underline{.06\text{eV}}$$

ΔI

$$k = 1 \quad r_0 = 2.1 \quad Z = 3$$

$$\Delta\mu = .11 \times 9.16^2 - \frac{2.94}{4.3} \frac{\delta}{(1+\delta)^2} - \frac{.4 \cdot 2.08 - .3}{2.1} \delta$$

$$= 116^2 - .685 \frac{\delta}{(1+\delta)^2} - .254\delta$$

for $\delta = +.1$

$$\Delta\mu = .01 - .0565 - .0254 = \underline{-.07 \text{ Rydberg}} = \underline{-.95\text{eV}}$$

for $\delta = 0.1$

$$\Delta\mu = \underline{.026} \text{ Rydberg} = \underline{1.2eV}$$

$\delta = .05$

$$.0025 - .031 - .012 = \underline{0.04} \text{ Rydberg} = \underline{-.54eV}$$

$$\delta = -.05 \quad \underline{+.045} \text{ Rydberg} = \underline{+.6eV}$$

W

$$k = 5 \quad r_s = 2 \quad z = 5$$

$$\Delta\mu = .88 \times 5\delta^2 - \frac{2.94}{4} \frac{\delta}{(1+\delta)^2} - \frac{.87}{2} \delta$$

$$4.4\delta^2 - .735 \frac{\delta}{(1+\delta)^2} - .435\delta$$

for $\delta = +.1$

$$\Delta\mu = .044 - .06 - .043 = \underline{-.05} \text{ Rydberg} = \underline{-.8eV}$$

$\delta = -.1$

for $\delta = .05$

$$\Delta\mu = .011 - .033 - .0217 = \underline{-.024} \text{ Rydberg} = \underline{-.32eV}$$

$$\delta = -.05 \quad \underline{.046} \text{ Rydberg} = \underline{+.62eV}$$

for $\delta = .01$

$$\Delta\mu = .00044 - .0072 - .0044 = \underline{-.011} \text{ Rydberg} = \underline{-.13eV}$$

NOTE ON DENSITY DETERMINATIONS
ON AMORPHOUS MATERIALS

Robert Gomer

Abstract

A possible method of measuring the density of amorphous films is described.

NOTE ON DENSITY DETERMINATIONS
OF AMORPHOUS MATERIALS

Robert Gomer

It appears that the determination of the density of amorphous films, for example of Ge, is of some importance in the interpretation of the optical electrical and structural properties of such materials. The following scheme may be useful in this connection. In order to determine the density it suffices to know the mass and volume of the film. The former can be determined with great accuracy and sensitivity by means of a quartz oscillator microbalance¹. The volume can be obtained from a knowledge of the area of the quartz crystal onto which the film is evaporated or sputtered and the thickness of the film. If the film is thin enough for optical transmission its thickness can be determined accurately by ellipsometry².

For opaque (but reflecting) films interferometric techniques also exist³ for determining film thickness. In order to compensate for the roughness of the original quartz surface, which may well be of the order of 1000\AA , it might be necessary to evaporate a pre-deposit of this order of thickness, of the material to be studied, heat it to near melting, and thus obtain an artificially smoothed surface. If a mask of somewhat smaller dimensions is now put in front

of the quartz crystal, the weight and additional thickness of any further deposit may be determined with respect to the smoothed substrate. The degree of roughness of the original quartz and of the smoothed surface can of course be determined interferometrically.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHCl5-71-C-0253 with The University of Michigan.

References

1. C. S. Stockbridge, Vacuum Microbalance Techniques, 5, 193 (1965), Plenum Press, New York.
2. W. A. Pliskin in Physical Measurement and Analysis of Thin Films, 1, Plenum Press, New York (1969), (from Progress in Analytical Chemistry, Vol. 2).
3. Multiple Beam Interferometry, S. Tlonasky, Oxford Press (1948); A. Weill, Structure and Properties of Solid Surfaces, 309 University of Chicago Press (1953).

MODIFIED NULL-FLUX MAGNETIC SUSPENSION
AND PROPULSION SYSTEM FOR HIGH-SPEED TRANSPORTATION

M. Tinkham and P. L. Richards

Abstract

A detailed Fourier analysis is given of a new hybrid system for magnetic suspension and propulsion of high-speed trains (~300 mph). The hybrid system combines the advantages of null-flux (low drag and feasible synchronous propulsion) with those of the image-force scheme (a cheap smooth track, typically 1/4" aluminum sheet); its disadvantage is that it requires a double set of opposing train magnets (as does one of the null-flux system already proposed by Powell and Danby). The analysis shows that a drag/lift ratio as low as 1/80 should be obtained with reasonable parameter values, far superior to any simple image-force system, and much less than air drag over the useful speed range. Another advantage of this system to the image-force scheme is that its strong magnets facilitate synchronous propulsion. It is shown that the active track required for this propulsion could be energized in sections of several miles without undue power loss. The stability of this drive is studied, and shown to offer no serious problems for small fluctuations, but there may be a serious problem in main-

Preceding page blank

taining synchronism during planned or accidental or accelerations
and decelerations.

**MODIFIED NULL-FLUX MAGNETIC SUSPENSION
AND PROPULSION SYSTEM FOR HIGH-SPEED TRANSPORTATION**

M. Tinkham and P. L. Richards

Table of Contents

	<u>Page</u>
Introduction	1
General Expressions for Lift and Drag Forces	2
Determination of Fields	4
Computation of Forces	8
Magnet Requirements	12
Optimization of Parameters	14
Linear Synchronous Propulsion	17
Transverse Stability	23
Longitudinal Stability	25
Acceleration to Speed	31
Appendix: Calculation of \bar{H}_z for a Sequence of Rectangular Magnet Loops	

Introduction

Richards and Tinkham have presented an analysis* of magnetic levitation of high speed trains which compared two proposed systems: 1) a periodic array of magnets moving over a conducting ground plane (the "image-force" scheme) and, 2) the Powell-Danby "null-flux" scheme. The latter uses a symmetric system of coupled coils for the track to obtain lower drag, but at the expense of a more complicated track which would increase costs to some degree as well as introduce a certain amount of vibration in the suspension.

It has been suggested** that a modified null-flux scheme using a thin conducting sheet (rather than coil) between two sets of opposing train magnets might combine the low-drag feature of the null-flux scheme with the simple smooth track construction of the ground plane scheme, at the expense of doubling the magnet requirement on the train. Since there will be hundreds of miles of track, and relatively few trains, such a tradeoff could be quite attractive. A further advantage of the null-flux schemes is that synchronous propulsion by an active track is facilitated by the stronger magnets which such schemes use to get a stiffer, lower-drag suspension than is feasible in image-force suspensions. Since dc superconducting magnets use no power, it is evidently desirable to capitalize

*Last year's ARPA Report.

**By the Ford group (Reitz and Borcherts), indirect private communication.

on the cheap availability of such strong fields.

In this report, we present results of an analysis of this modified null-flux scheme which confirms that such a scheme offers some very desirable features.

General Expressions for Lift and Drag Forces

We assume throughout that all fields vary as e^{ikx} , with $k = 2\pi/\ell_x$, ℓ_x being the magnet period. Then the lift and drag forces per unit area are given by

$$F_L = F_z = -\frac{1}{c} \operatorname{Re} \int J_y^* B_x dz \quad ,$$

$$F_D = F_x = \frac{1}{c} \operatorname{Re} \int J_y^* B_z dz \quad ,$$

where r.m.s. values are implied throughout, and the integration is through the thickness of the conducting sheet. It is convenient to re-express these results in terms of the field only, using Maxwell's equations $\nabla \times \vec{B} = \frac{4\pi}{c} \vec{J}$ and $\nabla \cdot \vec{B} = 0$. Thus

$$\begin{aligned} F_L &= -\frac{1}{4\pi} \operatorname{Re} \int \left(\frac{\partial B_x^*}{\partial z} + ik B_z^* \right) B_x dz \\ &= -\frac{1}{8\pi} \int \left\{ \frac{\partial}{\partial z} |B_x|^2 + 2 \operatorname{Re} ik B_z^* \left(\frac{1}{ik} \frac{\partial B_z}{\partial z} \right) \right\} dz \\ &= -\frac{1}{8\pi} \int \frac{d}{dz} \left(|B_x|^2 + |B_z|^2 \right) dz \\ &= \frac{1}{8\pi} \left\{ |B_{x\ell}|^2 - |B_{xu}|^2 + |B_{z\ell}|^2 - |B_{zu}|^2 \right\} \quad , \end{aligned}$$

where ℓ, u refer to the lower and upper sides of the plate.

This result could be written down by inspection as the difference in the magnetic pressures on the two sides of the plate. Under all useful regimes, $|B_z| \ll |B_x|$ at the surface, so that this can be simplified to

$$F_L = \frac{|B_{x\ell}|^2 - |B_{xu}|^2}{8\pi} \quad (1)$$

$$= \frac{1}{4\pi} \operatorname{Re} \bar{B}_x^* \Delta B_x \quad (1a)$$

where $\Delta B_x \equiv (B_{x\ell} - B_{xu})$ measures the net current in the sheet and $\bar{B}_x \equiv (B_{x\ell} + B_{xu})/2$ is the appropriate average field.

Similarly, the drag force is

$$F_D = \left(\frac{1}{4\pi}\right) \operatorname{Re} \int \left(\frac{\partial B_x^*}{\partial z} + ikB_z^* \right) B_z dz \quad ,$$

The second term is purely imaginary and drops out. Using $\nabla \cdot \vec{B} = 0$ and the skin effect equation

$$\frac{d^2 B}{dz^2} = \left(\nabla^2 - \frac{\partial^2}{\partial x^2} \right) B = \left(\frac{2i}{\delta^2} + k^2 \right) B \quad ,$$

where

$$\delta^2 = c^2/2\pi\sigma\omega = c^2/2\pi\sigma kv \quad ,$$

the first term can be transformed so that, after dropping an imaginary term and dropping the minus sign to get the magnitude of the force, we have

$$\begin{aligned}
 F_D &= \frac{1}{2\pi k \delta^2} \int |B_z|^2 dz \\
 &= \frac{\sigma v}{c^2} \int |B_z|^2 dz \quad (2)
 \end{aligned}$$

$$= \frac{1}{2\pi} \frac{v}{v_0} \langle |B_z|^2 \rangle_{av} \quad (2a)$$

where $v_0 = c^2/2\pi\sigma d$ is the characteristic velocity (~15 mph for a 1/4" aluminum plate) introduced in our previous report. This result can be obtained more simply (but with loss of sign information) by considering the power dissipation $\int \sigma E_y^2 dz$, where $E_y = (v/c) B_z$, and equating it to $F_D v$.

The results (1) and (2) are general formal results, not limited to the thin sheet case.

Determination of Fields

The fields in the sheet are determined by matching them to the external fields at the boundaries. These external fields, being periodic in x , are exponential in z , and have the general form

$$B = B_+ e^{kz} + B_- e^{-kz} .$$

The magnet currents uniquely determine the strength of the term which decays toward the sheet. This, in turn, fixes a linear combination of the field and its derivative. For example,

$$B + \frac{1}{k} \frac{dB}{dz} = 2B_+ e^{kz} .$$

If we specify the strength of the source in terms of the field

B_0 at the height h of the magnet above the sheet, then $B_+ e^{kh} = B_0$. Thus, the fields just outside the upper and lower surfaces (at $z = \pm d/2$) must satisfy

$$\begin{aligned} B + \frac{1}{k} \frac{dB}{dz} \Big|_u &= 2B_{0u} e^{-kh_u} \equiv b_u, \\ B - \frac{1}{k} \frac{dB}{dz} \Big|_l &= 2B_{0l} e^{-kh_l} \equiv b_l. \end{aligned} \quad (3)$$

Using $\text{div } \vec{B} = 0$ and $\text{curl } \vec{B} = 0$, $B'_z = -ikB_x$ and $B'_x = ikB_z$. Eliminating the normal derivatives in this way, the boundary conditions become

$$\begin{aligned} B_{xu} + iB_{zu} &= b_u, \\ B_{xl} - iB_{zl} &= b_l. \end{aligned} \quad (3a)$$

The conditions (3a) are more convenient than (3), since \vec{B} is continuous across the boundary, whereas dB_x/dz is not.

We now must find the field inside the sheet consistent with these boundary conditions. Both components of this field obey the equation

$$\nabla^2 B = \frac{2i}{\delta^2} B,$$

so that

$$\frac{d^2 B}{dz^2} = \left(\frac{2i}{\delta^2} + k^2 \right) B \approx \frac{2i}{\delta^2} B, \quad (4)$$

since $(k\delta)^2 \approx (2\pi\delta/\ell_x)^2$ is typically ≤ 18 . The solutions of

this second-order differential equation have two adjustable constants, which can be taken to be the value and slope at $z = 0$, the midplane of the sheet. Then the solutions have the form

$$\begin{aligned}
 U(z) = & U(0) \left[1 + \frac{1}{2} \frac{z^2}{\delta^2} - \frac{z^4}{6\delta^4} + \dots \right. \\
 & \left. + U'(0)z \left[1 + \frac{1}{3} \frac{z^3}{\delta^3} - \frac{1}{30} \frac{z^5}{\delta^5} + \dots \right] \right.
 \end{aligned} \tag{5}$$

Moreover, the Maxwell equations relate the x and z components as follows:

$$B'_x = \left(\frac{2v}{v_0} + ikd \right) \frac{B_z}{d} = \frac{2v}{v_0 d} B_z \quad , \tag{6a}$$

$$B'_z = -ik B_x \quad , \tag{6b}$$

where again we have dropped kd compared to v/v_0 , since for useful ranges of parameter values, $kdv_0/v < 0.01$. Thus, the internal fields are completely specified by any suitable pair of parameters among the set $B_x(0)$, $B'_x(0)$, $B_z(0)$, $B'_z(0)$.

In applying the boundary condition (3a) to determine these parameters, it is convenient to form the sum and difference:

$$\begin{aligned}
 b_u + b_l &= (B_{xu} + B_{xl}) + i(B_{zu} - B_{zl}) \quad , \\
 b_u - b_l &= (B_{xu} - B_{xl}) + i(B_{zu} + B_{zl}) \quad ,
 \end{aligned} \tag{7}$$

or, using (5) and carrying only the leading terms in d^2/δ^2

$$b_u + b_l = 2B_x(0) \left(1 + \frac{1}{4} \frac{d^2}{\delta^2}\right) + 11 \frac{d}{2} \left(1 + \frac{1}{12} \frac{d^2}{\delta^2}\right),$$

$$b_u - b_l = B_x'(0) d \left(1 + \frac{1}{12} \frac{d^2}{\delta^2}\right) + 2i B_z(0) \left(1 + \frac{1}{4} \frac{d^2}{\delta^2}\right).$$

Replacing the derivatives using (6), this becomes

$$\begin{aligned} b_u + b_l &= 2B_x(0) \left(1 + \frac{1}{4} \frac{d^2}{\delta^2}\right) + kd B_x(0) \left(1 + \frac{1}{12} \frac{d^2}{\delta^2}\right), \\ &= 2B_x(0) \left(1 + \frac{1}{4} \frac{d^2}{\delta^2}\right), \end{aligned} \quad (8a)$$

since $kd \ll 1$. Similarly, one finds

$$b_u - b_l = 2B_z(0) \frac{v}{v_0} \left(1 + \frac{1}{12} \frac{d^2}{\delta^2}\right) + \left(1 - \frac{d^2}{4\delta^2}\right). \quad (8b)$$

Now, if we drop the correction terms in d^2/δ^2 , (8) leads to the simple results:

$$\begin{aligned} B_x(0) &= \frac{b_u + b_l}{2}, \\ B_z(0) &= \frac{b_u - b_l}{2(v/v_0 + 1)}, \end{aligned} \quad (9)$$

which completely determine the field (and hence the force) in terms of the source strengths b_u , b_l . Since detailed examination of the effects of the d^2/δ^2 terms shows that they lead to corrections in the forces which are typically $\leq d^2/10\delta^2$, we expect these simple results to be quite reliable so long as $d \leq \delta$. Although the detailed formulation given here is needed to investigate how good the approximation is, these simple results can be obtained by considerably more elementary means.

Computation of Forces

Using (1) and carrying only leading terms, the lift force F_L can be written

$$\begin{aligned}
 F_L &= \frac{1}{4\pi} \operatorname{Re} \bar{B}_x^* \Delta B_x = \frac{d}{4\pi} \operatorname{Re} B_x(0)^* B_x'(0) \\
 &= \frac{d}{4\pi} \left(\frac{2v}{v_0 d} \right) \operatorname{Re} B_x^*(0) B_z(0) \\
 &= \frac{v}{8\pi v_0} \operatorname{Re} (b_x + b_l)^* \frac{(b_u - b_l)}{v/v_0 + 1} \\
 &= \frac{b_u^2 - b_l^2}{8\pi} \frac{1}{1 + v_0^2/v^2} \quad ,
 \end{aligned} \tag{10}$$

(The change in sign convention relative to (1) arises since we want the lift force on the magnet structure, not the track.) We take b_u and b_l as real, which can be done so long as the upper and lower magnet structures are not relatively displaced along the x-axis. In view of the definitions of b_u and b_l in (3), we see that this reduces to our earlier result for the single-sided case:

$$F_L \Big|_{\substack{\text{single-} \\ \text{sided}}} = \frac{B_0^2}{2\pi} \frac{e^{-2kh}}{1 + v_0^2/v^2} = \frac{F_m}{1 + v_0^2/v^2} \tag{11}$$

(The old formula had $B_0^2/4\pi$ rather than $B_0^2/2\pi$ because B_0 was a peak value; here it is an RMS value.)

Now let us consider instead of a symmetric arrangement, one with displacement Δh from the symmetry plane, the configuration which provided the motivation for this analysis. Then

$$\begin{aligned}
b_u &= 2B_0 e^{-k(h_0 - \Delta h)} , \\
b_l &= 2B_0 e^{-k(h_0 + \Delta h)} ,
\end{aligned}
\tag{12}$$

so that

$$\begin{aligned}
F_L &= \frac{B_0^2}{\pi} e^{-2kh_0} \frac{\text{sinc } 2k\Delta h}{1+v_0^2/v^2} \approx \frac{2}{\pi} B_0^2 e^{-2kh_0} \frac{k\Delta h}{1+v_0^2/v^2} \\
&\equiv F_\infty \frac{k\Delta h}{1+v_0^2/v^2}
\end{aligned}
\tag{13}$$

since $k\Delta h$ will be small in reasonable applications. As expected in a null-flux geometry, the lift force increases linearly with the displacement from the symmetry plane.

Using (2), the drag force is

$$\begin{aligned}
F_D &= \frac{1}{2\pi} \frac{v}{v_0} \langle |B_z|^2 \rangle \approx \frac{1}{2\pi} \frac{v}{v_0} \frac{1}{d} \int_{-d/2}^{d/2} |B_z(0) + B_z'(0)|^2 dz \\
&= \frac{1}{2\pi} \frac{v}{v_0} |B_z(0)|^2 + |B_z'(0)|^2 d^2/12 \\
&= \frac{1}{2\pi} \frac{v}{v_0} |B_z(0)|^2 + \frac{k^2 d^2}{12} |B_x(0)|^2 ,
\end{aligned}
\tag{14}$$

$$\text{so } F_D = \frac{1}{8\pi} \frac{v}{v_0} \frac{(b_u - b_l)^2}{1+v_0^2/v^2} + \frac{k^2 d^2}{12} (b_u + b_l)^2 .
\tag{15}$$

Again, this reduces properly to the usual single-sided result if one sets $b_l = 0$ and drops the negligible term in $k^2 d^2$. But for the nearly symmetric case, this term cannot be dropped since it provides irreducible minimum drag even when $b_u = b_l$.

Using (12), the drag force becomes

$$F_D = \frac{2}{\pi} B_0^2 e^{-2kh_0} \frac{v}{v_0} \frac{\sinh^2 k\Delta h}{1+v^2/v_0^2} + \frac{k^2 d^2}{12} \cosh^2 k\Delta h$$

$$\approx \frac{2}{\pi} B_0^2 e^{-2kh_0} \frac{v}{v_0} \frac{k^2 (\Delta h)^2}{1+v^2/v_0^2} + \frac{k^2 d^2}{12} \quad , \quad (16)$$

for small displacements Δh .

To carry the analysis further, it is convenient to examine the drag normalized to the lift, which will of necessity have the constant value Mg , where M is the mass/area of the train. (This is just the reciprocal of the more commonly used lift-to-drag ratio.) Combining (13) and (16), we have

$$\frac{F_D}{F_L} = \frac{v}{v_0} \frac{k^2 (\Delta h)^2}{1+v^2/v_0^2} + \frac{k^2 d^2}{12} \frac{1+v_0^2/v^2}{k\Delta h}$$

$$= k\Delta h \left[1 + \frac{d^2}{12(\Delta h)^2} \frac{v_0}{v} + \frac{kd^2}{12\Delta h} \frac{v}{v_0} \right] \quad .$$

In the useful regime where $v > v_0$ and $d \lesssim \Delta h$, the correction term $d^2/12(\Delta h)^2$ is always small. Thus we can write

$$\frac{F_D}{F_L} = k\Delta h \frac{v_0}{v} + \frac{kd^2}{12\Delta h} \frac{v}{v_0} \quad . \quad (17)$$

This normalized drag has a minimum value

$$(F_D/F_L)_{\min} = \frac{kd}{\sqrt{3}} = \frac{2\pi}{\sqrt{3}} \frac{d}{\ell_x} \sim \frac{1}{80} \quad , \quad (18)$$

at a characteristic velocity

$$v = 2\sqrt{3} v_0 \frac{\Delta h}{d} = \frac{\sqrt{3}}{\pi} \frac{c^2 \Delta h}{\sigma d^2} \sim 50 \text{ m.p.h.} \quad (19)$$

(The representative values quoted are for a 1/4" aluminum sheet, a magnet period of 6 feet, and $\Delta h \approx 1/4"$.) Thus, (17) has the normalized form

$$\frac{F_D/F_L}{(F_D/F_L)_{\min}} = \frac{1}{2} \left(\frac{v}{V} + \frac{V}{v} \right) \quad (20)$$

Note that the drag is within 25% of its minimum value over the velocity range from $V/2$ to $2V$. Thus, if $2V$ were chosen to be the full speed of the train, say 300 mph, then the drag from 75 mph to 300 mph would be nearly constant and well below the air resistance over most of this range of velocity. Even down at 30 mph, where $v \approx 2v_0$, so that F_d has reached 80% of its high-speed value, allowing the magnetic suspension to take over from the wheels, the drag would be only about 3 times its minimum value, giving a very acceptable lift/drag ratio of about 25. On the other hand, the lower value of V in the above example has the advantage of lowering the magnetic drag at low velocities (where it is the dominant drag), while allowing it to rise at higher velocities, where it will still be dominated by air drag. This is illustrated in Figure 1.

Magnet Requirements

The above very favorable estimates hinge on the assumption that the magnet strength required by $\Delta h \approx 1/4"$ is reasonable. This is determined by the relation (13), which, for $v \gg v_0$, can be written

$$Mg = F_L = F_w = k\Delta h \quad , \quad (21)$$

where

$$F_w = (2/\pi) B_0^2 e^{-2kh_0} \quad . \quad (21a)$$

In this, B_0 is the rms strength of the fundamental Fourier component of the x (or z) component of the field in the magnet plane, and h_0 is half the separation of the upper and lower magnet conductors. B_0 , in turn, is $(2\pi/c) \mathcal{J}_0$, where \mathcal{J}_0 is the rms value of the fundamental Fourier component of the magnet current. That is

$$B_0 = \frac{2\pi}{c} \mathcal{J}_0 = \frac{2}{\sqrt{2}} \frac{2\pi}{c} \frac{1}{l_x} \int_0^{l_x} \mathcal{J}_y(x) e^{-ikx} dx \quad . \quad (22)$$

(The curious factor $2/\sqrt{2}$ is inserted to take account of the existence of equal components $-e^{ikx}$ in the real field, and to convert to rms values.) A simple example is a sequence of magnet loops ($l_x/2$) long, producing alternately up and down fields. If the current in each is I_0 , the transverse currents add, and are equivalent to currents $\pm 2I_0$ spaced $l_x/2$ apart.

Then

$$B_0 = \frac{8\sqrt{2}}{c} \pi \frac{I_0}{l_x} = 4\sqrt{2} k I_0 / c \quad . \quad (23)$$

Thus B_0 is simply proportional to the transverse magnet current per unit length of train. From (21) we have

$$\Delta h = \frac{Mg}{kF_w} = \frac{\pi}{2} \frac{Mg e^{-2kh_0}}{B_0^2 k} \quad . \quad (24)$$

Holding B_0 (i.e., the average current density along the train) constant while varying k , the minimum Δh (i.e., the strongest lift) is found for

$$k = \frac{1}{2}h_0 \quad \text{or} \quad l_x = 4\pi h_0 \quad , \quad (25)$$

for which case

$$\frac{\Delta h}{h_0} = \pi e \frac{Mg}{B_0^2} \quad . \quad (26)$$

It is worth noting that for this value of k , the fundamental Fourier component accounts for about 87% of the total lift. Hence the single Fourier component approximation should be quite good, even for this simple current pattern, so long as l_x is not greater than $4\pi h_0$. If we separated the successive loops by $l_x/6$ (as outlined in the appendix of our earlier report) to cancel the 3rd harmonic content, the fundamental component would give about 98% of the lift at this spacing ($l_x = 4\pi h_0$). If we now take a reasonable magnet current $I_0 = 300,000$ amperes, $B_0 = 35,000$ Gauss/ l_x (feet) = 6,000 Gauss. Assume a weight of 100,000 lbs. supported by a magnet structure 100 feet long and 3 feet wide; then $Mg = 160,000$ dynes/cm². Using (26), we find

$$\frac{\Delta h}{h_0} = 0.04 \quad , \quad (27)$$

so that $\Delta h = 0.04h_0 = 0.04l_x/4\pi = 0.2"$, very near our design target. Thus, there seems to be no difficulty in finding a

reasonable combination of parameter values.

Optimization of Parameters

Now let us explore the possibilities for optimizing the parameter choices, subject to the constraints of the problem. These include the following: the minimum drag velocity V should have a set value in the range 50-150 mph. Since $V = \Delta h / \sigma d^2$, this combination is thus constrained. For given B_0 and Mg , $\Delta h = \frac{\pi}{2} \frac{Mg}{B_0^2} \frac{e^{2kh_0}}{k}$. To avoid excess harmonic content, with its higher drag/lift, we require $\frac{2\pi}{k} = \lambda_x \leq 4\pi h_0$, the value giving maximum lift for a given B_0 . Subject to these constraints, we want to minimize $(F_D/F_L)|_{\min} = \frac{kd}{\sqrt{3}}$ by varying d , σ , h_0 , Δh .

Making appropriate substitution, we have

$$\left(\frac{F_D}{F_L}\right)_{\min} = \frac{kd}{\sqrt{3}} = k \left(\frac{c^2 \Delta h}{\sqrt{3}\pi\sigma V}\right)^{\frac{1}{2}} = \left(\frac{Mg c^2 k e^{2kh_0}}{B_0^2 2\sqrt{3} \sigma V}\right)^{\frac{1}{2}} \quad (28)$$

Clearly this would be minimized by making k as small as possible. But if $k < 1/(2h_0)$, the harmonic components of the field increase the drag, and also the lift for given B_0 falls off. Thus, we are limited by that value, which also has the desirable property of giving maximum lift. Thus:

$$\left(\frac{F_D}{F_L}\right)_{\min} \geq \left(\frac{Mg e c^2}{B_0^2 4\sqrt{3} h_0 \sigma V}\right)^{\frac{1}{2}}$$

According to this, one could lower the drag indefinitely by

increasing h_0 . But this is limited by the fact that we have neglected the effect of the finite width of the magnet structure, which limits the validity of our one-dimensional solution. As shown in the Appendix of this report, however, the one-dimensional solution is quite a good approximation so long as $w \geq l_x/2 \approx 2\pi h_0$, where w is the width of the magnet in the y -direction. Thus, we want to have $h_0 \leq w/2\pi$. The width also affects the weight Mg per unit area of magnet, since $M'g = wMg$, the weight per unit length of train, is really the fixed quantity. Using these considerations, (29) becomes

$$\left. \frac{F_D}{F_L} \right|_{\min} \geq \left(\frac{\pi e}{2\sqrt{3}} \frac{M'g c^2}{B_0^2 \sigma V w^2} \right)^{1/2} \quad (30)$$

If w is limited to ≈ 3 feet, the values used in our numerical estimates ($l_x = 6$ feet, $h_0 = 6$ inches, $\Delta h = 0.25$ inches, $V = 50$ mph, σ (aluminum) effectively reaches this absolute limit for any given $(M'g/B_0^2)$.

Summarizing the optimum design relations:

$$l_x/2 \approx 2\pi h_0 \approx w \quad , \quad (31)$$

so the coils are approximately square. With these relations,

$$\Delta h \approx \frac{e}{2} \frac{M'g}{B_0^2} \quad . \quad (32)$$

Of course, considerations of overall system optimization need not coincide with minimizing the magnetic drag alone. For

example, efficient synchronous propulsion favors a slightly smaller value of λ_x .

Linear Synchronous Propulsion

We can take advantage of the powerful magnets in the train to make a simple, efficient linear synchronous drive system using an active track structure to produce a traveling magnetic field which pulls the train magnets along. The phase velocity v_ϕ of the field is $v\lambda_x$, where v is the frequency of the field and λ_x is its spatial period (the same as that of the train magnets). If we restrict attention for simplicity to a single-phase drive, the track winding could have the form indicated in Figure 2. A current I makes a complete circuit, out and back, from a source at one end.

For a specific example, let us assume $\lambda_x = 2w$, so the loops are square. In that case, the total length of conductor is $4x$ the length of track. If the conductor were $1" \times 1"$ square aluminum bus bar, for example, the round trip resistance would be $\sim 0.25\Omega$ /mile of track. For comparison, a typical effective train resistance is 1.5Ω (corresponding to $0.1g$ drive force at 300 mph, which requires 6MW, and assuming a drive current of 2000 amperes.) Thus, with such a heavy conductor, track losses would be less than the useful power even in a track segment extending for 5 miles in either direction from a power source. The track inductance will be ~ 0.006 H/mile, which gives a reactance of $\sim 2.5\Omega$ /mile at 60 Hz, a representative drive

frequency (corresponding to 250 mph and $\ell_x = 6$ feet). Thus, in a track run of a mile or more, the inductive reactance would dominate the impedance seen by the source despite the load due to the train, giving a poor power factor and requiring a high source voltage. These considerations might make relatively short runs, using thinner drive loop conductor, supplied with power by high impedance transmission lines through step-down transformers, and of course only switched on when a train is passing over the section, more economical in terms of total costs. In any case, it seems almost certain that the track inductance will effectively filter out any harmonic content in the output of the variable-frequency power source, so that the drive current may be assumed to be a pure sine wave.

The instantaneous drive force on the train is equal and opposite to the total force exerted on the track loop conductors by the field of the train magnets. Because the spatial period of track currents and train field is the same, the force on each track conductor is the same. Thus, we may restrict attention to a single one, of length w , carrying a current I in the y -direction. Since the local drive force is $F_x = I B_z/c$, the total force per conductor is $(I/c) \int B_z dy = I w \bar{B}_z/c$, where \bar{B}_z is the spatial average of the field of the train magnet along the track conductor. Since $I(t)$ is a pure sine wave, only the fundamental component in the Fourier analysis of $\bar{B}_z(t)$ (which arises from the fundamental spatial Fourier component) gives a

non-zero time average in the product. Thus, we can write the time-averaged propulsion force per conductor as

$$F_p = \frac{w}{c} I_{rms}(\omega) \bar{B}_z(\omega) \cos \phi ,$$

where ϕ is the phase angle between the current and field, and $\omega = kv = 2\pi v/\lambda_x$ for synchronous operation.

In our simple Fourier analysis of the fields, we have no variation of B_z with y , so that RMS values are given by

$$\bar{B}_z = B_z = B_0 e^{-kh} = 4\sqrt{2} \frac{kI_0}{c} e^{-kh} , \quad (34)$$

where B_0 specifies the train magnet strength in its own plane and I_0 is the current in the train magnet loops (as discussed earlier) and h is the separation of the magnet plane from the plane of the track currents. Thus, the average propulsive force per track conductor is

$$F_p = 4 \frac{\sqrt{2} w k e^{-kh}}{c^2} I_0 I_{rms} \cos \phi . \quad (35)$$

This result may appear unfamiliar because of the exponential dependence arising from the Fourier analysis. However, if one takes the separation h to be the same as the $h_0 = 1/2k$, which was optimum for the levitation case, then the factor $ke^{-kh} = \frac{e^{-1/2}}{2h}$, and one recovers the more familiar form $F_p \sim I_0 I_{rms} w/h$.

Since the actual conductor array in the train magnets does not extend to $y = \pm\infty$, as assumed in the simple model, but rather consists of approximately square loops, one might question the quantitative reliability of (34) as an expression

for the strength of the fundamental Fourier component of \bar{B}_z . To test this, we have developed an exact formula (see Appendix) for \bar{B}_z from a rectangular current loop of the same width w , and we have numerically estimated the strength of the fundamental and third harmonic Fourier components for the special case of a series of square loops. The results indicate that the Fourier coefficients computed for these two different current patterns differ by only about 10%.

Some insight into this result is given by the following observations: The field in the center of a square loop of edge w carrying current I is $8\sqrt{2} I/cw$ whereas the field midway between two infinite straight wires separated by w and carrying I is only $8 I/cw$. Thus, closing the loop gives a higher field than extending the parallel wires to infinity. On the other hand, if we take account of the presence of two wires on each side (because of the contiguous loops), the parallel wires given $16 I/cw$. Similarly, taking account of the two adjacent loops (by a dipole approximation) in the loop configuration, each contributes $\mu/w^3 = I w^2/cw^3 = I/cw$, for a total of $(8\sqrt{2} + 2) I/cw = 13.3 I/cw$. If we continue both calculations to convergence for an infinite series of parallel wires or square loops, the results are $\sim 12.5 I/cw$ and $\sim 13.1 I/cw$, respectively, again differing by less than 10%.

Returning now to the real case, in which we are interested in the strength of the fundamental Fourier component averaged over the width w in a plane separated from the magnet

plane by a distance h , rather than the value of the field at the single point in the center of a loop, the corresponding numerical results are ~ 15 I/cw for the parallel wires and ~ 17 I/cw for the square loops. [In these estimates we have taken $h = w/6$, a typical value; also, they must be reduced by a factor of $\sqrt{2}$ to correspond to the RMS value (34).] The fact that these coefficients are greater by some 30% than those found above reflects the fact that the third harmonic amplitude is $\sim 1/3$ the fundamental amplitude, and the third harmonic tends to cancel midway between conductors, while adding to the fundamental near either side. This causes the field to sag in the middle, as indicated in Figure 2(b). We conclude that for a typical geometry the simple Fourier analysis result may underestimate the drive force by $\sim 10\%$, but that it is good enough for preliminary estimates.

For these estimates, it is convenient to characterize the propulsive force F_p by an equivalent acceleration g' . If M' is the mass of the train per unit length, then

$$M'g' = F_p \frac{4}{\ell_x} ,$$

or

$$\frac{g'}{g} = \frac{32\pi\sqrt{2}w}{M'g\ell_x^2} e^{-2\pi h/\ell_x} \frac{I_0 I_{rms}}{2} \cos \phi . \quad (36)$$

Taking the reasonable values $M'g = 100,000$ lbs/100 feet, $w = 3$ feet, $\ell_x = 6$ feet, $h = 6$ inches, $I_0 = 300,000$ amperes, this leads to

$$\frac{g'}{g} = 0.5 \times 10^{-4} I_{\text{rms}} (\text{amps}) \cos \phi \quad . \quad (36.1)$$

Thus, a drive current of 2000 amperes would provide up to $\pm 0.1g$, depending on the phase of the train magnets relative to the current in the track. Evidently the drive current requirement could be reduced further by increasing the width (w) or strength of the train magnets (I_0) or by reducing l_x and/or h . For a typical clearance $h = w/2\pi$ between train and track, given train current density (I_0/l_x), and given track dissipation per mile $\sim I_{\text{rms}}^2 (1 + 2w/l_x)$, the spacing l_x giving maximum propulsion is $l_x = \sqrt{2}w = 2\sqrt{2}\pi h$. This is not very different from the condition (25), $l_x = 4\pi h_0$, for maximum lift at separation h_0 between magnets and the conducting plane. Thus, if $h_0 = h$, as is reasonable, the same value of l_x will be close to being optimal for both lift and propulsion.

Transverse Stability

The tight magnetic suspension of the wide magnet coils all along the train will provide a very strong restoring force against uniform vertical motion, pitch, and roll.

A magnetic restoring force against uniform sidewise motion and yaw can be established by vertical conductors along the edges of the conducting sheet track as shown in Figure 3. Since this conductor would not enjoy the benefit of a null-flux configuration, its contribution to the drag will be significant if the restoring force is made very strong. This drag contri-

tribution can be estimated as follows: The force F_y will vary as e^{2ky} . Thus,

$$\frac{1}{F_y} \frac{dF_y}{dy} = 2k \quad .$$

The associated drag force (assuming the vertical conductor is thicker than the skin depth δ) will be given by

$$F_D = k\delta F_y \quad ,$$

drawing on the results of our previous report. Thus

$$\frac{dF_y}{dy} = \frac{2}{\delta} F_D \quad .$$

If we want a restoring force to give an acceleration g' at displacement Δy from equilibrium, we have

$$Mg' = \frac{2}{\delta} F_D \Delta y \quad .$$

Dividing by the weight of the train, and rearranging, we have

$$\frac{F_D}{F_L} = \frac{\delta}{2\Delta y} \frac{g'}{g} \quad . \quad (37)$$

Taking the representative values $\delta = 0.4"$ and $g'/g = 0.2$ for a maximum excursion of $\Delta y = 4"$, this gives $F_D/F_L = 1/100$. Since this drag would double (because δ doubles) on reducing the speed from 300 mph to 75 mph, this drag is not negligible. A careful analysis should be made to find how strong the restoring force must be for proper operation. If the drag from this form of lateral stability scheme is excessive, one could

instead derive the restoring force from a more complex arrangement of null-flux conductors. For example, one could use a pair of magnet arrays, canted up from the horizontal on either side of the center line, to derive a centering force from gravity. Alternatively, one could use a separate null-flux arrangement with vertical coils to provide a transverse restoring force. This scheme has the attractive feature that the drag would be essentially zero in the centered position, and considerably smaller than (37) when a restoring force was generated in an off-center position.

Longitudinal Stability

It is well-known that there is a potential instability associated with a drag force which decreases with increasing velocity. In general, we have

$$M \frac{dv}{dt} = F_p - F_D(v) \quad , \quad (38)$$

where F_p and F_D are the propulsion and drag forces, respectively. If one attempts to maintain a constant velocity v' by setting $F_p = F_D(v')$, we have, for small changes in velocity,

$$M \frac{dv}{dt} = F_p - F_D(v') - F'_D(v')(v-v') \quad .$$

Now, if F_p remains constant, this leads to

$$M d(v-v')/dt = -F'_D(v')(v-v') \quad ,$$

so that

$$(v-v') \sim e^{t/\tau} \quad , \quad (39)$$

where the time constant τ is given by

$$\tau = -M/F'_D(v') \quad . \quad (39a)$$

Thus fluctuations in velocity will grow exponentially if $\tau > 0$; that is, if $F'_D(v') < 0$, as is always the case for the magnetic drag in conventional image-force or null-flux schemes. In our hybrid scheme, however, the magnetic drag increases for $v > V$. When the air drag is added in, all these systems become stable in this sense at sufficiently high velocities. The hybrid scheme is again most favorable, becoming stable somewhat below V , which might be ~ 50 mph. In fact this instability is not very serious because τ is so long. The expression (39a) can be re-written in the form

$$\tau = \frac{v}{g} \frac{F_L}{F_D} \frac{-v}{F_D} \frac{dF_D}{dv}^{-1} \quad , \quad (40)$$

from which we see that at 100 mph, τ is typically 200 seconds. This is slow enough to permit stabilization by a simple feedback system.

More fundamentally, the above analysis is incomplete because it is based on the assumption that the propulsive force F_p remains constant, independent of v . This might be a fair approximation for propulsion by a jet engine, but it certainly is inappropriate for either of the two magnetic drive systems which are under consideration.

In the induction drive system, the propulsive force is of the form

$$F_p = 2F_0 \frac{\Delta v/v_0}{1+(\Delta v/v_0)^2} \quad (41)$$

where F_0 is the maximum propulsive force, which occurs when the slip velocity $\Delta v = v_\psi - v$ relative to the traveling magnetic field has the value $v_0 = c^2/2\pi\sigma d$ (~15 mph for 1/4" aluminum). So long as $\Delta v < v_0$, (41) gives a driving force tending to stabilize the velocity. Instability could result only from fluctuations large enough to decrease v to a value near $(v_\psi - v_0)$, where the propulsive force reaches a maximum, allowing the destabilizing effect of $F'_D < 0$ to reassert itself. By operating at a slip velocity well below v_0 , e.g. with $F_p \sim F_0/2$, such fluctuations can be excluded. Thus, with a suitable choice of v_0 for the reaction rail, one could certainly produce a stable propulsion system using a linear induction motor.

The stability analysis of the linear synchronous drive is more complicated since the propulsive force F_p depends on the instantaneous displacement between the driving field and the train magnets, rather than on their relative velocities. That is, our equation of motion becomes

$$M \frac{dv}{dt} = F_0 \sin k (v_\phi t - x) - F_D(v) \quad (42)$$

In synchronous motion, $v = dx/dt = v_\phi$, and $x = -x_0 + v_\phi t$. Thus, the equilibrium lag x_0 is determined by

$$F_0 \sin kx_0 = F_D(v_\phi) \quad .$$

Expanding about this operating point,

$$\begin{aligned} M \frac{dv}{dt} &= -F_0 \cos k(v_\phi t - x) k \Delta x - F'_D(v_\phi) \Delta v \\ &= -F_0^2 - F_D^2(v_\phi)^{\frac{1}{2}} k \Delta x - F'_D(v_\phi) \Delta v \quad . \end{aligned}$$

If we now assume that Δx and Δv vary as $e^{i\omega t}$, this requires that

$$-M\omega^2 + iF'_D\omega + k(F_0^2 - F_D^2)^{\frac{1}{2}} = 0 \quad ,$$

or

$$\omega = -\frac{i}{2\tau} \pm \frac{k}{M}(F_0^2 - F_D^2)^{\frac{1}{2}} - \frac{1}{4\tau^2} \quad , \quad (43)$$

where $\tau = -M/F'_D(v_\phi)$ was defined in (39a). Note that if $F_0 = F_D'$, so that the propulsive force is constant at its maximum value, this reduces to the exponential solution (39), which describes an instability if $F'_D(v') < 0$. In normal operation, however, $F_D \sim F_0/2$, and ω becomes predominantly real. It is convenient to define a natural frequency of longitudinal oscillation by

$$\omega_0^2 = F_0 k/M = g_p k = 2\pi g_p / l_x \quad , \quad (44)$$

where $g_p = F_0/M$ describes the maximum acceleration capability of the propulsion. For the typical values $l_x = 6$ feet and $g_p = 0.2g$, $(\omega_0/2\pi) = 0.4$ Hz. Since $\omega_0 \gg 1/\tau$, we can approximate (43) by

$$\omega = 2\omega_0 [1 - (F_D/F_0)^2]^{1/2} - i/2\tau \quad . \quad (45)$$

Thus, if $F'_D(v_\phi) < 0$ so that $\tau > 0$, there will still be an instability, but in the form of an exponentially growing oscillatory motion.

Using the parameter estimates above, the amplitude would increase a factor of e in ~ 7 minutes, during which some 1000 oscillations about the equilibrium displacement x_0 would occur. Again, dynamic feedback could be used to damp out such a weak instability. However, there is the following automatic stabilizing effect from the induction force which arises if v departs from its synchronous value v_ϕ . (It is the analog of the use of "damper" windings in a rotary synchronous motor.) For small departures from v_ϕ , (41) becomes

$$F_p = -2F_I(v-v_\phi)/v_0 \quad , \quad (46)$$

where F_I is the maximum induction drive force from the eddy currents "incidentally" induced by the track coils in normal metal surrounding the superconducting train magnets. When (46) is added to the right member of (42), it replaces $-F_D(v_\phi)$ by $[-F'_D(v_\phi) - 2F_I/v_0]$ in the equation of motion about the operating point. Thus, the exponential will become a decaying rather than growing one provided that

$$F_I > -F'_D(v_\phi)v_0/2 \quad . \quad (47)$$

If one estimates $F_I \approx F_0(I_{\text{track}}/I_{\text{train}}) \approx F_0/100$, and

$-F_D'(v_\phi) \approx F_D/2v \approx F_O/4v$, this leads to the requirement that $v_o < v/12$. Since $v_o = 4$ mph for 1" aluminum plate, we see that this inductive stabilization effect could be quite substantial. The only troublesome regime will be the low velocity one, in which some sort of wheeled support would be used in any case, so that an analysis of specific systems would be required.

In summary, after taking account of air drag and the induction effect, the linear synchronous drive will be stable at most velocities, particularly with the hybrid levitation system, and any instability would be so weak that it could easily be cured by a simple feedback system.

Acceleration to Speed

In addition to the question of stability about a desired velocity, just discussed, there is the question of intentional acceleration and deceleration to change speed. This is relatively simple in the case of the induction motor, since there is substantial propulsion over a wide range of slip velocities. Thus, amplitude control could supplement frequency control to a considerable degree in controlling acceleration.

The problem is more severe with the linear synchronous motor, a reflection of the notoriously low starting torques of ordinary synchronous motors. This means that a variable frequency source to vary the synchronous speed $v_\phi = v\ell_x$ will be essential, at least during planned accelerations, and probably throughout the track for reasons of safety and con-

venience in unexpected situations. So long as exact synchronism is maintained, the full drive force given by (35) and (36) is available. But, how large is the velocity range from which the drive can pull the train into synchronism?

This can be estimated by using the equation of motion (42). (The inductive term (46) is negligible.) If $v \neq v_\phi$, the synchronous drive term averages to zero, except for the change in v induced by the force during a single cycle of slip. This change from the average v can be estimated by integrating (42) over the time interval $\Delta t = \ell_x / 4(v_\phi - v)$, with the result

$$\Delta v = \frac{F_o}{M} \frac{\ell_x}{2\pi} \left(1 - \frac{\pi}{2} \frac{F_D}{F_o} \right) \frac{1}{(v_\phi - v)} \quad . \quad (48)$$

Equating this to $(v_\phi - v)$, we get the condition for synchronization:

$$(v_\phi - v)^2 \leq \frac{F_o}{M} \frac{\ell_x}{2\pi} \left(1 - \frac{\pi}{2} \frac{F_D}{F_o} \right) \quad . \quad (49)$$

If we neglect drag, and set $F_o = Mg'$, we have

$$(v_\phi - v) \leq \left(\frac{g' \ell_x}{2\pi} \right)^{1/2} \sim 2 \text{ mph} \quad , \quad (49a)$$

for typical parameters. Unless this estimate is far from the mark, it will be necessary to maintain very fine control of v_ϕ (and hence of the frequency) to avoid going out of synchron-

ization velocity. This appears to be the major problem with the synchronous drive system.

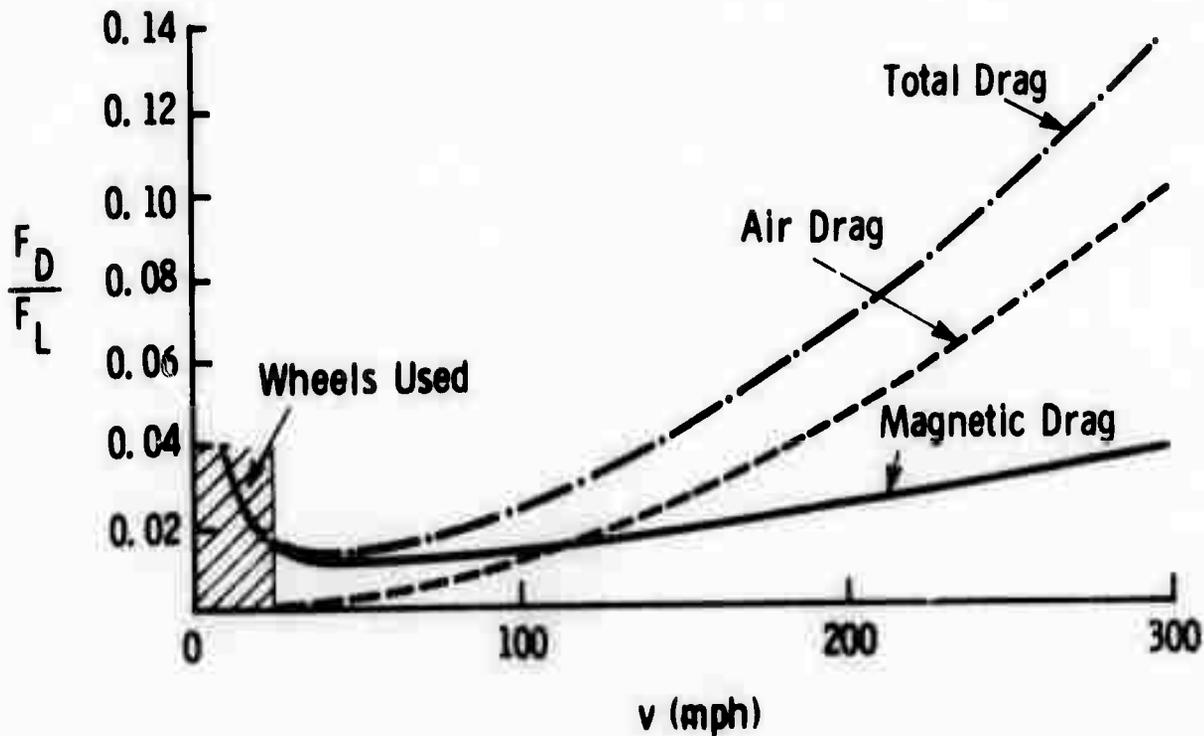
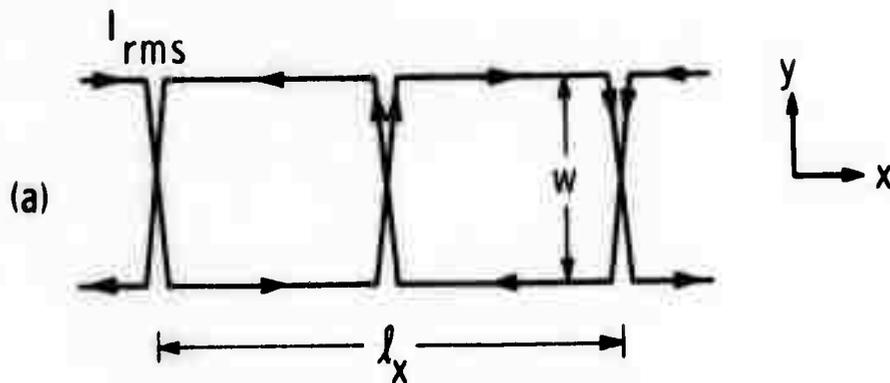
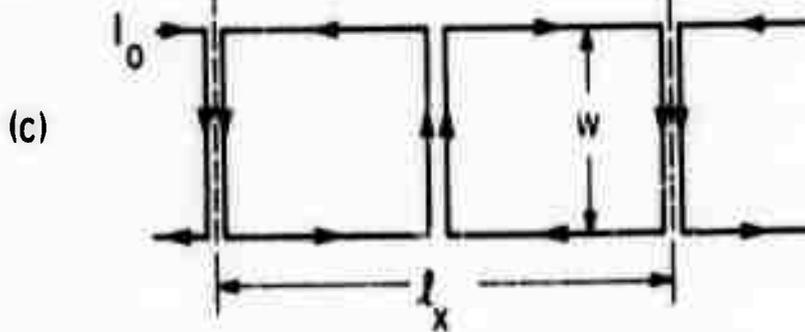
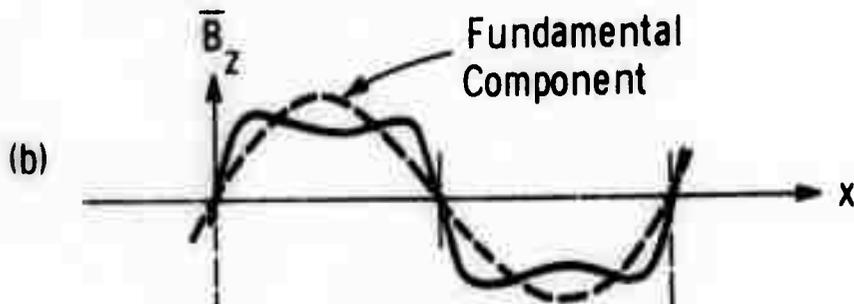


Figure 1. Velocity dependence of the magnetic drag in the hybrid system, compared with air drag. For this plot we have assumed $F_D/F_L|_{\min} = 2\pi d/\sqrt{3}\ell_x = 1/80$ and $V = \sqrt{3} c^2 \Delta h/\pi\sigma d^2 = 50$ mph, which are appropriate for an aluminum plate of thickness $d = 1/4$ ", a 6 foot magnet period ℓ_x , with the magnets strong enough to hold the sag of the train Δh to $1/4$ ". The air drag is assumed to rise as v^2 and reach 10% of the weight of the train at 300 mph.



THRUST LOOP CONFIGURATION



TRAIN MAGNET CONFIGURATION

Figure 2. (a) Configuration of thrust loop conductors for single-phase synchronous propulsion. (b) Spatial variation of \bar{B}_z , the average of B_z over width w , in the plane of the thrust loops due to the train loop currents. Note the rather square waveform due to the high 3rd harmonic content. (c) Train magnet current configuration which produces the field shown in (b).

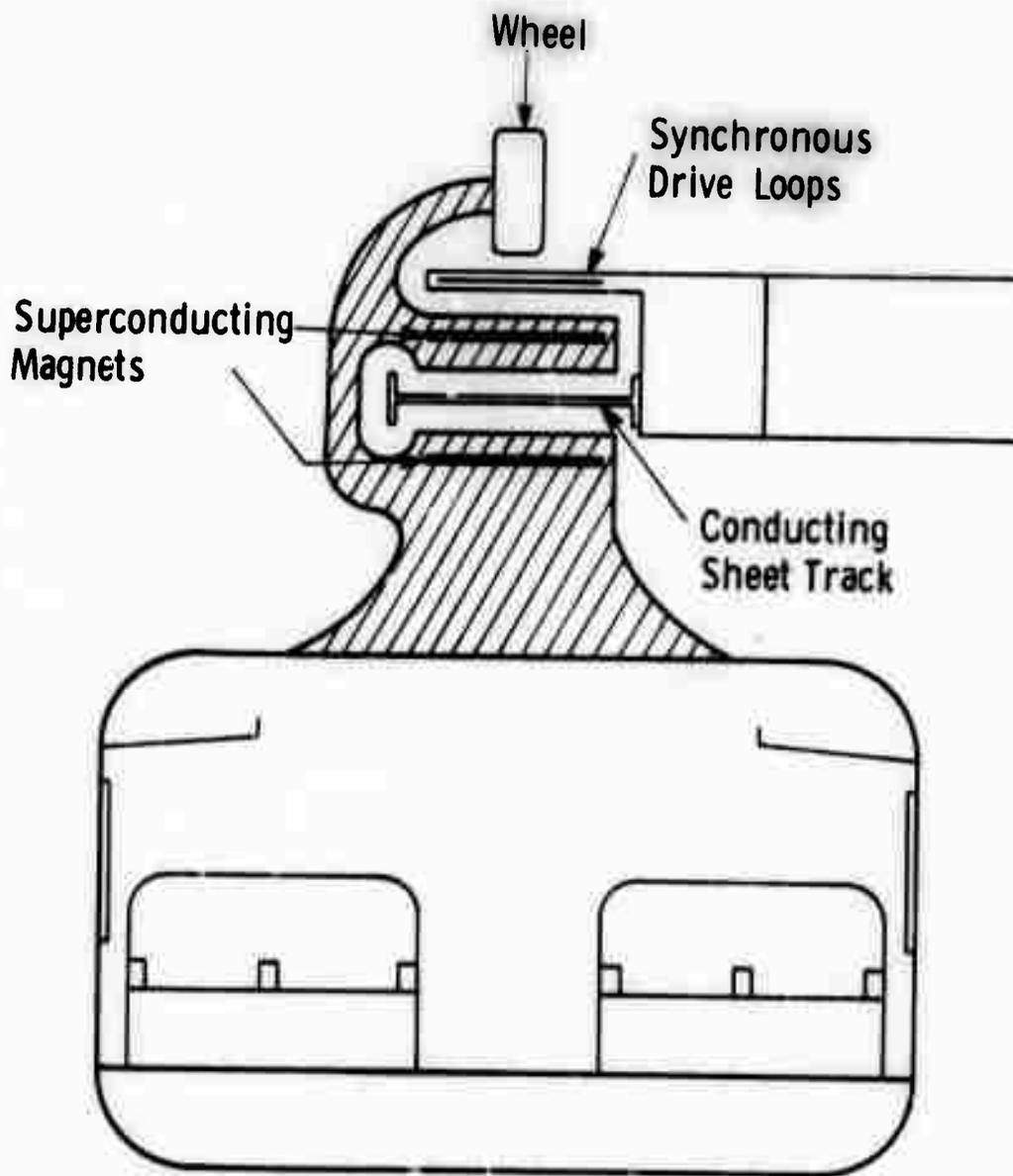


Figure 3. Schematic diagram of a possible train configuration using the hybrid system for levitation and propulsion. The conducting sheet track might be 1/4" aluminum supported mechanically by a strong non-conducting material such as fiberglass-epoxy. The vertical conductors on either side of it provide transverse stabilization forces (see text). Possible configurations for the magnets and drive loops are shown in Figure 2.

Appendix

Calculation of \bar{B}_z for a Sequence of Rectangular Magnet Loops

We first consider a single rectangular magnet loop bounded by $y = 0, w$ and $x = x_1, x_2$, where $x_2 = x_1 + l_x/2$, and carrying a current I in a counter-clockwise sense. We wish to compute \bar{B}_z , where the average is taken between $y = 0$ and $y = w$ along the line specified by $x = 0$ and $z = h$. This is done by first computing $B_z(y')$

$$B_z(y') = \oint dB_z = \frac{I}{c} \oint \frac{(d\vec{s} \times \vec{R}) \cdot \hat{z}}{R^3} \quad , \quad (A1)$$

where $\vec{R} = -x \hat{x} + (y'-y)\hat{y} + h\hat{z}$ is the vector distance from $d\vec{s}$ to $(0, y', h)$. Since $d\vec{s}$ is either along \hat{x} or \hat{y} , it is convenient to rearrange the triple product to $(\hat{z} \times d\vec{s}) \cdot \vec{R}$, so that only $R_y dx$ or $-R_x dy$ appears in the numerator. Thus, the contribution of the loop at x_1 is

$$\begin{aligned} [B_z(y')]_1 = (I/c) & \int_{x_1}^{x_2} \frac{y' dx}{(x^2 + y'^2 + h^2)^{3/2}} + \int_{x_2}^{x_1} \frac{(y'-w) dx}{[x_1^2 + (y'-y)^2 + h^2]^{3/2}} \\ & + \int_0^w \frac{x_2 dy}{[x_2^2 + (y'-y)^2 + h^2]^{3/2}} + \int_w^0 \frac{x_1 dy}{[x_2^2 + (y'-y)^2 + h^2]^{3/2}} \quad . (A2) \end{aligned}$$

The integrations are straightforward. After a second integration over y' to get the average, one finds

$$\begin{aligned}
(\bar{B}_z)_1 &= \frac{I}{cw} \frac{2x_2}{h^2+x_2^2} (h^2+w^2+x_2^2)^{\frac{1}{2}} - (h^2+x_2^2)^{\frac{1}{2}} \\
&\quad - \frac{2x_1}{h^2+x_1^2} (h^2+w^2+x_1^2)^{\frac{1}{2}} - (h^2+x_1^2)^{\frac{1}{2}} \\
&\quad + \ell n \frac{(h^2+w^2+x_1^2)^{\frac{1}{2}}+x_1}{(h^2+w^2+x_1^2)^{\frac{1}{2}}-x_1} \cdot \frac{(h^2+x_1^2)^{\frac{1}{2}}-x_1}{(h^2+x_1^2)^{\frac{1}{2}}+x_1} \\
&\quad \cdot \frac{(h^2+w^2+x_2^2)^{\frac{1}{2}}-x_2}{(h^2+w^2+x_2^2)^{\frac{1}{2}}+x_2} \cdot \frac{(h^2+x_2^2)^{\frac{1}{2}}+x_2}{(h^2+x_2^2)^{\frac{1}{2}}-x_2} \quad . \quad (A3)
\end{aligned}$$

By using this expression, with $x_{n+1} = x_n + \ell_x/2$, we can calculate the contribution to \bar{B}_z of each loop of the train magnet system, and sum them to get the total field. Because of the complexity of (A3), however, it is not convenient to make extensive calculations of that sort. Rather, we note that since the total \bar{B}_z must be periodic with period ℓ_x , it can be described by a Fourier series. Moreover, the symmetry forbids even harmonics. Thus, the dominant terms will be the fundamental and third harmonic terms, which can be characterized by just two numbers. These may be estimated by fitting the total \bar{B}_z computed at two points in the cycle. For this purpose, it is particularly efficient to choose the center of the loop $-x_1 = x_2 = \ell_x/4$ ($k|x_1| = \pi/2$), where the fundamental and third harmonic appear in full amplitude but out of phase, and the

point where $-5x_1 = x_2$ ($k|x_1| = \pi/6$) where the third harmonic adds to the fundamental, which comes in with half amplitude. Mathematically stated, if we neglect all higher harmonics,

$$\bar{B}_z(\pi/2) = \bar{B}_1 - \bar{B}_3 \quad ,$$

$$\bar{B}_z(\pi/6) = \bar{B}_1/2 + \bar{B}_3 \quad .$$

Thus:

$$\bar{B}_1 = \frac{2}{3} \bar{B}_z(\pi/2) + \bar{B}_z(\pi/6) \quad ,$$

$$\bar{B}_3 = \frac{2}{3} \bar{B}_z(\pi/6) - \frac{1}{2}\bar{B}_z(\pi/2) \quad . \quad (A4)$$

What remains is to calculate $\bar{B}_z(\pi/6)$ and $\bar{B}_z(\pi/2)$, taking account of all the loops, not just the central one which includes $x = 0$. Clearly the corrections from the distant loops are small, and one might hope to approximate the effect of each by the field of a dipole $\mu = Iwl_x/2c$. Neglecting the small z -displacement h , the contribution of the n -th loop becomes

$$(\bar{B}_z)_n = \frac{\pm Iwl_x/2c}{[(x_1+x_2)/2 + nl_x/2]^3} \quad . \quad (A5)$$

The \pm sign alternates along the line of \pm loops, the n is a positive or negative integer. For the $\pi/2$ case, $(x_1 + x_2)/2 = 0$, and it is $(l_x/6)$ for the $\pi/6$ case. The reliability of this procedure was tested by computing the exact contribution

of an adjacent loop for the $\pi/2$ case, assuming the representative values $w = \ell_x/2$ and $h = \ell_x/12$ and comparing with the dipole estimate. The results were $1.04 I/cw$ and I/cw , respectively, confirming the satisfactory accuracy of the procedure.

Using (A3) for the central loop, and the dipole approximation, summed over enough loops for convergence, we compute that

$$\begin{aligned}\bar{B}_z(\pi/2) &= 11.6 I/cw & , \\ \bar{B}_z(\pi/6) &= 13.8 I/cw & ,\end{aligned}\tag{A6}$$

where we have retained the square loop geometry $w = \ell_x/2$ and the height $h = \ell_x/12$. In both cases, the non-central loops contribute about 17% of the total. When these results are inserted in (A5), the results are

$$\begin{aligned}\bar{B}_1 &= 16.9 I/cw = 33.8 I/c\ell_x & , \\ \bar{B}_3 &= 5.3 I/cw = 10.6 I/c\ell_x & .\end{aligned}\tag{A7}$$

Since the corresponding results from a simple Fourier analysis of currents running to $y = \pm\infty$ are $B_1 = 30 I/c\ell_x$ and $B_3 = 10.4 I/c\ell_x$ for the same values of h/ℓ_x , but $w \rightarrow \infty$, we may presume that the finite value of w has little effect on the validity of the simple Fourier analysis at least so long as $w \geq \ell_x/2$.

MAGNETIC SUSPENSION AND PROPULSION SYSTEMS
FOR HIGH-SPEED TRANSPORTATION

P. L. Richards
M. Tinkham

Abstract

High speed transportation vehicles (trains) carrying superconducting magnets can be levitated by repulsion from diamagnetic currents induced in a conducting track. Various approximate methods are presented for calculating the lift and drag forces for such magnetic suspensions. Fourier analysis of periodic train magnet fields is used to analyze "image force" and "hybrid null flux" systems which involve homogeneous conducting sheet tracks. A lumped circuit analysis is used to discuss the "null flux" principle and related systems with structured tracks. The stability and efficiency of linear induction and linear synchronous motor propulsion systems are studied using related methods.

MAGNETIC SUSPENSION AND PROPULSION SYSTEMS
FOR HIGH-SPEED TRANSPORTATION

P. L. Richards
M. Tinkham

I. INTRODUCTION

Magnetic suspensions are being seriously considered for high speed ground transportation in several countries. The failure of adhesive drive systems such as steel wheels on steel rails at speeds in excess of 300 km/h has forced consideration of nonconventional approaches for both suspension and drive. Vehicles levitated by magnetic repulsion were first¹ considered in 1912 and discussed for transportation applications in the 1960's.² These are generally not practical without the large ratio of strength to weight available from modern magnets. Although repulsive levitation from a ferromagnetic track may be feasible^{3,4}, we confine our attention to systems in which levitation is obtained from the interaction of induced currents in a conducting (normal state) track with the field from superconducting magnets on the train.⁵ The optimal configuration of train magnets and track is being vigorously debated at the present time.⁵⁻²⁰ In order to select the system for a given application there is need for simple semi-quantitative calculations which will permit comparison of the relevant parameters for widely different geometries. In this paper we discuss various approximate methods for calcu-

lating the efficiency of magnetic suspensions.

In Section II, we Fourier analyze the periodic magnetic field and compute the lift and drag produced on a line of electromagnets above and below a flat conducting sheet as illustrated in Fig. 1. This method of Fourier analysis yields particularly simple physical formulas relevant to designs involving homogeneous sheet tracks. Detailed results are obtained for the case of a single line of train magnets above the track (the image force system). An analysis is then presented of a (hybrid null flux) system which has train magnets both above and below a sheet track that is thin compared with the skin depth. This system combines the homogeneous sheet track with some features of the "null flux" schemes of Powell and Danby,^{12,13} in that both lift and drag forces become small when the track is located at the symmetry plane between the magnets. Section III develops a complementary lumped circuit analysis which is useful for understanding the null flux track illustrated in Fig. 2(b) and (c). When the train magnet is located in the symmetry plane no flux threads the track loop so that lift and drag forces approach zero. Lift forces increase linearly and drag forces as the square of the (small) displacement Δh from the symmetry plane. The analysis developed in the preceding sections is used in Section IV to discuss the stability and efficiency of passive track induction and active track synchronous linear motor drive systems. A portion of this work has been made available previously.²¹

II. FOURIER ANALYSIS OF MAGNETIC SUSPENSION ABOVE A CONDUCTING GROUND PLANE

A. Formulation of the Method

In proposed magnetic suspension systems there is typically a periodic array of "up" and "down" superconducting magnets along the direction of motion. When the train moves over a conducting ground plane, the resulting time-varying fields induce eddy currents which interact with the magnetic field to produce both a lift and a drag force on the train. We present here a treatment which takes advantage of the periodicity of the fields to allow a straightforward solution of the electromagnetic equations over a wide range of conditions.

The geometry of the problem is sketched in Fig. 1, where for generality we have included the possibility of magnet structure both below and above the plane. This will allow us to model the "hybrid null-flux" as well as the "image force" schemes. The x-variation will be given by a sum of terms $\propto e^{ik_x x}$, with $k_x = n_x (2\pi/\ell_x)$, where n_x is restricted to the odd integers if the current distribution is symmetric in x about the center of a magnet coil and antisymmetric about a point midway between an up and down coil. Any variation in the y-direction is taken to have period ℓ_y , but this variation plays a less crucial role because it doesn't cause time-variation with a train velocity along x.

Since displacement current effects may be safely neglected, the magnetic field \vec{B} satisfies Laplace's equation

$\nabla^2 \vec{B} = 0$ in the non-conducting regions I, II, IV, and V of Fig. 1. Given a periodic variation in the xy plane as $e^{i(k_x x + k_y y)}$, this equation then requires real exponential variation with z, namely,

$$\vec{B} \propto e^{\pm kz} e^{i(k_x x + k_y y)}, \quad (1)$$

where $k^2 = k_x^2 + k_y^2$. Only exponential solutions which decay as $z \rightarrow \pm \infty$ are physically allowed in regions I and V. In regions II and IV, both exponential solutions are required to satisfy boundary conditions at the surfaces of the ground plane, but the inward decaying exponentials will dominate the solution. The coefficients of these terms are uniquely determined by the corresponding Fourier components of the current distributions in the magnet planes.

Because the higher Fourier components ($n_x = 3, 5, \dots$) decay more rapidly, we shall henceforth confine our treatment to the case of a single dominant Fourier component with $k_x = 2\pi/\ell_x$ with the understanding that it may be necessary to superpose results for two or more values of k_x to get a quantitative treatment in some cases. The relative size of the first Fourier component is illustrated in Fig. 3. The field variation in the y-direction only contributes to the exponential attenuation by making $k > k_x$ without any qualitative changes in the results. We give no further attention to it here except to retain the notation k in the exponential factors as a reminder.

The field in the metal satisfies the differential

equation $\nabla^2 \vec{B} = 2i\vec{B}/\delta^2$ where δ is the familiar skin depth defined by

$$\delta^2 = \frac{c^2}{2\pi\sigma\omega} = \frac{c^2}{2\pi\sigma k_x v} . \quad (2)$$

Since the magnet period is much greater than the skin depth, allowing neglect of k^2 compared to $1/\delta^2$, the solutions are well approximated by the form

$$\vec{B} \propto e^{\pm(1+i)z/\delta} e^{ik_x x} , \quad (3)$$

in the metal.

The most general solution can be built up from these parts with coefficients selected to satisfy the boundary conditions. Although such a solution is rather straightforward, it leads to complicated results needing numerical evaluation, with consequent loss of simplicity and insight. We sidestep this frontal attack. Instead, we first deduce convenient general expressions for lift and drag forces; then we treat the two limits in which simple solutions in closed form can be obtained, namely, the limits in which the thickness d of the ground plane is either much less than or much greater than the skin depth δ . Finally, we propose a way to piece these two solutions together to yield simple expressions which should also be reasonable approximations even in the cross-over region, where $d \approx \delta$. These results are used to explore the image force suspension which has a line of train magnets over a homogeneous conducting sheet track. They are then extended to treat a

novel form of the null-flux suspension with train magnets both above and below the sheet track.

B. General Expressions for Lift and Drag Forces

The Lorentz force per unit area is $(1/c) \int (\vec{J} \times \vec{B}) dz$, where the integration is over the thickness of the metal sheet. The forces on the magnet structure are equal and opposite. Using the Maxwell equations, simple expressions for the lift and drag forces may be deduced from this formula.

The lift per unit area on the magnets is simply the difference in the magnetic pressures due to the rms fields \vec{B}_u and \vec{B}_l at the upper and lower surfaces of the sheet:

$$F_L = \frac{|\vec{B}_u|^2 - |\vec{B}_l|^2}{8\pi} . \quad (4)$$

Taking advantage of the fact that $B_y \approx 0$ in our one-dimensional geometry and that either B_z is small ($\approx k_x \delta B_x \ll B_x$) or else $B_{uz} \approx B_{lz}$, this can be written to good approximation as

$$F_L \approx \frac{|B_{ux}|^2 - |B_{lx}|^2}{8\pi} = \text{Re} \frac{\bar{B}_x^* \Delta B_x}{4\pi} , \quad (5)$$

where $\bar{B}_x \equiv (B_{ux} + B_{lx})/2$ and $\Delta B_x \equiv (B_{ux} - B_{lx}) = 4\pi \int y/c$, the net sheet current density in the metal.

The magnitude of the drag force per unit area F_D can be inferred by equating the power dissipation in the plate $\sigma^{-1} \int |\vec{J}|^2 dz$ to $F_D v$. In the strictly one-dimensional case, the current flows only along \hat{y} , and it is given by $J_y = \sigma E_y = \sigma(v/c)B_z$ by Faraday's law of induction. In that case

$$F_D = \frac{\sigma V}{c^2} \int |B_z|^2 dz. \quad (6)$$

Now, if the width w of the sheet is finite, the requirement of closed current loops ($\text{div } \vec{j} = 0$) introduces terms in J_x^2 and J_z^2 , which are, respectively, corrections of order $(l_x/2w)^2$ and the smaller of $(d/w)^2$ or $(\delta/w)^2$. The first correction is of order unity, but a similar correction appears in the lift calculation, so it essentially cancels in the lift/drag ratio. The J_z^2 correction is generally negligible. Thus, we should get quite reliable results, especially for the lift/drag ratio, by using our one-dimensional approximations (5) and (6), even in realistic finite-width geometries. Accordingly, we shall use the one-dimensional approximations in the following without further discussion.

To use these formulas, we must relate the fields in the sheet to the magnet source strengths. The fields in the space between the sheet and the magnets are of the form $B = B^+ e^{kz} + B^- e^{-kz}$. For such fields,

$$2B_x' e^{kz} = B_x \cdot \frac{1}{k} \frac{d}{dz} B_x = B_x \cdot iB_z, \quad (7)$$

since $\text{curl } \vec{B} = 0$. Thus, the boundary conditions at the upper and lower surfaces can be expressed in terms of parameters $b_{u,\ell}$ which depend on both the strengths of the fields in the planes of the magnets $B_{0u,\ell}$ and the exponential attenuation over the distances $h_{u,\ell}$ separating the magnets from the surfaces:

$$b_u = 2B_{ou}e^{-kh_u} = B_{ux} + i B_{uy} , \quad (8)$$

$$b_l = 2B_{ol}e^{-kh_l} = B_{lx} - i B_{lz} .$$

The rms source strengths B_o are given directly in terms of the Fourier coefficients of the sheet current densities as

$$B_{ou,l} = (2\sqrt{2\pi}/cl_x) \int_0^{lx} j_{yu,l}(x) e^{-ikx} dx , \quad (9)$$

where the somewhat curious coefficient takes account of the presence of both $e^{\pm ikx}$ components and of the conversion to rms values from the dc value $j_y(x)$. As an example, a close spaced series of 3 ft. square loops carrying persistent currents of 300,000 A would give $B_o \approx 6,000$ G.

C. Thick Ground Plane or High-Velocity Limit

As our first application of the general method, we treat the case of a single set of magnets moving over a ground plane at sufficient speed that δ , the skin depth, is much less than d , the thickness of the sheet. In this case, $b_l = 0$, and we work only with the boundary condition at the upper surfaces; moreover, since the field in the sheet must decay downward, it must have the form $\vec{B} = \vec{B}_u e^{(1+i)z/\delta} e^{ikx}$. Using $\text{div } \vec{B} = 0$, $iB_z = k\delta B_x / (1+i)$. Thus, the boundary condition (8) yields $B_{ux} = b_u [1+k\delta/(1+i)]^{-1}$. To find the lift force, we simply insert this in (5), with the result

$$F_L = \frac{F_\infty}{1+k\delta} , \quad (10)$$

where

$$F_{\infty} = \frac{|b_u|^2}{8\pi} = \frac{B_{ou}^2 e^{-2kh_u}}{2\pi} \quad (11)$$

is the limiting value of F_L when $v \rightarrow \infty$ so that $k\delta \rightarrow 0$. In obtaining (10), we have dropped all terms of order $(k\delta)^2$ and higher. To get a more accurate result, one must include the $|B_z|^2$ term and also note that $d^2B/dz^2 = (v^2 - \partial^2/\partial x^2)B = (2i/\delta^2 + k^2)B$, so that a k -dependence enters the skin depth. When this is carried through consistently, one finds that the first corrections to (10) are actually of order $k^3\delta^3$, so (10) should be quite a useful approximation.

Now let us calculate the drag force using Eq. (6). It is

$$F_D = (\sigma v/c^2) |B_{uz}|^2 \int_{-\infty}^0 e^{2(1+i)z/\delta} dz = (\sigma v \delta / 2c^2) |B_{uz}|^2 \quad (12)$$

Using our relation between B_{uz} and B_{ux} , this becomes

$$F_D = k\delta |B_{ux}|^2 / 8\pi = k\delta F_L = F_{\infty} \frac{k\delta}{1+k\delta} \quad (13)$$

These results for the lift and drag forces are plotted vs velocity in Fig. 4. The velocity scale is normalized in terms of a characteristic velocity $v_2 = c^2/\ell_x \sigma$ at which $k\delta = 1$. (For the typical design parameters listed in Table I, $v_2 \approx \frac{1}{2}$ km/h.) Thus, the approximation $k\delta < 1$ is only valid for $v/v_2 \geq 1$, but in practical situations this is always satisfied.

Note the particularly important result that the normalized drag (or inverse of the lift/drag ratio) has the simple form

$$F_D/F_L = k\delta = 2\pi\delta/\ell_x = c/(\ell_x\sigma v)^{1/2} \quad (14)$$

Physically speaking, this is the small ratio imposed by the skin effect on the fields B_z and B_x , with which J_y interacts to give the two forces. Using numerical values from Table I, $F_D/F_L = 1/30$ for aluminum when $v = 500$ km/h. Even with the less favorable speed of 100 km/h, $F_D/F_L = 1/13$, which is still workable and comparable to the wind resistance at higher speeds. At these two speeds, $\delta \approx 1$ cm and 2.3 cm, respectively, so that the results should apply quite well for the case of a 2.5 cm aluminum ground plane.

D. Thin Ground Plane or Low Velocity Limit

We now consider the other simple limit, $d \ll \delta$, which always holds at sufficiently low velocities, and which is quite adequate at all speeds of interest for aluminum of $d \leq 0.6$ cm. First we treat the single-sided case, in which $b_\ell = 0$. For this case, it is a satisfactory approximation to replace the actual sheet by one of infinitesimal thickness but the same finite conductance per square σd as the actual sheet. The boundary values on the upper and lower surfaces are then related by $B_{uz} = B_{\ell z} = B_z$ and

$$\frac{c}{4\pi} (B_{ux} - B_{\ell x}) = \int_y = \sigma d E_y = \sigma d (v/c) B_z \quad (15)$$

Combined with (8), this gives

$$B_{uz} = B_{lz} = -i B_{lx} = (i + 2v/v_0)^{-1} B_{ux} = b_u / 2(i + v/v_0), \quad (16)$$

where

$$v_0 = c^2 / 2\pi\sigma d = 2cR_{\square} / Z_0, \quad (17)$$

is a characteristic velocity. Here, $R_{\square} + 1/\sigma d$ is the resistance per square of the plane, while $Z_0 = 4\pi/c$ is the impedance of free space (which is 377 Ω in practical units). The physical significance of this velocity is that when $v = v_0$, the angular frequency ω of the currents induced in the sheet equals their decay rate $\tau^{-1} = R/L$. For a 2.5 cm thick aluminum sheet, $v_0 = 6.7$ km/h.

When the relations (16) are used to compute the lift and drag forces using (5) and (6), the results are

$$F_L = F_{\infty} \frac{v^2/v_0^2}{1 + v^2/v_0^2}, \quad (18)$$

and

$$F_D = F_{\infty} \frac{v/v_0}{1 + v^2/v_0^2}, \quad (19)$$

where F_{∞} is as defined in (11). Thus the lift starts quadratically in velocity, reaching half its full value at v_0 , and saturates at the same value F_{∞} as for the thick ground plane. The drag, on the other hand, initially rises linearly, reaches a maximum of $F_{\infty}/2$ at v_0 , and then falls as v_0/v . The normalized drag again has a remarkably simple form:

$$F_D/F_L = v_0/v \quad (20)$$

so that below v_0 the drag force predominates, whereas above v_0 the lift force is dominant. Note that this result (20) is independent of the Fourier component, depending only on R_{\square} of the ground plane. Thus, for velocities sufficiently low that $\delta > d$, this result should have general validity for any field configuration.

E. Intermediate Velocities: General Discussion

Having worked out the behavior to be expected in the limiting cases in which the skin depth δ is either much greater or much less than the thickness d of the ground plane, we must now put them together to get an overall view. From (2), the characteristic velocity v_1 at which $d = \delta$ is $v_1 = v_0/k_x d$.

If we compare the two approximation schemes at $v = v_1$, where they should be approximately equally good, we find exactly the same lift/drag ratio for both, namely,

$$F_L/F_D|_{v = v_1} = \frac{v_1}{v_0} = \frac{k_x d}{2\pi} \quad (21)$$

Thus, we can simply piece together the two approximations, using (14) above v_1 and (20) below v_1 , and they will join together continuously. There will be a spurious discontinuous change of slope at v_1 due to the change in approximation, but generally the result should be reasonably close to that of the exact solution.

The situation is less favorable with the lift force itself since some errors which cancel in the ratio F_L/F_D do not do so in F_L itself. At $v = v_1$, the low and high velocity results for F_L/F_∞ are $[1+(v_0/v_1)^2]^{-1}$ and $[1+(v_0/v_1)]^{-1}$, respectively. So long as $v_1/v_0 \gg 1$, these are both nearly unity so the difference is not very important. For the values in Table 1, $v_1/v_0 \approx 12$, and F_L/F_∞ from the two limiting formulas differs by 4 percent at $v = v_1$. For smaller values of v_1/v_0 , the discrepancy is more significant.

Although straightforward numerical calculations would give the exact results, eliminating any uncertainty, it is convenient for semi-quantitative work to have simple analytic expressions. We suggest the following prescription: For $v < v_1$, we use the low velocity formulas (18) and (19). For $v > v_1$, we use these same formulae with d replaced by the skin depth δ . This has the effect of making the replacement $v \rightarrow (v_1 v)^{1/2}$ in the formulas for $v > v_1$. The ratio F_L/F_D is not changed from the value (14), but the lift force (10) becomes

$$F_L/F_\infty = \frac{1}{1+v_0^2/v_1 v} \quad , \quad (v > v_1) \quad , \quad (22)$$

which has the advantage of joining continuously to (18) at v_1 , but the disadvantage of a less exact asymptotic approach to unit. The latter disadvantage is not serious for train geometries where $v_1/v_0 \gg 1$ because $F_L/F_\infty \approx 1$ in any case for $v > v_1$.

These results are presented graphically in Fig. 5 for

several values of v_1/v_0 . Since we plot against v/v_0 , the low velocity regime ($v < v_1$) is independent of magnet configuration, following the heavy curves. Above the various values of v_1 , the various light curves are appropriate. Note that they vary more slowly than the heavy curves because $(v_1 v)^{1/2}$ replaces v in the formulae.

One consequence of this is the change of shape of the F_D curve from a sharply peaked one to one which rises quickly and falls slowly. The latter behavior is characteristic of the experimental data reported by Coffey, et al.^{7,8} Their data were obtained from a test geometry with $v_1/v_0 \approx 1$, so the above prescription is not expected to apply very accurately. Rather, the thick ground plane results, shown in Fig. 4, are probably more applicable over most of their velocity range. These results give an extremely slow fall of F_D from its maximum ($\approx 0.37 F_\infty$), in reasonable agreement with the data.

F. Double Magnet of "Null-Flux" Geometry

As our final application of this method, we treat a novel configuration in which the metal sheet lies between two rows of oppositely polarized train magnet coils, so that both b_u and b_ℓ are non-zero in (8). If the sheet is thick compared to δ , its two surfaces can be treated independently. This leads immediately to the conclusion that the drag forces add, while the lift forces tend to cancel, causing a degraded performance of little interest.

The interesting case is when $d \leq \delta$, so that the con-

tributions of the top and bottom magnets to B_z [which gives the drag, according to (6)] tend to cancel, as does ΔB_x [which gives the lift, according to (5)]. Since the drag vanishes as B_z^2 whereas the lift vanishes only linearly with ΔB_x , it is qualitatively obvious that the drag/lift can be made arbitrarily small by making the cancellation arbitrarily complete. This is essentially the same physical idea that underlies the null-flux scheme of Powell and Danby^{12,13}. We now develop this model in quantitative detail to show that it is feasible to obtain the low drag of the null-flux scheme without the need for the structured track of coils characteristic of the Powell-Danby configuration.

If we retain the approximation of replacing the sheet by one of infinitesimal thickness but the same R_{\square} , we have the boundary condition that B_z is continuous, so $B_{uz} = B_{lz}$, and B_x suffers a discontinuity $4\pi q_y/c = 2vB/v_0$ as above. Combined with the conditions in (8), imposed by the sources, this leads to the results

$$\bar{B}_x \equiv (B_{ux} + B_{lx})/2 = (b_u + b_l)/2 \quad ,$$

$$\Delta B_x \equiv B_{ux} - B_{lx} = 2vB_z/v_0 = (b_u - b_l)/(1 + iv_0/v). \quad (23)$$

Inserting these in our force equations (5) and (6), and taking b_u and b_l real, yields

$$F_L = \frac{b_u^2 - b_l^2}{8\pi} \frac{v^2/v_0^2}{1 + v^2/v_0^2}$$

and

$$F_D = \frac{(b_u - b_\ell)^2}{8\pi} \frac{v/v_0}{1+v^2/v_0^2} \quad (25)$$

Since $B_\ell > b_u$ as the sheet approaches a symmetric position, these results indicate that F_D/F_L can be made arbitrarily small. In fact the finite thickness of the sheet always causes a finite drag, and we must evaluate this to get a realistic result. This drag arises because $dB_z/dz \neq 0$ even if $B_z = 0$ in the symmetric position. Using $\text{div } \vec{B} = 0$, $B'_z(0) = dB_z/dz|_0 = -ikB_x|_0 = -ik\bar{B}_x = ik(b_u + b_\ell)/2$. If we approximate $B_z(z)$ in the sheet by $B_z(0) + B'_z(0)z$ when evaluating (6), the term in B'_z gives an additional contribution to the drag, proportional to $k^2 d^2$. Thus (25) is replaced by

$$F_D = \frac{v}{8\pi v_0} \left\{ \frac{(b_u - b_\ell)^2}{1+v^2/v_0^2} + \frac{k^2 d^2}{12} (b_u + b_\ell)^2 \right\} \quad (26)$$

All further correction terms are down by factors of order $d^2/10\delta^2$, and will be ignored hereafter.

It is now useful to re-express the parameters b_u and b_ℓ in terms of the magnet strengths and distances as given in (8). For simplicity consider a symmetric magnet structure, so that $B_{ou} = B_{ol} = B_0$. Further, define the "sag" of the train below the symmetry plane by $\Delta h = h_o - h_u = h_\ell - h_o$, where $2h_o$ is the separation of the upper and lower magnet planes. Then evaluating (24) and (26) carrying only leading terms in Δh , we obtain

$$F_L = \frac{2}{\pi} B_0^2 e^{-2kh_0} \frac{v^2/v_0^2}{1+v^2/v_0^2} k\Delta h \quad , \quad (27)$$

and

$$F_D = \frac{2}{\pi} B_0^2 e^{-2kh_0} \frac{v}{v_0} \left(\frac{k^2 (\Delta h)^2}{1+v^2/v_0^2} + \frac{k^2 d^2}{12} \right)$$

As expected qualitatively, $F_L \propto \Delta h$, whereas $F_D \propto (\Delta h)^2$ for $d \neq 0$. Also note that for fixed magnet strength B_0 , F_L is maximum for $k_x = 4\pi h_0$, which suggests that this relation be used as a design guide. It is consistent with the numerical values chosen in Table I.

To carry the analysis further it is convenient to consider the normalized drag F_D/F_L . After dropping a term $d^2/12 (\Delta h)^2$ compared to unity, this has the simple form

$$F_D/F_L = k\Delta h \frac{v_0}{v} + \frac{kd^2}{12\Delta h} \frac{v}{v_0} \quad (29)$$

This normalized drag has a minimum value

$$(F_D/F_L)_{\min} = \frac{kd}{\sqrt{3}} = \frac{2\pi}{\sqrt{3}} \frac{d}{k_x} = \frac{1}{80} \quad , \quad (30)$$

at a characteristic velocity

$$v_3 = 2\sqrt{3} v_0 \Delta h/d = \sqrt{3} c^2 \Delta h/\pi\sigma d^2 \approx 80 \text{ km/h} \quad . \quad (31)$$

Thus (29) has the normalized form

$$\frac{F_D/F_L}{(F_D/F_L)_{\min}} = \frac{1}{2} \left(\frac{v}{v_3} + \frac{v_3}{v} \right) \quad . \quad (32)$$

So long as F_L equals the weight of the train, it cancels out, and the ratio $F_D/(F_D)_{\min}$ will also be given by (32).

Since the drag force has a broad minimum near $v = v_1$, and since $(F_D)_{\min}$ is so small compared to the high speed air drag ($\approx 0.1 F_L$ at 500 km/h), it is not critical to minimize the magnetic drag at cruising speed. Rather, the parameters should be chosen to set v_1 well below the maximum speed of the train, so as to reduce the dominant magnetic drag at lower speeds. Fig. 6 illustrates this point with the parameter values in Table I.

Evidently, the potential performance of this system is very attractive compared to the single-sided, image-force systems, at least from the standpoint of low magnetic drag. Its primary disadvantage is the requirement of a double set of magnets, with the related structural complications. Also, as in the null flux system, the magnets must be considerably stronger than required in the single-sided system, because of the substantial cancellation of lift forces described above. This poses no serious difficulty; 300,000 A loops are adequate for typical train weights, and these are well within present capabilities. In fact, as Powell and Danby¹⁰ have pointed out, the stronger magnets offer the advantage of making possible an efficient linear synchronous propulsion system. A schematic diagram of such a combined levitation-propulsion system is shown in Fig. 7.

III. LUMPED PARAMETER ANALYSIS OF LEVITATION EFFICIENCY

A. General Formulation

We can gain very general insight into the comparative properties of various systems, such as the Powell-Danby "null-flux" suspension (which uses coils) and the simpler ground plane "image force", by a lumped parameter equivalent circuit analysis. Imagine that we are levitating one coil above another by passing alternating current I_1 through one coil which sets up a magnetic field which repels the current I_2 , it induces in the other coil. The force can be expressed in terms of the mutual inductance as $F = -I_1 I_2 \frac{dM}{dh}$. Since I_2 is induced by the coupling to I_1 , it is obtained by solving the differential equation

$$L_2 \frac{dI_2}{dt} + I_2 R_2 = M \frac{dI_1}{dt} \quad (33)$$

Taking an $e^{i\omega t}$ time dependence, and solving, we find

$$I_2 = \frac{(M/L_2) I_1}{1 + i/\omega\tau} \quad (34)$$

where $\tau = L_2/R_2$ is the ordinary time constant of the secondary circuit. Thus, the resulting lift force is

$$F_L = \frac{-|I_1|^2 M(dM/dh)}{L_2(1 + 1/\omega^2\tau^2)} \quad (35)$$

Here $|I_1|^2$ has the significance of the rms value, and h measures the vertical separation of the coils. Drive forces could be

obtained from the horizontal components of $\dot{V}M$ in a more complete model.

As in Section 2 we compute the drag force F_D by equating $F_D v$ to the dissipated power

$$P = I_2^2 R_2 = \frac{(M/L_2)^2 |I_1|^2 R_2}{1 + 1/\omega^2 \tau^2} ,$$

so that

$$\frac{F_L}{F_D} = v \tau \frac{1}{M} \left| \frac{dM}{dh} \right| . \quad (37)$$

Thus, F_L/F_D is proportional to velocity, insofar as M and τ are independent of frequency, that is, ignoring skin depth effects. Equation (37) is our fundamental result, which can be applied to any geometry of suspension with suitable evaluation of the parameters.

B. Suspension Over Ground Plane

In the levitation of a coil a distance h above a perfectly conducting ground plane, the field configuration and energies can be found by assuming an equal and opposite current in an image coil an equal distance below the surface. The mutual inductance M for two parallel conductors of length ℓ and separated by $2h \ll \ell$ varies as $\ell \ln(\ell/h)$ so that $(1/M) |dM/dh| \approx 1/h$. Generally, we expect the numerical coefficient to be of order unity. For a pair of square loops 90 cm on a side separated by 30 cm it is 1.1/h. Thus, the lift/drag ratio should be approximately

$$\frac{F_L}{F_D} \Big|_{\text{Image}} \approx \frac{v\tau}{h} . \quad (38)$$

We estimate τ by taking the inductance of a square loop of edge ℓ and diameter $\ll \ell$

$$L = \frac{8\ell}{c^2} (\ln \frac{\ell}{d} + .17) \approx \frac{20\ell}{c^2} \quad (39)$$

A rough estimate of the resistance of the current-carrying material is that of a loop of perimeter 4ℓ , width $2h$, and thickness d , namely $R \approx 2\ell/\sigma hd$, so that

$$\tau = \frac{L}{R} = \frac{10\sigma hd}{c^2} . \quad (40)$$

Thus, from (38)

$$\frac{F_L}{F_D} \approx \frac{10\sigma vd}{c^2} = 1.6 \frac{v}{v_0}$$

The good agreement of this result with (20) gives us some confidence in the present approach although it is not well suited for treating the case of a conducting ground plane, since it is far from a lumped constant situation.

Even further from the lumped parameter situation is the case of a periodic magnetic field over a homogeneous conducting sheet treated in Section II. In this case, the exponential variation of the field leads to $(1/M) |dM/dh|$

$= k = 2\pi/\lambda_x$, so the lift/drag ratio (40) becomes

$$F_L/F_D = 2\pi v \tau / \lambda_x = \omega \tau$$

where ω is the angular frequency of the ac field seen by the track. This agrees with our result $F_L/F_D = v/v_0$ (20) if $\tau = \sigma d \lambda_x / c^2$, which is reasonable considering that the sinusoidal currents flow in strips whose width is $\approx \lambda_x / 2\pi$. We may therefore consider the condition $v = v_0$ as equivalent to $\omega \tau = 1$. For lower velocities the track currents decay before the train moves on. For higher velocities current is left behind in the track.

C. Null-Flux Suspension

The real utility of the method comes in dealing with a system using coils such as the Powell-Danby null-flux suspension.^{12, 13} We represent this by three parallel square coils of edge l , sketched in Fig. 2(b) and (c). The outer two coils are connected in series opposition to form the secondary. Then, if the primary is centered, the mutual inductance to the two halves of the primary cancel. If it is displaced by a distance Δh , the net mutual inductance is given by

$$M = \frac{8\pi}{c^2} \ln \frac{h_0 + \Delta h}{h_0 - \Delta h} \approx \frac{16l\Delta h}{c^2 h_0}, \quad (\Delta h \ll h_0 \ll l).$$

Thus,

$$1/M |dM/dh| \approx 1/\Delta h \tag{43}$$

and

$$F_L/F_D \approx v \tau / \Delta h.$$

We can compare this result with the formally similar

(38) for the image force problem. Although differences do arise from the various values of τ in the different suspensions, the large benefits of the null-flux scheme arise from the fact that by increasing the train magnet current, Δh can be made arbitrarily small. There is no mechanical obstruction since Δh is measured from a reference plane in space. In the null-flux system, therefore, the suspension stiffness can be adjusted almost independently of the mechanical clearances and the stiffer suspensions have larger values of F_L/F_D . As is the case in the hybrid system discussed in Section II.F, this high efficiency arises from the fact that F_L is proportional to the product of train and track currents divided by their separation. The drag F_D , however, arises from dissipation in the track only. It decreases as the inverse square of the train magnet current when F_L is kept constant.

Because of the large critical currents available from modern superconductors it appears possible to achieve the increased efficiency of the null flux suspension in practical systems. The stiffer suspension would require finer track tolerance or a secondary suspension to protect passengers from unacceptable transverse accelerations. The suspension would also have to smooth out any vibration due to the loop structure in the track. (This problem would be eliminated in the configuration analyzed in Section II.F.) In order to maintain τ as large as for the homogeneous ground plane at low velocities, it would be necessary to use a comparable amount of metal

are much too large to permit the required $\approx 10 - 100$ Hz excitation of the train magnets. Therefore the linear induction drive motor must have normal state (copper) stator windings and be separated from the levitation system. The train then contains two linear induction motors pulling against each other. The design freedom obtained by separating the motors allows some improvements such as increasing the resistance of the reaction rail near stations to provide large starting and braking forces at large slip Δv . In the language of Section II, we increase v_0 to increase F_D at large v . This reduces the need for widely adjustable drive frequencies. For optimum drive force the motor should be operated at a slip velocity $\Delta v \approx v_0$ by varying the excitation frequency.

The drive motor must be strong enough to accelerate the train past the drag peak at $v = v_0$. It also requires a similar force to accelerate the train with $0.2g$ in the intermediate speed range. The force required to overcome air resistance at top speed will also be comparable. The drive motor, therefore, must be at least as strong as the levitation "motor" without its advantage of superconducting magnets. The use of the more efficient double-sided geometry²² with iron flux returns would partially compensate for the lower stator current densities. The magnetic reluctance of a large clearance gap, however, would still produce a poor power factor. Even with variable frequency drive it will be difficult for an induction motor to provide adequate drive force with the large train-

track clearances which are a desirable characteristic of superconducting magnetic suspensions. A more complicated system involving a separately suspended induction motor operating very close to a reaction rail appears to be the only solution. Even if the double suspension can be made to operate satisfactorily, onboard induction drive requires large amounts of power to be transmitted to the rapidly moving train. This appears to be a difficult technical problem.

B. Active Track

The obvious way to escape the troubles of the linear induction motor drive is to pull the train along by a traveling flux wave from an active track. Since this removes the heavy induction motor from the train, the train weight is substantially reduced and no drive power needs to be transferred to it. One could conceive of retaining the induction principle, but for any reasonable clearance, the efficiency of the system would be very low considering the need to activate miles of track. Thus we are led to give favorable consideration to a linear synchronous drive. Such a drive takes advantage of the existence of strong superconducting permanent magnets in the train to obtain adequate propulsive force with much lower track currents than would otherwise be possible. The principal disadvantage of the synchronous drive is the requirement of wide-range variable frequency power to give adequate propulsion at all speeds.

In principle the linear synchronous motor can be used with the homogeneous ground plane levitation system of Section II

for the track. At high velocity, on the other hand, an increase in r or a reduction in the amount of metal used (compared with the homogeneous ground plane) can be obtained if the coils are wound from many turns of wire to avoid skin effect losses.

IV. MAGNETIC PROPULSION

A. Passive Track

Any high speed vehicle needs substantial amounts of drive power. Linear electric motors are a favorable drive mechanism because they produce less environmental pollution than alternatives such as jet engines. Induction motors with the electro-magnets (stator) on the train and a conducting reaction rail (rotor) have been suggested for driving conventional wheeled trains as well as air-cushion and magnetically levitated vehicles. The theory developed in Section II for the drag force F_D of a passive magnetic suspension pulled over a plane homogeneous track is directly applicable to the design of a linear induction motor. The suspension can be considered as an induction motor with zero synchronous velocity. The curve of F_D versus velocity v in Fig. 5 can thus be interpreted as the drive force (torque) versus slip rate.

The simplest levitation and drive system is one in which the train magnets of Section II (Fig. 2(a)) are excited with suitably phased ac, so that a periodic magnetic field moves back along the train with a phase velocity v_ϕ . We then replace v by $\Delta v = (v - v_\phi)$ in our formulas and obtain a negative

F_D which is now the force driving the train. There would be adequate slip Δv to provide levitation during acceleration, but in the absence of any external drag force such as air friction the train would approach synchronous speed $v = v_\phi$ and drop onto the track. Even allowing for air drag, care would be required to design a system which would be stable at all drive speeds because the drive decreases with increasing slip in the range where levitation is strong. It appears that the stability problem could be solved by introducing standing wave as well as traveling wave excitation of the train magnets. Such a system could hover at rest, accelerate, and brake.

In the low velocity range $v < v_1$, the power required to levitate a train is $P_D = F_D v = F_L v_0 = Mgv_0$ from (20). It is independent of velocity because the same amount of track current is required to lift the train at any speed. The dissipated power is proportional to track resistance which increases only at high velocities where the skin depth limits the effective track thickness and (20) is no longer valid. We may compare the power required to overcome drag with the power required to accelerate the train at $0.1g$. This reversible power $P_R = 0.1 Mgv$ is greater than P_D for $v > 10 v_0$. Using the numerical values from Table I, $10 v_0 \approx 65$ km/h, so most of the power handling capacity of the motor would be required for acceleration, not levitation.

The idealized system described above is unfortunately not practical at present because ac losses in superconductors

if portions of the train loops are exposed so that the track drive loops can couple to them without excessive shielding from the ground plane. In practice, however, stronger train magnets are required for efficient drive than for image force levitation. Such strong train magnets are a feature of the various null-flux suspensions described above. Powell and Danby¹⁰ have studied specific examples of these systems. Here we re-examine some of the design considerations for linear synchronous propulsion, including questions of stability, from a more general point of view.

For simplicity, we restrict our attention to a single-phase drive, with a track thrust winding of the form shown in Fig. 8(a). The phase velocity of the wave is $v_\phi = v l_x$, where v is the frequency of the rms current I , and l_x is the spatial period. Such a configuration actually produces a standing wave which could drive a train in either direction. Additional windings, or more likely a full three-phase system, would be used to define a sense to v_ϕ . For a given direction of motion, however, the effects of the several windings are simply additive, and our simple analysis is readily adapted to give the appropriate numerical results.

For an initial orientation, it is useful to compare typical track impedance values with the load imposed by the train. For definiteness, consider a configuration with $l_x = 2w$, so that the loops are square. Then the total round-trip conductor length is 4 times the length of the track. If the

conductor were 0000 gauge aluminum (1.2 cm diameter), the resistance per km would be about 1.1 Ω ; 2.5 \times 2.5 cm square aluminum bus would give \approx 0.15 Ω /km. In either case, the inductance per km would be about 3.6 mH, corresponding to an inductive reactance of \approx 1.5 Ω /km at 60 Hz.

For comparison, we must know the load imposed by the train. The instantaneous back emf follows the variation of \bar{B}_z shown in Fig. 3, but only the fundamental Fourier component gives any time average effect if the track current is purely sinusoidal. This is nearly the case because of the highly inductive nature of the track impedance. The back emf induced by the train will be

$$e = (4wL/l_x)bv/c, \quad (44)$$

where L is the total length of the magnet structure. By conservation of energy, $F_p v = cI_1 \cos\phi$, where F_p is the propulsive force, I_1 is the track current, and ϕ is the phase angle between I_1 and e . Evidently F_p has its maximum value $F_{p,max}$ when $\cos\phi = 1$, in which case the impedance is purely resistive, and its value is

$$R = e/I_1 = c^2/F_{p,max} v = 16 B^2 w^2 L^2 / F_{p,max} c^2 l_x^2. \quad (45)$$

If less propulsive force is required, a phase shift $\phi \neq 0$ develops, and the train impedance acquires a reactive component.

The significance of (45) is that the impedance varies

as B^2 in comparing systems of the same physical size. Thus, the null-flux systems which typically have at least three times stronger magnets than image force suspensions will have impedances an order of magnitude larger. To relate the actual values of resistance to system parameters, it is convenient to assume that the clearance between magnet and drive winding is the same as that between the magnet and the track. One can then eliminate B^2 in terms of the weight, and deduce the result

$$\frac{R}{Z_0} = \frac{8 w L g v}{f l_x^2 g' c} \quad , \quad (46)$$

where $F_{p,max} = Mg'$, $Z_0 = 4\pi/c = 377 \Omega$ is the impedance of free space, and $f = 1$ for image force systems and $f = 2\Delta h/h_0 = 0.1$ for null-flux systems. Thus, at $v = 400$ km/h, we have $R \approx 0.1 \Omega$ for image-force and $R \approx 1 \Omega$ for null-flux suspensions. Since the drive power is ≈ 5 MW in our example, it is of considerable importance to maintain high efficiency. This requires that the track impedance of $0.1 - 1 \Omega/\text{km}$ be less than R . Thus, whatever system of suspension is used, a linear synchronous drive is made most practical by using the strongest possible train magnets. The strong magnets characteristic of null-flux systems give high enough impedance to allow reasonable efficiency with track segments of $1 - 2$ km on either side of a feed point to be activated at once. A track current $I_t \approx 2,000$ A would be required for a typical null-flux magnet strength of $I_m = 300,000$ A. The weaker magnets characteristic of image force systems would require ≈ 3

times the track current and thus ≈ 10 times the number of transformers and switches to have the same efficiency. Image force systems can of course be driven from an active track if extra train magnet strength is added which is not used for levitation.

An alternative system which is closely related to the linear synchronous motor is the linear dc motor. If current reversing switches controlled by the position of the train are placed on each drive loop of the track, a direct current can be used to propel the train. When compared with the above discussion of the linear synchronous motor, the importance of track inductance is decreased, but strong train magnets are still desirable for high efficiency. The practicality of the linear dc motor depends on the cost of the very large number of switches required.

C. Longitudinal Stability

It is well-known that there is a potential instability associated with a drag force which decreases with increasing velocity. In general, we have

$$M \frac{dv}{dt} = F_p - F_D(v) \quad , \quad (47)$$

where F_p and F_D are the propulsion and drag forces, respectively. If one attempts to maintain a constant velocity v_c by setting $F_p = F_D(v_c)$, we have, for small fluctuations in velocity δv , and assuming F_p remains constant,

$$M \frac{d(\delta v)}{dt} = -F_D'(v_c) \delta v \quad (48)$$

so that

$$\delta v = e^{t/\tau} \quad , \quad (49)$$

where the time constant τ is given by

$$\tau = -M/F'_D(v_C) \quad . \quad (50)$$

Thus fluctuations in velocity will grow exponentially if $\tau > 0$, that is, if $F'_D(v_C) < 0$, as is always the case for the magnetic drag in conventional image-force or null-flux schemes. In the hybrid scheme, however, the magnetic drag increases for $v > v_s$. When the air drag is added in, all these systems become stable at sufficiently high velocities. In fact this instability is not very serious because $\tau = 200$ sec is long enough to permit stabilization by a simple feedback system.

More fundamentally, the above analysis is incomplete because it is based on the assumption that the propulsive force F_p remains constant, independent of v . This might be a fair approximation for propulsion by a jet engine, but it certainly is inappropriate for either of the two magnetic drive systems which are under consideration.²⁶

In the induction drive system, the propulsive force has the form of F_D in (19) and Fig. 5:

$$F_p = 2F_I \frac{\Delta v/v_0}{1+(\Delta v/v_0)^2} \quad , \quad (51)$$

where F_I is the maximum propulsive force of the induction drive, which occurs when the slip velocity $\Delta v = v_\phi - v$ relative to the traveling magnetic field has the value $v_0 = c^2/2\pi\sigma d$ ($= 24$ km/h for 0.64 cm aluminum). So long as $\Delta v < v_0$, (51) gives a driving

force tending to stabilize the velocity. Instability could result only from fluctuations large enough to push the operating point over the peak in the drive curve. By operating at a slip velocity well below v_ϕ , one could certainly produce a stable propulsion system using a linear induction motor.

The stability analysis of the linear synchronous drive is more complicated since the propulsive force F_p depends on the instantaneous displacement between the driving field and the train magnets, rather than on their relative velocities. Thus, our equation of motion becomes

$$M \frac{dv}{dt} = F_0 \sin k(v_\phi t - x) - F_D(v) \quad (52)$$

where $F_0 = F_{p,max}$. In synchronous motion, $v = dx/dt = v_\phi$, and $x = -x_0 + v_\phi t$. Thus the equilibrium lag x_0 is determined by $F_0 \sin kx_0 = F_D(v_\phi)$. Expanding about this operating point,

$$M \frac{dv}{dt} = -[F_0^2 - F_D^2(v_\phi)]^{1/2} k \Delta x - F'_D(v_\phi) \Delta v \quad (53)$$

If we now assume that Δx and Δv vary as $e^{i\omega t}$ and define a natural frequency of longitudinal oscillation $\omega_0^2 = F_0 k/M$ we obtain

$$\omega = i \omega_0 [1 - (F_D/F_0^2)^{1/2}]^{1/2} - \frac{1}{2\tau} \quad (54)$$

since $1/\tau = F'_D(v_\phi)/M \ll \omega_0$. Note that if $F_0 = F_D$, so that

the propulsive force is constant at its maximum value, the exact form reduces to the exponential solution (49), which describes an instability if $F'_D(v_c) < 0$. In normal operation, however, $F_D = F_0/2$, and ω becomes predominantly real. Thus, if $F'_D(v_\phi) < 0$ so that $\tau > 0$, there will still be an instability but in the form of an exponentially growing oscillatory motion. For a maximum acceleration capability $F_0/M = 0.2g$, we find $\omega_0/2\pi = 0.4$ Hz.

The amplitude would increase a factor of e in ≈ 7 minutes, during which some 1000 oscillations about the equilibrium displacement x_0 would occur. Again, dynamic feedback could be used to damp out such a weak instability. However, there exists an automatic stabilizing effect from the induction force which arises if v departs from its synchronous value v_ϕ . (It is the analog of the use of "damper" windings in a rotary synchronous motor.) For small departures from v_ϕ , (51) becomes

$$F_p = -2F_I(v - v_\phi)/v_0, \quad (55)$$

where F_I is now the maximum induction drive force from the eddy currents induced by the track coils in normal metal surrounding the superconducting train magnets. When (55) is added to the right member of (52), it replaces $-F'_D(v_\phi)$ by $[-F'_D(v_\phi) - 2F_I/v_0]$ in the equation of motion about the operating point. Thus, the exponential will become a decaying rather than growing one provided that $F_I > |F'_D(v_\phi)|v_0/2$. If one estimates $F_I \approx F_0(I_{\text{track}}/I_{\text{train}}) \approx F_0/100$, and $-F'_D(v_\phi)$

$\approx F_D/2v \approx F_0/4v$, this leads to the requirement that $v_0 < v/12$. Since $v_0 = 6.5$ km/h for a 2.5 cm aluminum plate, we see that this inductive stabilization effect could be quite substantial. The only troublesome regime will be the low velocity one, in which some sort of wheeled support would be used in any case, so that an analysis of specific systems would be required.

In summary, after taking account of air drag and the induction effect, the linear synchronous drive will be stable at most velocities, particularly with the hybrid levitation system, and any instability would be so weak that it could easily be cured by a simple feedback system.

D. Acceleration to Speed

In addition to the question of stability about a desired velocity, just discussed, there is the question of intentional acceleration and deceleration to change speed. This is relatively simple in the case of the induction motor, since there is substantial propulsion over a wide range of slip velocities. Thus, amplitude control could supplement frequency control to a considerable degree in controlling acceleration.

The problem is more severe with the linear synchronous motor, a reflection of the notoriously low starting torques of ordinary synchronous motors. This means that a wide range variable frequency source will be essential. So long as exact synchronism is maintained, the full drive force is available. The velocity range over which the drive can pull the train into

synchronism can be estimated by using the equation of motion (52). (The inductive term (55) is, unfortunately, negligible in this context.) If $v \neq v_\phi$, the synchronous drive term averages to zero, except for the change in v induced by the force during a single cycle of slip. This change from the average v can be estimated by integrating (52) over the time interval $\Delta t = \ell_x/4(v_\phi - v)$. Equating the result to $(v_\phi - v)$, we get the following condition for synchronization:

$$(v_\phi - v)^2 \leq \frac{F_0}{M} \frac{\ell_x}{2\pi} \left(1 - \frac{\pi}{2} \frac{F_D}{F_0}\right) \quad (56)$$

For typical parameters we obtain the condition $(v - v_\phi) \leq 3$ km/h. It will clearly be necessary to maintain very fine feedback control of the drive frequency to avoid going out of synchronism.

VI. CONCLUSIONS

We can draw several conclusions about magnetic trains from the model calculations presented here. Of these, the first is that if a magnetic drive system (linear motor) is used for propulsion, the suspension and drive systems should be considered together in seeking an overall optimum system.

The image force suspension system appears to have many satisfactory features. The cost of both train and track is relatively low. Although it is less efficient than other systems, the magnetic drag can probably be made adequately low from the

viewpoint of power cost. This system is relatively difficult to drive, however, since losses in available superconducting materials prevent ac excitation of the train magnets. A conventional linear induction motor is required, which operates with close tolerance on a reaction rail. Such a motor would require an independent tight suspension system and also the transmittal of large amounts of power to the rapidly moving train.

Because they have stronger train magnets, it is practical to drive null-flux suspensions from an active track, eliminating the need to transmit power to the train. Image force suspension systems can also be driven from an active track if strong train magnets are used which are not efficiently coupled to the part of the track which causes levitation. Possible systems include synchronous drive from track loops excited with variable frequency alternating current (linear synchronous motor) or a dc system in which the direction of the current in the track loops is switched on command from the train (linear dc motor).

The Powell-Danby null-flux scheme uses a looped track to obtain cancellation of drag forces. The expense and possible vibration from a looped track can be avoided (while the low drag is retained) by a hybrid system which uses train magnets on either side of a thin homogeneous track. If the train magnet design is efficient, this hybrid null-flux system may prove to be an optimum one for many applications, especially those with long runs of track and relatively few trains.

The longitudinal stability of a train in the speed range in which total drag is decreasing with velocity depends on the drive system used. First-order analysis shows that the instabilities encountered in practical systems are weak and can be eliminated fairly easily with feedback control.

VII. ACKNOWLEDGEMENTS

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

REFERENCES

1. E. Bachelet, *The Engineer*, October 1912, p. 420.
2. G. R. Polgren, "New Applications of Modern Magnets" (McDonald, London, 1966).
3. A train with a servo-controlled attractive ferromagnetic suspension has been constructed by Messerschmitt-Bölkow-Blohm, GMBH, West Germany.
4. R. H. Borcherts, *J. Appl. Phys.* 42, 1528 (1971).
5. J. R. Powell and G. T. Danby, "High Speed Transport by Magnetically Suspended Trains," ASME Railroad Division Annual Meeting, New York, 1966, 66-WA/RR-5.
6. J. R. Powell and G. T. Danby, *Mech. Eng.* 89, 30 (1967).
7. R. T. Barbee, G. N. Bycroft, E. G. Chilton, F. M. Chilton, and H. T. Coffey, "The Hypervelocity Rocket Sled - A Design Analysis", Stanford Research Institute Project PMU-7014 (Sandia - AEC Contract AT(04-3)-115).
8. H. T. Coffey, F. M. Chilton, and T. W. Barbee, Jr., *Proceedings of the Royal Society Conference on Low Temperatures and Electric Power*, London, 1969.
9. C. A. Guderjahn, S. L. Wipf, H. J. Fink, R. W. Boom, K. E. MacKenzie, D. Williams, and T. Downey, *J. Appl. Phys.* 40, 2133 (1969).
10. H. T. Coffey, F. M. Chilton, and T. W. Barbee, Jr., *J. Appl. Phys.* 40, 2161 (1969).
11. C. A. Guderjahn and S. L. Wipf, in Advances in Cryogenic Engineering, K. D. Timmerhaus, ed., (Plenum, New York, 1970) v. 15, p. 117.
12. J. R. Powell and G. T. Danby, "Magnetically Suspended Trains for Very High Speed Transport", *Transactions of the 4th Annual Intersociety Energy Conversion Engineering Conference*, Washington, D.C. (1969).
13. J. R. Powell and G. T. Danby, *Cryogenics and Industrial Gases* 4, 19 (1969).
14. J. R. Reitz, *J. Appl. Phys.* 41, 2067 (1970).
15. J. R. Powell, in Recent Advances in Engineering Science, A. C. Eringen, ed. (Gordon and Breach, New York, 1970), v. 5, p. 141.

16. C. Guderjahn and S. L. Wipf, *Cryogenics* 11, 171 (1971).
17. J. R. Powell and G. T. Danby, *Cryogenics* 11, 192 (1971).
18. J. R. Powell and G. T. Danby, "The Linear Synchronous Motor and High Speed Ground Transport," (unpublished). Other recent unpublished reports by J. R. Powell and G. T. Danby include "Magnetic Suspension for Levitated Tracked Vehicles", and "Lift/Drag Ratios for Magnetic Suspensions".
19. J. R. Reitz, R. H. Borcherts, and L. C. Davis "Force on a Rectangular Coil Moving Above a Conducting Slab" Proceedings of the 1971 Conference on Magnetism and Magnetic Materials (to be published). Recent unpublished reports from the Ford group include R. H. Borcherts and J. R. Reitz "High Speed Transportation Via Magnetically-Supported Vehicles. A Study of the Magnetic Forces", L. C. Davis and J. R. Reitz, "Eddy Currents in Finite Conducting Sheets", and L. C. Davis and J. R. Reitz "Solution to Potential Problems in the Neighborhood of a Conducting Semi-Infinite Sheet or a Conducting Disk."
20. H. J. Fink and C. E. Hobrecht, *J. Appl. Phys.* (1971) discuss the instability of a single loop. Their rather alarming conclusions arise from the assumptions of constant force and freedom of the loop to rotate about a transverse horizontal axis. This freedom is severely restricted in most train designs.
21. P. L. Richards and M. Tinkham, "Magnetic Suspensions for High Speed Transportation", 1970 ARPA Materials Research Council Report, University of Michigan Contract DAHC-15-67-C-0062.
22. E. R. Laithwaite, Induction Motors for Special Purposes (Newnes, London, 1966).

TABLE I. Values Chosen for Numerical Examples

<u>Train Parameters</u>	<u>Track Parameters</u>
$x = 1.8 \text{ m (6 ft.)}$	$\rho = 2.8 \times 10^{-6} \text{ } \Omega \text{ cm (aluminum)}$
$w = 0.9 \text{ m (3 ft.)}$	$d = 2.5 \text{ cm (1 in.)}$
$L = 30 \text{ m (100 ft.)}$	$v_0 = 6.5 \text{ km/h (4 mph)}$
$M = 45,000 \text{ kg (100,000 lbs.)}$	$v_1 = 73 \text{ km/h (45 mph)}$
$g' = 0.1 - 0.2 \text{ g}$	$v_2 = 0.54 \text{ km/h (1/3 mph)}$
$v_{\text{max}} = 500 \text{ km/h (= 300 mph)}$	$d = 0.64 \text{ cm (1/4 in.)}$
$I_0 = \begin{cases} 100,000 \text{ A image force systems} \\ 300,000 \text{ A null-flux systems} \end{cases}$	$v_0 = 24 \text{ km/h (15 mph)}$
	$v_3 = 80 \text{ km/h (50 mph)}$
	$h = 15 \text{ cm (6 in.)}$
	$\Delta h = 0.64 \text{ cm (1/4 in.)}$

image force systems

null-flux systems

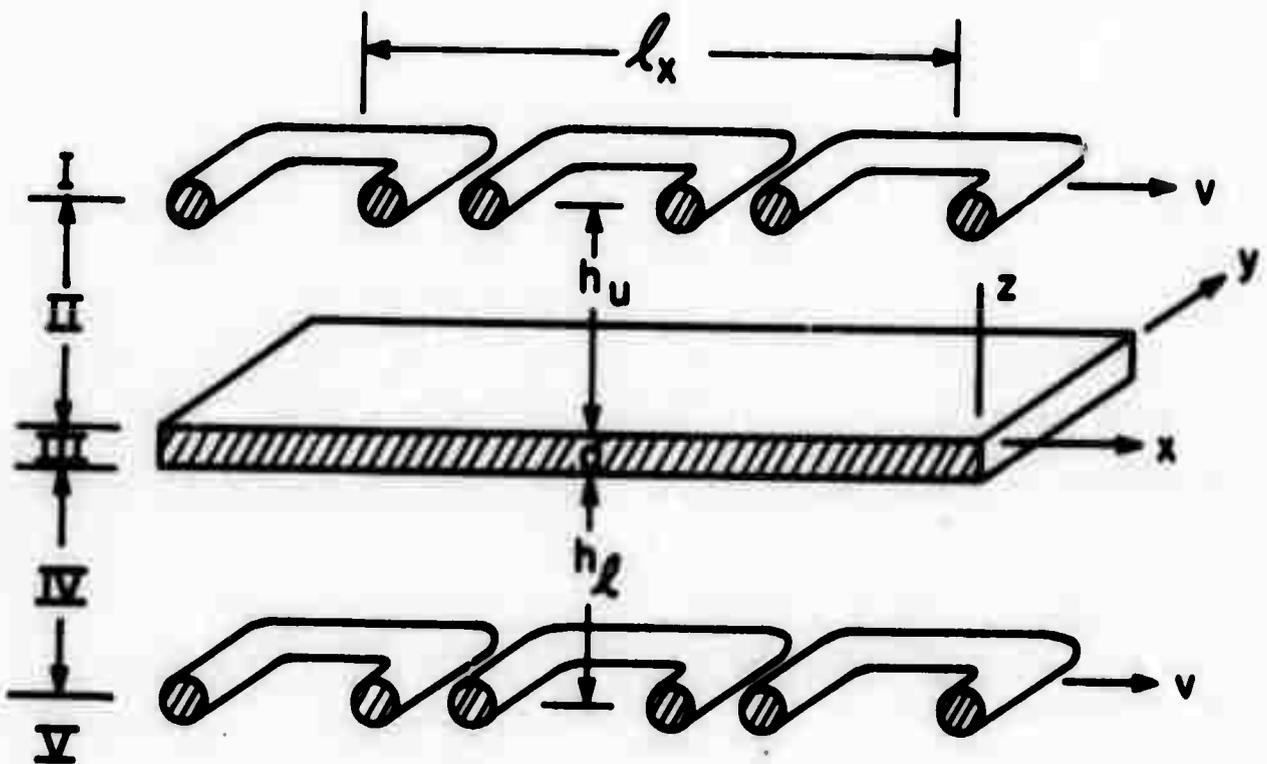


Figure 1. Schematic diagram showing periodic magnet currents at heights h_u and h_l above and below a conducting ground plane of thickness d . Magnet array is fixed in the train, and moving with velocity v along the x -direction.

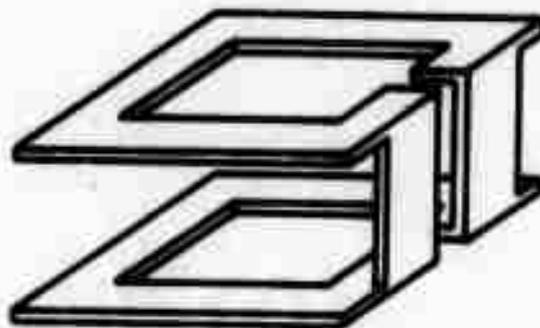
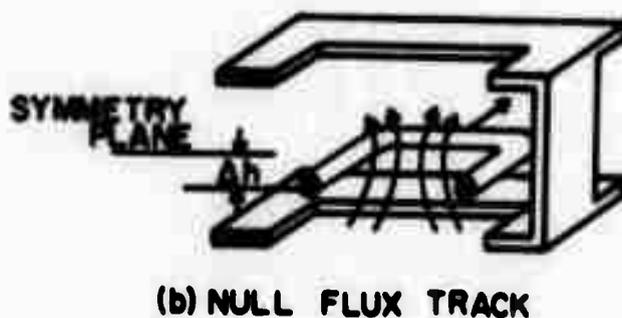
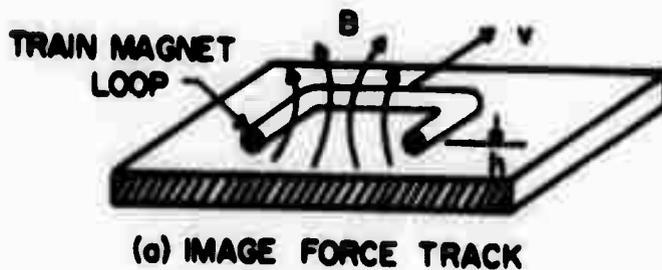


Figure 2. Schematic diagram contrasting the significance of the characteristic heights in the image force suspension (a) and the null-flux suspension (b). If the train magnet strength is increased, h increases in (a) raising the train higher. By contrast, Δh decreases in (b) giving a stiff efficient suspension. (c) shows a full null-flux track loop of a type which could be cheaply punched and folded from sheet. A more expensive design which would cause less drag at high velocities because of reduced skin effect losses consists of two multi-turn coils connected in series opposition. Both the magnet and track loops should be visualized as being extended periodically in the direction of train motion \vec{v} .

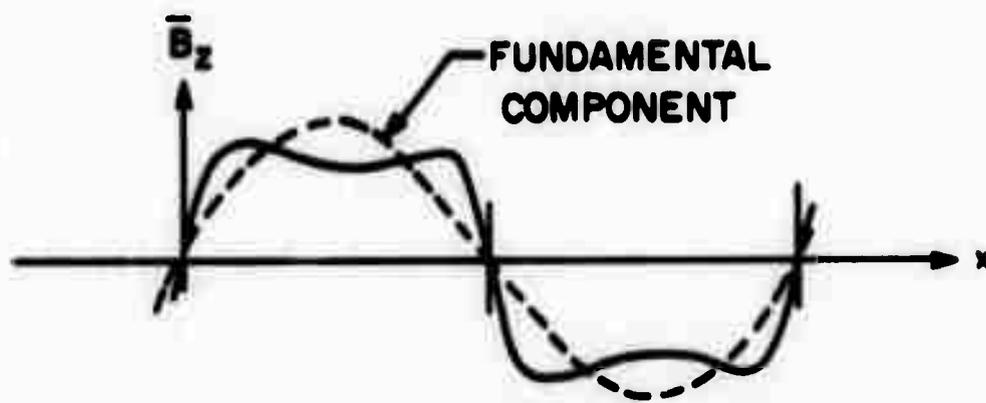


Figure 3. The z component of the average field due to a typical rectangular train magnet compared with its fundamental Fourier component. This field has been averaged over the width of the magnet in the y-direction.

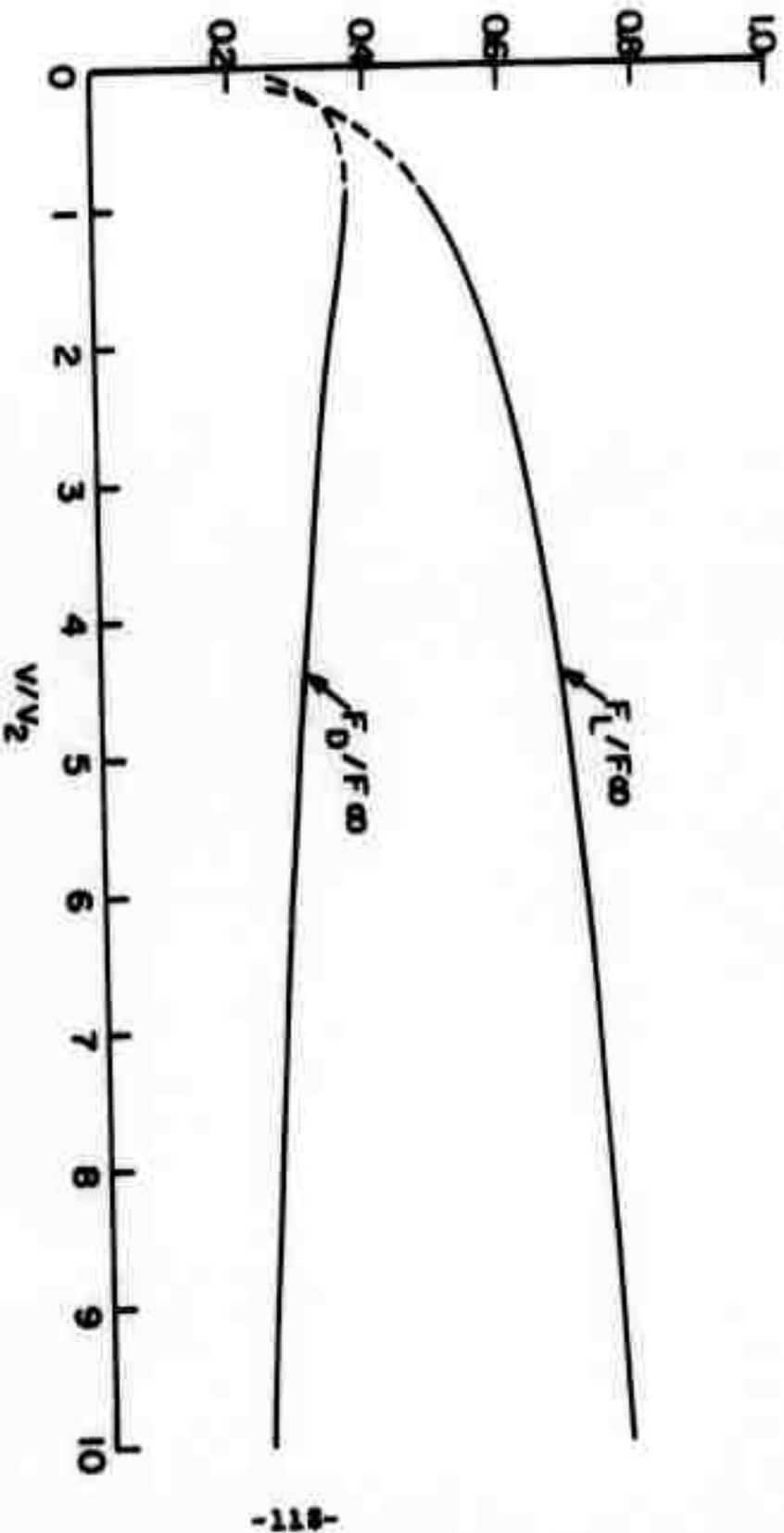


Figure 4.

Velocity dependence of normalized lift and drag forces at velocities high enough that the skin depth δ is small compared with both d , the ground plane thickness, and $l_x/2\pi$, where l_x is the period of the magnetic field. The approximation breaks down if $v \leq v_2 = c^2/l_x$, even if $d \rightarrow \infty$. To improve the accuracy of the drag curve near $v = v_2$, it has been plotted including a correction term of order $\approx k^2\delta^2$ to Eq. (13). At very low velocities, $F_D = v$ and $F_L = v^2$. Note that F_D/F_∞ never exceeds 0.37, and is very flat. Also, $F_L \rightarrow F_\infty$ very slowly.

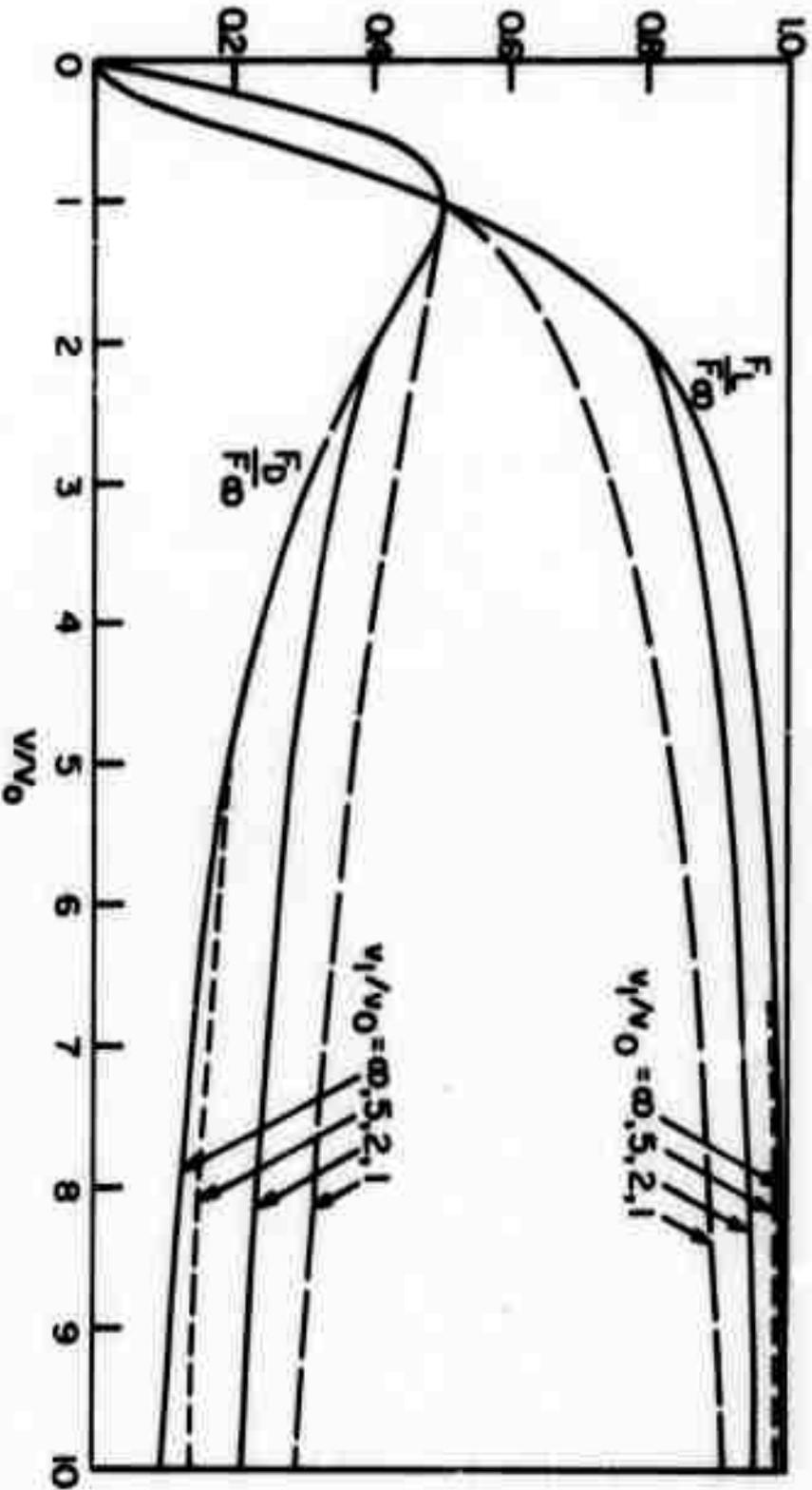


Figure 5. Velocity-dependence of lift and drag forces, normalized to F_∞ , the lift force for infinite velocity (or for a superconducting ground plane). v_0 is proportional to the resistance per square of the ground plane (see Eq. (7)) and is ≈ 6.5 km/h for 2.5 cm thick aluminum. v_1 is the velocity at which the skin depth equals the thickness d of the ground plane. The ratio $v_1/v_0 = l_x/2\pi d$, where l_x is the period of the magnet array.

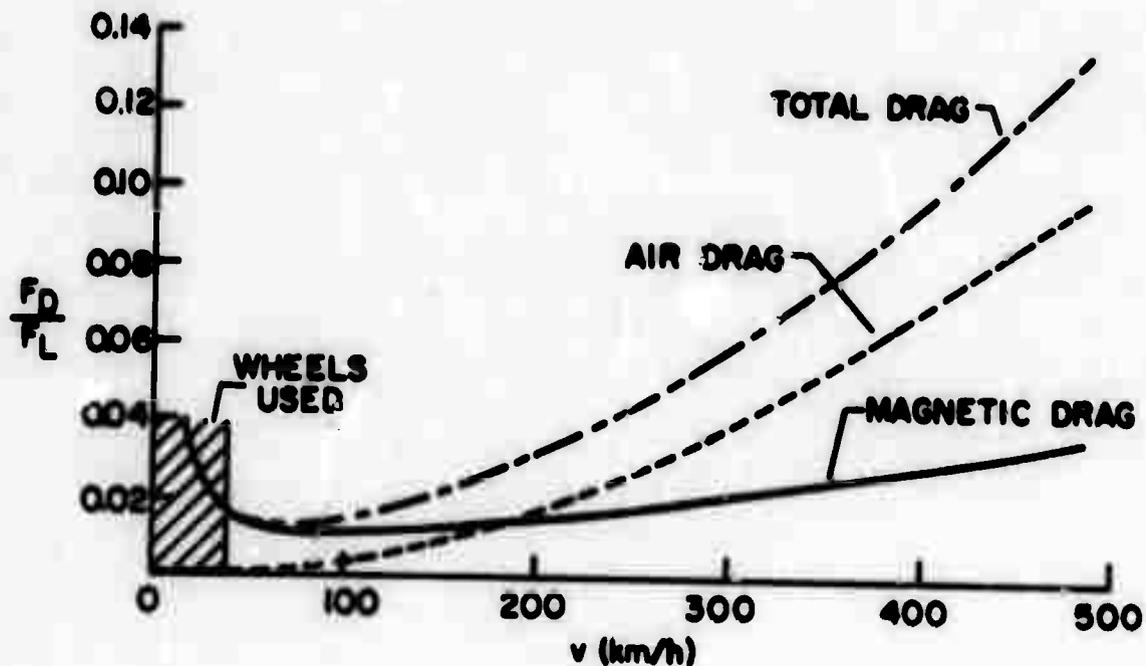


Figure 6. Velocity dependence of magnetic drag in the hybrid system of Section II.F, compared with the air drag. The parameters assumed are $F_D/F_L|_{\min} = 1/50$ and v_1 (the velocity for minimum magnetic drag) = 80 km/h. These values are appropriate for a 0.64 cm aluminum sheet, with $\Delta h = 0.64$ cm, and $l_x = 1.8$ m. The air drag is assumed to reach 0.1 of the weight at 500 km/h.

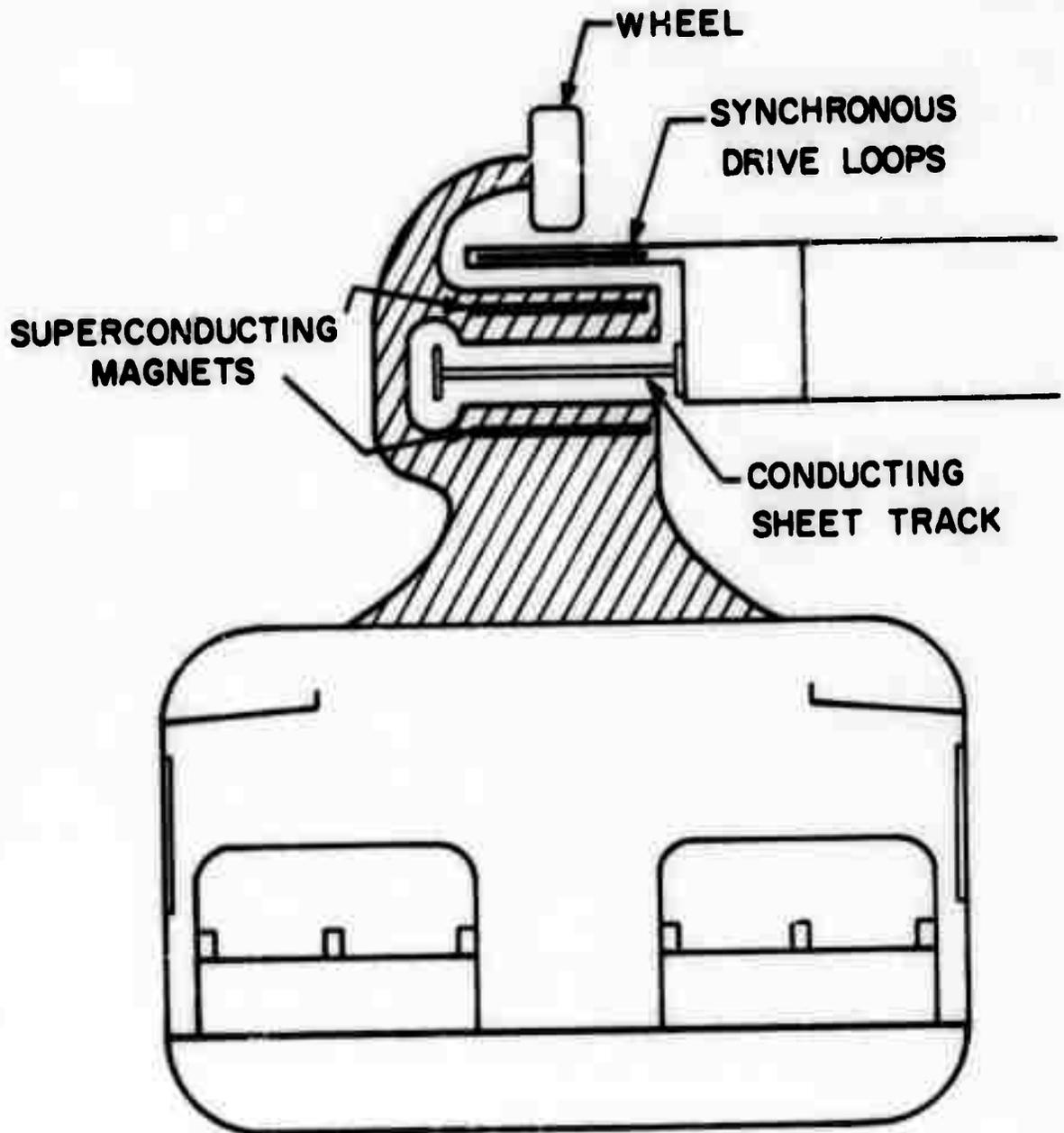
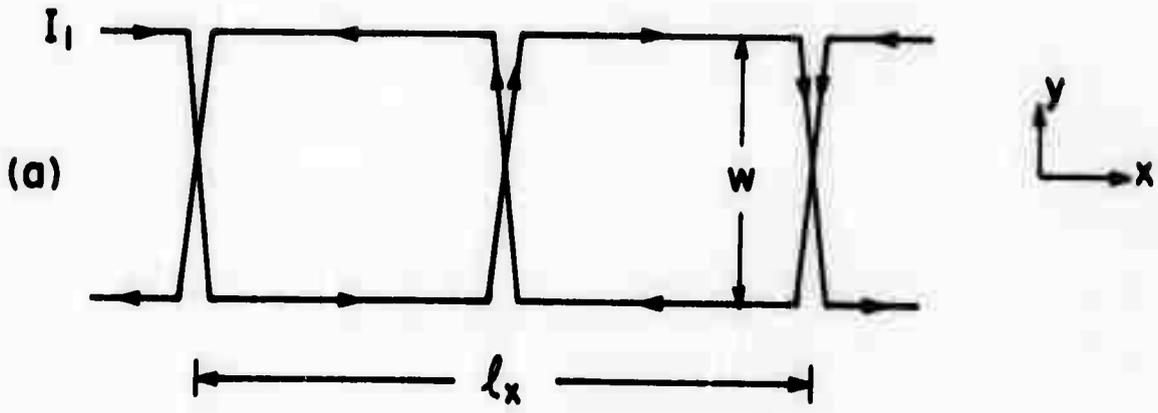
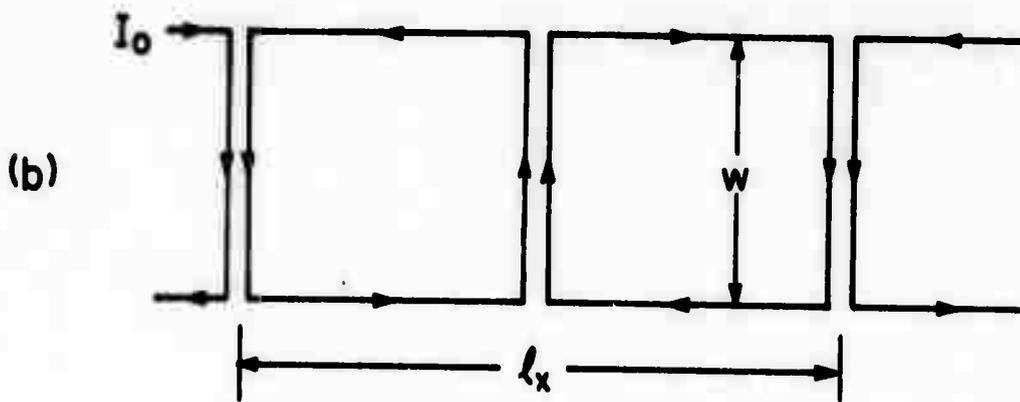


Figure 7. Schematic cross-section of train using hybrid null-flux levitation with a linear synchronous motor drive. Horizontal stability is maintained by the (non-null-flux) repulsion from the vertical pieces at the edges of the track. Alternatively, a "V" shaped magnet track configuration could be used to derive a centering force from gravity, or another orthogonal null-flux system could be used.



THRUST LOOP CONFIGURATION



TRAIN MAGNET CONFIGURATION

Figure 8. Schematic diagram of track thrust loop and train magnet configurations. An ac current is supplied to the normal metal track loops; a dc supercurrent circulates in the train loops.

CONSERVATION LAWS AND ENERGY RELEASE RATES

B. Budiansky
J. R. Rice

Abstract

New path-independent integrals recently discovered by Knowles and Sternberg are related to energy release rates associated with cavity or crack rotation and expansion. Complex-variable forms are presented for the conservation laws in the cases of linear, isotropic, plane elasticity. A special point concerning plastic stress distributions around cracks is discussed briefly.

CONSERVATION LAWS AND ENERGY RELEASE RATES

B. Budiansky
J. R. Rice

I. Introduction

The well-known J integral of fracture mechanics^{1,3,4} has been related to potential-energy release rates associated with moving or extending cracks in linear or non-linear elastic materials. Some new path-independent integrals (or conservation laws) have recently been discovered by Knowles and Sternberg². In this note these new laws are related to energy release rates associated with cavity or crack rotation and expansion rates. In addition, the conservation laws are displayed in complex-variable form for the case of linear, isotropic, plane elasticity. Finally, an implication concerning plastic stress distributions around cracks is discussed briefly.

II. Conservation Laws

Consider a two-dimensional deformation field for which the displacement vector u depends only on x_1 and x_2 . The J integral is $J = \oint_C (W dx_2 - T_i u_{i,1} dl)$ where C is a closed curve in the x_1, x_2 plane, W is the energy density, and T_i is the stress vector acting on the outer side of C . The J integral is actually the first component of the vector

$$J_k = \oint_C (W n_k - T_i u_{i,k}) d\ell \quad (1a)$$

where η is the unit outward normal to C , lying in the same plane. Each component of \underline{J} vanishes (trivially for J_3) for all closed paths C bounding a region in which W depends only on the strain $\eta_{ij} = (u_{i,j} + u_{j,i})/2$, and in which the stresses σ_{ij} , related to T_j on C by $\sigma_{ij} n_i$, satisfy

$$\sigma_{ij} = \frac{1}{2} \frac{\partial W}{\partial \eta_{ij}} + \frac{\partial W}{\partial \eta_{ji}},$$

and $\sigma_{ij,i} = 0$. This implies that J_α ($\alpha = 1, 2$) has the same value, not necessarily zero, for all paths that enclose a hole or crack.

The new Knowles-Sternberg integrals, in 2-D, are

$$L = \oint_C \epsilon_{sij} (W x_j n_i + T_i u_j - T_k u_{k,i} x_j) d\ell \quad (2a)$$

and

$$M = \oint_C (W x_i n_i - T_k u_{k,i} x_i) d\ell \quad (3a)$$

where ϵ_{ijk} is the alternating tensor. Under the same conditions specified for \underline{J} , L vanishes if, in addition, W depends only on the scalar invariants of η_{ij} . For M to vanish it is necessary that W be a quadratic function of the components of η_{ij} .

The results apply not only for any combination of plane and anti-plane straining, but also for the 2-D theory of generalized plane stress in which u and g denote thickness

averages. Also, with suitable redefinitions of the basic variables, J_α and L vanish when geometrical non-linearity is admitted^{2,3}; and if W is a homogeneous function of degree m in the strain components, M will still vanish if it is redefined by adding the quantity $(m-2)T_i u_i$ to its integrand.

In three dimensions the integrals generalize to

$$J_k = \int_S (W n_k - T_i u_{i,k}) dS \quad (1b)$$

$$L_k = \int_S c_{kij} (W x_j n_i + T_i u_j - T_l u_{l,i} x_j) dS \quad (2b)$$

$$M = \int_S (W x_i n_i - T_j u_{j,i} x_i - \frac{1}{2} T_i u_i) dS \quad (3b)$$

where S is a closed surface with outer normal \mathbf{n} , and analogous conservation theorems have been shown to hold by Knowles and Sternberg.

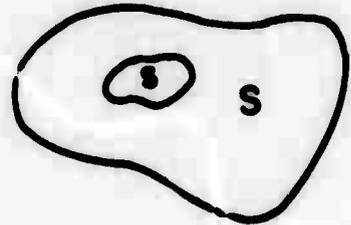
III. Energy Release Rates

Eshelby^{3,4} and Rice⁵ have shown that J_k can be interpreted as the energy release rate when a void or a crack tip is translated in position relative to a material body. We show that L_k and M have similar interpretations here.

Consider a 3-D elastostatic boundary value problem associated with the material contained within the surface $S+s$, for which the portion s of the boundary is traction-free, and external loading is imposed only by tractions on S . Without changing the boundary conditions on S , contemplate the contin-

ously varying sequence of static solutions for the displacements u generated as the spatial specification of s is varied with a time-like parameter t . The potential energy of the system at any time is

$$\Pi = \int_{V(t)} W \, dV - B[u]$$



where $V(t)$ is the volume enclosed by $S+s(t)$ and $B[u]$ is the potential of the loading specified on S . At each point in $V(t)$, W is a function of the time-varying strains compatible with u . Then

$$\frac{d\Pi}{dt} \equiv \dot{\Pi} = \int_{V(t)} \dot{W} \, dV - \frac{d}{dt} (B[u]) + \int_{s(t)} W \, v_j m_j \, ds$$

where v_i denotes the "velocity" of points on s and m_i is the current outward normal to s . (Note that only the normal component of v_i is determined uniquely by a given motion of the cavity boundary.) But by the principle of virtual work the first two terms in $\dot{\Pi}$ cancel (assuming that $\frac{du}{dt}$ is an admissible function in $V(t)$, so that

$$\dot{\Pi} = \int_s W \, v_j m_j \, ds$$

This result has been derived, less concisely, by Rice and Drucker⁶.

Next, suppose that s is the boundary of a cavity, and let $v_j = \delta_{ij}$; this corresponds to a conceptual translation of the cavity with a unit velocity in the i direction. Let $n_i = -m_i$ be the unit inward normal on the cavity surface; then $\xi_i \equiv -\dot{\Pi}$ is the rate of energy release per unit of cavity translation in the i direction, and is given by

$$\xi_i = \int_s W n_i ds$$

But by the first conservation law (1b), we have

$$\xi_i = J_i \quad (1c)$$

wherein the integral in J_i can be calculated on any closed surface surrounding the cavity. This last equation is Eshelby's³ result, and takes the same form as his equation for "forces" on point defects in solids⁴.

Next, consider a unit angular velocity about the i axis of the specified cavity wall; then $v_j = -\epsilon_{ijk} x_k$, and this leads to the result

$$\xi_i = -L_i \quad (2c)$$

for the rate of energy release per unit cavity rotation about the i axis (with the usual right-hand-rule sign convention for rotation.)

Finally, let the cavity boundary expand uniformly according to the rule $v_i = x_i$. This gives the energy release rate

$$\xi = M$$

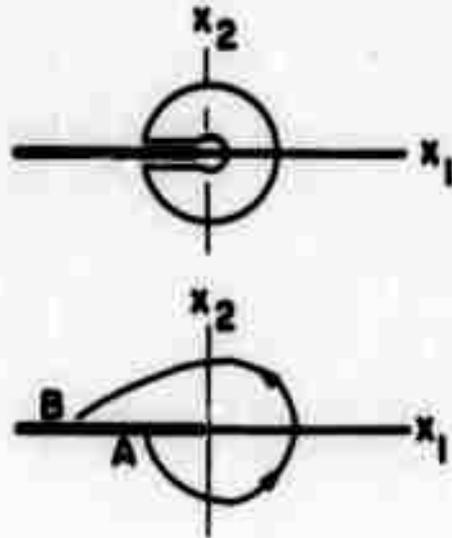
(3c)

Here the rate is with respect to relative scale change dl/l , where l is any characteristic length of the cavity.

The 2-D versions of these energy release relations, for plane or anti-plane conditions, are self-evident.

Since the final results are expressible as integrals over surfaces or curves off the cavity boundary, there is no reason to doubt their validity when the cavity is a crack.

In 2-D crack studies, the J ($=J_1$) integral has been exploited for closed paths of the type shown in the adjacent sketch. Since the integrand vanishes on the crack edges, it follows that the J integral has the same value for all open paths connecting any points A, B on opposite sides of the crack, as shown in the second sketch. In this case, the integral provides the energy release rate per unit crack tip extension^{1, 5}.



IV. Complex Variable Forms

In isotropic plane elasticity, the standard complex potentials $\phi(z)$ and $\psi(z)$ are analytic functions of $z = x+iy$ related to the stresses by

$$\sigma_{11} + \sigma_{22} = 2(\phi' + \bar{\phi}')$$

$$\sigma_{22} - \sigma_{11} + 2i\sigma_{12} = 2(\bar{z}\phi'' + \psi')$$

The following results have been derived for plane stress:

$$J_1 + iJ_2 = -\frac{2i}{E} \left[\oint_C (\phi')^2 dz - 2 \overline{\oint_C \phi' \psi' dz} \right] \quad (1d)$$

$$L = \frac{4}{E} \operatorname{Re} \oint_C z \psi \phi'' dz \quad (2d)$$

$$= \frac{4}{E} \operatorname{Re} \oint_C \psi' (\phi - z\psi') dz$$

$$M = \frac{4}{E} \operatorname{Im} \oint_C z \phi' \psi' dz \quad (3d)$$

where E is Young's modulus. For plane strain E should be replaced by $E/(1-\nu^2)$, where ν is Poisson's ratio. In the derivation of each formula, it was assumed that the region within C was free of any resultant forces, so that the potentials ϕ and ψ were single-valued.

If the integration path in (1a) is open, extending from A to B , the extra quantity

$$\frac{4i}{E} [z(\bar{\phi}')^2]_A^B$$

must be appended to (1d). Thus, the J integral, when taken around a crack tip, is just

$$J = \frac{2}{E} \operatorname{Im} \left(\int_A^B [(\phi')^2 + 2\phi' \psi'] dz + [\bar{z}(\phi')^2]_A^B \right)$$

Similarly, extra terms appear in (2d) and (3d) if the integration paths in (2a) and (3a) are open.

In the case of anti-plane shear, the relations

$$\sigma_{,2} + i\sigma_{,1} = w'(z)$$

$$u_3 = \frac{1}{2iG} (w - \bar{w})$$

provide the stresses and the displacement in terms of an analytic function $w(z)$; G is the shear modulus. Then formulas (1c), (2c), and (3c) can be written as

$$J_1 - iJ_2 = -\frac{1}{2G} \oint_C (w')^2 dz$$

and

$$L + iM = -\frac{1}{2G} \oint_C z (w')^2 dz$$

The same formulas also hold for open paths.

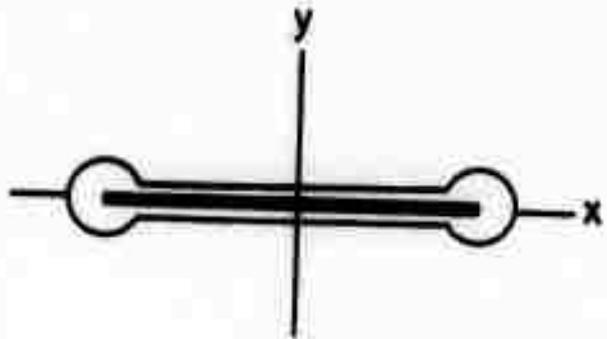
V. Crack-Tip Stresses in the Plastic Range

In conjunction with the simple deformation theory of plasticity, the J_1 integral has been used to calculate asymptotic plastic results^{5,7-9} for so-called "small-scale yielding" near the tips of cracks loaded in Modes I, II, and III. (Mode I, in the jargon of fracture mechanics, means crack opening, and Mode II is shearing parallel to the crack; in each case, the loading and geometry is symmetrical about the crack. Mode III is anti-plane shear.) In these solutions, the dominant part of the singular solution near the tip in the far plastic range

is determined to within a scalar factor; the factor is then found from the invariance of J_1 , evaluated for paths around the crack tip at small and large radii, where conditions are, respectively, purely plastic and purely elastic. But this method has been successful only for loadings that are purely in one of the three modes mentioned. Thus, for a mixture of Modes I and II, two scalar quantities are needed to establish the stress distribution near the crack tip, and the J_1 integral does not supply enough information for their determination.

Unfortunately, contrary to some initial hopes, use of the new L integral does not appear to supply the missing data in mixed-mode cases. It does, however, provide an unexpected result. The earlier pure-mode solutions indicated that W may be expected to vary inversely

with the distance r from the crack tip. Now consider the evaluation of L , for a mixed-mode situation on the path shown. Since L has a bounded,



unique value for all paths enclosing the crack, and since the contributions to L due to the small circles remain bounded as their radii shrink to zero, it follows that

$$\int_{\text{crack}} x[W_+ - W_-] dx$$

must be bounded, where W_+ and W_- are evaluated on the top and

bottom faces of the crack, respectively. This means, therefore, that the singular, $(1/r)$ parts of W must be equal at opposing points on either side of the crack. This is surprising when one considers arbitrary mixtures of Mode I and II. On the crack, W depends only on $|\sigma_x|$ in simple deformation theory. Hence, the equality of W_+ and W_- near the crack tip implies that the ratio of the dominant singularities in $(\sigma_x)_+$ and $(\sigma_x)_-$ must, in all mixed-mode cases, be either 1 or -1. This ratio is +1 for Mode I and -1 for Mode II. If, then, one imagines the loading mode to change continuously from I to II, this ratio will have to jump suddenly from 1 to -1 at some particular mode mixture.

The equality of the $(1/r)$ contributions to W on opposite sides of the crack must persist even if some Mode III, which has a τ_{xz} contribution to W , is also present. It may be noted, finally, that these conclusions could have been reached by a consideration of the J_2 integral.

Acknowledgements

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHCl5-71-C-0253 with The University of Michigan.

References

1. Rice, J. R.: "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks", J. of Applied Mechanics, June 1968, pp. 379-386.
2. Knowles, J. K. and Sternberg, Eli: "On a Class of Conservation Laws in Linearized and Finite Elastostatics", Technical Report No. 24, California Institute of Technology, July 1971.
3. Eshelby, J. D.: "The Energy Momentum Tensor in Continuum Mechanics", in Inelastic Behavior of Solids, Ed. by M. F. Kanninen, et al, McGraw-Hill, 1970.
4. Eshelby, J. D.: "The Continuum Theory of Lattice Defects", in Solid State Physics, Vol. III, Ed. F. Seitz and D. Turnbull, Academic Press, 1956.
5. Rice, J. R.: "Mathematical Analysis in the Mechanics of Fracture", in Fracture, An Advanced Treatise, Vol. II, Ed. by H. Liebowitz, Academic Press, 1968, pp. 191-311.
6. Rice, J. R. and Drucker, D. C.: "Energy Changes in Stressed Bodies due to Void and Crack Growth", Int. J. Fracture Mech., Vol. 3, 1967, pp. 19-27.
7. Rice, J. R. and Rosengren, G. F.: "Plane Strain Deformation Near a Crack Tip in a Power-Law Hardening Material", J. Mech. Phys. Solids, Vol. 16, 1968, pp. 1-12.
8. Hutchinson, J. W.: "Singular Behavior at the End of a Tensile Crack in a Hardening Material", J. Mech. Phys. Solids, Vol. 16, 1968, pp. 13-31.
9. Hutchinson, J. W.: "Plastic Stress and Strain Fields at a Crack Tip", J. Mech. Phys. Solids, Vol. 16, 1968, pp. 337-347.

APPLICATION OF A DEFECT MODEL
TO VITREOUS SOLIDS

R. A. Huggins

Abstract

The translation of the concepts of defect chemistry developed for crystalline solids to the description of the structure of one class of "amorphous" materials, the vitreous oxides, is discussed. Of specific interest are the types and concentrations of unusual local structural configurations and the influence of oxygen activity and solute concentrations.

APPLICATION OF A DEFECT MODEL
TO VITREOUS SOLIDS

R. A. Huggins

Introduction

Partly as a result of recent progress and controversies relating to various switching and memory effects in non-crystalline semiconductors, greatly increased attention is presently being directed toward understanding the structure and properties of a wide variety of "amorphous" materials. One of the obvious procedures that is being followed in many quarters involves the investigation of the extent to which the concepts and tools that have been developed to a considerable degree of refinement for use in other areas, such as the physics of crystalline metals and semiconductors, can be applied to the study of solids which do not have structures with long range order. Although there is much to be gained by building upon this knowledge base, one has to be quite careful to avoid blind translation from crystalline systems without giving attention to their appreciable differences, as well as their similarities.

One of the obvious important differences that is found between these two general classes of materials involves the relative importance of short range, compared to long range, interactions. As a result, models and tools that relate more

directly to short range or localized features and phenomena should be particularly appropriate to investigate with regard to their application to noncrystalline systems.

A general concept that has been especially valuable in dealing with many phenomena and properties in crystalline insulators and semiconductors is what is sometimes called the defect model of the solid state. The basic approach that is used involves the focussing of attention upon the description and behavior of structural perturbations; that is, upon deviations from the normal or "ideal" structure. Deviations from the normal regularly repeating structure within crystals involve linear defects (dislocations) and so-called point defects, such as vacancies, interstitials, substitutional atoms or ions, etc. In the case of "electrically-active" solids one of the powerful techniques that has been developed for describing and handling such structural features and the properties that relate to them is known as defect chemistry. In this approach, structural defects are treated as quasi-chemical species, and many of the standard techniques of chemical thermodynamics and kinetics applied to them. This has been a very useful methodology, and has helped achieve an understanding of many of the structure-dependent properties of crystalline solids. It has been particularly valuable and straightforward when dealing with materials in which the individual defects can be considered to be essentially non-interacting. Defect interactions can also be handled by making

use of the same procedures employed in concentrated solution theory, but at the price of a reduction in simplicity.

One group of amorphous materials to which this approach seems most appropriate involves simple vitreous, or glassy, oxides. On the basis of relatively recent experimental work using optical spectroscopy, EPR, NMR and sophisticated X-ray diffraction techniques, it is now well recognized that such vitreous materials contain very well-defined local configurations. These local structural arrangements are virtually identical to those found in similar crystalline solids; this has specifically been pointed out for silicates and borates.¹⁻⁵

A start has already been made toward the application of the principles of defect chemistry to such simple vitreous systems,⁶⁻⁸ using the "structon" as the elementary structural unit. Each type of structon⁹⁻¹² signifies a particular type of atom with specified kinds and numbers of close neighbors. The overall structure of a vitreous material can be expressed as the sum of all structons present, both those which are "normal" (analogous to the regular repeating structure of a perfect crystal) and those which are different (analogous to crystalline defects). The concepts and techniques of defect chemistry can be applied to these different structons, called "defect structons", as well as to electronic perturbations ("extra" electrons or holes). The equilibrium concentrations of the various types of defect structons, electrons and holes within the vitreous structure can be found as functions of the

macroscopic thermodynamic variables; specifically useful are the influence of temperature and the concentrations or chemical activities of the various chemical elements present.

In the case of silica, the structure of both the crystalline^{2,3} and vitreous¹ states can be expressed in terms of just two types of normal structons; a silicon surrounded tetrahedrally by four oxygens, and an oxygen bridging between two silicon neighbors. Both of these structon types can be treated as having a zero effective charge relative to the rest of the structure.

Deviations from this normal (or ideal) vitreous structure can be described either in terms of a set of defect structons centered about electropositive atoms (silicon) or a different group centered upon the electronegative species (oxygen). The total close-neighbor structure can be determined by either set of structons, as they are directly related to each other, but in vitreous oxides it is more convenient to deal only with oxygen-centered structons.

It has been shown^{6,7} that the most probable oxygen-centered defect structons in high purity vitreous silica are of just two types, oxygen atoms with either one (called a "non-bridging" oxygen) or three close silicon neighbors. In these cases the formal charges associated with the structons are -1 and +1, respectively. There is a large body of experimental evidence for the existence of non-bridging oxygens in silicates containing cations such as H^+ , Li^+ , Na^+ or K^+

in addition to the basic constituents silicon and oxygen. The second type of defect structure has not yet been reported experimentally in SiO_2 , undoubtedly because of the practical difficulty of obtaining silica without the presence of small quantities of such additional cations. Recent observations^{13,14} that GeO_2 can contain a number of percent excess Ge indicates that a positively charged structure should be observable in that material.

A set of symbols has been introduced⁷ to describe both the normal (background) structures and the defect structure species present in vitreous silica and dilute silicates. These are illustrated in Table I.

General Methodology Used in Calculation of Defect Equilibria

Analogous to the procedure normally followed in the defect chemistry of simple nonmetallic crystals, defect equilibrium in vitreous oxides can be expressed in terms of quasi-chemical reactions involving the pertinent defect structure species. Standard thermodynamic treatment results in law of mass action expressions in terms of the chemical activities or concentrations of such species.

If proper attention is given to the construction of these quasi-chemical reactions with regard to both mass and electrostatic charge balances, a completely rigorous result is obtained in terms of the activities of the pertinent species. Translation from activities to concentrations involves the

introduction of assumptions, as is the case with crystalline materials. Because of the relative importance of short range, rather than long range, interactions in the common vitreous materials, it is reasonable to expect simple linear relations between activities and concentrations in such systems to remain valid up to quite large concentrations.

The selection of the appropriate quasi-chemical reactions to be used to solve defect equilibrium problems is straightforward. The Gibbs phase rule indicates that at a specified value of total pressure and temperature a system in equilibrium is completely determined if a number of compositional variables are specified that is one less than the number of chemical components present. For this purpose, we can consider each chemical element to be a component. Thus, in a binary system such as pure silica, in which the components are silicon and oxygen, one composition-determining relation is needed to establish the thermodynamic state, and thus the concentrations of all the structural species present, if the pressure and temperature are specified. In a ternary system, such as sodium silicate, two such composition-determining relationships are needed in addition to the pressure and temperature.

These compositional relations, which can be considered as thermodynamic constraints, can be of various types. The concentration (or chemical activity) of one of the components (or one of the structural species) is one type of such

relation. The ratio of specie concentrations (or activities) is another. In some cases, the concentrations or activities of species are determined by equilibrium with an adjacent phase of known properties and the constraint can be expressed in terms of a quasi-chemical reaction across the phase boundary.

In addition to these composition-determining relations, other independent relationships are needed, as the total number of constraints must equal the number of structural species under consideration.

A second type of independent relation involves the formation of conjugate defect pairs. One of these is electron-hole pair formation. Another is the formation of a pair of structural defects. (In the application of defect chemistry to crystalline solids these latter species would be vacancies, interstitials, etc.) Obviously, the selection of the specific defects to be included in this defect pair formation reaction depends upon the disorder or structural model that is being used to describe the vitreous state.

One further constraint must always be considered when treating defects in nonmetallic materials. This is the requirement for electrostatic neutrality; that is, the deviation in electrostatic charge balance contributed by the presence of all the negatively charged defect species must be balanced by that contributed by the presence of the various positively charged species.

As a general rule, the total number of defect species

which are involved in the pertinent equilibria in simple two-component systems within any given range of chemical constitution is four. Two of these four are structural defects; the other two are electronic. As a result, four independent relations need to be solved simultaneously in order to determine the concentrations of this group of defects. One of these is always the electroneutrality condition; the second is the electronic defect pair formation reaction, the third is a structural defect pair formation reaction, and the fourth is some type of composition-determining relation. If other defect species are also assumed to exist (including defects with different ionization states or effective charges) further independent reactions are obviously required.

After an appropriate set of simultaneous equations has been selected, the concentrations of all pertinent defect species can be obtained as functions of the prescribed thermodynamic variables. For purposes of simplification, it is often found useful to approximate the electroneutrality condition by a simple equality between the concentrations of the two dominant defects (one negative and one positive) in any given range of composition.¹⁵ By writing this simplified electroneutrality condition, and expressing the law of mass action expressions for the appropriate quasi-chemical reactions in logarithmic form, a simple set of simultaneous linear equations results, which can easily be solved. A number of examples of the use of this method have been presented.¹⁶⁻²⁰

Application to Defect Equilibria in Pure Vitreous Silica

The calculation of the defect concentrations in pure silica⁷ involves the simultaneous solution of a set of four independent equations:

$$K_i = [e^-] [h^+], \quad (1)$$

$$K_d = [<s^->] [<s^+>], \quad (2)$$

$$K_o = [<s^->]^2 [h^+]^2 p^{-1/2}, \quad (3)$$

and

$$[<s^->] + [e^-] = [<s^+>] + [h^+], \quad (4)$$

where the K values are temperature-dependent equilibrium constants and the square brackets indicate concentrations. Simultaneous solution produces the following expressions for the pertinent defect concentrations:

$$[e^-] = K_i \left[\frac{K_d + F}{F(K_i + F)} \right]^{1/2}, \quad (5)$$

$$[h^+] = \left[\frac{F(K_i + F)}{K_d + F} \right]^{1/2}, \quad (6)$$

$$[<s^->] = \left[\frac{F(K_d + F)}{K_i + F} \right]^{1/2}, \quad (7)$$

and

$$[<s^+>] = K_d \left[\frac{K_i + F}{F(K_d + F)} \right]^{1/2}, \quad (8)$$

where

$$F = [K_o p^{1/2}]^{1/2} \quad (9)$$

and p represents the partial pressure of oxygen gas.

These equations can be simplified considerably under certain situations. For example, because of the large band gap, it is reasonable to assume that the concentrations of the two types of defect structures are much greater than those of either electrons or holes over a broad range of intermediate values of oxygen partial pressure. When this is the case, the electroneutrality relation can be approximated by

$$[<s^->] \approx [<s^+>] \quad (10)$$

and

$$K_d \gg F \gg K_i \quad ,$$

so that equations (5-8) simplify to:

$$[e^-] = \frac{K_i K_d^{1/2}}{F} \quad (11)$$

$$[h^+] = \frac{F}{K_d^{1/2}} \quad (12)$$

$$[<s^->] = K_d^{1/2} \quad (13)$$

and

$$[<s^+>] = K_d^{1/2} \quad (14)$$

Thus, we see that in this range of oxygen partial pressure the concentrations of both minority defects (electrons and holes) depend on the magnitude of the oxygen partial pressure, while the concentrations of the dominant defect structures do not.

Influence of the Presence of a Third Component

Even nominally pure silica typically contains at least

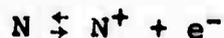
one other element in addition to silicon and oxygen, and many silicates contain appreciable concentrations of one or more other species. The most common case involves the presence of monovalent cations such as lithium, sodium, potassium, or hydrogen. The presence of the latter is related to the frequent observation of OH^- bands in the optical absorption spectrum. The presence of these other species in solution in the silica must be taken into account in considerations of the structure of the silicate in terms of its defect description. In order to specify the defect equilibria when an additional species is present, an additional thermodynamic constraint must also be stated, increasing the number of equations to be solved simultaneously. The following cases can be considered:

1. An additional ionic species is present in fixed concentration. This involves the use of an equation of the form

$$[\text{M}^+] = M \tag{15}$$

where M is a constant.

2. An additional species is present whose chemical activity is determined by equilibrium with an adjacent phase, either external or present as a precipitate. This involves the use of a quasi-chemical reaction of the form



which leads to an equilibrium relation of the form

$$[N^+] = \frac{K_N a_N}{[e^-]} \quad (16)$$

3. The activity of the oxide of the third component could be specified. In the case of a monovalent cation L^+ this would mean specification of the activity (related to the concentration) of L_2O . The oxide activity is related to the activities of both the cation and oxygen by the formation reaction



which leads to the law of mass action relation

$$K_{L_2O} = \frac{a_{L_2O}}{(a_L)^2 p^{1/2}} \quad (17)$$

Furthermore, if essentially all the L is ionized (present as L^+) we can write



and

$$K_L = \frac{[L^+][e^-]}{a_L} \quad (18)$$

Substituting into (17), we see that

$$K_{L_2O} = \frac{(a_{L_2O}) K_L^2}{p^{1/2} [L^+]^2 [e^-]^2} \quad (19)$$

which can be rearranged to give

$$[L^+] = \frac{1}{[e^-]} \left[\frac{K_L^2 a_{L_2O}}{K_{L_2O} p^{1/2}} \right]^{1/2} \quad (20)$$

The way that these different relations enter the defect equilibrium calculation can be seen by examination of the electroneutrality condition, which can be written as

$$[e^-] + [s^-] = [h^+] + [s^+] + [M^+] \quad (21)$$

where $[M^+]$ can be replaced by either $[N^+]$ or $[L^+]$. If substitutions are made to write this equation in terms of the electron concentration $[e^-]$, it can be written as

$$[e^-] \left[1 + \frac{F}{K_i} \right] = \frac{1}{[e^-]} \left[K_i + \frac{K_d K_i}{F} + S \right] \quad (22)$$

where $S_1 = M/[e^-]$ for case 1,

$S_2 = K_N a_N$ for case 2, and

$S_3 = \left[\frac{K_L^2 a_{L_2O}}{K_{L_2O} p^{1/2}} \right]^{1/2}$ for case 3.

Solving for $[e^-]$,

$$[e^-] = \left[\frac{K_i}{K_i + F} \right] (Y) \quad (23)$$

where

$$Y_1 = \left[\frac{M}{2} + \left(\frac{M^2}{4} + \frac{(K_d + F)(K_i + F)}{F} \right)^{1/2} \right] \quad \text{for case 1,}$$

$$Y_2 = \left[\frac{(K_d + F)(K_i + F)}{F} + \left(\frac{K_i + F}{K_i} \right) (K_N a_N) \right]^{1/2} \quad \text{for case 2, and}$$

$$Y_3 = \left[\frac{(K_d + F)(K_i + F)}{F} + \left(\frac{K_i + F}{K_i} \right) \left(\frac{K_L^2 a_{L_2O}}{K_{L_2O} p^{1/2}} \right)^{1/2} \right]^{1/2} \quad \text{for case 3.}$$

Similar relations can be derived for the concentrations of the other defect species, producing the following set:

$$[e^-] = \left(\frac{K_1}{K_1 + F} \right) [Y] \quad (23)$$

$$[h^+] = (K_1 + F) [Y]^{-1} \quad (24)$$

$$[<s^->] = \left(\frac{F}{K_1 + F} \right) [Y] \quad (25)$$

$$[<s^+>] = \frac{K_d (K_1 + F)}{F} [Y]^{-1} \quad (26)$$

and either

$$[M^+] = M, \quad (15)$$

$$[N^+] = (K_N a_N) \left(\frac{K_1 + F}{K_1} \right) [Y_2]^{-1}, \quad (27)$$

or

$$[L^+] = \left[\frac{K_L^2}{K_{L_2O}} \frac{a_{L_2O}}{p^2} \right]^{1/2} \left(\frac{K_1 + F}{K_1} \right) [Y_3]^{-1} \quad (28)$$

In the important central region of oxygen partial pressure,

$$K_d \gg F \gg K_1$$

and these relations can be greatly simplified, to give

$$[e^-] \approx \left(\frac{K_1}{F} \right) [Z] \quad (29)$$

$$[h^+] \approx F [Z]^{-1} \quad (30)$$

$$[< s^- >] = [Z] \quad (31)$$

$$[< s^+ >] = K_d [Z]^{-1} \quad (32)$$

where

$$z_1 = \frac{M}{2} + \left[K_c + \frac{M^2}{4} \right]^{1/2}$$

$$z_2 = \left[K_d + \left(\frac{F}{K_1} \right) K_N a_N \right]^{1/2}$$

$$z_3 = \left[K_d + \left(\frac{F}{K_1} \right) \left(\frac{K_L^2}{K_{L_2O}} \frac{a_{L_2O}}{p^{1/2}} \right)^{1/2} \right]^{1/2}$$

Also, in this range

$$[M^+] = M, \quad (15)$$

$$[N^+] = \left(K_N a_N \right) \left(\frac{F}{K_1} \right) [z_2]^{-1}, \quad (33)$$

and

$$[L^+] = \left[\frac{K_L^2}{K_{L_2O}} \frac{a_{L_2O}}{p^{1/2}} \right]^{1/2} \left(\frac{F}{K_1} \right) [z_3]^{-1} \quad (34)$$

Influence of Oxygen Partial Pressure on Solubility

From these relations one can readily obtain information about the influence of the oxygen partial pressure upon the solubility of a constituent in a ternary silicate system. If the species N precipitates as elemental N, for which the value of a_N is defined as unity, the solubility of N is directly found from equation (33) to be

$$[N^+]_{\max} = \left(\frac{K_N F}{K_1} \right) \left[K_d + \frac{K_N F}{K_1} \right]^{-1/2} \quad (35)$$

This can be written as

$$[N^+]_{\max} \approx \left(\frac{K_{NF}}{K_i} \right)^{\frac{1}{2}} \quad (36)$$

if $\left(\frac{K_{NF}}{K_i} \right) \gg K_d$

and

$$[N^+]_{\max} \approx \frac{K_{NF}}{K_i K_d^{\frac{1}{2}}} \quad (37)$$

if $K_d \gg \left(\frac{K_{NF}}{K_i} \right)$

There is also a temperature dependence to the solubility, as each of the K values is expected to vary with temperature according to a relation of the form

$$K = K_0 \exp(-Q/RT) \quad (38)$$

On the other hand, if the solute precipitates in the form of its oxide, its solubility is found from equation (34) by setting the oxide activity equal to unity. Using the symbol L,

$$[L^+]_{\max} \approx \left[\frac{K_0^{\frac{1}{2}} K_L}{K_i K_{L_2O}^{\frac{1}{2}}} \right]^{\frac{1}{2}} \quad (39)$$

if $\left(\frac{K_0^{\frac{1}{2}} K_L}{K_i K_{L_2O}^{\frac{1}{2}}} \right) \gg K_d$

and

$$[L^+]_{\max} \approx \frac{K_o^{1/2} K_L}{K_i K_{L_2O} K_d^{1/2}} \quad (40)$$

if $K_d \gg \left(\frac{K_o^{1/2} K_L}{K_i K_{L_2O}} \right)$

Thus it is seen that the oxygen partial pressure has an important influence upon the solubility of a third component in silicates if that component precipitates in elemental form. On the other hand, if it precipitates as an oxide, the solubility will be independent of oxygen partial pressure.

Influence of One Solute on the Solubility of Another

The presence of one solute can influence the solubility of another. We can consider the influence of a fixed concentration of M^+ upon the solubility of N^+ . The latter species can be assumed to precipitate as elemental N, so that its solubility is reached when a_N is equal to unity.

In this case there are six defect concentrations to consider, and therefore must be six independent relations.

These are

$$K_i = [e^-][h^+] \quad (1)$$

$$K_d = [s^-][s^+] \quad (2)$$

$$K_o = [s^-]^2 [h^+]^2 p^{-1/2} \quad (3)$$

$$[M^+] = M \quad (15)$$

$$[N^+] = \frac{K_N a_N}{[e^-]} \quad (16)$$

and

$$[e^-] + [s^-] = [h^+] + [s^+] + [M^+] + [N^+] \quad (41)$$

Simultaneous solution of these relations results in the following expression for the solubility of N,

$$[N^+]_{\max} = \frac{K_N (K_i + F)}{K_i} \left[\frac{M}{2} + R^{1/2} \right]^{-1} \quad (42)$$

where

$$R = \frac{M^2}{4} + \frac{(K_d + F)(K_i + F)}{F} + \frac{K_N (K_i + F)}{K_i} \quad (43)$$

This can be simplified somewhat within the important practical range in which $K_d \gg F \gg K_i$

to give

$$[N^+]_{\max} \approx \frac{K_N F}{K_i} \left[\frac{M}{2} + \left[\frac{M^2}{4} + K_d + \frac{K_N F}{K_i} \right]^{1/2} \right]^{-1} \quad (44)$$

Further, if

$$\frac{M^2}{4} \gg K_d$$

and

$$\frac{M^2}{4} \gg \frac{K_N F}{K_i}$$

this reduces to

$$[N^+]_{\max} \approx \frac{K_N F}{K_i M} \quad (45)$$

Polyvalent Ions

Although the cases that have been described here all involve monovalent solute ions, similar considerations apply for polyvalent species. The results, although different, will be similar to those found for monovalent species, and will not be included here.

Redox Reactions in Vitreous Materials

It has long been known that many ions can undergo changes in valence or effective charge in oxide glasses. Summaries of the early work in this area have been compiled^{21,22}, but as was pointed out by Johnston²³, insufficient attention was given to the achievement of equilibrium in many cases. A list of multivalent ions that have been shown to undergo redox reactions in silicates is presented in Table II.

The factors involved in the change of ionization state of solute species within vitreous materials should be amenable to treatment by the same techniques described on the previous pages. This will not be pursued here, however, but will be included in a subsequent report.

Discussion

A general methodology has been developed for the treatment of structural equilibria within vitreous materials such as the oxide silicates analogous to the defect chemistry which has been successfully used for many crystalline materials. Several examples have been presented of the use of these

techniques; others will follow in later reports. One of the conclusions that results is the significant influence that the oxygen partial pressure during melting should have upon the structure and the properties of vitreous oxides. Although this factor has not received a great deal of attention in glass technology to date, the few studies have been reported tend to confirm the validity of this general approach.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

TABLE I

Structural Units in Vitreous Silica and Dilute Vitreous
Silicates Containing Oxides of Monovalent Cations

<u>Structural Species</u>	<u>Structon Formula</u>	<u>Structural Diagram</u>	<u>Formal Charge</u>	<u>Simplified Notation</u>
<u>Normal Structons</u>				
1. Predominant Si-centered structon	$\langle \text{Si}(4 \cdot \text{O}''') \rangle$	$\begin{array}{c} \\ \text{O} \\ \\ -\text{O}-\text{Si}-\text{O}- \\ \\ \text{O} \\ \end{array}$	0	
2. Predominant Oxygen-centered structon (bridging oxygen)	$\langle \text{O}''(2\text{Si}) \rangle$	Si-O-Si	0	
<u>Defect Structons</u>				
1. Silicon-centered structon with one oxygen neighbor nonbridging	$\langle \text{Si}(\text{O}', 3 \cdot \text{O}''') \rangle$	$\begin{array}{c} \\ \text{O} \\ \\ -\text{O}-\text{Si}-\text{O}^- \\ \\ \text{O} \\ \end{array}$	-1	
2. Silicon-centered structon with one oxygen neighbor forming a 3-way bridge	$\langle \text{Si}(3 \cdot \text{O}''', \text{O}'') \rangle$	$\begin{array}{c} \\ \text{O} \\ \\ -\text{O}-\text{Si}-\text{O} \begin{array}{l} / \\ \backslash \end{array} \\ \\ \text{O} \\ \end{array}$	+1	
3. Nonbridging oxygen structon	$\langle \text{O}'(\text{Si})^- \rangle$	Si-O ⁻	-1	$\langle \text{S}^- \rangle$
4. Most probable oxygen-centered structon with positive charge (3-way bridging oxygen)	$\langle \text{O}'''(3\text{Si})^+ \rangle$	$\begin{array}{c} \text{Si} \\ / \\ \text{Si}-\text{O}^+ \\ \backslash \\ \text{Si} \end{array}$	+1	$\langle \text{S}^+ \rangle$
5. Monovalent cation (eg. Na, K)	$\langle \text{M}^+ \rangle$		+1	$\langle \text{M}^+ \rangle$
<u>Electronic Defects</u>				
1. Electron			-1	e ⁻
2. Hole			+1	h ⁺

TABLE II

Ions With Variable Valence Within Silicates

<u>Ion</u>	<u>References</u>
Ce ³⁺ - Ce ⁴⁺	22-27
Co ²⁺ - Co ³⁺	23, 28
Cr ²⁺ - Cr ³⁺	22, 24, 25, 29
Eu ²⁺ - Eu ³⁺	30
Fe ²⁺ - Fe ³⁺	22-25, 31-34
Mn ²⁺ - Mn ³⁺	22-24
Ni ²⁺ - Ni ³⁺	23
Ti ³⁺ - Ti ⁴⁺	23
V ³⁺ - V ⁴⁺	23, 35, 36
V ⁴⁺ - V ⁵⁺	23, 35, 36
As ³⁺ - As ⁵⁺	22, 24
Sb ³⁺ - Sb ⁵⁺	22-24
Sn ²⁺ - Sn ⁴⁺	22-24

References

1. R. L. Mozzi and B. E. Warren, *J. Appl. Cryst.* 2, 164 (1969).
2. M. L. Huggins, *Macromolecules* 1, 184 (1968).
3. M. L. Huggins, *Inorg. Chem.* 7, 2108 (1968).
4. M. L. Huggins, "The Structon Theory. Stabilities of Different Structon Types", presented at Bragg Symposium, London (1970).
5. M. L. Huggins, *Inorg. Chem.* 10, 791 (1971).
6. R. A. Huggins, "Structural Model of Vitreous Oxides Based Upon Concepts of Defect Chemistry", Report of ARPA Materials Research Council (1969).
7. R. A. Huggins and M. L. Huggins, *J. Solid State Chem.* 2, 385 (1970).
8. R. A. Huggins, "On the Platinum Precipitation Problem in Laser Glass", Report of ARPA Materials Research Council (1970).
9. M. L. Huggins, *J. Phys. Chem.* 58, 1141 (1954).
10. M. L. Huggins, *J. Amer. Ceram. Soc.* 38, 172 (1955).
11. M. L. Huggins, *Bull. Chem. Soc. Japan* 28, 606 (1955).
12. M. L. Huggins, *J. Amer. Chem. Soc.* 77, 3928 (1955).
13. J. de Neufville, Drummond and D. Turnbull, *Phys. & Chem. of Glasses* (1970).
14. J. de Neufville and D. Turnbull, to be published in *Disc. of Faraday Soc.*
15. G. Brouwer, *Philips Res. Rep.* 9, 366 (1954).
16. F. A. Kröger and H. J. Vink, *Solid State Physics* 3, 307 (1956).
17. H. J. Vink, *Proc. Int. School of Physics, Enrico Fermi, Course 22*, R. A. Smith, ed., Academic Press (1963).
18. F. A. Kröger, Chemistry of Imperfect Crystals, North-Holland Pub. Co., Amsterdam (1964).

19. H. J. Vink, *Festkoerperprobleme* 4, 205 (1965).
20. F. A. Kröger, in *Physical Chemistry*, Vol. X, ed. by Eyring, Henderson and Jost, Academic Press (1970).
21. W. A. Weyl, *Coloured Glasses*, Dawsons of Pall Mall, London (1959).
22. H. J. Tress, "A Thermodynamic Approach to Redox Equilibria in Glasses", *Phys. Chem. Glasses* 1, 196 (1960).
23. W. D. Johnston, "Oxidation-Reduction Equilibria in Molten $\text{Na}_2\text{O} \cdot 2\text{SiO}_2$ Glass", *J. Am. Ceram. Soc.* 48, 184 (1965).
24. C. Kühl, H. Rudow and W. Weyl, "Oxydations - und Reduktionsgleichgewichte in Farbgläsern", *Sprechsaal* 71, 118 (1938).
25. R. W. Douglas, P. Nath and A. Paul, "Oxygen Ion Activity and Its Influence on the Redox Equilibrium in Glasses", *Phys. Chem. Glasses* 6, 216 (1965).
26. A. Paul and R. W. Douglas, "Cerous-Ceric Equilibrium in Binary Alkali Borate and Alkali Silicate Glasses", *Phys. Chem. Glasses* 6, 212 (1965).
27. V. Gottardi, G. Paoletti, and M. Tornati, "The Ratio of $\text{Ce}^{3+}/\text{Ce}^{4+}$ in the Melting of Different Glasses and Its Influence on Their Properties", *Advances in Glass Technology. VI International Congress on Glass, Washington, D.C. (1964)* p. 412.
28. A. Dietzel and M. Coenen, "Über dreiwertiges Kobalt in Gläsern hohen Alkaligehaltes", *Glastech. Ber.* 34, 49 (1961).
29. P. Nath and R. W. Douglas, " $\text{Cr}^{3+} - \text{Cr}^{6+}$ Equilibrium in Binary Alkali Silicate Glasses", *Phys. Chem. Glasses* 6, 197 (1965).
30. J. H. Mackey and J. Nahum, "Spectral Study of the Interconversion of Eu^{2+} and Eu^{3+} in Silicate Glass", *Phys. Chem. Glasses* 9, 52 (1966).
31. W. D. Johnston, "Oxidation-Reduction Equilibria in Iron-Containing Glass", *J. Am. Ceram. Soc.* 47, 198 (1964).
32. T. Baak and E. J. Hornyak, Jr., "Iron-Oxygen Equilibrium in Glass: Effect of Platinum on $\text{Fe}^{2+}/\text{Fe}^{3+}$ Equilibrium", *J. Am. Ceram. Soc.* 44, 541 (1961).

33. E. T. Turkdogan and P. M. Bills, "Thermodynamic Study of FeO-Fe₂O₃-SiO₂, FeO-Fe₂O₃-P₂O₅, and FeO-Fe₂O₃-SiO₂-P₂O₅ Molten Systems", J. Iron Steel Inst. (London) 186, 329 (1957).
34. A. Paul and R. W. Douglas, "Ferrous-Ferric Equilibrium in Binary Alkali Silicate Glasses", Phys. Chem. Glasses 6, 207 (1965).
35. G. J. Kakabadse and E. Vassiliou, "The Isolation of Vanadium Oxides in Glasses", Phys. Chem. Glasses 6, 33 (1965).
36. W. D. Johnston, "Optical Spectra of the Various Valence States of Vanadium in Na₂O·2SiO₂ Glass", J. Am. Ceram. Soc. 48, 608 (1965).

THERMODYNAMIC PROPERTIES OF LIQUID METALS

J. L. Margrave

Abstract

As a continuation of the project started earlier, an extended effort has been made to collect the best available data on the thermodynamic properties of liquid metals, including heats and entropies of fusion, heat capacities and surface tensions. Attention has been concentrated on the heats and entropies of fusion and the heat capacities in a search for parameters which would allow correlation and make possible reasonably reliable predictions for the many high melting elements yet unstudied -- Cr*, Hf, Ir, Mn*, Os, Re, Rh, Ta*, W, and Zr -- as well as the hundreds of important refractory compounds and alloys. A mutually consistent table of such liquid metal properties has been assembled and various correlations are presented graphically.

Special attention has been given to correlations of the entropies of fusion with structure parameters and the literature has been surveyed for information regarding theo-

*Currently under study.

retical approaches for explaining the rise of the specific heats of both elemental solids and liquids far beyond the $3R$ predicted by either the Debye or the Born-von Karman theories. The predominant literature explanation is "anharmonic effects" but new experimental data* show strong correlations with electronic configurations and suggest that new calculations which take into account the actual densities of states at high temperatures might yield more realistic electronic contributions to the specific heat than those usually predicted by the equation

$$C_{el} = \left(\frac{Ti^2 Nk}{2T_F} \right) T = \gamma T$$

when applied in the evaluation of γ from low-temperature heat capacity data.

*Various publications making use of levitation calorimetry in High Temperature Science and in Journal of Chemical Thermodynamics.

THERMODYNAMIC PROPERTIES OF LIQUID METALS

J. L. Margrave

An exploration of the chemical and physical literature reporting theoretical and experimental studies of the high temperature properties of liquid metals has been conducted. A number of basic references were located and these are cited at the end of this report.¹⁻³

It is also worthwhile to recapitulate the conclusions from a paper⁴ which represented a summary through 1970 of the available literature data for the heats and entropies of fusion and heat capacities of the various liquid metals. Basically, the report concluded that one could find relatively little experimental data on metals melting above 2000° except for the very recent work emanating from Rice University by the technique of levitation calorimetry. The data were relatively few but did indicate the possibility of some periodic correlations for estimating "best" values when no experimental data were available. During the interim, more information from levitation studies has become available⁵ and also a number of correlations have been tested for the various parameters, including correlations between heats of fusion and atomic numbers as contrasted with melting points and atomic numbers between entropies of fusion and atomic

numbers, and between heat capacities of ideal gases or heat capacities of liquid metals and the Gibbs free energy function for ideal gaseous atoms with a standard reference temperature of 2000°K being used for the gaseous species. (Figures 1-8). These correlations and conclusions therefrom, form the basis for a new table of thermodynamic properties which is included in this report, Table 1.

In the course of investigation it has become clear that there are several fundamental types of questions about heat capacity variations and high temperatures which are not answerable by means of currently published theory. The questions can be best understood by referring to Fig. 9, a diagram for elemental Li, taken from a review article by Borelius¹ who has been actively calling attention to this problem for the past twenty years. As one advances up the C_p vs T curve for a typical solid he expects to find increasing C_p as is consistent with either the Einstein, Debye or the more refined Born-von Karman theories and eventually at temperatures which are high compared with the Debye temperature one predicts C_p should approach the classical limit of $3R$. This seems not to be the case for practically all of the metals. In looking at the tabulated data, about the only elements for which C_p stays anywhere near $3R$ at the melting points are those whose solid lattice is diamond structure -- C, Ge, and Si. These elements show C_p values for the liquid in the range of 6-6.6⁴ and it is conceivable that current theory would predict deviation from

the classical 3R at high temperatures of the order of 10-20 percent so that these values are at least consistent.

It is important to recognize that the sources of the 10-20 percent deviations allowed by current theory and these are the contributions of anharmonic effects due to the unusually large lattice vibrations expected at high temperatures. These deviations, which in magnitude are about 10 percent, should lead to an approximately linear variation in C_p as a function of temperature and one would anticipate a still smaller rate of change for the derivative of C_p as a function of temperature due to the higher-order anharmonicity terms.² The other contribution to the specific heat of all temperatures is the electronic contribution given normally by the formula $C_p = \gamma T$ where γ is composed of a group of constants including the Fermi-temperature in the denominator and typical values for γ run in the range of 10^{-4} calories degrees⁻². Again, this term leads to a linear variation in the heat capacity with a total change in the absolute value of C_p of approximately a few tenths per thousand degree temperature change. Neither of these factors as currently interpreted gives any kind of explanation for the tremendous deviations of C_p which are observed in real metals even at relatively low temperatures like 300-500°C. For example, the alkali metals which melt in the range 300-500°C, all show C_p values of greater than 7. The situation is even worse for the high melting metals and C_p values for solids and liquid metals ranging as high as 12

calories degrees⁻¹ mole⁻¹ are currently known from reliable experimental studies.

Borelius¹, who recognized and has discussed in great detail this general problem, tries to explain the uptake of heat at rates greater than predicted by the classical 3R heat capacity on the basis of pre-melting and post-melting phenomena. Thus, in Fig. 1, he identifies the areas C_{sm} and C_{fm} as indicative of these premelting and postmelting processes and he accounts for the high values of C_p by suggesting that the structure of the crystal is undergoing some preliminary breaking up before the experimental melting point is reached and that some continued structural changes occur after melting. The additional electronic contribution is C_e and the harmonic lattice contribution is C_a .

There appear to be several distinguishable problem areas:

- (1) The rise of C_p (solid) above 3R in the range below T_M . A plot of $(C_p(\text{real}) - 3R)$ as T for several real metals should be informative about the type of excess-function needed to explain this.
- (2) The fairly good correlation of ΔS_f with structure, except that the qualitative trend is opposite that which one would predict. Thus,

$$\Delta S_{\text{fus}}(\text{bcc}) < \Delta S_{\text{fus}}(\text{hcp}) < \Delta S_{\text{fus}}(\text{ccp})$$

while the change in coordination number is greatest for (bcc \rightarrow l).

(3) The large and essentially constant value for $C_p(l)$ for many very high melting liquid metals over ranges of 500-600°K to a precision of ±0.5%. Thus, if one considers the maximum deviations, the range for C_p is much less than one might have predicted if either anharmonic or electron contributions were contributing the usual linear terms to the high-temperature heat capacity.

(4) The periodic variation of C_p 's for transition metals shows a double-hump and correlates best with functions dependent on energy level configurations and the electronic partition function like C_p (ideal gas) at 2000°K or

$$\left(\frac{G_T^0 - H_{2,98}^0}{T} \right)_{2000^\circ\text{K}}$$

for the ideal gas atoms.

(5) The general failure of mass spectrometer methods to detect large concentrations of polymers over liquid metals⁶. Typical studies suggest 1-2% or less dimer in equilibrium vapors, yet polymeric species would have to be predominant in liquids to give the huge premelting and postmelting effects observed. Also, polymers in the liquid would be subject to dissociation with increasing temperature in a homogeneous equilibrium and one would not expect C_p to be constant for a liquid of rapidly changing complexity. C_p should decrease almost exponentially with T , at least for an appreciable range, after the melting point if the molecular complexity is really decreasing in the liquid.

In summary, more experimental data and more detailed theoretical work will be required before a complete picture of liquid metal thermodynamic properties can be developed. However, meaningful correlations can be made between heats of fusion and melting point and between C_p (liquid metals) and the Gibbs free-energy factor, $\frac{G_T - H_{298}}{T}$. Predictions are possible but crucial measurement needs are apparent. Work is especially needed on elements from the groups [Zr, Rh and Ru]; (Hf, Ir, Os, Ta, Re and W); and on Sm, Ho and Lu to make extrapolations and/or interpolations more reliable. When pure metals are understood, one can then consider these same properties for alloys and for refractory compounds like carbides, borides, nitrides and silicides.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAH15-71-C-0253 with The University of Michigan.

References

1. G. Borelius, *Solid State Phys.* 6, 65- (1958).
ibid. 15, 1-51 (1963).
2. G. Leibfried and W. Ludwig, *Solid State Phys.* 12, 275-444 (1961).
3. A. R. Ubbelohde, "Melting and Crystal Structure," (Clarendon Press, Oxford), 1965.
4. J. L. Margrave, *High Temperatures-High Pressures*, in press, 1971.
5. (a) J. L. Margrave and J. A. Treverton, *Proc. Symp. on Thermophys. Prop.*, Boston, October 1970. (Data on Mo).
(b) J. L. Margrave and J. A. Treverton, *J. Chem. Thermo.*, in press, (1971). (Data on Fe, Ti and V).
(c) J. L. Margrave and J. A. Treverton, *J. Phys. Chem.*, in press, (1971). (Data on Co and Pd).
(d) A. J. Valerga and J. L. Margrave, *Abstract, American Chem. Sci. Meeting, Washington, D.C.*, (1971). (Data on Nb and Zr).
6. J. L. Margrave, Ed., "Chemistry of HT Vapors", (Wiley, New York), 1967.

TABLE 1.

Thermodynamic Properties of Liquid Metals

Element	MP K	ΔH_{fus} Kcal mole ⁻¹	ΔS_{fus} Cal deg ⁻¹ mole ⁻¹	C_p (l) at MP Cal deg ⁻¹ mole ⁻¹
Ag	1234	2.65±0.1	2.15	8.00 (H)
Al	933	2.58±0.03	2.76	7.59 (H)
Au	1336	3.05±0.1	2.29	7.0 (H)
B	2303	(4.4-6.0)	(1.9-2.6)	(7.8) (H)
Ba	983	1.83±0.1	1.86	(7.5±0.5)
Be	1557	3.5±0.2	2.25	7.04 (H)
Bi	544	2.6±0.05	4.78	7.30 (H)
C (gr)	(4073)	(25±2)	(6±1)	(6.2) (H)
C (dia)	(4073)	(28±2)	(7±1)	
Ca	1116	2.0±0.1	1.79	7.00 (H)
Cd	594	1.53±0.04	2.58	7.10 (H)
Ce	1071	1.305±0.05	1.24	9.01 (H)
Co	1765	3.76±0.04	2.13	11.25 (LC)
Cr	2130	(4.0)	(1.9)	(9.4±0.5)
Cs	303	0.50±0.01	1.65	7.62 (H)
Cu	1356	3.110±0.075	2.29	7.9 (LC)
Dy	1682	2.643±0.2	1.57	11.93 (H)
Er	1795	4.757	2.65	9.25 (H)
Eu	1090	2.20±0.05	2.02	9.11 (H)
Fe	1809	3.24±0.06	1.79	10.35 (LC)
Ga	303	1.336±0.01	4.41	6.65 (H)
Gd	1585	2.403±0.1	1.52	8.88 (H)
Ge	1210	8.8±0.3	7.27	6.60 (H)
Hf	2493	(5.7-6.8)	(2.3-2.7)	(9.0±0.5)
Hg	234	0.55±0.005	2.35	6.806 (H)
Ho	1743	(2.7)	(1.5)	(10.5) (H)
In	429	0.78±0.02	1.82	7.05 (H)
Ir	2727	(6.2-7.6)	(2.3-2.8)	(9.0±0.5)
K	337	0.571±0.005	1.69	7.68 (H)
La	1193	1.481±0.05	1.24	8.20 (H)
Li	453	0.70±0.03	1.54	7.26 (H)
Lu	1936	(3.0-4.8)	(1.5-2.5)	(11±1)
Mg	923	2.1±0.1	2.28	7.8 (H)
Mn	1517	(3.0)	(2.0)	(9.0±0.5)
Mo	2839	8.32±0.07	2.88	8.2 (LC)
Na	371	0.63±0.02	1.70	7.61 (H)
Nb	2740	7.9±0.2	2.88	9.4 (LC)
Nd	1289	1.707±0.1	1.32	11.66 (H)
Ni	1726	4.1±0.08	2.38	10.30 (H)
Os	3045	(6.9-8.6)	(2.3-2.8)	(9.0±0.5)

TABLE 1 (continued)

Element	MP K	ΔH_{fus} Kcal mole ⁻¹	ΔS_{fus} Cal deg ⁻¹ mole ⁻¹	$C_p(l)$ at MP Cal deg ⁻¹ mole ⁻¹
Pb	600	1.15±0.03	1.92	7.32 (H)
Pd	1825	4.01±0.05	2.19	9.1 (LC)
Po	519	(1.2-1.6)	(2.3-3.0)	(9±1)
Pr	1204	1.646±0.100	1.36	10.27 (H)
Pt	2042	5.30±0.15	2.61	8.7 (LC)
Ra	973	(1.9)	(1.9)	(7.0±0.5)
Rb	312	0.525±0.01	1.68	(7.5) (H)
Re	3453	(8.0-10.1)	(2.6-2.9)	(8.0±0.5)
Rh	2233	(5.1-5.7)	(2.3-2.6)	(9.5±0.5)
Ru	2583	(6.1-7.2)	(2.3-2.8)	(9.5±0.5)
Sc	1812	3.369	1.86	(10.57) (H)
Se	493	1.5±0.3	3.04	(7.0)
Si	1685	12.1±0.4	7.19	6.10 (H)
Sm	1345	2.13±0.1	1.58	(12.0) (H)
Sn	505	1.69±0.03	3.55	7.10 (H)
Sr	1043	(1.9)	(1.3)	(7.0±0.5)
Ta	3250	(7.5-9.3)	(2.3-2.9)	(9.0±0.5)
Tb	1630	2.58±0.10	1.59	11.11 (H)
Tc	2413	(5.6-6.5)	(2.3-2.7)	(8.5±0.5)
Te	723	4.18±0.3	5.78	9.00 (H)
Th	2028	(4.0-5.0)	(2.0-2.5)	(11±1)
Ti	1940	3.78±0.06	1.95	11.9 (LC)
Tl	576	1.03±0.03	1.79	7.2 (H)
Tm	1818	4.025±0.05	2.22	9.89 (H)
U	1405	2.036	1.45	11.45 (H)
V	2193	4.05±0.15	1.85	12.0 (LC)
W	3680	(8.4-10.8)	(2.3-2.9)	(8.0±0.5)
Y	1793	2.75±0.1	1.53	10.30 (H)
Yb	1097	1.830±0.008	1.67	8.79 (RSD)
Zn	692	1.74±0.03	2.49	7.5 (H)
Zr	2125	(4.8-5.2)	(2.3-2.5)	(10.2±0.2)

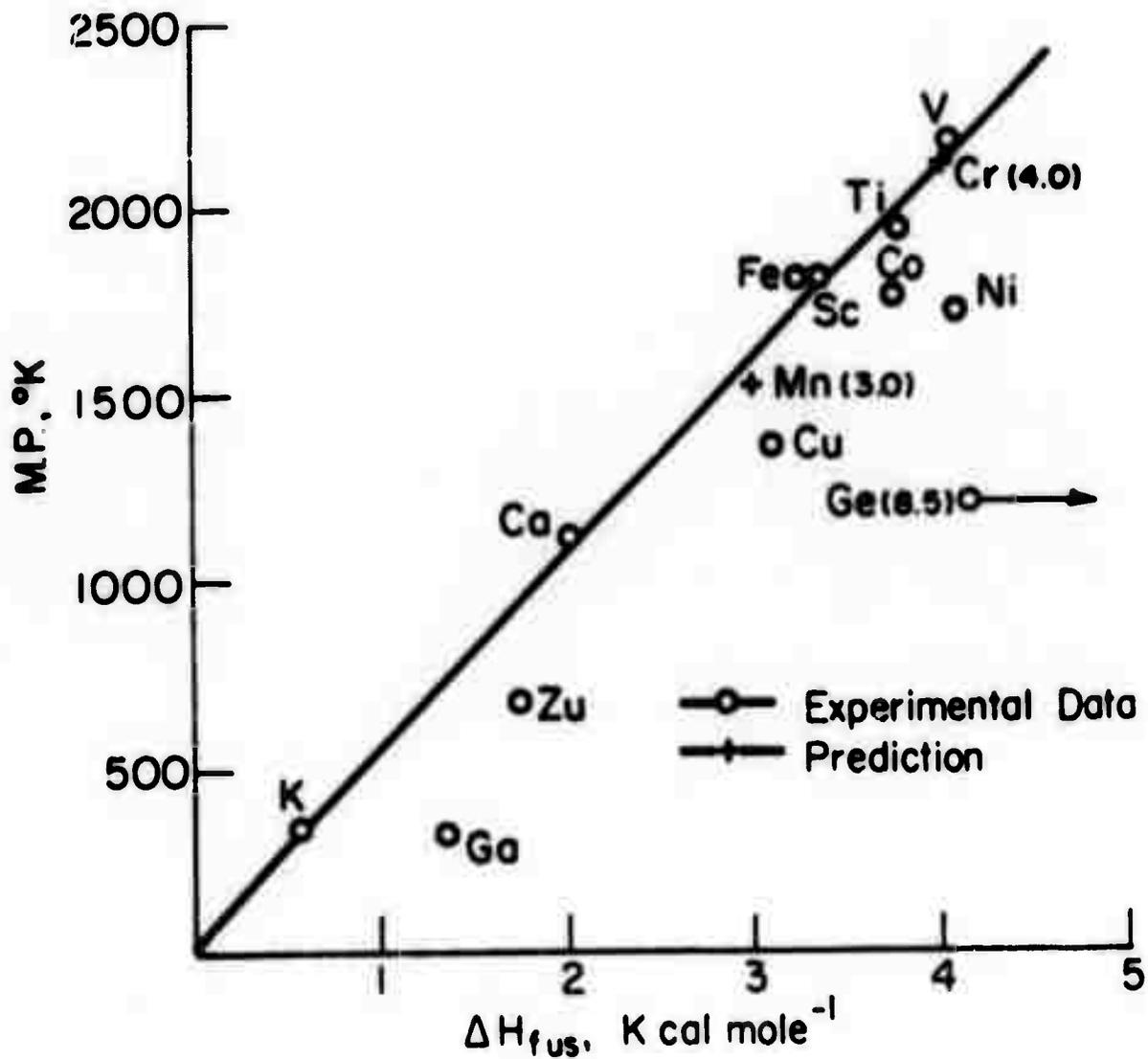


Figure 1. Correlation of Heats of Fusion and Melting Points for First-Row Transition Metals.

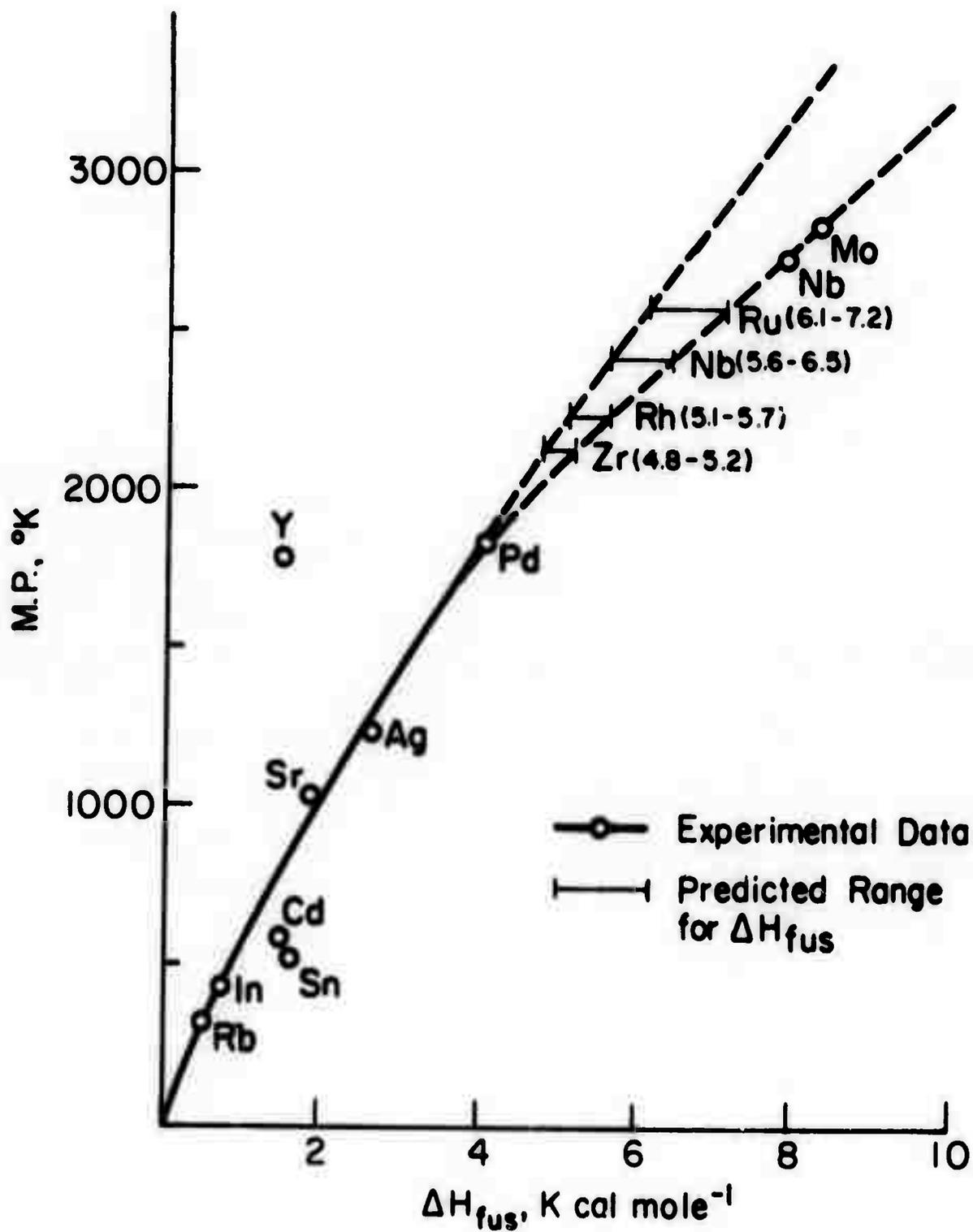


Figure 2. Correlation of Heats of Fusion and Melting Points for Second-Row Transition Metals.

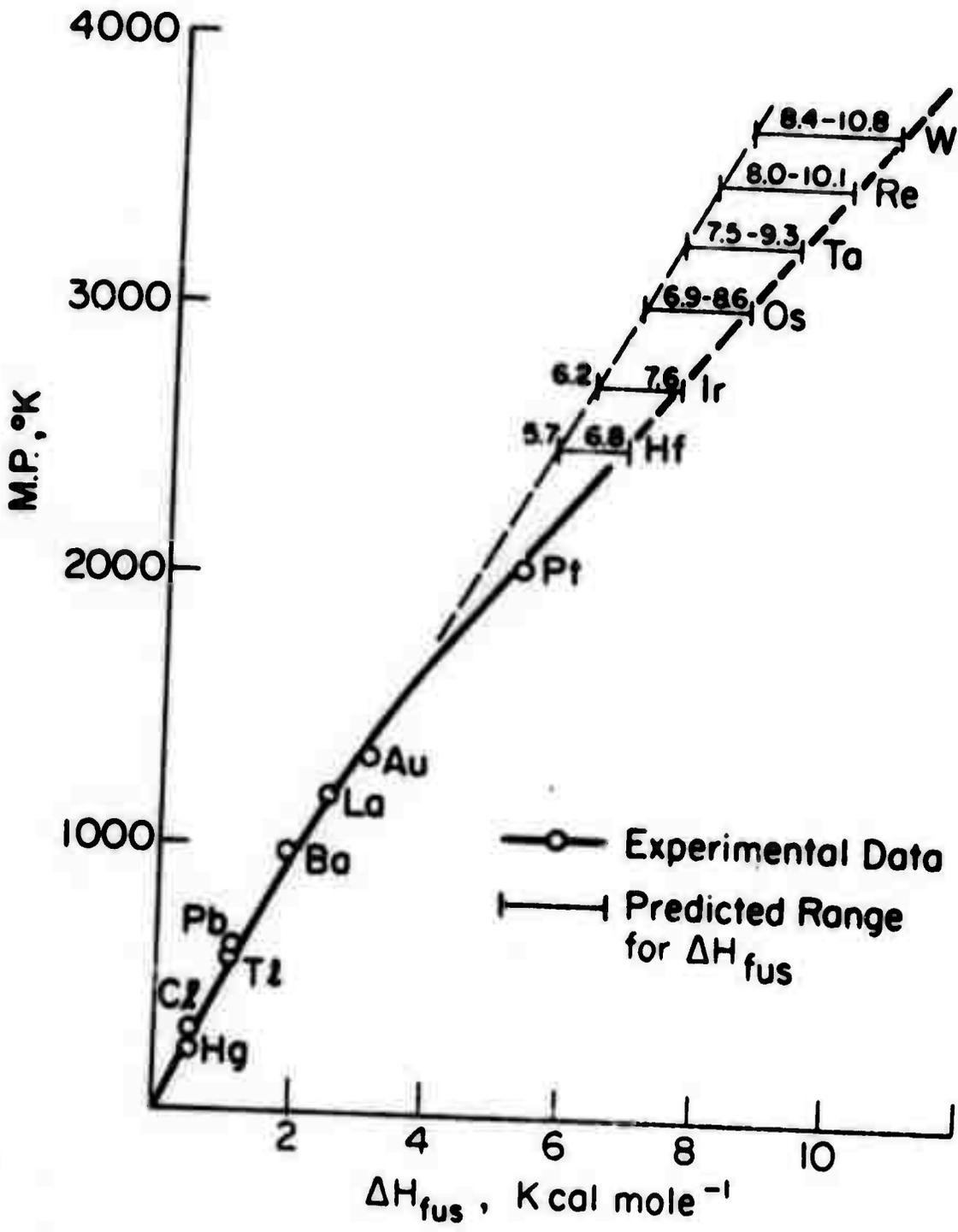


Figure 3. Correlation of Heats of Fusion and Melting Points for Third-Row Transition Metals.

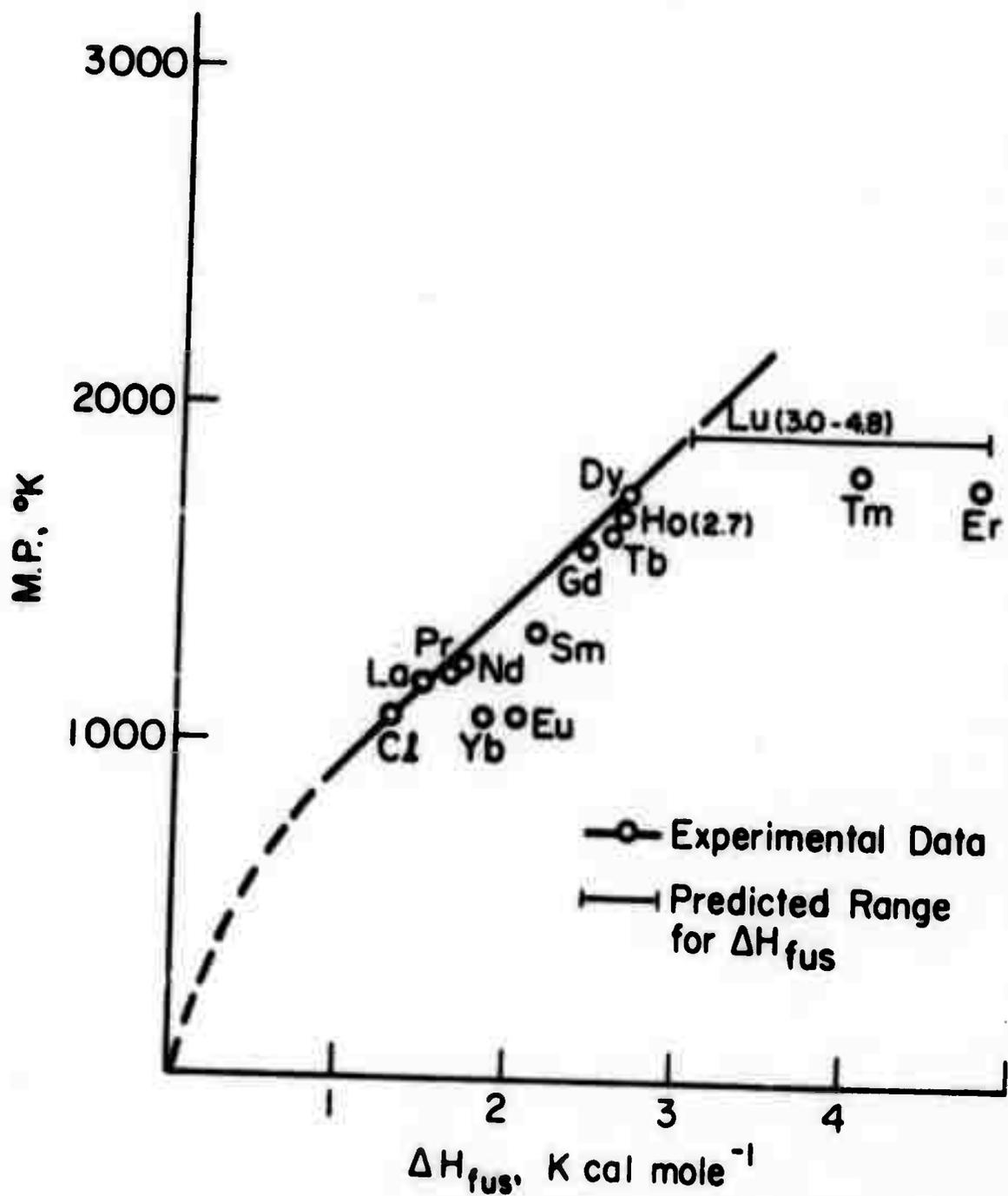


Figure 4. Correlation of Heats of Fusion and Melting Points for Rare-Earth Metals.

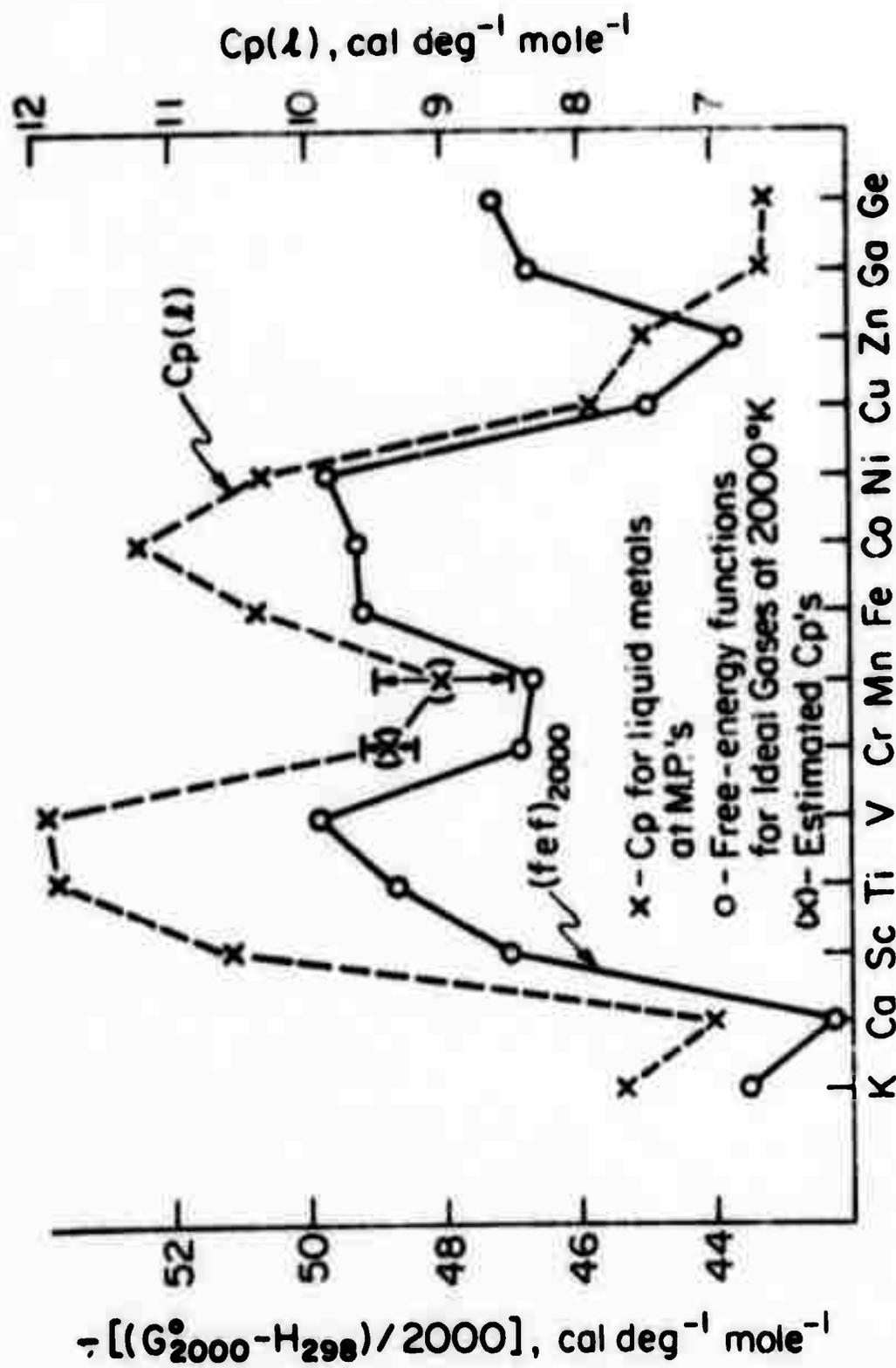


Figure 5. C_p Correlation for First-Row Transition Metals.

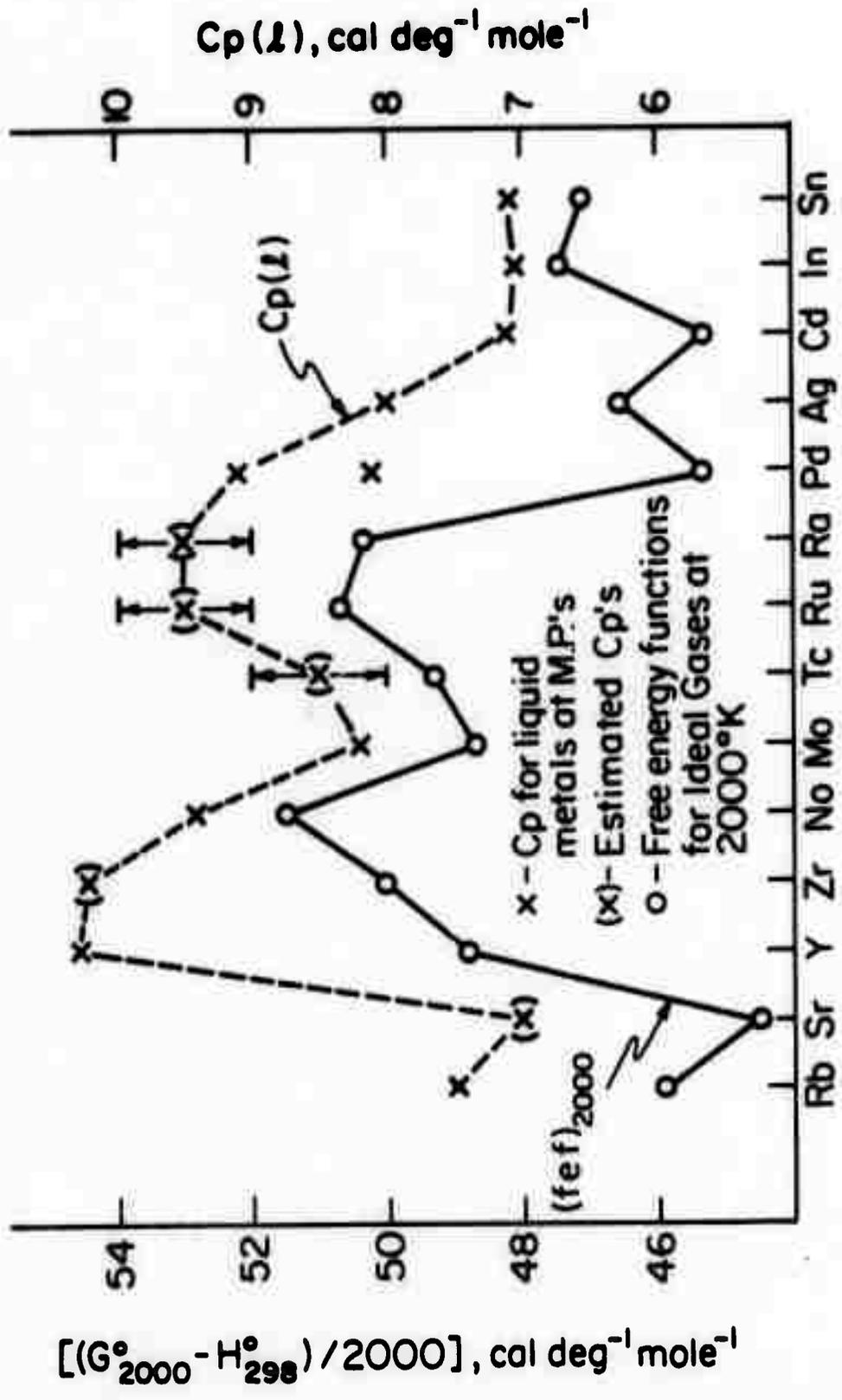


Figure 6. C_p Correlation for Second-Row Transition Metals.

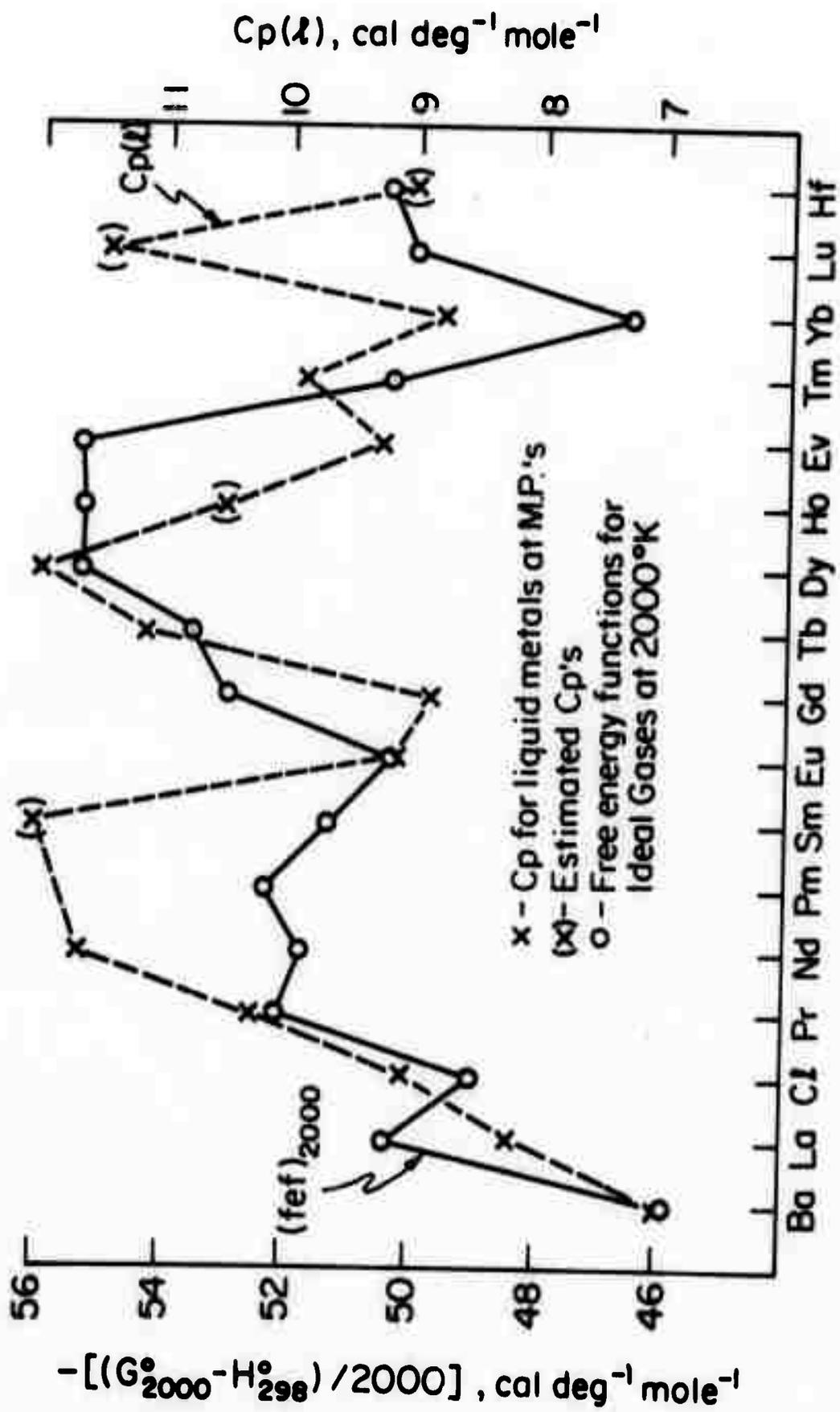


Figure 7. C_p Correlation for Rare Earth Elements.

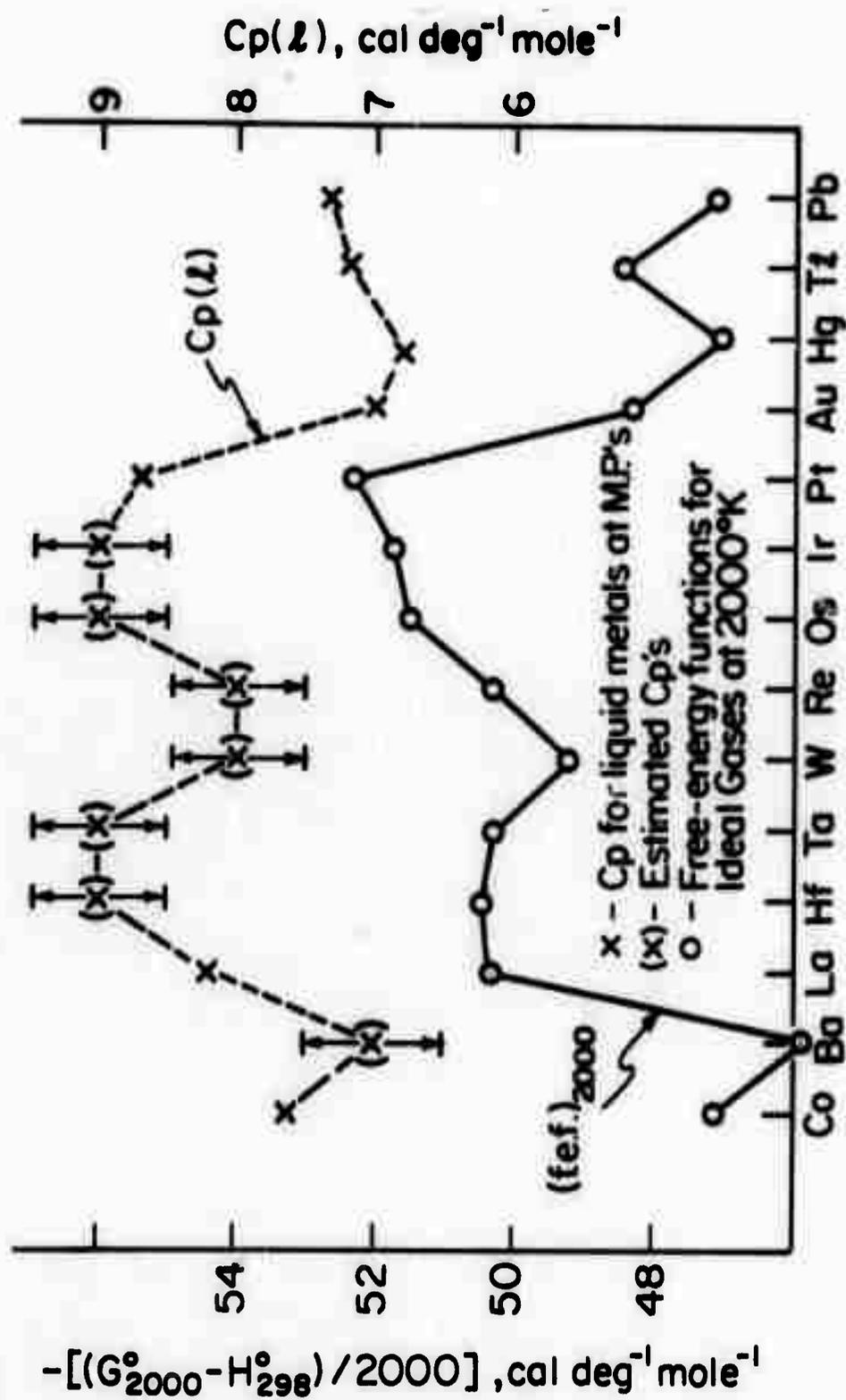


Figure 8. C_p Correlation of Third-Row Transition Metals.

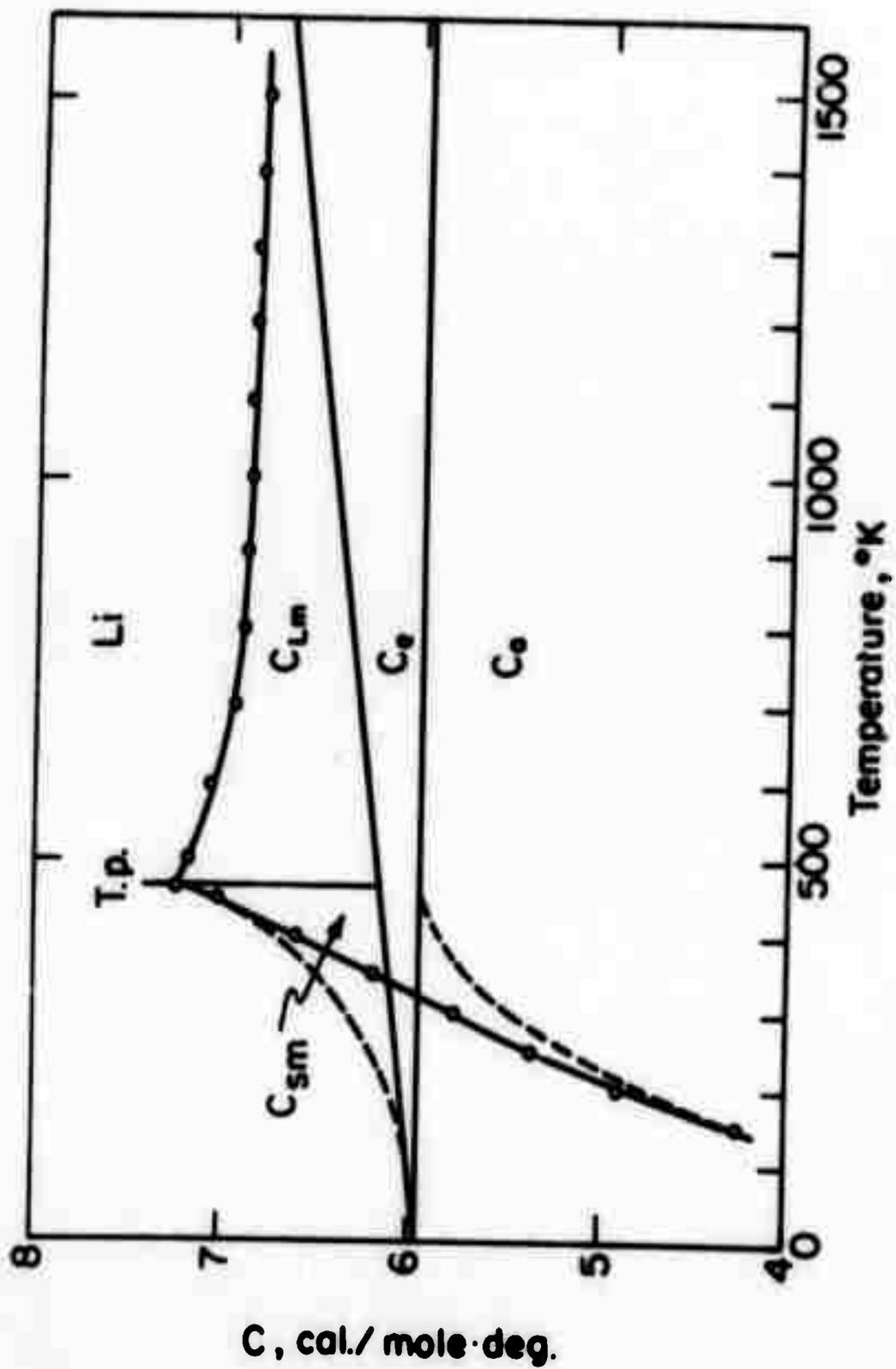


Figure 9. Specific Heat of Lithium.

STRESS CORROSION CRACKING IN PLASTIC SOLIDS
INCLUDING THE ROLE OF HYDROGEN

J. J. Gilman

Abstract

Small changes in surface environments can change the energy needed to create a surface shear step. Increases in this energy tend to shift a delicate balance between glide and cleavage initiation at a crack tip. By inhibiting plastic deformation this causes an increased tendency for cleavage. Thus a material that is ductile in a vacuum can become quite brittle in the presence of certain surface active environments; particularly atomic hydrogen.

The quantitative criterion for deciding whether flow or cleavage will prevail is the ratio of the appropriate glide plane's surface energy to the cleavage plane's surface energy.

A survey of the strengths of the interactions between hydrogen and metals has been made. Throughout the periodic table strong diatomic interactions occur. At surfaces the interactions remain strong although they are somewhat weaker. In solid hydrides they tend to be much weaker or nonexistent. Thus the strength of the interaction

depends on the metal-metal distance. It is shown that for a typical transition metal the interaction with hydrogen is strong enough to readily cause embrittlement.

STRESS CORROSION CRACKING IN PLASTIC SOLIDS
INCLUDING THE ROLE OF HYDROGEN

J. J. Gilman

I. Introduction

The phenomenon of stress-corrosion cracking depends sensitively on various parameters of a system, such as the magnitude of the stress, the temperature, applied potentials, composition of the environment, etc. Therefore, it is a "critical" phenomenon. That is, it is associated with a balance between counter-vailing forces and small changes in the state of a system can shift it dramatically away from the balanced situation (Latanision and Westwood, 1970).

Another characteristic is that the counter-vailing forces are very localized. This may be concluded because stress-corrosion cracking sometimes occurs with great speed (of order 10^3 cm/sec), and because the fractured surfaces are often very smooth on a local scale (at magnifications of 20,000 and more). Both of these facts indicate that the volume of materials that is involved in the critical process is very small.

Stress-corrosion cracking is observed in purely elastic solids (completely brittle); in elastic-plastic solids (semi-brittle); and in viscoelastic solids. In the first class, which is typified by glass, the effect consists

essentially of stress-induced chemical decomposition. This has recently been discussed by Tong and Gilman (1971) and will not be discussed further here. The second case, which is typified by metallic alloys, is the one of primary present interest.

In discussing stress-corrosion cracking, it is important to keep in mind that cracking cannot occur unless two conditions are simultaneously satisfied. One is the static energy balance or Griffith condition. The other is the kinetic condition that the local driving force (normal stress) must be great enough to cause the solid to "break" at an observable rate. Satisfaction of either of these conditions alone does not necessarily lead to fracture. In the discussion that follows it will be assumed that the Griffith condition is always satisfied so that the behavior of the system depends on whether the local stress does or does not exceed the cohesive stress of the solid in the appropriate direction.

For a purely elastic solid, when the Griffith condition is satisfied the local normal stress at the crack tip approximately equals the cohesive stress of the solid. (Keer, 1964) However, if the solid is ideally plastic this stress instantly relaxes to approximately twice the yield shear stress. Or, if the solid is ideally viscoelastic, the local stress begins to relax toward zero at a rate that depends on the viscosity coefficient.

In a real solid, plastic or viscous relaxation occurs via the motion of dislocation lines. The resulting plastic strain-rate that tends to relax the elastic stress depends on the instantaneous values of the local stress and the plastic strain (amount of plastic relaxation that has already occurred). The result of this effect is that the local stress at a crack tip will not exceed the cohesive stress unless:

- a) there are no dislocations present locally and they cannot be nucleated rapidly,
- b) there are dislocations present but the local shear stress cannot move them, or they lie on planes that do not have the line of the crack tip as their zone axis,
- c) dislocations are present and moving but the crack tip is moving at a velocity that does not allow enough time for adequate stress relaxation at the crack tip (Gilman, 1966).

Since stress-corrosion does not necessarily require high crack velocities and this discussion concerns plastic solids, cases (b) and (c) above are ruled out. Case (a) remains for further consideration. The key parts of statement (a) are the need for dislocations locally at the crack tip, and the question of whether the high local stresses associated with a sharp crack can or cannot nucleate dislocations at the tip which move outward from it and thereby absorb energy and blunt the shape.

II. Importance of the Crack Tip

For reasons given in the second paragraph of the Introduction, occurrences at and very near a crack tip are especially important for stress-corrosion. It is difficult to separate the phenomena there because the local stresses which cause plastic flow are in equilibrium with the local surface forces. Therefore an observed effect that is induced by a change in the environment might result either from a change in the surface forces or from a change in the plastic flow rate at constant stress. Perhaps this can be understood better by considering an approximate expression for the fracture surface energy (Gilman, 1959);

$$\gamma_f = \left(\frac{\tau_{\text{cohesive}}}{\tau_{\text{yield}}} \right) \gamma_e$$

Here, τ_{coh} is the cohesive shear stress required to permanently shear the perfect structure, τ_y is the shear stress needed for rapid plastic flow; and γ_e is the "elastic" surface energy. In a vacuum γ_e equals the intrinsic surface energy of the solid but it can be modified by surface adsorption in a gaseous or liquid environment. The yield stress, τ_y , might also be influenced by the environment. Since $(1/\tau_y)$ multiplies, γ_e , the effect of an environment on γ_f cannot be uniquely related to either individual factor without an independent microscopic measurement. If γ_e becomes small, it does not matter if τ_{coh}/τ_y is large because γ_f will be small and the

material brittle. On the other hand, if $\tau_y = \tau_{coh}$, then $\gamma_f = \gamma_e$ and the material becomes elastic and, hence, brittle.

III. Glide vs. Cleavage at a Crack Tip

An environment cannot suppress flow in a sizeable region around a crack tip because the interaction energy at a surface is not large enough to change the energy density within a sizeable volume of material. Therefore, special modes of deformation must be the ones that are suppressed. The obvious mode for attention is the flow that emerges from a crack tip. This mode has been observed experimentally (Gilman, 1957), and studied extensively by Burns and Webb (1970). It is the most important mode because it has the most effect on the sharpness of a crack tip; and, therefore, on its local stress distribution. This is clarified by means of Figure 1, which shows the configuration near the tip of a critical crack in a simple cubic structure (Goodier, 1968).

A critical crack is one for which the Griffith propagation condition is just satisfied. The tip configuration is then nearly independent of the crack length because the maximum stress at the tip must always equal the cohesive stress (Keer, 1964). Thus, approximately the same set of strains is always present and, hence, the same shape. Plastic shears are effective in blunting such a crack only if they pass within one or two atom distances of the tip, as shown in Figure 2. In this figure, a shear at the very tip (A) has changed the

tip shape markedly, but a shear somewhat away from the tip (B) has changed the shape relatively little. Thus small changes in plasticity very near a crack tip can markedly influence the subsequent behavior, but general plastic flow in the vicinity may have little effect.

The critical competition that takes place during cracking in general and during stress-corrosion cracking in particular is between the tendency for a crack to become longer by cleavage and its tendency to become blunted by plastic shears. This problem was first considered explicitly by Kitajima (1965) and later by Kelly, Tyson and Cottrell (1967). Recently, Rice and Thomson (1971) have considered it in some detail. It will be considered here from a different viewpoint that allows the essential result to be expressed in terms of relative surface energies. This form of the result is particularly compact and can be readily applied to stress-corrosion problems.

Because of the small scales of the crucial events at a crack tip, a detailed analysis of the physical processes that occur is very complex, but simple models serve to define the relationships between the principal physical features. Figure 3 shows schematically how dislocations can be used for this purpose.

First, a crack can be modeled by means of an array of edge dislocations (Friedel, 1959) because the summation stress field of an appropriate array is closely similar to

that of a crack. In Figure 3A a crack that is modeled in terms of dislocations is shown. Figures 3B and 3C focus attention on the region near the very tip of a crack.

Now the competitive processes can be clearly defined. First, the elementary cleavage process consists of an advance of the leading crack-type dislocation, (Figure 3B) by an atomic distance in the x-direction. Second, the start of plastic blunting consists of the creation of two glide dislocations, (1) and (2) which move out to a distance, x , from the crack tip.

The first process requires an energy per unit length of approximately $2\gamma_c b$; where γ_c is the specific surface energy of the cleavage plane, b is an interatomic distance, and two is the number of new surfaces.

An estimate of the energy of the other process is somewhat more complex. However, it is greatly simplified by the fact that the stress field of one of the two dislocations does not exert a force on the other dislocation if the medium is isotropic and the dislocations lie on planes inclined at 45° with respect to the cleavage plane. Also, their interaction energy is constant (independent of x), and the sum of their self-energies equals the energy of the leading crack-type dislocation. Furthermore, their stress fields approximately cancel the stresses on the crack surfaces so the image forces that tend to pull them back into the crack are small.

When the dislocations are created two small surface

steps of specific energy, γ_g , appear at the tip of the crack as indicated in Figure 3C. These exert a tension, $(+2\gamma_g)$ to inhibit the creation of the dislocations for small x . A driving force, $(-2\tau b)$ tends to push the dislocations into the medium, where the stress, τ decreases with x as follows:

$$\tau(x) = \tau_0 \left(\frac{b}{x} \right)^{1/2} \quad (1)$$

Since the interaction force is zero, the total force equilibrium is:

$$2\gamma_g - 2\tau_0 b^{3/2} x^{-1/2} = 0 \quad (2)$$

and solving for x yields:

$$x = b^3 \left(\frac{\tau_0}{\gamma_g} \right)^{1/2} \quad (3)$$

This means that the dislocations tend to move further into the material as the applied stress increases or the surface energy of the glide plane decreases.

The conclusion at this point is that (for a critical crack) both cleavage and glide may be spontaneous processes depending on the conditions very near a crack tip.

For significant spontaneous glide to occur, x must be equal to or greater than b so that:

$$\tau_0 b > \gamma_g \quad (\text{for glide}) \quad (4)$$

For cleavage to occur the local tensile stress must be equal to or greater than the cohesive stress, σ_{coh} which is approximately:

$$\sigma_{coh} = \frac{2\gamma_c}{b} \quad (5)$$

Then since $\sigma_0 \approx 2\tau_0$ the condition for the cleavage mode to predominate is:

$$\gamma_c \approx \gamma_g \quad (6)$$

This condition is approximate so the main conclusion is that the mode of propagation of a crack is very sensitive to the ratio: (γ_g/γ_c) and anything that tends to raise γ_g relative to γ_c will favor cleavage. It is consistent with the behavior of layered crystals for which the ratio of γ_g to γ_c tends to be large.

IV. The Role of Surface Layers

The application of this result to stress corrosion requires a recognition of the fact that very thin layers of material can substantially change the ratio of the energy of formation of a glide surface steps to the energy of a cleavage surface, and thereby enhance cleavage. This is illustrated in Figure 4A which shows how the formation of a thin surface compound (chemisorption) can increase the energy needed to make a surface step; it can also markedly increase the activation energy.

The presence of the surface compound is indicated schematically by the cross-hatched circles. Initially (the dotted lines) each equivalent surface position is occupied by a second atomic species. After the shear, the positions

are indicated by the solid lines and it may be seen that the configuration around the cross-hatched atom at the center of the tip has changed considerably. Thus, the surface compound has essentially been decomposed which requires an energy approximately equal to its free energy of formation, ΔF . In order for a dislocation to form at this site spontaneously then, the work done during shearing must approximately equal ΔF per molecule. That is:

$$\tau b^3 = \Delta F \quad (7)$$

and since $\tau = G/15$ at the tip of a critical crack, if ΔF is greater than about $Gb^3/15$ spontaneous blunting cannot occur. Taking typical values of 10^{12} d/cm² and 2.5×10^{-6} cm for G and b , respectively, this means that spontaneous blunting is not expected if ΔF exceeds about 0.6 e.v. or 15 Kcal/mol.

The activation energy for this event may be equal to or even greater than the total energy change because the surface symmetry is destroyed when the glide displacement is $b/2$ so half the final shear has occurred and atoms have been moved to "wrong" positions.

In addition to suppressing plastic flow of the critical blunting type, chemisorption lowers the effective cleavage surface energy and thereby tends to enhance cleavage. The change in surface energy is approximately $\Delta F/b^2$ and if this equals or exceeds γ_c , then spontaneous cleavage is expected provided the chemisorption species can continue to

reach the crack tip.

Another way to discuss this is to consider the effects on Equation (6). The chemisorption energy per unit surface is $\sim \Delta F/b^2$ which enhances cleavage so it subtracts from γ_c ; and adds to γ_g since it tends to suppress glide. Thus the cleavage condition becomes:

$$\left(\gamma_c - \frac{\Delta F}{b^2} \right) \lesssim \left(\gamma_g + \frac{\Delta F}{b^2} \right) \quad (8)$$

or cleavage is expected if:

$$\Delta F > (\gamma_c - \gamma_g)b^2/2 \quad (9)$$

Suppose that $(\gamma_c - \gamma_g) = 10^3$ ergs/cm² and again $b = 2.5 \times 10^{-8}$ cm. Then cleavage will occur for $\Delta F \approx 0.2$ e.v. or ~ 5 Kcal/mol. Even the physical adsorption of a highly polarizable species such as water can cause this much energy change.

Figure 4B illustrates schematically how local shearing affects the configuration when a localized surface dipole layer is present; or a species that adsorbs external to the substrate surface. When shearing occurs, the local configuration must again change substantially. In this case whether or not its energy of formation must be provided depends on how strongly the adsorbed atoms interact laterally. If this interaction is weak there may be little effect.

V. Example of Hydrogen Embrittlement

The bond energies of the gaseous hydrides are shown for a majority of the elements in Figure 5. These numbers indicate that atomic hydrogen interacts strongly with most elements (and, hence, with most alloys). In fact, the interaction energies are typically large compared with the surface energies (per atom) of the elements. Therefore, it can be expected that if hydrogen can reach a crack tip in nearly any element (or alloy) it will tend to cause embrittlement.

The interaction with hydrogen is strongest for elements with low atomic numbers and large electronegativities but it is substantial for most of the elements except the rare gases. Since these bond energies are comparable with the heats of chemisorption, these are also high for most elements. The widespread existence of hydrogen embrittlement of metals is therefore not surprising but the detailed mechanisms continue to need elucidation.

One of the first problems to be considered is the hydrogen-surface interaction because the interaction of a metal with hydrogen depends strongly on the state of the metal. The data given in Figure 6 show this quite clearly for a few illustrative metals. Metals bond most strongly with hydrogen in diatomic molecules and least strongly with hydrogen in diatomic molecules and least strongly as stoichiometric solid hydrides. At free surfaces the strength of bonding is inter-

mediate. This suggests that the interaction of a metal with hydrogen depends strongly on the distance between metal atoms as illustrated in Figure 7. Strain of a metal surface can therefore be expected to strongly affect the chemisorption of hydrogen; especially the large strains found near the tips of cracks.

If repulsion effects determine the interaction energy then it is reasonable to approximate the curve of Figure 7 by:

$$E_b = E_b^\infty [1 - e^{-(d-d_s/\delta)}] - E_b^s \quad (10)$$

where d is the metal-metal distance, E_b^∞ is the interaction energy for full separation (dilute gas), E_b^s is the interaction energy in the solid hydride, d_s is the metal-metal distance in the solid hydride, and δ is a constant. The initial slope is:

$$\left(\frac{\delta E_b}{\delta d}\right) = \frac{E_b^\infty}{\delta} \quad [(d-d_s) \ll \delta] \quad (11)$$

Since E_b^∞ and E_b^s are often known from thermodynamic data, if the initial slope can be estimated this allows an estimate to be made of the binding energy of hydrogen at appropriate points near the tip of a crack.

Because of the strong interactions of hydrogen with free atoms and surfaces it can be expected to chemisorb at all available sites nearby a crack tip in most metals. Both glide suppression and cleavage enhancement may occur. Either will result in embrittlement of the metal.

VI. Summary

The importance of events within the atomic structure at the tip of a critical crack in an elastic-plastic solid is emphasized. Changes in the surface environment can change the energy needed to create a dislocation and the energy needed for cleavage. Such changes can shift the delicate balance between the tendencies toward glide and cleavage initiation at a crack tip. Inhibition of plastic deformation increases the tendency for cleavage, and numerical estimates of this effect are given which show that relatively small adsorption energies can have marked effects. The case of hydrogen adsorption is emphasized and it is concluded that since hydrogen interacts strongly with most elements it can readily cause embrittlement.

Acknowledgements

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

References

1. S. J. Burns and W. W. Webb - J. Appl. Phys. 41, 2078 and 2086 (1970).
2. J. Friedel - Fracture, Ed. by Averbach, et al., p. 498, J. Wiley and Sons, New York (1959).
3. J. J. Gilman - Trans. AIME 209, 449 (1957).
4. L. M. Keer - J. Mech. Phys. Sol. 12, 149 (1964).
5. R. M. Latanision and A. R. C. Westwood - Advances in Corrosion Sci. and Tech. 1, 51 (1970).
6. J. J. Gilman - Proc. Int. Conf. on Fracture, Sendai, Japan (1965), Vol. 2, p. 733.
7. J. N. Goodier - Fracture, Ed. by Liebowitz, Academic Press, New York, p. 44 (1968).
8. K. Kitajima - Proc. 1st. Int. Conf. on Fracture, Sendai, Japan (1965), Vol. 2. p. 647.
9. A. Kelly, W. R. Tyson, A. H. Cottrell - Phil. Mag. 15, 567 (1967).
10. J. R. Rice and R. Thomson - to be published.
11. H. C. Tong and J. J. Gilman - J. Appl. Phys., August (1971).

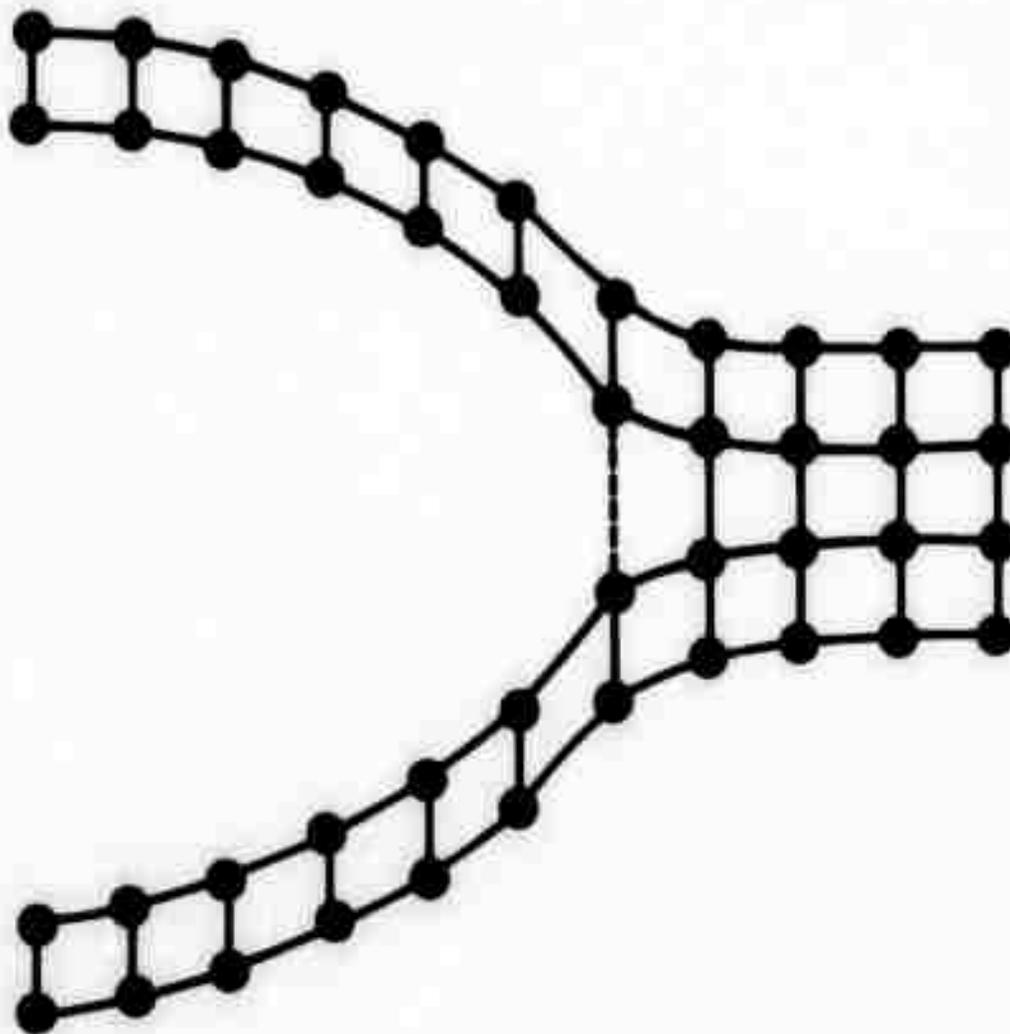


FIGURE 1

Configuration near tip of a crack in a simple cubic crystal.

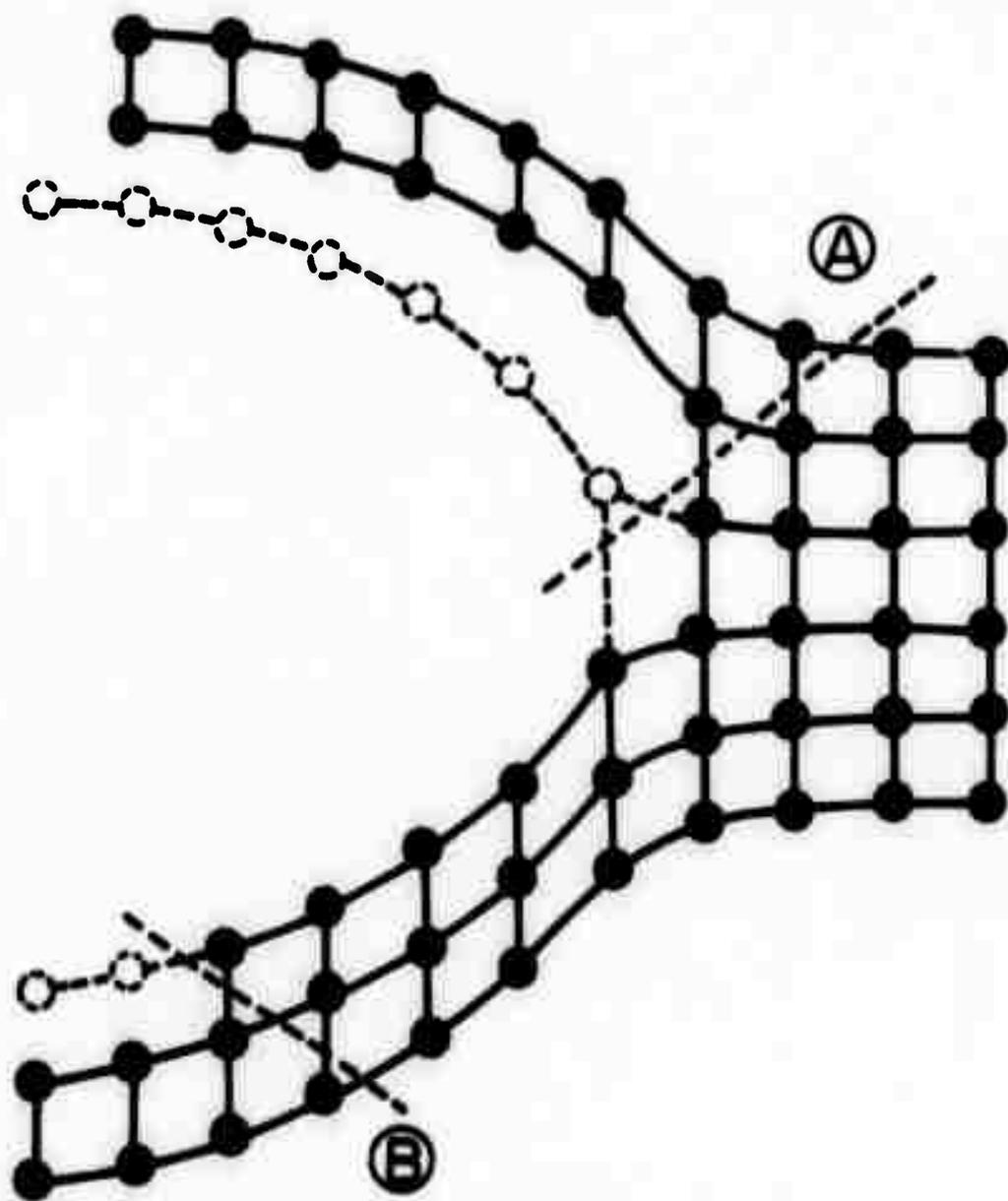
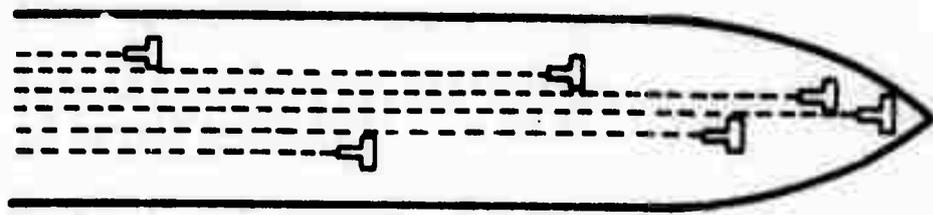


FIGURE 2

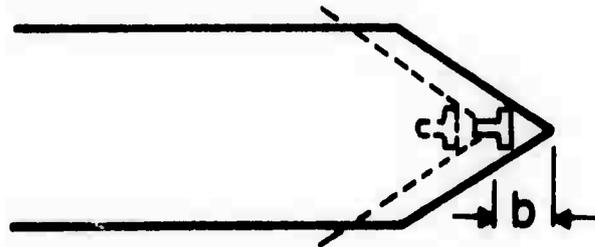
Schematic effects of plastic shears on the configuration near the tip of a critical crack.

(A)



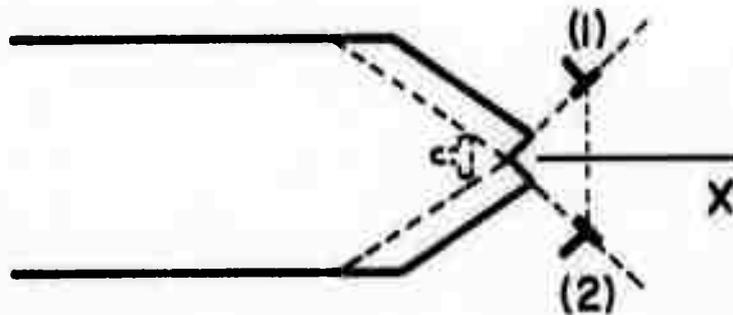
Array of crack dislocations representing crack.

(B)



Advance of crack by cleavage.

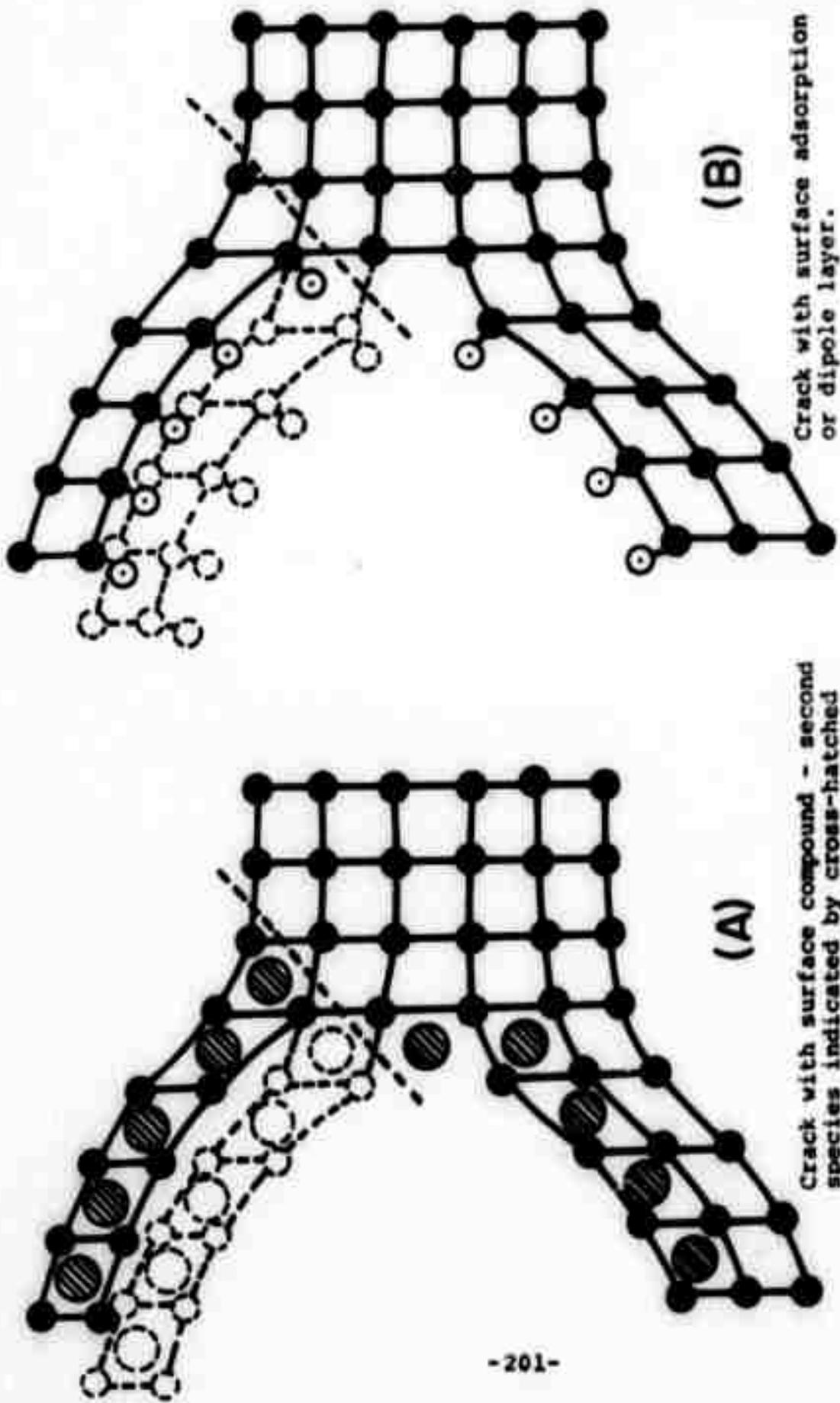
(C)



Blunting of crack tip by plastic shears

FIGURE 3

Dislocation models of crack tip processes



Crack with surface adsorption
or dipole layer.

Crack with surface compound - second
species indicated by cross-hatched
circles. Dotted lines show crack
configuration before shear.

FIGURE 4

Schematic effects of surface layers on plastic shears near crack tips.

H 104																	He —																												
Li 57	Be 53															B 78	C 80	N 74	O 102	F 135	Ne																								
Na 47	Mg 46															Al 67	Si 71	P 81	S 82	Cl 102	Ar —																								
K 43	Ca 39	Sc	Ti 30	V 9	Cr 66	Mn 55	Fe (64)	Co	Ni 60	Cu 66	Zn 20	Ga 67	Ge 74	As	Se 73	Br 87	Kr																												
Rb 40	Sr 39	Y 25	Zr 34	Nb 11	Mo (72)	Tc	Ru	Rh (82)	Pd (80)	Ag 53	Cd 16	In 58	Sn 62	Sb	Te 64	I 71	Xe —																												
Cs 41	Ba 41	La 50	Hf 27	Ta (83)	W (74)	Re	Os	Ir	Pt 83	Au 74	Hg 9	Tl 44	Pb 41	Bi 58	Po	At	Rn																												
Fr	Ra	Ac	<table border="1"> <tr> <td>Ce 49</td> <td>Pr 49</td> <td>Nd 48</td> <td>Pm 47</td> <td>Sm 44</td> <td>Eu</td> <td>Gd</td> <td>Tb</td> <td>Dy</td> <td>Ho</td> <td>Er</td> <td>Tm</td> <td>Yb 37</td> <td>Lu</td> </tr> <tr> <td>Th 35</td> <td>Pa</td> <td>U 31</td> <td>Np</td> <td>Pu 37</td> <td>Am</td> <td>Cm</td> <td>Bk</td> <td>Cf</td> <td>Es</td> <td>Fm</td> <td>Md</td> <td>No</td> <td>Lw</td> </tr> </table>															Ce 49	Pr 49	Nd 48	Pm 47	Sm 44	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb 37	Lu	Th 35	Pa	U 31	Np	Pu 37	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lw
Ce 49	Pr 49	Nd 48	Pm 47	Sm 44	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb 37	Lu																																
Th 35	Pa	U 31	Np	Pu 37	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lw																																

↑ TENDENCY FOR INCREASING STABILITY
→

DATA FROM: A.G. GAYDON - "DISSOCIATION ENERGIES"
 CHAPMAN & HALL, LTD., LONDON (1968)
 B.E. CONWAY & J.O. BOCKRIS - J. CHEM. PHYS. 26, 532 (1957)
 G.G. LIBOWITZ - "THE SOLID-STATE CHEMISTRY
 OF BINARY METAL HYDRIDES"
 W.A. BENJAMIN, NEW YORK (1965)

FIGURE 5

Bond energies of diatomic hydrides (kcal/mol)

- XX - From spectroscopic data (Gas)
- (XX) - Calculated from Pauling formula
- XX - Enthalpy of formation of solid hydride

STATE \ METAL	Ni	Pd	Ta
DIATOMIC GASEOUS HYDRIDE	-60	-60	-83
CHEMISORPED SURFACE LAYER	-30	-26	-45
SOLID HYDRIDE	+6.0*	-2.7	-9.5

* FREE ENERGY AT R.T. - THE ENTHALPY IS NEGATIVE

FIGURE 6

Effect of the state of typical metals on their interaction energies (enthalpies) with atomic hydrogen (kcal/mol).

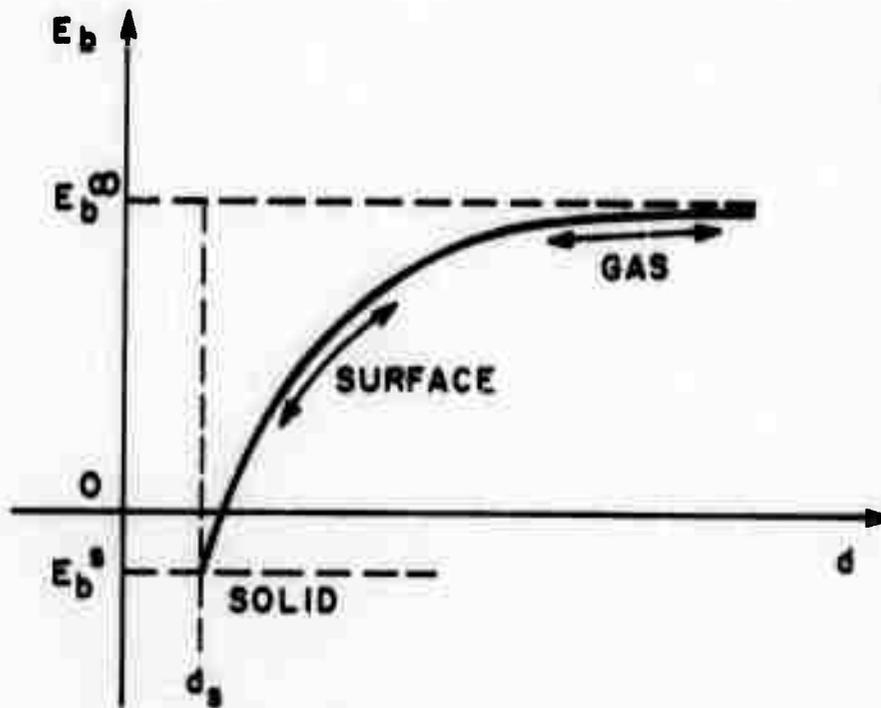


FIGURE 7

Effect of the distance between metal atoms on their interaction with atomic hydrogen (schematic).

A UNIFIED THEORY FOR THE FREE ENERGY
OF INHOMOGENEOUS SYSTEMS

L. A. Swanger
G. M. Pound
J. P. Hirth

Abstract

Recent theories for the free energy of inhomogeneous systems are considered and shown to be consistent when certain apparent discrepancies in coordinate systems are removed. The question of reduction of symmetry of a crystal to that of the group $C_{\infty v}$ in the presence of a field gradient is discussed. Conditions for the existence of odd-order gradient terms in a free energy expression are presented.

A UNIFIED THEORY FOR THE FREE ENERGY
OF INHOMOGENEOUS SYSTEMS

L. A. Swanger
G. M. Pound
J. P. Hirth

Introduction

The free energy of non-uniform systems has been expressed in terms of thermodynamic variables and their gradients by Landau and Lifshitz¹ and has been applied in detail to the case of concentration in binary solutions by Cahn and Hilliard². The latter treatment is valid for the case of cubic or isotropic materials with a concentration gradient that is small compared to the reciprocal of the intermolecular distance. Their result for the free energy density at a point

$$f = f + K(\nabla c)^2 + \dots \quad (1)$$

contains only even powers of the even concentration gradients. Combining this result with an expression for the strain energy of a two phase medium, Cahn² derived the modern description of the early stages of spinodal decomposition in the defect-free material.

An extension to the above theory was presented by Tiller, Pound, and Hirth (TPH)³ and Hirth, Tiller, and Pound (HTP)⁴. Their conclusion was that, in cases outside the rather restrictive assumptions of Cahn and Hilliard, there is no reason

to exclude the dependence of free energy on all powers of all orders of concentration gradients, specifically the first derivative which is also referred to as the linear term.

Indeed, the presence of such odd order terms produced by polarizations has been known for some time in crystal physics under the category of morphic effects.

A subsequent paper by TPH⁸ re-emphasized the original contribution of these authors: that a concentration gradient will, in general, induce gradients in other thermodynamic entities via relaxational accommodations. Such additional gradients could, in a phenomenological description, be assigned as properties of the material and thus cause linear terms to occur in an expression for the free energy in terms of concentration and its gradients.

Recently Morris⁹ has combined the above contribution from TPH⁸ with a statement from HTP⁴. There the authors stated that in general the total spatial derivative of energy is

$$\frac{dE}{dx} = \left(\frac{\partial E}{\partial c} \right) \left(\frac{\partial c}{\partial x} \right) + \left(\frac{\partial E}{\partial x} \right) \quad (2)$$

The first term just tabulates the dependence of energy on concentration but the second term admits the possibility of an intrinsic directionality in the material. Morris incorporated this concept by demonstrating that a material with symmetry derivable from the group $C_{\infty v}$, which is the case for the material properties discussed by TPH^{4,8}, will indeed have

directed axis and will include in a free energy expression a term that is linear in concentration gradient.

The former concept of induced fields was included by Morris in a set of parameters ϕ which are arguments of an energy-density functional. By including the set of parameters in a description of a reference state to one derivable from C_{av} . By using ϕ to describe the induced fields identified by TPH⁰, such reduction of material symmetry is accomplished.

Mathematical Models

By adopting a particular mathematical viewpoint, the concepts of the various contributors to this topic can be analyzed, organized and compared.

First consider an idealized class of binary materials that deviate from complete thermodynamic homogeneity only by variations in concentration. Here external fields, induced fields, and defect fields are excluded, leaving only concentration fields. For such materials the logical choice of reference state is one of uniform composition. The concentration is constant and equals the average value in the actual state. Then the free energy density at a point is expressible as a function of the independent variables concentration, first gradient of concentration, second gradient of concentration, etc.

$$f = f(c, Vc, V^2c, V^3c, \dots) \quad (3)$$

For such a system a set of "partial response curves" can be defined. These curves describe the dependence of the free energy density on each of the independent variables, with the other variables held constant at their reference state values. Figure 1 shows how these curves might look for a binary solution. Such curves are conceptually available either experimentally, by making measurements varying one parameter at a time, or by calculation assuming exact expressions for all interactions are available.

Figure 1 shows the various curves for a one dimensional crystal in laboratory coordinates, the basis for the following examples. The various imbedded coordinates and generalized three-dimensional cases are discussed in a later section.

Now consider the partial response curves for a cubic or isotropic medium, Fig. 2. Particularly note that the curve for $\partial c/\partial x = 0$. This is because, being isotropic, the material responds identically to a concentration gradient of given magnitude, independent of the direction of the gradient. Thus, f_2 is an even function of $\partial c/\partial x$ and as such can be expressed as a sum of a set of even orthogonal functions. A suitable set is the sequence of even powers. The method for determining the coefficients of each term is by Taylor's theorem:

$$f_2 = f_2 \Big|_{\frac{\partial c}{\partial x}=0} + \frac{\partial f_2}{\partial \left(\frac{\partial c}{\partial x}\right)} \Big|_{\frac{\partial c}{\partial x}=0} \cdot \frac{\partial c}{\partial x} + \frac{\partial^2 f_2}{2 \left[\partial \left(\frac{\partial c}{\partial x}\right)^2 \right]} \Big|_{\frac{\partial c}{\partial x}=0} \cdot \left(\frac{\partial c}{\partial x}\right)^2 + \dots \quad (4)$$

In this equation $\partial f_2^n / \partial^n (\frac{\partial c}{\partial x}) = 0$, $n = 1, 3, 5, \dots$. This is the origin of the statement by CH¹ that odd powers of the concentration gradient do not contribute to the free energy. This case may include the effects of induced gradients if the species of the induced concentration profile, e.g. vacancies in a crystal, are everywhere at equilibrium with the material. In this situation, partial response curves like those in Figure 2 would of course apply. Whenever this condition is not met, the state to be described is inaccessible by the usual Taylor expansion about $\partial c / \partial x = 0$, linear and higher odd-order forms do arise and the required phenomenological expression must be regarded as empirical. This important class of cases will be discussed later in the present section.

The partial response curve f_1 is in general not symmetric about $\partial^2 c / \partial x^2 = 0$, even for an isotropic material. Thus, an expression for this quantity in terms of a power series will include terms of all powers, even and odd. Cahn and Hillard's¹ coefficient K_1 is just the slope of f_1 at $\partial^2 c / \partial x^2 = 0$.

The partial response curves given in Fig. 3, still in laboratory coordinates, can arise in two ways: (i) the material itself is of a symmetry derived from $C_{\infty v}$, or (ii) a more symmetrical crystal has had its symmetry reduced by inclusion of an external gradient in the system of the reference state. Because it is no longer symmetric about $\partial c / \partial x = 0$, f_2 now needs to be expressed as a power series containing terms of all orders, with coefficients calculated using Taylor's theorem.

Expressions of this type were presented by Morris⁹.

Let us next consider the very important class of cases which arise when the species of the induced concentration profile, e.g., vacancies and/or dislocations in a crystal, are not at equilibrium with the material. A prime example of this situation occurs in spinodal decomposition of solids where vacancies are being pumped by high gradients in the regions of short half wave length ($\sim 100 \text{ \AA}$) between extrema which differ greatly in composition. It is clear that such states are inaccessible by the usual Taylor expansion about the normal state of uniform composition of the type given by Eq. (4). Instead of expanding about $\partial c/\partial x = 0$, the expansion of f_2 , if meaningful, must be about a value of $\partial c/\partial x$ for which the non-equilibrium effects become appreciable. Such a situation is depicted in Fig. 4 for both imbedded (Lagrangian) and the usual laboratory (Eulerian) coordinates. The former are preferred, because it adds nothing to consider the reverse direction of an external coordinate when describing the effect of an internal field. Note that curves such as that for $\partial c/\partial x$ in Fig. 3 are impossible to obtain in the induced field case. The reason that the usual Taylor expansion about $\partial c/\partial x = 0$ is not valid for the outer branches of the curves in Fig. 4 is the singularity at (i) where the differential coefficient of $\partial c/\partial x$ is discontinuous (in the present example) and the differential coefficient of $\partial^2 c/\partial x^2$ is non-existent. In simpler terms, the two positions of each of the curves in Fig. 4 are represented by different

functions. The function for the outer branch is different from the inner branch. The inner branch represents the situation where the induced concentration profiles are in equilibrium with the matrix. The outer branch is the sum of the function for the inner branch and another function which expresses the effect of the induced non-equilibrium concentration profile. One notes that this effect becomes appreciable only above some value (i) of the imposed gradient, which corresponds to the "critical" gradient for which the crystal symmetry is reduced by the induced field to the $C_{\infty v}$ group. The second term of Eq. (2) is now appreciable: with such effects, there must be linear and higher order odd terms in the expression representing the free energy density in spinodal decomposition and other important cases, although their magnitudes have not yet been estimated. As pointed out by TPH³ and HTP⁴, the required expansion must be regarded as empirical.

A specific mathematical form corresponding to Fig. 4 would be the Heavside function discussed by TPH and associated with dislocation formation. With effects of this type associated with vacancies, electron polarization and the like, the effect can in principle occur for very small gradients. On the scale of experimental observation, the slope discontinuity could be effectively at the origin as shown in Fig. 5. This case where $\partial f/\partial(\partial c/\partial x) \neq 0$ as $\partial c/\partial x \rightarrow 0$ is the specific case leading to concentration faceting treated by HTP⁴, CH⁵, and TPH⁶. In fact, the case of excess vacancy formation caused

by diffusion in a region of crystal which is free of dislocations or other vacancy sources is a good example of Fig. 5. In the absence of sources and sinks at quasi-steady state there can be no divergence of the vacancy flux which will thus be determined by boundary conditions removed from the region in question: hence, the local vacancy concentration will differ from both local equilibrium value and the value in the presence of a gradient but with local vacancy sources and sinks. From the viewpoint of diffusion theory, diffusivities for processes such as the early stages of spinodal decomposition or Nabarro-Herring creep, where no local vacancy sources are anticipated, should differ.

The calculation from first principles of the various coefficients in the free energy expression given above is not very illuminating. As pointed out by HTP⁴, and clarified above, determination of the coefficients is a process of parametric curve fitting, and does not involve a one-to-one correspondence between physical effects and coefficients.

All the above has been consistent with Morris'⁹ assumption of the existence of a functional \hat{F} that relates f to concentration and its gradients, temperature, and the set of thermodynamic parameters ϕ , which is continuous to all orders. It is only with the existence of this continuity that a Taylor's expansion about $\partial c/\partial x = 0$ to determine a power series representation for f is meaningful. By going outside that assumption, the topic of singularities becomes accessible.

HTP⁴ wrote their free energy function as a phenomenological expression rather than an analytic expansion in order to include situations outside the domain of the description of CH and Morris.

The present formulation can also be used to investigate the claim of HTP⁴ that two different coefficients are needed for $\partial^2 c / \partial x^2$, depending on the sign of the curvature (Eq. 5 in HTP⁴). If f_3 is an analytic function of $\partial^2 c / \partial x^2$, then such dual coefficients are meaningless. The curves for f_3 in Figs. 1, 2, and 3 are all expressible as

$$f_3 = K_1 \left(\frac{\partial^2 c}{\partial x^2} \right) + K_2 \left(\frac{\partial^2 c}{\partial x^2} \right) + K_3 \left(\frac{\partial^2 c}{\partial x^2} \right) + \dots \quad (5)$$

K_1 is just the slope of f_3 at $\partial^2 c / \partial x^2 = 0$. On the other hand, if it is desired to express the effect of $\partial^2 c / \partial x^2$ phenomenologically, as done by HTP⁴ in their Eq. (5), then two different coefficients are needed for $\partial^2 c / \partial x^2$. Further, these coefficients are a function of x .

The discussion on this point is in HTP⁴ at the bottom of their page 122, and the supporting diagram, their Fig. 2b. Then, given the two composition profiles of HTP⁴ Fig. 2a, indicated by their curvature on Fig. 7, the free energy curves of their Fig. 2b should be separated by an amount $f_3(u) - f_3(l)$ at x_0 . Furthermore, under the assumption that

$$\frac{df}{dx} = \left(\frac{\partial f}{\partial c} \right) \left(\frac{\partial c}{\partial x} \right) + \left(\frac{\partial f}{\partial x} \right) \quad \text{with} \quad \frac{\partial f}{\partial x} > 0 \quad (6)$$

the solid line of their Fig. 2b would be replaced by those shown in Fig. 8.

From the above viewpoint, it can be seen that Cahn and Hilliard treated only high symmetry systems, without external fields or such induced internal fields as non-equilibrium vacancy gradients and describable by Morris' functional (i.e., no singularities). Working within the limits of his functional, Morris extended the work of CH to include low symmetry cases, non-equilibrated internal fields, and external fields, all of which require terms linear in composition gradient to describe their free energy density. However, Morris' work was preceded by the concepts of HTP⁴ and TPH^{3,6}. Not only did the latter authors anticipate Morris' careful treatment of the continuum model, but they also recognized the role that singularities could play.

Since there has been considerable discussion of coordinates^{3,6,9} some comment on the extension of the above discussion to three dimensions is in order; a two dimensional case suffices to indicate the three dimensional case. Fig. 9 shows x, y imbedded coordinates, x', y' laboratory coordinates with x'', y'' crystal symmetry coordinates. HTP⁴ treated the case where x, y and x'', y'' coincided, $\delta = 0$. Hence, the scalar nature of the free energy (invariance to coordinate transformation) is automatically satisfied in their work, contrary to the assertion of Morris⁹ which arose because of a misunderstanding of the above choice. Moreover, HTP⁴ Eq. (2) is

is correct, again contrary to Morris' assertion, and corresponds exactly to Morris' Eq. (10). CH⁵ purported to show that HTP⁴ Eq. (2) violated symmetry requirements but they assumed that the equation was written in laboratory coordinate x', y' . In x', y' coordinates, the free energy must be invariant to rotation $\Delta\phi$; then, of course, the form of HTP⁴ Eq. 2 would not hold. For example, the linear term would have to be cast in the invariant form $\alpha|\nabla c|$ as indicated by CH⁵. Also, more meaningful from the crystal physics viewpoint, the free energy must be invariant to rotation $\Delta\delta$ as noted by Morris⁹. Again, the form HTP⁴ Eq. (2) would have to be recast for the case $\delta \neq 0$, and the linear term would become $\alpha|\nabla c|$ as noted by Morris⁹.

Hence, Eq. (2) of HTP⁴ is correct and certainly applies to the various one-dimensional applications^{3,4}. Morris' formulation better reflects the crystal symmetry influence and is more useful in three-dimensional cases. However, physical cases where this model applies correspond directly to those predicted for the simpler geometry discussed by HTP⁴.

Kinetic Effects

As mentioned in the Introduction, HTP⁴ suspected and CH⁵ demonstrated that singularities describable by $\alpha|\partial c/\partial x|$ will result in concentration facets during diffusional processes. One way to illustrate this result is by considering the volume integral of the free energy density of a fluctuated system. Considering just the contribution from the $\alpha|\partial c/\partial x|$ term:

$$F_2 = \int_{x_1}^{x_2} a_0 \left| \frac{\partial c}{\partial x} \right| dx = a_0 \left| \int_{x_1}^{x_2} dc \right| = a_0 |c_2 - c_1|$$

To make physical sense, this integral must be performed piecewise between local extrema. The only variations that lower F_2 are those that reduce the deviation of local extrema. So the only diffusive driving force is at the extrema, where $\partial c / \partial x = 0$, thus flat concentration facets are produced.

The technique can be applied to assess the consequences of singularities in $\partial^2 c / \partial x^2$. If the dual-coefficient statement of HTP⁴ is combined to yield a term in $|\partial^2 c / \partial x^2|$:

$$K_1 \left(\frac{\partial^2 c}{\partial x^2} \right)_+ + K_2 \left(\frac{\partial^2 c}{\partial x^2} \right)_- = K \left| \frac{\partial^2 c}{\partial x^2} \right| + (K_1 + K_2) \left(\frac{\partial^2 c}{\partial x^2} \right)_- \quad \begin{array}{l} K_1 > 0 \\ K_2 < 0 \end{array}$$

then an integration of that portion may be performed.

$$\int_{x_1}^{x_2} \left| \frac{\partial^2 c}{\partial x^2} \right| dx = \left| \frac{\partial c}{\partial x} \right|_{x_2} - \left| \frac{\partial c}{\partial x} \right|_{x_1}$$

Again, this integration must be performed piecewise between local extrema in $\partial c / \partial x$ (where the second derivative is zero), and the result would be faceted concentration gradient profiles. Faceted gradient profiles can also arise from linear gradient dependence of the diffusivity or mobility as discussed by TPH⁵.

For the case of a linear term of the type $a \cdot \partial c / \partial x$, another case considered by Morris⁶, integration need not be piecewise since negative contributions would cancel positive ones. Consider Fig. 10 with five local extrema shown.

$$\int_{x_1}^{x_3} \alpha_0 \frac{\partial c}{\partial x} dx = \alpha_0 \int_{x_1}^{x_2} dc = \alpha_0 (c_3 - c_1)$$

The high energy gradient from A to 2 just cancels the low energy gradient from 2 to 3. The energy of this region is independent of c_2 , so there is no driving force for diffusion. Integrating over the whole sample:

$$\int_0^L \alpha_0 \frac{\partial c}{\partial x} = \alpha_0 \int_0^L dc = \alpha_0 (c_L - c_0)$$

shows that the only diffusive driving force is at the boundaries of the specimen. Thus, there is no internal faceting predicted for this type of linear term.

Time Dependent Linear Gradient Terms

Another interesting topic is the transient effect of induced fields. Let us consider the example of transient non-equilibrium vacancy gradients. It is known that in systems that exhibit a Kirkendall effect, net transport of atoms in one direction is accompanied by transport of vacancies in the opposite direction. In processes such as spinodal decomposition, during which no steady-state is reached for some considerable time, the vacancy profiles left behind by previous concentration profiles will provide a time-dependent reference state, or a history effect. Computer simulations of spinodal decomposition by Swanger, Cooper and Gupta¹⁰ and Gupta and Cooper¹¹ have shown that as time progresses, the concentration gradients at

a point can actually change sign (direction). Thus, concentration gradients can temporarily interact with vacancy gradients, opposing those they would induce themselves at steady state.

The vacancy gradient could be treated as a time-dependent member of Morris' set ϕ^s , and thus yield a term $\alpha \cdot \partial c / \partial x$ in the free energy expression. As shown above, this would not normally produce any diffusive driving force. However, the unique character of the vector α describing the vacancy gradient is that its direction is position dependent within the material. In Fig. 11 a vacancy profile and partial response curves for three different regions are shown. Possible concurrent concentration profiles to be imposed on the material are shown in Fig. 12, along with the resulting free energy densities. Considering the variations in concentration that would lower the integral free energy, one would expect profile I to facet at the top, profile II to cusp at the bottom, and profile III to shift to the left and steepen on its right side. These tendencies are shown as dotted lines. These are transient effects, operative only as long as the vacancy profile is intact.

Conclusions

By adopting a general viewpoint it is possible to reconcile the various postulates made about the free energy of homogeneous solutions, as well as to deduce further aspects of their nature. It is hoped that by making the nature of the

expansions and empirical series describing free energy density in terms of concentration and its gradients quite explicit, the contributions of all the workers in this field can be appreciated.

Acknowledgements

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan. Dr. J. W. Morris, Jr. provided helpful discussions which served to remove apparent discrepancies between earlier works on this subject with regard to coordinate systems.

References

1. Cahn and Hilliard, J. Chem. Phys. 28, 258 (1958).
2. Cahn, Acta Met. 9, 795 (1961).
3. Tiller, Pound and Hirth, Acta Met. 18, 225 (1970).
4. Hirth, Tiller and Pound, Phil. Mag. 22, 117 (1970).
5. Cahn and Hilliard, Acta Met. 19, 151 (1971).
6. Cook and Hilliard, J. Appl. Phys. 40, 2191 (1969).
7. Philofsky and Hilliard, J. Appl. Phys. 40, 2198 (1969).
8. Tiller, Pound and Hirth, Acta Met. 19, 475 (1971).
9. Morris, Phil. Mag. 23, 1041 (1971).
10. Swanger, Cooper and Gupta, Acta Met. 18 9 (1970).
11. Gupta and Cooper, Phil. Mag. 21, 611 (1970).

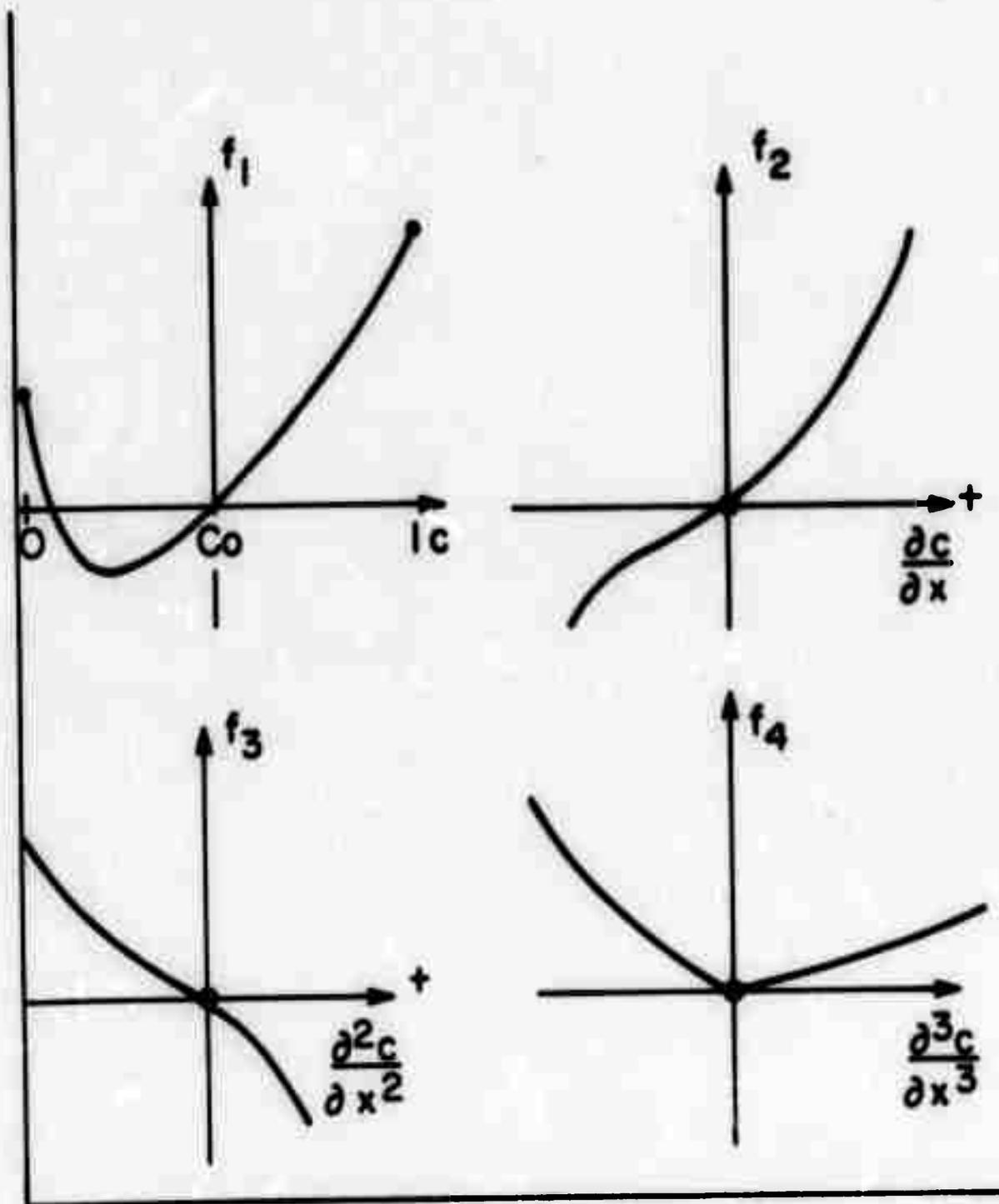


Figure 1. Partial response curves for an arbitrary binary solution.

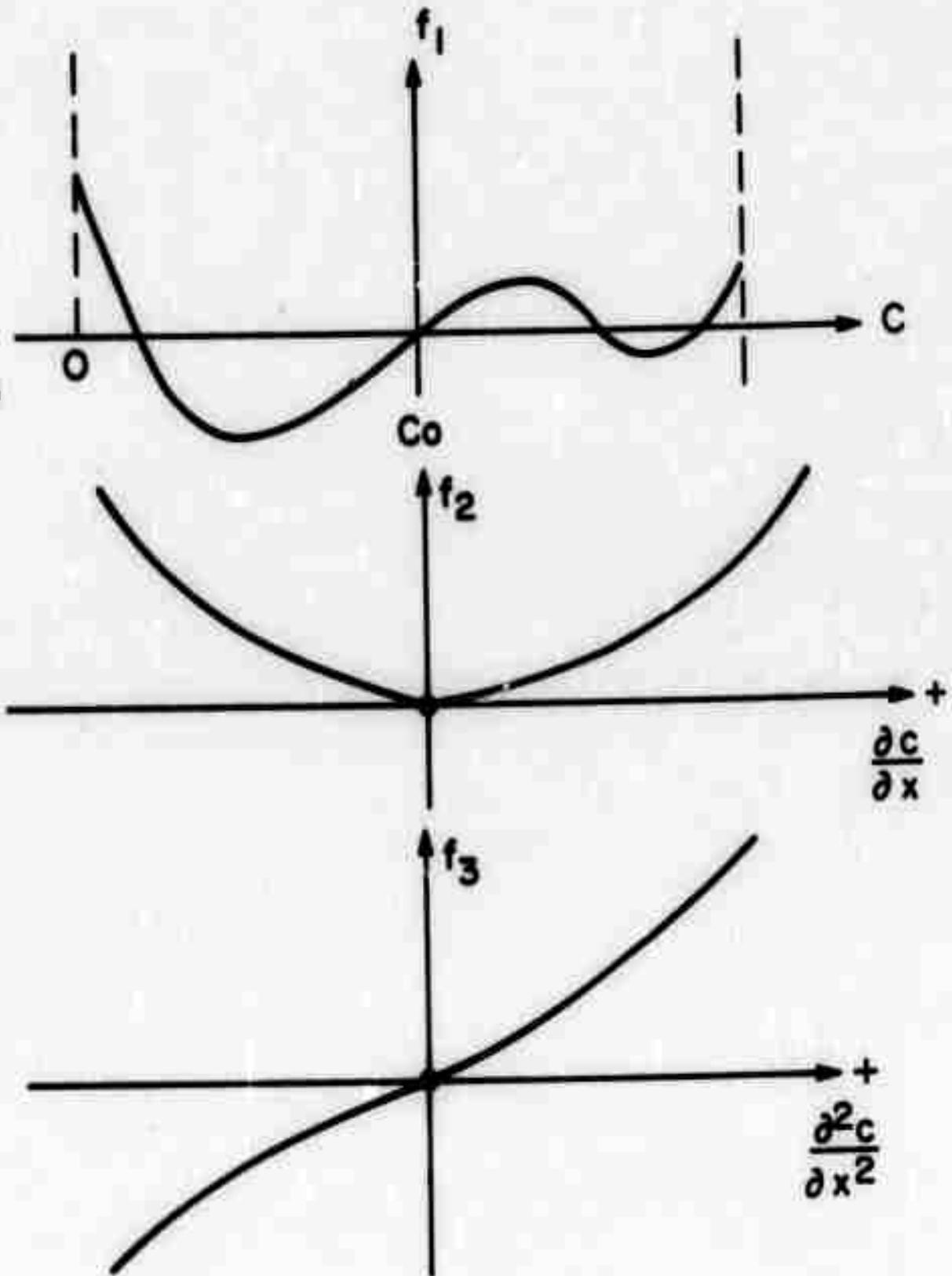


Figure 2. Partial response curves for an isotropic material.

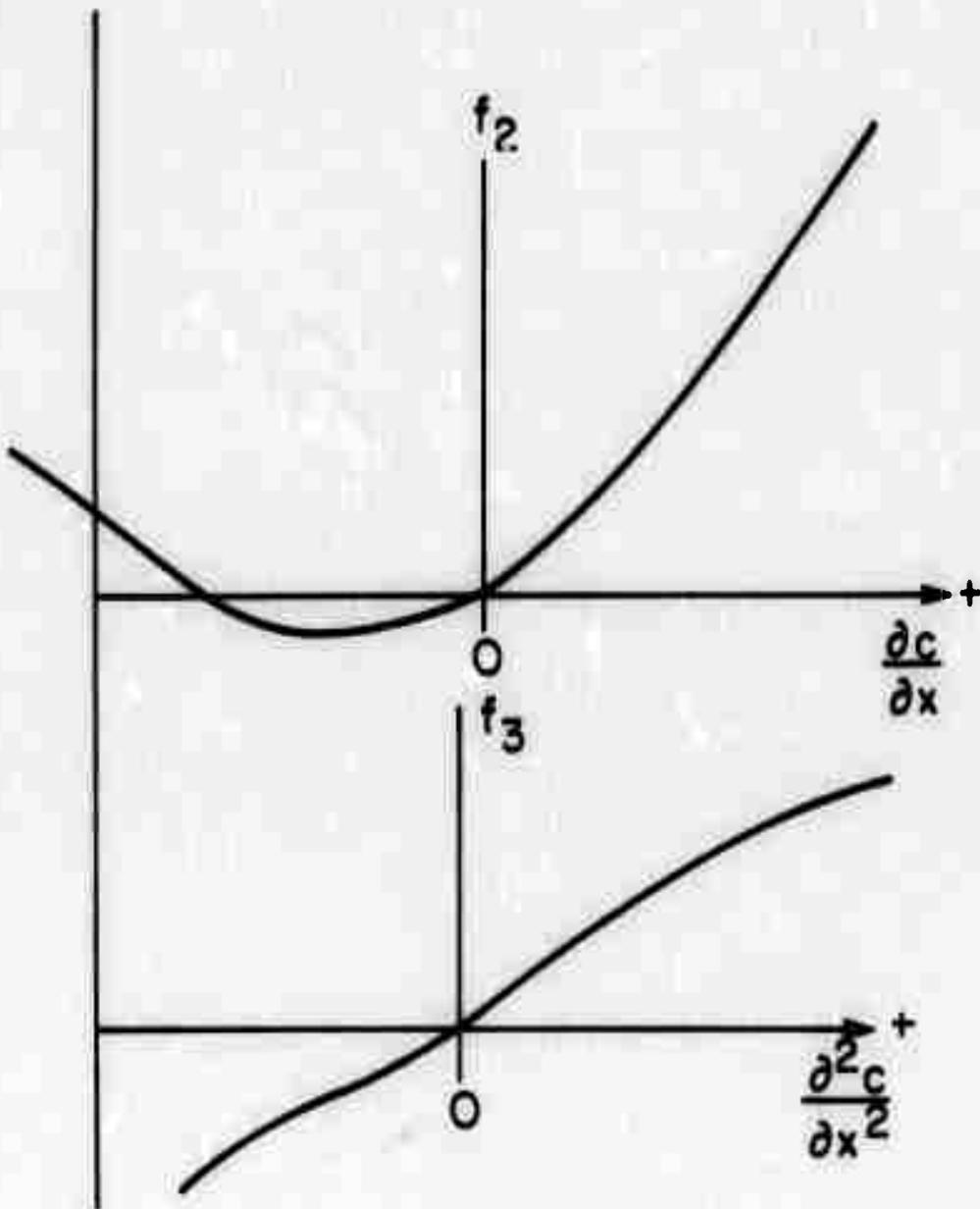


Figure 3. Partial response curves for a solid with symmetry $C_{\infty V}$, or a system including directed external fields.

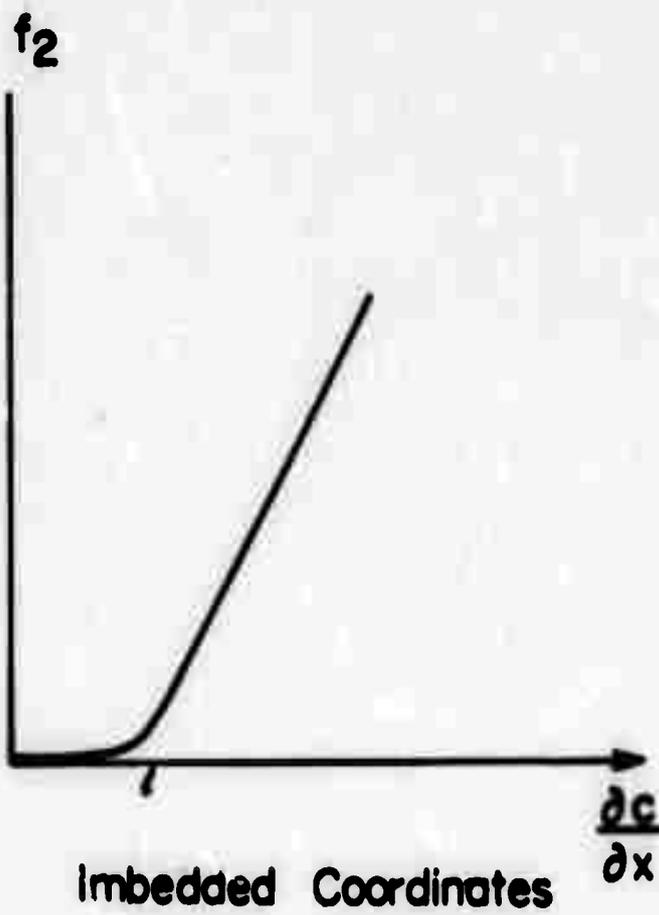
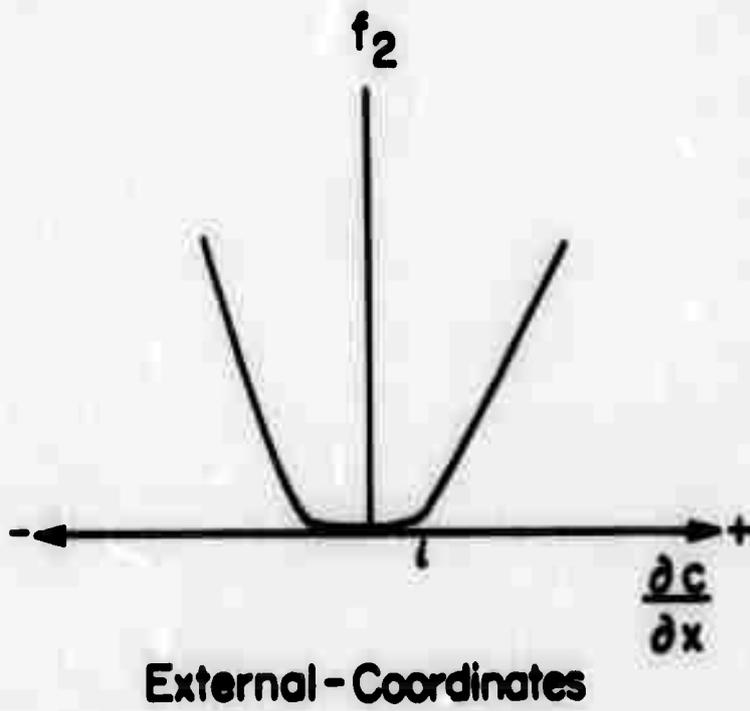


Figure 4. Partial response curves for induced internal fields.

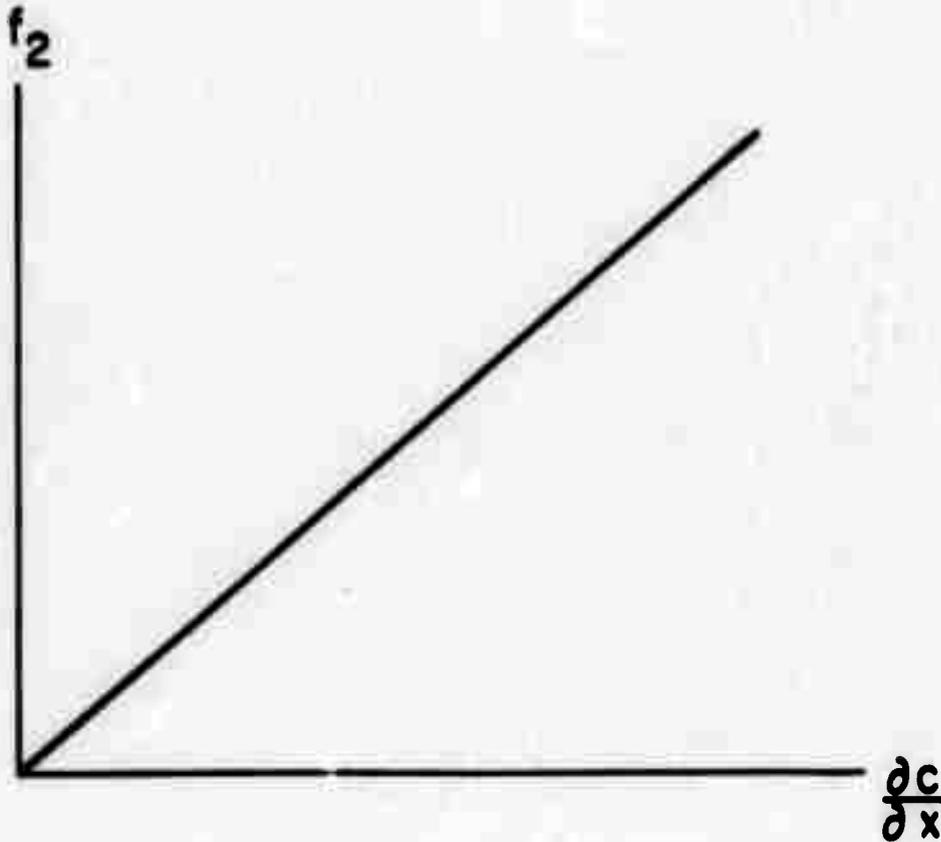


Figure 5. Plot of free energy density versus gradient for induced non-equilibrium vacancy concentration profiles.

Diffusion Current
or
 $\frac{\partial C_{v_{xs}}}{\partial x}$

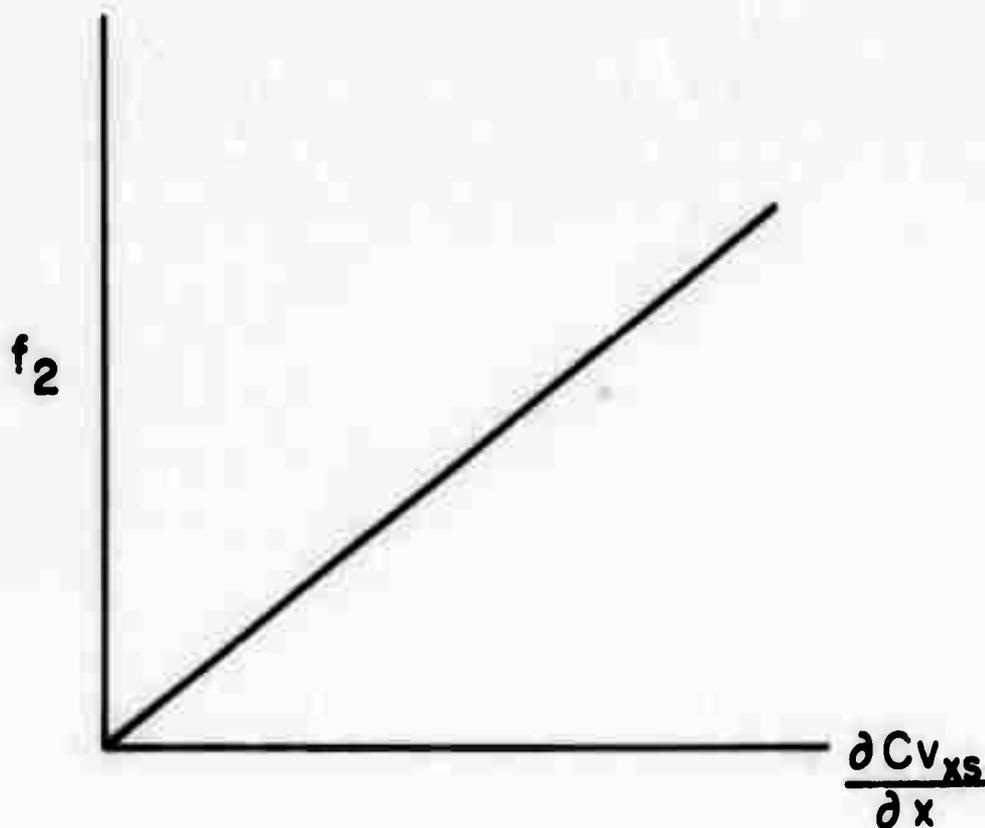
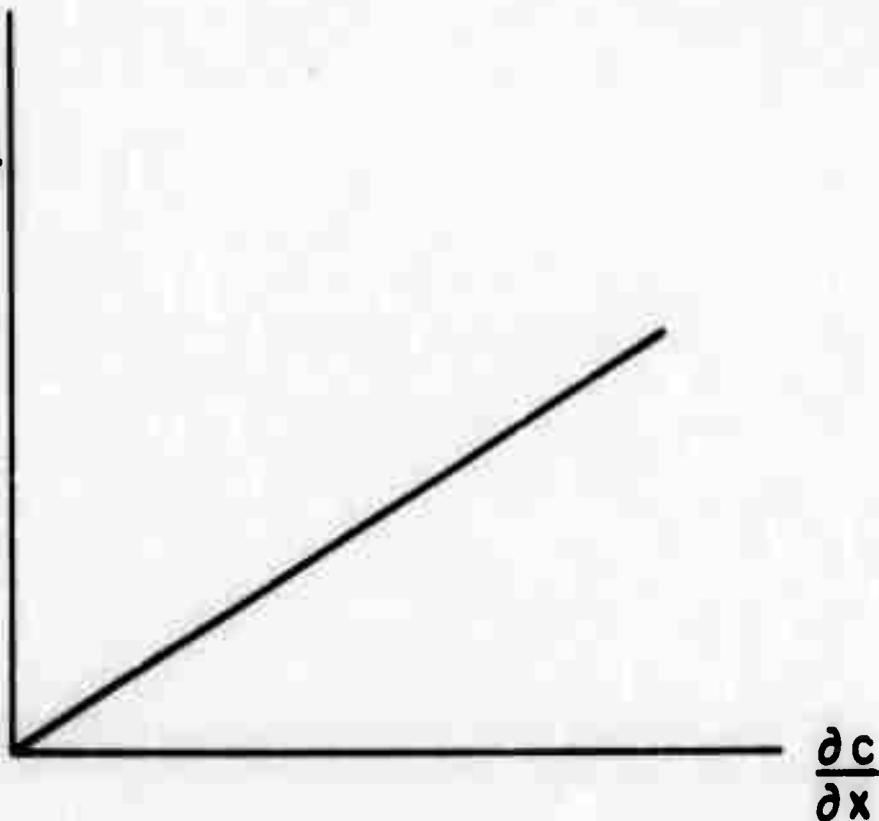


Figure 6. Diffusion current or excess vacancy concentration gradient $\frac{\partial C_{v_{xs}}}{\partial x}$ versus $\frac{\partial c}{\partial x}$ and f_2 versus $\frac{\partial C_{v_{xs}}}{\partial x}$.

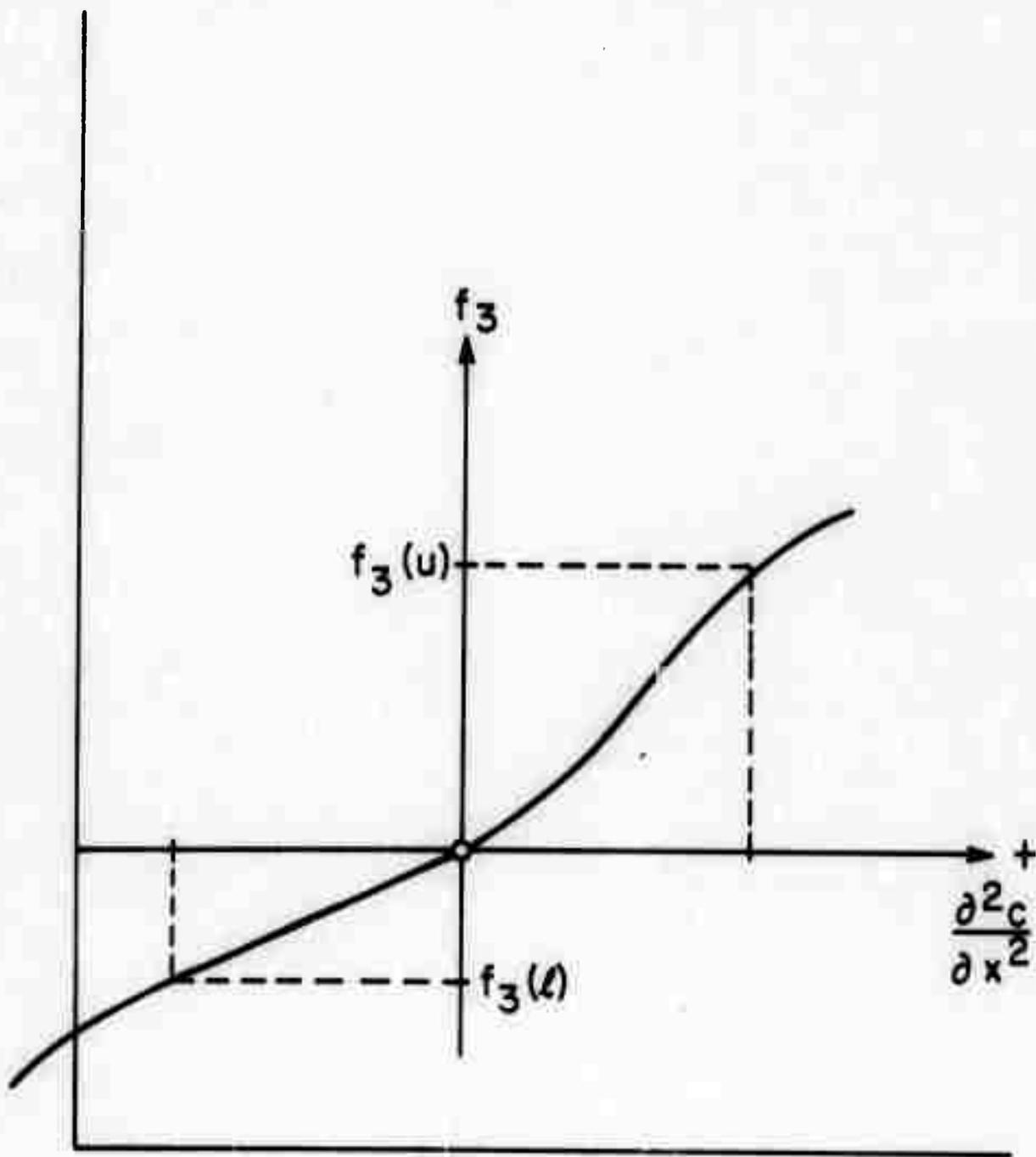


Figure 7. General analytic partial response curve for $\partial^2 c / \partial x^2$.

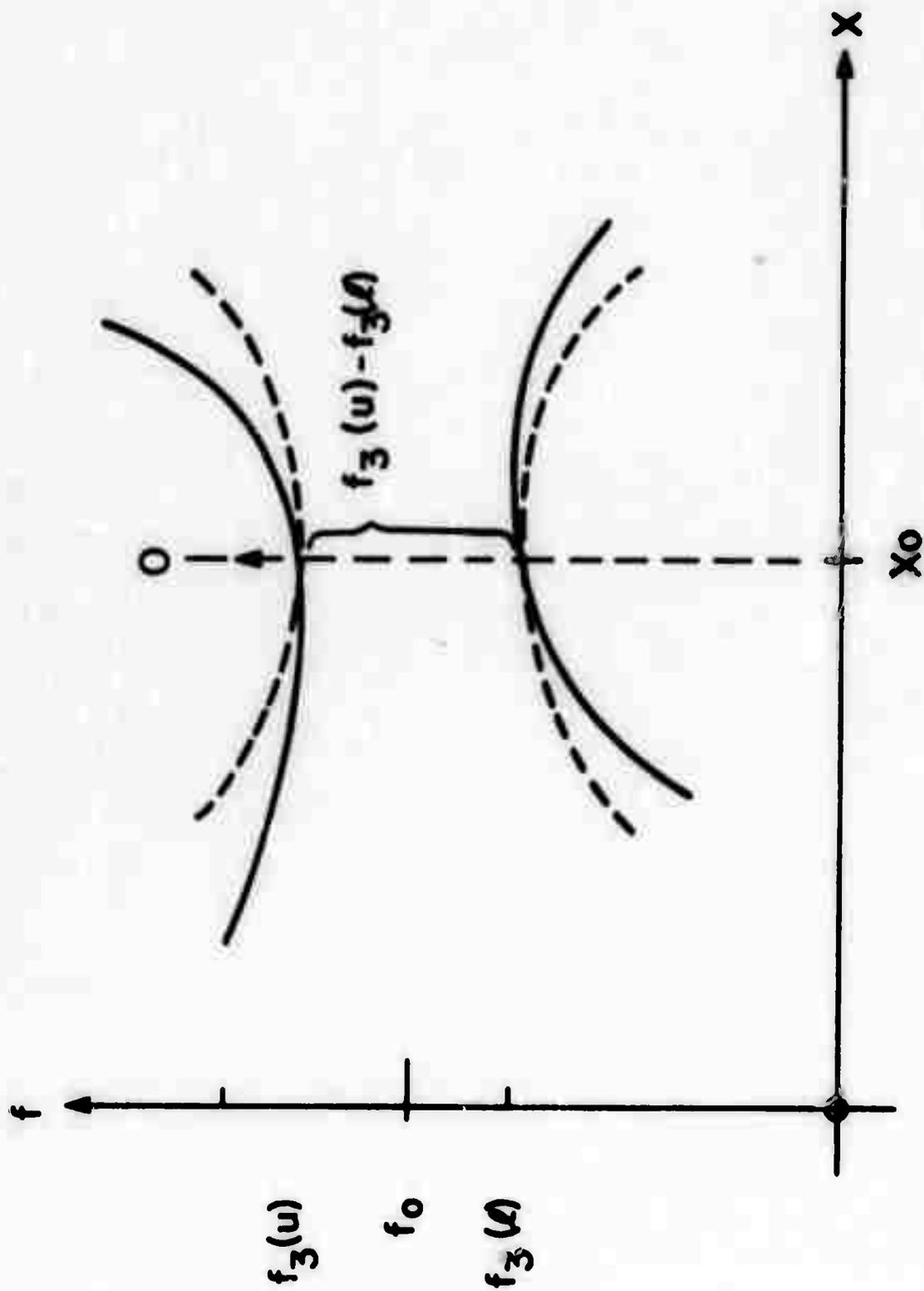


Figure 8. Correction of HTP(4)'s Figure 26.

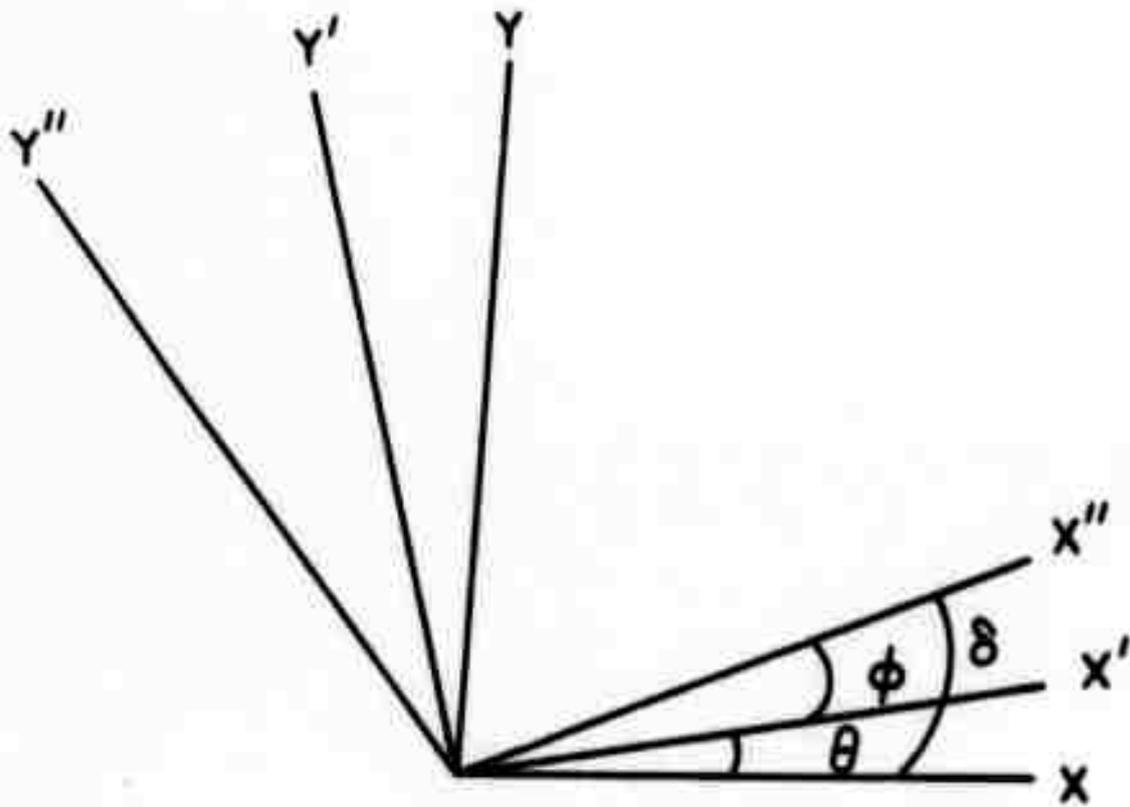


Figure 9. Various coordinates for a crystal; x, y are imbedded, x', y' are laboratory and x'', y'' are crystal symmetry axes, respectively.

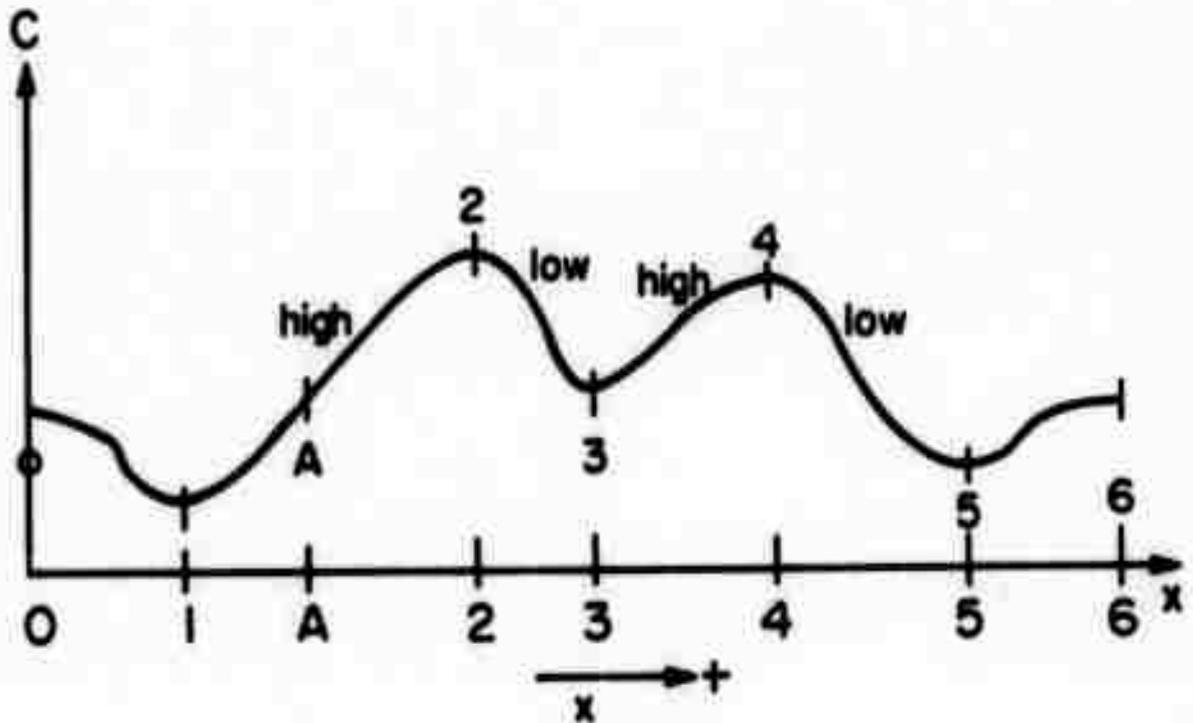


Figure 10. Complete profile for illustration of effect of $\frac{\partial C}{\partial x}$.

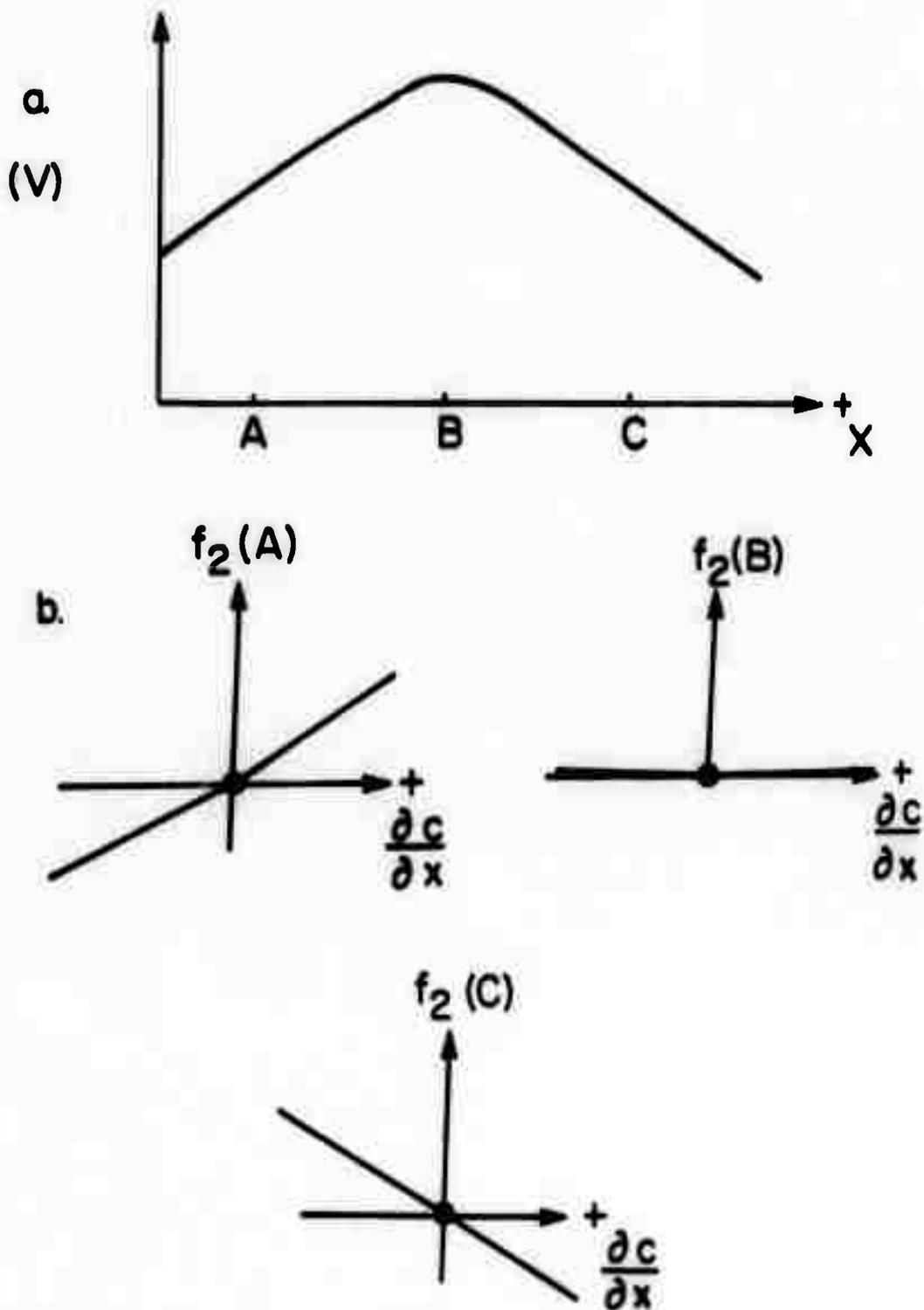


Figure 11. a. Vacancy profile
 b. Partial response curves

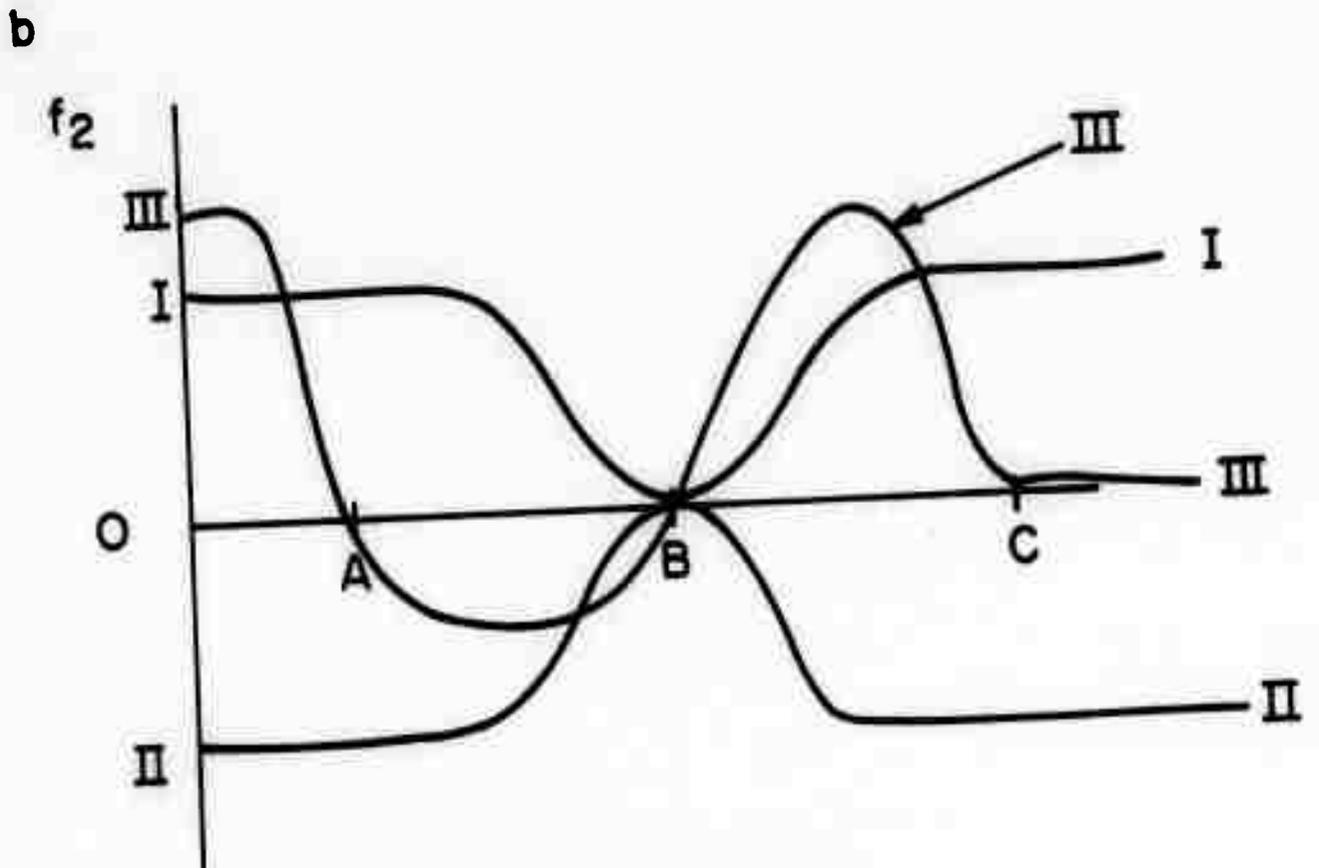
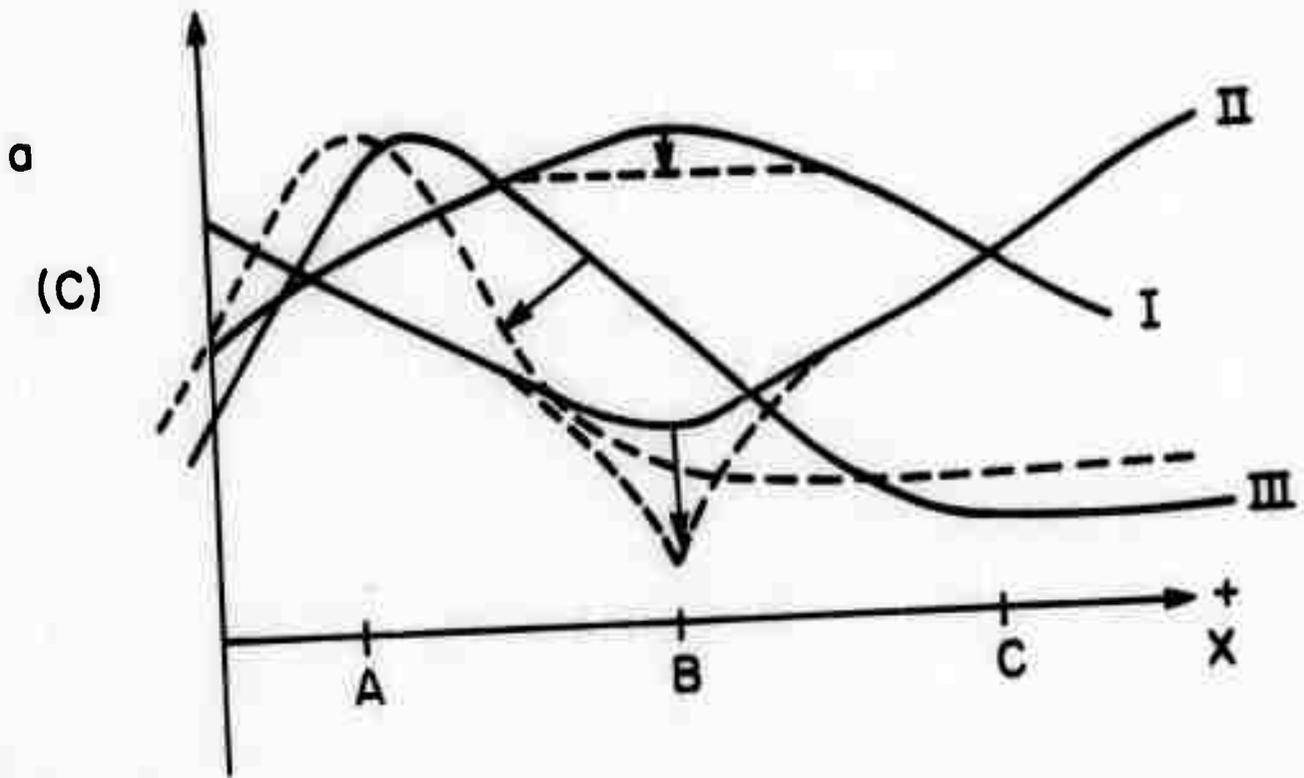


Figure 12. a. Concentration profiles
b. Energy densities

THEORY OF IONIC TRANSPORT
IN CRYSTALLOGRAPHIC TUNNELS

W. H. Flygare and R. A. Huggins

ABSTRACT

A theoretical model has been developed for the treatment of the motion of ions through crystallographic tunnels, as are found in materials that are interesting solid electrolytes. This model, which includes consideration of both point charge and higher order attractive terms as well as overlap repulsion effects, allows the calculation of the minimum energy positions of mobile ions and the activation energy barrier that they must surmount to move through the lattice. Calculations have been made for ions of different sizes in the AgI lattice which show the experimentally observed dependence of mobility on ionic size, and initial steps have been taken toward extensions to more complicated structures.

THEORY OF IONIC TRANSPORT IN CRYSTALLOGRAPHIC TUNNELS

W. H. Flygare and R. A. Huggins

Introduction

Experiments have shown that unusually large values of diffusion coefficient and ionic conductivity are found in several groups of materials which have crystal structures that are characterized by the existence of linear or nearly linear tunnels. Under proper conditions certain ions (e.g., the alkali ions) can readily move along such tunnels under the influence of chemical or electrical forces. The magnitude of the resulting transport fluxes in some electronic insulators make them particularly interesting because of their potential use as solid electrolytes in new types of batteries or fuel cells.

One of the interesting aspects of the available data is that there appears to be an optimum size for the ion translating through the tunnel. Ions significantly smaller than the transverse dimensions of the tunnel, as well as much larger ions both appear to have lower values of mobility than ions of intermediate size.

The Model and Calculations

The purpose of this paper is to examine a model to calculate the energy (as a function of position) of a mobile

ion in a crystal lattice. This model may allow a calculation of the minimum energy position for the mobile ion as well as the activation energy or energy barrier between minimum positions.

The general approach will be to examine increasingly complex crystal lattices in the following order:

1. AgI, Ag^+ ion motion
2. TiO_2 , Alkali ions (Li^+ , Na^+ , K^+ , Rb^+ , and Cs^+) in motion
3. β -alumina, $\text{Na Al}_{11}\text{O}_{17}$, Na^+ ions in motion

The energy of an ion in a lattice is described by a sum of electrostatic and overlap repulsion terms. The electrostatic term includes the sum over point charge interactions (Madelung sum) and the higher order terms which include the point charge-induced dipole, point charge-induced quadrupole, dipole-dipole (point charge induced), and dipole-dipole (dispersion) terms. The general summations used in alkali halide lattices have been examined by Quigley and Das¹ to show that the Li^+ substitutional impurity in KCl has an energy minimum off-center along both the $\langle 111 \rangle$ and $\langle 100 \rangle$ axes. The activation energy for diffusion of interstitial lithium, copper, and silver impurities in the diamond lattice has also been evaluated by Weiser². Both of these calculations used techniques developed by others³.

In the present work we will use the following electro-

static energy for the j^{th} point charge ion in a lattice:

$$E = e^2 \sum_j \frac{q_i q_j}{r_{ji}} - \frac{e^2}{2} \sum_j \frac{\alpha_j}{r_{ji}^3} q_i \quad (1)$$

The sum over j is over all lattice ions, q_i and q_j are the fractions of charge of the mobile ion and fixed lattice charges respectively, α_j is the dipole polarizability of the i^{th} fixed lattice atom and r_{ij} is the distance from the mobile ion to the j^{th} lattice ion. We are assuming that the polarizability of the mobile ion is much less than the ions in the host lattice.

The second term in Eq. (1) is the polarization self-energy of a non-polarizable mobile ion in the lattice of fixed charges. The orientational dipole-dipole terms,

$$E_{\mu} = \frac{\underline{\mu}_1 \cdot \underline{\mu}_2}{(r_{12})^3} - \frac{3(\underline{\mu}_1 \cdot \underline{r}_{12})(\underline{\mu}_2 \cdot \underline{r}_{12})}{(r_{12})^5} \quad (2)$$

are neglected. These terms are normally smaller ($\mu = e a$, $E = e q_i / r^2$) than the corresponding scalar self-energies by a factor of $1/r^2$ for the dominant near-neighbor terms. In any event, the orientational energy in Eq. (2) will not be included in this calculation. This approximation is consistent with the method used by Weiser². The overlap repulsion term between closed-shell ions is given by⁴

$$E_{\text{rep}12} = \beta_{12} e^{3(r_1 + r_2 - r_{12})} \quad (3)$$

where r_1 and r_2 are the ionic radii of ions one and two, r_{12} is the interionic radii, and B_{12} is a multiplicative factor depending on the nature of the interacting ions⁵. The parameters used in the repulsive term in Eq. (3) and the polarization term in Eq. (2) are given in Table I. The final energy of the i^{th} mobile ion is

$$E = e^2 \sum_j \frac{q_i q_j}{r_{ji}} - \frac{e^2}{2} \sum_j \frac{\alpha_j q_i}{r_{ji}^4} + \sum_j B_{ij} e^{3(r_1+r_2-r_{12})} \quad (4)$$

Eq. (4) diverges as $r_{ij} \rightarrow 0$ which is sometimes referred to as the "polarization catastrophe". The energy will converge by making α_i dependent on the interionic distance to lower α_i as the ions approach each other or by converting the exponential repulsive term to a $1/r_{ij}^{12}$ term when the ions are closer than the sum of ionic radii. These refinements are easy to apply, but Quigley and Das² have shown that the minimum positions for the substitutional Li^+ impurity in KCl were predicted with Eq. (4). Thus, we will use Eq. (4) directly without correcting for the polarization catastrophe at small r_{ij} . Our analysis will be valid for $r_1+r_2 < r_{12}$.

We will now attempt to apply this model to a number of crystal lattices.

The AgI Structure

Ag^+ ions have unusually high mobilities and diffusion constants in the AgI lattice. The crystal structure of AgI has been given by Stroock⁶. A diagram of the AgI lattice is

given in Fig. 1. According to Strock, the I^- ions are much larger than the Ag^+ ions. Thus, the I^- atoms dominate the crystal structure, which is body-centered cubic in I^- . There are several near-equivalent positions for the Ag^+ ions as indicated in the diagram. The large number of energetically near-equivalent and vacant sites for Ag^+ in AgI leads to the high mobility. Apparently, tunnels of relatively constant potential energy occur throughout the lattice.

The positions of the two types of I^- ions in the lattice are (see Fig. 1):

$$R_1 = ((x-Ia)^2 + (y-Na)^2 + (z-Ma)^2)^{1/2}$$

$$R_2 = ((x-a/2-Ia)^2 + (y-a/2-Na)^2 + (z-a/2-Ma)^2)^{1/2} \quad (5)$$

The energy of an Ag^+ ion was calculated as a function of x , y , and z in the host I^- lattice according to Eq. (4), the R_i values in Eq. (5), and the parameters in Table I. I , N , and M are the multipliers of the lattice distance ($a=5.034$) along the x , y , and z axes. A sample calculation with various values of x , y , and z is listed in Table II. As only I^- ions are included in the sum, the energy will increase with increasing size of the lattice which is included in the calculation ($NMAX$, $IMAX$, and $MMAX$). However, the energy differences should converge with increasing values of $NMAX$, $IMAX$, and $MMAX$. $NMAX = IMAX = MMAX = 9$ values were found large enough to achieve convergence. The scan along the x axis (see Fig. 1) in Table II is symmetric

and shows definite minima in the energy at the following values:

<u>x</u>	<u>y</u>	[z = 2.517 = a/2]
0.000	0.2517	
0.252	0.2517	
0.503	0.5034	
0.755	0.5034	
1.007	0.7551	
1.258	1.0068	
1.510	1.2585	
1.762	1.2585	

These minima positions correspond roughly with the \odot positions shown in Fig. 1.

A sequence of calculations was then performed as a function of RPLU (the cation size) with the following general conclusions.

1. The energy at $x=0, y=0, z=2.5170$ (positive ion lattice point) is lower for the light positive ions because of the smaller repulsive contribution (RPLU small).
2. At $y=0, z=2.5170$, the ΔE along the x-axis increases with RPLU.
3. A minimum was found in a location near the \odot positions in Fig. 1 for positive ion (RPLU) values below 1.0 and a maximum was found near the \odot positions for $RPLU > 1.0$. The cross-over point where there are energy minimum paths of $\Delta E \approx 0$ appears to be about $RPLU = 1.1$. A diagram of the path of minimum energy in the $z=2.5170$ xy plane is shown in Fig. 2.

The TiO₂ Structure

Several of the positive alkali ions (as impurities) have unusually high mobilities in TiO₂. The TiO₂ lattice is shown in Fig. 3. There are four O²⁻ ions and two Ti⁴⁺ ions in each unit cell and the coordinates of these 6 ions are (see Fig. 3 for numerical notation).

cations

$$R_1 = ((x-aI)^2 + (y-aN)^2 + (z-cm)^2)^{1/2}$$

$$R_2 = ((x-a/2-aI)^2 + (y-a/2-aN)^2 + (z-c/2-cm)^2)^{1/2}$$

anions

$$R_3 = ((x-u-aI)^2 + (y-u-aN)^2 + (z-cm)^2)^{1/2} \quad (6)$$

$$R_4 = ((x+u-aI)^2 + (y+u-aN)^2 + (z-cm)^2)^{1/2}$$

$$R_5 = ((x-u-a/2-aI)^2 + (y-a/2+u-aN)^2 + (z-c/2-cm)^2)^{1/2}$$

$$R_6 = ((x+u+a/2-aI)^2 + (y+a/2-u-aN)^2 + (z+c/2-cm)^2)^{1/2}$$

The energy of a number of positive ions in the tunnel region of this structure (Fig. 3) can be obtained as a function of x , y , and z as described previously for the AgI case. These calculations are still underway, and the results will be presented in a subsequent report.

Discussion

This model represents a new approach to the explanations of the unusually high values of ionic mobility found in certain types of solids. Initial calculations using the AgI structure indicate that it successfully corresponds to the important experimental observations. Further work should be done to apply

this approach to other more complex structures and to investigate its application as a tool to aid the search for new solid electrolytes.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

References

1. R. J. Quigley and T. P. Das, *Solid State Comm.* 5, 487 (1967).
2. K. Weiser, *J. Phys. Chem. Solids* 17, 149 (1960), and *Phys. Rev.* 126, 1427 (1962).
3. C. J. F. Bottcher, Theory of Electric Polarization (Elsevier Publishing Co., Inc., New York, 1952), N. F. Mott and M. J. Littleton, *Trans. Faraday Soc.* 34, 485 (1938), E. S. Rittner, R. A. Hutner, and F. K. dePré, *J. Chem. Phys.* 17, 198 (1949).
4. M. Born and K. Huang, Dynamical Theory of Crystal Lattices (Oxford, 1954).
5. L. Pauling, *Zeit. f. Kryst.* 67, 377 (1928).
6. L. W. Strock, *Zeit. f. Phys. Chem.* 25, 441 (1934).

TABLE I. Parameters used in Eqs. (2) and (3)

<u>ion</u>	<u>r</u>	<u>a</u>
Li ⁺	0.67	0
Na ⁺	1.0	0
K ⁺	1.3	1.3
Rb ⁺	1.5	2.0
Cs ⁺	1.7	3.3
F ⁻	1.3	0.6
Cl ⁻	1.8	3.7
Br ⁻	2.0	4.8
I ⁻	2.2	6.4
Ti ⁺⁺⁺⁺		0
O ⁼		2.4
Al ⁺⁺⁺		0

$$B_{+-} = 1.0$$

$$B_{\uparrow-} = 1.0$$

$$B_{\uparrow\uparrow} = 1.0$$

$$B_{++} = 1.25$$

$$B_{--} = 0.75$$

$$B_{==} = 0.40$$

$$B_{\uparrow\uparrow\uparrow} = 1.60$$

TABLE II. Sample calculation of Ag⁺ ion energy in AgI lattice (see Fig. 1).

>ARPA MATERIALS RESEARCH COUNCIL 1971, MHFPROG9
 >FLYGARE AND HUGGINS CRYSTAL LATTICE PROGRAM, WITH

> ESQ = 23.0000
 > ALPHA = 3.2000
 > LAMBDA = 0.0
 > RHO = 0.3330
 > A = 5.0360
 > B = 5.0360
 > C = 5.0360
 > NX = 10
 > NY = 5
 > NZ = 5
 > XLON = 0.0
 > XHIGH = 2.5170
 > YLOW = 0.0
 > YHIGH = 1.2585
 > ZLOW = 2.5170
 > ZHIGH = 2.5170
 > NMAX = 9
 > IMAX = 9
 > MMAX = 9
 > CB = 1.0000
 > RPLU = 0.9000
 > RMIN = 1.7500

>ENERGY MATRIX NOW FOLLOWS

	X =	0.0	0.252	0.503	0.755	1.007	1.258	1.510	1.762	2.014	2.265	2.517
> Y												
> Z												
> 0.0	2.5170	-1760.81	-1761.27	-1761.68	-1762.02	-1762.25	-1762.34	-1762.26	-1762.03	-1761.69	-1761.27	-1760.81
> 0.2517	2.5170	-1761.27	-1761.74	-1762.15	-1762.47	-1762.66	-1762.70	-1762.57	-1762.23	-1761.85	-1761.37	-1760.88
> 0.5034	2.5170	-1761.68	-1762.15	-1762.52	-1762.81	-1762.91	-1762.79	-1762.63	-1761.84	-1761.13	-1760.43	-1759.85
> 0.7551	2.5170	-1762.02	-1762.47	-1762.81	-1762.96	-1762.96	-1762.37	-1761.66	-1760.14	-1758.65	-1757.33	-1756.53
> 1.0068	2.5170	-1762.26	-1762.66	-1762.91	-1762.96	-1762.32	-1761.07	-1754.92	-1755.96	-1752.62	-1749.81	-1748.42
> 1.2585	2.5170	-1762.34	-1762.70	-1762.79	-1762.37	-1761.07	-1758.37	-1753.81	-1747.40	-1740.15	-1734.12	-1731.51

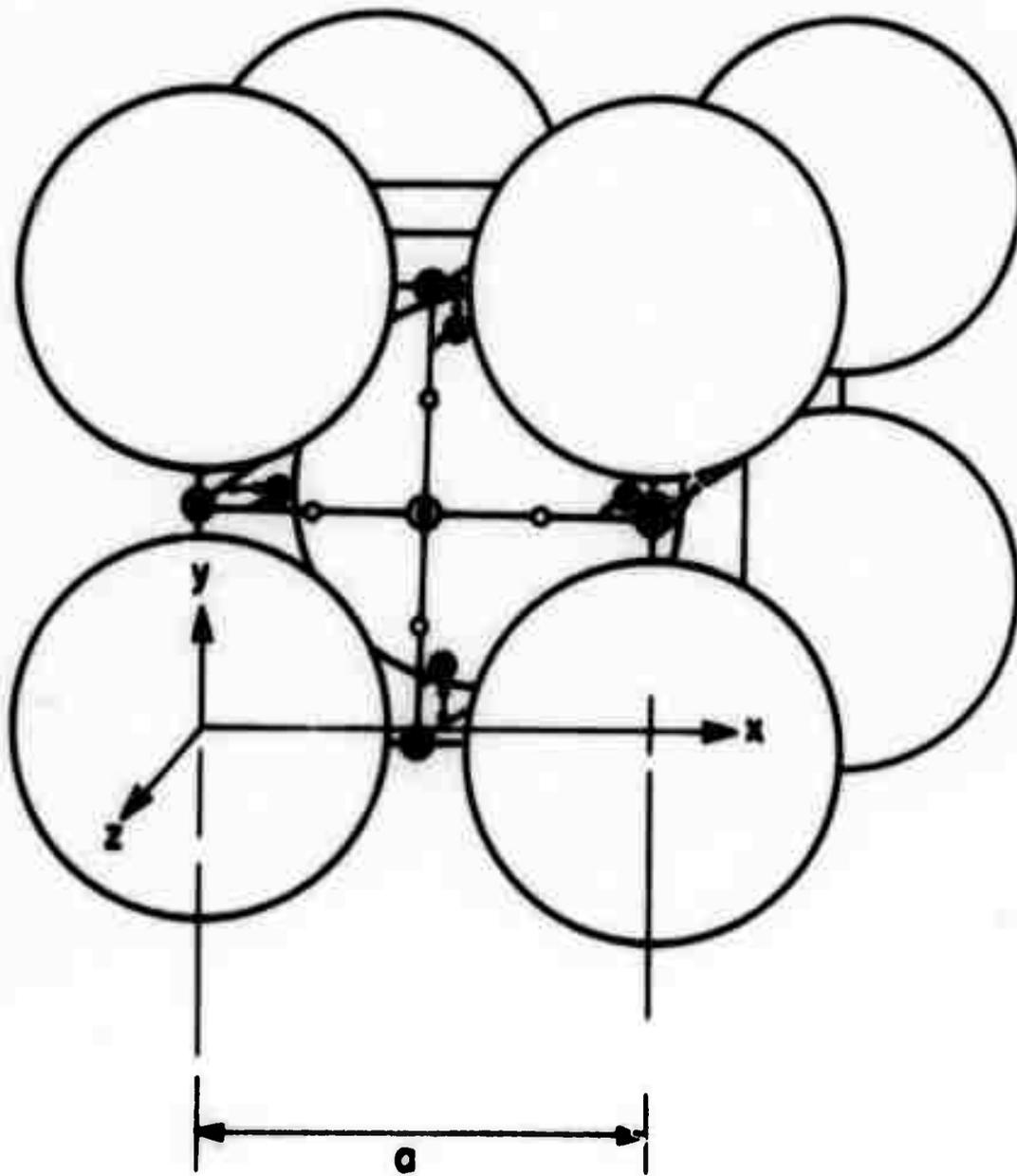
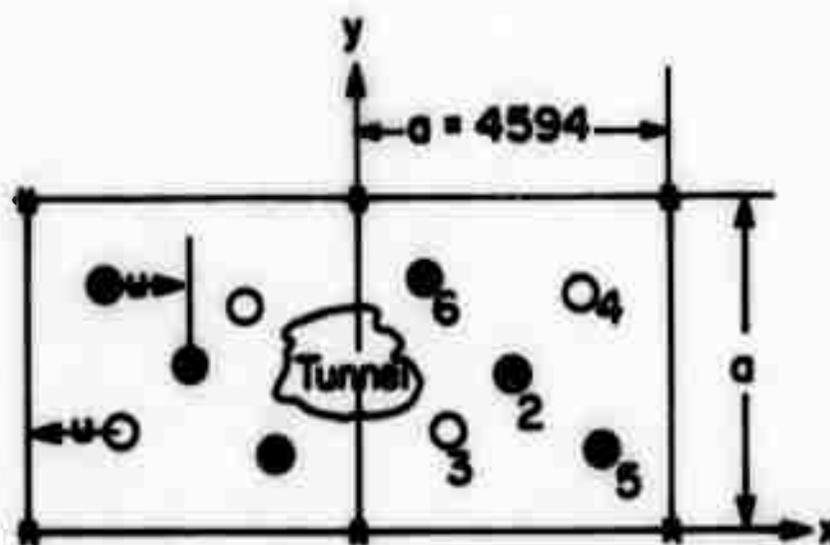


Figure 1. The AgI lattice. The large ions are I⁻ and the small ions represent Ag⁺ ion positions. There are 2 I⁻ ions and 6 , 6 , 12 , and 12  Ag⁺ positions.

$c = 2.958$
 $u = 0.3053$



- * Cation in-plane
- Cation $c/2$ out-of plane
- Anion in-plane
- Anion $c/2$ out-of plane

Figure 2. TiO_2 lattice showing two adjacent unit cells with a tunnel between them.

$Z = 2.5170$

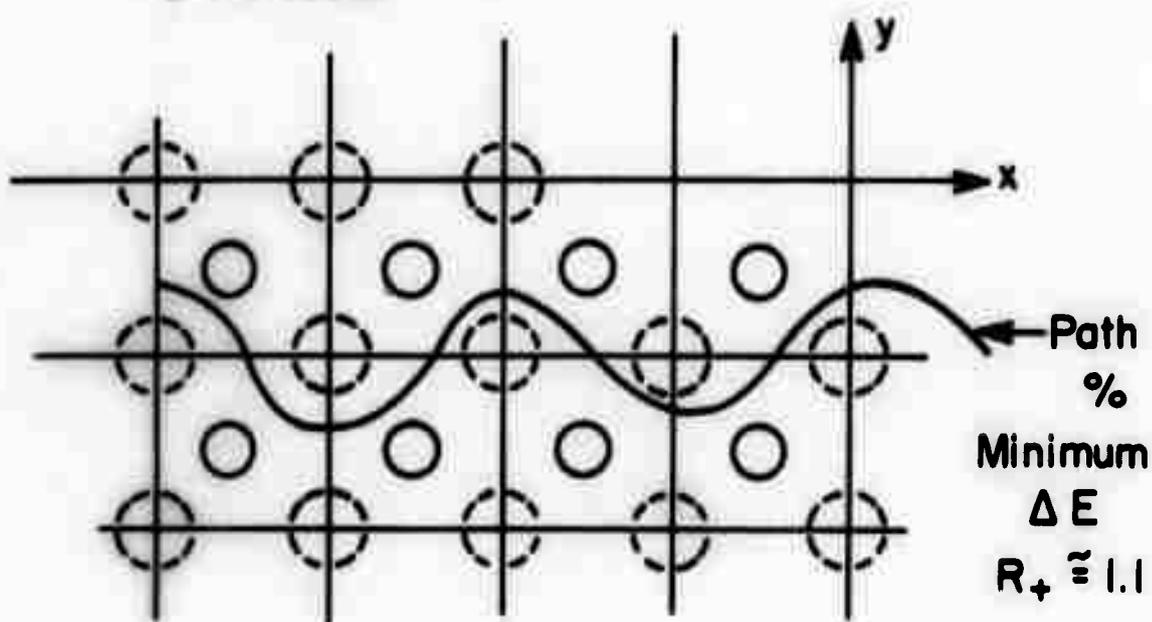


Figure 3. Path of minimum energy in the $z = 2.5170$ xy plane. ΔE is a minimum when $R_+ = 1.1$.

EFFECT OF STRESS ON ELECTROCHEMICAL DISSOLUTION

R. Gomer
R. Thomson

Abstract

Some simple thermodynamic considerations are applied to the dissolution of material from a surface under stress such as obtained at a crack tip. We estimate that the EMF generated at the crack tip in iron is of the order of four millivolts.

EFFECT OF STRESS ON ELECTROCHEMICAL DISSOLUTION

R. Gomer
R. Thomson

In view of the high stresses possible at the tip of cracks in stress corrosion, it seemed worthwhile to make an estimate of the change in (open circuit) oxidation potential of simple dissolution reactions of the type



where n represents the oxidation state of the ion. We consider for simplicity uniaxial stress only at this point, although the stresses at crack tips will generally be more complicated.

Consider the change in Helmholtz free energy ΔF per gram-atom of metal M when a uniaxial stress σ_z is applied in the z direction, as shown in Figure 1.

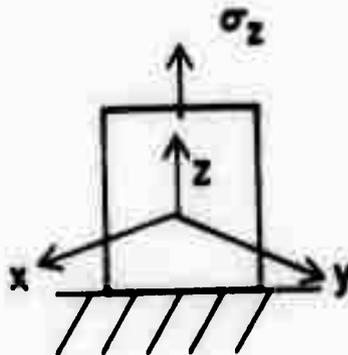


Figure 1.

$$\Delta F = \bar{V}_0 \int \sigma_z d\epsilon = \bar{V}_0 k_z \epsilon^2 / 2 \quad (2)$$

where k_z is Young's modulus and $\epsilon = \Delta z/z$ is strain and \bar{V}_0 the gram atomic volume at $\sigma = 0$. ΔF may be interpreted as the work done on a gram atom of M when it is transferred reversibly, say by vaporization, from a block of unstrained material, to a different block, and the latter then reversibly strained. Since F is a state function the same ΔF must result when this process is carried out (reversibly) by means of an electrochemical cell, consisting of two blocks of M, strained and unstrained, immersed in a solution containing M^{+n} ions. ΔF will again be the reversible work done on the system when a gram-atom of M is transferred electrochemically from $M(\epsilon=0)$ to $M(\epsilon)$. If the deposition takes place on the xz or yz planes of the stressed electrode, no mechanical work is done, and

$$\Delta F = n F E_{||} \quad (\text{joules}) \quad (3)$$

where $F = 96,500$ coulombs and $E_{||}$ is the cell emf in volts. Consequently

$$E_{||} = \frac{\bar{V}_0 k_z \epsilon^2}{2nF} \quad \text{volts} \quad (4)$$

with $k_z \epsilon^2$ in joules. If electrodeposition occurs on the xy planes however, the total reversible work consists of electrical plus mechanical work, done on the system, the latter being

$$W_m = \pm \frac{\bar{V}}{A_{xy}} \quad f_z = \pm \bar{V} \sigma_z \quad (5)$$

where A_{xy} is the area of the xy plane and f_z the force in the z direction; \bar{V} is the gram-atomic volume (at stress σ_z) in first approximation equal to \bar{V}_0 for solids or liquids. The plus sign applies to tension, the minus sign to compression, since in the latter case the addition of matter to $M(\epsilon)$ requires the system to do work on the surroundings and our convention here is to take ΔW negative for this case. Since the change in F must still equal that given by Eq. (2) we have

$$E_{\pm} = \frac{\bar{V}_0}{nF} [kc^2/2 \pm (\bar{V}/\bar{V}_0)\sigma_z] = \frac{V_0 ck}{nF} [c/2 \pm 1] \quad (6)$$

where the minus sign refers to tension, the plus sign to compression. The subscripts on E in Eqs. (4) and (6) refer to deposition (or dissolution) parallel and orthogonal to the applied stress, and a positive sign implies spontaneous dissolution of the stressed material, i.e., the latter being the negative electrode of our hypothetical cell.

We shall assume that Eq. (4) applies to the crack tip. If the latter is very sharp, the effect of surface tension must be taken into account. The addition of an atom of M to a small particle under stress now requires an amount of work

$$W = \Omega kc^2/2 + \left(\frac{\partial A}{\partial n}\right) [\gamma_0 + k_s \epsilon_s^2/2] \quad (7)$$

where Ω is the atomic volume, γ_0 the surface tension at zero stress, k_s and ϵ_s surface Young's modulus and surface stress,

respectively, and

$$\frac{\partial A}{\partial n} = \frac{2\Omega}{r} \quad (8)$$

with r the effective radius of the particle. The expression in brackets in Eq. (7) can also be considered the surface tension at bulk stress σ_z , and we so consider it. The analogue of Eq. (4) now becomes

$$E_{||} = \frac{\bar{V}}{nF} [k\epsilon^2/2 + 2\gamma/r] \quad (9)$$

If r is negative as at the tip of a crack, $E_{||}$ is reduced. This is intuitively clear, since the effect of surface tension would be to force the crack shut (i.e., increase its radius). Thus there is a minimum crack radius for electrochemical dissolution (relative to unstrained identical material) given by

$$r = 4\gamma/k\epsilon^2 \quad (10)$$

Eq. (10) is somewhat misleading of course, in that γ is itself a function of ϵ . In a very crude way we may assume that γ is given by the number of "dangling bonds" per unit area and that these decrease linearly with the dilation. Then, for a stress applied along the z direction, with unconstrained sides

$$\gamma = \gamma_0 \left(\frac{1}{1 + .7\epsilon_z} \right) \quad (11)$$

where we have assumed a Poisson's ratio of 0.3.

With the sides constrained, we would have

$$\gamma \approx \gamma_0 (1 - \epsilon_z) \quad (12)$$

Since $\epsilon < .2$ in any case, it is improbable that the effect on γ is very great.

Numerical Estimates

To get some idea of the magnitude of the emf caused by strain we make the following assumptions, appropriate to Fe:

$$\bar{V} = 7.13 \text{ cm}^3$$

$$k = 2.4 \cdot 10^{12} \text{ erg/cm}^3 = 2.4 \cdot 10^5 \text{ joule/cm}^3$$

$$\epsilon = 0.1$$

$$\gamma = 2000/\text{erg/cm}^2 = 2 \cdot 10^5 \text{ joule/cm}^2$$

It should be noted that the value of γ may be substantially reduced by the presence of even a single atom layer of adsorbate, in view of the Gibbs adsorption equation. We then find

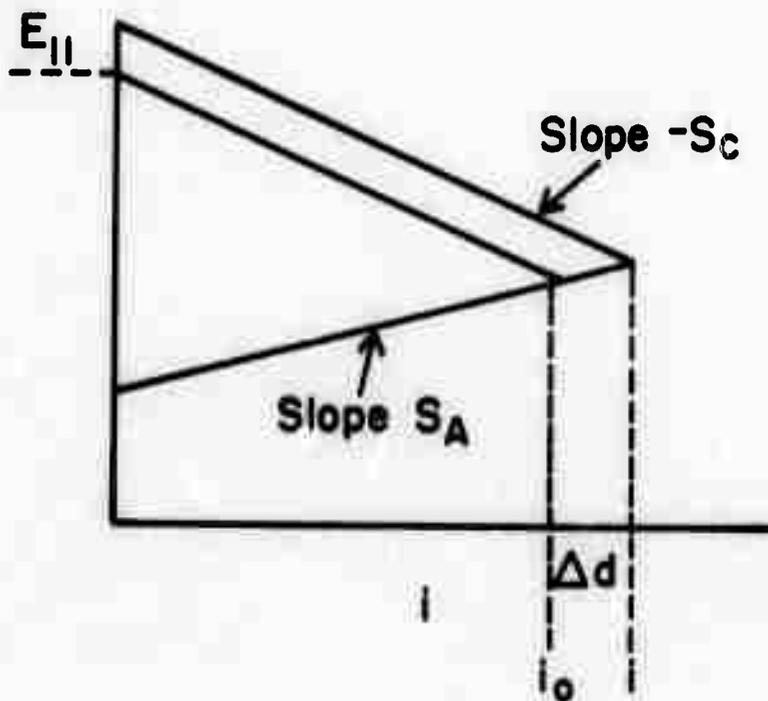
$$E_{||} \approx 3.7 \times 10^{-5} [125 - 4 \cdot 10^{-4}/r] \quad (13)$$

so that the critical radius is $\sim 30 \text{ \AA}$. For $r > 100 \text{ \AA}$ (10^{-6} cm) the surface energy term becomes unimportant and

$$E_{||} \approx 3.7 \text{ millivolt} \quad (14)$$

This value is of course extremely small and suggests that preferential electrochemical dissolution at the crack tip due

to strain is negligible. A quantitative answer can be found as follows. Since the stressed region is electrically short-circuited to the unstressed parts of the metal, the current densities i and i_0 of the dissolution reaction at the stressed and unstressed regions can be found from an Evans diagram:



whence

$$i - i_0 = \Delta i = E_{||} (S_A - S_C) \quad (15)$$

with $E_{||}$ given by Eq. (9) viz. (13).

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

AMORPHOUS METALLIC ALLOYS

P. E. Duwez

Abstract

During the last ten years, a relatively large number of amorphous alloys with metallic properties have been obtained by very rapid cooling from the liquid state. This paper presents a brief review of the chemical compositions of these alloys and the factors which appear to be important in obtaining the amorphous state after quenching from the melt. The unusual electrical and magnetic properties of these alloys are discussed. There is some strong indication that these properties are unaffected by rather high doses of radiation. References are given to 52 papers published on the subject before June 1971.

Preceding page blank

AMORPHOUS METALLIC ALLOYS

P. E. Duwez

I. Introduction

Metallic alloys have been obtained in the amorphous state by three different methods; vapor deposition, electrolytic or electroless deposition and quenching from the liquid state. Vapor deposited copper-silver (Mader, et al., 1967) and gold-cobalt amorphous alloys (Mader and Nowick, 1965) have been described in the literature. Amorphous electrolytic deposits have been obtained in binary nickel-phosphorus and cobalt-phosphorus alloys. The same technique could probably be applied to palladium-phosphorus alloys and to ternary and quaternary alloys containing Ni, Co, Pd and about 15 to 20 at.% of phosphorus. Until now, vapor deposition and electrolysis have lead to a relatively small number of amorphous alloys, but an increased effort in this field is likely to uncover additional interesting thin films with unusual properties. The technique of rapidly quenching a liquid alloy has received more attention and this brief review will be limited to amorphous alloys obtained by this technique.

II. Liquid Quenching Techniques

Various techniques for rapid quenching from the liquid state have been described in the literature. (Duwez and Willens, 1963; Pietrokovsky, 1963; Willens and Buehler, 1966; Duwez, 1968;

Pond and Maddin, 1969; Chen and Miller, 1970) All these techniques are based on the same principle, namely cooling of a thin layer of liquid by conduction on a substrate made of a good heat conductor, such as copper. The rate of cooling required for producing amorphous alloys is between 10^3 and 10^6 °C/sec. For obvious reasons, the high rates of cooling can be obtained only if the specimen is rather thin, generally between a few μm and about $50\mu\text{m}$.

III. Experimental Evidence for the Amorphous State

The question of amorphous versus microcrystalline structure is still a controversial one. The amorphous state does not mean complete disorder in the atomic arrangement and local or short range order is always present. If the extent of this short range order is comparable to that found in the liquid state, the term amorphous solid is justified.

The methods used to define the atomic arrangement by means of an atomic distribution function are based on diffraction by either X-rays, electrons or neutrons. Since neutron diffraction experiments require about 1 cm^3 of material, it is impractical, because it would involve a very large number of quenched foils. So far no neutron diffraction experiments have been reported on amorphous metallic alloys. Electron diffraction techniques are particularly useful in the case of vapor deposited thin films, which are too thin for X-ray diffraction. The ordinary electron diffraction patterns obtained on a photographic film, however, are not suitable for

obtaining a precise radial distribution function and rather complicated energy analysers must be used. Hence, X-ray diffraction remains the most practical method for studying the atomic arrangement in amorphous alloys.

Many papers have been published on the analysis of X-ray diffraction data obtained on amorphous solids. A recent study concerned mostly with liquid quenched alloys (Giessen and Wagner, 1971) gives a summary of most of the radial distribution functions of binary and ternary amorphous alloys recently published. It is interesting to note that all radial distribution functions differ only by details of secondary importance. A direct comparison between the radial distribution function of an amorphous alloy and that of the same alloy in the liquid state (deduced from measurements made with the same X-ray diffractometer) would be the most convincing proof of the similarity between atomic arrangements in both the amorphous and the liquid state. So far, such a comparison exists only between an amorphous $\text{Au}_{30}\text{-Si}_{20}$ alloy and liquid gold (Dixmier and Guinier, 1967). The two radial distribution functions are slightly different, but it is not known how much the presence of Si in the alloy contributes to this difference.

The question of amorphous versus microcrystalline structure has been reviewed recently by Dixmier and Guinier (1971). A study was made of the structure of evaporated platinum on an amorphous carbon film. This evaporated platinum is generally considered as amorphous because its interference

function is apparently amorphous, but a detailed study of its structure could be explained only by assuming that it consisted of close-packed hexagonal microcrystals containing about 13 atoms. In contrast with these results, the interference function of an electro-deposited amorphous nickel-phosphorus alloy could not be explained by assuming a microcrystalline structure. This study, as well as that of Cargill (1970), demonstrates that in the case of amorphous metallic alloys, a very careful analysis of X-ray diffraction data taken under the best experimental conditions can indeed differentiate between amorphous and microcrystalline solids.

The lack of contrast in transmission electron microscopy has been often considered as a proof of the amorphous nature of a solid. Since the resolving power of a modern electron microscope is better than 10 \AA , microcrystals ranging from 7 to 17 \AA in size (which is the range observed for a relatively large number of liquid quenched amorphous alloys, Sinha and Duwez, 1970) should be clearly detected in transmission electron-microscopy. The problem, however, is not as simple as it looks. It was pointed out recently (Giessen and Wagner, 1971) that resolving crystalline particles less than about 10 \AA in size can be achieved only if these particles do not overlap and these authors conclude, "It has been pointed out by Coslett (1966) that the resolution found in amorphous carbonaceous films is probably of the order of 10% of the film thickness. This would mean that for film thickness of less than 1000 \AA

used in transmission electronmicroscopy of metallic films, the relatively poor resolution for amorphous alloys would not allow the observation of particles or grains of sizes as small as those deduced from their diffraction patterns which show only a few diffuse halos."

Transmission electronmicroscopy is probably the best technique for studying the early stages of crystallization of amorphous alloys. (Willens, 1962; Crewdson, 1966; Rastogi and Duwez, 1970). The studies reported so far, however, are only qualitative in nature and deal with the shape and size of the crystalline phase, or phases, growing out of the amorphous matrix by heating the specimen in the microscope. Quantitative studies of the kinetics of nucleation and growth of the crystalline phases are difficult because of the present limitations in measuring the actual temperature of the small region of the specimen under observation.

Perhaps the most convincing proof of the existence of the glassy state is the measurement of a glass transition temperature T_g . This temperature can be detected by measuring the specific heat which shows a sharp increase around T_g . Such measurements have been performed on amorphous Au-Si (Chen and Turnbull, 1967) and Pd-Si alloys (Chen and Turnbull, 1969). In both cases, T_g is very close to the temperature at which crystallization of the amorphous alloy is very rapid. For $Au_{77}-Ge_{14}-Si_9$, for example, $T_g = 285-290^\circ K$ and crystallization temperature is about $297^\circ K$. Another experimental fact confirming

the amorphous nature of quenched Au₇₇-Ge₁₃-Si, is a viscous flow observed in the vicinity of T_g. The viscosity varied from 10¹¹ to 10⁹ poise in the temperature interval between 285 and 305°K.

IV. Composition of Amorphous Alloys

A list of amorphous alloys mostly obtained by liquid quenching is given in Table I which includes references. The amorphous alloys which have received the most attention are Au-Si and Pd-Si. Electrolytic and vapor deposited Ni-P alloys have also been extensively studied.

After the first amorphous phase in Au-Si alloys was obtained by quenching from the liquid state (Klement, et al, 1960), it was pointed out (Cohen and Turnbull, 1961) that the composition of this alloy was close to that corresponding to a eutectic which has a very low melting point (370°C) compared with that of gold (1063°C). This observation served as a guide to discover other binary alloys likely to be amorphous after quenching. Pd-Si binary alloys, with a eutectic at about 16 at. % Si and a eutectic temperature of about 800°C, compared with a melting temperature of 1550°C for Pd, turned out to be the easiest to quench by both the gun and the piston and anvil technique and have been the subject of numerous investigations. (Duwez, et al., 1965) The Pd-Si amorphous alloys are quite stable, with a rapid crystallization temperature of about 420°C. Kinetics data indicate that no crystallization can be detected after more than 6 months at 200°C (Duwez, 1967).

Given two elements in the periodic table, it is not yet possible to predict with certainty whether or not they will lead to an amorphous alloy after quenching. The eutectic condition mentioned above seems to be satisfied for all amorphous alloys reported so far. When a survey is made of binary alloys having low eutectics (in the sense of Au-Si alloys) it is found that these occur mostly between transition metals, or Cu, Ag and Au, and semi-metals of valence 3, 4 or 5. These include B, C, Si, Ge, As and P. In many of these low eutectic binary alloys the quenched foils often consist of an amorphous phase in which a small number of microcrystals are imbedded. In these cases the X-ray diffraction patterns consist of a series of broad halos typical of the amorphous state to which are superimposed a small number of very weak but relatively sharp Bragg's crystalline peaks. This type of mixed structure can be confirmed by transmission electromicroscopy. When two of these almost completely amorphous alloys are taken to form a pseudo-binary system, the ternary alloys can be obtained very often in the amorphous state. This is the case for the ternary alloys Pd-Ni-P, Pd-Fe-P, Pt-Ni-P, Pd-Ni-B, Fe-P-C, and Mn-P-C listed in Table I.

Several models have been proposed to describe the atomic arrangement in amorphous alloys. For alloys containing a transition metal and a semi-metal an attractive model is based on a Bernal's type random packing of the transition metal atoms with the metalloid atoms occupying the largest voids in

the structure. In $\text{Pd}_{80}\text{-Si}_{20}$ for example, the Si atoms could fit into the Bernal's voids if their size is approximately that of Si in the crystalline silicide Pd_3Si . This model also accounts for the fact that in all amorphous alloys between transition metals and semi-metals, the concentration of the semi-metal ranges from 15 to about 27 at.% and does not exceed the theoretical number of large voids in the model.

Microcrystalline alloys have been reported in Nb-Ni and Ta-Ni systems (Ruhl, et al., 1967) quenched from the liquid state. The compositions of these alloys are in the vicinity of eutectics in the equilibrium diagrams. These alloys might be amorphous, but neither a detailed analysis of the radial distribution function nor a measurement of their glass transition temperatures have been reported. If they were amorphous they would not satisfy the above mentioned criterion requiring the presence of a semi-metal as a constituent of the alloys. Further experimental studies are required to establish the true amorphous state of alloys containing transition metals only.

V. Properties of Amorphous Metallic Phases

Amorphous metallic alloys are likely to have unusual properties because of their unusual structure. In general, the electrical resistivity of an amorphous alloy has a very small temperature coefficient between 1.5°C and the temperature at which crystallization takes place. This small coefficient results because most of the scattering is due to the structural atomic disorder and the contribution of the thermal scattering

is very small. When a transition metal with a magnetic moment such as chromium is substituted for Pd in a Pd-Si alloy, a minimum in the resistivity vs. temperature curve is observed (Tsuei, et al., 1969). The temperature at which the minimum occurs increases with increasing amounts of Cr and is about 200°C for 7 at.% which is about the maximum concentration above which the amorphous state cannot be obtained. Between the minimum temperature and 1.6°K, the resistivity increases steadily with decreasing temperature and this alloy constitutes a very sensitive resistance thermometer for low temperatures. (Tsuei and Duwez, 1970).

The existence of a minimum in the electrical resistivity-temperature curve of amorphous alloys containing a magnetic atom is related to the Kondo effect observed in crystalline alloys. One major difference, however, is that the Kondo effect in crystalline alloys is generally found only for low concentrations of the magnetic atom, and the temperature at which the minimum occurs is very low. In amorphous alloys the magnetic impurity content may be as high as 12 at.%. Existing theories based on the hypothesis of a dilute impurity concentration are not applicable and the problem is an important one in theoretical solid state physics.

Although the existence of ferromagnetism in non-crystalline solids has been predicted by Gubanov (1960) no such ferromagnets were found until recently. Several amorphous alloys obtained by quenching from the liquid state are ferromagnetic. In

particular the $\text{Fe}_{75}\text{-P}_{15}\text{-C}_{10}$ alloy (Duwez and Lin, 1967) has a saturation magnetization above 12,000 G, a cohesive force of 3 Oe and a Curie temperature of 320°C. Many questions are still unanswered on the exact nature of ferromagnetism in non-periodic structures. It is not yet clearly established whether or not domain boundaries exist in amorphous ferromagnets, and if so, what is their origin, their width and their mobility. These new alloys present a challenge to experimental and theoretical solid state physicists.

Little attention has been given so far to the mechanical properties of metallic glasses. As expected, these materials exhibit very small plastic flow in tension. The following properties have been recently reported for a $\text{Pd}_{80}\text{-Si}_{20}$ amorphous alloy (Masumoto and Madrin, 1971); yield strength 87.2 kg/mm², fracture strength 136.4 kg/mm², elongation 0.11%, Young's modulus, 6.8×10^3 kg/mm². Slightly higher elongation values were found in a $\text{Pd}_{80.5}\text{-Si}_{19.5}$ alloy in which some Pd was replaced by 14 at.% of U, Ag or Cu (Chen and Wang, 1971). The elongation in this case was between 1 and 3%. In spite of this apparent brittleness in a tension test, $\text{Pd}_{80}\text{-Si}_{20}$ glassy alloys have been rolled at room temperature from 40 down to about 13μm thick without cracking (Duwez, 1970, unpublished). No change in the X-ray diffraction pattern nor in the electrical resistivity could be detected after rolling.

Only one study of the effect of radiation on the structure and properties of a glassy alloy has been reported (Lesueur, 1968).

A foil of $\text{Pd}_{80}\text{-Si}_{20}$ amorphous alloy obtained by the piston and anvil technique was covered with a layer of U^{235} and subjected to a neutron flux in a reactor. The damage due to fission products was monitored by both X-ray diffraction and electrical resistivity. After having been subjected to a total dose of 5.6 fission products/ cm^2 , no detectable changes were observed in either the diffraction pattern or the electrical resistivity measured at both room and liquid nitrogen temperature. An interesting experiment, also performed by Lesuer, was to expose a specimen of the same alloy $\text{Pd}_{80}\text{-Si}_{20}$ in the crystalline equilibrium state to the same accumulated dose of fission products. In this case, the intensity of the Bragg diffraction peaks progressively decreased and the intensity of the halos typical of the amorphous phase increased as irradiation proceeded. The electrical resistivity of this sample increased smoothly and was approaching that of the amorphous state at the final dose of 5.6×10^4 fission products/ cm^2 .

VI. Conclusions

Amorphous metallic alloys constitute a relatively new class of solids. At the present time one of the reasons for studying these materials is their possible contribution to the understanding of the metallic state in disordered solids and the relationship between such solids and the liquid state. Very few practical applications have been suggested for this new class of materials (Tsuei and Duwez, 1971). Their unusual electrical properties, and also perhaps magnetic properties

combined with their resistance to radiation damage (Lesueur, 1968) are likely to be exploited in special purpose electronic devices.

Some of the important problems related to amorphous metallic alloys are:

1. Structural investigations of the atomic arrangement in the alloys. It is already established that most of the alloys reported so far are really amorphous (in the sense of liquid-like) and not microcrystalline. This has been deduced from accurate studies of the radial distribution function obtained from X-ray diffraction data, and comparing these results with those obtained on liquid alloys. A perhaps more definite proof of the amorphous or glassy state is the measurement of a glass transition temperature deduced from specific heat determination or viscous flow measurements. The now widely accepted model for the structure of amorphous alloys is a Bernal's packing of the transition metal atom with the semi-metal located in the larger voids of this packing. Additional work is needed in this area to establish this model on firmer grounds.
2. What are the alloy compositions susceptible to be quenched into an amorphous solid by quenching from the liquid state is an important and still unsolved problem. The empirical rules proposed are that the main element in the alloy is a transition metal forming a low temperature eutectic with a semi-metal of valence 3, 4 or 5. Amorphous alloys are generally found around the eutectic composition and the semi-metal concentration is

from about 15 to 27 at.%. This concentration range is compatible with the number of large voids in the Bernal's packing of hard spheres. Theoretical work in this field is required.

3. The electrical resistivity of amorphous metallic alloys is temperature insensitive down to 2°K or lower. This is qualitatively understood since most of the scattering is due to the high degree of disorder and the relatively small contribution of phonon scattering. What is not understood is that the temperature coefficient of resistivity can be changed from positive to negative by, for example, changing the Pt to Ni ratio in a ternary alloy Pt-Ni-P (Sinha, 1970) or by changing the phosphorus content in a Pd-Ni-P alloy and keeping the Pd to Ni ratio constant (Boucher, 1971). This problem requires an analysis of the band structure in amorphous metallic alloys. A minimum in the electrical resistivity vs. temperature curves of amorphous alloys containing magnetic elements is similar to a Kondo effect in crystalline alloys. However, in amorphous alloys this minimum persists as the concentration of the magnetic element is increased up to 12 at.% and may occur at temperatures as high as 200°C. Amorphous alloys constitute an interesting class for the study of the resistivity minimum problem in alloys containing a relatively high concentration of magnetic elements for which the d-d interaction cannot be neglected. More experimental and theoretical work is needed in this field.

4. The existence of ferromagnetism in amorphous alloys has been demonstrated, but few measurements have been made of the

magnetic properties. This field is of fundamental interest and the possibilities for practical applications are still unexplored.

5. Interest in the mechanical properties of amorphous alloys is just beginning. Preliminary results, however, already bring up puzzling questions. While an amorphous alloy appears to lead to a brittle fracture in tension, it can be rolled to about 1/3 of its thickness at room temperature, which is several hundred degrees below the glass transition temperature. The origin of this apparent plasticity (or viscous flow) should be a rewarding field of study.

6. Preliminary experiments have shown that amorphous metals are insensitive to heavy doses of radiation. More accurate experiments should be performed to detect the exact nature of the effect of heavy particles (especially fission products) on the atomic configuration of the glassy structure. It has also been shown that when a crystalline alloy such as $\text{Pd}_{80}\text{-Si}_{20}$ which can be obtained in the amorphous state by quenching from the liquid state is bombarded by fission products, it becomes amorphous. This raises the question of knowing if an alloy susceptible to become amorphous by liquid quenching is also susceptible to becoming amorphous by radiation damage. If so, what are the reasons for this behavior?

In summary, a large number of amorphous alloys covering a wide range of compositions have been obtained during the last ten years. Classical solid state physics of metals is based

on the existence of a periodic lattice. Amorphous alloys open the door to an extension of the classical concept on disordered systems, which so far was limited to liquid metallic alloys.

Acknowledgment

This study was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253.

TABLE I
Amorphous Alloys

<u>Alloys</u>	<u>Range</u>	<u>Reference</u>
$Au_{100-x}Si_x$	-	Klement et al. (1960) Dixmier and Guinier (1967) Chen and Turnbull (1967)
$Au_{77}Si_{14}Ge_9$	-	Bagley and Turnbull (1968) Chen and Turnbull (1970)
$Pd_{100-x}Si_x$	$16 < x < 22$	Duwez et al. (1965) Crewdson (1966) Chen and Turnbull (1969)
$Ni_{100-x}P_x$ (Electroless) (Electrolytic) (Vapor Deposited)	$15 < x < 20$ $15 < x < 24$ -	Dixmier et al. (1969) Cargill (1970) Bagley and Turnbull (1970)
$(Pd_{100-x}T_x)_{100}Si_x$ T=Cr, Mn, Fe, Co or Ni	$x < 10$	Tsuei and Duwez (1966) Tsuei and Hasegawa (1969) Hasegawa (1970)
$(Pd_{100-x}Ni_x)P_{20}$	$15 < x < 90$	Maitrepierre (1969) Maitrepierre (1970)
$(Pd_{100-x}Fe_x)P_{20}$	$17 < x < 55$	Maitrepierre (1969)
$(Pt_{100-x}Ni_x)P_{25}$	$20 < x < 60$	Sinha (1970) Sinha and Duwez (1970)
$(Pd_{50}Ni_{50})B_{10}$	-	Liang and Tsuei (1971)
$(Pd_{50}Ni_{50})_{100-x}P_x$	$16 < x < 26.5$	Dixmier (to be published)
$Fe_{75}P_{15}C_{10}$ to $Fe_{100}P_{15}C_7$	-	Lin and Duwez (1969) Lin (1969a) Lin (1969b) Rastogi and Duwez (1970)
$Mn_{75}P_{15}C_{10}$	-	Hasegawa (1971)
$(Fe_{100-x}Mn_x)_{75}P_{15}C_{10}$	$0 < x < 100$	Sinha (1971)

References

- Bagley, B. G. and Turnbull, D., (1968) *J. Appl. Phys.*, 39, 5681.
- Bagley, B. G. and Turnbull, D., (1970) *Acta Met.* 18, 857.
- Boucher, B., (1971) To be published.
- Cargill, G. S. III, (1970) *J. Appl. Phys.* 41, 12.
- Cargill, G. S. III, (1970) *J. Appl. Phys.* 41, 2248.
- Chen, H. S. and Turnbull, D., (1967) *J. Appl. Phys.*, 38, 3646.
- Chen, H. S. and Turnbull, D., (1968) *J. Chem. Phys.* 48, 2560.
- Chen, H. S. and Turnbull, D., (1969) *Acta Met.*, 17, 1021.
- Chen, H. S. and Miller, C. E., (1970) *Rev. Sci. Instr.*, 41, 1237.
- Chen, H. S. and Turnbull, D., (1970) *Acta Met.* 18, 261.
- Chen, H. S. and Wang, T. T., (1971) *J. Appl. Phys.* 41, 5338.
- Cohen, M. H. and Turnbull, D., (1961) *Nature* 189, 131.
- Coslett, W. E., (1956) *Brit. J. Appl. Phys.*, 7, 10.
- Crowdson, R. C. (1966) Ph.D. Thesis, California Institute of Technology
- Dixmier, J. and Guinier, A., (1967) *Rev. Metallurgie*, 64, 53.
- Dixmier, J. and Guinier, A., (1970) *Fizika*, 2, Supplement 2, 9.1.
- Doi, K. and Guinier, A., (1969) Physics of Non-Crystalline Solids, Ed. J. A. Prins, North Holland Publishing Co.
- Duwez, Pol and Willens, R. H., (1963) *Trans. AIME*, 227, 362.
- Duwez, Pol, Willens, R. H. and Crowdson, R. C., (1965) *J. Appl. Phys.*, 36, 2267.
- Duwez, Pol, (1967) in Phase Stability in Metals and Alloys, p. 523. Eds. P. S. Rudman, J. Stringer and R. I. Jaffee, McGraw-Hill.
- Duwez, Pol and Lin, S. C. H., (1967) *J. Appl. Phys.*, 38, 4096.
- Duwez, Pol, (1968) Techniques in Metals Research, Part I, Ed. Bunshah, p. 347.

- Giessen, B. C. and Wagner, C. N. J., (1971) "Structure and Properties of Non-Crystalline Metallic Alloys Produced by Rapid Quenching of Liquid Alloys" to be published in Liquid Metals, S. Beer, Ed., Marcel Dekker, New York.
- Graczyk, J. F. and Moss, S. C., (1969) Rev. Sci. Instr., 40, 424.
- Gubomov, A. (1960) Soviet Physics, Solid State, 2, 468.
- Hasegawa, R., (1970) J. Appl. Phys., 41, 4096.
- Hasegawa, R., (1971) Phys. Rev. B, 3, 1631.
- Klement, W. Jr., Willens, R. H. and Duwez, Pol, (1960) Nature, 187, 869.
- Lesueur, D., (1968) C. R. Ac. Sci. Paris, 226B, 1038.
- Liang, V. K. C. and Tsuei, C. C., (1971) Sol. St. Comm., 9, 579.
- Lin, S. C. H. and Duwez, Pol, (1969) Phys. St. Solidi, 34, 469.
- Lin, S. C. H., (1969a) J. Appl. Phys., 40, 2173.
- Lin, S. C. H., (1969b) J. Appl. Phys., 40, 2175.
- Mader, S. and Nowick, A. S., (1965) Appl. Phys. Letters, 7, 57.
- Mader, S., Nowick, A. S. and Widmer, H., (1967) Acta Met. 15, 203.
- Masumoto, T. and Maddin, R. H. (1971) Acta Met. 19, 725.
- Maitrepierre, P. L., (1969) J. Appl. Phys., 40, 4826.
- Maitrepierre, P. L., (1970) J. Appl. Phys., 41, 498.
- Pietrokowsky, P., (1963) Rev. Sci. Instr., 34, 445.
- Polk, D. E., (1970) Scripta Metallurgica 4, 117.
- Pond, R., Jr. and Maddin, R., (1969) Trans. AIME, 245 2475.
- Rastogi, P. K. and Duwez, Pol, (1970) J. Non-Cryst. Solids, 5, 1.
- Ray, R., Giessen, B. C. and Grant, N. J., (1968) Scripta Metallurgica 2, 357.
- Ruhl, R., Giessen, B. C., Cohen, M. and Grant, N. J. (1967) Acta Met. 15, 1693.

- Sinha, A. K., (1970) Phys. Rev. B, 1, 4541.
- Sinha, A. K., (1971) J. Appl. Phys., 42, 338.
- Sinha, A. K. and Duwez, Pol, (1971) J. Phys. Chem. Solids, 32, 267.
- Tsuei, C. C. and Duwez, Pol, (1966) J. Appl. Phys. 40, 435.
- Tsuei, C. C., Longworth, G. and Lin, S. C. H., (1968) Phys. Rev. B, 170, 775.
- Tsuei, C. C. and Hasegawa, R., (1969) Sol. St. Comm. 7, 1581.
- Tsuei, C. C. and Duwez, Pol, (1970) J. of Physics, E. 4, 466.

DIFFUSION THROUGH ANISOTROPIC POLYMER SYSTEMS

John D. Ferry

Abstract

An experiment is proposed to provide information about the effects of large deformations on the local molecular mobility in rubbery polymers, and the mechanism of translatory motion of foreign molecules, either small (molecular weight about 200) or polymeric (10^4 - 10^5). Diffusion of a radioactively tagged penetrant from a point source in a stretched strip is monitored by autoradiography, giving the ratio of diffusion coefficients in the two principal directions. The absolute coefficient in the stretch direction is obtained by diffusion from a line source, and compared with that for the isotropic polymer. Free volume fluctuation theory can be used to provide a coefficient which measures the relative ease of penetration in different directions, probably influenced also by the shape and flexibility of the penetrant molecule. Polymeric penetrants may reveal additional effects associated with entanglement coupling, related also to the density of cross-links in the polymer matrix.

DIFFUSION THROUGH ANISOTROPIC POLYMER SYSTEMS

John D. Ferry

I. Introduction

The diffusion coefficients of small molecules (i.e., molecular weights of the order of 200) through rubbery polymers can be measured by use of radioactively tagged penetrants in trace amounts.¹⁻³ They reflect the local mobilities of polymer segments and can be closely correlated with viscoelastic relaxation times. The dependence of diffusion coefficient D , and of the friction coefficient $\zeta_1 = kT/D$, on temperature, chemical composition of polymer, and concentration of diluent can be described in terms of the effects of these variables on the fractional free volume.⁴ A central parameter in the free volume treatment is a dimensionless coefficient B_d which may be qualitatively interpreted as the ratio of the volume opening required for translatory motion of a penetrant molecule (or mobile segment thereof) to the volume of a mobile polymer segment. This parameter is of the order of unity but depends somewhat on the shape and size of the penetrant molecule.⁵

There is very little known about local molecular mobility in rubbery polymers in highly strained states where the molecules are substantially oriented but still in rapid thermal motion. Measurements of diffusion anisotropy in strained polymers are a potential source of such information and also

perhaps of understanding the mechanism of translatory motion of small foreign molecules. Some preliminary experiments⁶ failed to find any difference between the diffusion coefficients of n-hexadecane and 1,1-diphenyl ethane through biaxially stretched rubbers (perpendicular to the stretch direction) and through the same rubbers in the isotropic state. However, these polymers were far above their glass transition temperatures (T_g), and it can now be anticipated that D will be more sensitive to anisotropy at temperatures only slightly above T_g (see below). By use of uniaxial rather than biaxial extension, D can be conveniently measured in two directions.

II. Proposed Experiment

A strip of cross-linked rubber is subjected to simple extension (a relative stretch of 2 or more) and diffusion of a radioactively tagged penetrant from a point source is observed by autoradiography. The ratio D_1/D_2 for diffusion parallel and perpendicular to the stretch directions is readily obtainable as $(x_1/x_2)^2$, where the x's are the axes of the ellipses of constant concentration at a given elapsed time.⁷ (The analogous problem in heat conduction was studied by de Sénarmont⁸ in 1848.) Calculation of absolute values from this geometry is quite complicated, but D_1 can be obtained by autoradiography of diffusion from a line source applied perpendicular to the stretch direction; if c is the concentration at distance x and time t , a plot of $\log c\sqrt{t}$ against x^2/t gives a line with slope inversely proportional to D_1 .⁹ Moreover, a similar experiment with an

unstretched strip yields the isotropic diffusion coefficient D_0 (though this is also obtainable by other convenient methods³). It is assumed that the gradient of concentration across the film thickness becomes negligible in a relatively short time and that autoradiographic exposures can be obtained in short intervals.

III. Interpretation for Small Penetrant Molecules

Phenomenological. The relations among D_0 , D_1 and D_2 are of interest. If there is negligible volume change with deformation, analogy with refractive index suggests that for small deformations $D_0 = (D_1 + 2D_2)/3$. Possible relations for large deformations have not yet been examined.

Molecular. The free volume theory provides the simple relation⁴

$$\ln(D/T) = (B_d/f) + \text{const.}$$

where f is the average fractional free volume, a quantity which can be estimated from the temperature dependence of D by this equation (and in other ways). If there is no volume change on deformation, f should not be affected, and the anisotropy in D should reflect anisotropy in the parameter B_d , which may be regarded as a measure of the relative ease of penetration.

$$\ln(D_1/D_2) = (1/f)(B_{d1} - B_{d2})$$

$$\ln(D_1/D_0) = (1/f)(B_{d1} - B_{d0})$$

Near the glass transition temperature, f will be small and the differences will be greatest.

Alternatively, if D_1 and D_2 are measured over a range of temperatures, it can be shown that a plot of $\log D_1$ vs. $\log D_2$ gives a line with slope B_{d1}/B_{d2} . (An analogous treatment for comparing different penetrant molecules is given in Reference 5.)

The dependence of B_d on direction, magnitude of strain, and size, shape, and rigidity of penetrant should provide some insight into the mechanism of motion through the polymer.

IV. Diffusion of Oligomers and Polymers

It is also of interest to examine the diffusion of linear oligomers (molecular weight of the order of 2000) or moderately high polymers (50,000) through rubbery polymers in both uncross-linked and cross-linked states. There have been very few studies of this kind even for isotropic systems. By use of rubbers with very high internal mobility (polydimethyl siloxane, 1,4-polybutadiene), experiments within a reasonable time scale should be possible. For oligomers, it is probable that the effect of deformation will be very similar to that for smaller molecules. For polymers long enough to experience entanglement coupling with the matrix, additional effects may appear.

The diffusion of long linear molecules through a cross-linked network has been recently treated theoretically by de Gennes.¹⁰ He predicts, somewhat surprisingly, that the diffusion coefficient is independent of the distance between

cross-links provided this is small compared with total molecular length but large compared with the length of a moving segment. However, entanglement coupling is not taken into account; this will probably be strongly affected by cross-link density.

Acknowledgment. This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253.

References

1. I. Auerbach, S. D. Gehman, W. R. Miller, and W. C. Kuryla, *J. Polymer Sci.*, 28, 129 (1958).
2. R. S. Moore and J. D. Ferry, *J. Phys. Chem.*, 66, 2699 (1962).
3. S. P. Chen and J. D. Ferry, *Macromolecules*, 1, 270 (1968).
4. H. Fujita, *Adv. Polymer Sci.*, 3, 1 (1961).
5. C. P. Wong, J. L. Schrag, and J. D. Ferry, *J. Polymer Sci.*, A-2, 8, 991 (1970).
6. C. P. Wong and J. L. Schrag, *Macromolecules*, 3, 468 (1970).
7. W. Jost, "Diffusion," Revised Edition, Academic Press, 1960.
8. H. de Sénarmont, *Compt. rend.*, 22, 179 (1848).
9. J. Crank, "Mathematics of Diffusion," Oxford, 1956, p. 27.
10. P. de Gennes, private communication.

**REMARKS ON MONTROLL'S
NON-LINEAR WAVE EQUATION**

George H. Vineyard

Abstract

A solvable non-linear integro-differential equation which is a conceivable model for shock wave phenomena is examined. Conditions for the existence of an energy integral of the equation are determined and found to impose severe restrictions on the coefficients.

Preceding page blank

REMARKS ON MONTROLL'S
NON-LINEAR WAVE EQUATION

George H. Vineyard

Montroll¹ has given the remarkable result that the non-linear integro-differential equation

$$\frac{\partial^2 U(r,t)}{\partial t^2} = \int_{\underline{}} S_2(r-r')U(r;t)dr' + \iint_{\underline{}} S_3(r-r'-r'')U(r;t)U(r'';t)dr'dr'' + \iiint_{\underline{}} S_4(r-r'-r''-r''')U(r;t)U(r'';t)U(r''';t)dr'dr''dr''' + S_1(r), \quad (1)$$

where $S_1(r)$, $S_2(r)$, $S_3(r)$, and $S_4(r)$ are arbitrary functions of the spatial variable, can be solved exactly*.

$\frac{\partial^2 U}{\partial t^2} = \nabla^2 U$ is a special case of this equation, obtained by putting $S_2(r) = \delta''(r)$ and $S_1=S_3=S_4=0$, and since non-linear wave equations are almost never susceptible to exact solutions, the result is of interest and in particular suggests that some shock wave phenomena might be illuminated by solutions of this equation.

The equation was postulated for reasons of mathematical accessibility, and possible physical meanings of its non-linear terms are not obvious. It may also be noted that a vector or

¹E. Montroll, to be published. The solution is expressible as a Fourier transform of an algebraic function of Weierstrass elliptic functions. The parameters of the elliptic function depend on Fourier transforms of the kernel functions $S_n(r)$.

*Our definition of $S_1 \dots S_4$ differs from Montroll's by numerical factors.

tensor field may be used in place of the scalar U , spatial derivatives of U of all orders are implicitly included, and other time derivations may be imported, although at the cost of making the complete solution more complicated. Also, a discrete version has been given by Montroll for a field defined on a cyclic lattice. This is equally solvable.

We have investigated conservation principles for this equation, and in particular, the conditions for the existence of an energy integral. Since the second time derivative at r appears on the left, the terms on the right hand side can be considered to be the "force" exerted on the field at r by the fields and sources at all other points. To avoid dealing with functional derivatives, it is more convenient to consider the lattice version of Montroll's equation. $U(r,t)$ is then replaced by $U(n,t)$, where n denotes a lattice point, and $S_j(r)$ is replaced by $S_j(n)$. Integrals in (1) are replaced by sums over the irreducible region of the lattice. A lattice of arbitrary dimensionality is implied. Then the "force" on the dynamical variable associated with the lattice site n , $F(n)$, can be defined as

$$\begin{aligned}
 F(n) = & \sum_j S_1(n-j)U(j) + \sum_{j,k} S_2(n-j-k)U(j)U(k) \\
 & + \sum_{j,k,l} S_3(n-j-k-l)U(j)U(k)U(l) + S_1(n) \quad . \quad (2)
 \end{aligned}$$

The time variable in U has been suppressed; all sums are over

irreducible regions of the lattice.

The necessary and sufficient conditions for $F(n)$ to be derivable from a potential are

$$\frac{\partial F(n)}{\partial U(m)} = \frac{\partial F(m)}{\partial U(n)}, \quad \text{all } n \text{ and } m. \quad (3)$$

Then,

$$\frac{\partial F(n)}{\partial U(m)} = S_2(n-m) + 2 \sum_j S_3(n-m-j)U(j) + 3 \sum_{j,k} S_4(n-m-j-k)U(j)U(k).$$

Eq. (3) will be satisfied if and only if

$$S_2(n-m) = S_2(m-n), \quad (4)$$

$$S_3(n-m-j) = S_3(m-n-j), \quad (5)$$

$$S_4(n-m-j-k) = S_4(m-n-j-k), \quad (6)$$

for all $n, m, j,$ and k .

Eq. (4) thus requires $S_2(n)$ to be symmetric in n , and (5) and (6) require S_3 and S_4 to be constants if there is to be an energy integral. The somewhat disappointing result thus emerges that the non-linear terms of the equation will not conserve energy unless their kernel functions are limited to constants. The result is equally true for the continuum version of the equation.

The energy, in cases where these conditions are satisfied, can be written as the sum of kinetic and potential parts:

$$\begin{aligned}
E = & \frac{1}{2} \int_{\infty} \left(\frac{\partial U}{\partial t} \right)^2 dr + \frac{1}{2} \int_{\infty} \int_{\infty} S_2(r-r') U(r) U(r') dr dr' + \frac{S_3}{3} \int_{\infty} \int_{\infty} \int_{\infty} U(r) U(r') U(r'') dr dr' dr'' \\
& + \frac{S_4}{4} \int_{\infty} \int_{\infty} \int_{\infty} \int_{\infty} U(r) U(r') U(r'') U(r''') dr dr' dr'' dr''' + \int_{\infty} S_1(r) U(r) dr . \quad (7)
\end{aligned}$$

The question whether other integrals of the equation exist has not been answered.

Acknowledgment

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253. The stimulation and advice of Elliott Montroll is gratefully acknowledged.

"RANDOM CLOSE PACKING" OF SPHERES

George H. Vineyard

Abstract

Elementary considerations of the density of a "random close packed" arrangement of hard spheres are given.

Preceding page blank

"RANDOM CLOSE PACKING" OF SPHERES

George H. Vineyard

Bernal and Finney¹ and Scott and Kilgore² have experimented with assemblages of identical hard spheres and have found that, in addition to the familiar crystalline close-packed arrangements, there exists an apparently well defined pseudo-random arrangement called by these authors "random close packing". This arrangement is looser than close packing, possesses a wide variety of coordination numbers and has no long range order. Assemblages of ball bearings can be made to assume this packing when thoroughly shaken up in containers of irregular form. Any flat surfaces on the containers nucleate ordered arrangements and prevent the achievement of random close packing.

The apparently well defined character of the random arrangement is surprising. The arrangement is of interest as a prototype of liquid and amorphous structures. It also offers a challenging set of problems in statistical geometry.

One of the striking properties of random close packing is the apparently well defined mean density. Scott and Kilgore, working with up to 25,000 $\frac{1}{4}$ " steel ball bearings and carrying out careful extrapolations to eliminate surface effects found a mean density (volume occupied by material divided by total volume)

$$\rho_{\text{rcp}} = 0.6366 \pm .0005;$$

this is to be compared with the mean density of close-packed arrangements

$$\rho_{\text{cp}} = \frac{\sqrt{2}\pi}{6} = 0.7405.$$

In this note, a very simple argument will be suggested by which the approximate density of a random close packed arrangement can be estimated. The argument may be useful for extension to other pseudo-random systems, and for further inquiry into the nature of random close packing.

Since the models with which random packing has been investigated are static, we are concerned with the geometry of static assemblages of spheres. The volume of a large assemblage is a minimum at close packing, and reaches a larger, metastable value in random close packing. We start by supposing that the random arrangement can be produced by successive elemental disorderings applied to an arrangement that is initially close packed. Suppose that there is a parameter σ which measures the disorder of an arrangement. Define σ to lie between 0 and 1, with $\sigma = 0$ corresponding to complete order. $\sigma = 1$ corresponding to complete disorder. σ might be taken as a measure of the fraction of the atoms that have been moved off lattice sites. The volume, V , of an arrangement may be supposed to depend only on σ . The derivative of volume with respect to σ will then be a function of σ , and we write

$$\frac{1}{V_0} \frac{dV}{d\sigma} = f(\sigma) \quad , \quad (1)$$

where V_0 is the volume of the ordered assembly. It seems clear that $f(\sigma)$ will be a maximum at $\sigma = 0$, and will decline monotonically to zero at $\sigma = 1$, for in the perfectly ordered arrangement, the introduction of an element of disorder will cause the maximum expansion; the more disordered the arrangement the less the volume increase that will occur with introduction of further disorder, and at complete disorder (random close packing) disordering operations cannot increase the volume.

The simplest form of $f(\sigma)$ which possesses the above properties is

$$f(\sigma) = \alpha(1-\sigma)^p \quad (2)$$

where α and p are constants yet to be determined.

Starting from the close-packed arrangement of hard spheres, removal of a single sphere does not allow any disordering, but removal of a pair of near neighbors from close packed plane B in the stacking sequence ABA will. Suppose this pair be moved to the surface of the set, increasing the volume by $2v_0$ (where v_0 is an atomic volume). In Fig. 1 this is the pair 1 2. Then ball 3 can be moved to the point 4 in the close packed, plane, producing an elementary region of stacking fault with no further increase in volume. The triplet 1 2 3 has been involved in this elementary disordering. If a small

number n of triplets of this kind have been introduced, the volume increase has been $2nv$, and the fraction of atoms removed from lattice sites (including those moved to the surface) is $\Delta\sigma = 3n/N$. Thus,

$$\frac{\Delta V}{\Delta\sigma} = \frac{2}{3} v_0, \text{ or}$$

$$\frac{1}{v_0} \frac{dV}{d\sigma} = \frac{2}{3} \quad (\sigma \ll 1) \quad . \quad (3)$$

This identifies the constant α in Eq. (2) as $\frac{2}{3}$.

We have attempted to determine p by considering the second order correction to the volume change caused by introducing further elements of disorder into a set containing a finite number of triplets. The results are complex and have not lead to a convincing value of p . Further study of a systematic way of defining σ and averaging over the various kinds of interacting clusters that can be introduced may allow a direct determination of p . One calculation has yielded different definition of σ than that given $\alpha = \frac{2}{3}$ and $p = 3$, but it is not clear that the calculation is correct.

Eq. (2) can be integrated to give

$$\frac{V-V_0}{V_0} = \alpha \frac{(1-\sigma)^{p+1}}{p+1} \quad . \quad \text{For the transformation}$$

from close packing to random close packing this becomes

$$\frac{V_{rcp}-V_0}{V_0} = \frac{\alpha}{p+1} \quad . \quad (4)$$

Since $V \sim \frac{1}{\rho}$, the experimental results of Scott and Kilgore

quoted earlier give

$$\frac{V_{rcp} - V_0}{V_0} = \frac{\rho_{cp}}{\rho_{rcp}} - 1 = 0.1632 \pm .0007$$

Eq. (4) then requires the exponent p to be about 3, in agreement with the earlier (but not firm) result of a direct calculation.

Acknowledgment

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253.

References

1. Bernal, King, and Finney, Disc. Faraday Soc., 43, 62 (1967). Earlier references are cited here.
2. Scott and Kilgore, Brit. J. Appl. Phys. (J. Phys. D.), 2, 863 (1966).

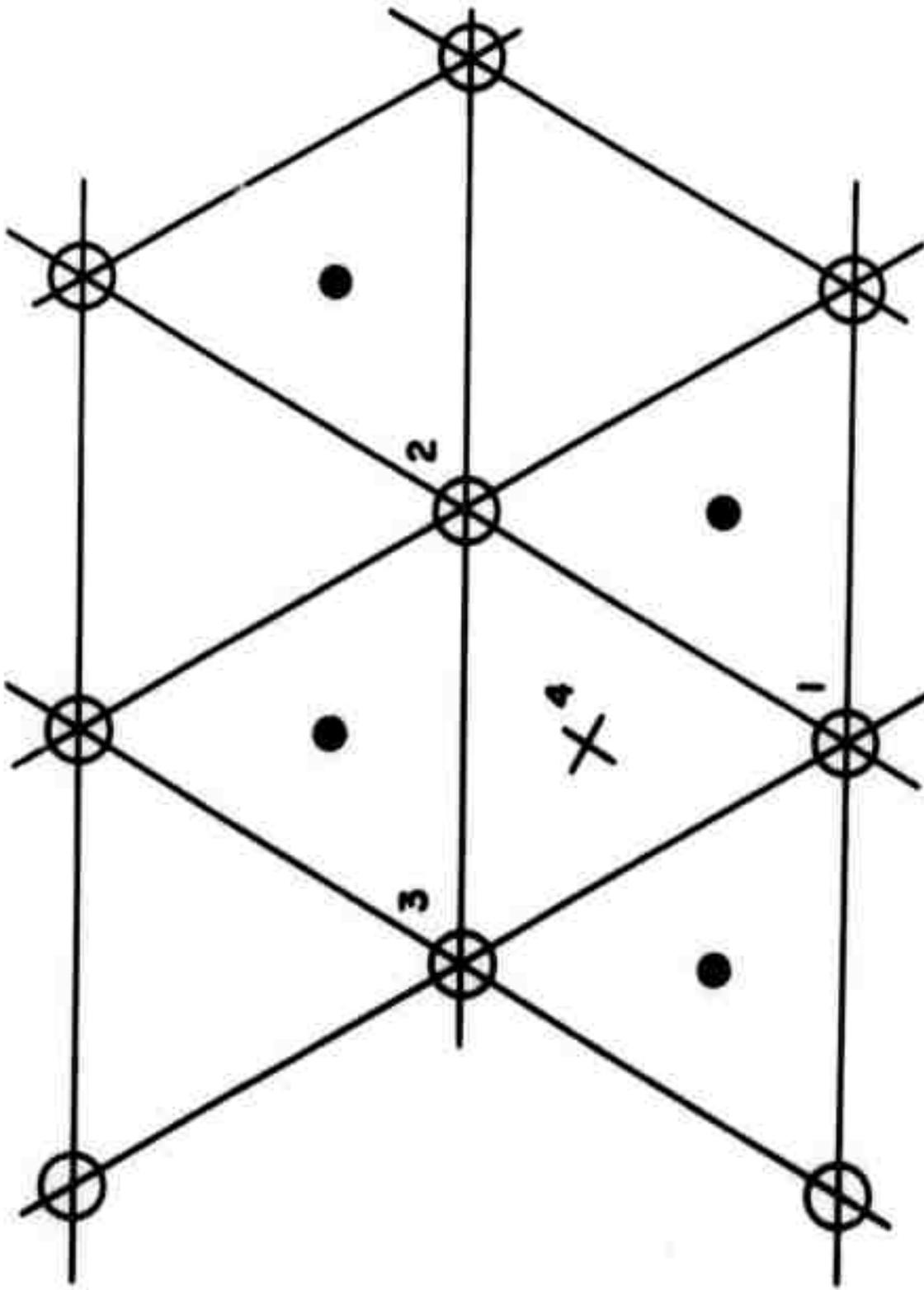


Figure 1. Lattice sites in close packed plane. Sites in the plane denoted by circles. Sites in plane below denoted by small dots. An elementary stacking fault is produced by removing balls from 1 and 2 to a distant surface and moving ball 3 to point 4.

NOTE ON IR WINDOWS TRAVERSED BY
SHORT LASER PULSES*

N. Bloembergen

Abstract

Considerable attention has been devoted to the thermal loading with resultant deformation and beam distortion, when an infrared laser beam traverses a window material. In comparing window materials, it has typically been assumed that the laser beam would be on for a period of many milliseconds or longer. It is pointed out in this note that a different situation prevails when the duration of the laser pulse is shorter than the time required by a sound wave to traverse the window. In the limit of very short pulses the beam deformation due to residual absorption in the window is largely determined by $(\partial n/\partial T)_{\text{strain}}$. Since the peak powers in the short pulses will be orders of magnitude higher than in pulses lasting milliseconds or seconds, attention should be paid to other possible modes of window failure in this regime.

Preceding page blank

NOTE ON IR WINDOWS TRAVERSED BY
SHORT LASER PULSES

N. Bloembergen

I. Introduction

It is well established that the transmission of high power diffraction limited IR laser beams through windows is limited by the optical distortion introduced in the window by the (low) residual absorption losses of the window material. The performance of different materials has been analyzed and compared on the basis that the IR laser beam is switched on for a relatively long time, typically on the order of one second. This time is short enough that thermal diffusion may be ignored and the temperature rise in each volume element in the window may be calculated from the total energy absorbed in that element. The on-time of the laser beam is long enough, however, that the deformation of the window resulting from the thermal stresses may be calculated in a static manner. For the many detailed results that have been obtained for this mode of operation we refer to a number of recent technical reports¹⁻⁴.

Recently, high-power IR lasers with good optical beam quality have been operated in short pulsed modes. These lasers are typically of the transversely excited discharge type with the active gas at relatively high pressure. The optical uniformity of the gas characteristics is improved by pre-ionization

techniques. It may therefore be anticipated that relatively high energy IR pulses of short duration, say one microsecond or less, will become available. It should be noted that the total energy required to induce material damage in some applications may be reduced up to an order of magnitude for short pulses, as the damage mechanism changes from heat ablation to spallation.

Consequently the thermal loading of IR windows would be reduced by the same factor. It turns out that the previous analyses¹⁻⁴ require considerable modification if the pulse duration becomes comparable to or shorter than the time required for a sound wave to traverse the window dimension. In that case inertial effects become important in the elastic behavior and the quasi-static deformation analysis loses its validity. In the next section a qualitative discussion of window distortion by a single very short pulse is given, while other possible failure mechanisms besides linear energy absorption loss are mentioned in the third section. The final section summarizes this note in a number of conclusions and recommendations.

II. Thermal Beam Distortion of a Very Short Laser Pulse in an IR Window

Consider a plane parallel IR window, as shown in Figure 1, of thickness L_0 and radius ρ_0 , and a very short incident laser pulse of duration $t_p \ll L_0/c_\ell$. Here c_ℓ is the propagation velocity of the fastest acoustical wave mode that

propagate in the crystal. During the duration of the laser pulse, the window material has no time to deform. The amount of heat deposited by residual absorption loss mechanism will heat each volume element at constant strain. The phase front of the wave will only be affected by the intrinsic dependence at constant strain and volume of the index of refraction. One has

$$\Delta n(\rho, t) = \left(\frac{\partial n}{\partial T} \right)_{\text{strain}} \Delta T(\rho, t) \quad (1)$$

$$\Delta T(\rho, t) = \frac{\beta}{C_V} \int_0^t I_0(\rho, t') dt' \quad (2)$$

Here C_V is the specific heat at constant volume, β is the absorption coefficient, which is assumed small, $\beta L_0 \ll 1$, so that the beam pulse intensity $I_0(\rho, t)$ is independent of z .

The phase variation at the exit plane is

$$\Delta \phi(\rho, t) = 2\pi (L_0/\lambda) \Delta n(\rho, t) \quad (3)$$

Note that there is no window distortion on this time scale, nor has there been sufficient time lapse for stress birefringence to develop.

The relevant material constant $(\partial n/\partial T)_{\text{strain}}$ may be deduced for cubic crystals from known variation of the index of refraction with the temperature and pressure P , the thermal volume expansion and the compressibility.

$$\left(\frac{\partial n}{\partial T}\right)_P = \left(\frac{\partial n}{\partial T}\right)_{\text{strain}} + \left(\frac{\partial V}{\partial T}\right)_P \left(\frac{\partial n}{\partial V}\right)_T$$

$$\left(\frac{\partial n}{\partial P}\right)_T = \left(\frac{\partial n}{\partial V}\right)_T \left(\frac{\partial V}{\partial P}\right)_T \quad (4)$$

It may be expected that $(\partial n/\partial T)_{\text{strain}}$ is small for the ionic crystals with relatively large energy gaps. The candidates with the lowest product $\beta(\partial n/\partial T)/C_V$ would again be NaCl, KCl and KBr. The covalent materials and materials with very heavy polarizable ions would be less favorable from this point of view.

After the passage of the very short laser pulse the window would deform in a time interval on the order of ρ_0/c_L to the deformed state previously assumed¹⁻⁴. Then cooling mechanisms described previously¹⁻⁴, such as supercooling by gas jets on the window surfaces, should restore the initial uniform window condition in a time less than a few seconds, after which another short pulse could be fired.

It is of some interest to make an admittedly very qualitative explanation about the partial cancellation of $(\partial n/\partial T)_{\text{stress}}$ and the effect of window distortion for the alkali halides mentioned in the earlier reports. Consider the hypothetical case that the window has had time to deform elastically in the z-direction as indicated by the dotted lines in the Figure 1b, but there is no deformation in the radial direction. This situation may be approximated physically in a time interval $c_L/L_0 < t < c_L/\rho_0$. Since a light ray

parallel to the z-direction still traverses the same total amount of material, i.e., it passes by the same number of positive and negative ions, there is no change in phase shift, if changes in local field corrections due to the deformation could be ignored. Thus the part of $(\partial n/\partial T)_{\text{stress}}$ due to expansion is compensated by the curvature of the window surfaces. The uncompensated part is $(\partial n/\partial T)_{\text{strain}}$ and this quantity is more fundamental from a physical point of view. Clearly, the actual situation will be further complicated by changes in local field corrections and by subsequent expansions in the radial direction (Poisson's modulus).

Although it is clearly important to extend the analysis to a more quantitative treatment of the transient development of phase distortion in the window, the argument shows that for a given amount of total energy in the laser pulse, the distortion for short pulses will certainly not be worse than for longer pulses. It might be considerably better, if a material could be found with a very small value of $(\partial n/\partial T)_{\text{strain}}$. The main obvious advantage of short pulses would result from the fact that the total energy per pulse for a required effect would be considerably reduced with a concomitant reduction in wave front distortion by the loss mechanisms.

III. Other Failure Mechanisms

When the pulse duration is short, it is necessary to consider the consequences of the use of higher peak power densities, say up to 10 megawatts/cm². It is conceivable that

limitations other than thermally-induced optical beam distortion become important. The following nonlinear optical or mechanical properties must be considered.

The sudden heat production will produce acoustical disturbances and possibly some shock waves after the light pulse has passed through. It is precisely by these mechanisms that the window will approach its equilibrium elastic deformation as discussed above. It should be pointed out that the thermal shock is very mild corresponding to maximum temperature differentials of about 1°C in low absorption window materials. It is well established that the intrinsic mechanical and electric breakdown threshold for bulk alkali halide crystals lies in the range of 10^{10} to 10^{11} watts/cm² for 10.6μ radiation⁵. Although the contemplated peak power density levels of 10^7 watts/cm² do not seem to represent a serious threat to the bulk material, it is nevertheless important to obtain experimental data on the probability of breakdown in bulk materials by repeated pulses, and on the influence of surfaces and impurities in the regime of short intense pulses.

Nonlinear absorption mechanisms which should be considered include the following:

1. Raman scattering, followed by subsequent absorption of the stokes (or anti-stokes) shifted light.
2. Two photon absorption and multiphoton absorption processes.
3. Harmonic generation with subsequent linear absorption of the harmonic. Rough estimates indicate that, at power flux

densities of 10^6 watts/cm², these processes are less important than the linear absorption mechanism with $\beta \sim 10^{-3}$ cm⁻¹. The nonlinear mechanisms will become of importance relatively sooner in the semiconductors with large nonlinear susceptibilities than in the alkali halide crystals. Clearly, quantitative results for these nonlinear mechanisms in different materials must be tabulated.

Particular attention should be paid to the phenomenon of self-focusing. It is well known that this mechanism presents a limitation in high power Nd-glass laser rods. Since it leads to a large increase in the effective power flux density, it may constitute a threshold for material breakdown. Since the quadratic Kerr-type nonlinearities are known^{6,7} for many alkali-halide and semiconductor materials of groups IV, III-V and II-VI, the effect may be readily estimated. It is not expected to constitute a serious breakdown limitation for the range of beam sizes, power flux densities and window thicknesses of interest.

Nonlinear index changes could, however, have some effect on the optical beam distortion. This may be seen by comparing the nonlinear change in index of refraction proportional to the intensity with the thermally-induced change in index given by Eq. (1),

$$\Delta n_{\text{nonlinear}} = n_2 |E|^2 = 2\pi\chi^{(3)} |E_L|^2/n_0 \quad (5)$$

Here $\chi^{(3)}$ is the third order nonlinear susceptibility, E_L is

the amplitude of the electric field in the beam. $\chi^{(3)}$ has been measured for several semiconductors⁶ for frequency mixing at 10.6μ . Its magnitude is on the order of 10^{-11} esu for Ge, GaAs and other semiconductors with a sufficiently low concentration of conduction electrons. A power flux density of 10^7 watts/cm² corresponds to $|E_L|^2 \sim 3 \times 10^4$ esu. Thus $\Delta n_{\text{nonlinear}} \approx 10^{-6}$. This is still one or two orders of magnitude smaller than the thermally-induced effect in these materials. Fortunately, the nonlinear index change has even less relative importance in the alkali-halides, since for these $\chi^{(3)} \approx 10^{-14}$ to 10^{-13} esu. Thus the nonlinear effects decrease by a larger factor than the thermally-induced linear effect. It is clear that for the highest peak power pulses alkali halide windows have a decisive advantage over covalent materials because of the smallness of their nonlinear coefficients.

IV. Conclusions

1. It is desirable to analyze in more detail the behavior of IR window materials subjected to short laser pulses, with $t_p < 10^{-8}$ sec.
2. Figures of merit of different materials will be quantitatively different in this regime, but the alkali halide crystals are expected to be significantly better than semiconductor crystals and glasses, from the standpoint of beam distortion.
3. At peak power density levels higher than 10^7 watts/cm², serious attention must be paid to the nonlinear optical

properties of the window materials.

Acknowledgment

This study was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253. The author is indebted for stimulating discussions to participants in the Woods Hole meeting on IR windows, July 12-13, 1971, in particular to Dr. C. M. Stickley.

References

1. M. Sparks, Immediate Needs of the High Power Infrared Window and Methods of Reducing Thermally Induced Optical Distortion. Rand Corporation, Report WN-7243-PR, April 1971.
2. M. Sparks, Physical Principles, Materials Guidelines and Materials List for High Power 10.6 Windows. Rand Corporation, Report WN-7396-PR, April 1971. An other reports quoted therein.
3. F. A. Horrigan and R. I. Rudko, Materials for High Power CO₂ Lasers. Raytheon, Final Technical Report AD 691331. Contract No. DA-AH01-69-C-0038, May 1969.
4. F. A. Horrigan and T. F. Deutsch, Research in Optical Materials and Structures for High-Power Lasers, Raytheon Quarterly Progress Reports under Contract No. DA-AH01-70-C-1251, ARPA Order No. 1180, March 1971.
5. E. Yablonovitch, private communication, and Appl. Phys. Letters, to be published.
6. J. J. Wynne, Phys. Rev. 178, 1295 (1969).
7. N. Bloembergen, W. K. Burns and M. Matsuoka, Optics Commun. 1, 195 (1969).

LINE TENSION AND KINKS ON FRACTURE CRACKS

J. J. Gilman
J. P. Hirth
R. M. Thomson

Abstract

The concept of line tension for a crack is explored and applied to the atomic kink on a crack. The line tension turns out to be so dependent upon shape that a simple use of it as in dislocation theory is not possible. For atomic kinks, the argument leads to the prediction of an abrupt kink shape. The activation energy of the kink is not unique, but has values over a range from zero to a maximum value. This range of activation energies results from the lattice trapping plateau for the crack.

LINE TENSION AND KINKS ON FRACTURE CRACKS

J. J. Gilman
J. P. Hirth
R. M. Thomson

I. Introduction

Like the dislocation, the fracture crack is a line defect in a solid, and both are special cases of the general Somigliana dislocation. Three dimensional theories of both are difficult, but both can be treated with a fair degree of completeness and rigor in two dimensions, in large part because of the power of complex function theory. Consequently, when it was noticed very early in dislocation theory¹ that a dislocation line in three dimensions possessed many of the attributes of a flexible string, the string model of the dislocation became an important part of the lore of dislocation theory.

In the string analogy (whether applied to a crack or to a dislocation) the assumption is made that the energy of the line defect can be represented as an energy density seated in the geometric line representing the defect, and that other forces act on the line as point functions of the line position. The total potential energy of the solid is then a functional of the line position and length,

$$V = \int (\epsilon dl + \phi(x) dx) \quad (1)$$

where ϵ is the energy density per unit length of line and

$\phi(x)$ is any contribution to the potential energy of a segment of line at the point x . In a general case, c may also be a function of position, and of slope, curvature, etc. of the line. For the dislocation, the energy density is roughly

$$c = \mu' b^2 \ln R/b \quad (2)$$

where μ' is some kind of modulus, b is the Burgers vector, and R is some kind of outer cut off. In fact, R depends upon the shape of the line, and hence (2) violates the basic assumption that the energy density is a point function. The success of the line tension approximation rests upon the fact that the shape dependent term sits under a logarithm, and for many applications, the line tension is simply taken to be a constant. As we shall see for the crack the analogous situation is not so fortuitous.

One of the important applications of the line tension concept in dislocation theory is in working out the response of the dislocation to the Peierls energy. In this case, one is led to a theory of kinks², and in this paper we shall follow a similar line of argument.

Recently, lattice trapping of fracture cracks has been suggested in direct analogy to the Peierls crystallographic energy barriers of dislocations³. Computer calculations have also been reported of kinks on cracks in discrete calculations⁴, and indicate significant crack kink energies. This paper is an attempt to assay the importance that line tension may be

able to play in discussions of thermal kinks on a crack line.

II. Line Tension Approximation

As in dislocation theory, we must identify which terms in (1) for cracks are proportional to the line length, and which are a function of the crack position. We work from the Griffith theory of fracture for straight cracks. Then, if y is the crack displacement,

$$\frac{d}{dy} \left(\begin{array}{c} \text{total} \\ \text{system} \\ \text{energy} \end{array} \right) = \frac{d}{dy} \left(\begin{array}{c} \text{internal} \\ \text{elastic} \\ \text{energy} \end{array} \right) + \frac{d}{dy} \left(\begin{array}{c} \text{external} \\ \text{load} \\ \text{energy} \end{array} \right) + \frac{d}{dy} \left(\begin{array}{c} \text{surface} \\ \text{energy} \end{array} \right) = 0 \quad (3)$$

is the condition for equilibrium for straight cracks.

The various terms are respectively,

$$E_{\text{elastic}} = \pi \sigma^2 y^2 / 4Y \quad (4a)$$

$$E_{\text{external}} = -\pi \sigma^2 y^2 / 2Y \quad (4b)$$

$$E_{\text{surface}} = 2\gamma_0 y + f(y) \quad (4c)$$

σ is the applied stress, Y is Young's modulus, γ_0 is the average surface energy, and $f(y)$ is a periodic surface energy function which reflects the atomic character of the lattice. The function, $f(y)$ reflects the atomicity, leads to lattice trapping, and is discussed in reference (3).

Clearly, the surface energy term is simply proportional to the surface area swept out by the crack. The elastic term is proportional to the length of a crack for a straight crack.

For the moment, we shall simply assume the same form (4a) for the curved crack too. The external work term reflects the average strain throughout the sample and if the crack is convoluted in some manner, it is only the average position of the crack which specifies the average strain. We shall thus take the second term to be a "potential term" rather than a length term in (1). Even though the elastic energy and external work done have the same form in (4) for straight cracks they are physically different, because the elastic energy is enhanced by a highly convoluted crack line because of the larger number of atoms with high strain energy around the convoluted crack tip. The external work is affected only by the average crack depth, not its convolutions. There are questions of cut-off procedures for a convoluted line to which we return in later paragraphs. Thus, from (4) and (1), we have

$$V = \int \left\{ \lambda y^2 \sqrt{1+y'^2} + 2\gamma_0 y + f(y) - 2\lambda y^2 \right\} dx \quad (5)$$

$$\lambda = \pi \sigma^2 / 4Y$$

If $\sqrt{1+y'^2} = (1 + \frac{1}{2}y'^2)$, the differential equation for the string shape is

$$\lambda y^2 y'' - 2\gamma_0 - f'(y) + 2\lambda y = 0 \quad (6)$$

Equation (6) has an immediate solution representing the straight crack which contains useful information. If $y'' = 0$, then the modified Griffith condition becomes

$$\lambda y = \left(\gamma_0 - \gamma_1 \sin \frac{2\pi y}{a} \right) \quad (7)$$

where the periodic function $f(y)$ has been assumed to have the simple form

$$f = \frac{\gamma_1 a}{\pi} \left[1 - \cos \frac{2\pi y}{a} \right] \quad (8)$$

a is the lattice spacing and γ_1 is the amplitude of the surface energy variation. The standard Griffith condition is obtained from (7) by setting $\gamma_1 = 0$.

The solution (7) is depicted in Fig. 1. In general, there are multiple solutions stretching over a stable region of the crystal, in contrast to the standard Griffith solution which predicts a single unstable solution. The Griffith value for y is at the midpoint of the stable region. For different values of the stress, the stable region for the crack shifts back and forth.

The length of the stable region is obtained by noting that the upper limit of the stable region is where $\sin 2\pi y/a = +1$ and the lower limit is where $\sin \frac{2\pi}{a}y = -1$. Hence if y_+ is the maximum stable crack position and y_- is the minimum stable crack position, then

$$\delta y = y_+ - y_- = 2\gamma_1/\gamma_0 \quad (9)$$

If we designate the "Griffith depth", y_g , as that value of y given by (7) when $\gamma_1 = 0$, then

$$\frac{\delta y}{y_g} = 2\gamma_1/\gamma_0$$

(10)

$$y_g = \gamma_0/\lambda$$

This equation states the important result that for straight cracks, δy is a macroscopic quantity when γ_1 is not many orders of magnitude less than γ_0 . The ratio γ_1/γ_0 depends drastically upon the size of the cohesive region or core width, of the crack, and in the computer calculations for iron⁴, the core width appears atomically narrow. In this case the stable region would be of the same order of magnitude as the crack depth itself.

Another point of some consequence is immediately noticeable from Fig. 1. If a crack is at some stable point within δy , the excess stress to overcome the lattice barriers depends upon the starting position within δy . Near the forward limit of the region, a very small stress is necessary to overcome the lattice barriers, and convert the crack into a dynamically moving crack, while a maximum stress is necessary to overcome the barrier near the rear limit of the stable region. This point will be important in the discussion of kinks.

III. Kinks

Within the region, δy , although the straight crack is mechanically trapped by the atomic periodicity of the surface energy, the crack can move by the local kink generation due to thermal fluctuations. We now try to solve for the saddle point

energy configuration of a kink on the crack line where the kink is in unstable equilibrium under the applied stress. We return to (5) and assume that the crack is straight everywhere except for a local kink of small excursion. Hence,

$$y = y_0 + \xi \quad (11)$$

and we keep only linear terms in $\xi(x)$. Hence (6) becomes

$$y_0^2 \lambda \xi'' + 2\lambda \xi - f'(\xi) = 2\gamma_0 - 2\lambda y_0 \quad (12)$$

We arbitrarily choose y_0 on a microscopic scale so that $f'(\xi + y_0) = f'(\xi)$. At the midpoint of the stable region where the Griffith condition is satisfied, the right side is zero in (12). At other points within the stable region, there will be a "driving term" which will bias the thermal fluctuations so that on balance the crack either closes or opens.

In the case of the sinusoidal surface term, (12) becomes a nonlinear equation, the precise solution of which is not worth the trouble to obtain numerically. We learn all that is necessary about the solution by writing the equation for $\xi \ll a$ which will give us an adequate estimate for the width of the kink. Then (12) becomes

$$y_0^2 \lambda \xi'' - 2 \left(\frac{2\pi\gamma_1}{a} - \lambda \right) \xi = g \quad (13)$$

$$g = 2(\gamma_0 - \lambda y_0)$$

For a single kink which crosses over from one lattice through to another at $g = 0$, the two asymptotic solutions are

$$\begin{aligned} \xi &= a e^{-x/l} & x >> l \\ \xi &= a(1 - e^{x/l}) & x \ll -l \end{aligned} \tag{14}$$

where

$$\frac{a^2}{l^2} = 2 \frac{a^2}{y_0^2} \left(2\pi \frac{y_g}{a} \frac{y_1}{y_0} - 1 \right) \tag{15}$$

We have used the standard Griffith condition in the form $y_g = \gamma_0/\lambda$ to write (15). So long as the crack is in the middle of the stable region, and assuming that the crack depth is sufficient (see Fig. 1), the first term in (15) is dominant. If $y_0 = y_g$ in (15) then the kink width becomes simply

$$\frac{l}{a} = \frac{1}{2} \sqrt{\frac{1}{\pi} \frac{y_g}{a} \frac{y_1}{y_0}} \tag{16}$$

which for macroscopic crack depths is very large, and leads to extremely large kink energies.

Equation (16), however is an unphysical solution because in it the entire crack depth, y_0 , contributes to the line tension. Some numbers will show the nature of the difficulty. In (16) if $y_0/y_1 = 10$, corresponding to a fairly narrow crack core, and if the crack is one micron deep, $y_g/a\pi = 10^{+3}$, then $l/a = 10^2$. Thus the distance, y_g , contributing to the line tension is an order of magnitude greater than the kink width itself. According to St. Venant's principle, the strain field of a local disturbance should extend in

the medium to about a distance equal to the dimension of the strain source itself. The troublesome term, y_0 or y_g , first enters the elastic energy term in Eq. (4a), and contributes to the line tension of a uniformly straight crack because the crack opening displacement contributes significant elastic energy density over regions of the crystal at large distances from the crack tip. In a sense, the strain field of a crack is "long range", not because of the elastic misfit near the crack tip, but because of the long range "misfit" over the macroscopic open surface of the crack. However, when a macroscopically deep crack possesses an atomic kink, the local kink strain field must not depend upon the macroscopic crack depth as predicted by (16), but must depend only upon the details of the local atomic forces. After all, the displacements of the atoms in the immediate vicinity of the crack tip depend only upon these same atomic force details, and not upon the total crack depth, and there is no reason to suppose a kink to behave otherwise.

The line tension approximation must then be subjected to a cut-off procedure. There are several ways to affect this. The first is to attempt to set $y_0 = \ell$ in (16), but this leads to no solution at all. The second is to return to (4) and set $y_0 = R$ in those terms which appear as coefficients of y' in the potential functional, where R is the local radius of curvature of the line. Care must be used in doing so to retain (7) as the limit for straight cracks in the subsequent equations.

This approach leads to highly nonlinear equations, and high derivatives from which physically meaningful results are not easy to abstract. A third and more successful technique is give up the differential equation for the kink and instead to minimize the line energy of a segmented kink as shown in Fig. 2.

Since the midpoint of the kink is stationery, the external work term is constant, and we can write (1) as

$$\delta V = \ell^2 \lambda \left[\frac{\ell}{\cos \theta} - 1 \right] + \int \frac{\gamma \ell a}{\pi} \sin \frac{2\pi y}{a} dx \quad (17)$$

(We have shifted the origin in our use of $f(y)$ from its original form, (8), in a trivial manner). ℓ is the width of the kink shown in Fig. 2, and θ is the angle of the kink. From Fig. 2, $y = x \tan \theta$, $a = \ell \tan \theta$.

$$\delta V = \frac{\lambda a^3}{\tan^3 \theta} \left(\frac{1}{\cos \theta} - 1 \right) - \frac{a^2}{2\pi^2 \tan \theta} \quad (18)$$

which has a minimum value at $\theta = \pi/2$, corresponding to a sharp or abrupt kink.

Double Kinks

The proper procedure for the string model at this stage would be to solve for the saddle point configuration for the double kink in a properly modified equation, (12), under the influence of the driving term. However, we have just demonstrated that the simple minded concept of line tension breaks down, and that in fact the kinks are sharp. In the limit where the driving term, g , is zero, the energy E_k , to form a sharp

double kink is simply the energy to ride over the hump in the surface energy given by (8).

$$E_{k0} = \frac{2\gamma_1 a}{\pi} l \quad (19)$$

We assume that l , the kink length, in general may be a few atomic distances instead of strictly one, in view of our inability to take account properly of the core effects.

We remember, however, that this kink activation energy is not unique, and represents only a maximum in a continuous spectrum depending upon where the kink is in the stable region. From Fig. 1, the height of the barrier decreases linearly to zero from the midpoint of the stable region where the Griffith condition is exactly satisfied out to the front edge of the stable plateau, where the dynamic crack condition becomes satisfied.

IV. Conclusions

Our result is that the line tension of a crack line depends markedly on the crack shape. The average line tension for extension of a straight crack is large, and the distance, y_0 , appearing in the line tension, (4a), is the total depth of the crack. However, if we wish to bow out the crack line over a distance appreciably smaller than the depth of the crack, the line tension then decreases sharply through the use of St. Venant's principle, so that $y_0 \approx R$ where R is the linear dimension of the bow out. In the case of the atomic kink, this theorem leads one to expect an atomically sharp kink, with

the details of the force laws on the atoms in the vicinity of the crack tip determining the final result. The kink energy is then simply related to the parameter γ_1 , which measures the sinusoidal variation in the surface energy as a function of crack position.

We can comment on what light the recent computer calculations of Kanninen et al⁴ throw on these results. The computer calculations apparently qualitatively verify our findings of sharp kinks. To be sure, they are not atomically sharp for reasons which are beyond our approximations, and have to do with atomicity. Our result shows that unlike dislocations, the kink width should not vary much with f , except of course as $f \rightarrow 0$, the crack line should slowly straighten out.

Finally, the activation energy for crack mobility through the stable region will show a dense uniform spectrum with a maximum value given approximately by the energy to form two single kinks, and a minimum value of zero.

Acknowledgment

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253.

References

1. J. S. Koehler, *Imperfections in Nearly Perfect Crystals*, John Wiley & Sons, (1952) p. 197.
2. A. Seeger, *Phil. Mag.* 1, 651 (1956).
3. R. Thomson, C. Hirth, V. Rana, *J. Appl. Phys.* 42, 3154 (1971).
4. P. Gahlen, M. Kanninen, *Forthcoming Proceedings, Battelle Northwest Symposium*, 1971.

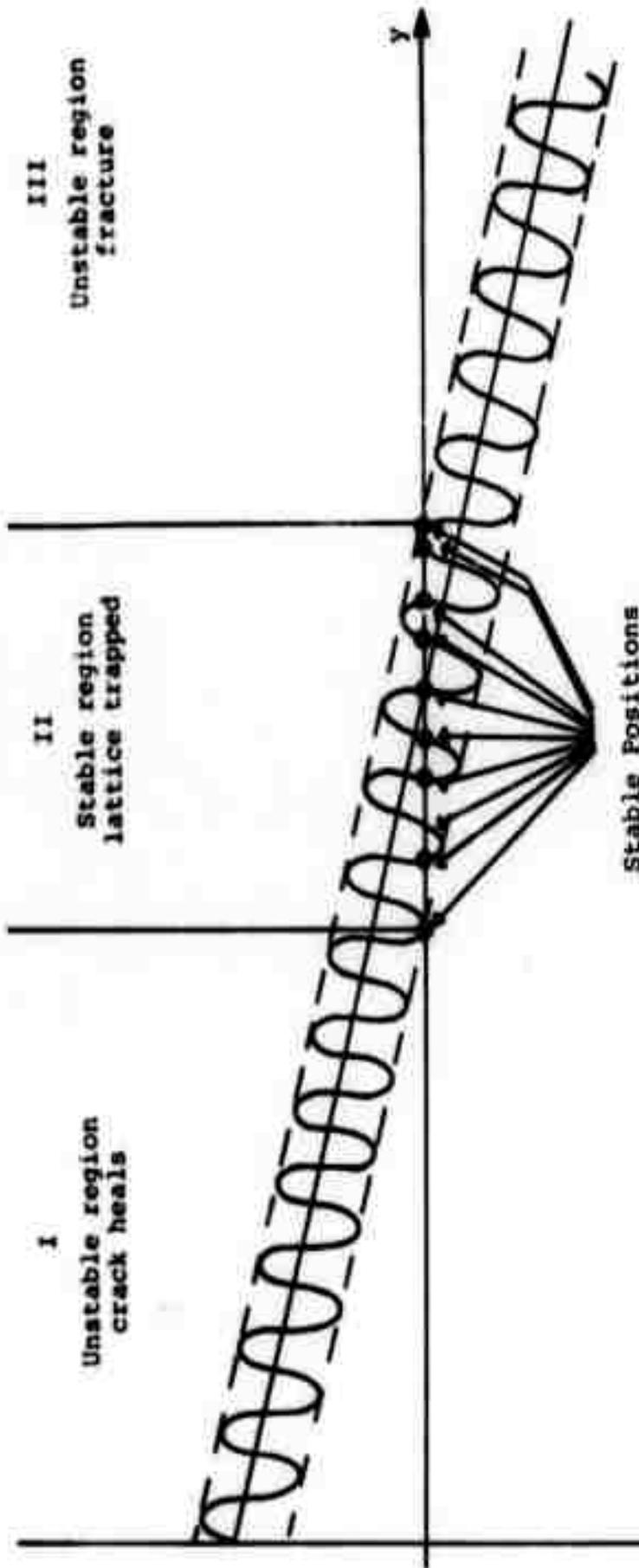


Figure 1. Plot of the function $\gamma_c + \gamma_i \sin \frac{2\pi y}{\lambda} - \lambda y$, showing three regions. For small cracks, the stress is insufficient to hold the crack in the solid. In region II, the crack is mechanically stable, while in region III the crack moves dynamically. Stable positions are those points of intersection of the axis by the curve.

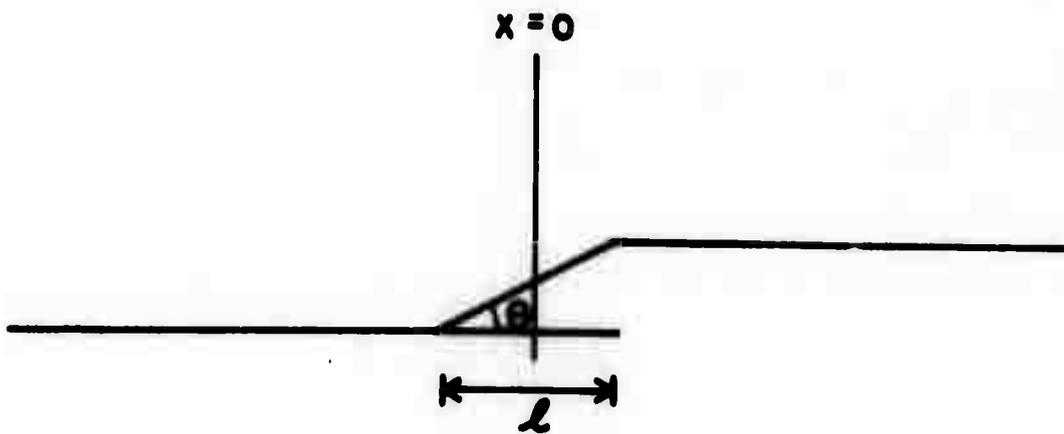


Figure 2. Kink as straight segments.

STATES OF EASE OF POLYMERIC ENTANGLEMENT NETWORKS CROSS-LINKED IN STRAINED STATES

John D. Ferry

Abstract

An experiment is proposed to provide information about the nature of "entanglement coupling", a somewhat vague concept which is widely applied to interpret mechanical and other physical properties of polymers of high molecular weight. A sample of such a polymer, visualized as a network of entanglement loci, is deformed in simple extension and subjected to chemical cross-linking; upon release it contracts and approaches a first state of ease in which the forces exerted by the effective networks of entanglement loci and of cross-links are balanced. If some of the entanglements remain untrapped by the cross-links and can rearrange, the sample then more slowly extends again and approaches a second state of ease. The dimensions in these states have been evaluated numerically for various cases, based on existing theories for the related problem for chemical cross-links in two successive stages. Experimental measurements of equilibrium dimensions, as well as the kinetics of dimensional change, and also changes in dimensions on subsequent swelling in solvent, can clarify the nature of entanglement coupling and trapping.

STATES OF EASE OF POLYMERIC ENTANGLEMENT NETWORKS CROSS-LINKED IN STRAINED STATES

John D. Ferry

I. Introduction

Many mechanical and certain other physical properties of uncross-linked polymers of high molecular weight have been interpreted in terms of rather strong coupling between molecules at widely separated points, commonly termed entanglements.¹ It is fairly clear that the coupling arises from topological restrictions; its effects on viscosity, steady-state compliance, and other mechanical properties have been adumbrated in some detail,^{2,3} and the average number of moles of effective entanglement junctions per unit volume, c , can be roughly estimated from various rheological measurements.⁴ Qualitatively, there is a range of time scale within which portions of a molecule between entanglement loci rearrange their configurations frequently but long-range configurational changes are prevented by the extreme slowness of entanglement slippage. Within this time scale, the mechanical properties resemble those of a cross-linked network. However, the entanglement concept is somewhat vague because there can be no specific loci on a molecule of uniform chemical composition; the precise nature of entanglement coupling remains speculative.

It should be possible to obtain further information by a type of experiment which is a modification of a procedure

used some years ago by investigators of cross-linked rubbers¹, involving cross-linking in strained states.

II. Description of Proposed Experiment and Anticipated Results

A polymer of sufficiently high molecular weight (so that entanglement slippage is minimized within a convenient laboratory time scale) is deformed in simple extension with a stretch ratio λ_0 . Chemical cross-links are introduced in this fixed strained state; then the sample is unclamped and allowed to retract, still within a time scale such that there is no entanglement slippage, to its first state of ease⁴. This corresponds to a state of minimum strain energy in which, qualitatively, the original entanglement network and the chemical cross-link network are pulling in opposite directions.

At longer times, some of the entanglements may slip and rearrange to allow resumption of a random distribution of configurations for certain portions of the network. Certain properties of lightly cross-linked networks have been interpreted in terms of a dichotomy between entanglements which are trapped between cross-links and cannot rearrange, and others on dangling structures which are untrapped⁷; the proportion of trapped entanglements increases with the degree of chemical cross-linking. If untrapped entanglements can subsequently rearrange, the sample will subsequently change its dimensions slowly in the opposite direction to a second state of ease.

III. Theory for First State of Ease

Equilibrium. The theory for the related experiment in which one starts with some chemical cross-links introduced in the isotropic state and adds more cross-links in a strained state has been developed by several authors,^{5,6,7} on the basis of the stress-strain relations of neo-Hookean elasticity, i.e., retractive force in simple extension proportional to $\lambda - 1/\lambda^2$ where λ is the relative length in the direction of stretch. If we substitute entanglement loci for first-stage cross-links, assuming that all linkages are tetrafunctional and loose ends can be neglected, the relative length λ_s in the first state of ease is given by

$$\lambda_s = \left(\frac{1 + \lambda_0 v_c / c}{1 + v_c / c \lambda_0^2} \right)^{1/3} \quad (1)$$

where v_c is the moles cross-links per cc. introduced at a fixed stretch ratio λ_0 and c is the moles of entanglement loci per cc. In Figure 1, λ_s is plotted against λ_0 for several different values of the ratio v_c/c . The separations of the curves from the dashed line with slope of unity represent the amounts of retraction from the initial stretched state to the state of ease.

The neo-Hookean stress-strain relation does not hold strictly in simple extension for $\lambda > 1.2$. Calculations have been made for the more general Mooney-Rivlin relation which introduces another empirical constant⁸, and for non-Gaussian strands with limited extensibility¹⁰, but they are very compli-

cated and these refinements are probably not necessary for present purposes. Comparison of experimental data with Figure 1 should provide a new method of determining the entanglement density ϵ and also of looking at its possible dependence on λ .

If Young's modulus (or the shear modulus) is now measured in small deformations from the state of ease, it is predicted to differ from the original modulus E_0 by the factor⁵

$$E_s/E_0 = (\lambda_0^2 - \lambda_0^{-1})/(\lambda_0^2/\lambda_s^2 - \lambda_s/\lambda_0) \quad (2)$$

In Figure 2, this ratio is plotted against v_c/ϵ for various values of λ_0 . For small extensions, E_s/E_0 approaches $1 + v_c/\epsilon$ (additivity of entanglements and cross-links), but at higher extensions it exceeds this value. It should be noted that the moduli E_s and E_0 refer to a time scale within which there is no entanglement slippage ("plateau" modulus), and may be difficult to specify experimentally, but in principle their ratio affords an alternative method for determining ϵ and its possible dependence on λ .

Kinetics. The approach to the first state of ease is essentially a creep recovery process for the entanglement network combined with a creep process for the cross-link network, with time-dependent stresses equal and opposite. In principle, its course should be predictable from linear viscoelasticity data; comparison with experiment may provide further clues to the nature of the entanglement coupling. For large deformations, the problem is considerably more complicated but

may be approached with a suitable constitutive equation^{11,12}.

IV. Theory for Second State of Ease

Equilibrium. Relaxation of untrapped entanglements corresponds, in the theory for two successive stages of cross-links, to removal of a portion of the first-stage cross-links. Flory⁶ has calculated the final state of ease in this situation; with suitable substitutions, the relative length λ_{gr} can be obtained from equation 1 by replacing ν_c and c with the following effective values:

$$c_r = T_e c + \phi \nu_c \quad (3)$$

$$\nu_{cr} = \nu_c - \phi \nu_c \quad (4)$$

$$\text{where } \phi = \frac{2}{\phi_2} \left(1 + \frac{1-\phi_2}{\phi_2} \ln (1-\phi_2) \right) \left(1 + \frac{1-\phi_1}{\phi_1} \ln (1-\phi_1) \right) \quad (5)$$

and $\phi_2 = \nu_c / (c + \nu_c)$, $\phi_1 = 1 - T_e$, where T_e is the fraction of entanglements that are trapped and cannot relax. In Figure 3, ϕ is plotted against ϕ_2 for several values of ϕ_1 . These parameters are not independently variable, since increased cross-linking increases ϕ_2 and diminishes ϕ_1 to a degree which depends also on the initial molecular length¹¹. Two representative cases may be chosen for illustration, suitable for 1,2-polybutadiene of molecular weight 1×10^5 , which has about 30 entanglements per molecule:

$$\nu_c/c = 1, \phi_1 = 0.075, \phi_2 = 0.50, \phi = 0.06, (\nu_c/c)_r = 1.033$$

$$\nu_c/c = 0.2, \phi_1 = 0.42, \phi_2 = 0.17, \phi = 0.26, (\nu_c/c)_r = 0.268$$

In Figure 4, λ_s and λ_{sr} are compared for these two cases as functions of λ_0 . Evidently, the change from first to second state of ease is relatively small, but if it can be measured it can test the hypothesis of untrapped entanglements which has been evoked to explain very slow relaxation mechanisms in lightly cross-linked networks^{13,14}.

Kinetics. The change from first to second state of ease is expected to be quite slow if the relaxing entanglements are on dangling branched structures, and the rate should decrease markedly with decreasing v_c/ϵ ; thus, the larger the effect, the more slowly it will be accomplished. Even qualitative observations should clarify the roles of trapped and untrapped entanglements. In particular, it has been assumed^{13,14} that a trapped entanglement contributes as much to the strain entropy change, and hence to stress, as a chemical cross-link (or differs by a constant effectiveness factor). The possibility that its effectiveness depends on both time and magnitude of deformation should be probed. It has been suggested¹⁵ that the effectiveness of trapped entanglements should decrease with increasing λ .

V. Swelling

Further information may be obtained from swelling the network in solvent after it has reached its first state of ease. (This experiment is primarily of interest for networks with sufficiently large v_c/ϵ that there is no further change to a second state of ease.) If the effectiveness of trapped

entanglements is unaffected by swelling, the dimensional changes should be isotropic. (In networks with two states of chemical cross-linking, slight anisotropies in swelling have been observed⁹ and attributed to deviations from neo-Hookean behavior, but they are minor.) If swelling diminishes the effectiveness of entanglements, the increase in length in the stretch direction should be greater than the cube root of the volume ratio, and from it a new value of effective entanglement density can be calculated by a modification of Equation 1.

Acknowledgment

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253. Data for the figures were evaluated numerically by Mr. Robert Marrone.

References

1. J. D. Ferry, Proc. 5th Intern. Cong. Rheology, Vol. 1, p. 3, University of Tokyo Press, 1969.
2. W. W. Graessley, J. Chem. Phys., 47, 1942 (1967).
3. W. W. Graessley, J. Chem. Phys., 54, 5143 (1971).
4. J. D. Ferry, "Viscoelastic Properties of Polymers", Second Edition, Wiley, New York, 1970, Chapter 13.
5. J. P. Berry, J. Scanlan, and W. F. Watson, Trans. Faraday Soc., 52, 1137 (1956).
6. P. J. Flory, Trans. Faraday Soc., 56, 722 (1960).
7. Reference 4, Chapter 14.
8. R. D. Andrews, A. V. Tobolsky, and E. E. Hanson, J. Appl. Phys., 17, 352 (1946).
9. A. Greene, K. J. Smith, Jr., and A. Ciferri, Trans. Faraday Soc., 61, 2772 (1965).
10. K. J. Smith, Jr., A. Ciferri, and J. J. Hermans, J. Polymer Sci., A2, 1925 (1964).
11. H. Leaderman, Trans. Soc. Rheology, 6, 361 (1962).
12. B. Bernstein, E. A. Kearsley, and L. J. Zapas, Trans. Soc. Rheology, 7, 391 (1963).
13. N. R. Langley, Macromolecules, 1, 348 (1968).
14. N. R. Langley and J. D. Ferry, Macromolecules, 1, 353 (1968).
15. B. M. E. van der Hoff, J. Macromol. Sci. (Chem.), A1, 747 (1967).

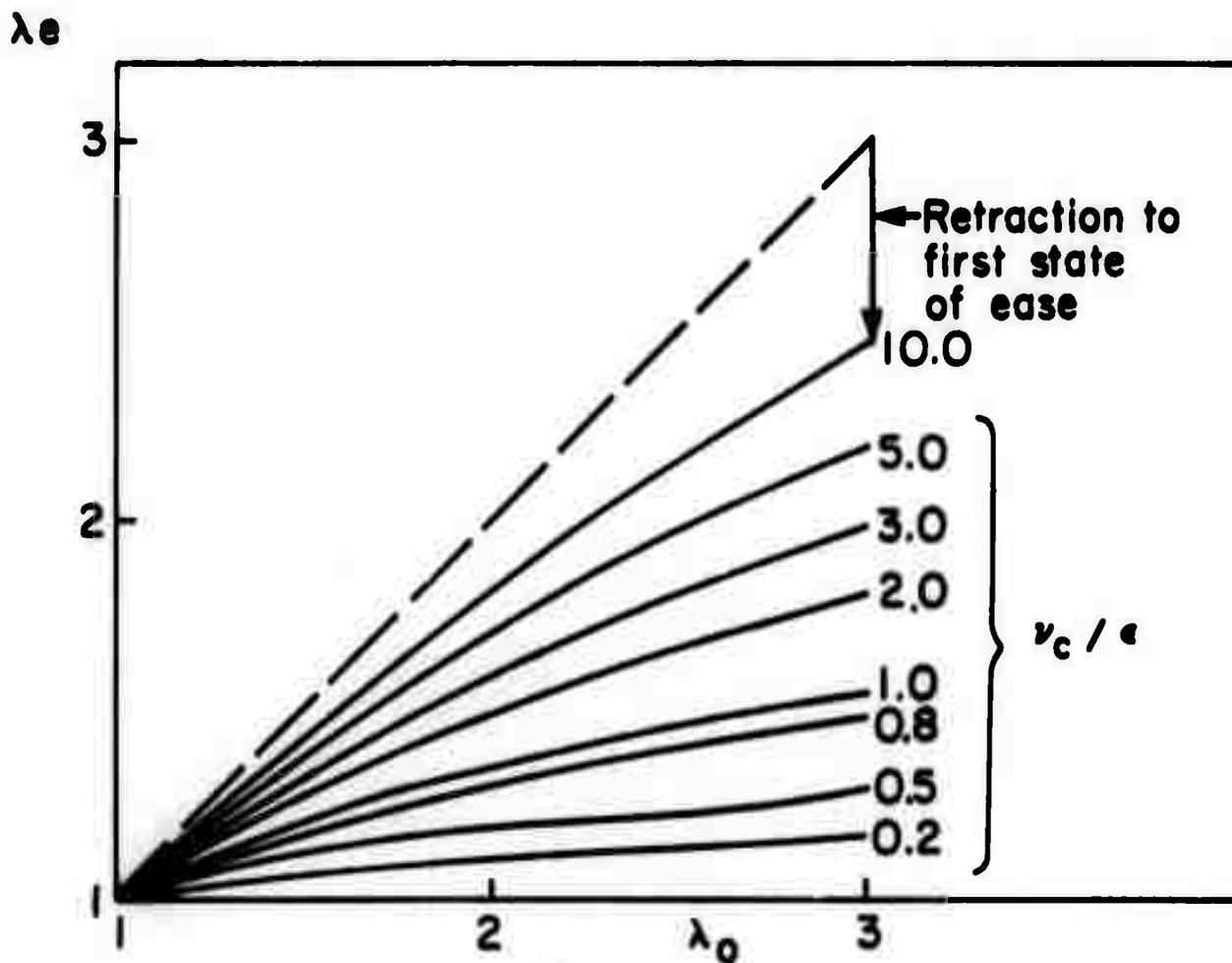


Figure 1. Relative length in first state of ease for simple extension plotted against relative length λ_0 at which cross-links are introduced, for various values of ν_c / ϵ .

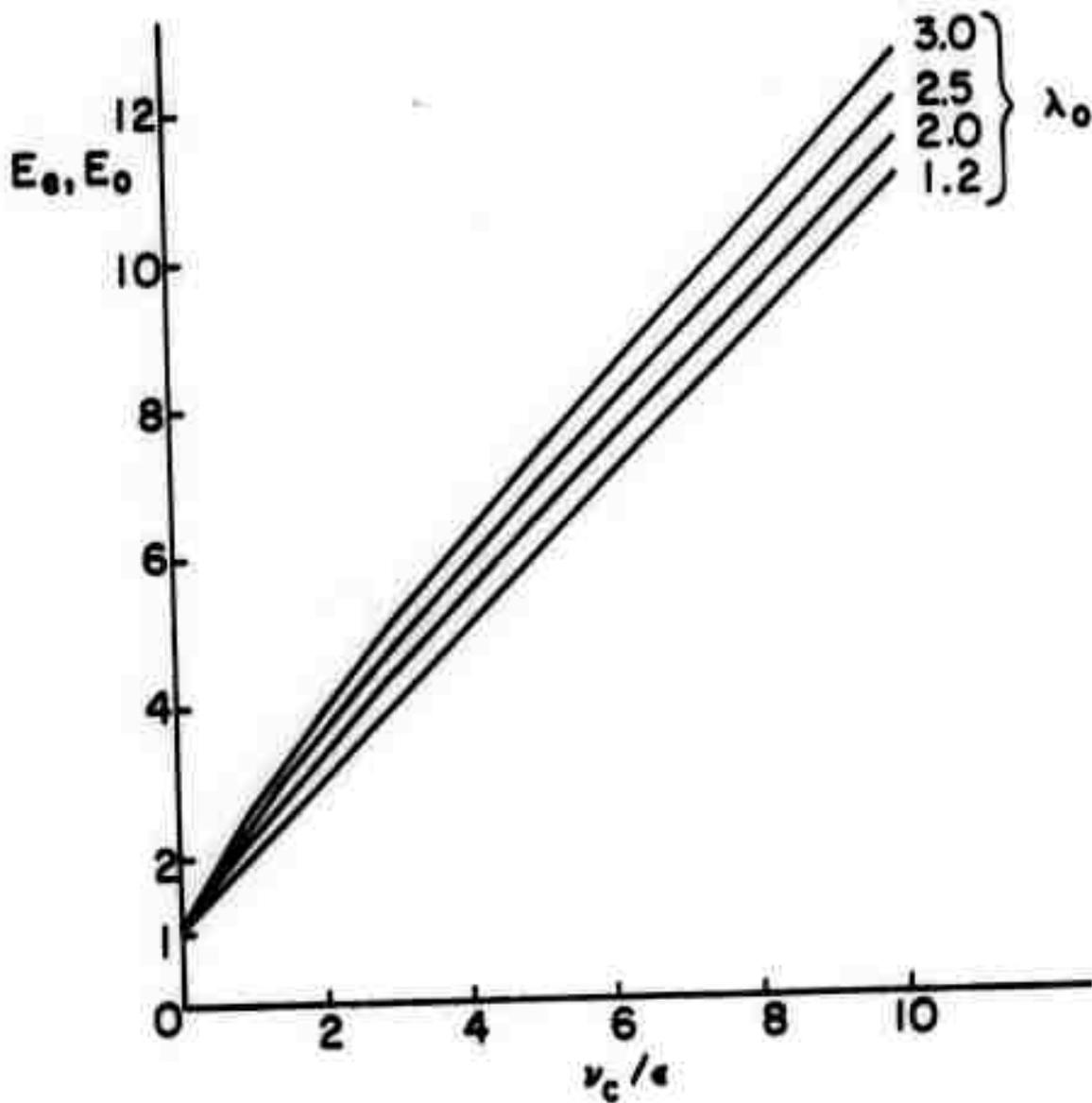


Figure 2. Ratio of Young's modulus in first state of ease to plateau modulus before cross-linking, plotted against ν_c / ϵ for various values of λ_0 .

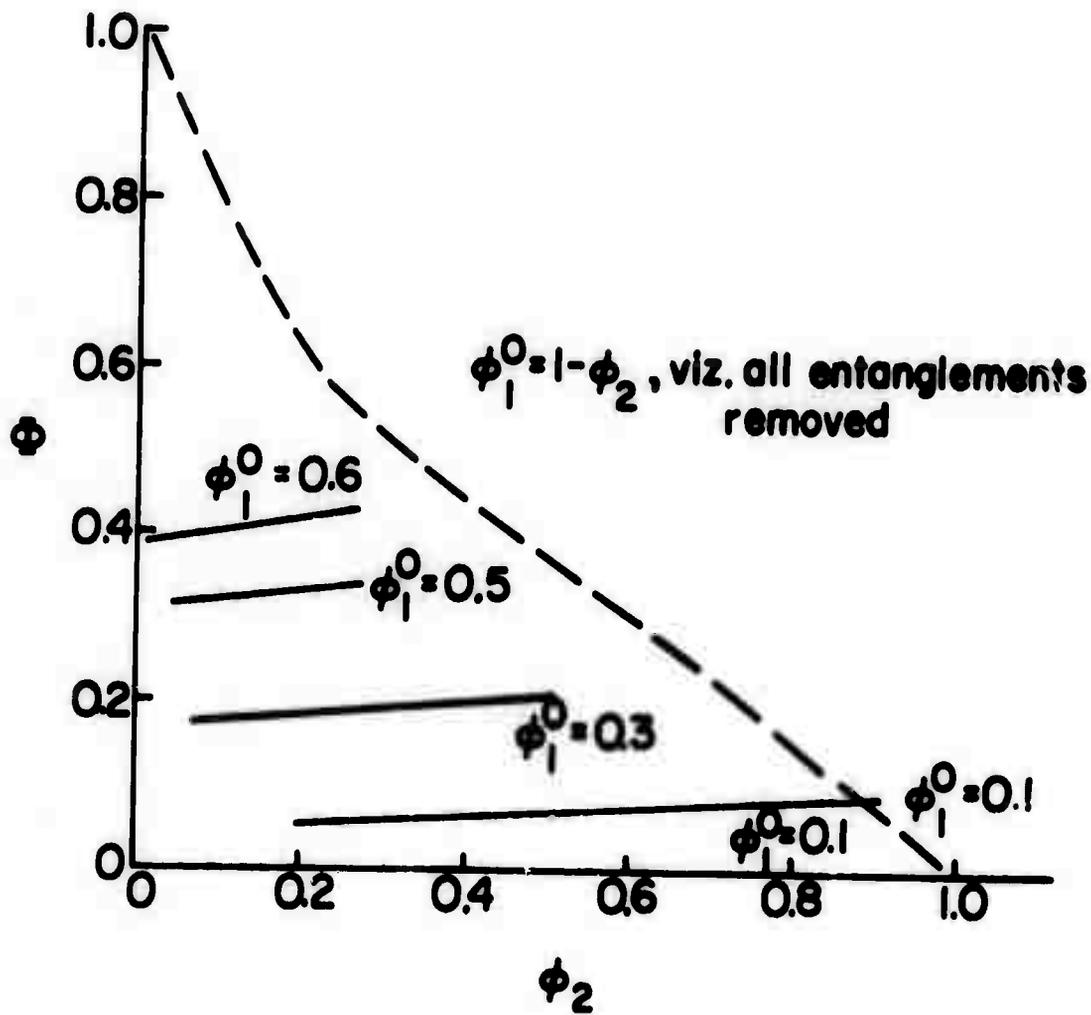


Figure 3. Parameter ϕ describing effect of entanglement relaxation, plotted against ϕ_2 , for various values of ϕ_1^0 (see text).

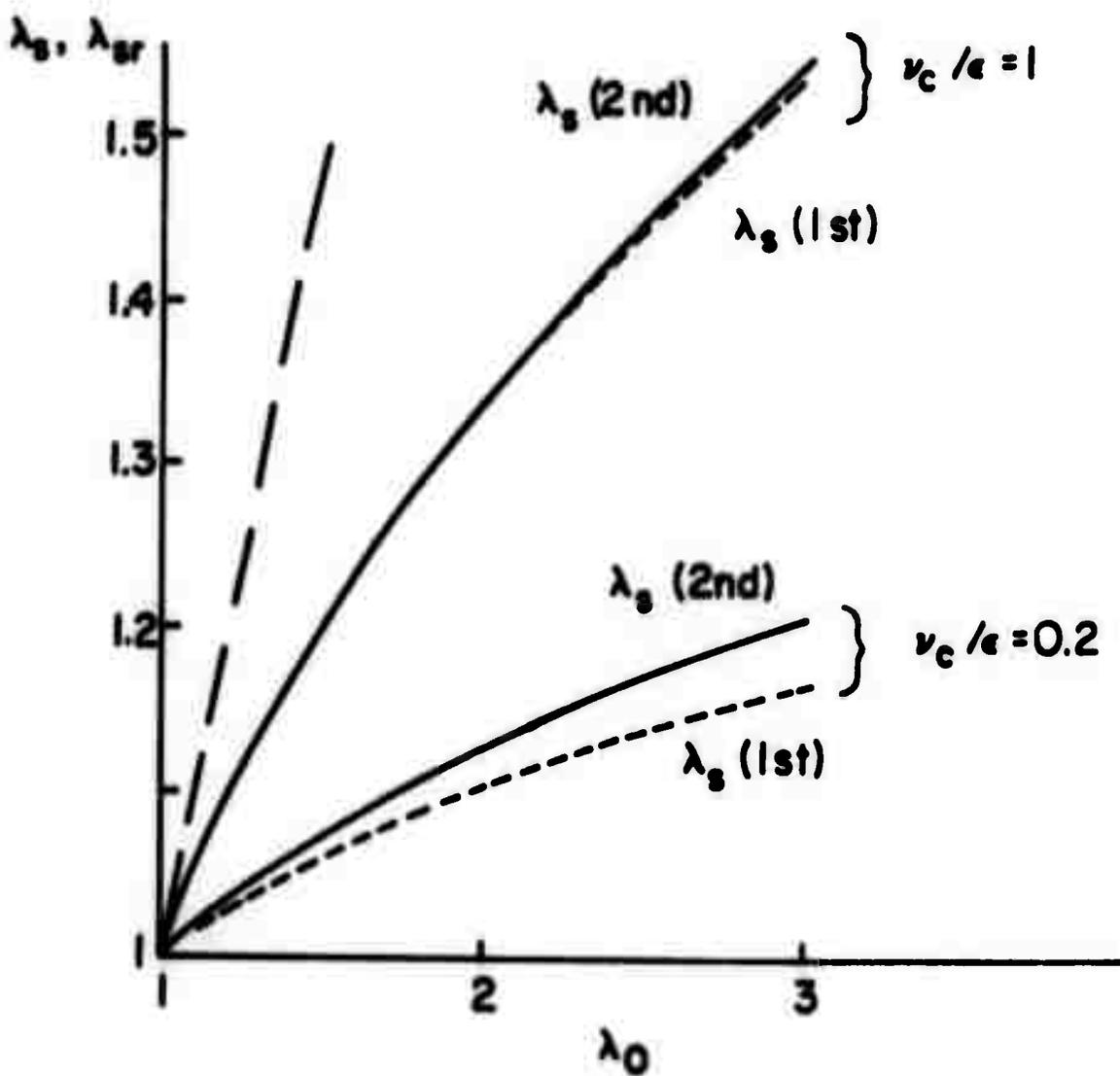


Figure 4. Comparison of first and second states of ease for two illustrative choices of ν_c/ϵ , ϕ_2 , and ϕ_1 (see text).

DIFFUSION THROUGH COMPOSITE
POLYMER SYSTEMS

John D. Ferry

Abstract

The factors influencing the diffusion coefficient for a small molecule through a composite structure of permeable (polymeric) and impermeable domains are discussed. Calculations of the geometrical structure factor based on various models with a continuous permeable phase are reviewed; they are similar enough so that deviations of experimental data therefrom may be interpreted in terms of other features such as modification of the local molecular mobility in the continuous phase, failure of contact between phases, and gross alterations of morphology.

Preceding page blank

DIFFUSION THROUGH COMPOSITE POLYMER SYSTEMS

John D. Ferry

I. Introduction

Reinforcement of rubbery polymers by incorporation of solid particles has long been an essential feature of rubber technology, and in recent years composite polymer systems, in which two or more polymer components with very different mechanical properties are blended in various ways, have become very important. The relation of domain structure and other aspects of morphology to physical properties is a problem of high priority in the polymer industry.

Since the diffusion of small foreign molecules (molecular weight of the order of 200) through polymer systems can be easily measured by use of radioactively tagged penetrants in trace amounts¹⁻³, it is of interest to examine the possibility of obtaining information from this type of experiment. In a homogeneous rubbery polymer, the diffusion coefficient of a penetrant such as n-hexadecane reflects the local mobility of polymer segments. In a rubber filled with particles of carbon black or calcium carbonate, or a block copolymer with domains of glassy polystyrene and rubbery polybutadiene, the hard or glassy domains can be considered impermeable to such a penetrant. The ratio D/D_0 of the diffusion coefficients through the composite system and through the unmodified rubbery polymer will

depend on (1) the geometrical effects of inclusion of impermeable domains, (2) possible modification of the local molecular mobility of the rubbery polymer in the vicinity of the impermeable domains, (3) interface anomalies, especially failure of contact or existence of voids.

II. Geometrical Effects of Impermeable Domains

Many theoretical treatments of suspensions of impermeable particles in an impermeable medium have been reviewed by Barrer⁵. We choose a number which are of possible interest and compare the dependence of D/D_0 on the volume fraction of the impermeable component, v_1 .

- (1) Cubical particles in a simple cubic array⁵:

$$D/D_0 = 1 - v_1^{2/3}$$

- (2) Spherical particles in a simple cubic array (de Vries³):

$$D/D_0 = \frac{2 - 2v_1 + (3\alpha/4)v_1^{10/3}}{2 + v_1 + (3\alpha/4)v_1^{10/3}}, \quad \alpha = 1.31$$

- (3) Spherical particles in a body-centered cubic array⁵: same as (2) with $\alpha = 0.129$

- (4) Spherical particles in a face-centered cubic array⁵: same as (2) with $\alpha = 0.075$

- (5) Elongated ellipsoidal particles, one-third oriented parallel to each principal direction (Rayleigh⁶, Barrer⁵):

domain interfaces there may be a mixture of segments⁹. In a structure with continuous B and discontinuous S domains, diffusion may reveal that the molecular mobility is lower in such interfaces than in the B phase. Interfacial mixing would be presumably enhanced by progressive shortening of the block lengths, down to a limit of random copolymer¹⁰. If the diffusion constant in the random copolymer is known as a function of S content, it may be possible to make deductions concerning the composition and extent of the interfacial zone. (The microstructure of the polybutadiene blocks must also be known, since D is influenced by this variable as well.)

IV. Interface Anomalies

If adhesion between a polymer and filler particles is poor, voids may exist. These apparently enhance the permeability to small gas molecules⁴, but their effect with the type of penetrant discussed here is uncertain. When subjected to large deformations, filled rubbers may experience substantial detachment of the particles from the matrix (Mullins effect¹¹). The effect of this phenomenon on diffusion might be explored.

V. Changes in Morphology

In block copolymers, the connectivities of the two phases depend greatly upon the nature of the solvent from which a solid structure is deposited by evaporation¹², and upon subsequent annealing¹³. With increasing S content in SBS copolymers, the S domains may join through tapered leg structures¹⁴ without at first interrupting the connectivity

of the B domains. When such materials are subjected to large deformations, gross alterations in morphology occur¹². These features can also be investigated by diffusion measurements.

Acknowledgment

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

References

1. I. Auerbach, S. D. Gehman, W. R. Miller, and W. C. Kuryla, *J. Polymer Sci.*, 28, 129 (1958).
2. R. S. Moore and J. D. Ferry, *J. Phys. Chem.*, 66, 2699 (1962).
3. S. P. Chen and J. D. Ferry, *Macromolecules*, 1, 270 (1968).
4. R. M. Barrer, in "Diffusion in Polymers," edited by J. Crank and G. S. Park, Academic Press, 1968, Chapter 6.
5. D. A. de Vries, *Med. Landbouwhogeschool Wageningen*, 52(1), 1 (1952).
6. Lord Rayleigh, *Phil. Mag.*, 34, 481 (1892).
7. S. Prager, *J. Chem. Phys.*, 33, 122 (1960).
8. W. P. Slichter, D. D. Davis, and S. Kaufman, *Bull. Amer. Phys. Soc.*, Ser. II, 16, 365 (1971).
9. D. H. Kaelble, *Trans. Soc. Rheology*, 15, 235 (1971).
10. T. Kotani and Y. Kuwabara, *Kōgyō Kagaku Zasshi*, 74, 151 (1971).
11. L. Mullins, *J. Rubber Res.*, 16, 275 (1947); *J. Phys. Colloid Chem.*, 54, 239 (1950).
12. J. Moacanin, G. Holden, and N. W. Tschoegl, Editors, "Block Copolymers," *J. Polymer Sci.*, C26, (1969).
13. N. W. Tschoegl, private communication.
14. J. F. Beecher, L. Marker, R. D. Bradford, and S. L. Aggarwal, *J. Polymer Sci.*, C26, 117 (1969).

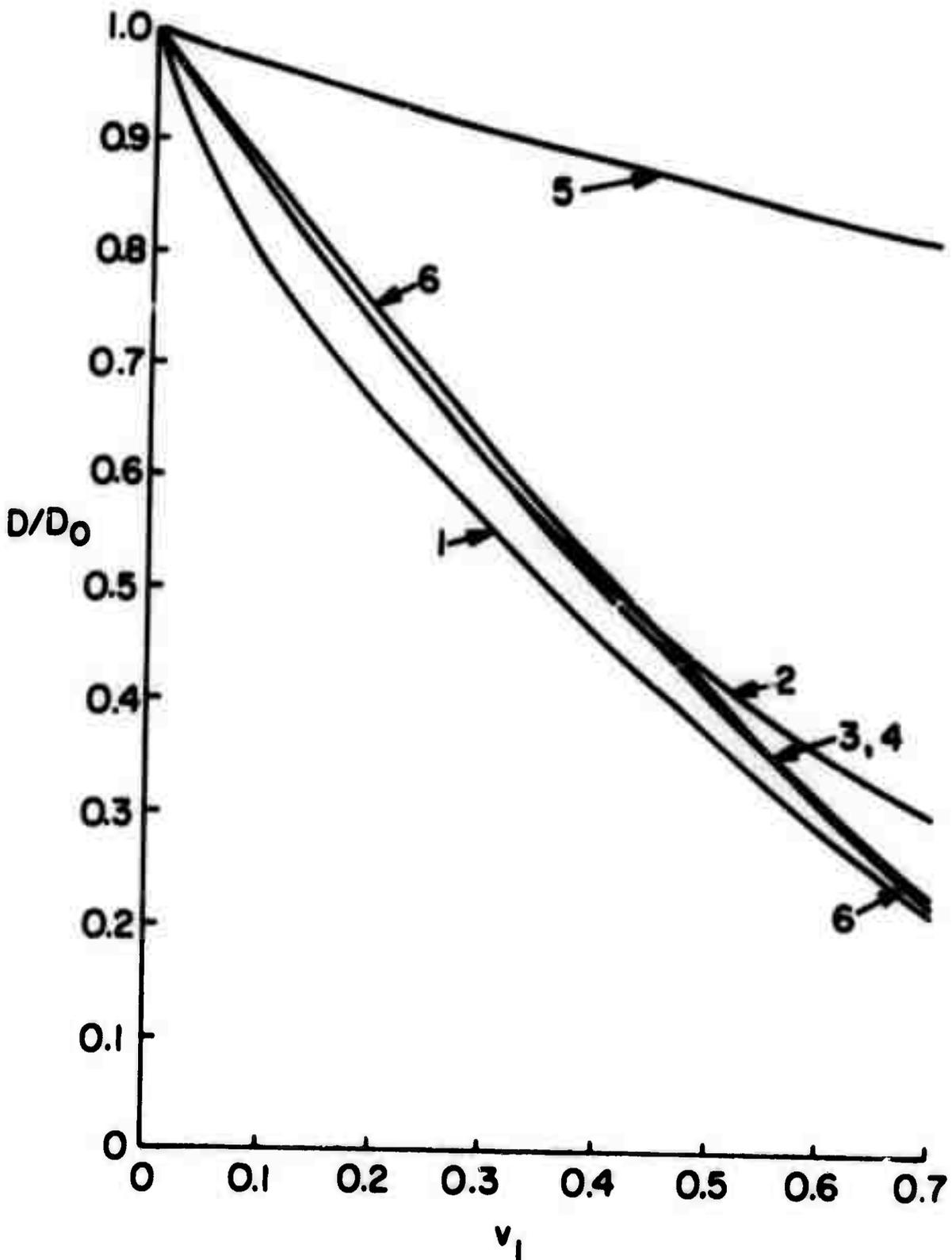


Figure 1. Ratio of diffusion coefficients in composite system and in the permeable matrix, plotted against volume fraction of impermeable inclusions, numbered to correspond to the 6 equations in the text.

SOME PROBLEMS IN BULK POLYMERIC SYSTEMS

H. Reiss

Abstract

A condensed summary of some of the statistical techniques used in the study of polymer chain configurations is presented.

In connection with the theory of polymers containing rings, a new variation principle is developed which can be used in connection with associated configurational problems. This variation principle leads to a "self-consistent field" solution of the configurational problem. The theory of rubber elasticity is discussed, and the possibility of applying the variation principle to this phenomenon considered.

Glassy polymers and the glass transition are discussed, and some ideas connecting the glass transition and what appears to be a related phenomenon involving a fluid-solid transition observed in a system of hard spheres are considered. The hard sphere transition is really a glass-avoiding transition for which we have coined the term "hyalo-phobic transition".

Block copolymers and their participation in the formation of heterophase polymer bulk polymers are also dis-

cussed. It is suggested that the fairly well developed field of the statistical thermodynamics of curved surfaces might be of use in connection with the heterophase transition.

Inhomogeneous bulk polymers resulting from the formation of inhomogeneous networks are discussed. Some corrections are applied to the "gel ball" theory of Labana, Newman, and Chomppff. Polymer crystals are also treated.

Finally, frictional and dynamical effects in bulk polymers and polymer solutions are considered. The Rouse model is outlined and possible future directions are indicated.

SOME PROBLEMS IN BULK POLYMERIC SYSTEMS

H. Reiss

I. Introduction

Polymers have provided useful materials of application for many years. The most important have usually involved the polymer in the "bulk" rather than the "free molecular" state, i.e., in the glassy, rubbery, or even in the crystalline state. In spite of the technological importance of the bulk state, most of the "science" of polymers has concerned the free or quasi-free molecular state, e.g., polymers in solution. The reasons for this are clear--the molecular state is more easily amenable to quantitative theoretical analysis.*

In spite of the lesser technological importance of the molecular state, research in this area has been important for several reasons, among which is the fact that polymers are best characterized in solution. Measurements of molecular weight, polymer configuration, reactivity, and relaxation mechanisms are more easily performed in solution. After having been characterized in this manner, the polymers can then be converted into various bulk conditions whose properties are then related to the known molecular features. Of course, certain features have no independent existence outside of the bulk state and it is in

*A comprehensive and excellent description of the theory of polymers in solution can be found in H. Yamakawa, *The Modern Theory of Polymer Solutions* (Harper and Row, New York, 1971).

these areas that much new research needs to be done.

Another bequest of solution research is the set of methodologies connected with the molecular theory of polymers. These methods can be appropriated and extended to the treatment of polymers in bulk. Progress in this direction has already been made.

Unlike other fields of materials science, the polymer area has remained almost the exclusive property of chemists and chemical engineers. This point is emphasized by the fact that ARPA itself has supported very little work in polymers. On the other hand, there are many points of possible contact with physicists, especially solid state physicists. This is even more true for polymers in bulk. One of the goals of this somewhat informal (and perhaps disorganized) discussion of certain aspects of the field will be to acquaint the physicist with a few problems where his help would be both appreciated and effective.

The following text contains several streams of ideas. These include (1) a somewhat qualitative, and at best semi-quantitative, description of some conventional well established ideas in polymer science; (2) some research ideas -- carried through initial steps of activation -- which have emerged from a month of summer study; (3) designation of those ideas and researches which may be of interest to physicists. These categories will not always be clearly separated, but appropriate comments will be made wherever clarification seems necessary.

II. The Configuration of Polymer Molecules

The classic approach to the theory of polymer molecule configuration has involved representation of the molecule by a number, n , of freely orienting segments where n need not be identical with P , the degree of polymerization which denotes the number of monomer groups bound together to form the polymer. With this model the allowable configurations are in one to one correspondence with the paths of a random flight in which the individual steps are all of length a , a being the length of a segment.

It follows¹ that the end to end distance, \vec{R} , of the polymer has the distribution

$$P(\vec{R}) = C \left(\frac{3}{2\pi na^2} \right)^{3/2} e^{-\frac{3R^2}{2na^2}} \left\{ 1 + \frac{3R^2}{2n^2 a^2} - \frac{9R^4}{20m^3 g^4} + \dots \right\}, \quad (1)$$

where $P(\vec{R}) d\vec{R}$ is the probability that the end to end distance lies in $d\vec{R}$ and C is a normalizing constant. For large n it proves sufficient to approximate $P(\vec{R})$ by the Gaussian prefactor in eq. (1), and after normalization,

$$P(\vec{R}) = \left(\frac{3}{2\pi na^2} \right)^{3/2} e^{-\frac{3R^2}{2na^2}}. \quad (2)$$

For this model the mean square end to end distance proves to be

$$\langle R^2 \rangle = na^2 \quad (3)$$

and is linear in the number of segments, n .

Another configurational quantity more closely related to experimental measurement (for example, to light scattering is the so-called radius of gyration, s , which is the root mean square distance of a polymer segment (averaged over all segments) from the center of mass of the molecule. The mean square radius of gyration is obtained by averaging the square of the radius of gyration over all configurations. Simple analytical expressions for the distribution of the radii of gyration are not available. However, it can be shown that the distribution is approximately gaussian.³ A good approximate closed expression for $P(S)$, where S is the radius of gyration, is due to Flory and Fisk⁴ and has the form

$$P(S) = \frac{K}{\langle S^2 \rangle^{3/2}} \left(\frac{6S^2}{na^2} \right)^m e^{-(m+1/2) \frac{6S^2}{na^2}} \quad (4)$$

The quantity m is usually set equal to 3.

Although a simple analytical expression for $P(S)$ is not available, formulas for the various moments of S can be derived with relative ease. For example, for the molecule with freely orienting segments, the mean square radius of gyration is

$$\langle S^2 \rangle = \frac{1}{6} na^2 = \frac{1}{6} \langle R^2 \rangle \quad (5)$$

so that the mean square radius of gyration is 1/6 the mean square end to end distance.

The configurational statistics of polymer molecules can be developed expeditiously through the application of

Markoff's method⁵. This method employs the Fourier representation of the three-dimensional delta function and may be applied to the distribution of a sum of vectors

$$\vec{\phi} = \sum_{i=1}^M \vec{\phi}_i \quad , \quad (6)$$

where the $\vec{\phi}_i$ are vector functions of the individual bond vectors. Thus, $\vec{\phi}_i(\vec{r}_i)$ is a function of \vec{r}_i where \vec{r}_i is the vector representing the i th segment of the polymer. If \vec{r}_i itself has a distribution $\tau_i(\vec{r}_i)$ then it may be shown that $P(\vec{\phi})$, such that $P(\vec{\phi})d\vec{\phi}$ is the probability that $\vec{\phi}$ lies in $d\vec{\phi}$, is given by

$$P(\vec{\phi}) = \frac{1}{(2\pi)^3} \int K(\vec{\rho}) e^{-i\vec{\phi} \cdot \vec{\rho}} d\vec{\rho} \quad , \quad (7)$$

where, the Fourier transform of $P(\vec{\phi})$, $K(\vec{\rho})$, is given by

$$K(\vec{\rho}) = \prod_{i=1}^M \int \tau_i(\vec{r}_i) e^{i\vec{\rho} \cdot \vec{\phi}_i} d\vec{r}_i \quad , \quad (8)$$

where it will be recalled that $\vec{\phi}_i$ is a function of \vec{r}_i . In particular $\vec{\phi}_i$ can be \vec{r}_i so that $\vec{\phi}$ becomes the end to end distance.

Markoff's method has been extended by Wang and Uhlenbeck⁶ to multivariate distributions in the particular case that

$$\tau_i(\vec{r}_i) = \left(\frac{3}{2\pi a_i^2} \right)^{3/2} e^{-\frac{3r_i^2}{2a_i^2}} \quad , \quad (9)$$

i.e., to the case in which each bond or segment is gaussian distributed. This method is applicable to sets of vectors $\vec{\phi}_k$ such that

$$\vec{\phi}_k = \sum_j \phi_{kj} \vec{r}_i \quad , \quad (10)$$

where the ϕ_{kj} are constants, i.e., to vectors which are linear combinations of the bond or segment vectors. If there are s vectors in the set and the set is denoted by $\{\phi_s\}$ then the simultaneous or multivariate distribution of this set is given by

$$P(\{\vec{\phi}_s\}) = \left(\frac{3}{\pi \langle a^2 \rangle} \right)^{3s/2} \frac{1}{|\zeta|^{3/2}} e^{-\left(\frac{3}{2 \langle a^2 \rangle |\zeta|} \right) \sum_{k=1}^s \sum_{\ell=1}^s C^{k\ell} \vec{\phi}_k \cdot \vec{\phi}_\ell} \quad (11)$$

In this formula ζ is the matrix with $k\ell$ element

$$C_{k\ell} = \sum_{j=1}^m \phi_{kj} \phi_{\ell j} \frac{a_j^2}{\langle a^2 \rangle} \quad , \quad (12)$$

where $\langle a^2 \rangle$ is the mean square bond length

$$\langle a^2 \rangle = \frac{1}{n} \sum_{j=1}^n a_j^2 \quad . \quad (13)$$

Furthermore, $|\zeta|$ is the determinant of ζ and $C^{k\ell}$ is the cofactor of the element $C_{k\ell}$ in this determinant.

Eq. (11) has many uses. It can be used expeditiously for the evaluation of the distribution of the radius of gyration since if \vec{S}_j is the distance of a segment from the molecular center of mass the radius of gyration, S , is given by

$$S^2 = \frac{1}{n} \sum_{j=1}^n S_j^2 \quad . \quad (14)$$

It may be shown⁷ that

$$\vec{S}_j = \sum_{i=1}^m \phi_{ji} \vec{r}_i \quad (15)$$

with

$$\phi_{ui} = u(j-i) \frac{i}{n+1} - 1 \quad , \quad (16)$$

where $j(j-i)$ is the unit step function. Thus the \vec{S}_j are of the form exhibited in eq. (10). Because of this, both S^2 and $\langle S^2 \rangle$, (where the brackets indicate configurational averaging so that $\langle S^2 \rangle$ is the mean square radius of gyration) can be evaluated using eq. (11).

The bond distribution, eq. (9) is not a serious restriction since it may be shown⁸ that as long as gaussian statistics are applicable to the overall chain configuration the use of

$$\tau = \delta(\vec{r} - \vec{a}) \quad , \quad (17)$$

or

$$\tau = \left(\frac{3}{2\pi a^2} \right)^{3/2} e^{-3r^2/2a^2} \quad , \quad (18)$$

where $a^2 = \vec{a} \cdot \vec{a}$, lead to the same formulas.

It can be shown⁹ that the effective freely orienting segment model is an accurate representation of any chain in which the correlation between successive actual segments extends only over a finite number of such segments. Under this circumstance, one can always find an effective segment length so that the freely orienting model is applicable. On the other hand, by its very nature, such a model conceals most of the effects which are specific to the chemical structure of the particular polymer

under consideration. Furthermore, the actual correlation between successive segments involves energetic effects and so one cannot properly account for the thermodynamics of chain molecules without its explicit consideration. For example, the freely orienting molecule involves entropy only and on the average adopts a coiled form with a short end to end distance because so many more configurations are available to it in that state. In fact, the elastic modulus of the molecule reflects these entropy requirements, rather than being rooted in potential energy. The molecule behaves like an entropy spring!

Recently, considerable work has been done, especially by Flory¹⁰, in an effort to treat chain molecules on an actual segment basis. The technique involves introduction of the actual potentials which determine the relative orientations of successive segments. The problem is reducible to a form which is an extension of the Ising model of the one-dimensional ferromagnet. Averages involving the configurations of the polymer are obtained in much the same way that the average magnetization is obtained in the case of the ferromagnet. The "matrix method" for the treatment of the ferromagnet, first introduced by Kramers and Wannier can be used very expeditiously in the present case.

None of the foregoing methods is capable of dealing with the so-called "excluded volume" problem. This problem, of central importance to many aspects of polymer behavior, may be

described as follows. The freely-orienting segment model possesses statistics identical with those of the problem of random flights. However, the path of a given random flight may cross itself any number of times. In contrast, with a real polymer, since two segments cannot simultaneously occupy the same position in space, a configuration corresponding to such a "crossing path" must be excluded. At first sight this effect may seem negligible since the polymer configurations are known to be quite open, and therefore the number of paths excluded quite small. However, the polymer molecule is very long and it turns out that the chance of getting at least one crossing in such a long path is appreciable. Since even one crossing is grounds for exclusion, the number of excluded paths becomes in fact quite large, and both the mean square end to end distance and radius of gyration may depart appreciably from the values predicted by the theory of random flights.

Many authors have attempted to treat the excluded volume effect. Early treatments which did not account for correlations extending over the entire length of the chain succeeded only in deducing that, for very long chains, the mean square end to end distance will still be proportional to the number of segments.

The most reliable technique for examining the excluded volume problem is founded on the canonical distribution of configuration given by statistical mechanics. Thus, if R_i represents the location of the i th segment, a potential is

introduced and denoted by

$$W(\{\vec{R}_n\}) = \sum_{ij} u(\vec{R}_{ij}) \quad (19)$$

where the sum is over all ij pairs of segments and \vec{R}_{ij} is the distance between the i th and j th segments. $\{\vec{R}_n\}$ stands for the set of \vec{R}_i . When the molecule is in solution this potential is actually a potential of "mean force" and therefore is itself a free energy, subsuming various solvent effects. The pairwise additivity indicated in eq. (19) is then only valid when the segment density is low (a condition approximated by the open structures of most polymers). The potential which ties the segments into a sequential chain is not included in eq. (19). This can be written as

$$W_0 = \sum_i u_0(\vec{R}_{i,i+1}) \quad (20)$$

the sum on the right including interactions between a given segment and its successor only. The probability of chain configuration is then given by

$$P(\vec{R}_N) = \frac{1}{Z} \exp\left(-\frac{1}{kT}(W_0(\{\vec{R}_N\}) + W(\{\vec{R}_N\}))\right) \quad (21)$$

in which Z is the configurational partition function.

$$Z = \int \exp\left(-\frac{1}{kT}(W_0+W)\right) d(\{\vec{R}_N\}) \quad (22)$$

$$\text{The quantity } e^{-W_0/kT} = Z_0 P_0(\{\vec{R}_N\}) = \prod_{i=1}^n \tau_i(\vec{r}_i) \quad (23)$$

where Z_0 is the configurational partition function with W suppressed, and thus $e^{-W_0/kT}$ is proportional to P_0 the probability of configuration for the freely orienting chain.

Approximations to configurational statistics including the excluded volume effect have been derived in a number of ways, beginning with eq. (21). One of the earliest approaches¹¹ (and still one of the most successful) is due to Flory who wrote eq. (21) in the form

$$P(S) = \frac{1}{Z} P_0(S) e^{-V(S)/kT} \quad (24)$$

where S is the radius of gyration. Actually, eq. (24) represents a partial integration of eq. (21) over all configurations $\{\vec{R}_N\}$ consistent with a given S . Of course Z is now given by

$$Z = \int P_0(S) e^{-V(S)/kT} dS \quad (25)$$

For the freely orienting chain function, P_0 , the Gaussian form

$$P_0(S) = \left(\frac{3}{2\pi \langle S^2 \rangle_0} \right)^{3/2} e^{-3S^2/2\langle S^2 \rangle_0} 4\pi S^2 \quad (26)$$

is assumed. The basis of the Flory theory is then the assumption that $P(S)$ as well as $P_0(S)$ is of a Gaussian form with parameters different from those of $P_0(S)$. With this assumption $V(S)$, which is really the free energy of a polymer constrained to have the radius of gyration S , can be computed in terms of the parameters of this new Gaussian distribution. Then, in effect, the free energy over all configurations is minimized by finding the maximum of Z , given by eq. (25), with respect to those parameters.

This extremalization fixes the new parameters, and therefore $P(S)$. Flory's method, involving as it does, a variation principle on the free energy is really an example of the self-consistent field method discussed more thoroughly in the next section. The result predicts a general increase of the average polymer dimension, and a mean square radius of gyration, or mean square end to end distance, going as

$$\langle R^2 \rangle \quad \text{or} \quad \langle S^2 \rangle \propto n^{6/5}, \quad n \rightarrow \infty \quad . \quad (27)$$

Thus the ratio of mean square radius of gyration, with excluded volume, to that without excluded volume, should go as $n^{1/5}$; a very serious departure as $n \rightarrow \infty$.

Other approaches^{1,2}, following Flory have included the application of perturbation methods. Several such approaches are analogs of the Ursell-Mayer "cluster theory" approach to the theory of imperfect gases. These theories, beginning with the canonical ensemble, endeavor to expand $P(\{\vec{R}_N\})$ in a power series in the degree of polymerization P . All of them expose even in their lowest order terms, the true many body aspect of the problem and the fact that $\langle R^2 \rangle / n$ increases with n .

Numerical computations of nonintersecting walks on lattices have also been performed using computers¹³. For three-dimensional lattices (irrespective of type) eq. (27) seems to be indicated as $n \rightarrow \infty$. Thus the ratio $n^{1/5}$ mentioned above seems to be connected to the dimensionality of the walk rather than to the specific nature of $w(\{\vec{R}_N\})$.

The most modern theories applicable at $n \rightarrow \infty$ are of the self consistent field variety mentioned above. Since this is in an area currently subject to vigorous research and since it also represents an especially appropriate point of contact for physicists, we discuss it in more detail in the following section. Here we merely note that the results of such calculations also yield eq. (27) so that the validity of that result receives further support.

As a final topic, in this section, we discuss problems which arise with polymers which are branched or which may contain rings. Branched polymers are usually classifiable not only by their degrees of polymerization, but also by the number of branch points, the functionalities (number of chains emanating from a given branch point) of the branch points, the distribution of chain lengths between branch points, and, in general, the connectivity of the network. For branched or ring-containing polymers it is obvious that the mean square end to end distance has little meaning, and that we must concentrate on the radius of gyration.

Usually interest centers on a quantity

$$g = \langle S^2 \rangle_b / \langle S^2 \rangle_l \quad (28)$$

which is the ratio of the mean square radius of the branched molecule to that of a linear chain possessing the same number, n ; of segments. The quantity, g , can be calculated for branched molecules (containing no rings) in the special cases in which

either (1) all chains between branches have the same length, or (2) chain lengths between branches are distributed completely at random. In the first case (uniform distribution) if there are m branch units of functionality, f , and p subchains.

$$g = \frac{3p-2}{p^2} + \frac{1}{p} (f-1)^2 m(m^2-1) \quad (29)$$

For the second case (random distribution)

$$g = \frac{6}{p(p+1)(p+2)} [p^2 + h(f,m)] \quad (30)$$

where

$$h(f,m) = \frac{1}{2} (f-1)^2 m(m-1) \sum_{v=1}^{m-1} \frac{(fm-m-v)!(m-2)!}{(fm-m)!(m-v-1)!} \{ (f-1)^{v-1} (fv-2v+f)v \} \quad (31)$$

For molecules with rings g is much more difficult to compute and we make some suggestions in this respect in the next section.

III. Self Consistent Field Approaches to Excluded Volume and Molecules with Rings; Some New Research Ideas

The self-consistent field approach to polymer configuration can be derived from a variation principle not unlike the method employed in the derivation of the Hartree-Fock equations in quantum mechanics. However, since entropy contributions must be accounted for in the polymer case the variation principle involves finding an extremum of the free energy, instead of, as in the quantum mechanical case, the matrix element

of the Hamiltonian. It is in just this area where the methodologies of the field theorist and solid state physicist should find application and indeed one of the earliest practitioners, S. F. Edwards, is among other things a field theorist. Paradoxically, Edwards did not, however, derive his self-consistent field from a variation principle.

To understand this variation principle, suppose that the probability of finding the system in the i th quantum state is P_i . Then the entropy of the system is

$$S = -k \sum_i P_i \ln P_i \quad , \quad (32)$$

where the sum is over all quantum states accessible to the system. If E_i is the energy of the i th quantum state, then the internal energy of the system is

$$E = \sum_i P_i E_i \quad ; \quad (33)$$

and since the Helmholtz free energy is $\Lambda = E - TS$, we have

$$\Lambda = \sum_i (P_i E_i + kT P_i \ln P_i) \quad (34)$$

If we now vary the P_i so as to minimize Λ subject to no constraints other than those which maintain the E_i spectrum fixed, and the requirement,

$$\sum_i P_i = 1 \quad , \quad (35)$$

we find

$$P_i = \frac{e^{-E_i/kT}}{\sum_j e^{-E_j/kT}} \quad , \quad (36)$$

which is the usual distribution in the canonical ensemble.

Now it is well known that other ensembles and other expressions for P_i can be generated by adding additional constraints to the system. For example, the grand ensemble is obtained (in the case of open systems) when the chemical potential of a transferable species is fixed. In the case of a polymer molecule the constant force ensemble is obtained when the molecule is subjected to a constant tensile force.

Even more general constraints can be imposed. For example, in eq. (21) where the configuration (\hat{R}_N) may be thought of as corresponding to one of the quantum states with (\hat{R}_N) replacing i , so that $P((\hat{R}_N))$ takes the place of P_i we may insist the $W((\hat{R}_N))$ be restricted to the form

$$W((\hat{R}_N)) = \sum_i v_i(R_i) \quad (37)$$

rather than have the far more complicated form given by eq. (19). This would correspond to restricting $P((\hat{R}_N))$ to the form

$$P((\hat{R}_N)) = \frac{1}{Z} \prod_i \tau_i(\hat{r}_i) \phi_i(\hat{R}_i) \quad , \quad (38)$$

(ϕ_i not to be confused with ϕ_i appearing in eq. 16)

where now

$$Z = \int \prod_i \tau_i(\dot{R}_i) \phi_i(\dot{R}_i) d(\dot{R}_N) \quad (39)$$

Thus, $e^{-W(\{\dot{R}_N\})/kT}$ is now "constrained" to be represented by $\prod_i \phi_i(\dot{R}_i)$, a product of one segment distribution functions. This restriction is reminiscent of the one electron wave function appearing in the Hartree-Fock method. The functional forms of the ϕ_i or v_i can be determined by minimizing A in eq. (34) subject to constant spectrum E_i and the constraint that P_i is the product $\prod_i \phi_i$. The function v_i determined in this manner is the "self consistent" potential and gives rise to the "self consistent" field.

To be more precise, for the polymer case Reiss¹⁴ replaced eq. (34) by

$$A = \int [W_0(\{\dot{R}_N\}) + W(\{\dot{R}_N\})] P(\{\dot{R}_N\}) d(\dot{R}_N) + kT \int P(\{\dot{R}_N\}) \ln P(\{\dot{R}_N\}) d(\dot{R}_N) \quad (40)$$

where A is now just the configurational free energy and $W_0 + W$ is the exact configurational energy given by eqs. (19) and (20). However $P(\{\dot{R}_N\})$ is the "constrained" distribution function given by eq. (38). When the variation is performed, it turns out

$$v_i(\dot{R}_i) = \sum_{j \neq i} \int u(\dot{R}_{ji}) P(\dot{R}_{ji} | \dot{R}_i) d\dot{R}_j \quad (41)$$

where the sum is over all segments j except the i th, and

$P(\dot{R}_{ji}|\dot{R}_i)$ is the conditional probability that the i and j segments are separated by the vector distance \dot{R}_{ji} when it is known that the i th segment is located at \dot{R}_i . Since the $P(\dot{R}_{ji}|\dot{R}_i)$ can be expressed as functionals of the various $v_i(\dot{R}_i)$, eq. (41) is actually a set of equations for the self-consistent determination of the v_i .

At present this hierarchy of equations has not been solved. Instead a function $\hat{P}(\dot{R}_j)$ is defined as

$$\hat{P}(\dot{R}_j) = \int \prod_{i=1}^j v_i(\dot{r}_i) e^{-v_i(\dot{R}_i)/kT} \frac{d\dot{r}_i}{d\dot{R}_j} \quad , \quad (42)$$

where the notation $d\dot{r}_i/d\dot{R}_j$ implies integration keeping \dot{R}_j constant. Note that $\hat{P}(\dot{R}_j)$ will give the end to end distribution of the polymer as $j, n \rightarrow \infty$. In deriving the differential equation two assumptions are made. The first expresses $u(R_{ij})$ as

$$u(R_{ij}) = kT\epsilon(R_{ij}) \quad , \quad (43)$$

where ϵ is an appropriate segment interaction constant, and the second is the assumption that the distribution including excluded volume effects is still gaussian. The differential equation may be solved, and using this solution it is possible to compute the mean square end to end distance. As indicated earlier eq. (27) is once again recovered.

Edwards¹⁵ self consistent field was not derived from a variation principle and is equivalent to the assumption

$$P(\dot{R}_{ji} | \dot{R}_j) = P(\dot{R}_j) \quad . \quad (44)$$

He also is able to derive a differential equation for $P(\dot{R}_j)$, the solution of which again recovers eq. (27). Freed^{16,17} has shown that Edward self consistent field and the more general result of eq. (41) are successive approximations in a functional integral representation.

Much work still needs to be done along the lines of solving the hierarchy of equations specified by eq. (14), and represents a challenging task for the near future.

A new application of the self consistent field method, in connection with which some initial research has been performed during this month of study, will now be discussed. The application involves a solution (approximate) to the problem of computing the radius of gyration of a polymer molecule containing rings. The difficulties involved in this computation were mentioned at the end of the previous section.

Consider a polymer molecule containing rings and branches something like the example illustrated in Figure 1.



Figure 1.

As in the Figure, the branch points are located by the coordinate vectors \dot{R}_i . The branches are then labeled by the branch points which they connect. Thus the branch lying between \dot{R}_i and \dot{R}_j will be the ij branch. If two branches connect the same pair of branch points an additional index will be necessary to distinguish them. In the interest of a simple notation, however, we will assume in this exposition that such a complication does not arise. Now the number of segments in the ij branch will be denoted by n_{ij} and the total number of segments in the molecule will be n ,

$$n = \sum_{ij} n_{ij} \quad , \quad (45)$$

The coordinate (in some laboratory frame of reference) of the k th segment in the ij th branch will be denoted by $\dot{r}_{ij,k}$. The center of mass of the ij th branch will be \dot{c}_{ij} while the center of mass of the entire molecule is \dot{c} . Finally, the radius of gyration of the ij th branch is denoted by S_{ij} while the radius of gyration of the molecule will be S .

Within each branch the k th segment can be assigned a coordinate $\xi_{ij,k}$ relative to the center of mass of that branch. Thus

$$\xi_{ij,k} = \dot{r}_{ij,k} - \dot{c}_{ij} \quad , \quad (46)$$

where (if the segments are all assumed to have the same mass)

$$n_{ij} \dot{c}_{ij} = \sum_k \dot{r}_{ij,k} \quad , \quad (47)$$

from which it follows that

$$\sum_k \xi_{ij,k} = 0 \quad . \quad (48)$$

Now, from the definition of the radius of gyration, it follows that

$$n_{ij} S_{ij}^2 = \sum_k |\dot{\xi}_{ij,k}|^2 \quad . \quad (49)$$

Furthermore

$$n\dot{\vec{c}} = \sum_{ij} n_{ij} \dot{\vec{c}}_{ij} = \sum_{ij} \sum_k \dot{\vec{r}}_{ij,k} \quad . \quad (50)$$

We can also assign coordinates to every segment in the molecule relative to the molecular center of mass. Thus to the k th segment in the ij th branch we assign

$$\dot{\xi}_{ij,k} = \dot{\vec{r}}_{ij,k} - \dot{\vec{c}} \quad . \quad (51)$$

Then for the radius of gyration of the entire molecule we obtain the relation,

$$\begin{aligned} nS^2 &= \sum_{ij,k} |\dot{\xi}_{ij,k}|^2 \\ &= \sum_{ij,k} (|\dot{\vec{r}}_{ij,k}|^2 + |\dot{\vec{c}}|^2 - 2\dot{\vec{c}} \cdot \dot{\vec{r}}_{ij,k}) \quad . \end{aligned} \quad (52)$$

Eliminating $\dot{\vec{r}}_{ij,k}$ from this expression with the aid of eq. (46) and using eq. (48) we obtain

$$nS^2 = \sum_{ij,k} (|\dot{\xi}_{ij,k}|^2 + |\dot{\vec{c}}_{ij}|^2 - 2\dot{\vec{c}} \cdot \dot{\vec{c}}_{ij} + |\dot{\vec{c}}|^2) \quad . \quad (53)$$

Now from eq. (50)

$$\sum_{ij,k} \dot{c} \cdot \dot{c}_{ij} = \sum_{ij} \dot{c} \cdot n_{ij} \dot{c}_{ij} = \dot{c} \cdot \sum_{ij} n_{ij} \dot{c}_{ij} = n |\dot{c}|^2 \quad (54)$$

With this relation, the last two terms in eq. (53) may be expressed as

$$-2 \sum_{ij,k} \dot{c} \cdot \dot{c}_{ij} + \sum_{ij,k} |\dot{c}|^2 = -n |\dot{c}|^2 \quad (55)$$

Using eqs. (49) and (55) in eq. (53) we obtain

$$S^2 = \sum_{ij} \frac{n_{ij}}{n} (S_{ij}^2 + |\dot{c}_{ij}|^2) - |\dot{c}|^2 \quad (56)$$

Next we average over all configurations of the molecule keeping the locations of the branch points fixed. We denote this average by $\langle \rangle_R$. This, from eq. (56)

$$\begin{aligned} \langle S^2 \rangle_R &= \sum_{ij} \frac{n_{ij}}{n} \langle S_{ij}^2 \rangle_R + \sum_{ij} \frac{n_{ij}}{n} \langle |\dot{c}_{ij}|^2 \rangle_R \\ &\quad - \sum_{ij} \sum_{lm} \frac{n_{ij} n_{lm}}{n^2} \langle \dot{c}_{ij} \cdot \dot{c}_{lm} \rangle_R \end{aligned} \quad (57)$$

where we have used eq. (50) to rewrite the last term in eq. (56). Now $\langle S_{ij}^2 \rangle_R$ depends only on $|\vec{R}_j - \vec{R}_i|$. In fact, if the branch is a freely orienting chain, it can be shown that

$$\langle S_{ij}^2 \rangle_R = \frac{n_{ij} a^2}{12} \left(1 + \frac{|\vec{R}_j - \vec{R}_i|^2}{n_{ij} a^2} \right) \quad (58)$$

where a is the mean segment length. Furthermore, since with

the branch points fixed, the chance of observing a given configuration in one branch is independent of the configuration in another branch

$$\langle \vec{c}_{ij} \cdot \vec{c}_{lm} \rangle_R = \langle \vec{c}_{ij} \rangle_R \cdot \langle \vec{c}_{lm} \rangle_R \quad (59)$$

This result can be substituted into the last term on the right of eq. (57). From symmetry it is clear that

$$\langle \dot{\vec{c}}_{ij} \rangle_R = \frac{1}{2}(\dot{R}_j - \dot{R}_i) \quad (60)$$

Furthermore, $\langle |\dot{\vec{c}}_{ij}|^2 \rangle_R$ will also depend, only on R_i and R_j so that $\langle S^2 \rangle_R$ given by eq. (57) depends only on the coordinates of the branch points. The evaluation of the complete mean square radius of gyration, namely $\langle S^2 \rangle$, involves averaging $\langle S^2 \rangle_R$ over all configurations of the branch points. Thus all that is necessary for the computation of $\langle S^2 \rangle$ is the distribution giving the probability of configuration of the branch points. We develop a variation principle (self consistent field method) in order to approximate this distribution function.

First we note that if every branch consists of a freely orienting chain at low enough extension so that Gaussian statistics apply, the distribution function for the branch points can in principle be evaluated in closed form. Thus, if $p_{ij}(|\vec{R}_j - \vec{R}_i|) = p_{ij}(R_{ij})$ is gaussian function for the ij branch, the branch point distribution function

$$P(\{\vec{R}_B\}) d\{\vec{R}_B\} = K \prod_{ij} p_{ij}(R_{ij}) d\{\vec{R}_B\} \quad (61)$$

where K is a normalizing constant and \dot{R}_B refers to the coordinates of the set of branch points. When the p_{ij} are gaussian it is possible to evaluate

$$\frac{1}{K} = \int_{ij} p_{ij} d(\dot{R}_B) \quad (62)$$

explicitly, by a coordinate transformation which decomposes (62) into factor integrals each of which can be computed in closed form. The matrix diagonalization involved in this procedure however, depends sensitively on the connectivity of the molecule so that in most cases it is impractical to follow this route. Furthermore, the method fails for short branches where the assumption of gaussian behavior is invalid.

We therefore introduce an approximate method, based on a variation principle, which is capable of handling the non-gaussian case, and which, in spite of its dependence, still, on connectivity, may be easier to apply to some "average" molecule characterized only by the degree of polymerization and the number of branch points. For the purpose of illustration, we work out the case of a specific simple branched molecule containing rings. This molecule is shown in Figure 2.

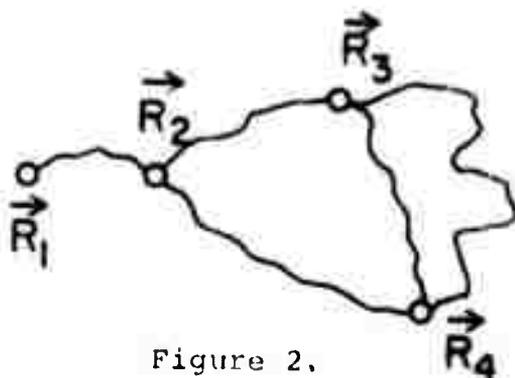


Figure 2.

It has 4 branch points and 5 branches, including 2 branches between points \vec{R}_1 and \vec{R}_4 . Since we are only interested in the relative configurations of the molecule we choose \vec{R}_1 as the origin.

Now with each branch point fixed the free energy of say the ij th branch may be denoted by

$$A_{ij} = A_{ij}(R_{ij}) \quad . \quad (63)$$

Presumably the value of A_{ij} can be evaluated exactly since a linear chain is involved. If potential energy is important the methods discussed in Flory's book¹⁰ may be used. In the simplest case, for the freely orienting chain, $A_{ij} = -TS_{ij}$ where S_{ij} is the entropy consistent with a given value of R_{ij} . Furthermore,

$$P_{ij}(R_{ij}) = K_{ij} e^{-A_{ij}/kT} \quad , \quad (64)$$

where K_{ij} is a normalizing constant.

Corresponding to a given configuration $\{\vec{R}_B\}$ the free energy $\Lambda(\{\vec{R}_B\})$ is given by

$$\Lambda(\{\vec{R}_B\}) = \sum_{ij} A_{ij}(R_{ij}) \quad . \quad (65)$$

Furthermore, $P(\{\vec{R}_B\})$ is given by eq. (61) which may be written in the alternative form

$$P(\{\vec{R}_B\}) = \frac{Z(\{\vec{R}_B\})}{Z} \quad , \quad (66)$$

where $Z(\{\vec{R}_B\})$ is the configurational partition function corres-

ponding to fixed $\{\vec{R}_B\}$ and Z is the total configurational partition function,

$$Z = \int Z(\{\vec{R}_B\}) d\{\vec{R}_B\} \quad . \quad (67)$$

It is also true that

$$A(\{\vec{R}_B\}) = -kT \ln Z(\{\vec{R}_B\}) \quad , \quad (68)$$

and that the total configurational free energy is

$$A = -kT \ln Z \quad . \quad (69)$$

From the last two equations,

$$P(\{\vec{R}_B\}) = e^{[A - A(\{\vec{R}_B\})]/kT} \quad . \quad (70)$$

Thus, the possession of $A(\{\vec{R}_B\})$ will not by itself determine $P(\{\vec{R}_B\})$ because A is not known.

Suppose, however, that we wish to make maximum use of information on $A(\{\vec{R}_B\})$ in order to estimate A . To this end, we can guess at the function $P(\{\vec{R}_B\})$ in the hope that some further principle will be available to assist with the process of guessing. We denote the guessed value of $P(\{\vec{R}_B\})$ by $P_*(\{\vec{R}_B\})$. Now, since, within this approximation, the fraction of time which the system spends in $d\{\vec{R}_B\}$ is given by $P_*(\{\vec{R}_B\})d\{\vec{R}_B\}$, the contribution, A_* , to the estimated A from this configuration, might, at first, be taken as

$$P_*(\{\vec{R}_B\})A(\{\vec{R}_B\})d\{\vec{R}_B\} \quad . \quad (71)$$

However, the full value of A_* must also include entropy contributions arising from the uncertainty as to which configuration, $\{\vec{R}_B\}$, the system is actually in. Thus, we must include a term

$$kT \int P_*(\{\vec{R}_B\}) \ln P_*(\{\vec{R}_B\}) d\{\vec{R}_B\} \quad (72)$$

in the estimate of the total free energy. Adding the contributions defined by eqs. (71) and (72), we get

$$A_* = \int P_*(\{\vec{R}_B\}) A(\{\vec{R}_B\}) d\{\vec{R}_B\} + kT \int P_*(\{\vec{R}_B\}) \ln P_*(\{\vec{R}_B\}) d\{\vec{R}_B\}. \quad (73)$$

It is easy to show, for any choice of $P_*(\{\vec{R}_B\})$ which satisfies the normalization condition,

$$\int P_*(\{\vec{R}_B\}) d\{\vec{R}_B\} = 1 \quad , \quad (74)$$

that

$$A_* \geq A \quad . \quad (75)$$

Herein lies the variation principle. To prove eq. (75), we substitute, in eq. (73), using eqs. (66), (67), (68), (69), and (74). The result is

$$A_* = A + kT \int P_*(\{\vec{R}_B\}) \ln \frac{P_*(\{\vec{R}_B\})}{P(\{\vec{R}_B\})} d\{\vec{R}_B\} \quad . \quad (76)$$

Now eq. (74), together with the equivalent normalization of $P(\{\vec{R}_B\})$, can be shown to require the integral on the right to be positive or zero. This assures the satisfaction of eq. (75).

The best choice of $P_*(\{\vec{R}_B\})$ will be that which therefore

minimizes A_* . If there are no constraints on $P_*(\{\vec{R}_B\})$ beyond those of normalization, then the extremal is supplied by

$$P_*(\{\vec{R}_B\}) = P(\{\vec{R}_B\}) \quad , \quad (77)$$

which, by eq. (76), gives $A_* = A$. On the other hand, we may constrain the form of $P_*(\{\vec{R}_B\})$ to be such as to simplify the cooperative problem, in which case eq. (77) will not hold.

As indicated earlier $A(\{\vec{R}_B\})$, specified by eq. (65) can prematurely be evaluated more or less exactly. We therefore replace the true $P(\{\vec{R}_B\})$ by a product of "single branch point" functions,

$$P_*(\{\vec{R}_B\}) = \prod_i \phi_i(\vec{R}_i) \quad , \quad (78)$$

where normalization now requires that each single branch point function satisfy the constraint

$$\int \phi_i(\vec{R}_i) d\vec{R}_i = 1 \quad . \quad (79)$$

We then substitute eq. (78) into eq. (73) and choose the set of $\phi_i(\vec{R}_i)$ which minimize A_* subject to the constraints imposed by eq. (79). Thus we look for the solution to

$$\delta A_* = \delta \left\{ \int \prod_i \phi_i(\vec{R}_i) \left[\sum_{ij} A_{ij}(\vec{R}_{ij}) d\{\vec{R}_B\} + kT \int \prod_i \phi_i(\vec{R}_i) \ln \phi_i(\vec{R}_i) d\{\vec{R}_B\} \right] \right\} = 0$$

$$\delta \left\{ \int \phi_i(\vec{R}_i) d\vec{R}_i \right\} = 0 \quad . \quad (80)$$

For the particular molecule illustrated in Figure 2 we

have

$$\begin{aligned}
 A_* = & \int \phi_2(\vec{R}_2) \phi_3(\vec{R}_3) \phi_4(\vec{R}_4) [A_{12}(R_{12}) + A_{23}(R_{23}) + A_{24}(R_{24}) + A_{34}(R_{34}) \\
 & + A_{43}(R_{34})] dR_2 dR_3 dR_4 \\
 & + kT \int \phi_2(\vec{R}_2) \phi_3(\vec{R}_3) \phi_4(\vec{R}_4) \ln\{\phi_2(\vec{R}_2) \phi_3(\vec{R}_3) \phi_4(\vec{R}_4)\} d\vec{R}_2 d\vec{R}_3 d\vec{R}_4. \quad (81)
 \end{aligned}$$

Application of eq. (79) gives

$$\begin{aligned}
 A_* = & \int \phi_2 A_{12} d\vec{R}_2 + \int \phi_2 \phi_3 A_{23} d\vec{R}_2 d\vec{R}_3 + \int \phi_2 \phi_4 A_{24} d\vec{R}_2 d\vec{R}_4 + \int \phi_3 \phi_4 [A_{34} + A_{43}] d\vec{R}_3 d\vec{R}_4 \\
 & + kT \int \phi_2 \ln \phi_2 d\vec{R}_2 + kT \int \phi_3 \ln \phi_3 d\vec{R}_3 + kT \int \phi_4 \ln \phi_4 d\vec{R}_4, \quad (82)
 \end{aligned}$$

where we have not shown the arguments of the various functions under the assumption that they are obvious. Performing the variation indicated by eq. (80), we obtain the following set of equations:

$$\begin{aligned}
 \phi_2 &= K_2 \exp \left\{ -\frac{1}{kT} [A_{12} + \int \phi_3 A_{23} dR_3 + \int \phi_4 A_{24} dR_4] \right\} \\
 \phi_3 &= K_3 \exp \left\{ -\frac{1}{kT} \left[\int \phi_2 A_{23} d\vec{R}_2 + \int \phi_4 [A_{34} + A_{43}] d\vec{R}_4 \right] \right\} \\
 \phi_4 &= K_4 \exp \left\{ -\frac{1}{kT} \left[\int \phi_2 A_{24} d\vec{R}_2 + \int \phi_3 [A_{34} + A_{43}] d\vec{R}_3 \right] \right\}. \quad (83)
 \end{aligned}$$

Eqs. (83) are to be solved simultaneously for the functions ϕ_2 , ϕ_3 , and ϕ_4 . The solution can be substituted into eq. (78) to obtain an approximate distribution function, $P(\{\vec{R}_B\})$, for the branch points with the aid of which the various averages $\langle \rangle_R$ appearing in eq. (57) can be further averaged to give $\langle S^2 \rangle$, the mean square radius of gyration.

Just how easily eq. (78) can be solved, or how adaptable it is to the treatment of "average" branched molecules characterized only by the number of branch points, their functionalities, and the degree of polymerization remains to be seen. At the moment an interative solution of eq. (83) seems within the realm of possibility.

IV. Rubber Elasticity

As indicated in Section I, the technological importance of polymers lies overwhelmingly in the bulk state. In this condition the polymer may be rubbery, gel-like, glassy, or crystalline. Usually, but not always, it is in a cross-linked state, forming a three-dimensional network. The branched and ring molecules discussed in the last section have many of the characteristics of such infinite network structures, and the variational principle suggested for the treatment of such molecules could find important application in infinite networks as well. In the dense bulk state excluded volume effects are bound to be important and self consistent field approaches may be of considerable utility.

Among the earliest attempts at the treatment of the mechanical properties of bulk polymers is the theory of rubber elasticity. This theory¹⁰ has undergone very little improvement, since the early efforts of Guth, James, and Wall some thirty years ago, and there is still room for considerable advance.

The essence of the present theory is as follows. Consider an isotropic cube of rubber of edge L_0 and volume L_0^3 in the undeformed state, and make the following assumptions:

- (1) It consists of a network structure each branch of which contains the same number of freely orienting segments (no potential energy).
- (2) The distribution of end to end distances for the branches is the same as if the branches were free molecules. Therefore the fraction of branches of length R in $d\vec{R}$ is given by eq. (2) which we now rewrite as

$$P_0(x, y, z) dx dy dz = \left(\frac{3}{2\pi n a^2} \right)^{3/2} e^{-\frac{3(x^2 + y^2 + z^2)}{2na^2}} dx dy dz, \quad (84)$$

where we have used $R^2 = x^2 + y^2 + z^2$

- (3) When the rubber is deformed the distribution function, eq. (84), undergoes an affine transformation. This means that when the edges go from all of them equal to L_0 to L_x , L_y , and L_z , respectively $P_0 \rightarrow P$ such that

$$P(x, y, z) = \frac{1}{\alpha_x \alpha_y \alpha_z} \left(\frac{3}{2\pi n a^2} \right)^{3/2} e^{-\frac{3}{2na^2} \left[\frac{x^2}{\alpha_x^2} + \frac{y^2}{\alpha_y^2} + \frac{z^2}{\alpha_z^2} \right]}, \quad (85)$$

where

$$\alpha_x = L_x/L_0, \quad \alpha_y = L_y/L_0, \quad \alpha_z = L_z/L_0, \quad (86)$$

(4) Finally, the rubber is incompressible,

$$\alpha_x \alpha_y \alpha_z = 1 \quad (87)$$

This assumption is not absolutely necessary.

Since there is no potential energy the change in internal energy upon isothermal deformation is zero and the combined first and second laws of thermodynamics require that, the change in entropy

$$dS = \frac{1}{T} (DW) \quad (88)$$

where DW is the mechanical work done by the system during deformation. If we choose $\alpha_x = \alpha$ and $\alpha_y = \alpha_z$, then with eq. (87) the state of deformation is described by α alone. Furthermore DW in eq. (88) can be expressed in terms of the tensile stress τ and strain α . Equations (84) and (85) allow one to enumerate the change in the number of configurations accessible to the network upon deformation and so to compute dS. When this result is substituted into eq. (88) the following stress-strain relation is obtained

$$\tau = \frac{NkT}{L_0} \left(\alpha - \frac{1}{\alpha^2} \right) \quad (89)$$

where N is the number of branches or more properly, strands, in the network.

Actually, besides the entropy change deduced from eqs. (84) and (85), another contribution must be considered, associated with a quantitative difference, between the un-

deformed and deformed state, in the restrictions imposed on configuration by network branch points. However, this contribution does not change the form of eq. (89).

Assumption (1) and (2) above are quite ad hoc in their characters and it is desirable to find means of eliminating them. Also it is of some importance to introduce the effects of excluded volume (this time-of an inter-chain nature) and potential energy. The self-consistent field approach may have much to offer in accomplishing these refinements.

V. Glassy Polymers and the Glass Transition - Some New Research Ideas

Most polymer masses undergo a transition from a rubbery to glassy state as the temperature is lowered. This may occur independent of whether or not a network structure exists. The onset of this state is generally signaled by dramatic changes in viscoelastic properties, especially in the spectrum of relaxation times characterizing them. However, there are often significant thermodynamic or quasi-thermodynamic effects, such as heat capacity anomalies.

For some time it had been thought that the transition to the glassy state did not involve a thermodynamic transition but merely corresponded to a gradual "quenching in" as relaxation processes slowed-down with decreasing temperature. This point of view received support from the fact that longer

time experiments designed to explore slower relaxation processes led to lower glass transition temperatures. Recently however, there has been a growing conviction that the transition is, fundamentally, of a thermodynamic nature. Gibbs and DiMarzio¹⁹ have set forth a theory for this effect which, though obviously incomplete, is convincing.

According to this theory, there are energetic effects which act to establish a disordered configurational state as being of lower energy than any ordered one. Furthermore, as the temperature is reduced the system, if allowed to come to equilibrium, condenses into a disordered state with a finite degeneracy. This degeneracy is finite in the sense that if it is represented by a number of states, W , then, $\ln W = 0$, relative to, say, Avogadro's number. As a result all configurational entropy vanishes at the glass transition even though the system remains disordered (in the sense that a periodic lattice is not formed).

Actually, the Gibbs-DiMarzio theory is highly modelistic, and is based on a lattice model of the disordered state. In addition, several other ad-hoc assumptions concerning the energetics of the system are made. It would, therefore, be of significant value if a theory could be developed without appeal to a lattice.

In this connection, the present author would like to suggest that another system exists in which a transition has been studied which on closer examination has many features

seem to be related to the glass transition. In fact, one might say it undergoes a "hyalo-phobic" transition and therefore avoids the glassy state and loss of configurational entropy. The system in question is the well known fluid of rigid spheres studied by many authors. It is known that as the volume of the fluid is decreased, the number of spheres remaining constant, a point of density is reached well below the density of close-packing, at which the system undergoes a transition to the crystalline state. This transition has been observed both in approximate analytical theories^{20,21}, with and without lattice models, and by computer methods, either by means of Monte Carlo techniques or through integration of the dynamical equations of motion. Although there is no absolute proof, the existence of the transition is regarded as well established.

Now since there is no potential energy the thermodynamics of the rigid sphere system is based entirely on considerations of entropy. Regarded from this viewpoint, the "causative" mechanism of the transition is very likely the following.

As the volume is decreased and the system remains in the disordered fluid state the number of distinct configurations available to the collection of rigid spheres is progressively decreased. It is well known that spheres can be packed more efficiently in an ordered rather than a disordered array. As a result the system begins to "jam" at a volume larger than

that of close packing. As this "jamming" continues with further decrease in volume it is accompanied by a simultaneous decrease in the number of accessible disordered configurations, until a point is reached at which only one, or at the most a relatively finite few, configurations are available. Thus the entropy of the fluid approaches zero as in the case of the glass treated by Gibbs and DiMarzio.

However, unlike the case of the glass it costs the sphere system no energy to adopt an ordered state while at the same time it can realize an appreciable gain in entropy. Thus if a small volume of the fluid crystallizes, the more efficient packing which occurs "loosens up" the remaining fluid by providing it with more volume and more configurations. Thus crystallization leads in this case to an overall increase in entropy and the transition occurs. The rigid sphere system is thus able to avoid a glass transition by becoming ordered. The transition may then be thought of as being induced by an aversion to the glassy state -- hence the suggested terminology "hyalo-phobic".

Since this hyalo-phobic transition can be studied without introducing a lattice it may be worthwhile to start the search for a non-lattice theory of the glass transition by building further restrictions which forbid ordering into the hard sphere system. This represents an interesting challenge for the future.

It is worth noting that such a modified hard sphere model might form a useful theoretical basis for the study of amorphous metals and semiconductors.

Another area of increasing importance to polymer technology combines both the features of the glassy and rubbery state in a single molecule; a block-copolymer. Such molecules involve a polymer of say type A connected to a chain of type B which may in turn be connected to A, i.e., of the form ABA. Now it has been observed that a bulk block-copolymer can undergo a transition, as the temperature is lowered, such that the B portion becomes glassy and separates (except for the AB and BA bond) from the A portion which remains rubbery. One therefore achieves a composite structure in which the "filament" and the matrix are bound by a covalent bond! Furthermore, the glassy regions often form a quasi-ordered array. Obviously a great variety of specialized mechanical features may be "engineered" in such substances.

Such "engineering" would be considerably advanced if it were possible to predict the conditions of temperature and composition under which specific such "heterophase" structures would develop. Some crude beginnings in this direction have been made by Williams²² in which he treats the heterophase state as a true two-phase system with an interfacial region between them. It seems, however, that much more can be done in the way of a statistical mechanical theory employing the developing theory of the statistical thermodynamics of curved

interfaces¹³ together with a better treatment of the statistical thermodynamics of the glassy state.

Here again the self-consistent field approach may be of value. For example, with the block copolymer of type ABA the bonding points between the blocks might be treated like the branch points \vec{R}_i in eq. (75) while the Λ_{ij} might refer to the free energies of A and B blocks, respectively.

VI. Inhomogeneous Networks

Three dimensional polymer networks are frequently considered as homogeneous structures and their ultimate properties are generally related to the properties of such a continuum. Actually, if such networks were homogeneous their theoretical strengths should be about one hundred times higher than is ordinarily observed. This suggests that these networks are in reality quite inhomogeneous -- regions having a high degree of cross linking being interlaced with those which are less heavily cross-linked. From the practical point of view such inhomogeneities may at times even be desirable, when they can confer on the system some of the features sought after in composites.

Recently Labana, Newman, and Chomppf²⁴ have developed a theory (and performed experiments aimed at testing it) which indicates that the development of inhomogeneities is a natural outgrowth of the kinetics of the cross-linking process, and that indeed homogeneous networks are not to be expected.

In such a cross-linking process there is a stage,

depending upon the functionality of the branch points, at which cross-linking has reached a level such that a "gel" is said to have formed. For example, if the branch points are f -functional then, when on the average $(f-1)^{-1}$ of the potential branches of each branch unit are utilized, the probability of formation of infinite network becomes effectively unity²⁵. This is the gel point. Usually it is assumed that the gel thus formed is a uniform network.

Labana, Newman, and Chompff obtain their important result, contradicting this premise, as follows. They consider a single non-cross linked linear molecule in bulk polymer mass, about to be crosslinked. They then assume as in the case of eq. (26) that the distribution of segments around the center of mass of this molecule (including the effects of excluded volume) is gaussian. Thus, they write

$$\rho(r) = n \left(\frac{3}{2\pi \langle R^2 \rangle} \right)^{3/2} e^{-\frac{3r^2}{2\langle R^2 \rangle}} \quad (90)$$

in which n is the number of segments in the molecule $\langle R^2 \rangle$ is the mean square radius of gyration and $\rho(r)$ is the segment density at a distance r from the center of mass. They define

$$\rho(0) = n \left(\frac{3}{2\pi \langle R^2 \rangle} \right)^{3/2} \quad (91)$$

For a linear molecule $\langle R^2 \rangle$ is obtainable from eq. (5) and for a branched molecule, through the application of eqs. (29) and (30), but for a molecule containing rings another pro-

cedure, described below, is necessary.

As the initial unlinked molecule undergoes crosslinking it is assumed that the distribution maintains the form given by eq. (90) but that both n and $\langle R^2 \rangle$ change. With this assumption it is possible to study the relative rate at which the molecule links with itself as compared to the rate of forming cross links with other molecules. The analysis indicates that the rate of intramolecular cross linking gradually accelerates relative to the intermolecular process so that an inhomogeneity develops. The system gels locally before the extended gel point is reached. This result is summarized by the statement that "gel balls" are formed, and that these local highly cross-linked regions, separated from each other by more lightly cross-linked material, are natural features of the network forming process.

The actual analysis goes as follows. At any stage of the process the probability that a segment of the molecule, located in the volume element $d\vec{r}$, reacts with another segment from the same molecule is given by an expression which reflects the bi-molecular nature of the rate process, namely,

$$Q[\rho(r)]^2 d\vec{r} \quad , \quad (92)$$

where Q is proportional to the rate constant for cross-linking. From this it follows that the total probability of intramolecular reaction is

$$P_{\text{internal}} = Q \int_{\infty}^0 [\rho(r)]^2 4\pi r^2 dr \quad . \quad (93)$$

By the same line of reasoning the total probability of forming and intermolecular link is

$$P_{\text{external}} = Q \int_0^{\infty} [\bar{\rho}_0 - \rho(r)] \rho(r) 4\pi r^2 dr \quad , \quad (94)$$

where $\bar{\rho}_0$ is the average segmental density in the bulk polymer mass and $\bar{\rho}_0 - \rho(r)$ is supposed to represent the density of segments in $4\pi r^2 dr$, arising from other molecules besides the one in question. The use of $\bar{\rho}_0 - \rho(r)$ in this manner represents a first order attempt to account for the other-molecule segmental density on the basis that part of $\bar{\rho}_0$ in the volume element $4\pi r^2 dr$ is accounted for by $\rho(r)$ from the primary molecule under consideration. As we shall show, this is a somewhat inconsistent approach since $\bar{\rho}_0$ and $\rho(r)$ are not defined in the same coordinate system. $\bar{\rho}_0$ is defined in the laboratory frame of reference while $\rho(r)$ is in the center of mass system of the molecule.

In the present author's opinion eq. (94) should be replaced by

$$P_{\text{external}} = Q \int_0^{\infty} \bar{\rho}_0 \rho(r) 4\pi r^2 dr \quad . \quad (95)$$

The argument for this will be presented later. Also, Labana, Newman and Chompff insist that $\rho(0)$ defined by eq. (91) cannot exceed $\bar{\rho}_0$. For very similar reasons the present author considers this to be an incorrect statement. On the other

hand, in spite of these inaccuracies the substance of conclusions drawn by Labana et al. is not altered and their theory remains significant.

Upon substituting eqs. (90) and (91) into eqs. (93) and (94) and performing the indicated integrations, one finds (using an obvious abbreviation of notation)

$$\frac{p_{in}}{p_{ex} + p_{in}} = \frac{\rho(o)}{2^{1/2} \bar{\rho}_o} \quad (96)$$

If, in fact $\rho(o) \leq \bar{\rho}_o$, then, according to eq. (96), $p_{in}/(p_{ex} + p_{in})$ could not exceed 0.354.

According to eq. (91) $\rho(o)$ depends on n and $\langle R^2 \rangle$, and these change during network formation. With branched molecules, obtained by random intermolecular reaction of the central molecule with its neighbors, it can be shown, using the statistical methods for the calculation of the radius of gyration outlined in Section II, that $\rho_o(o)$ changes only slightly as the molecule grows -- the distribution broadens but does not grow in amplitude. Thus according to eq. (96) the ratio $p_{in}/(p_{ex} + p_{in})$ stays almost the same during the branching process.

The computation of $\langle R^2 \rangle$ for a molecule containing rings is another matter. As indicated at the end of Section II simple techniques do not suffice. Using an indirect method which we shall discuss later, Champff²⁶ is able to estimate $\langle R^2 \rangle$ as the cross linking proceeds to the point where rings actually form.

As rings form, $\rho(o)$ undergoes a drastic increase, and $p_{in}/(p_{ex} + p_{in})$ also increases abruptly. Therefore the process of ring formation increases the chance for more ring formation and $\rho(o)$ increases rapidly towards $\bar{\rho}_o$, then the "gel ball" which the growing cross linked molecule represents must grow outward (i.e., must become more extensive). In this manner inhomogeneities are produced.

Before discussing the question of possibly improving the estimate of $\langle R^2 \rangle$ for molecules with rings we comment on the validity of eq. (95) versus eq. (94).

We return to the primary molecule whose center of mass defines the origin of the coordinate system, and we denote by $G(R)$ the radial distribution function of the centers of mass of other polymer molecules about this origin. The simplest estimate of $G(R)$ is that it is constant (since molecules can be interleaved and since no segment need lie at the center of mass, two molecules may even have their centers of mass at the same point) such that

$$G(R) = \bar{\rho}_o / M \quad . \quad (97)$$

Note that $G(R)d\vec{R}$ gives the number of mass centers in the volume element $d\vec{R}$. As before $\rho(r)$ is used to represent the density of segments about the center of mass of a given molecule. $\rho(r)$ should already contain interference effects between molecules because it refers to the bulk state. Now we compute the

probability of a link forming between two different molecules in the volume element dr about the center of mass of the primary molecule. The situation is diagrammed in Figure 3.

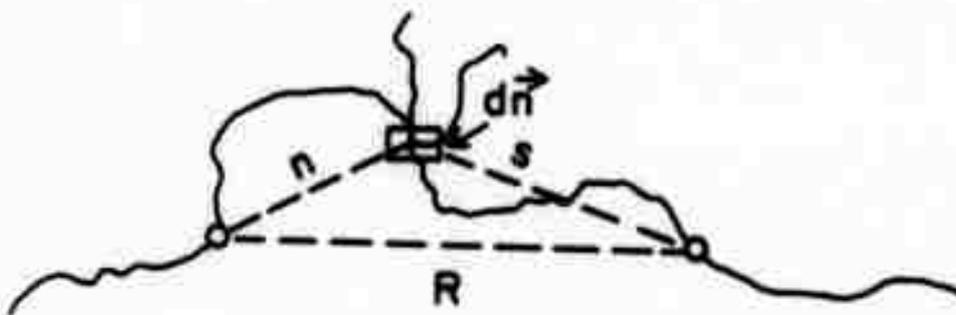


Figure 3.

The distances r , s , and R are self-explanatory. We employ s and R as bipolar coordinates with an aximuthal angle for rotations about the axis r . Then in place of $\bar{\rho}_0 - \rho(r)$ in eq. (94) we must use

$$\frac{1}{r} \int d\phi \int s ds \int G(R) \rho(s) R dR \quad , \quad (98)$$

where the range of integration for ϕ must be $(0, 2\pi)$ while the integrations over s and R must be taken in two parts as follows:

$$\int_0^r ds \int_{r-s}^{r+s} dR + \int_r^\infty ds \int_{s-r}^{r+s} dR \quad , \quad (99)$$

The expression for p_{ex} now becomes

$$p_{ex} = Q \int \rho(r) d\vec{r} \int \frac{d\phi}{r} \int s ds \int G(R) \rho(s) R dR \quad , \quad (100)$$

where the limits in (99) are to be observed. Employing eq. (97) for $G(R)$ and carrying out the integrations over ϕ and the bipolar coordinates eq. (100) reduces immediately to eq. (95) which proves the point.

The new p_{ex} given by eq. (95) is

$$p_{ex} = Q\bar{\rho}_0 n \quad (101)$$

and the ratio

$$\frac{p_{in}}{p_{ex} + p_{in}} = \frac{\rho(o)}{2^{3/2} \bar{\rho}_0 \rho(o)} \quad (102)$$

Unlike the expression given in eq. (96), even if $\rho(o)$ exceeds $\bar{\rho}_0$, as it may, $p_{in}/(p_{ex} + p_{in})$ cannot exceed unity. As a matter of fact this is the upper limit for the ratio and indicates that the instability in reaction (intramolecular versus intermolecular) can grow until no further reaction outside the growing molecule occurs. Of course $\rho(o)$ will probably reach some limit but it need not be $\bar{\rho}_0$.

Turning now to Chompff's²⁶ estimate of $\langle R^2 \rangle$ for molecules containing rings it is sufficient to describe his method briefly. In another context Chompff was able to generalize the Rouse model, to be discussed later, to compute the viscosity contribution, in solution, of molecules containing rings. In this way he was able to develop a relationship between the structure and degree of polymerization of the molecule and the viscosity. An alternative relationship existed between viscosity and radius of gyration. Thus by eliminating the viscosity

between these two expressions he was left with the relation between $\langle R^2 \rangle$ and structure.

This seems like an awkward method for deriving a geometric quantity like radius of gyration. It amounts to computing an equilibrium quantity by means of an irreversible (viscosity) process! There should be a more direct route, involving only equilibrium considerations. The self-consistent field approach for ring molecules, described between eq. (45) through eq. (83) seems to offer promise in this connection.

As a final note we should record the fact that the discussions during the current summer study have indicated that glassy carbon seems also to be comprised of an inhomogeneous network. In fact, the inhomogeneities seem to have dimensions similar to those predicted by the foregoing theory and, indeed, apparently observed by Labana et al. Since the pyrolysis process leading to glassy carbon structures probably involves cross linking, the advances made in connection with inhomogeneous polymer networks may be relevant to the study of glassy carbon.

VI. Polymer Crystals

Crystalline polymers are also known in the bulk state. Frequently they can be observed growing within amorphous polymer masses as spherulites, and they have been known for many years as micelles in polymer fibers. Until the 1960's the "fringe" theory of polymer crystals was well accepted. According to

this theory a single polymer molecule could run its course through both crystalline micelles and amorphous regions, and the fraction of crystallinity could vary with external factors such as state of strain, etc.

It was assumed that the polymer molecules within micelles lined up with their axes parallel to one another.

During the 1960's, especially under the influence of Hoffman and his associates, vigorous experimental and theoretical work²⁷ was performed on polymer single crystals, especially polyethylene. The cross referencing of theory and experiment has made it clear that the polymer chains in a crystal (and also in spherulites and micelles) are highly folded back upon themselves, and are anything but straight and aligned as in a bundle of sticks.

Furthermore, recent work on polymer morphology seems to indicate that spherulites in bulk masses are tied to each other by thin fibrils so that the polymer represents a very complex composite indeed. Much work needs to be done in this area.

Perhaps one of the most interesting bits of information uncovered by the interdisciplinary exchange of the summer study is the fact that non chain-folded polyethylene crystals can be produced by allowing a polymeric melt to crystallize under conditions of flow, e.g., while being stirred. Such crystals which seem to have the molecules aligned parallel to their long axes are found to be very strong (although they have some

(laws).

It is possible that they could form ideal carbon filaments for use in carbon composites! Exploration of this possibility is strongly recommended.

VIII. Frictional and Dynamical Effects

The incremental viscosity imparted by polymer molecules to a solution has been investigated at great length. These effects are usually discussed in terms of the intrinsic viscosity defined as

$$[\eta] = \lim_{c \rightarrow 0} \frac{\eta - \eta_0}{\eta_0 c} \quad , \quad (103)$$

where η is the viscosity of the solution, η_0 is the viscosity of the pure solvent, and c is the concentration of polymer solute in g/cc. This limiting quantity (η is a function of c) represents the relative increase in viscosity per solute molecule.

The methods developed for the treatment of visco-dynamical effects for polymer molecules in solution are to a large extent applicable to the treatment of relaxation effects in bulk network polymers. In fact, a beginning has already been made by Chomff and his collaborators in this respect. We therefore outline some of these methods and indicate some directions of research which might be fruitful in the immediate future. In this area, in particular, techniques of the solid state physicist, especially those

connected with the lattice dynamics of disordered lattices, seem to hold a great deal of promise.

In the discussion it is convenient to have in mind a particular flow field. We choose this such that

$$v_x = gy, \quad v_y = v_z = 0 \quad , \quad (104)$$

where g is a constant, x, y, z are the coordinates of position and $v_x, v_y,$ and v_z the components of flow velocity. This flow could, for example, be generated by the motion of one of two parallel plates (with fluid between them) perpendicular to the y axis.

Let $\vec{\sigma}_v$ and $\vec{\sigma}_v^0$ be the stresses, in the solution and solvent respectively, exerted across a unit area Σ whose unit normal is \vec{e}_v . We evaluate the increase in stress, $\vec{\sigma}_v - \vec{\sigma}_v^0$, (at constant rate of shear) due to the introduction of polymer molecules. This fundamental derivation is due to Kramers who recognized that the increase of stress is due to the tension in the segments of the polymer chain which intersect Σ . He considered a molecule with $n + 1$ segments. The terminus of the j th segment is located by the vector \vec{R}_j while the vector length of the j th segment is given by $\vec{r}_j = \vec{R}_j - \vec{R}_{j-1}$. The tension along the j th segment is denoted by \vec{T}_j . If ρ is the number of polymer molecules per unit volume, then by averaging over all configurations of molecules intersecting Σ it is possible to show that

$$\vec{\sigma}_v - \vec{\sigma}_v^0 = \rho \sum_j \langle \vec{T}_j (\vec{r}_j \cdot \vec{e}_v) \rangle \quad (105)$$

where the brackets indicate the average over all configurations of a molecule.

The total force exerted on the j th segment is derived from the "gradient" of \vec{T}_j from segment to segment and the frictional force \vec{F}_j on the segment. Thus the equation of motion is

$$m \frac{d^2 \vec{R}_j}{dt^2} = \vec{T}_{j+1} - \vec{T}_j - \vec{F}_j \quad (106)$$

Since inertial effects may be neglected the left side of eq. (106) may be set equal to zero and the result substituted into eq. (105). After a small amount of manipulation, one obtains

$$\vec{\sigma}_v = \vec{\sigma}_v^0 - \rho \sum_j \langle \vec{F}_j (\vec{R}_j \cdot \vec{e}_v) \rangle \quad (107)$$

where the boundary condition

$$\vec{T}_0 = \vec{T}_{n+1} = 0 \quad (108)$$

has been used. If we set $\vec{e}_v = \vec{e}_y$, the x-component of $\vec{\sigma}_v$ becomes

$$\sigma_{xy} = \sigma_{xy}^0 - \rho \sum_j \langle (\vec{F}_j \cdot \vec{e}_x) (\vec{R}_j \cdot \vec{e}_y) \rangle \quad ; \quad (109)$$

and since, for the flow field described by eq. (104)

$$\sigma_{xy} = \eta g \quad , \quad (110)$$

eq. (109) can be converted into

$$\eta - \eta_0 + \frac{\rho}{g} \sum_j \langle (\vec{F}_j \cdot \vec{e}_x) (\vec{R}_j \cdot \vec{e}_y) \rangle \quad ; \quad (111)$$

and, finally, it can be shown using (111) that

$$[\eta] = - \frac{N_A}{M\eta_c g} \sum_j \langle F_{jx} y_j \rangle \quad , \quad (112)$$

where N_A is Avogadro's number, M the polymer molecular weight, and $F_{jx} = \vec{F}_j \cdot \vec{e}_x$ is the x-component of \vec{F}_j while y_j is the y component of \vec{R}_j .

This extremely important basic equation is used in a very central fashion in the evaluation of the intrinsic viscosity. In the form shown in eq. (112) it appears disarmingly simple. It must be remembered however that the evaluation of the average indicated by the brackets in eq. (112) requires the possession of the molecular configurational distribution function in the nonequilibrium state characteristic of the flow. In essence, then, the remainder of the problem involves the determination of this distribution function.

The distribution function is a density $P(\vec{R}_0, \vec{R}_1, \dots, \vec{R}_n)$ in the space defined by the various \vec{R}_j . If it is assumed that each segment interacts dissipatively with the solvent through a friction coefficient, ζ , such that the frictional force on the segment is

$$\vec{F}_j = \zeta (\vec{u}_j - \vec{v}_j) \quad , \quad (114)$$

where \vec{v}_j is the velocity of the surrounding fluid at \vec{R}_j while \vec{u}_j is the velocity of the segment, then a Fokker-Planck equation

can be derived from which it may be shown that

$$\vec{J}_i = P \vec{v}_i + \frac{P}{\zeta} (-\nabla_i U + \vec{X}_i) - \frac{kT}{\zeta} v_i P \quad , \quad (114)$$

In this equation \vec{J}_i is the flux density in the \vec{R}_i direction in the generalized configuration space. In other words, \vec{J}_i measures the average rate at which P changes its generalized position because the i th segment changes its position. v_i symbolizes the gradient with respect to \vec{R}_i , while \vec{X}_i represents the force on the i th segment arising outside the molecule and $-\nabla_i U$, the force from within the molecule. The continuity equation requires

$$\frac{\partial P}{\partial t} = - \sum_{i=0}^n \nabla_i \cdot \vec{J}_i \quad . \quad (115)$$

Since the current density is also expressible as

$$\vec{J}_i = P \vec{u}_i \quad , \quad (116)$$

use of eq. (113) and (116) in eq. (114) gives

$$\vec{F}_i = -\nabla_i U + \vec{X}_i - kT \nabla_i \ln P \quad , \quad (117)$$

so that the frictional force which must be used in eq. (112) is seen to be expressible in terms of both intramolecular ($-\nabla_i U$) and external (\vec{X}_i) forces and, not surprisingly, an entropy generating "statistical" force, $-kT \nabla_i \ln P$, associated with intramolecular diffusion (Brownian motion). Thus P in eq. (112) is not only involved in taking the average represented by the brackets, but also in F_{jx} .

Kirkwood has attempted a very general evaluation of P in Riemann space. However, this approach, in spite of its systematic beauty, has not been easy to apply to the computation of data concerning real systems, and another, more heuristic approach is more frequently employed. This is the so-called Rouse "spring and bead" model²⁸. It may be applied in the so-called "free draining" and "non free draining" cases. The former refers to the situation in which the perturbation in the flow field at segment i due to the influence of other segments, say segment j , etc., are ignored because they are assumed small, while in the latter case such perturbations are included in the treatment.

Although we shall limit our considerations to the free draining case it is pertinent to mention that a sort of Green's function tensor, the Oseen tensor, is used in the non free draining case which expresses the effect on the flow at segment i at \vec{R}_i of the perturbation due to a "point" segment at \vec{R}_j . The Oseen tensor formalism is a fruitful point of interaction with physicists and continuum mechanicians.

The Rouse model may be described as follows. We start with a linear chain molecule and formerly divide it into sub-molecules, all of the same size, which are to constitute the n molecular segments. Each segment is assumed to be a freely orienting molecule having an end to end distribution of the gaussian form typified by eq. (9). Using eq. (64) we can write for the i th segment

$$\tau(\vec{r}_i) = Ke^{-A_i/kT} = \left(\frac{3}{2\pi\langle a^2 \rangle}\right)^{3/2} e^{-\frac{3r_i^2}{2\langle a^2 \rangle}} \quad (118)$$

or

$$A_i(\vec{r}_i) = kT \ln K \left(\frac{3}{2\pi\langle a^2 \rangle}\right)^{3/2} + \frac{3kT r_i^2}{2\langle a^2 \rangle} \quad (119)$$

In both of these equations K is a normalizing constant.

The tensile force \vec{T}_i in the i th segment is obtained by substituting A_i given by eq. (119) into the thermodynamic relation

$$\vec{T}_i = -\frac{\partial A_i}{\partial \vec{r}_i} \quad ; \quad (120)$$

and the result is

$$\vec{T}_i = -\frac{3kT}{\langle a^2 \rangle} \vec{r}_i = -\kappa(\vec{R}_i - \vec{R}_{i-1}) \quad , \quad (121)$$

so that the segment acts like a spring of modulus

$$\kappa = \frac{3kT}{\langle a^2 \rangle} \quad . \quad (122)$$

In eq. (114), $-V_i u$ is the intramolecular force on the i th segment. In our model this force originates in the effective springs to which eq. (121) applies and it follows that

$$\begin{aligned} V_i u &= \kappa(\vec{R}_0 - \vec{R}_1) \quad , \quad i = 0 \\ &= \kappa(-\vec{R}_{i-1} + 2\vec{R}_i - \vec{R}_{i+1}) \quad , \quad i < i < n-1 \\ &= \kappa(-\vec{R}_{n-1} + \vec{R}_n) \quad i = n \quad . \end{aligned} \quad (123)$$

Since we are dealing with the "free draining" case we set X_i in eq. (114) equal to zero.

The friction coefficient ζ in eq. (114) is still a semi-macroscopic quantity. Thus the present model does not treat the molecular dissipation process directly. Nevertheless, it does allow all relaxation times to be expressed in terms of this single dissipative quantity ζ . In general, we take ζ proportional to the size of the submolecule or segment. With this assumption it turns out that the final result of the theory is independent of the size of the submolecule, so the choice of size is not critical.

Now we define vectors and matrices in the complete configuration space spanned by the set of vectors \hat{R}_i . The vectors are

$$\begin{aligned}
 \hat{R} &= (\hat{R}_0, \hat{R}_1, \dots, \hat{R}_n) \\
 \hat{v} &= (\hat{v}_0, \hat{v}_1, \dots, \hat{v}_n) \\
 \hat{J} &= (\hat{J}_0, \hat{J}_1, \dots, \hat{J}_n) \\
 \hat{v} &= (\hat{v}_0, \hat{v}_1, \dots, \hat{v}_n)
 \end{aligned}
 \tag{124}$$

while the matrix is the tridiagonal array

$$\begin{aligned}
\tilde{\Lambda} = & \begin{matrix} 1 & -1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ -1 & 2 & -1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & -1 & 2 & -1 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & -1 & 2 & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ 0 & 0 & \cdot & \cdot & \cdot & -1 & 2 & -1 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & -1 & 1 \end{matrix} \quad (125)
\end{aligned}$$

With these definitions, eq. (124), and with \dot{X}_1 set equal to zero, the set of eqs. (109) may be expressed compactly as

$$\ddot{J} = p\ddot{v} - \frac{p}{\zeta} \kappa \tilde{\Lambda} \cdot \ddot{R} - \frac{kT}{\zeta} \ddot{v} p \quad , \quad (126)$$

and eq. (115) becomes

$$\ddot{v}^T \cdot p\ddot{v} + \frac{\partial p}{\partial t} = -\ddot{v}^T \cdot p\ddot{v} + \frac{1}{\zeta} \ddot{v}^T \cdot (\kappa p \tilde{\Lambda} \cdot \ddot{R}) + \frac{kT}{\zeta} \ddot{v}^T \cdot \ddot{v} p \quad , \quad (127)$$

where \ddot{v}^T is the transpose of \ddot{v} .

Eq. (127) can be put into an easily solvable form by a similarity transformation which diagonalizes $\tilde{\Lambda}$. We denote the matrix of this transformation by Q . Then the diagonal matrix is

$$\tilde{\Lambda} = Q^{-1} \tilde{\Lambda} Q \quad . \quad (128)$$

The transformed (normal) coordinates are

$$\dot{z} = Q^{-1} \cdot \dot{R} \quad , \quad (129)$$

while the transformed velocity is

$$\vec{v}_\xi = Q^{-1} \cdot \vec{v} \quad (130)$$

Furthermore, the transformed gradient operator is

$$\vec{\nabla}_\xi = Q \cdot \vec{\nabla} \quad (131)$$

where $\vec{\nabla}_\xi = (v_{\xi_0}, v_{\xi_1}, \dots, v_{\xi_n})$

Eq. (127) transforms to

$$\frac{\partial P}{\partial t} = -\vec{\nabla}^T \cdot \vec{v}_\xi P + \frac{1}{\zeta} \vec{\nabla}_\xi^T \cdot (\kappa P \underline{\Lambda} \cdot \vec{\xi}) + \frac{kT}{\zeta} \vec{\nabla}_\xi^T \cdot \vec{\nabla}_\xi P \quad (132)$$

Because $\underline{\Lambda}$ is a diagonal matrix, this equation contains no cross terms and can be solved for P. In practice, eq. (132) is used not to solve for P directly, but rather to write an equation for $\langle F_{jx} y_j \rangle$ in Eq. (112), using for F_{jx} the component derived from Eq. (117). This equation can be solved, and the $[\eta]$ is computed from Eq. (112).

The matrix $\underline{\Lambda}$ contains along its diagonal the eigenvalues of A , and so the problem can only be completed when these eigenvalues have been found. These eigenvalues are, in fact, easily found. We may denote them by λ_j .

Usually the measurement of intrinsic viscosity is carried out in a non-steady flow field, oscillating at some frequency ω . In this case, it proves desirable to introduce a complex intrinsic viscosity $[\bar{\eta}]$ which reduces to $[\eta]$ at zero frequency. Without going into any further detail, it turns out that

$$[\eta] = \frac{RT}{M\eta_0} \sum_{j=1}^n \frac{\tau_j}{1+i\omega\tau_j} \quad , \quad (133)$$

where R is the gas constant and i is $\sqrt{-1}$. The τ_j are the relaxation times related to the eigenvalues λ_j by

$$\tau_j = \frac{\zeta}{2k\lambda_j} \quad (134)$$

We have presented this rather superficial outline of the theory of polymer viscosity in order to emphasize the central role played by the diagonalization of the matrix Λ . We have not considered effects due to "non-free draining", cross-linking, excluded volume, or potential energy within the submolecules. As indicated earlier, Chompff and co-workers have shown how the same theory can be applied to the calculation of the relaxation times of bulk polymer networks, and it is therefore quite important to refine the model.

At higher frequencies, shorter relaxation times become important and this implies shorter segments. Thus, we will not be at liberty to choose the lengths of submolecules as arbitrarily as before. Furthermore, potential energy may have to be introduced into the segmental springs in an explicit manner. A more careful examination of the true meaning of the model may then be in order, and it is quite probable that a distribution of submolecule sizes will have to be used.

Under these conditions, the normal mode problem (from which the relaxation times are derived) would become similar

to that encountered with solids when dealing with disordered lattices. Certainly a straightforward diagonalization of the matrix, equivalent to \tilde{A} , will no longer be possible. It is also possible, and probable, that band gaps will appear in the spectrum of relaxation times. It seems as though this problem is ideally suited for the attention of the solid state physicists, especially those concerned with disordered lattices.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHCl5-71-C-0253 with The University of Michigan.

REFERENCES

1. H. Yamakawa, *Modern Theory of Polymer Solutions*, Chapter 1, Section 5 (Harper and Row, New York, 1971).
2. B. H. Zimm, *J. Chem. Phys.* 16, 1093 & 1099 (1948).
3. H. Yamakawa, *Modern Theory of Polymer Solutions*, p. 34 (Harper and Row, New York, 1971).
4. P. J. Flory and S. Fisk, *J. Chem. Phys.* 44, 2243 (1966).
5. S. Chandrasekhar, *Rev. Mod. Phys.* 15, 1 (1943).
6. M. C. Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.* 17, 323 (1945).
7. H. Yamakawa, *Modern Theory of Polymer Solutions*, p. 21 (Harper and Row, New York, 1971).
8. *Ibid.*, p. 16.
9. W. Kuhn, *Kolloid Z.* 87, 3 (1939).
10. P. J. Flory, *Statistical Mechanics of Chain Molecules* (Interscience, New York 1969).
11. P. J. Flory, *J. Chem. Phys.* 17, 303 (1949); P. J. Flory and T. G. Fox, *J. Am. Chem. Soc.* 73, 1904 (1951).
12. H. Yamakawa, *Modern Theory of Polymer Solutions*, pp. 79-93 (Harper and Row, New York, 1971).
13. *Ibid.*, pp. 120-127.
14. H. Riess, *J. Chem. Phys.* 47, 186 (1971).
15. S. F. Edwards, *Proc. Phys. Soc.* 85, 613 (1965).
16. K. F. Freed, *J. Chem. Phys.* 55, 3910 (1971).
17. K. F. Freed, *J. Chem. Phys.* 54, 1453 (1971).
18. T. L. Hill, *Introduction to Statistical Thermodynamics*, Chapter 21 (Addison-Wesley, Reading, 1960).
19. J. H. Gibbs and E. A. DiMarzio, *J. Chem. Phys.* 28, 373 & 807 (1958).
20. W. G. Hoover and F. H. Ree, *J. Chem. Phys.* 49, 3609 (1968).

21. J. G. Kirkwood, E. K. Maun, and H. J. Alder, *J. Chem. Phys.* 18, 1040 (1950); J. G. Kirkwood and E. M. Hogan, *J. Chem. Phys.* 10, 394 (1942).
22. M. C. Williams, *Polymer Letters* 8, 115 (1970).
23. S. Ono and S. Kondo, *Handbuch der Physik*, Vol. 10 (Springer, Berlin, 1960), ed. by S. Flügge.
24. S. S. Labana, S. Newman, and A. J. Chompff, *Polymer Networks, Structure, and Mechanical Properties* (Plenum Press, New York, 1971); ed. by A. J. Chompff and S. Newman.
25. P. J. Flory, *Principles of Polymer Chemistry*, p. 353 (Cornell University Press, 1953).
26. A. J. Chompff, *J. Chem. Phys.*, to be published.
27. J. D. Hoffman, *SPE trans.* 4, 315 (1964).
28. P. E. Rouse, Jr., *J. Chem. Phys.* 21, 1272 (1953);
D. H. Zimm, *J. Chem. Phys.* 24, 269 (1956).

A NOTE ON THE GROUND STATE ENERGY
OF AN ASSEMBLY OF INTERACTING ELECTRONS

Akira Ishihara*
E. W. Montroll

Abstract

The ground state energy of an assembly of charged particles of density ρ imbedded in a continuum of charge of the other sign in an electrically neutral system is considered. Asymptotic formulae for the ground state energy of such systems are known in the high and low density regimes. An interpolation formula covering the entire density range is derived using the method of two point Padé approximants. A phase transition from an electron lattice to an electron gas seems to occur at $r_s \approx 14$, r_s being the radius of a sphere which, on the average, contains a single charge, in units of the Bohr radius of the electron in a hydrogen atom.

*State University of New York at Buffalo

Preceding page blank

A NOTE ON THE GROUND STATE ENERGY
OF AN ASSEMBLY OF INTERACTING ELECTRONS

Akira Ishihara
E. W. Montroll

A perennial many-body problem common to the theories of metals and of the internal constitution of stars (especially the white dwarfs) is the determination of the ground state energy of an assembly of point-charged particles as a function of their number density. In a traditional model, the electrical neutrality of the assembly is maintained by imbedding the point charges, which we identify as electrons (in the theory of white dwarfs, the positively charged nuclei are considered to be the point charges), in a continuum of positive charge. In the ground state (0°K), the only parameters which appear in a non-relativistic theory are Planck's constant h , the charged particle density ρ , mass m , and charge e . The basic dimensionless quantity produced from these parameters is $\pi^2/\hbar^3 \rho^{1/3}$. An important variable proportional to this combination is $r_s = r_0/r_H$ where r_0 is related to the density ρ through the definition $(4/3)\pi r_0^3 = \rho^{-1}$ and r_H is the Bohr radius, \hbar^2/me^2 .

Since the energy of a perfect Fermi gas is, per particle,

$$E_F/N = (3/2)\hbar^2 (3\rho/\pi)^{2/3}/5m \quad (1)$$

the introduction of the coulomb interactions changes the ground state energy by a factor which depends on r_s so that $E/D_f = (1 + w(r_s))$. The central problem then becomes the determination of the functional form of $w(r_s)$. In metals $2 < r_s < 6$ while in stars its range is much broader.

The high and low density limit behavior of our assembly of interest is somewhat different from that observed in molecular systems. In their high density regime, atoms and molecules become localized in a lattice structure whose zero point vibrations are quite small. For very light electrons, localization would, through the uncertainty principle, force the particles to have large momentum fluctuations which, in turn, correspond to a high kinetic energy. The total energy, kinetic plus potential, might then be higher than in the electron gas state in which the division of the two forms of energy might be better balanced. However, at low densities, lattice localization need not be so severe, so that the kinetic and potential energies might take up more equal shares of the total. We have estimated that a phase transition from the low density lattice state to the high density gas state should occur at $r_s \approx 14$.

Wigner¹ showed that among the cubic electron lattices, the body-centered one had the lowest lattice energy, its value being $U_{bcc} = -1.79186 r_s^{-1}$ Rydbergs. Since the lattice must suffer some small vibrations from its equilibrium state, a correction for the energy of zero point oscillations must be added to this value of the lattice energy in the estimation of

the ground state energy of the electron lattice. The first estimate of the correction, made by Fuchs² in 1935, was improved by Caldwell-Horsefall and Maradudin³, who found $E_H = 2.638 r_s^{-3/2}$ Rydbergs while Carr⁴ had independently reported a value of $2.66 r_s^{-3/2}$. Carr has also estimated the lowest order anharmonic contribution which, when added to the above terms, yields the beginning of an asymptotic expansion for the ground state energy (in Rydbergs) in the low density (large r_s) regime:

$$E_G = -1.79186 r_s^{-1} + 2.638 r_s^{-3/2} - 0.73 r_s^{-2} + \dots \quad (2)$$

At high densities, the limiting unperturbed state is that of the perfect Fermi gas. As the density increases, exchange terms and electron correlations of various orders become important. These effects were investigated in a flurry of activities in the late 1950's and early 1960's. Gell-Mann and Brueckner⁵ calculated the lowest order correlation term while Dubois⁶ obtained the higher order ones. The resulting small r_s expansion is

$$E_G \sim 2.21 r_s^{-2} - 0.916 r_s^{-1} + 0.0622 \ln r_s - 0.096 + r_s (0.0049 \ln r_s + C) + \dots \quad (3)$$

where $C \approx -0.02$. Progress in series expansions essentially stopped when it seemed that the calculation of another term or two would require a tremendous effort which would probably

not lead to a significant improvement in the understanding of the electron correlation phenomenon.

The state of affairs is reminiscent of that which existed in the investigation of the 3D Ising model in the 1950's. By counting the number of ways appropriate graphs could be constructed on lattices, a number of terms in a power series expansion in reciprocal temperature could be obtained for the thermodynamic properties of the Ising ferromagnet. The state was reached at which the determination of a new term required tremendous effort. Until Domb⁷ and his collaborators and G. Baker⁸ showed how to extract information from the coefficients, it was hardly worth expending any more effort to obtain more coefficients.

The Baker technique is to systematically express a power series $f(x) = c_0 + c_1x + c_2x^2 + \dots$ as the ratio of two polynomials

$$f_n(x) = \frac{a_0 + a_1x + \dots + a_nx^n}{1 + b_1x + \dots + b_mx^m} \quad (4)$$

by using the method of Padé approximants^{7,8}. The a's and b's are determined from the c's. The nature of singularities in $f(x)$ are found as well as their location. With 10 to 20 terms, one obtains results of considerable precision.

Our individual electron assembly expansions have fewer terms but, together, the large and small r_s expansions do have a considerable number of terms. We assume that they are two asymptotic forms of the same expression and attempt to use the

Padé expansion to lead us to an approximation of that expression. One of the difficulties in doing so is that our small r_s expansion includes logarithmic terms and the small and large r_s expansions are not of the same type. This fact makes a simple application of the Padé approximant method impossible. However, one can circumvent this difficulty by differentiating E_G twice with respect to r_s and expressing the results in terms of a new variable $x = r_s^2$. Then, one finds

$$r_s^4 d^2 E_G / dr_s^2 = \begin{cases} 13.258 - 1.832x^2 - 0.0622x^2 + 0.018x^6 + \dots & \text{if } x \ll 1 \\ -3.584x^2 + 9.8925x - 4.38 - tx^{-1} + \dots & \text{if } x \gg 1 \end{cases} \quad (5)$$

$$\text{if } x \gg 1 \quad (6)$$

where Eq. (5) corresponds to Eq. (3) and Eq. (6) to Eq. (2). In the latter equation an additional term tx^{-1} has been introduced with an unknown coefficient t .

Note that Eqs. (5) and (6) are counterparts of each other for the two limiting domains of x . Since the right-hand sides are simple power series in x or x^{-1} , it is now possible to apply the Padé approximants to

$$f_n(x) = r_s^4 (d^2 E_G / dr_s^2) \quad (7)$$

In the low density regime as $x \rightarrow \infty$ in Eq. (6), the first term dominates for very large x . Therefore, in Eq. (4), $m = n + 2$.

In the traditional applications of Padé approximants, one considers the expansion of a function about one point,

while here we wish to take advantage of the fact that we have independent information concerning the behavior of E_G at two points, $x = 0$ and $x = \infty$. A similar situation has arisen in the polaron problem. Sheng and Dow⁹ have applied two point Padé approximants to that problem.

The systematic procedure for the determination of the a_i 's and b_i 's in (4) from asymptotic formulae such as (5) and (6) is discussed in Appendix I. The values of these constants for the case $n = 2$ and $m = 4$ are

$$a_0 = 13.2594, \quad a_1 = 2.6600, \quad a_2 = -1.3617, \quad r_3 = -0.3676$$

$$a_4 = -0.12729, \quad b_1 = 0.2006, \quad b_2 = 0.03552$$

while those for the case $n = 3$ and $m = 5$ are

$$a_0 = 13.2594, \quad a_1 = 11.848, \quad a_2 = 1.0146$$

$$a_3 = -1.2165, \quad a_4 = -0.4557, \quad a_5 = -0.1137$$

$$b_1 = 0.8954, \quad b_2 = 0.21474, \quad b_3 = 0.03173$$

The expression for $f(x)$ given by (4) with these constants must be integrated twice in order to obtain a formula for the ground state energy of an electron gas, E_G . A systematic discussion of this integration and the resulting equations are presented in Appendix II. The computed curves for E_G as a function of r_s based on Eqs. (7) and (6) of Appendix II for $n = 2$ and 3 are plotted as (i) and (ii) in Fig. 1. The zeros of the denominator

of $f(x)$ of Eq. (4) for the $n = 2$ and $n = 3$ cases have no positive poles. Hence these approximations to E_G have no singularities in the r_s range $0 < r_s < \infty$.

Since the electron lattice is the stable form of an assembly in the low density regime while the electron gas is the stable state in the high density regime, one expects a phase transition, the lattice melting, to occur at some intermediate r_s value. Various estimates which range from $r_s = 5$ to $r_s = 2000$ have been made for the critical r_s value. Our two approximations to E_G discussed above seem to be unsatisfactory because they give no indication of the existence of such a critical value.

If we try to make the next ($n = 4$) approximation to E_G we are faced with the difficulty that we need to know the value of the coefficient t of x^{-1} in Eq. (6). This corresponds to the second order anharmonicity contribution of the electron lattice to $r_s^3 d^2 E_G / dr_s^2$. This has not yet been calculated, but we can investigate its influence by assuming that it falls in the interval $-2 < t < 7$ and plotting E_G as a function of r_s for several values of t in this range. The striking new feature of the $n = 4$, $m = 6$ approximation to $f(x)$ is that it yields a pole in $d^2 E_G / dr_s^2$ at $r_s = 14$.

The denominator of the left hand side of (7) in our present approximation is a fourth degree polynomial with one positive, one negative, and two complex roots. It is the positive root which leads to the singularity at $r_s = 14$.

Equation (7) can be rewritten as

$$d^2 E_G / dr_s^2 = (a_0 + \sum_{i=1}^6 a_i x^i) / b^4 x^3 P_6(x) \quad (9a)$$

with

$$P_6(x) = (x-a)(x+b)(x^2-2ax+c^2) \quad (9b)$$

$$c^2 = a^2 + \beta^2 \quad (10)$$

The integration of (9a) is discussed in Appendix 1. The result is

$$\begin{aligned} E_G \sim & K_1 r_s^{-2} + K_2 r_s^{-1} - (K_3 - K_4 r_s) \ln r_s - K_5 r_s^{1/2} \\ & + K_6 ((r_s - a^2) \ln(r_s^{1/2} - a)^2 - 2ar_s^{1/2}) \\ & + K_7 ((r_s - b^2) \ln(r_s^{1/2} + b)^2 + 2br_s^{1/2}) \\ & + K_8 ((r_s + (\beta^2 - a^2)) \ln[(r_s^{1/2} - a)^2 + \beta^2] - 2ar_s^{1/2}) \\ & + K_9 (4r_s^{1/2} + 2a \ln(r_s^{1/2} - a) + \beta^2) \\ & + K_{10} r_s (\tan^{-1}[(r_s^{1/2} - a)/|\beta|] - \frac{1}{2}\pi + |\beta| r_s^{-1/2}) \\ & + K_{11} \tan^{-1}[(r_s^{1/2} - a)/|\beta| - \frac{1}{2}\pi] + K_{12} \end{aligned} \quad (11)$$

The various coefficients and roots are given in Table 1 for various values of t . The parameters a_0 , b_0 , K_1 , K_2 , K_3 , and K_4 , are completely independent of t while K_G and K_{10} have the strongest t dependence. The sensitivity is greatest for those terms which are large for large r_s . However, through certain basic cancelations, E_G itself is very small for large r_s and

TABLE 1

Coefficients and Roots of the Fourth Order Padé Approximant

t	7	3	2	0	-2
a ₀	13.2594				
a ₁	1.0315	0.9720	0.9571	0.9274	0.8976
-a ₂	1.6864	1.6887	1.6893	1.6905	1.6917
-a ₃	0.2701	0.2629	0.2611	0.2575	0.2539
-a ₄	0.1406	0.1399	0.1398	0.1397	0.1395
-a ₅	0.008609	0.008360	0.008297	0.008173	0.008049
-a ₆	0.1561	0.01562	0.01562	0.01562	0.01563
b ₀	1.0	"	"	"	"
b ₁	0.07780	0.07331	0.07219	0.06994	0.06770
b ₂	0.01103	0.01086	0.01081	0.01072	0.01063
-b ₃	0.009674	0.009694	0.009714	0.009752	0.009791
-b ₄	0.00435	0.004357	0.004358	0.004359	0.004361
a	3.8158	3.7957	3.7907	3.7807	3.7706
b	4.4629	4.4814	4.4919	4.5035	4.5150
-α	0.7807	0.7672	0.7638	0.7571	0.7503
(±i)β	3.5882	3.5901	3.5906	3.5916	3.5926
K ₁	2.2099	"	"	"	"
K ₂	-0.91633	"	"	"	"
K ₃	0.0622	"	"	"	"
K ₄	0.08558	0.08627	0.08644	0.08678	0.08712
K ₅	0.00977	0.00988	0.00987	0.00982	0.00977
K ₆	0.000753	0.000773	0.000778	0.000789	0.000799
-K ₇	0.002441	0.002396	0.002385	0.002363	0.002341
-K ₈	0.008289	0.008264	0.008258	0.008245	0.009231
-K ₉	0.001158	0.001543	0.001679	0.001830	0.002020
K ₁₀	0.000645	0.000860	0.000913	0.001019	0.0011250
K ₁₁	0.100797	0.101628	0.101834	0.102241	0.102644
K ₁₂	0.062200	"	"	"	"

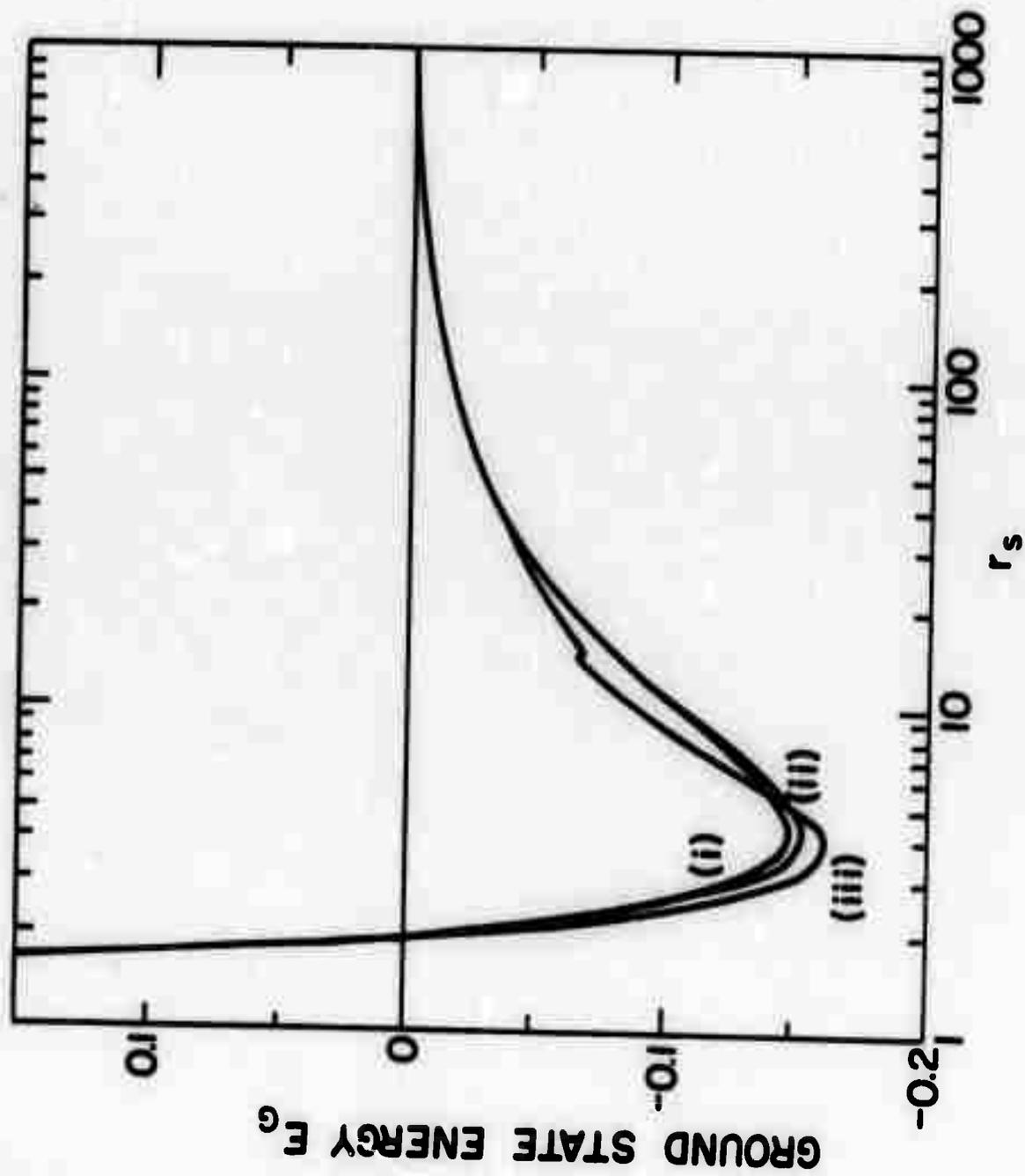
it is insensitive to t in that range. In the small r_s regime, E_G is essentially independent of t .

Our estimate of the ground state energy is plotted in Fig. 1. The most striking feature of the curve is the anomaly at $r_s \approx 14.4$. The location of the singularity depends slightly on t . That dependence on t , as well as that of the energy at the minimum and r_s value at the minimum of E_G is exhibited in Table 2. Curve (iii) corresponds to $t = \tau$.

Table 2. Certain E_G Characteristics as a Function of t

t	7	3	2	0	-2
E_G minimum	-0.16342	-0.16311	-0.16303	-0.16287	-0.16271
r_s at minimum	3.77	3.76	3.76	3.76	3.76
location of singularities	14.56	14.41	14.37	14.29	14.22

Note that while the location of the singular point depends only weakly on t , the dependence is roughly linear. At the singular r_s , the second derivative $d^2 E_G / dr_s^2$ diverges, but E_G itself is continuous. In the immediate neighborhood of singularity, the curve shows a van der Waals type loop. By using an argument similar to that due to Maxwell for the interpretation of the van der Waals loop, the Maxwell equal area construction implies that a first order phase transition occurs at the singular r_s . For large values of r_s , the curve is close



to that reported by Carr, Caldwell-Horsefall and Maradudin*. The minimum in E_G for all t values investigated in the 4th order Padé approximant lies somewhat deeper than in the 3rd order case. The rising part of the curve in the range $3.8 < r_s < 14$ might be attributed to local ordering which persists at densities higher than the melting point.

The melting of an electron lattice may be treated in accordance with Lindemann's criterion. For a bcc lattice it is $r_s = 0.4054\delta^{-4}$ where δ is the ratio of the amplitude of lattice vibration to the interlattice spacing. Using the value $r_s = 14.37$, it was found that δ is approximately 0.4.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with the University of Michigan. It was also partly sponsored by ARPA and monitored by ONR (N00014-17-C-0308).

References

1. E. Wigner, Trans. Parad. Soc. 34:678 (1938).
2. K. Fuchs, Proc. Roy. Soc. (London) 151:585 (1935).
3. R. A. Coldwell-Horsefall and A. A. Maradudin, Phys. Rev. 1:395 (1960).
4. W. J. Carr, Jr., Phys. Rev. 122:1437 (1961).
W. J. Carr, Jr., R. A. Coldwell-Horsefall and H. A. Fein, *ibid* 124:747 (1961).
5. M. Gell-Mann and K. Bruechner, Phys. Rev. 106:364 (1957).
6. D. P. DuBois, Ann. Phys. 7:174 (1959).
7. C. Domb, Phil. Mag. Suppl. 9:149 (1960).
8. G. Baker, Jr., Phys. Rev. 124:768 (1961).
9. P. Sheng and J. D. Dow, Phys. Rev. 24:1343 (1971).

Appendix I

Two Point Padé Approximants

Let us examine a situation in which two asymptotic forms are known for a function $f(x)$

$$f(x) \sim \begin{cases} A_0 + A_1x + A_2x^2 + \dots & \text{as } x \rightarrow 0 & (1a) \\ B_0x^2 + B_1x + B_2 + \dots & \text{as } x \rightarrow \infty & (1b) \end{cases}$$

We discuss a systematic manner of approximating $f(x)$ by a ratio of two polynomials

$$f_n(x) = [n+2, n] = \frac{a_0 + a_1x + \dots + a_{n+2}x^{n+2}}{1 + b_1x + \dots + b_nx^n} \quad (2)$$

Consider first the small x expansion of this ratio.

By elementary division, if the ratio is to be consistent with

(1a)

$$\begin{aligned} a_0 &= A_0, \quad a_1 = A_0B_1 + A_1 \\ a_2 &= A_0b_2 + A_1b_1 + A_2 \\ a_3 &= A_0b_3 + A_1b_2 + A_2b_1 + A_3, \quad \text{etc.} \end{aligned} \quad (3)$$

As an example, let $n = 2$ in (2). Then $b_3 = b_4 = \dots = 0$ and

$$\begin{aligned} a_0 &= A_0, \quad a_1 = A_0b_1 + A_1 \\ a_2 &= A_0b_2 + A_1b_1 + A_2 \end{aligned}$$

$$a_2 = A_1 b_2 + A_2 b_1 + A_3 \quad (4)$$

$$a_1 = A_2 b_1 + A_3$$

Similar equations can be constructed for $n = 3$, etc. The connection between the a 's, b 's, and B 's is obtained by writing (with $y = x^{-1}$).

$$(n+2, n) = (x^2 a_{n+2} / b_n) \left\{ \frac{1 + (a_{n+1}/a_{n+2})y + \dots + (a_0/a_{n+2})y^{n+2}}{1 + (b_{n-1}/b_n)y + \dots + (1/b_n)y^n} \right\} \quad (5)$$

When $n = 2$, the first two terms in the expansion imply

$$a_2 = B_0 b_2 \quad (6)$$

$$a_1 = B_0 b_1 + B_1 b_2$$

When $n = 3$, the first three terms imply

$$a_3 = B_0 b_3, \quad a_2 = B_0 b_2 + B_1 b_3 \quad (7)$$

$$a_1 = B_0 b_1 + B_1 b_2 + B_2 b_3$$

and, when $n = 4$, the first four terms imply

$$a_4 = B_0 b_4, \quad a_3 = B_0 b_3 + B_1 b_4 + B_2 b_5 + B_3 b_6 \quad (8)$$

$$a_2 = B_0 b_2 + B_1 b_3 + B_2 b_4$$

$$a_1 = B_0 b_1 + B_1 b_2 + B_2 b_3$$

The pattern is obvious for larger values of n

We are now in a position to discuss the evaluation of the (a_j) and (b_j) when $(n+3)$ A 's and n B 's are given. This is

sufficient input data to find the $(2n+3)$ constants required for the $(n+2, n)$ approximation.

When $n = 2$, the given data values would be $B_0, B_1, \Lambda_0 - \Lambda_1$. The basic equations would be the sets (6) and (4). The values of b_1 and b_2 would be determined by equating both expressions for a_0 and a_1 ,

$$a_0 = B_0 b_2 = \Lambda_1 b_2 + \Lambda_2 b_1 + \Lambda_3 \quad (9)$$

$$a_1 = B_0 b_1 + B_1 b_2 + \Lambda_1 b_2 + \Lambda_2 b_1 + \Lambda_3$$

Then b_1 and b_2 are expressible as the ratio of the two (2×2) determinants. The quantity $1 + b_1 x + b_2 x^2$ which appears in the denominator of (2) can then be written as the ratio of a (3×3) and a (2×2) determinant.

$$1 + x b_1 + x^2 b_2 = \begin{vmatrix} 1 & \Lambda_1 & \Lambda_2 \\ -x & B_0 - \Lambda_2 & 0 - \Lambda_1 \\ -x^2 & B_1 - \Lambda_1 & B_0 - \Lambda_2 \end{vmatrix} \bigg/ \begin{vmatrix} B_0 - \Lambda_2 & 0 - \Lambda_1 \\ B_1 - \Lambda_1 & B_0 - \Lambda_2 \end{vmatrix} \quad (10)$$

as the reader can easily verify. Clearly the combination $a_1 + a_2 b_1 + a_3 b_2$ has the corresponding determinant representation

$$a_1 + a_2 b_1 + a_3 b_2 = \begin{vmatrix} a_1 & \Lambda_1 & \Lambda_2 \\ -a_2 & B_0 - \Lambda_2 & 0 - \Lambda_1 \\ -a_3 & B_1 - \Lambda_1 & B_0 - \Lambda_2 \end{vmatrix} \bigg/ \begin{vmatrix} B_0 - \Lambda_2 & 0 - \Lambda_1 \\ B_1 - \Lambda_1 & B_0 - \Lambda_2 \end{vmatrix} \quad (11)$$

The numerator of (2) has the form (using 4)

$$\begin{aligned}
& a_0 + xa_1 + x^2a_2 + x^3a_3 + x^4a_4 \\
& = (A_0 + xA_1 + x^2A_2 + x^3A_3 + x^4A_4) \\
& \quad + b_1(xA_0 + x^2A_1 + x^3A_2 + x^4A_3) \\
& \quad + b_2(x^2A_0 + x^3A_1 + x^4A_2)
\end{aligned}$$

Hence

$$F_2(x) = [4, 2] = \begin{array}{c|ccc} & \Sigma_0^4 A_j x^j & A_3 & A_4 \\ \hline -x \Sigma_0^3 A_j x^j & & B_0 - A_2 & 0 - A_3 \\ -x^2 \Sigma_0^2 A_j x^j & & B_1 - A_1 & B_2 - A_2 \\ \hline 1 & & A_3 & A_4 \\ x & & B_0 - A_2 & 0 - A_3 \\ -x^2 & & B_1 - A_1 & B_2 - A_2 \end{array} \quad (12)$$

When $n = 3$, the basic equation for the determination of the b 's as a ratio of two determinants follow from (3) with $b_4 = b_5 \dots = 0$. And (7):

$$a_5 = A_2 b_3 + A_3 b_2 + A_4 b_1 + A_5 = B_0 b_3 \quad (13a)$$

$$a_4 = A_1 b_3 + A_2 b_2 + A_3 b_1 + A_4 = B_0 b_2 + B_1 b_3 \quad (13b)$$

$$a_3 = A_0 b_3 + A_1 b_2 + A_2 b_1 + A_3 = B_0 b_1 + B_1 b_2 + B_2 b_3 \quad (13c)$$

Once the b 's are written as a ratio of determinants, the same can be done for the denominator of (2) when $n = 3$. Finally, $F_1(x)$ is easily shown to have the form

$$F_3(x) = [5, 3] = \frac{(S_5, -xS_4, -x^2S_3, -x^3S_2)}{(1, -x, -x^2, -x^3)} \quad (14a)$$

where we define $S_n = \sum_{j=0}^n A_j x^j$ and

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \equiv \begin{vmatrix} \alpha_1 & A_3 & A_4 & A_5 \\ \alpha_2 & B_0 - A_2 & 0 - A_3 & 0 - A_4 \\ \alpha_3 & B_1 - A_1 & B_0 - A_2 & 0 - A_3 \\ \alpha_4 & B_2 - A_0 & B_1 - A_1 & B_0 - A_2 \end{vmatrix} \quad (14b)$$

Also,

$$F_4(x) = [6, 4] \equiv \frac{(S_6, -xS_5, -x^2S_4, -x^3S_3, -x^4S_2)}{(1, -x, -x^2, -x^3, -x^4)} \quad (15a)$$

$$(\alpha_1, \dots, \alpha_5) \equiv \begin{vmatrix} \alpha_1 & A_3 & A_4 & A_5 & A_6 \\ \alpha_2 & B_0 - A_2 & 0 - A_3 & 0 - A_4 & 0 - A_5 \\ \alpha_3 & B_1 - A_1 & B_0 - A_2 & 0 - A_3 & 0 - A_4 \\ \alpha_4 & B_2 - A_0 & B_1 - A_1 & B_0 - A_2 & 0 - A_3 \\ \alpha_5 & B_3 - 0 & B_2 - A_0 & B_1 - A_1 & B_0 - A_2 \end{vmatrix} \quad (15b)$$

The generalization to any $F_n(x)$ is obvious. Similar expressions exist for two point Padé approximants for $[n+\mu, n]$.

Appendix II

Integration of d^2E_G/dr^2

In this appendix, we are concerned with the integration of expressions (with $x = r_s^{1/2}$ and $b_0 = 1$) such as

$$d^2E_G^{(n)}/dr^2 = x^{-2} \left(\sum_0^{n+2} a_j x^j \right) / \left(\sum_0^n b_j z^j \right) \quad (1)$$

The basic integrals required are

$$I_m(1) = \int x^{-m} dx / (1+x)$$

Note that

$$I_0(1) = \log(1+x); \quad I_1(1) = \log[x/(1+x)]$$

$$I_1(1) = -x^{-1} - \log[x/(1+x)]$$

while

$$I_m + I_{m-1} = \int (x^{-m} + x^{-m+1}) dx / (1+x) = 0x^{-(m-1)} / (m-1),$$

$$I_m = (I_m + I_{m-1}) - (I_{m-1} - I_{m-2}) + \dots + (-1)^m (I_2 + I_1) - (-1)^m I_1$$

$$= \sum_{s=1}^{m-1} (-1)^{m+s} s^{-1} x^{-s} - (-1)^m \log[x/(1+x)]$$

Furthermore,

$$I_m(y) = \int x^{-m} dx / (x+y)$$

$$= y^{-m} \left((-1)^{m+1} \log[x/(y+x)] + \sum_{s=1}^{m-1} (-1)^{m+s} (y/x)^s / s \right) \quad (2)$$

The quantities $[g_n(x)]^{-1} \equiv (1+b_1x + \dots + b_nx^n)^{-1}$ in Eq. (1) have the following partial fraction representations for $n = 2, 3, 4$ respectively:

$$b_2^{-1} \{ [(x-\alpha_1)\Delta_{12}]^{-1} + [(x-\alpha_2)\Delta_{21}]^{-1} \}; \Delta_{ij} \equiv \alpha_i - \alpha_j \quad (3a)$$

$$b_3^{-1} \{ [(x-\alpha_1)\Delta_{12}\Delta_{13}]^{-1} + [(x-\alpha_2)\Delta_{21}\Delta_{23}]^{-1} + [(x-\alpha_3)\Delta_{31}\Delta_{32}]^{-1} \} \quad (3b)$$

$$b_4^{-1} \{ [(x-\alpha_1)\Delta_{12}\Delta_{13}\Delta_{14}]^{-1} + \dots + [(x-\alpha_n)\Delta_{n1}\Delta_{n2}\Delta_{n3}]^{-1} \} \quad (3c)$$

The $\{\alpha_j\}$ for a given n are the roots of $g_n(\alpha) = 0$.

Since $dE_G/dr_s \rightarrow 0$ as $r_s \rightarrow \infty$, the integration of (1) from r to ∞ yields, employing (2) and (3),

$$\begin{aligned} dE_G^{(n)}/dr = & (2/b_n) \sum_{m=0}^{n+2} a_m \frac{J_{7-m}(r, -\alpha_2)}{\Delta_{12}\Delta_{13}\dots\Delta_{2n}} + \frac{J_{7-m}(r, -\alpha_2)}{\Delta_{21}\Delta_{23}\dots\Delta_{2n}} \\ & + \dots + \frac{J_{7-m}(r, -\alpha_n)}{\Delta_{n1}\Delta_{n2}\dots\Delta_{n,n-1}} \end{aligned} \quad (4)$$

if $n > 4$, where

$$J_k(r, -\alpha) = -\alpha^{-k} \log[r^{1/2}/(r^{1/2}-\alpha)] + \sum_{p=1}^{k-1} p^{-1} \alpha^{p-k} r^{-p/2} \quad (5)$$

This function is easily integrated and the constant of integration can be chosen so that the integral vanishes as $r \rightarrow \infty$ as is required by our problem. If α is real and positive, and $r^{1/2} < 0$, the term $(r^{1/2}-\alpha)$ should be written as $|r^{1/2}-\alpha|$. The form of the integral when $k > 4$ is

$$F_k(r, -\alpha) = 2 \sum_{p=1}^{k-1} \frac{\alpha^{p-k} r^{1-4p}}{p(2-p)} + r^k \alpha^{1-k} + \left(\frac{1}{2}\right) \alpha^{2-k} + \alpha^{-k} (r-\alpha^2) \log[(r^k-\alpha)/r^k] \quad (6a)$$

When $k = 2$ and 3 , the term expressed as a sum over p is omitted, and when $k = 1$,

$$F_1(r, -\alpha) = \alpha^{-1} (\alpha r^k - \alpha^2 \log r^k + (r-\alpha^2) \log[(r^k-\alpha)/r^k]) \quad (6b)$$

Again, when α is real and positive and $r^k < \alpha$, the terms $(r^k-\alpha)$ in (6a) and (6b) should be replaced by $|r^k-\alpha|$. The final expression for $E_G^{(n)}(r_s)$ is

$$E_G^{(n)}(r_s) = (2/b_n) \sum_{m=0}^{n+2} a_m D_{n-m}^{(n)}(r_s) \quad (7a)$$

$$D_k^{(n)}(r) \equiv \frac{F_k(r, -\alpha_1)}{\Delta_{12} \dots \Delta_{kn}} + \frac{F_k(r, -\alpha_2)}{\Delta_{21} \Delta_{23} \dots \Delta_{2n}} + \dots + \frac{F_k(r, -\alpha_n)}{\Delta_{n1} \dots \Delta_{n,n-1}} \quad (7b)$$

Several technical points are to be observed in the application of these formulae. First, various symmetrical functions of the roots α_j of $g_n(\alpha) = 0$ occur naturally. For example,

$$\alpha_1 + \alpha_2 + \dots + \alpha_n = -b_1/b_n; \quad \alpha_1 \alpha_2 \dots \alpha_n = (-1)^n/b_n, \text{ etc.} \quad (8)$$

A recurrence formula exists for the combination

$$\delta_{-j}^{(n)} = \frac{\alpha_1^{-j}}{\Delta_{12} \dots \Delta_{kn}} + \dots + \frac{\alpha_n^{-j}}{\Delta_{n1} \dots \Delta_{n,n-1}}, \quad (9a)$$

$$\delta_{-j}^{(n)} = -b_1 \delta_{-j+1}^{(n)} - b_2 \delta_{-j+2}^{(n)} \dots - b_n \delta_{-j+n}^{(n)} \quad (9b)$$

This is easily deduced from the relation

$$\alpha^{-j} = -b_1 \alpha^{-j+1} - b_2 \alpha^{-j+2} \dots - b_n \alpha^{-j+n} \quad (10)$$

When $n = 2, 3, 4$ the initial values to be used with (9b) are, respectively,

$$\delta_0^{(2)} = 0, \quad \delta_{-1}^{(2)} = -b_2 \quad (11a)$$

$$\delta_1^{(3)} = \delta_0^{(3)} = 0, \quad \delta_{-1}^{(3)} = -b_3 \quad (11b)$$

$$\delta_2^{(4)} = \delta_1^{(4)} = \delta_0^{(4)} = 0, \quad \delta_{-1}^{(4)} = -b_4 \quad (11c)$$

In the case in which one pair of roots, say $\alpha_1 = -a+ib$ and $\alpha_2 = -a-ib$ are complex, one notes that

$$\log(r^h + a - ib) = \frac{1}{2} \log[(a+r^h)^2 + b^2] - i \tan^{-1}[b/(a+r^h)] \quad (12)$$

Through the appropriate combinations of α_1 and α_2 , which appear in our problem, no imaginary terms appear in final formulae for $E_G^{(n)}$. Finally, some care has to be taken in the integration which yields logarithms when a given α_i is real and positive. However, the singularity at α_i leads to no difficulty in performing the required integrations. As was stated earlier in this case, the logarithm terms in (6) are to be replaced by $\log(|r^h - \alpha|/r^h)$.

A RIGID-PLASTIC MODEL OF SPALL
FRACTURE BY HOLE GROWTH

F. A. McClintock

Abstract

A numerical procedure is formulated for calculating spall fracture by hole growth, as approximated by the spherical growth of holes in rigid-plastic material. An order-of-magnitude estimate indicates that inertia effects are more important than rate effects and that this model is likely to be appropriate for 6061-T6 aluminum alloy at stresses up to twice the static strength.

Preceding page blank

A RIGID-PLASTIC MODEL OF SPALL
FRACTURE BY HOLE GROWTH

F. A. McClintock

Introduction

In some metals, including 6061-T6 aluminum alloy, spall fracture occurs by the growth and coalescence of holes. This mode is shown by both scanning electron micrographs (Joyce et al., 1968) and metallographic cross-sections obtained by Viola (see McClintock et al., 1968). Furthermore, the sides of the growing holes are so smooth, even though curved, that it seems relevant to calculate the hole growth from continuum mechanics, starting from some initial array of holes. Under dynamic loading the resulting calculations of elastic-plastic wave interactions require detailed numerical methods and computer times of the order of hours. An approximate analysis is therefore desirable as a first step in investigating the generalized empirical fracture criterion of Tuler and Butcher (1968). This criterion predicts that fracture occurs at a stress $\underline{\sigma}$ and a time \underline{t}_f when

$$\int_0^{\underline{t}_f} (\sigma - \sigma_0)^\lambda dt = K \quad (1)$$

The object of the present analysis is to test the form of the above relation and to propose another or to derive the magnitudes

of the constants σ_0 , λ , and K from metallographic data on hole size and spacings, as well as the stress-strain behavior.

Conceivably the time dependence of the fracture criterion (1) could be due to strain rate effects. Typically reported values for the exponent λ are of the order of 2 which would require a much greater strain rate effect than is normally observed. It seems more likely that the time dependence arises from inertia effects on the stress distribution around the growing holes. If so, the limit load calculations are not applicable.

For very high stress levels, elastic strains will be important. At a somewhat lower stress level, not only will the elastic strains be relatively unimportant but also there will be sufficient time before fracture for the elastic stress waves to run back and forth, giving a quasi-static stress distribution from the elastic point of view. We therefore consider a rigid-plastic material with inertia terms in the equilibrium equations.

For a first approximation neglect strain hardening. Consider layers of holes normal to the direction of wave motion as shown in Figure 1. Interactions between the holes in a layer will make each hole behave as if it were in a cylindrical tube of constant cross-section with zero shear stress. Such boundary conditions would give a two or three dimensional problem in plasticity. For simplicity it is desirable to assume spherical growth of the holes, as also appears justified from the metallographic cross-sections. The question then arises as to what

boundary conditions on the outside of the spheres would plausibly simulate the actual problem. The applied load consists of plane waves traversing the layers of holes. As the holes grow and make a given layer porous, there are also unloading waves which affect neighboring layers. In other words, the decrease in traction exerted by the net area of one layer will appear as an unloading wave at another layer at a time later by an amount of the order of the distance travelled divided by the velocity of an elastic wave. (This appears to be the one point at which the elastic wave velocity should appear.)

Finally, fracture occurs when the radius of the hole becomes approximately the initial half-distance between holes, R . In a spherical shell the boundary conditions would be now applied at the somewhat greater current external radius and fracture would not yet have occurred. In the actual case of course, material is flowing axially, normal to the spall plane, so the approximation of fracture when $R = a$ should be reasonable.

Analysis

The initial and boundary conditions are indicated for a simple case in Figure 2, in which the positions of various layers of holes are plotted versus time. For convenience we speak of a row of holes rather than a series of layers. Initially, positive or negative pulses $\delta\sigma_p$, $\delta\sigma_n$ are specified, according to the direction of travel along the row. In the diagram, they travel to the right diagonally upwards or down-

wards, respectively. The units of time are the travel time of a wave with velocity c across the half-spacing R of the holes, $\delta t = R/c$. Initial values of the applied stress \underline{u}_R on a spherical shell of outer radius R are given at one time unit before the beginning of the actual calculation. Initial values of the hole radius \underline{a} are given just at the start of the calculation. There is a possibility that stress pulses will enter the field of interest at later times, so $\underline{\delta\sigma}_p$ must be defined below the lower row of holes and $\underline{\delta\sigma}_n$ above the upper. The normalized times \underline{i} at which these initial conditions are to be set are determined from the later discretization of the equations.

$$\begin{aligned}
 \text{Given: } & a(3,j) \quad j = 2, j_{\max}^{-1} \\
 & \sigma_R(2,j) \quad j = 2, j_{\max}^{-1} \\
 & \delta\sigma_p(1,j), \delta\sigma_n(1,j) \quad j = 2, j_{\max}^{-1} \\
 & \delta\sigma_p(i,1), \delta\sigma_n(i,j_{\max}) \quad i = 1, i_{\max}^{-1} \text{ by 2's.}
 \end{aligned} \tag{2a}$$

In addition to the initial conditions to be read in, it turns out to be necessary to set to zero the initial radial velocity of the hole, \dot{a} , the initial radial acceleration of the hole, \ddot{a} , and the unloading wave from prior hole growth, $\delta\sigma_g$.

$$\begin{aligned}
 \dot{a}(3,j) &= 0 \quad j = 2, j_{\max}^{-1} \\
 \ddot{a}(2,j) &= 0 \quad j = 2, j_{\max}^{-1} \\
 \delta\sigma_g(1,j) &= 0 \quad j = 2, j_{\max}^{-1}.
 \end{aligned} \tag{2b}$$

The main computational loop, carried out for each time step $\underline{2i}$ and hole position \underline{j} , is now begun. First, the radial

velocity of the hole at the middle of the next time step is estimated from the prior acceleration. For brevity, where j is constant in an equation, denote $a(i,j)$ by a_i , for example.

$$\dot{a}_i = \dot{a}_{i-1} + \ddot{a}_{i-1} \delta t. \quad (3)$$

If the holes appear to be shrinking their velocity is set to zero. Later a check will be made to insure that the yield strength in compression has not been exceeded.

As after all further calculations of the hole radius a check is made to insure that fracture has not yet occurred in the sense that

Next the hole radius is estimated from the velocities

$$a_i = a_{i-1} + \frac{1}{2}(\dot{a}_{i-1} + \dot{a}_i)\delta t. \quad (4)$$

As after all further calculations of hole growth, a check is made that fracture has not yet occurred in the sense that $\underline{a} = \underline{R}$. If $\underline{a} \geq \underline{R}$, the program is terminated.

Now calculate the stress pulses that will change the applied stress at \underline{R} . As indicated in Fig. 3, an incoming plastic stress pulse will be partly reflected by a hole. For conservation of momentum the fraction transmitted must be unity - the fraction reflected as a negative wave. The fraction transmitted is assumed to be proportional, with the factor \underline{R}_f to the area fraction of the holes in that layer, assuming the holes in a square array:

$$T_r = 1 - R_f^2 a^2 / R^2. \quad (5)$$

The positive- and negative-travelling waves can now be calculated:

$$\begin{aligned} \delta\sigma_p(i,j) &= T_r \delta\sigma_p(i-2, j-1) + (1-T_r)\delta\sigma_n(i-2, j+1) \\ \delta\sigma_n(i,j) &= (1-T_r) \delta\sigma_p(i-2, j-1) + T_r \delta\sigma_n(i-2, j+1). \end{aligned} \quad (6)$$

It would seem that in some way a loss term due to radial dispersion of waves in the plastically deforming field should be taken into account, but that will be left as a matter for future work. Next calculate the current stress in the element due to the pulses that have just passed it. In addition there will be a stress pulse due to the previous growth of holes which, it turns out, is most easily and accurately calculated at odd intervals of time. For instance, the hole growth at time δt will radiate out to the radius R at time $2\delta t$ and the effect will reenter the hole at time $3\delta t$, so it is included in the calculation at this time:

$$\sigma_R(i,j) = \sigma_R(i-2,j) + \delta\sigma_p(i,j) + \delta\sigma_n(i,j) + \delta\sigma_g(i-3,j) \quad (7)$$

Next check that the resulting radial stress is greater (less negative) than that required for compressive yielding from a static condition:

$$\sigma_R = 2Y \ln(a/R) \quad (8)$$

If the static compressive strength is exceeded, the calculation is stopped.

The stress-acceleration relation for a rigid-plastic spherical shell of non-hardening material with yield strength Y is given by Hopkins (1960). For a material with density ρ , hole radius a , hole velocity \dot{a} , acceleration \ddot{a} , and with zero applied stress at the surface of the hole, the radial stress σ_r at the radius r of the hole is

$$\sigma_r = \rho(\ddot{a}a + 2\dot{a}^2)(1 - a/r) - \frac{\rho\dot{a}^2}{2}(1 - (a/r)^4) + 2Y \ln(r/a). \quad (9)$$

The acceleration can be found from Eq. 9 in terms of the radial stress σ_R at R . In this and subsequent equations it is convenient to normalize distance in terms of R , velocity by the uniaxial strain wave velocity c , time by R/c , and stress by ρc^2 , which is the modulus of elasticity for one component of normal strain alone. The normalized acceleration of the hole is then

$$\frac{R\ddot{a}}{c^2} = \left\{ \left[\frac{\sigma_R}{\rho c^2} + \frac{2Y}{\rho c^2} \ln \frac{a}{R} + \frac{1}{2} \frac{a^2}{c^2} (1 - (a/r)^4) \right] / (1 - a/R) - \frac{2\dot{a}^2}{c^2} \right\} / \left(\frac{a}{R} \right). \quad (10)$$

Because Eq. 10 involves the radius and radial velocity, which in turn depends on the acceleration, Eqs. 3, 4 and 10 must be solved iteratively. k_{\max} iterations will be carried out. The first step is to integrate to determine the velocity from a modified form of Eq. 3, but assuming a step change in acceleration at $i - 1$, and a constant acceleration until i :

$$\dot{a}_i = \dot{a}_{i-1} + \ddot{a}_i \delta t. \quad (11)$$

Once again, if the hole appears to be shrinking, make it rigid and set the acceleration to zero. Next, for calculating the hole radius, Eq. 4 still applies and can be stated as

$$a_i = a_{i-1} + \dot{a}_{i-1} \delta t + \frac{1}{2} \ddot{a}_i \delta t^2. \quad (12)$$

With the improved estimates of hole radius and rate, return to Eq. 10 for another iteration, until Eqs. 10 through 12 have been evaluated k_{\max} times. After calculating the radius, check for fracture as before.

As a matter of interest, estimate the mean traction carried across a section containing the hole. If the holes have neither velocity nor acceleration, the mean traction would be the same as σ_R . With a growing hole, however, the two tractions will be different. As long as the hole is growing,

$$\sigma_\theta = \sigma_r = Y \quad (13)$$

Substitution of Eq. 13 into Eq. 9, integration, dividing by the cross sectional area πR^2 to get the stress, and normalizing yield

$$\frac{(\sigma_\theta)_{av}}{\rho c^2} = \left(\frac{R\ddot{a}}{c^2} \frac{a}{R} + \frac{2\dot{a}^2}{c^2} \right) \left(1 - \frac{a}{R} \right)^2 + \left(\frac{a}{c} \right)^2 \left(1 - \left(\frac{a}{R} \right)^2 \right)^2 / 2 - \frac{ZY}{\rho c^2} \ln \frac{a}{R} \quad (14)$$

An estimate of the unloading pulses due to hole growth can be made from the volumetric rate of hole growth, assuming

that an increment in volumetric rate \dot{V} divided by the cross sectional area gives rise to a uniaxial velocity increment δv in either direction:

$$\delta v = \frac{1}{2} \delta \dot{V} / \pi R^2 = \frac{1}{2\pi R^2} \frac{d^2 V}{dt^2} \delta t = \frac{1}{2\pi R^2} \frac{d^2}{dt^2} \left(\frac{4\pi a^3}{3} \right) \delta t;$$

$$\delta v = \left[2 \left(\frac{a}{R} \right)^2 \ddot{a} + 4 \frac{a \dot{a}^2}{R^2} \right] \delta t \quad (15)$$

The velocity increment δv in a medium with a modulus for uniaxial normal strain of $E_c = \rho c^2$ produces an unloading wave of magnitude $-\delta \sigma_g$:

$$-\delta \sigma_g = E_c \delta v / c = \rho c \delta v \quad (16)$$

Combining Eqs. 15 and 16 and normalizing gives

$$\frac{\delta (\sigma_0)_{av}}{\rho c^2} = - \left[\left(\frac{a}{R} \right)^2 \left(\frac{\ddot{a} R}{c^2} \right) + 4 \frac{a}{R} \left(\frac{\dot{a}}{c} \right)^2 \right] \frac{\delta t}{R/c} \quad (17)$$

In preparation for the next time step, calculate for the hole rate and radius of the following odd interval of time using Eqs. 11 and 12 with $2\delta t$ substituted for δt . Again, checks are made that the holes do not appear to be shrinking and that they have not yet grown to fracture.

With the calculation completed for hole j at time i , attention is directed to hole $j+1$ and so on to $j_{\max}-1$, at which time the time index is incremented by 2 intervals, $2\delta t = 2R/c$.

The Fortran computer program is listed in the Appendix. Output consists of hole radius a and applied stress σ_R for each hole at each interval of time step $2R/c$. If desired, a complete set of output data is available.

Elastic Effects in Slow Expansion of Spherical Holes

As one indication of the range of validity of the preceding analysis, consider the stress levels required for slow expansion of spherical holes in the absence of inertia terms, but taking elasticity into account. Hill (1950) gives the following equation for the radial stress required to expand infinitesimal holes in a non-hardening elastic-plastic material with modulus of elasticity E and yield strength Y :

$$\frac{\sigma_R}{Y} = [1 + \ln(E/1.5 Y)]/1.5 \quad (18)$$

Values of δ_R/Y are presented in Table 1 for a variety of values of Y/E . In a rigid-plastic continuum, an infinite stress would be required to expand infinitesimal holes ($a/R \rightarrow 0$). Therefore, the rigid-plastic analysis becomes invalid when the holes are small enough so that the applied stress σ_R required to expand the holes according to Eq. 8 exceeds that to expand infinitesimal holes in an elastic medium, given by Eq. 18. Limiting values of a/R according to this criterion are given in Table 1. For a hole ratio of 0.2 the calculations might be valid to $Y/E = 0.008$, which represents fairly high strength aluminum, titanium or steel

alloys, whereas for a hole radius ratio of 0.1 only relatively soft materials with $Y/E = 0.001$ to possibly 0.002 can be calculated.

Order-Of-Magnitude Estimate of Inertia

For 6061-T6 aluminum alloy, taking $a/R = 0.1$ to 0.2 and a yield strength Y of 4×10^9 dyne/cm² gives a static strength of 13 to 18×10^9 dyne/cm². Consider an applied stress of, say, double this value. From Eq. 14, at high velocities most of the stress will be used in maintaining a velocity rather than in further acceleration. Then from the first term of Eq. 14, with $\rho c^2 = 10^{12}$ dyne/cm², $c = 0.6 \times 10^6$ cm/sec, taking a mean value of $a/R = \frac{1}{2}$ during growth,

$$\frac{\dot{a}}{c} = \sqrt{\frac{2(\sigma_{\theta av}/\rho c^2)}{3}} \approx \sqrt{.01} \approx 0.1 \quad (19)$$

The fracture time is then of the order of $R/\dot{a} = R/0.1 c$. For $R \approx 10^{-3}$ cm, the time is then

$$t_f = \frac{10^{-3}}{(0.1)(0.6 \times 10^6)} = 1.6 \times 10^{-9} \text{ sec.} \quad (20)$$

This is of the order of that observed by Viola in shot 930.

Order-Of-Magnitude Estimate of Strain Rate Effect

6061-T6 alloy exhibits almost negligibly small strain rate effect under ordinary conditions. Here, let us assume a power law relation $\sigma = \sigma_1 (\dot{\epsilon}/\dot{\epsilon}_1)^m$, with $m = 1/60$ to $1/120$ for

steels and brass (MacGregor and Fisher, 1945). The strain rates of the order of unity in 10^{-6} sec. instead of 10^3 sec., correspond to a rise in stress by a factor of

$$10^{11}/(60 \text{ to } 120) = 1.5 \text{ to } 1.24$$

Thus strain rate effects may in some metals be comparable to inertia effects.*

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

*C. R. Hoggatt and R. F. Recht, in "Dynamic Stress-Strain Relationships Obtained from an Expanding Ring Experiment", Paper 1487 1969 SESA Spring Meeting, show a 20% effect for 6061-T6 on changing the rate by 10^6 , corresponding to an exponent of $1/77$, as reported by several workers.

REFERENCES

1. Chadwick, P., Cos, A. D., Hopkins, H. G., 1964, "Mechanics of Deep Underground Explosions", Phil. Trans. Roy. Soc. A256, 235-300.
2. Cristescu, N., 1968, "Dynamic Plasticity", Appl. Mech. Rev. 21, 658-659.
3. Hopkins, H. G., 1960, "Dynamic Expansion of Spherical Cavities in Metals", Progress in Solid Mechanics 1, 85-164.
4. Hunter, S. C., 1957, "The Propagation of Spherically Symmetric Disturbances in Ideally Plastic Materials", Conf. on The Properties of Materials at High Rates of Strain, Institution of Mechanical Engineers, London, 147-155.
5. Joyce, J., Viola, J. T., McClintock, F. A., 1968, Report 17, Preliminary Reports, Memoranda, and Technical Notes of the ARPA Materials Summer Conference, Materials Science Division, Advanced Research Projects Agency.
6. Levy, N., 1969, "Application of the Finite Element Method to Large Elastic-Plastic Problems of Fracture Mechanics", Ph.D. Thesis, Division of Engineering, Brown University.
7. MacGregor, C. W., Fisher, J. C., 1945, "Tension Tests at Constant True Strain Rates", J. Appl. Mech., 12.
8. McClintock, F. A., Viola, J. T., Wilkins, M., 1969, "Proposed Analysis of Dynamic Hole Growth to Fracture", Paper 18, Preliminary Reports, Memoranda, and Technical Notes of the ARPA Materials Summer Conference, Materials Science Division, Advanced Research Projects Agency.
9. Mok, 1967, "The Dynamic Expansion of a Spherical Cavity in an Elastic-Perfectly Plastic Material", Ballistic Research Laboratories, Report 1357, Aberdeen Proving Ground, Maryland.
10. Rice, J. R., 1968, "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks", J. Appl. Mech. 35, 379-386.

TABLE 1

Stress Ratios for Static Expansion of Holes in
 an Infinite Elastic-Plastic Medium and in a
 Finite Rigid-Plastic Medium

Infinite Elastic-Plastic Medium, from Eq. 18

Y/E	0.001	0.002	0.003	0.008
σ_R/Y	5.00	4.53	4.07	3.61

Finite Rigid-Plastic Medium, from Eq. 8

a/R	0.2	0.1	0.05	0.025
σ_R/Y	3.22	4.61	5.99	7.38

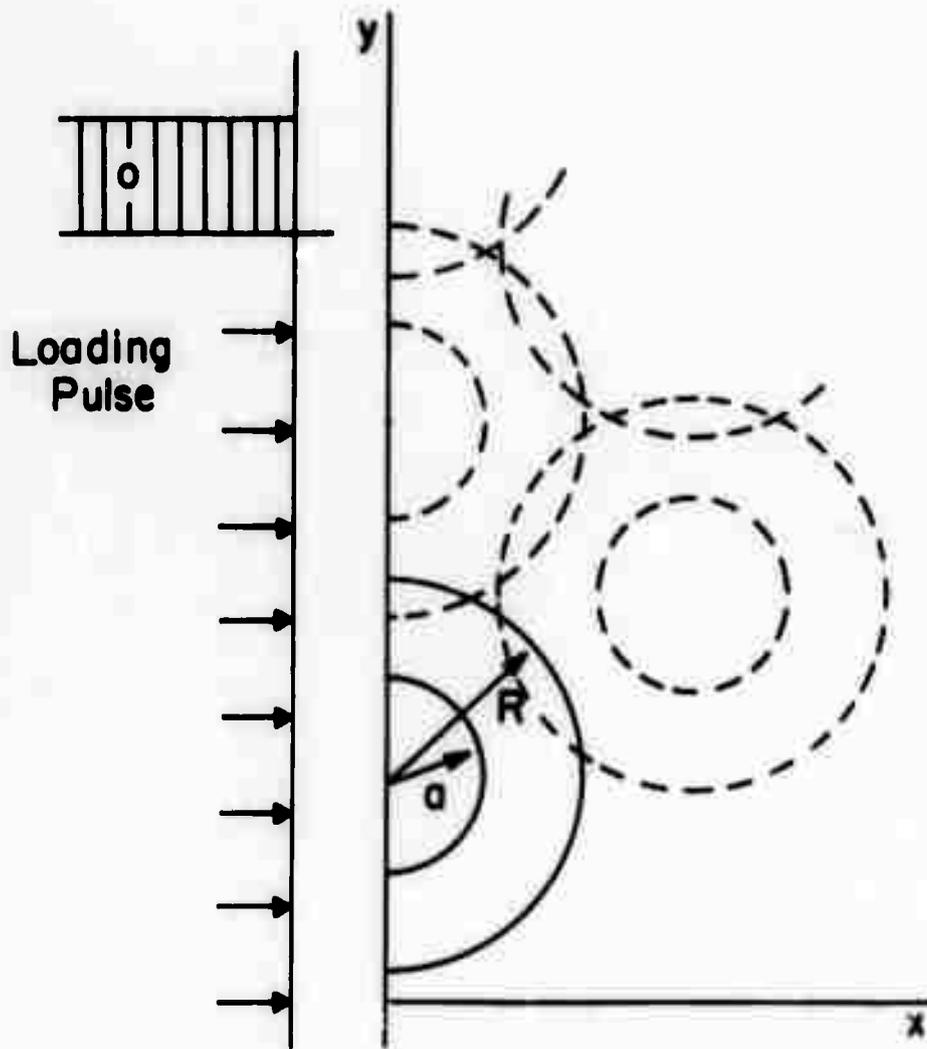


Figure 1: Approximation of spall by growing spherical holes of internal radius \underline{a} and external radius \underline{R} .

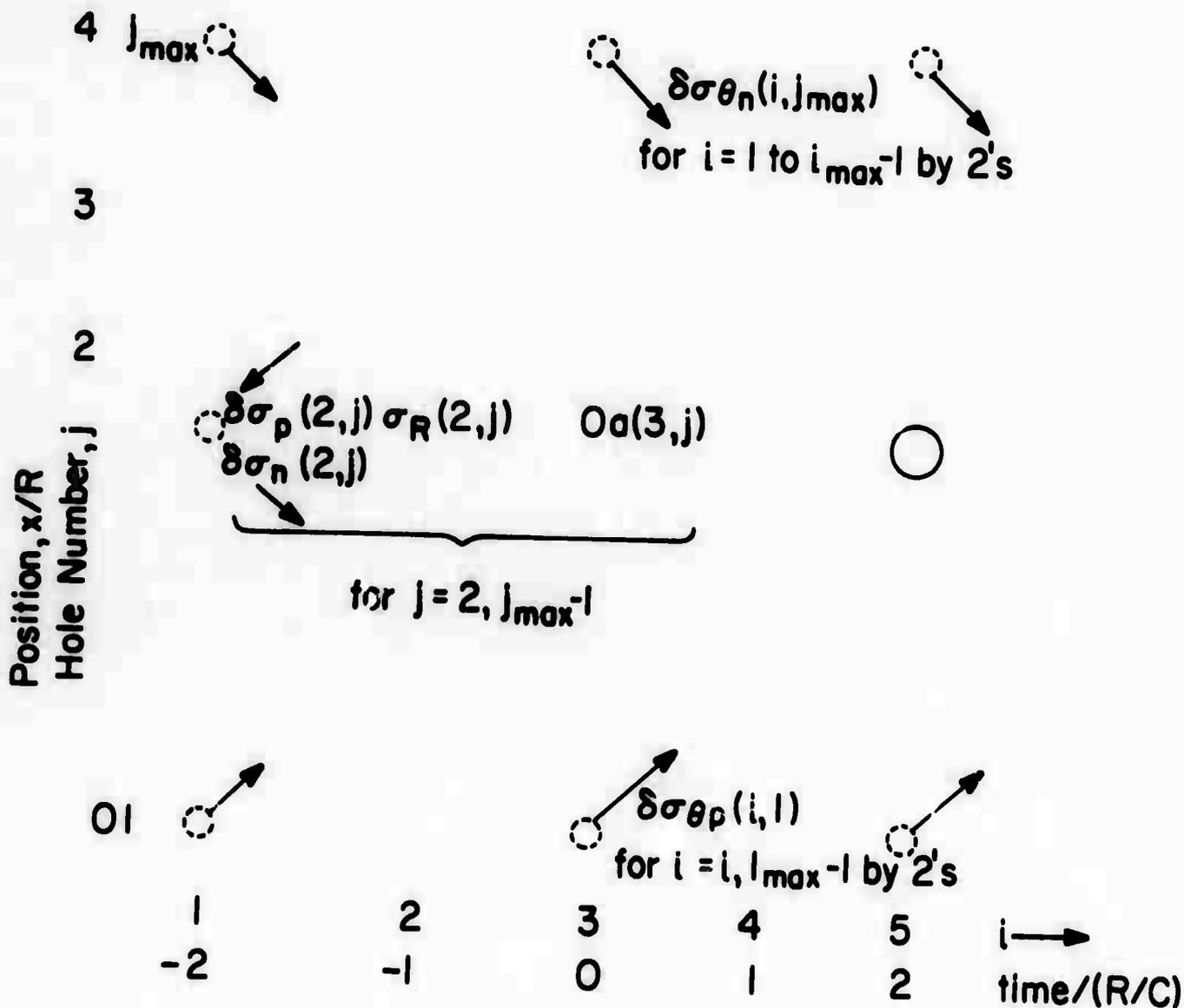


Figure 2: Array of input data for two planes of holes ($j_{max} = 4$).

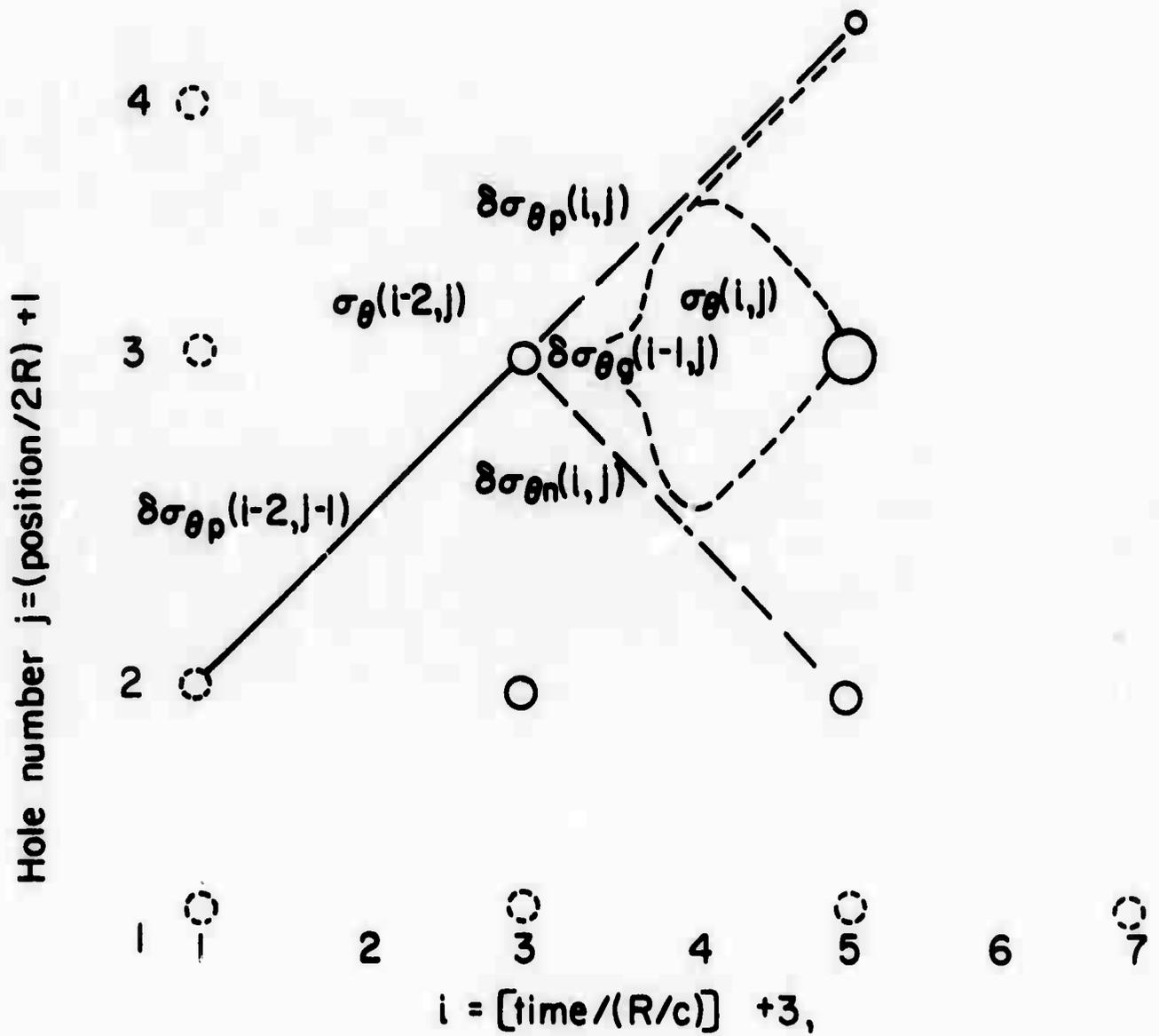


Figure 3: Transmission and reflection of positive-traveling pulse $\delta\sigma_{\theta p}$ and addition of pulse due to hole growth, $\delta\sigma_{\theta g}$.

SD EFFECT

D. C. Drucker

Abstract

The distinction between true and apparent strength-differential (SD) effects must be sharpened and definitive experiments performed to evaluate recent data and place plasticity theory in proper context.

SD EFFECT

D. C. Drucker

Additional attention has been devoted, with Morris Cohen and John Hirth, to ever increasing reports of a yield and flow strength differential (SD) for martensitic steels and other materials with the stress-strain curve for simple compression significantly above the curve for simple tension.

A difference between the yield strength in tension and compression can result from a wide variety of causes and be strongly dependent on the offset chosen to define yield. Prior deformations, or phase changes, or differential temperature contractions may produce large residual stresses on a macroscopic or on a dislocation scale or both which favor shear in one direction over its reverse. Microcracks or decohesion between particles and matrix favor tension yielding over compressive while Bauschinger and allied effects can go either way. Residual effects tend to decrease with increasing plastic strain and decreasing plastic modulus. Void opening should lead to an increasing SD with plastic straining.

SD may be thought of simply as the difference between the magnitude of the yield strength in compression and in tension, but the physical effect most authors seem to wish to

describe is the effect of hydrostatic pressure on the flow strength in shear, an effect examined experimentally by Bridgman some years ago and Rauch, et al recently. The customary assumption, based on Bridgman's results primarily, is that the SD is negligible in the range of hydrostatic pressure of \pm the yield strength. Yet Rauch, et al report an enormous influence of hydrostatic pressure on martensitic steel.

Chait reports a small but not negligible SD effect (10% or so) which agrees with Rauch and Leslie. Olsen and Ansell give results for TD-Nickel, and Rari and Gibala for niobium with dissolved oxygen, all of which reopen the question.

Except for the remarkable and totally surprising direct observation by Rauch, et al that the addition of somewhat over 100,000 psi hydrostatic pressure raised the magnitude of the compressive yield strength and flow curve by almost 100,000 psi, most comparisons are of the flow curves in simple tension and compression. True SD effects are distinguished in the literature from apparent by their persistence as plastic strain increases well beyond the 0.002 plastic strain offset which arbitrarily defined yield strength. Residual stress effects are uniformly described as apparent rather than true SD.

An attempt will be made to sharpen the distinction between true and apparent SD, to indicate effects which lead

to some confusion, to describe the severe limitations on behavior imposed by conventional plasticity theory of both the simple and complex work-hardening variety, and to suggest some more definitive experiments in both the elastic and plastic domain. Although the giant SD effects reported by Rauch, et al can be brought within the framework of plasticity theory, every indication is that the results of experiments not yet run will limit those true SD effects which lie within plasticity theory to the far more moderate range reported by Rauch and Leslie and confirmed by Chait (say 20,000 psi for a 250,000 psi yield strength steel).

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

ENERGY TO CREATE NEW SURFACE
DURING CRACK PROPAGATION

D. C. Drucker

Abstract

An understanding of the effect of microstructural and atomic features and properties along with the influence of corrosive and other environments depends critically on the energy needed to create new surface, an energy which is neither the thermodynamic equilibrium surface energy nor the energy needed to deform bulk material.

ENERGY TO CREATE NEW SURFACE
DURING CRACK PROPAGATION

D. C. Drucker

An incisive experimental determination of the likelihood of brittle or quasi-brittle crack initiation and propagation in the prototype requires a duplication of the local conditions of stress and strain history in a test specimen. This enormously complicated problem in mechanics of continua reduces in the case of small scale yielding (plastic zones of small size constrained by elastic surroundings which follow a fully linear elastic solution) to duplication of the appropriate K values. A remarkably simple result and a startling achievement of continuum mechanics at a fundamental level, it does bring order out of chaos when followed in studies of crack propagation in beneficial, neutral, and adverse environments.

However, the very simplicity and generality of the requirement that the appropriate K be matched leaves completely open the details of the process of separation as the crack extends. In an energy balance calculation based upon experimental results for crack propagation, the term which appears as surface energy depends strongly upon the method of calculation of the work of the applied forces, the strain energy stored, and the

energy dissipated and radiated away as kinetic energy. Yet, an understanding of the effect of microstructural and atomic features and properties along with the influence of corrosive and other environments depends critically on the ability to compute the actual energy needed to create new surface separately from the energy required for deformation of the bulk of the material around the crack tip. Thermodynamic equilibrium surface energy clearly is a lower bound and an excellent approximation when sufficient thermal energy is available to overcome the activation energy barrier to forming new surface. In general, however, this will not be so as Robb Thomson has shown for a simplified picture of atomic separation at a crack tip. Energy lost through radiated kinetic energy or inelastic deformation in the immediate region of the surface of separation can be many times larger than the equilibrium surface energy. This is a genuine energy to create new surface and is not to be confused with the often far greater amount of energy dissipated in the plastic zone around the crack tip associated with the stable continuum behavior. As Jim Rice has emphasized earlier, the genuine energy to create surface governs. When it reduces to zero a sharp crack runs at zero applied load.

Instability of the local force-separation relations leads to the running crack. How the unstable portions of the force-separation relations alter with microstructure, atomic structure, and environment warrants intensive additional study

at both the atomic and microstructural level to permit reliable prediction in advance of tests.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAH15-71-C-0253 with The University of Michigan.

ON THE CHARACTERISTICS OF GRADIENT MATERIALS

M. B. Bever

Abstract

Materials with compositional and structural gradients are of theoretical and practical interest. Such materials have been widely used and probably will find new applications. This memorandum deals with the general characteristics of gradients in metallic, ceramic, polymeric, composite and complex materials. The mathematical analysis of gradients and the systematic consideration of the effects of gradients on properties will be the subjects of later memoranda.

ON THE CHARACTERISTICS OF GRADIENT MATERIALS

M. B. Bever

1. Introduction

Materials with compositional or structural gradients have been widely used in the past. They probably will find many new applications in the future. These materials, however, have not been systematically investigated and analyzed. This memorandum will deal with the general characteristics of gradient materials and will also briefly consider the effects of gradients on some properties. The mathematical analysis of gradients and the detailed consideration of their effects on properties will be the subjects of subsequent memoranda.

Gradient structures are found in metals and alloys, ceramics, polymers, composites and complex materials. Among the oldest examples of gradient materials are surface hardened steels^{1,2}. Steels can be surface hardened by various methods, but they all have transition zones or gradients between the hard but brittle surface layer and the softer but tough interior.

A recent application of the gradient principle is represented by graded cermets which combine a ceramic and a metal. Such cermets have been proposed for jet engine parts^{3,4} and are being considered for armor^{5,6}. Some other gradient materials are graded seals, which serve as a transition between glass and

metal or between two different glasses, and tempered glasses in which compressive stresses are present in the surface layer.

Polymeric materials may have many different kinds of gradients⁷. In one proposed application the values of the elastic moduli change continuously with location in the polymer; such a gradient should make it possible to combine a soft layer, which is compatible with natural tissue, and a hard layer, which can be anchored to a strong material⁸.

Many different types of gradient structures can occur in composite and other complex materials⁹. Gradient structures are also found in nature, especially in biological and geological materials.

The essential characteristic of a gradient material is the spatial variation of a compositional or structural feature. "Gradient materials" must be distinguished from other inhomogeneous materials. In a gradient material, a single inhomogeneity occurs on essentially a macroscopic or global scale. This criterion separates gradient materials from the products of spinodal decomposition, in which numerous concentration gradients are present on a fine scale. Another criterion of a gradient material is that the inhomogeneity must be gradual - a sharp discontinuity is an interface rather than a gradient. The criteria of scale and continuity are relative and deserve further consideration. In summary a gradient material may be defined as a material in which a compositional or structural feature changes continuously in space on a macroscopic scale.

A gradient can serve different functions. At one extreme, the gradient region provides a transition between two materials or components; the gradient is merely incidental to the combination and constitutes a necessary compromise. At the other extreme, the gradient region provides a desirable spatial variation in properties. Transitions between the two extremes are possible.

The features which may have gradients are not necessarily the same in metallic, ceramic, polymeric, composite and complex materials. Gradients characteristic of these major classes of materials will now be considered.

2. Gradients in Metals and Alloys

2.1 Single-Phase Metals and Alloys

Single-phase metallic systems may have gradients in their composition, microstructure, crystal orientation (texture) and internal stresses. Composition gradients may occur in solid solution alloys. For example, a surface layer may differ in composition from the bulk metal; such a material may be especially resistant to corrosion or to high-temperature oxidation. Nabarro has postulated a gradient of the concentration of interstitial solute atoms, which has theoretical interest¹⁰.

The grain size is the most important microstructural feature of a single-phase metal in which a gradient is possible. Such a gradient can be produced by differential mechanical or thermal treatments or by suitable alloying (including the formation of a minute amount of a second phase). A grain size

gradient results in gradients of properties, especially mechanical properties.

Gradients in grain shape and in grain shape anisotropy can be conceived; their effects probably would be subtle. Gradients in crystal orientation - both in the type and intensity or a preferred orientation or texture - are possible¹¹. They will be associated with gradients in the anisotropy of some local properties such as the elastic properties in cubic metals and also the thermal expansion and electrical conductivities in noncubic metals. Gradients in preferred orientation, however, would be difficult to produce¹².

Internal stresses are a special kind of structural feature. They may have macroscopic gradients. For example, compressive stresses can be introduced at the surface of single-phase metals and alloys by surface working processes; these stresses are necessarily balanced by subsurface tensile stresses, thus involving gradients.

2.2 Multiphase Metals and Alloys

The structure of a multiphase metal or alloy depends on the volume fractions of the constituent phases and the configuration of these phases, especially their microstructural and crystallographic arrangements. Three major types of microstructural arrangements can occur:

(i) one continuous and one (or more) dispersed phase(s), that is a matrix and dispersed phase; in such an arrangement either the major or the minor phase may be continuous;

(ii) two (or more) discontinuous phases arranged in a mixed grain structure or a lamellar structure;

(iii) two (or more) continuous and interpenetrating phases; in such an arrangement each phase has a high connectivity. Each of these types affords the possibility of structural gradients.

Matrix and Dispersed Phase - If the matrix is the major phase, it displays most of the characteristics of a typical single-phase polycrystalline metal. In principle it can, therefore, have the gradients in composition, microstructure, crystal orientation and internal stresses discussed in 2.1 above.

The characteristics of the dispersed phase (or phases) which may have gradients are the volume fraction and the size, shape and orientation of the particles of the dispersed phase. A gradient in the volume fraction can result from either a gradient in the initial composition of the parent phase or from a differential thermal treatment which produces a gradient in the nature of the precipitation reaction. An interesting possibility is the preferential stimulation of precipitation in a surface layer by surface working.

Gradients in the size of the dispersed particles can be produced by suitable gradients in composition or suitable thermal treatments such as a temperature differential causing differences in space of the rates of nucleation and growth of the precipitate. Since little work appears to have been done in exploring such possibilities, opportunities for development probably exist.

Gradients in particle shape, particle shape anisotropy and crystallographic orientation of dispersed particles suggest themselves, but would probably be difficult to produce.

Mixed-Grain Structures - Such structures occur in two-phase regions such as the ($\alpha + \beta$) region in the copper-zinc system. Gradients of the grain sizes of the two phases and perhaps also of their mutual arrangement are possible.

Lamellar Structures - Lamellar structures have been investigated extensively because of their role in pearlitic steels and some eutectic systems. The degree of fineness of the lamellae and the size of the pearlite colonies can probably be changed with position, but such gradients do not seem to have been investigated. Their investigation should be undertaken.

Continuous Interconnected Phases - Gradients in the composition and scale of the interconnected phases appear to be possible in these structures.

Examples from the Metallurgy of Steels - The quench-hardening of steel presents a classical example of a gradient structure. If a steel part is cooled at a rate not sufficiently fast to yield martensite throughout, pearlite forms in the interior and in an intermediate zone pearlite and martensite form as a mixture; this results in a gradient from an all-martensitic to an all-pearlitic structure. The phenomena are well understood and pertinent tests have been developed. Parts with martensite-pearlite gradients can be used in technical applications.

The metallurgy of steels suggests the possibility of

producing structural gradients by a unique mechanism: since the decomposition of austenite and the resulting structure are affected by the grain size of the austenite, the decomposition products will have gradients if the austenite has a grain size gradient. Experimental work on this subject seems to be desirable.

The properties of a quench-hardened (martensitic) steel are modified by tempering (holding at relatively low temperatures). Tempering in a temperature gradient may be expected to produce gradients in the structure and properties of the steel. The low temperatures involved would make the control of such a differential procedure relatively easy.

Various heat treatments and thermomechanical treatments of plain carbon and alloy steels such as austempering, martempering and ausforming provide opportunities for producing other gradient structures.

In the case hardening of steel by carburizing, the carbon is supplied to a surface layer. The steel, therefore, can form effective martensite only in this layer^{1,2}. The transition from the high-carbon case to the low-carbon core is relatively steep. Troiano and coworkers^{13,14} have shown that the resulting abrupt change in local properties causes a "metallurgical notch" at the interface and hence a susceptibility to fracture initiation. The obvious solution is a more gradual transition due to a less steep carbon gradient, but such a gradient appears to be difficult, if not impossible, to produce.

In case hardening of steel by nitriding, the hardness of the case results from the precipitation of alloy nitrides and decreases gradually at the bottom of the nitrided layer^{1,2}. While this gradual decrease has been deplored because it tends to reduce the case depth and hardness at the surface, it probably improves the mechanical behavior of nitrided steels by reducing the notch effect present in carburized steel. Induction hardened and flame hardened steels provide the possibility of producing deeper cases with more extended transition zones than nitrided steels^{1,2}. The characteristics and effects of the transition zone between case and core in carburized, nitrided, induction hardened and flame hardened steels merit further investigation.

Other Metallurgical Gradients - Surface treatments can be applied to metals other than steels. The protection against corrosion of aluminum alloys by a surface layer of pure aluminum is an example.

Porosity - which is usually undesirable in metals - will be discussed in connection with ceramics (see Section 3). The generation of compressive stresses by surface working was mentioned in connection with single-phase metals (see 2.1 above). This process is also applicable to multiphase metals and especially to steels.

3. Ceramic Materials

The compositional and structural features which characterize ceramic materials may have spatial gradients. Many of these

features are analogous to features which are also characteristic of metals and which were discussed in Section 2. Section 3 will deal with features of special interest for ceramic materials.

3.1 Single-Phase Ceramics

A major variable in polycrystalline single-phase ceramics is the grain size. This has important effects on mechanical and other properties of ceramics¹⁵.

A grain size gradient has recently been observed in aluminum oxide ceramics: the grain size averaged about 1.6 μm in the interior of a typical specimen and 0.8 μm at the surface¹⁶. The investigator attributed the grain size gradient to an impurity gradient. He believed that the grain size gradient contributed to the high strength properties observed in this material.

Porosity has been investigated extensively in ceramic materials¹⁵. The possibility of porosity gradients obviously exists. Gradients may occur in the total density (or volume) of the pores, the size and shape of the pores, and the degree of their connectivity. An example is the open porosity deliberately introduced at the surface of certain ceramic materials which are being considered for use in surgical implants^{17, 18}.

The chemical tempering by ion exchange provides an example of a gradient in a glass. In one of several possible processes ions at the surface are replaced by larger ions; this is carried out below the annealing temperature of the glass. After cooling to room temperature, the concentration gradient is associated with a gradient in the compressive stresses.

3.2 Multiphase Ceramics

The most common structure of multiphase ceramic materials is probably the dispersion of one or more phases in a matrix. The matrix may be a crystalline phase or a glass. The dispersed particles may have a variety of shapes and different orientation relations. Many possibilities exist and in principle each may yield a structural gradient.

A matrix containing dispersed particles is not the only multiphase structure occurring in ceramic systems. For example, spinodal decomposition may produce a structure of two interpenetrating phases. While the concentration gradients occurring on a microscale in a spinodal decomposition product do not qualify for a "gradient material" as discussed in the Introduction (see 1 above), a spinodal decomposition product may have a gradient on a macroscale due to a gradient in the composition or in the applied process variables.

The effects of various possible arrangements of the phases on the properties of multiphase ceramics have been analyzed for several schematic arrangements. Various mixture rules apply for such properties as the electrical and thermal conductivities. These rules can be adapted to gradient ceramic structures.

4. Polymeric Materials*

The structure and properties of most polymeric materials can be varied over wide ranges, but to date little effort has

*The writer acknowledges the contributions made by M. Shen to the contents of Section 4.

been devoted to developing gradient polymeric materials. Ferry⁶ has analyzed some features of gradient polymers for biomedical applications. Shen and Bever⁷, in a concurrent memorandum, consider systematically the structure and properties of polymeric systems susceptible to spatial gradients and discuss some potential applications. The present discussion is based on their paper, which should be consulted for details.

Gradients in the structure of polymeric systems may be generated by varying (i) the chemical nature of the monomers, (ii) the molecular constitution of the polymers and (iii) the supramolecular structure or morphology of the polymers. Gradients in each of these categories are possible in single-phase as well as heterophase systems.

4.1 Single-Phase Polymers

The nature of a polymer depends first of all on the constituent monomers. Various monomers A and B with different properties can be combined to form random copolymers ranging in composition from all A to all B. A gradient polymer with a spatial variation in composition can therefore be prepared. An example is the possible combination in a gradient structure of methyl methacrylate, which forms a tough, hard polymer, and methyl acrylate, a soft rubber. Polymer chemistry offers many possibilities for creating similar composition gradients by copolymerization.

Crosslinking is another characteristic feature of a polymer which can be varied. In principle, techniques available

for controlling crosslinking can be adapted to producing gradients.

Differences in the nature of the constituent monomers and differences in crosslinking can cause differences in equilibrium swelling. This opens the way to creating gradients in the swelling tendency.

The average molecular weight and the molecular weight distribution are other variables affecting single-phase polymers. Since they can be manipulated, corresponding gradients may be achieved. Gradients in the degree and the direction of the orientation of polymers are other possibilities.

4.2 Heterophase Polymers

The variables discussed in the foregoing for single-phase systems can also yield gradients in heterophase systems. Heterophase polymeric systems, however, offer additional possibilities for gradients.

In crystalline polymers the ratio of the crystalline phase to the amorphous phase can be varied. The degree of crystallinity can be altered by thermal treatments. The degree of crystallinity also depends on the degree of tacticity and the ratio of cis to trans isomers of a given polymer. Copolymerization with an appropriate comonomer is another method for changing the degree of crystallinity. These methods should make it possible to establish gradients in the degree of crystallinity. Gradients in the orientation and size of the spherulites appear also to be attainable.

An important class of heterophase polymers can be prepared

by mixing polymers. Different techniques yield "polyblends", block copolymers and graft polymers.

As in metals and alloys and ceramic materials, the morphology of heterophase polymer mixtures depends on the amount and configuration of the constituent phases. In particular, the matrix-dispersoid structures, the lamellar structures and the highly connected structures are to be considered. (See Sub-section 2.2 for details.) In polymers these configurations can be produced by varying the proportions of the components or the method of preparation. In this way gradients in morphology can also be established.

5. Composite Materials

The large number of compositional and structural features which characterize a composite material make a large number of gradients possible. Bever and Duwez⁹ have systematically considered these gradients and their effects on the properties of composites; their paper also discusses reported and potential applications of gradient composites.

The gradients in a composite may involve the matrix or the reinforcing phase. The same gradients may be present in the matrix of a composite as in the matrix material in the absence of a dispersed phase. For example, if the matrix is a metallic or ceramic material, gradients may occur in the size, shape or orientation of the grains. (See Sections 2 and 3). Polymeric matrices of composites also can have gradients (See

Section 4).

Gradients may also involve the dispersed phase in a composite material. We assume first that the dispersed phase consists of unidirectionally arranged filaments. In this case the most obvious possibility is a gradient of the concentration (density) of filaments. Other gradients can involve the orientation of the filaments and the length of discontinuous fibers.

The concentration gradient may lie in a plane normal to the filament direction. In a plate, such a gradient can lie in the plane of the plate or be normal to it. In a rod, the concentration of filaments may change in a radial direction. The concentration of discontinuous filaments can also change in the direction of the filaments.

If a composite contains more than one type of filament various concentration gradients are possible. These gradients may alter the ratio of the local concentrations of several types of filaments.

A dispersed phase in a composite may have shapes other than filaments of circular cross section, such as ribbons. In such a case additional possibilities of gradients involving the orientation of the ribbon (with respect to its width as well as to its length) are introduced. Similarly, gradients in a dispersed phase consisting of flakes or platelets may involve orientation in addition to such variables as concentration, size and shape of the particles. If several dispersed phases are

present, the number of possible combinations and corresponding gradients is further increased.

Details concerning gradients in composites and a discussion of their effects on the local and global properties of composites may be found in Reference 9.

Composite materials can be designed on a coarser scale than that of the composites considered in the foregoing. For example, various kinds of laminates have been used. They lend themselves to gradient construction.

6. Complex Materials

Materials belonging to different classes can be combined in one structure. As mentioned in the Introduction, a cermet combines a metal and a ceramic. The ratio of the components and other characteristics in a cermet can be varied and this introduces the possibility of gradients. The example of armor application has been mentioned.

A graded glass/metal seal is a type of material which solely depends on the gradient for its usefulness. The specific nature of this gradient depends on the thermal expansion coefficients of the glass and metal.

Polymers can be combined with nonpolymeric materials. Reinforcing fillers, which cause marked changes in the properties of polymers, are usually not polymers; an example is carbon black in natural rubber. Doping can change a polymer from an insulator to a semiconductor; the addition of iodine to polyethylene

causes such a change. The use of fillers and doping agents makes concentration gradients possible. Another potential development is the joining of a polymer to an inorganic material; the joint can be made if the inorganic material has a porosity gradient at the surface and the pores are impregnated with the polymer.

7. Conclusion

This memorandum has concentrated on the characteristics of gradients which can occur in the major classes of materials. Details on quantitative aspects of gradients and the effects of gradients on properties, particularly in polymeric and composite materials, may be found in References 7 and 9. The mathematical analysis of compositional and structural gradients and the systematic consideration of the effects of gradients on properties will be the subjects of later memoranda.

Acknowledgements

The writer first became interested in the subject of this memorandum while spending a sabbatical year at Harvard University as a guest of the Division of Engineering and Applied Physics. Later he collaborated with Professor P. E. Duwez on a related subject involving composites and with Professor Mitchel Shen in the area of polymers; these joint efforts have contributed to the present study. Support by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 is gratefully acknowledged.

References

1. M. B. Bever and C. F. Floe, Chapters on Case Hardening of Steel by Carburizing, Carbonitriding, Nitriding and Diffusion Treatments for Wear Protection, in: Surface Protection Against Wear and Corrosion, American Society of Metals, 1954, pp. 63-130.
2. M. B. Bever and C. F. Floe, Surface Hardening of Steel, in: McGraw-Hill Encyclopedia of Science and Technology, 3rd ed., Vol. 13, 1971, pp. 329-332.
3. C. G. Goetzel and J. B. Adamec, Metal Progress, 70, No. 6 (1956) 101.
4. C. G. Goetzel and H. W. Lavendel, Infiltrated Powder Components for Power Plant and Propulsion Systems, in: Plansee Proceedings, 1964, ed. F. Benesovsky, p. 149.
5. M. L. Wilkins, Lawrence Radiation Laboratory, Livermore, Cal., private communication.
6. J. J. Stiglich, Jr., New Materials and Processes for Ceramic Armor, DCIC Report 69-1, Part 1, 1969.
7. M. Shen and M. B. Bever, On Gradients in Polymeric Systems, Preliminary Reports, Memoranda and Technical Notes of the ARPA Materials Summer Conference, July, 1971; also J. Mater. Sci. (in press).
8. J. D. Ferry, Control of Mechanical Properties of Swollen Hydrophilic Network Polymers; Layer and Gradient Structures, Preliminary Reports, Memoranda and Technical Notes of the ARPA Materials Summer Conference, July, 1970, pp. 611-627.
9. M. B. Bever and P. E. Duwez, Gradients in Composite Materials, Preliminary Reports, Memoranda and Technical Notes of the ARPA Materials Summer Conference, July, 1970, pp. 117-140; also Mater. Sci. Eng. (in press).
10. F. R. N. Nabarro, Theory of Crystal Dislocations, Oxford University Press, 1967, p. 575.
11. G. Wassermann and J. Grewen, Texturen Metallischer Werkstoffe, 2nd ed., Springer, 1962, pp. 7, 13.
12. P. S. Mathur, Mechanical Contributions to the Plane Strain Deformation and Recrystallization Textures of Al-killed Steel, Sc.D. Thesis, Department of Metallurgy and Materials Science, Massachusetts Institute of Technology, 1972.

13. L. J. Ebert, F. T. Krotine and A. R. Troiano, A Behavioral Model for the Fracture of Surface Hardened Components, Trans. ASME, 1965, Paper No. 65-Met-2.
14. F. T. Krotine, M. F. McGuire, L. J. Ebert and A. R. Troiano, The Influence of Case Properties and Retained Austenite on the Behavior of Carburized Components, Trans. ASM, 62, 1969, 829,
15. W. D. Kingery, Introduction to Ceramics, John Wiley & Sons, 1960.
16. R. E. Mistler, private communication.
17. C. D. Talbert, A Basic Investigation into the Potential of Ceramic Materials as Permanently Implantable Skeletal Prostheses, Master's Thesis, Department of Ceramic Engineering, Clemson University, May, 1969.
18. S. F. Hulbert, F. A. Young, R. S. Mathews, J. J. Klawitter, C. D. Talbert and F. H. Stelling, Potential of Ceramic Materials as Permanently Implantable Skeletal Prostheses, College of Engineering, Clemson University. Paper presented at February 1970 National Meeting of American Institute of Chemical Engineers.

ANALYSIS OF STRESS INTENSITY FACTORS IN A PLATE
WITH ANY GIVEN DISTRIBUTION OF CRACKS

M. Ishida

Regular Member
Aeronautics and Astronautics Technology Laboratory
1880 Jindaiji-chō, Chōfu
Tokyo, Japan

April 2, 1968
Presented at the 45th Session of the Regular Symposium
Received April 15, 1969

Nihon Kikai Gakkai Rombunshu
(Trans. Japan Soc. Mech. Engineers),
35 (1969), 1815-1822

Translated by:

Mr. Coe Ishimoto
M.I.T.

for

ARPA Materials Research Council
Summer Study Group
Woods Hole, Massachusetts
July 1971

Preceding page blank

ANALYSIS OF STRESS INTENSITY FACTORS IN A PLATE
WITH ANY GIVEN DISTRIBUTION OF CRACKS

M. Ishida

1. Introduction

Since the introduction of the concept of the crack tip stress intensity factor, K , by Irwin¹, many analyses of the theoretical value of K in various new problems have been carried out parallel with applications to the problems of [[brittle]] fracture and fatigue. For the mutual interference among more than two cracks, exact solutions exist only for the case of cracks lined up on a straight line. For other cases, in general, elaborate analyses are necessary, but recently Yokobori et al.² have obtained important results for parallel arrays of cracks.

This paper, in further generality, describes the method of solution by iteration for a plate with an array of cracks of arbitrary number, positions, lengths, and directions subjected at a long distance to arbitrary biaxial tensile and shear stress and bending moment. The method is applied to typical problems, providing new data on K . The range of validity of the method and the accuracy of the results are carefully examined.

2. Theory

2.1 Stress function. We consider an infinite plate with a distribution of an arbitrary array of an arbitrary number (N) of linear cracks subject to stress at a long distance. We fix a reference coordinate system (X_0, Y_0) on the plate, and let each crack have center O_j , length $2L_j$, angle with respect to the X_0 -axis α_j and polar coordinates from O_j in the reference coordinate system (R_j, β_j) ($j = 1, 2, \dots, N$). For use later in the analysis, we establish N rectangular coordinate systems (X_j, Y_j) with origin at the center of each crack and with the X-axis in the direction of the crack.

We further set a standard length d and define dimensionless rectangular coordinate systems (x_j, y_j) , complex coordinates z_j and parameters l_j and r_j as in the following equations:

$$\left. \begin{aligned} x_0 &= \frac{X_0}{d} \quad , \quad y_0 = \frac{Y_0}{d} \quad , \quad z_0 = x_0 + iy_0 = \frac{X_0 + iY_0}{d} \\ x_j &= \frac{X_j}{d} \quad , \quad y_j = \frac{Y_j}{d} \quad , \quad z_j = x_j + iy_j = \frac{X_j + iY_j}{d} \\ l_j &= \frac{L_j}{d} \quad , \quad r_j = \frac{R_j}{d} \end{aligned} \right\} \quad (1)$$

$(j = 1, 2, \dots, N)$

Then the relative dimensionless position vectors $O_j O_k$ of the crack centers are given by

$$r_{jk} e^{i\beta_{jk}} = r_k e^{i\beta_k} - r_j e^{i\beta_j} \quad (2)$$

Assume that the plate is infinite in extent, that there are no external forces on the boundary of any of the cracks, and that a stress of the following form consisting of a combination of biaxial tension, shear stress and bending moment is applied at a long distance.

$$\left. \begin{aligned} \sigma_x^\omega &= \sigma \left(\alpha + \frac{\kappa Y_0}{d} \right) \\ \sigma_y^\omega &= \sigma \left(\beta + \frac{\mu X_0}{d} \right) \\ \tau_{xy}^\omega &= \sigma \gamma \end{aligned} \right\} \quad (3)$$

Here, σ is some standard stress and α , β , γ , κ and μ are constants (Fig. 1).

For our analysis, we first assume the following form for Airy's stress function in this problem.

$$\chi = \chi_0 + \sum_{k=1}^N \chi_k \quad (4)$$

χ_0 corresponds to the stress condition at a long distance (Eq. 3), and is expressed in the following form using complex potentials $\phi_0(z_0)$ and $\psi_0(z_0)$:

$$\left. \begin{aligned} \chi_0 &= \sigma d^2 \operatorname{Re} \{ \bar{z}_0 \phi_0(z_0) + \psi_0(z_0) \} \\ \phi_0(z_0) &= \frac{1}{4} (\beta + \alpha) z_0 + \frac{1}{8} (\mu - i\kappa) z_0^2 \\ \psi_0(z_0) &= \frac{1}{4} (\beta - \alpha + 2i\gamma) z_0^2 + \frac{1}{24} (\mu + i\kappa) z_0^3 \end{aligned} \right\} \quad (5)$$

Next, consider all the cracks as limiting cases of elliptical holes of large eccentricity. Then, in general χ_k is a stress function which has singularities only inside the crack k , and may be expanded as follows:

$$\left. \begin{aligned} \chi_k &= \sigma d^2 R_e \{ \bar{z}_k \phi_k(z_k) + \psi_k(z_k) \} \\ \phi_k(z_k) &= \sum_{n=0}^{\infty} (F_{n,k} \cdot + i F_{n,k}') z_k^{-(n+1)} \\ \psi_k(z_k) &= -D_{0,k} \cdot \log z_k + \sum_{n=1}^{\infty} (D_{n,k} \cdot + i D_{n,k}') z_k^{-n} \end{aligned} \right\} \quad (6)$$

Here, coefficients bearing the signs " \cdot " and " $'$ " express the real parts and the imaginary parts respectively, of the corresponding unknown complex coefficients.

2.2 Coordinate transformations. In determining these unknown coefficients from the boundary conditions, we only need to consider the conditions such that all the edges of the cracks are free, since the boundary conditions (Eq. 3) at infinite distance are satisfied by χ_0 alone and the χ_k 's ($k = 1, 2, \dots, N$) are of the form such that the stress vanishes at a long distance.

Since each element χ_k of the stress function is expressed in a different coordinate system, in considering the boundary conditions for a given crack j , it is convenient to express everything in the coordinate system $z_j = (x_j + iy_j)$ whose origin is at the center O_j of this crack. For χ_0 , using the relation

between z_0 and z_j

$$z_0 = z_j e^{i\alpha_j} + r_j e^{i\beta_j} \quad (7)$$

In Eq. 5, we obtain the following form:

$$\left. \begin{aligned} \chi_0 &= \sigma d^2 R_e \{ \bar{z}_j \phi_0^*(z_j) + \psi_0^*(z_j) \} \\ \phi_0^*(z_j) &= \frac{1}{4} \{ \beta + \alpha + r_j e^{i\beta_j} (\mu - i\kappa) \} z_j + \frac{e^{i\alpha_j}}{8} (\mu - i\kappa) z_j^2 \\ \psi_0^*(z_j) &= \frac{e^{2i\alpha_j}}{4} \{ \beta - \alpha + 2i\gamma + r_j (\mu \cos \beta_j - \kappa \sin \beta_j) \} z_j^2 \\ &\quad + \frac{e^{3i\alpha_j}}{24} (\mu + i\kappa) z_j^3 \end{aligned} \right\} \quad (8)$$

Also, for χ_k ($k \geq 1$), we put

$$z_k = z_j e^{i(\alpha_j - \alpha_k)} - r_{jk} e^{i(\beta_{jk} - \alpha_k)} \quad (9)$$

into the first of Eq. 6 and obtain an expression in terms of z_j

$$\left. \begin{aligned} \chi_k &= \sigma d^2 R_e \{ \bar{z}_j \phi_k^*(z_j) + \psi_k^*(z_j) \} \\ \phi_k^*(z_j) &= e^{i(\alpha_k - \alpha_j)} \phi_k [z_j e^{i(\alpha_j - \alpha_k)} - r_{jk} e^{i(\beta_{jk} - \alpha_k)}] \\ \psi_k^*(z_j) &= \psi_k [z_j e^{i(\alpha_j - \alpha_k)} - r_{jk} e^{i(\beta_{jk} - \alpha_k)}] \\ &\quad - r_{jk} e^{i(\alpha_j - \beta_{jk})} \phi_k(z_j) \end{aligned} \right\} \quad (10)$$

and then substitute the second and third of Eq. 6 into Eq. 10, and again expand χ_k in a power series of z_j . Using the result together with Eq. 8 in Eq. 4, the entire stress function may be rewritten in terms of the z_j 's as follows:

$$\chi = \sigma d^2 R_e \{ \bar{z}_j \phi_j(z_j) + \psi_j(z_j) \}$$

$$\phi_j(z_j) = \sum_{n=0}^{\infty} \{ (F_{n,j} \cdot + iF_{n,j}') z_j^{-(n+1)} + (M_{n,j} \cdot + iM_{n,j}') z_j^{n+1} \}$$

$$\psi_j(z_j) = -D_{0,j} \cdot \log z_j + \sum_{n=1}^{\infty} (D_{n,j} \cdot + iD_{n,j}') z_j^{-n}$$

$$+ \sum_{n=0}^{\infty} (K_{n,j} \cdot + iK_{n,j}') z_j^{n+2}$$
} (11)

$$M_{n,j} \cdot = \frac{\Delta_0 n}{4} (\beta + \alpha + r_j (\mu \cos \beta_j + \kappa \sin \beta_j)) + \frac{\Delta_1 n}{8} (\mu \cos \alpha_j + \kappa \sin \alpha_j)$$

$$+ \sum_{p=0}^{\infty} \sum_{k \neq j}^N (e_{n,j}^{p,k} F_{p,k} \cdot + f_{n,j}^{p,k} F_{p,k}'))$$

$$M_{n,j}' = \frac{\Delta_0 n}{4} r_j (\mu \sin \beta_j - \kappa \cos \beta_j) + \frac{\Delta_1 n}{8} (\mu \sin \alpha_j - \kappa \cos \alpha_j)$$

$$+ \sum_{p=0}^{\infty} \sum_{k \neq j}^N (-f_{n,j}^{p,k} F_{p,k} \cdot + e_{n,j}^{p,k} F_{p,k}'))$$

$$K_{n,j} \cdot = \frac{\Delta_0 n}{4} ((\beta - \alpha) \cos 2\alpha_j - 2\gamma \sin 2\alpha_j + r_j \cos 2\alpha_j (\mu \cos \beta_j - \kappa \sin \beta_j))$$

$$+ \frac{\Delta_1 n}{24} (\mu \cos 3\alpha_j - \kappa \sin 3\alpha_j) + \sum_{p=0}^{\infty} \sum_{k \neq j}^N (a_{n,j}^{p,k} D_{p,k} \cdot$$

$$+ b_{n,j}^{p,k} D_{p,k}' + c_{n,j}^{p,k} F_{p,k} \cdot + d_{n,j}^{p,k} F_{p,k}'))$$
} (12)

$$\begin{aligned}
K_{n,j}' = & \frac{\Delta_0^n}{4} ((\beta - \alpha) \sin 2\alpha_j + 2\gamma \cos 2\alpha_j + r_j \sin 2\alpha_j (\mu \cos \beta_j - \kappa \sin \beta_j)) \\
& + \frac{\Delta_1^n}{24} (\nu \sin 3\alpha_j + \kappa \cos 3\alpha_j) + \sum_{p=0}^{\infty} \sum_{k \neq j}^N (-b_{n,j}^{p,k} D_{p,k} \\
& + a_{n,j}^{p,k} U_{p,k} - d_{n,j}^{p,k} F_{p,k} + c_{n,j}^{p,k} F_{p,k}')
\end{aligned}$$

where

$$a_{n,j}^{0,k} = \frac{\cos((n+2)(\beta_{jk} - \alpha_j))}{(n+2)(r_{jk})^{n+2}}, \quad b_{n,j}^{0,k} = \frac{\sin((n+2)(\beta_{jk} - \alpha_j))}{(n+2)(r_{jk})^{n+2}}$$

$$\left\{ \begin{array}{l} a_{n,j}^{p,k} \\ b_{n,j}^{p,k} \end{array} \right\} = (-1)^p \binom{n+p+1}{n+2} \frac{\left\{ \begin{array}{l} \cos \\ \sin \end{array} \right\} (p(\alpha_j - \alpha_k) + (n+p+2)(\beta_{jk} - \alpha_j))}{(r_{jk})^{n+p+2}} \quad (p \geq 1)$$

$$\left\{ \begin{array}{l} c_{n,j}^{p,k} \\ d_{n,j}^{p,k} \end{array} \right\} = (-1)^p \binom{n+p+2}{n+2} \frac{\left\{ \begin{array}{l} \cos \\ \sin \end{array} \right\} ((p+2)(\alpha_j - \alpha_k) + (n+p+4)(\beta_{jk} - \alpha_j))}{(r_{jk})^{n+p+2}}$$

$$\left\{ \begin{array}{l} e_{n,j}^{p,k} \\ f_{n,j}^{p,k} \end{array} \right\} = (-1)^{p+1} \binom{n+p+1}{n+1} \frac{\left\{ \begin{array}{l} \cos \\ \sin \end{array} \right\} ((p+2)(\alpha_j - \alpha_k) + (n+p+2)(\beta_{jk} - \alpha_j))}{(r_{jk})^{n+p+2}}$$

(13)

and Δ_0^n, Δ_1^n are Kronecker deltas.

2.3 Boundary conditions on the crack boundaries. We have now an expression for the stress function in terms of only the z_j 's. We next consider the conditions which make the boundary of crack j free. We have already expressed the stress function

of an infinite plate with elliptical holes in the form of Eq. 11, and we have given the relationship which must hold among the coefficients of the stress function if the edges of the holes are to be free. We need to find the limiting expressions as the minor axis approaches zero. By appropriately rewriting some variables and constants in this result, the condition that the stress function Eq. 11 makes the boundary of crack j free is found as the following relations which give the $\log z_j$ and the coefficients of the odd-powered terms in z_j as linear functions of coefficients of the even-powered terms in z_j :

$$\left. \begin{aligned}
 D_{n,j}' &= \sum_{p=0}^{\infty} 1_j^{n+p+2} a_p^n (k_{p,j}' + m_{p,j}'), \\
 F_{n,j}' &= - \sum_{p=0}^{\infty} 1_j^{n+p+2} c_p^n (k_{p,j}' + m_{p,j}'), \\
 D_{n,j}' &= - \sum_{p=0}^{\infty} 1_j^{n+p+2} (e_p^n k_{p,j}' + f_p^n m_{p,j}'), \\
 F_{n,j}' &= \sum_{p=0}^{\infty} 1_j^{n+p+2} (c_p^n k_{p,j}' + g_p^n m_{p,j}')
 \end{aligned} \right\} \quad (14)$$

($n=0, 1, 2, \dots$)

where

$$\begin{aligned}
a_{2p}^0 &= \frac{p+1}{2^{2p}} \binom{2p+1}{p} \\
a_{2p}^{2n} &= \frac{1}{2^{2n+2p+2n}} \left\{ 4(p+1) \binom{2p+1}{p} \binom{2n-1}{n} + 2p \sum_{m=0}^{p+1, n} \binom{2p+2}{p-m+1} \binom{2n}{n-m} \right\} \quad (n \geq 1) \\
c_{2p}^{2n} &= \frac{p+1}{2^{2n+2p+1} (2n+1)} \sum_{m=0}^p (2m+1) \binom{2p+1}{p-m} \binom{2n+1}{n-m} \\
e_{2p}^{2n} &= \frac{p+1}{2^{2n+2p+1} n} \sum_{m=0}^{p+1, n} \binom{2p+2}{p-m+1} \binom{2n}{n-m} \\
f_{2p}^{2n} &= \frac{p}{p+1} c_{2p}^{2n} \\
g_{2p}^{2n} &= \frac{1}{2^{2n+2p+2} (n+1)} \left\{ (p+1) \binom{2p+1}{p} \binom{2n}{n} - \frac{n+1}{2n+1} \sum_{m=0}^p (2m+1) \binom{2p+1}{p-m} \binom{2n+1}{n-m} \right. \\
&\quad \left. + (2p+1) \sum_{m=0}^{p, n+1} \binom{2p}{p-m} \binom{2n+2}{n-m+1} \right\} \\
a_{2p+1}^{2n+1} &= \frac{1}{2^{2n+2p+4}} \left\{ \frac{p+2}{n+1} \binom{2p+3}{p+1} \binom{2n}{n} + \frac{2p+1}{2n+1} \sum_{m=0}^{p+1, n} \binom{2p+3}{p-m+1} \binom{2n+1}{n-m} \right\} \\
c_{2p+1}^{2n+1} &= \frac{2p+3}{2^{2n+2p+4} (n+1)} \sum_{m=0}^p (m+1) \binom{2p+2}{p-m} \binom{2n+2}{n-m} \\
e_{2p+1}^{2n+1} &= \frac{1}{2^{2n+2p+4}} \left\{ -\frac{p+2}{n+1} \binom{2p+3}{p+1} \binom{2n}{n} + \frac{2p+3}{2n+1} \sum_{m=0}^{p+1, n} \binom{2p+3}{p-m+1} \binom{2n+1}{n-m} \right\} \\
f_{2p+1}^{2n+1} &= \frac{2p+1}{2p+3} c_{2p+1}^{2n+1} \\
g_{2p+1}^{2n+1} &= \frac{1}{2^{2n+2p+4}} \left\{ -\frac{1}{n+1} \sum_{m=0}^p (m+1) \binom{2p+2}{p-m} \binom{2n+2}{n-m} \right. \\
&\quad \left. + \frac{2p+2}{2n+3} \sum_{m=0}^{p, n+1} (2m+1) \binom{2p+1}{p-m} \binom{2n+3}{n-m+1} \right\}
\end{aligned} \tag{15}$$

The other coefficients (such as a_p^n and c_p^n when $p + n$ is odd) all vanish. In these equations, the upper limits of the summations are taken as the smaller of the two given.

We have considered crack j so far, but in our problem all the cracks must have free boundaries, so we have all the relationships for $j = 1, 2, \dots, N$ in Eq. 14.

2.4 Determination of unknown coefficients. Summarizing the foregoing, the stress function is given in terms of the z_j 's by Eq. 11, in which appear the unknown coefficients $D_{n,j}''$, $D_{n,j}'$, $F_{n,j}''$ and $F_{n,j}'$ which are determined by substituting Eq. 12 into Eq. 14. However, it is more convenient to consider $D_{n,j}''$, $D_{n,j}'$, $F_{n,j}''$, $F_{n,j}'$, $K_{n,j}''$, $K_{n,j}'$, $M_{n,j}''$, and $M_{n,j}'$ ($n = 0, 1, 2, \dots; j = 1, 2, \dots, N$) as the unknowns and to solve the conditions of Eqs. 12 and 14 simultaneously.

Equation 12 is completely determined by the relative positions of the cracks from the constants $\alpha, \beta, \gamma, \kappa$ and μ given in the long-distance stress condition Eq. 3 and from Eq. 13. Equation 14 is a power series in l_j ($j = 1, 2, \dots, N$), so if we write $l_j = \epsilon_j \lambda$ where the constants ϵ_j give the ratio of lengths of the cracks, they can be expressed as power series of λ alone. The positive-valued λ may be chosen arbitrarily, but here we use for our calculations

$$\lambda = \max \left[\frac{l_j + l_k}{r_{jk}} \right] \quad (16)$$

for a reason which will be explained later.

Next, in solving the simultaneous Eqs. 12 and 14, we use the iterative method, as in the author's previous paper⁵. That is, we assume all the unknown coefficients to be a power series in λ , substitute them into Eqs. 12 and 14, and set the terms of the same powers in the two expressions equal to each other, giving a set of relationships from which the coefficients of the assumed series may be automatically found progressively from the low-power terms to the high-power terms.

2.5 The stress intensity factor. In general, if we set the rectangular coordinates with the origin at the crack tip as (x,y) and the related polar coordinates as (r,θ) as in Eq. 2, the stress components in the neighborhood of the crack tip in a plane can always be given in the following form¹:

$$\left. \begin{aligned}
 \sigma_x &= \frac{K_1}{\sqrt{2r}} \cos \frac{\theta}{2} \left(1 - \sin \frac{\theta}{2} \sin \frac{3\theta}{2} \right) \\
 &\quad - \frac{K_2}{\sqrt{2r}} \sin \frac{\theta}{2} \left(2 + \cos \frac{\theta}{2} \cos \frac{3\theta}{2} \right) \\
 \sigma_y &= \frac{K_1}{\sqrt{2r}} \cos \frac{\theta}{2} \left(1 + \sin \frac{\theta}{2} \sin \frac{3\theta}{2} \right) \\
 &\quad + \frac{K_2}{\sqrt{2r}} \sin \frac{\theta}{2} \cos \frac{\theta}{2} \cos \frac{3\theta}{2} \\
 \tau_{xy} &= \frac{K_1}{\sqrt{2r}} \sin \frac{\theta}{2} \cos \frac{\theta}{2} \cos \frac{3\theta}{2} \\
 &\quad + \frac{K_2}{\sqrt{2r}} \cos \frac{\theta}{2} \left(1 - \sin \frac{\theta}{2} \sin \frac{3\theta}{2} \right)
 \end{aligned} \right\} \quad (17)$$

K_1 and K_2^* are related to the shape, boundary shape and load conditions of the crack, and represent the intensity of the stress field. In this sense, in a planar problem, they are called the crack tip stress intensity factors and are used extensively in the analysis of problems of brittle fracture and fatigue. K_1 and K_2 are respectively called the symmetrical and the antisymmetrical components of the stress intensity factor.

In the present problem, the stress function is given in the form of Eq. 11 in terms of the z_j 's. Here, the stress intensity factor at the tips A_j and B_j of crack j is computed from the following equation:

$$K_1^{(A_j, B_j)} - iK_2^{(A_j, B_j)} = 2\sqrt{2} \sigma \sqrt{d} \lim_{\lambda \rightarrow 0} [\sqrt{z_j + \epsilon_j \lambda} \phi_j'(z_j)] \quad (18)$$

After interchanging the order of summation in Eq. 18 in which the expansions in λ of $F_{n,j}^*$, $F_{n,j}'$, $M_{n,j}^*$, $M_{n,j}'$ obtained by the iterative method described earlier are substituted into closed form into $\phi_j(z_j)$ of Eq. 11, we sum up by a method similar to that of the author's previous paper⁵ the first series and then pass to the limit. The resulting crack tip stress intensity factors are given as power series in λ as follows:

Editor's note: In the USA, K_1 and K_2 are more commonly k_1 and k_2 or $K_{II}/\sqrt{\pi}$ and $K_{III}/\sqrt{\pi}$.

$$\begin{aligned}
K_1^{(\Lambda_j)} &= \sigma \sqrt{L_j} f_1^{(\Lambda_j)}(\lambda), & f_1^{(\Lambda_j)} &= \sum_{n=0}^M P_n \lambda^n \\
K_1^{(B_j)} &= \sigma \sqrt{L_j} f_1^{(B_j)}(\lambda), & f_1^{(B_j)} &= \sum_{n=0}^M P_n (-\lambda)^n \\
& & &= f_1^{(\Lambda_j)}(-\lambda) \\
K_2^{(\Lambda_j)} &= \sigma \sqrt{L_j} f_2^{(\Lambda_j)}(\lambda), & f_2^{(\Lambda_j)}(\lambda) &= \sum_{n=0}^M Q_n \lambda^n \\
K_2^{(B_j)} &= \sigma \sqrt{L_j} f_2^{(B_j)}(\lambda), \\
f_2^{(B_j)}(\lambda) &= \sum_{n=0}^M Q_n (-\lambda)^n = f_2^{(\Lambda_j)}(-\lambda)
\end{aligned} \tag{19}$$

3. Numerical Computations

The large-scale computer HITAC 5020K at the Aeronautics and Astronautics Technology Laboratory was used for the numerical computations. A Fortran program is available which prints the values of the coefficients $f_1^{(\Lambda_j)}$, $f_1^{(B_j)}$, $f_2^{(\Lambda_j)}$, $f_2^{(B_j)}$ of each crack tip intensity factor directly from the stress parameters α , β , γ , κ , ν , the polar coordinates (r_j, θ_j) of the centers of the cracks, the directions of the cracks α_j , and the constants c_j ($j = 1, 2, \dots, N$) expressing the ratios of the lengths of the cracks.

There are no theoretical limits to the number of cracks N and the order M of the series of Eq. 19. However, using the present program which avoids the use of magnetic tape in order

to save computation time, the maximum value of M diminishes as the number of cracks N increases because of the limitation of the core capacity. The relationship between N and M is shown in Table 1.

Results of the numerical computation for typical cases are shown in Figures 3-10. Figures 3 and 4 are the horizontal and vertical tensions of a plate with two cracks of equal length at an angle to each other. In general, both K_1 and K_2 are finite but we show only for the fracture-mechanically important K_1 its dependence on the length and indentation of the cracks. The solid and broken lines correspond respectively to the inner and outer crack tips A and B . $f_{1,B}$ does not vary much with the length of the crack, but $f_{1,A}$, which corresponds to the inner tips, increases sharply under [[vertical]] tension as the crack lengthens whereas it decreases under [[horizontal]] tension and shows the stress-moderating effect due to alignment of the cracks.

Figures 5 and 6 show the results for a plate with two parallel cracks of equal length subjected to both tensile and shear stress. $f_{1,A}$ and $f_{2,A}$ refer to the inner tips of the cracks, and show a rather complex behavior with respect to variations in c/d and λ . In both cases the value does not increase significantly even if λ increases if c/d is small, but if c/d is large the values increase sharply with increase in λ , converging to the result for cracks lined up on a straight

line (curve for $c/d \rightarrow \infty$).

Figure 7 shows $f_{1,A}$ and $f_{2,A}$ of the inner crack tips for a plate with two cracks lined up on a straight line subjected to both tension and shear force. For this case, there are exact solutions⁷ which involve elliptic integrals, but since numerical solutions are not available we have shown the results in the figure. The solid line for $\alpha = 0^\circ$ in Fig. 3, $c/d \rightarrow \infty$ in Figs. 5 and 6 and $L_1/L_2 = 1$ in Fig. 7 are the same curve, and $\alpha = 90^\circ$ in Fig. 4 and $c/d = 0$ in Fig. 5 are the same case, and independent computations of these results, of course, are in agreement.

Figure 8 shows $f_{1,A}$ and $f_{2,A}$ of the middle crack for a plate with an odd number of cracks of equal length lined up on a straight line subjected to both tensile and shear stress. The limiting expression as N becomes infinitely large is known in closed form as

$$f_{1,A} = f_{2,A} = \sqrt{\frac{2}{\pi\lambda} \tan \frac{\pi\lambda}{2}} \quad (20)$$

and is included in the figure. One can well see the convergence of the curves as N increases.

Figure 9 is for a plate with an odd number of evenly-spaced cracks of equal length. f_1 for tension is shown in the bottom half of the figure. The solid lines are for the outermost cracks for which K_1 is greatest and the broken lines are for the middle crack for which K_1 is least. All of these values

decrease as the number of cracks N or λ increases. On the other hand, f_2 for the shear force is greatest for the middle crack, so it is shown in the top half of the figure and it increases as the number of cracks N or λ increases. As N approaches infinity, all groups of curves should converge to a limiting curve.

Figure 10 is for an example of the most general case of crack distribution and stress conditions. f_1 is indicated above the crack tip and f_2 below the crack tip.

Finally, we discuss the range of validity and the accuracy of this method. In general, Eq. 19 obtained by this method is nothing but the Maclaurin expansions of the exact solutions, so the error arises only from the omission of higher-order terms. For example, for a plate with two cracks of equal length subject to both tensile and shear stress, corresponding to $L_1/L_2 = 1$ in Figure 7, the exact solution⁷

$$f_{1,A} = f_{2,A} = \frac{(1+\lambda)^2 \frac{E(k)}{K(k)} - (1-\lambda)^2}{2\lambda\sqrt{1-\lambda}}$$

$$f_{1,B} = f_{2,B} = \frac{(1+\lambda)^{3/2} \left\{ 1 - \frac{E(k)}{K(k)} \right\}}{2\lambda} \quad (21)$$

$$k = \frac{2\sqrt{\lambda}}{1+\lambda}$$

is known. Here, $K(k)$ and $E(k)$ are respectively elliptic integrals of the first and second kinds. If Eq. 21 is expanded in λ , they

indeed agree with the author's calculated Eq. 19. The numerical values are as compared in Table 2, and the results of the present method agree with the exact solutions to more than 7 significant figures when $\lambda \leq 0.8$, to 4 significant figures when $\lambda = 0.9$ and to within only 1.7% even when $\lambda = 0.96$.

A mathematically rigorous examination of the convergence of Eq. 19 and of the errors for the general case has not been done. However, it is known from all the problems of stress concentration analyzed by the author by the iterative method that there is no great error in estimating the convergence and error from the terms in the series actually computed⁵. Also, we recognize the general decrease in absolute value of the coefficients of the series with increasing order of the terms for all cases analyzed if we take λ as in Eq. 16, in the present problem. Considering these two facts, we may consider the present general method of solution to be valid and accurate as long as λ is not very close to 1, that is, as long as the circles whose diameters are the cracks do not overlap and are not very close to one another.

Now, in the actual computation it is convenient to compare partial sums of the series in order to estimate the accuracies of the results. Table 3 gives the partial sums of the series $f_{1,A}$ for $c = 5d$ in Figure 5. Lines in the table indicate that thereafter there are no changes in the last digit. The number of accurate places for each λ may be estimated from

the behavior of convergence of the partial sums. Again, we see that the present method has practical validity in the range $\lambda \leq 0.94$. In the computation program the partial sums of the series Eq. 19 may also be found if necessary, and the curves of Figures 3 through 9 represent analyses of several typical cases with the accuracy just described.

4. Conclusion

A general method of solution by iteration has been proposed for the crack tip stress intensity factors K_1 and K_2 for an infinite plate with an arbitrary array of a number of linear cracks, distributed in arbitrary positions, with arbitrary lengths and orientations, subject to arbitrary biaxial tension, shear stress and bending moment at a long distance. A Fortran computer program was made which directly gives K_1 and K_2 , given mechanical and geometrical constants, and numerical results were obtained for typical cases. It may be considered that this method of solution is valid as long as all the circles having the cracks as diameters do not come very close to each other.

However, the cases in which the cracks are close to each other are also fracture-mechanically important, and a general solution for those cases are also desirable. The author wishes to report his attempt when it is completed. In closing, the author thanks Mr. Amakusa of the Aeronautics and Astronautics Technology Laboratory, who has collaborated in the numerical computation and the production of the figures.

Table 1

Relationship between the number of cracks N
and the order M of the series (19)

N	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
M	57	57	57	55	47	41	35	31	29	25	23	21	21	19	17	17

Table 2

Tension and shear force on a plate with two cracks
of equal length lined up on a straight line

λ	f_A		f_B	
	(19)	(21)	(19)	(21)
0.80	1.228 935	1.228 935	1.081 067	1.081 067
0.85	1.312 429	1.312 439	1.096 714	1.096 716
0.90	1.453 53	1.453 87	1.117 38	1.117 41
0.92	1.543 1	1.544 5	1.128 0	1.128 1
0.94	1.669 6	1.676 2	1.140 7	1.141 1
0.96	1.861	1.893	1.157	1.158

Table 3

Example ($c = 5d$ in Figure 5) of partial sums
of the series $f_{1,\lambda}$

$\lambda \backslash M$	9	19	29	39	49	54	57
0.40	1.032 3	---	---	---	---	---	---
0.50	1.058 1	1.058 2	---	---	---	---	---
0.60	1.099 2	1.099 9	---	---	---	---	---
0.65	1.128 3	1.130 1	---	---	---	---	---
0.70	1.165 6	1.169 5	1.169 2	---	---	---	---
0.75	1.213 4	1.221 2	1.220 2	---	---	---	---
0.80	1.275 1	1.289 8	1.285 5	1.285 6	---	---	---
0.85	1.354 8	1.380 4	1.363 2	1.364 3	1.364 8	1.364 7	---
0.90	1.458 0	1.498 4	1.433 4	1.442 5	1.448 4	1.446 9	1.446 0
0.92	1.507 5	1.553 9	1.445 0	1.465 7	1.480 5	1.475 8	1.473 0
0.94	1.562 4	1.614 0	1.433 1	1.479 0	1.515 7	1.501 1	1.491 6
0.96	1.623 2	1.678 1	1.379 8	1.480 0	1.568 9	1.525 1	1.494 3

BIBLIOGRAPHY

1. G. R. Irwin, Handbuch der Physik, 6 (1958), 551, Springer-Verlag.
2. T. Yokobori et al., Rep. Res. Inst. Strength and Fracture of Materials, Tohoku Univ., Sendai, 1-1 (1965); 1-2 (1965); 3-1 (1967).
3. Ishida, Nihon Kikai Gakkai Rombunshū, 21-107 (July 1955), 502.
4. Hayashi, Nihon Kikai Gakkai Rombunshū, 25-159 (November 1959), 1133, paper (8).
5. Ishida, Nihon Kikai Gakkai Rombunshū, 21-107 (July 1957), 507; 22-123, (1958), 809; 23-131 (1959), 474; Records of the Aeronautical Society, 10-100 (May 1962), 141; Trans. A.S.M.E. Ser. E 33-3 (September 1966), 674.
6. H. M. Westergaard, J. Appl. Mech. 6 (1939), A-49.
7. F. Erdogan, Proc. 4th U.S. Nat. Congr. Appl. Mech., (1962), 547, Pergamon Press. (SIC)

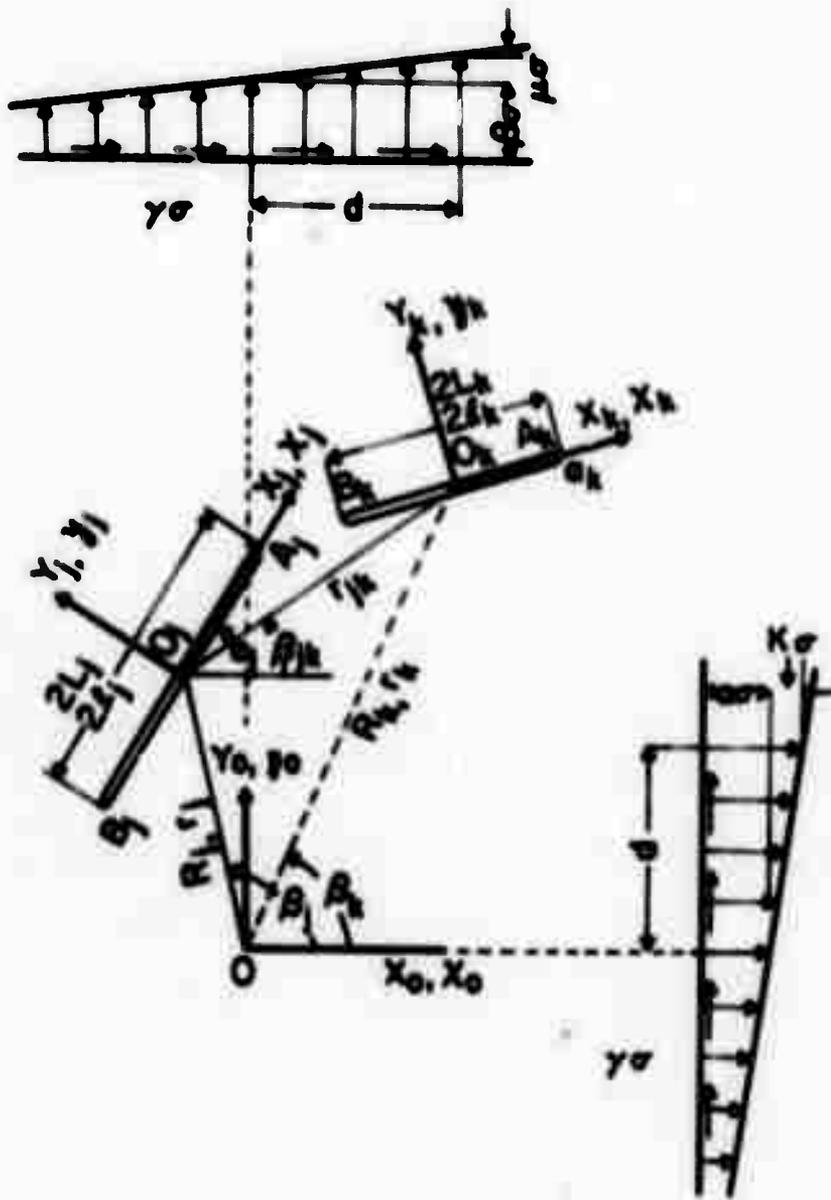


Figure 1.

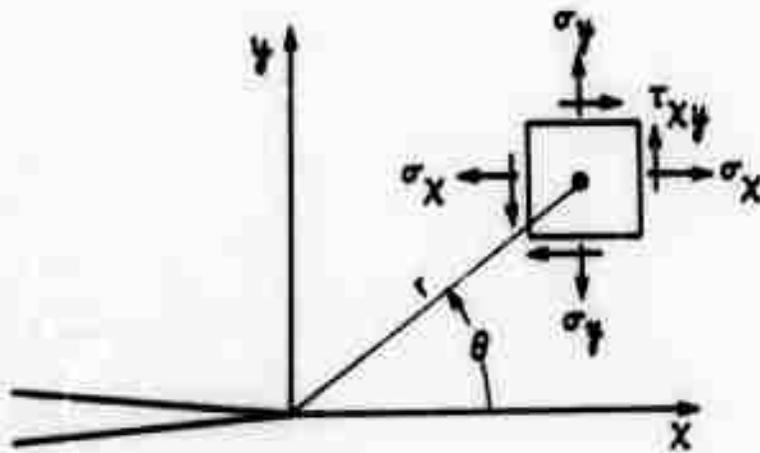


Figure 2.

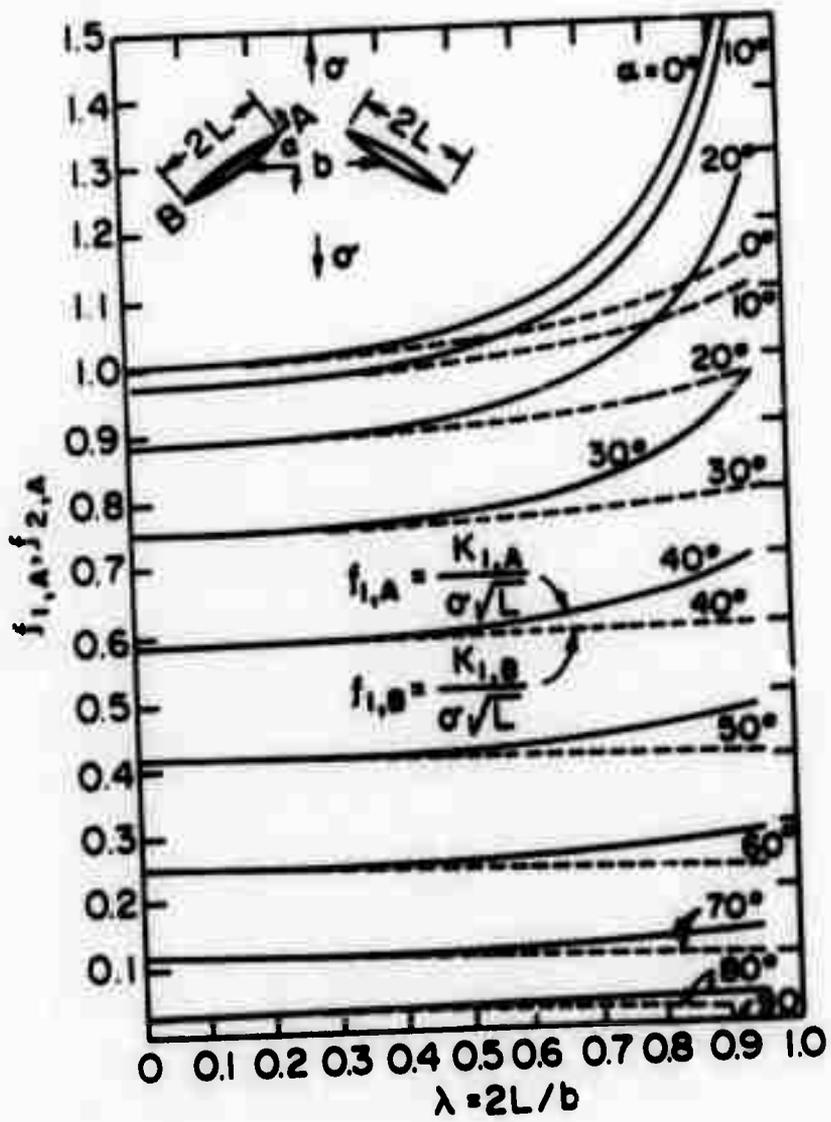


Figure 3.

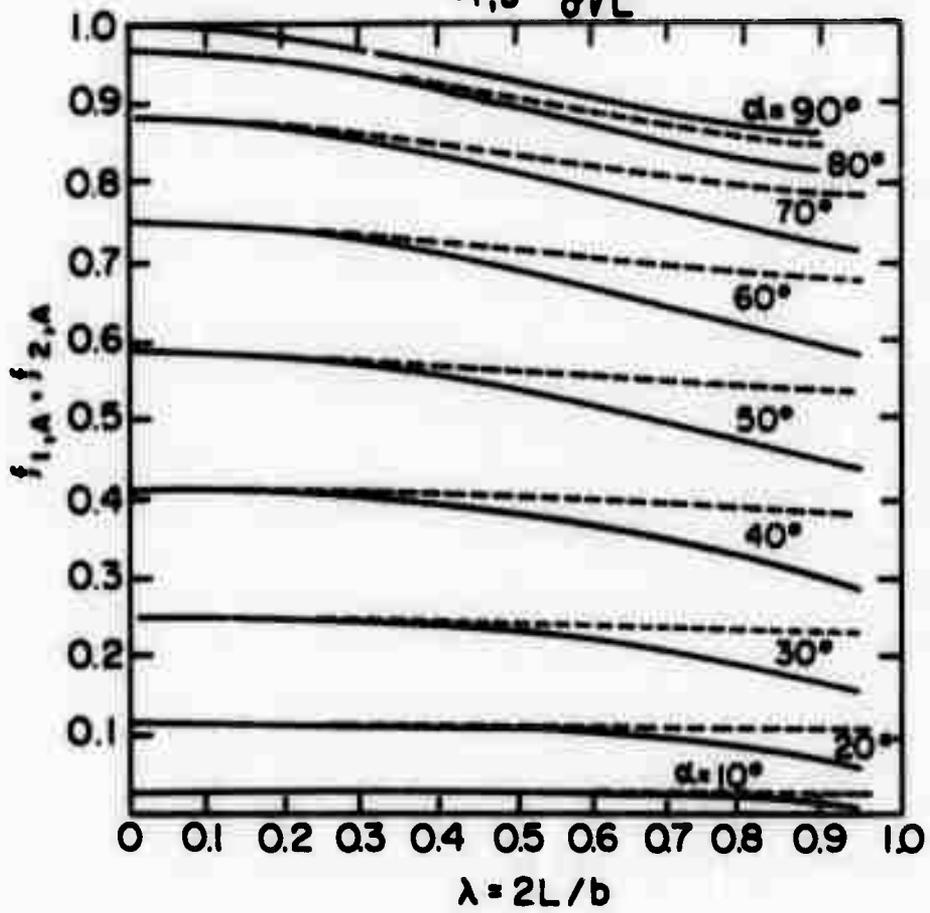
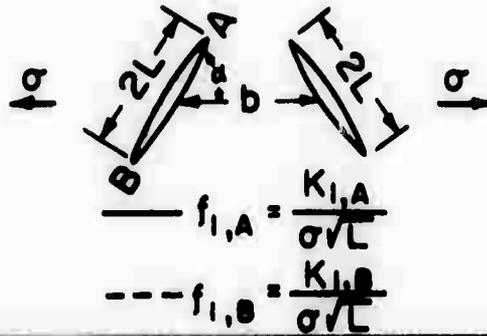


Figure 4.

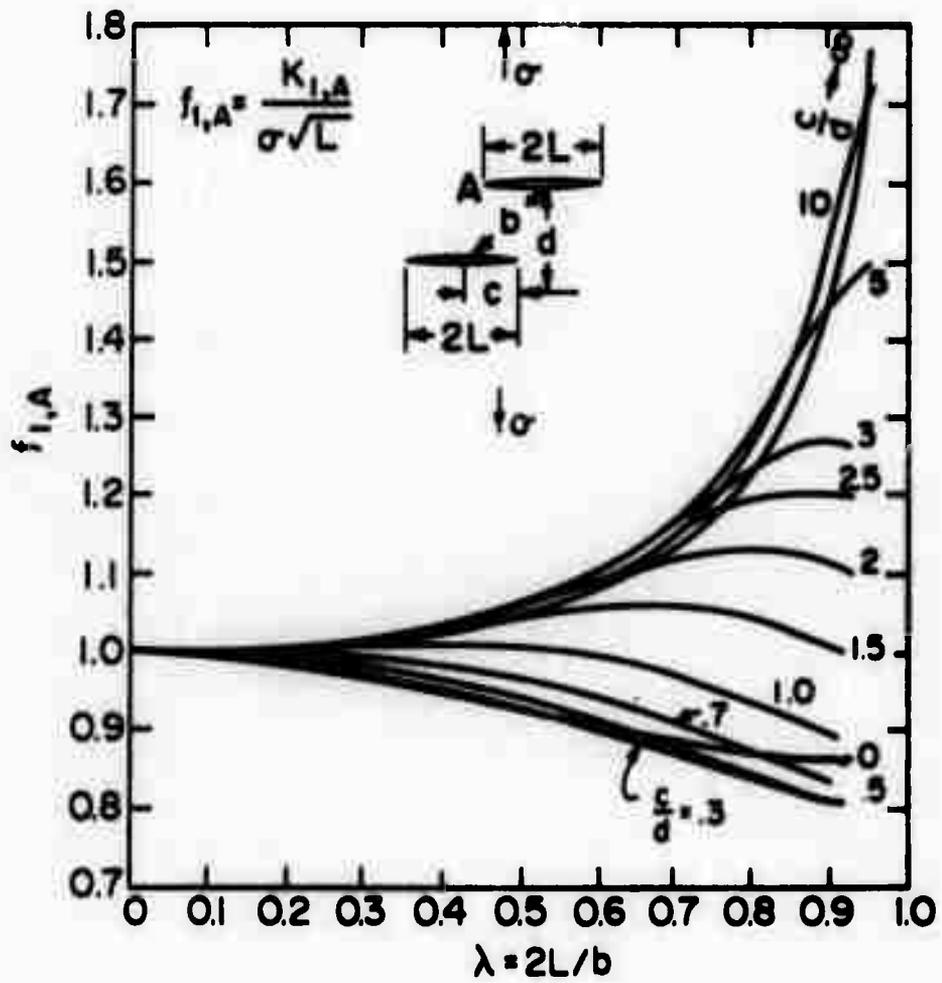


Figure 5.

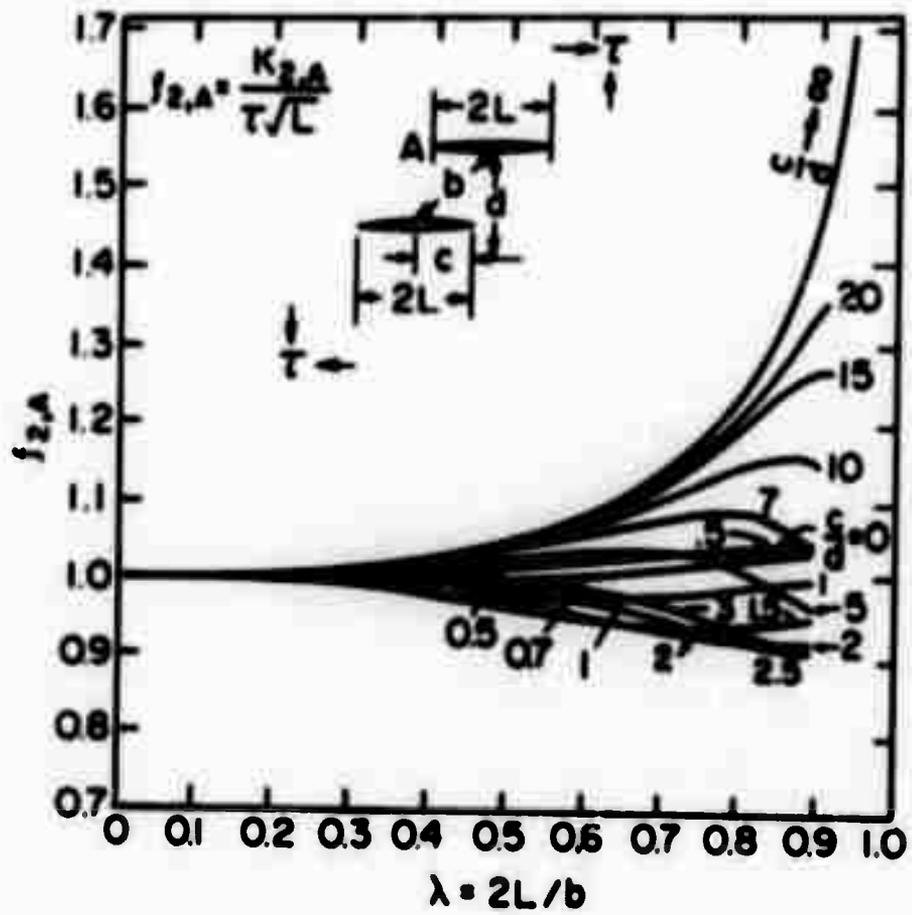


Figure 6.

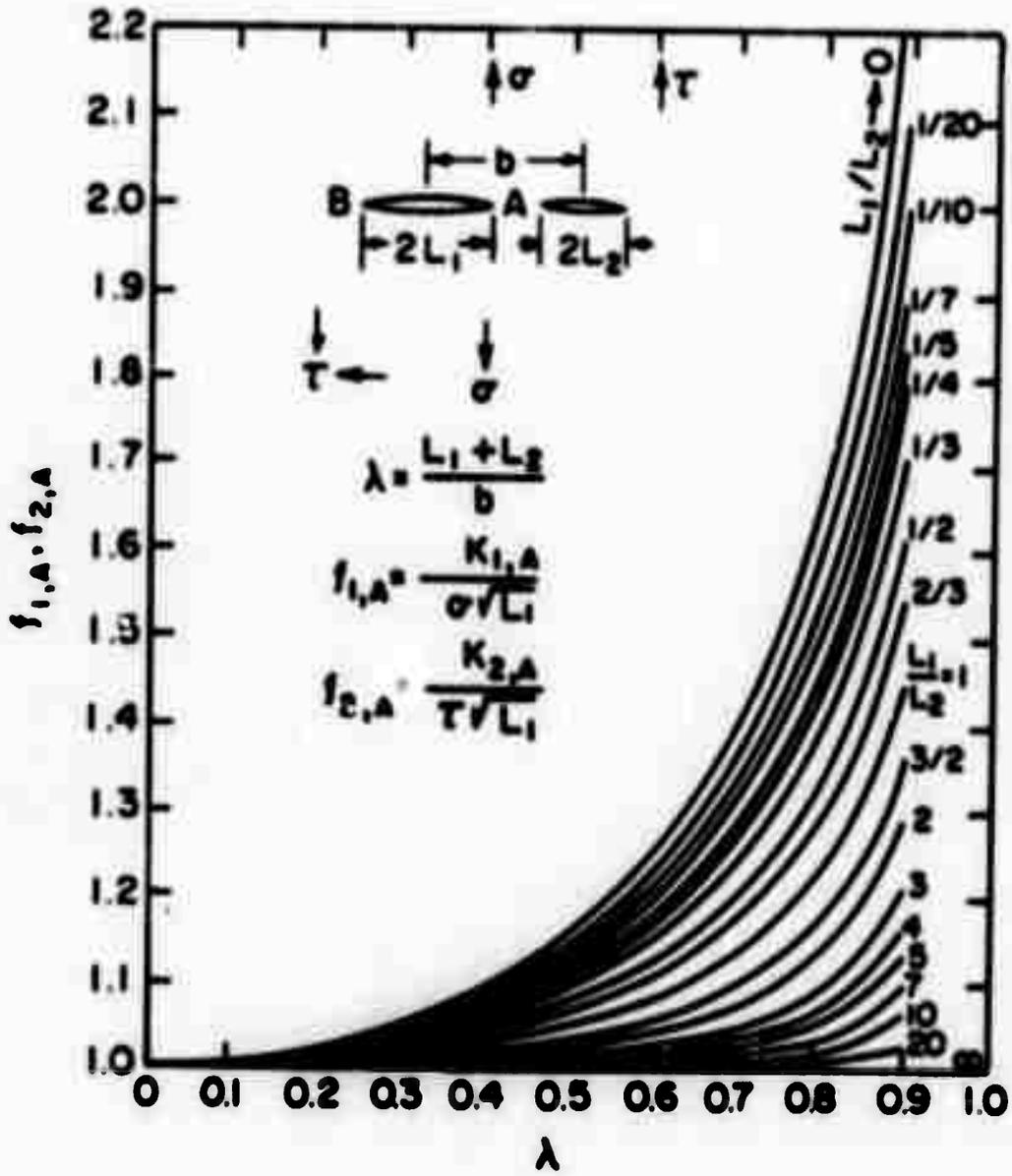


Figure 7.

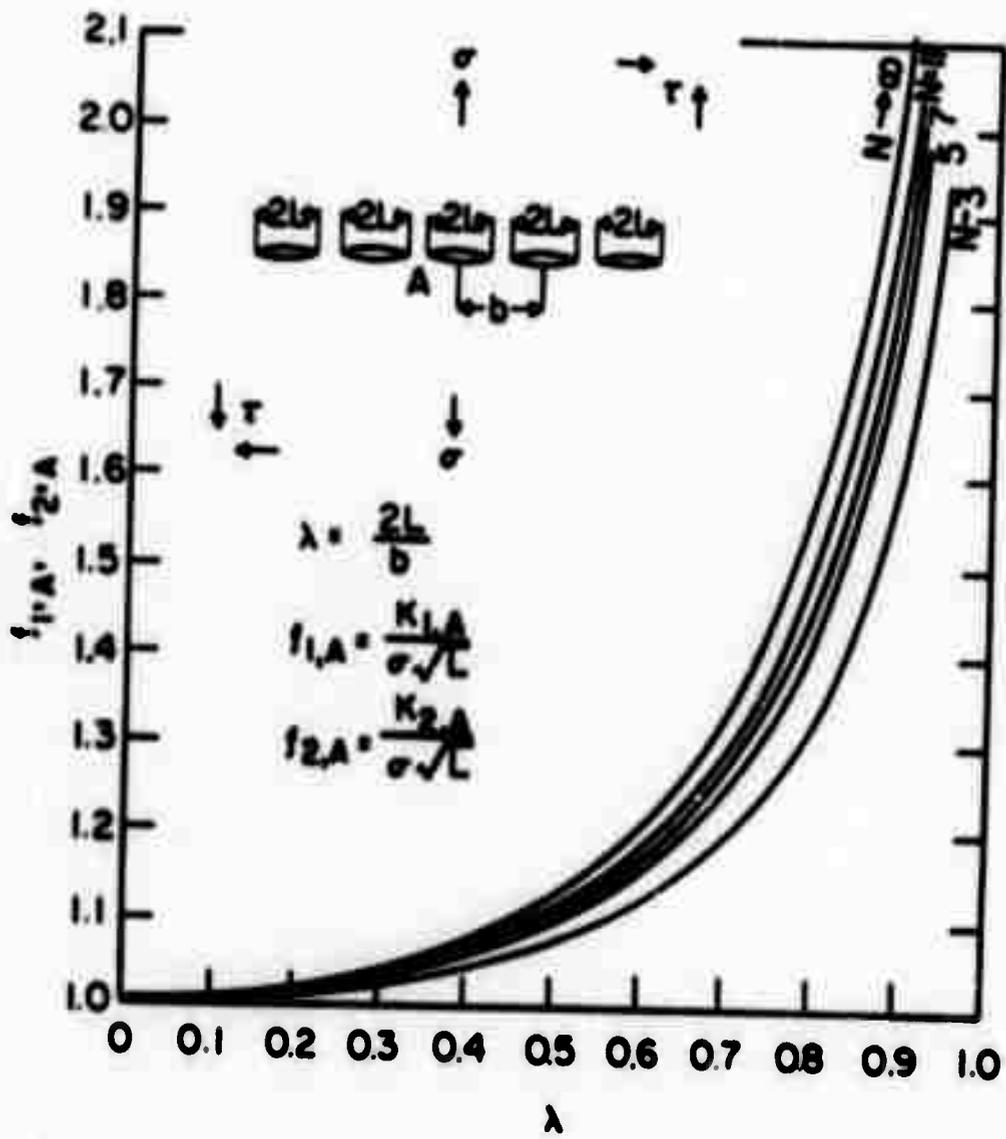


Figure 8.

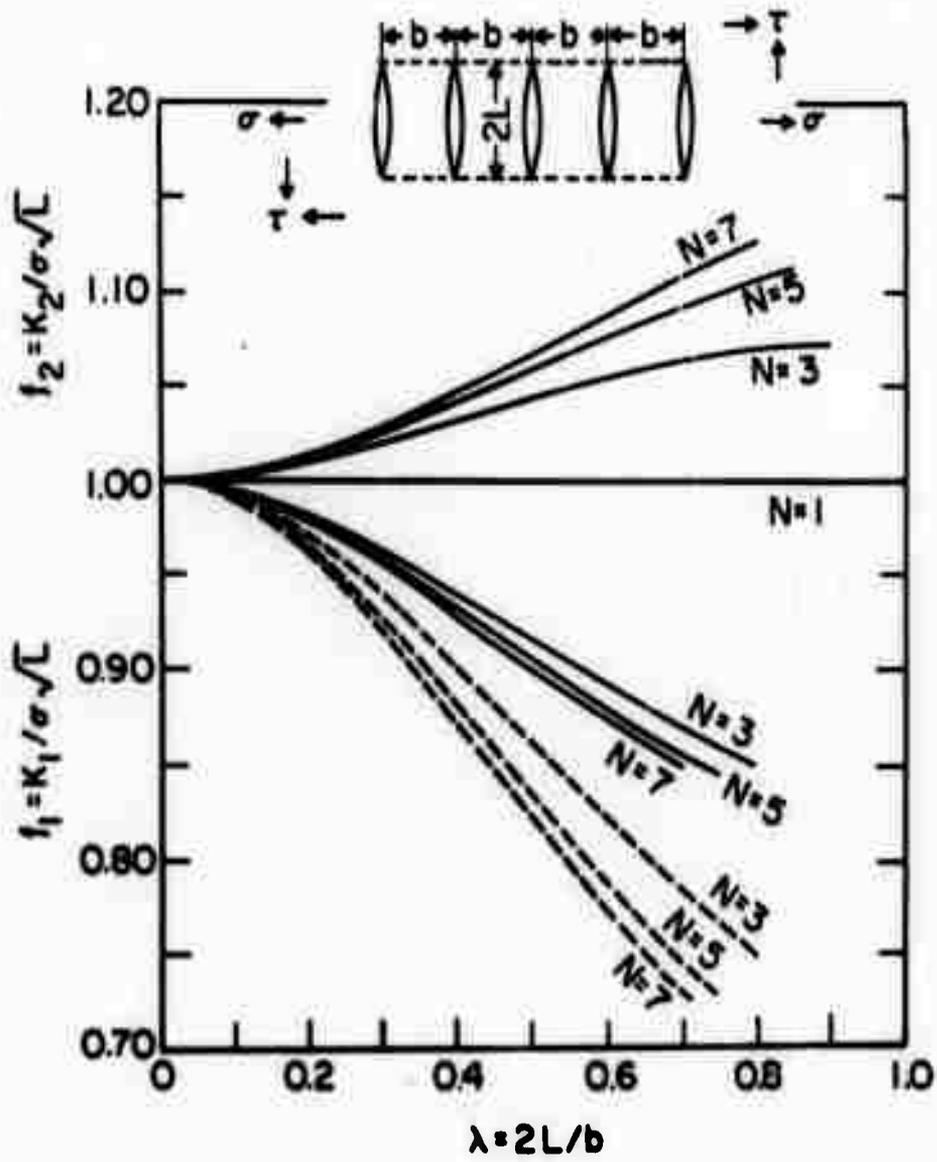


Figure 9.

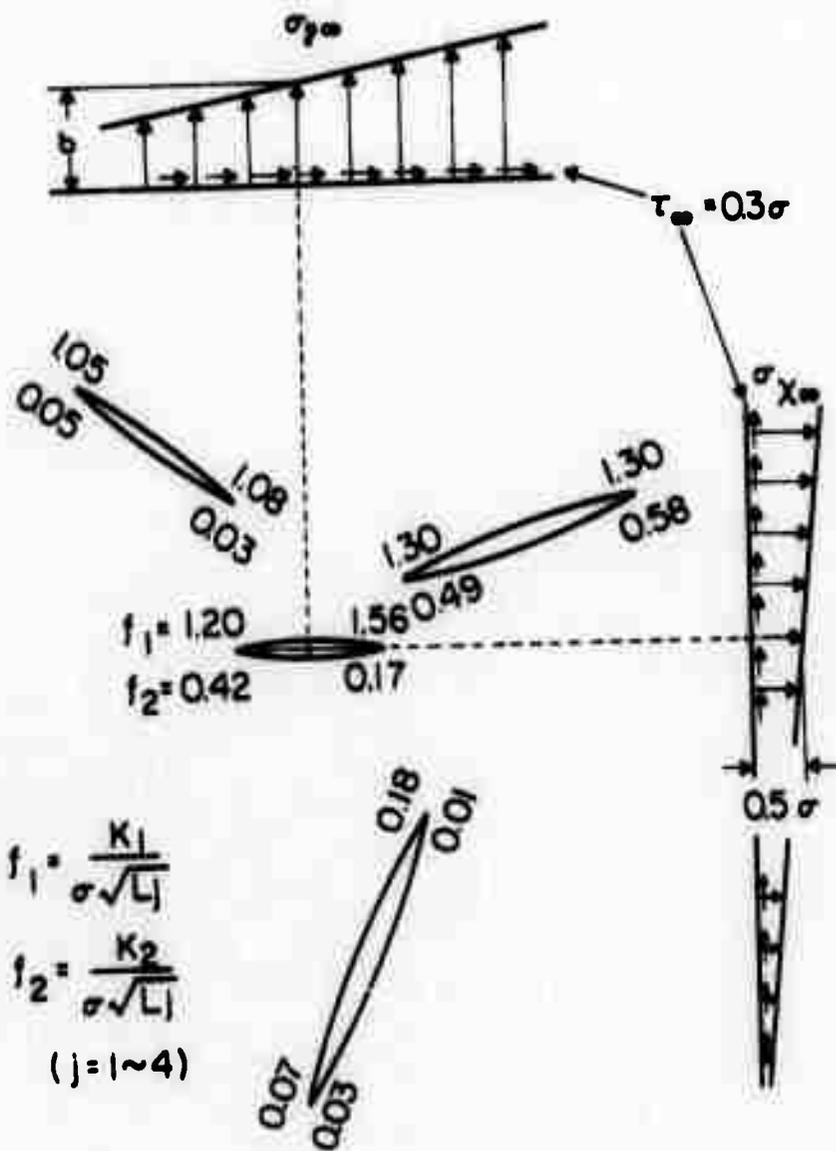


Figure 10.

EDGE DISLOCATION ARRAYS AROUND A CRACK
UNDER TENSION

F. A. McClintock

Introduction

Currently, one of the principal needs in fracture mechanics is the stress and strain field around a crack growing normal to the direction of tension (i.e., in Mode I) in an elastic-plastic material. Pending an exact solution, a representation making use of dislocation arrays may be useful in shedding light on the role played by residual stress and on the difference between growing and monotonically loaded cracks. A continuum solution has been found for shear parallel to the leading edge of the crack (Mode III) by Chitaley and McClintock (1971). There the effect of residual stress is especially important where the rows of screw dislocations left behind an advancing crack correspond to incompatible strains. In Mode I, however, the strain associated with trailing rows of edge dislocations having Burger's vectors perpendicular to the flank is not incompatible, and the effects of residual stress may be much less.

For simplicity, assume plastic flow to be confined to a single pair of planes emanating from the tip of the crack. Non-uniform slip along these planes gives a dislocation distribution,

as shown in Fig. 1. Crack growth leaves these dislocations behind in a wake region. Under Mode III shear Chitaley found that secondary flow in the wake region could be neglected. A similar assumption will be made and later examined here.

Such a study was made feasible by a solution obtained by J. R. Rice for the stress at a point z near the tip of a crack due to an arbitrary edge dislocation at the point z_d near the tip (Appendix 1). These results should also be obtained from the general anisotropic solution by Atkinson (1966), applied to a particular crystal by Heald and Atkinson (1967). It should be emphasized that the dislocations considered here are only the resultants of the real dislocations; these are a discrete model of the continuum model of the real dislocations.

The Yield Condition Under Loading Without Growth

It turns out to be convenient to use complex variables. Let the crack be along $z = x < 0$. Denote the deviatoric part of the stress tensor, referred to the x-y coordinates and normalized using the yield strength in shear, by

$$\tau_{dx} \equiv \left[\frac{\sigma_{yy} - \sigma_{xx}}{2} + i\sigma_{xy} \right] / k \quad (1)$$

Referred to an $r-\theta$ coordinate system, the stress deviator is

$$\tau_{dr} \equiv \left[\frac{\sigma_{\theta\theta} - \sigma_{rr}}{2} + i\sigma_{r\theta} \right] / k = \tau_{dx} e^{i\theta} \quad (2)$$

Normalize the coordinates with the height of the initially active plastic zone, H_1 :

$$\xi = x/H_1, \quad \eta = h/H_1, \quad \zeta = (x+iy)/H_1 = z/H_1 \quad (3)$$

Define a generalized Burgers vector and normalize it, using the slip height H_1 and the yield strain in shear, k/G .

$$\beta = (b_x + ib_y)/4\pi i(1+\nu)H_1(k/G) \quad (4)$$

For slip on planes at an angle θ to the x axis,

$$\beta/|\beta| = e^{i\theta} \quad (5)$$

The stress deviator at z due to a dislocation at z_d having a Burgers vector β can be represented in the form

$$\tau_{dx} = |\beta| F(\zeta, \zeta_d, \theta) \quad (6)$$

The function $F(\zeta, \zeta_d, \theta)$ will be given explicitly in a later section.

The resolved shear stress on the slip plane $\theta > 0$ due to a pair of dislocations at $\zeta_d, \bar{\zeta}_d$ can be stated from Eqs. 2 and 5 as

$$\text{Im}(\tau_{dr}) = \text{Im}(e^{2i\theta} |\beta| [F(\zeta, \zeta_d, \theta) + F(\zeta, \bar{\zeta}_d, -\theta)]) \quad (7)$$

The resolved shear stress due to the applied stress is given in terms of the stress intensity factors K_I and K_{II} for normal loading and in-plane shear:

$$\text{Im}(\tau_{ar}) = \text{Im} \left\{ e^{i\theta} \left[\frac{iK_{II}}{\sqrt{2\pi} k z^{3/2}} + \frac{K_I - iK_{II}}{4\sqrt{2\pi} k} \left(\frac{z-\bar{z}}{z^{3/2}} \right) \right] \right\} \quad (8a)$$

With symmetrical loading $K_{II} = 0$. Express K_I in terms of a normalized, nominal plastic zone size

$$R = \frac{1}{2} \left(\frac{K_I}{\sqrt{\pi} k} \right)^2 / H_1 \quad (8b)$$

With $K_{II} = 0$, substitution of Eq. 8b into 8a gives

$$\text{Im}(\tau_{ar}) = \text{Im} \left\{ e^{i\theta} \left\{ \sqrt{R} \frac{\text{Im} \bar{z}}{z^{3/2}} \right\} \right\} \quad (8c)$$

The condition for incipient motion of all dislocations along the upper slip plane is that the resolved shear stress due to the applied stress plus that due to the dislocations distributed along the initial slip plane, with density per unit height B_{1n} , out to the limit $h/H_1 = 1$, attain the normalized value of unity.

$$-1 = \text{Im} \left\{ e^{i\theta} \left\{ \frac{\sqrt{R} \text{Im} \zeta}{\zeta^{3/2}} + \int_0^1 |B_{1n}| \left[F(\zeta, \zeta_D, 0) + F(\zeta, \bar{\zeta}_D, -0) \right] dn \right\} \right\} \quad (9)$$

Equation 9 is a Fredholm integral equation of the first kind for the initial dislocation density $|B_{1n}|$, having a solution for only one value of the normalized plastic zone size R .

As the applied stress is increased, the distributions of stress and strain will remain geometrically similar, in view of the homogeneity of both the equations of continuum mechanics

and the boundary conditions (as long as the plastic zone is contained within a boundary along which the stress and strain are described by K_I). The coordinates of homologous points will be proportional to K_I^2 in view of the constancy of the parameter R from Eq. 9. Therefore any parameter involving scale provides an appropriate criterion for fracture initiation; extent of plastic zone, $H_1/\cos \theta$, stress intensity factor K_I , crack opening displacement $\int \beta_{1n} dn$ or stress at some particular point. It is the condition for crack growth that requires an understanding of the physical mechanisms involved. First consider the mechanics of steady-state growth.

Steady-state Stress and Strain Distributions

The steady-state growth of a crack, under a constant stress intensity factor, is of direct interest in stress corrosion cracking, for example, but may also provide a useful estimate of the onset of instability in plane stress Mode I cracking (e.g., Frisch, 1958; McClintock, 1958; Broek, 1968) and longitudinal shear or Mode III loading (McClintock, 1965). Its applicability to plane strain Mode I cracking is one of the questions to be studied with the analysis being developed here.

As mentioned in the introduction, it will be assumed that the only active flow is in the leading edge of the plastic zone. The dislocation density is then only a function of the normalized y coordinate η . When η is a variable of integration it will be denoted by η_v . The location of an element dislocation is thus at

$$\zeta_v = \xi_v + i\eta_v \quad (10)$$

The total Burgers vector within an element of area depends only on η_v and is

$$d\beta = \beta_{\xi\eta}(\eta_v) d\xi_v d\eta_v$$

The steady-state solution is obtained by integrating the contributions of the trails of dislocations left behind the leading edge;

$$-1 = \text{Im} e^{2i\theta} \left[\frac{R \text{Im}\zeta}{\zeta^{3/2}} + \int_0^1 |\beta_{\xi\eta}(\eta_v)| \int_{-\infty}^{\eta_v \cot\theta} F(\zeta, \zeta_v, \theta) + F(\zeta, \bar{\zeta}_v, -\theta) \right] d\xi_v d\eta_v \quad (11)$$

Effect of a Decohering Region

In some cases, especially with multiple cleavage crack initiation ahead of the main crack, with the crazed region in polymers, or with the necking region in thin sheets under plane stress, the fracture does not occur abruptly, but rather as a result of the gradual loss in strength of a thin region of localized dilatational flow (Berg, 1970, 1971). As we shall see, the decohering type of fracture is likely to involve either slip on multiple planes or no plastic flow at all, adjacent to the crack. In neither case do we treat it in this study; the discussion is included simply as an aid to physical insight and to help define the limits of the slip plane model of plasticity

discussed here. In this model, the point at which the flow localization begins can only depend on the stress (or strain) ahead of the tip of such a zone. For a decohering zone caused by intermittent cleavage fracture, the beginning of decohering might require the cleavage stress on an unnotched specimen at a structural distance ahead of the tip of the decohering region that was of the order of the grain size.

Apply traction boundary conditions at $x < c_d$ whenever

$$\sigma > \sigma_c \text{ at } x = c_d + \rho_s \quad (12)$$

Within the decohering region the normal traction \underline{t} will in general depend on the amount of cracking or hole growth indicated by a porosity p and on the triaxiality σ in the surrounding matrix:

$$t = t(p, \sigma) \quad (13)$$

The porosity may rise suddenly to an initial value \underline{p}_0 at the beginning of decohesion; thereafter it will rise with increasing normal displacement \underline{v}

$$p = p_0 + \int_0^v \frac{dp}{dv} (s, p) dv \quad (14)$$

Satisfaction of Eq. 14 relating traction to displacement may well require slip simultaneously on a number of parallel planes. Alternatively, it is possible that the decohering region will so attenuate the stress concentration that the surroundings will

be entirely elastic, as pictured by Elliott (1947), Barenblatt (1964) and Dugdale (1960). In either case the problem is beyond the scope of this article, and is mentioned here only for the sake of completeness.

Transient Growth

Once the crack has grown a first increment δc , the dislocation distribution on the slip planes emanating from the current crack tip is found from Eq. 9, modified to include the translation of the original slip plane behind the crack and the different height of the new slip plane (see Fig. 1):

$$\begin{aligned}
 -1 = & \operatorname{Im} e^{2i\theta} \frac{\sqrt{R} \operatorname{Im} \zeta}{\zeta^{3/2}} \sqrt{\frac{H_1}{H_2}} \\
 & + \int_0^1 |\beta_{1\eta}| \left[F\left(\zeta, (\eta_V \cot\theta - \delta c + l\eta_V), \theta\right) \right. \\
 & \left. + F\left(\zeta, (\eta_V \cot\theta - \delta c - l\eta_V), -\theta\right) \right] d\eta_V \\
 & + \int_0^{H_2/H_1} |\beta_{2\eta}| \left[F(\zeta, \eta_V \cot\theta + l\eta_V, \theta) \right. \\
 & \left. + F(\zeta, \eta_V \cot\theta - l\eta_V, -\theta) \right] d\eta_V \tag{15}
 \end{aligned}$$

In the limit for a number of increments, as $\delta c \rightarrow 0$, Eq. 15 becomes a Volterra integral equation of the second kind, with an imbedded Fredholm equation of the first kind, in terms of the cross-derivative of the dislocation density, $\partial^2 \beta / \partial \eta \partial \xi$.

The Volterra integral equation of the second kind was encountered in anti-plane shear (Mode III) by McClintock (1958), but there it was possible to assume that the extent of the plastic zone was nearly constant, as shown by Chitale and McClintock (1971) for the steady state case.

Fracture Criteria

A criterion for cleavage fracture has been discussed above. For hole growth, Joyce (1968) has found a blunting followed by alternating sliding off on a microscopic scale that opens up a 90° groove. This is apparently the cause of the so-called "stretch zones" or more accurately "sheared zones" on overloading a fatigue crack. Fractographs show that after a certain extension, holes begin to occur on the surface, often abruptly, and the flank angle becomes much less than 90°. It is not known to what extent this transition is due to changing strain distribution due to a mixture of strain hardening and crack growth, to what extent it is due to more prestrain and hence lower ductility of the material approaching the current tip, and to what extent it is due to strain over a larger region. In crack growth on a zig-zag path, the shear cracking leads to a rapid crack extension per unit opening which may then decrease as the plastic zone becomes skewed. Blunting may occur, and in any case, there is subsequent rapid growth on the conjugate plane, as sketched in Fig. 2. Non-uniform development of the crack along its front leads to an average curve. Pending a

more complete understanding of such phenomena, it seems reasonable to take as a fracture criterion some average slope, representing crack growth per unit crack opening displacement. The crack opening at which the growth rate accelerates provides the scale factor essential to deducing a stress intensity factor, and the relevance of the crack opening angle has been discussed by McClintock (1968, 1969).

With such a fracture criterion, growth appears immediately on loading. For simplicity the crack growth in the blunting stage may be neglected, so all slip is confined to a single plane. Equation 9 then applies, with the condition for the beginning of more rapid growth being the crack opening displacement:

$$2v_1 = 2 \sin \theta \int_0^r |B_{1n}| dn \quad (16)$$

After an incremental crack growth $\delta c_1 = c_2 - c_1$, Eq. 15 applies with the new plastic zone found from

$$2(v_2 - v_1) = \frac{2\delta c_1}{dc/dv} = 2 \sin \theta \int_0^{H_2/H_1} |B_{2n}| dn_v \quad (17)$$

Very likely the critical question is whether instability occurs on the first growth beyond c_1 , i.e., whether K_I drops. If not, an integral equation could be developed for continued growth. The need for a numerical approximation to these integral equations is clear.

Discretization of Equations

For numerical solution the continuous dislocation arrays can be separated into discrete dislocations. To keep track of strain, it may be more convenient to regard the flow as being due to dislocation sources located at specific points, as shown in Fig. 3. The dislocations produced by these sources intersect (and nearly cancel) at the midpoints of the segments between the sources. The stress at each source is equated to the yield strength in shear. For instance, for initial yielding, consider I_{\max} dislocations with sources and dislocations respectively at

$$\begin{aligned} \zeta(I) &= \left(\frac{e^{i\theta}}{\sin \theta} \right) \left(\frac{I - \frac{1}{2}}{I_{\max} + \frac{1}{2}} \right) && \text{for } I=1 \text{ to } I_{\max}+1 \\ \zeta_d(J) &= \left(\frac{e^i}{\sin} \right) \frac{J}{I_{\max} + \frac{1}{2}} && \text{for } J = 1 \text{ to } I_{\max} \end{aligned} \quad (18)$$

The corresponding discrete form of Eq. 9 is

$$\begin{aligned} -1 = \operatorname{Im} e^{i\theta} \left\{ \frac{\sqrt{R} \operatorname{Im} \zeta(I)}{\zeta(I)^{3/2}} + \sum_{J=1}^{I_{\max}} |\beta(J)| \left[F(\zeta(I), \zeta_d(J), 0) \right. \right. \\ \left. \left. + F(\zeta(I), \bar{\zeta}_d(J), -0) \right] \right\} \end{aligned} \quad (19)$$

Equation 19 can be placed in a simple matrix form if the root of the normalized plastic zone size, \sqrt{R} , is regarded as $|\beta(I_{\max}+1)|$. Then one can write Eq. 19 as

$$-1 = \Lambda_{ij} |\beta_j| \quad i, j = 1 \text{ to } I_{\max} + 1 \quad (20)$$

As a first investigation, take $\theta = 90^\circ$, $I_{\max} = 1, 2, 3$, and 10.

For steady state growth, Eq. 11 becomes

$$-1 = \operatorname{Im} e^{i\theta} \frac{\sqrt{R} \operatorname{Im} \zeta(I)}{\zeta(I)^{1/2}} + \sum_{J=1}^{I_{\max}} |B_{\zeta}(J)| \int_{-\infty}^{J \cot / (I_{\max} + 1/2)} (F(\zeta(I), \zeta_v(J), 0) + F(\zeta(I), \bar{\zeta}_v(J), -0)) d\zeta_v \quad (21)$$

where $\zeta_v(J) = \zeta_v + i J / (I_{\max} + 1/2)$. The integrand approaches $\sim 1/\zeta^2$, so the integration can be carried out numerically to determine the coefficients in an equation of the form of Eq. 20.

Acknowledgement

This work was carried out at Brown University under the ARPA Grant E 744 and completed in the Materials Research Council Summer Study, July 1971, under the Advanced Research Projects Agency of the Department of Defense under Contract No. DAH15-71-C-0253 with The University of Michigan.

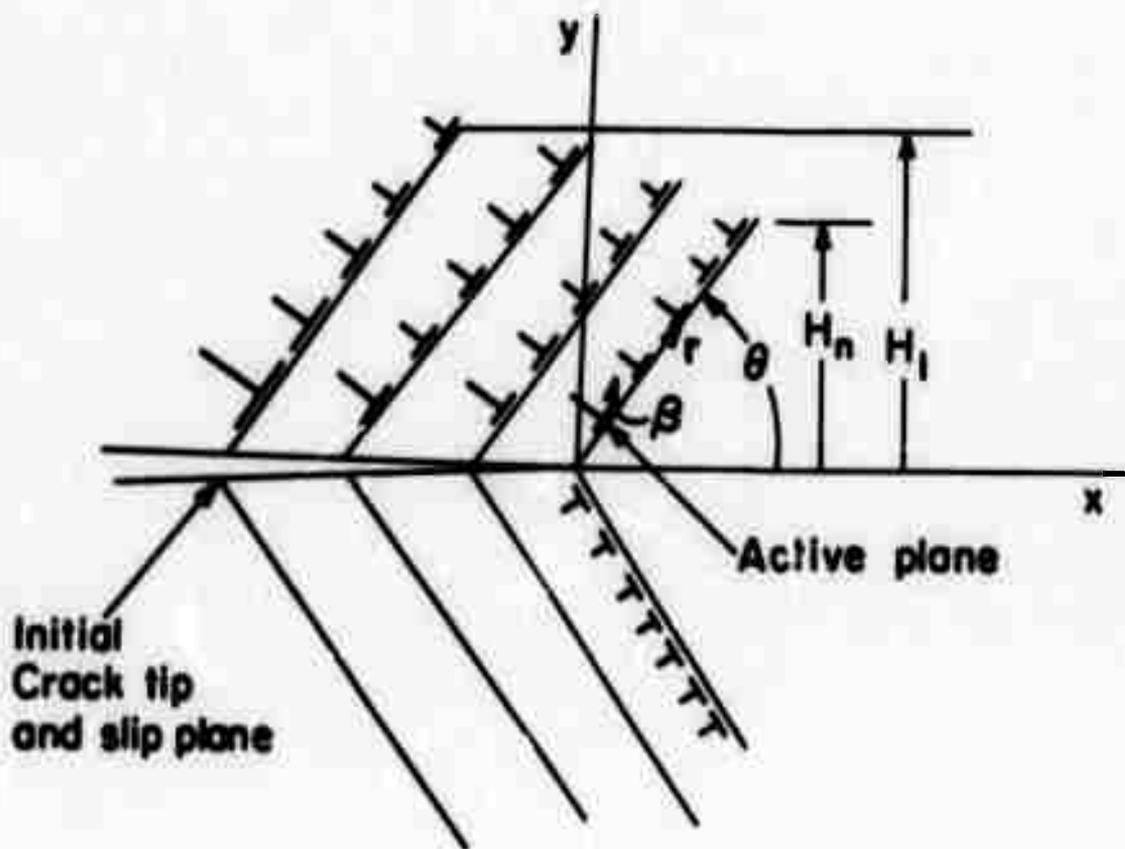
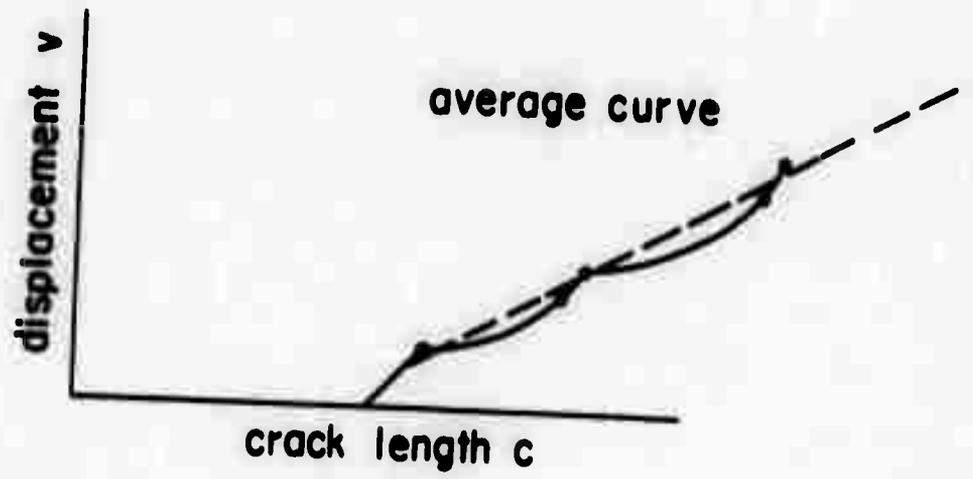
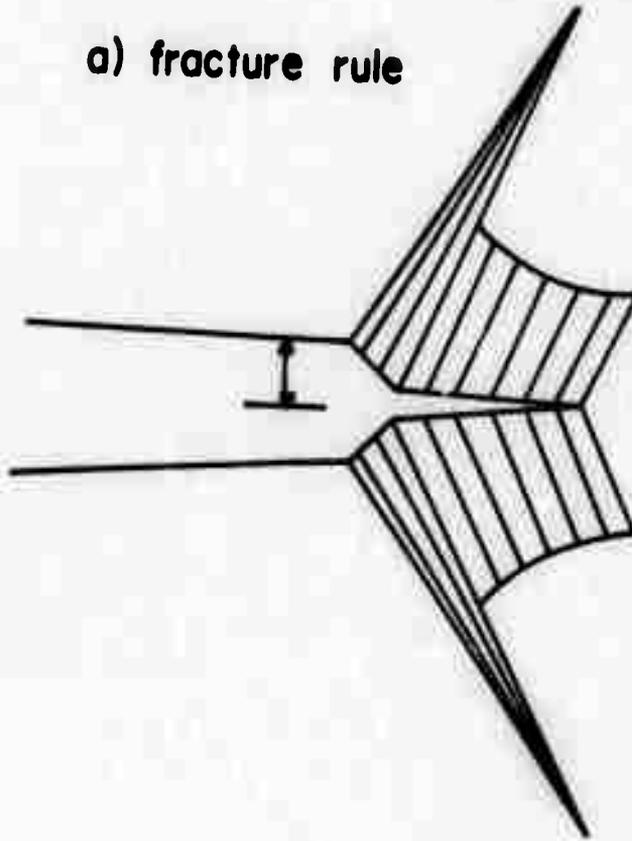


Figure 1: Dislocations on slip planes during crack growth.



a) fracture rule



b) Crack profile and slip lines

Figure 2: Ductile crack propagation

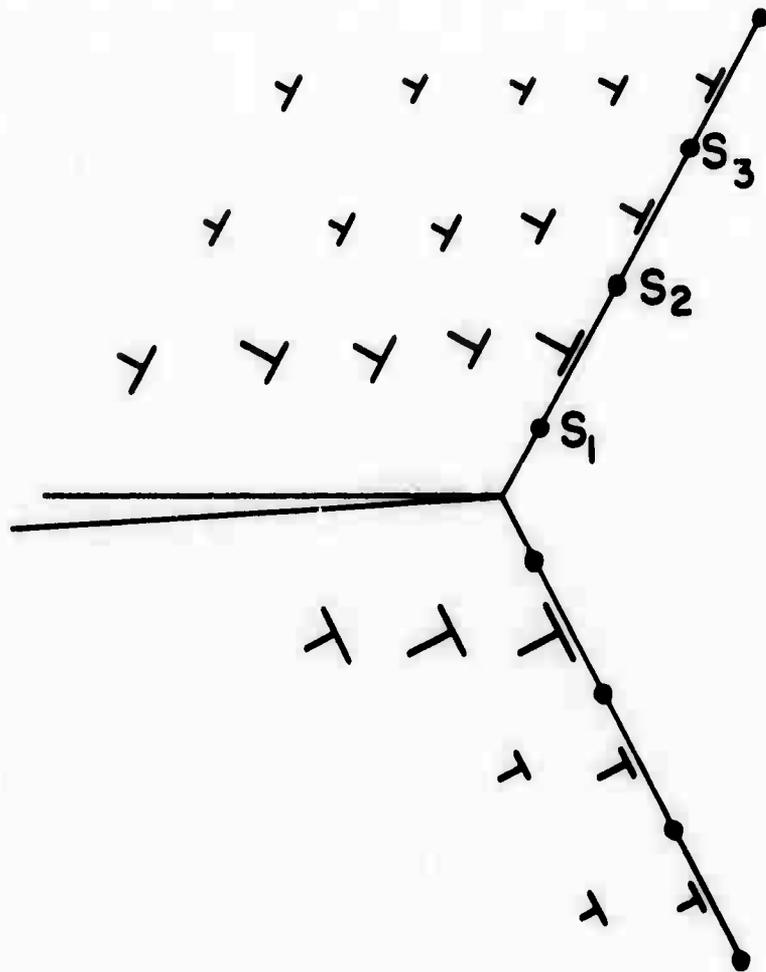


Figure 3: Trails of dislocations behind a crack tip in steady flow.

HIGH-TEMPERATURE STABILITY OF SILICON NITRIDE

J. L. Wood, G. P. Adams and J. L. Margrave

Abstract

Within the last decade, the level of interest in silicon nitride as a practical refractory has risen meteorically and it is somewhat surprising that high-temperature thermodynamic properties are so poorly known. Fluorine combustion of a variety of Si_3N_4 samples should lead to a reliable value for the standard heat of formation. This information, coupled with available and/or estimated thermodynamic functions for the material at elevated temperatures, can be used to calculate dissociation pressures for comparison with experimental data. Mass spectrometric studies allow one to demonstrate that only elemental species, $\text{Si}(\text{g})$ and $\text{N}_2(\text{g})$, are important in ordinary decomposition but other complex species are known-- Si_2N and SiN --which may play a role in forming techniques and establishing ultimate use limitations.

Preliminary calorimetric results establish the heat of formation of $\text{Si}_3\text{N}_4(\text{s})$ in the range -200 ± 10 kcal/mole, which implies a considerably greater thermal stability than predicted from earlier JANAF thermal data.

Preceding page blank

THE USE OF LEVITATION IN INORGANIC SYNTHESIS*

J. L. Margrave, J. A. Treverton and P. W. Wilson

Abstract

The possible uses of levitation as an inorganic synthetic technique have been investigated. It was found that levitation is a useful method for the preparation of all types of compounds excepting those that are both non-conducting and non-volatile. Observations suggest that if the product may have any of several oxidation states, the product will likely have the lowest oxidation state.

*Published in High Temperature Science, Vol. 3, (1971).

Preceding page blank

THE USE OF LEVITATION IN INORGANIC SYNTHESIS

J. L. Margrave, J. A. Treverton and P. W. Wilson

INTRODUCTION

The phenomenon of levitation of conducting and semi-conducting materials has been known for a considerable time. Unfortunately the technique has been more of a scientific curiosity than a useful approach to scientific problems. Some efforts have been made to exploit the unique properties of the levitation technique,¹⁻² but most of these were not very productive.

Probably the most useful work done to date has centered around measurements of the physical properties of the levitated materials. For example, emissivity measurements have been made, and heats of fusion of metals have been measured using a combination levitation-drop calorimeter apparatus.³⁻⁴ The metal is suspended in the magnetic field, heated and fused by interaction with the field, and then dropped into a calorimeter. The advantages of this method are obvious - high temperatures are easily attained, there is no container problem, and one may use either vacuum or controlled atmospheres.

There is one report in the literature of an experiment that made limited use of synthesis with a levitated metal.⁵

Cobalt was suspended in a field and CO flowed around it. Cobalt oxide was formed and the equilibrium constants for the system were determined. Except for this single experiment there appears to have been no application of the levitation technique for synthesis.

The advantages of the levitation technique are numerous. The most important are those mentioned above; high temperatures can be attained in an atmosphere of choice, and since the material is suspended, there is no container to react or contaminate. It is the combination of these conditions that should make the levitation technique broadly applicable. At high temperatures many materials are very reactive and finding suitable containers can be difficult. Indeed, with molten compounds at high temperatures, the container problem is usually the limiting factor, and certain experiments cannot even be performed.

Levitation techniques can be used for studying various types of reactions - solid-solid, solid-liquid, liquid-liquid, gas-solid or gas liquid reactions. Solid-solid reactions usually pose no particular problem using ordinary techniques and would not take full advantage of the unique qualities of levitation. Solid-liquid reactions or liquid-liquid reactions (where the liquids are probably molten solids) could be conducted to advantage in a levitating r.f. field. The mixture of reactants could be levitated, heated until reaction was completed, cooled and the product recovered. Provided the various compounds

involved were suitable for levitation, these reactions should be relatively easy to carry out. No experiments on these two types of reaction have yet been conducted since it was felt that the most interesting reactions would be gas-solid or gas-liquid reactions. In this field the range of reactants is broad (only one reactant has to be levitated) and so the range of product type is also broad. Reaction systems can be chosen that are adequately representative of various types. It is impossible to predict the course of these reactions without experimentation. For instance, it is hard to predict whether reaction will occur on the surface of a liquid, in the body of a liquid or in the gas phase as the reactant boils off the hot or molten levitated reactant.

To investigate these problems a number of gas-solid and gas-liquid reactions were studied. Reactions were chosen to represent as complete a range as possible. The aim of the study was to define the limits of the levitation technique and to determine those types of reactions that could be most successfully studied using levitation techniques. It was felt that the best approach was to study reactions that were already known, so that any advantage the levitation technique offered would be obvious.

EXPERIMENTAL AND RESULTS

The samples were levitated and heated using a General Electric 15 KW induction heater (Model HM15L3) operating at 450 KHz. The induction heater was coupled to the levitation

coil via a transformer. The chamber wherein the samples were levitated was a modified 500 ml glass bulb. A gas inlet, gas outlet, and a quartz pyrometer window were fitted to this bulb. Temperatures were measured using an optical pyrometer. All of this equipment has been described in detail previously.³

Temperature control was achieved by altering the atmosphere in the reaction chamber. In a pure argon atmosphere the samples could be heated to 2700°C. In a pure helium atmosphere the samples could be maintained at temperatures as low as 300-400°C. By judiciously mixing argon and helium, any temperature between these two extremes could be maintained.

Metallic samples were easy to levitate. The sample was placed in a silica cup on a pyrex rod. The induction heater was turned on and the sample was raised into the center of the coil where it immediately levitated. Some materials (for example, Si and Ge) are not sufficiently conducting at room temperature to levitate. For these materials, the silica cup was lined with a metal, such as molybdenum. In the coil, the molybdenum was heated; this heated the sample, and at some temperature (generally around 400 or 500°C) the sample became sufficiently conducting to levitate. The metal lined silica cup was then withdrawn.

All of the solids and gases used are commercially available and were used without further purification. Reactions were examined in the following manner. The solid starting material (2 gr, lump) was levitated in an argon atmosphere; helium was bled

into the reaction chamber to control the temperature, which was generally held at 2000°C. The gaseous reactant was then bled into the chamber and reaction was allowed to proceed until the levitated sample has been consumed, or until the accumulated products made further levitation of the sample impossible.

The products recovered were of two types: (1) volatile compounds that had condensed onto the glassware and the cooled copper levitation coil, (2) the residue of the levitated sample. Sometimes this residue was unreacted starting material, sometimes a mixture of reactant and product, and sometimes a pure product.

All of the products were identified using x-ray powder diffraction patterns. Samples were ground, placed in 0.3 capillary tubes and exposed for 16 hrs. in a conventional camera. The sample composition could be gauged to an accuracy of about 5% which was sufficient for this experiment.

Results

One of the first experiments run was to determine what type of sample can be levitated. It has been determined in this laboratory that almost all of the metals can be levitated. The following compounds were also levitated: CaSi, SiC, Al₄C₃, ZrB₂ and MoSi₂; B₄C would heat but not levitate. Attempts to levitate CaF₂ and NaCl were unsuccessful.

The results of the reactions tried are listed in Table 1.

TABLE 1

LEVITATION SYNTHESSES

Experiment No.	Reactants		Products	
	levitated compound ^a	gas	recovered from coil	residue
1	Si	air	SiO ₂ (100) ^b	none
2	Si	N ₂	Si (95), Si ₃ N ₄ (5)	Si (95), Si ₃ N ₄ (5)
3	Si	NH ₃	Si (95), Si ₃ N ₄ (5)	Si (50), Si ₃ N ₄ (5)
4	Si	CH ₄	none	SiC (100)
5	Ge	NH ₃	none	Ge (100)
6	Ge	N ₂ S	GeS (100)	Ge (100)
7	Al	NH ₃	AlN (100)	Al (40), AlN (60)
8	Al	CH ₄	none	Al ₃ C ₄ (100)
9	Ti	CH ₄	none	TiC (100)
10	Fe	HCl	FeCl ₂ (100)	none

^a The sample temperature was about 2000°C.

^b The number in parenthesis represents the percentage composition of the recovered product.

DISCUSSION

Generalizing from the above results it appears that metals, silicides, and some borides and carbides are suitable for levitation. It is not easy to levitate ionic salts, even at elevated temperatures.

Possibly the most important result of this work is the observation that many reactions occur between the gas and the molten solid, and not in the gas phase. This is consistent with the observation that products remained in the residue in experiments 3, 4, 7, 8 and 9. Had reaction occurred in the gas phase, all of the product would have been recovered from the coils and glassware; there would be no product in the residue. This observation is of importance where it might be difficult to anticipate the oxidation state of product. For example, in experiment 10 the product may have been either FeCl_2 or FeCl_3 . However, since the reaction occurred between the gas and the molten iron, the iron was always present in effective excess. This insures that the compound of lower oxidation state will be formed, as confirmed by the results. It appears that the use of levitation techniques will not lead to the formation of new halides of oxidation state higher than normally observed. Conversely, however, it should be an excellent technique for the production of halides of lower oxidation state, and be of use in the production of the lower halides of transition, rare-earth and actinide elements.

The other reactions and results can be most easily classified in terms of the physical properties of the products. The reaction of Si with NH_3 , (experiment no. 3) gave a product that was both non-volatile and non-conducting. Reaction proceeded until the accumulation of product lowered the conduction of the residue, prevented the sample from heating and levitating, and so extinguished the reaction. Levitation is not an especially useful technique where the product is non-volatile and non-conducting.

Another group of reactions (no. 1, 6, 7 and 10) gave products that were essentially pure. These were reactions where the product was volatile at the temperatures used. The reactions proceeded to completion in most cases and the levitation technique should be of great use in preparing compounds of this type.

Experiment no. 5 was run to see what would happen if the product was thermally unstable at the reaction temperature. It has been observed that products are often formed at temperatures at which they are unstable, but provided they are quickly quenched, they can sometimes be recovered. In this experiment, however, there was no evidence of any reaction. Either Ge_3N_4 did not form, or if it did, it did not volatilize before it decomposed.

The next group of successful reactions included those in which the product was non-volatile but it was conducting (experiments no. 4, 8, 9). These reactions proceeded to give complete conversion to a pure single product. This is another

field where levitation techniques should be useful.

To summarize, it appears that levitation is applicable to most types of reaction, except those where the product is both non-volatile and non-conducting. Levitation has been used successfully to prepare compounds that were either volatile or conducting. A result of consequence in proposing reactions that might be studied using levitation is the realization that if two products can be formed, of different oxidation state but similar thermal stability, the product actually formed will likely be the one of lower oxidation state. This occurs because reaction takes place in the presence of an effective excess of metal. Provided these criteria are considered when experiments are proposed, it can be anticipated that levitation techniques allow the preparation of a number of compounds not yet known and make the preparation of some compounds that are known much easier.

ACKNOWLEDGEMENTS

This research has been supported by the United States Atomic Energy Commission and the United States Army Office, Durham, North Carolina.

REFERENCES

1. D. M. Wroughton and E. C. Okress, *J. Electrochem. Soc.*, 99, 205 (1952).
2. E. C. Okress and D. M. Wroughton, *J. Appl. Phys.*, 23, 545 (1952).
3. A. K. Chaudhuri, D. W. Bonnell, L. A. Ford and J. L. Margrave, *High Temp. Sci.*, 2, 203 (1970).
4. J. A. Treverton and J. L. Margrave, *Proceedings of 5th Symposium on Thermophysical Properties, Boston, Mass., Oct. 2, 1970*, pp. 489-494.
5. G. W. Toop and F. D. Richardson, "Advanced in Extractive Metallurgy," Elsevier Publ. Co., Lausanne, 1968.

HIGH-TEMPERATURE PROPERTIES OF Nb AND Zr*

D. W. Bonnell, J. L. Margrave
and A. J. Valerga

Abstract

Various high-temperature data available on these two important elements will be presented including new data from levitation calorimetric studies. A complete set of thermodynamic properties will be presented for both the solid and liquid phases. The emissivities of both solid and liquid phases will be discussed.

*Taken in part from the Ph.D. Thesis of D. W. Bonnell, Rice University, May, 1972.

Preceding page blank

NIOBIUM

In an effort to begin filling in the second long transition metal period, and perhaps to obtain an estimate of the thermodynamic properties of tantalum, which had been intractable even to the levitation process, the study of niobium was undertaken. Coils of the solenoid design wound on a 0.375 in. form were used, and it was necessary to tighten the upper turns to an inner diameter of 0.250 in. in order to maintain a stable melt of niobium.

The results of these determinations are presented in Table 1. An extrapolation by Hultgren, et al.¹ of solid data from 1400 K to the melting point (2750 K) was used to obtain ΔH_{2750} (fusion) as 34570 ± 240 j/mole and ΔS_{2750} (fusion) as 12.57 ± 0.09 j/mole.K (3.00 cal/mole.K). This can be compared to the estimate, based on ΔS_{2045} (fusion) = 9.6 j/mole.K for the heat of fusion as 26415 j/mole. The heat capacity obtained was 40.6 ± 1.0 j/mole.K (9.70 cal/mole.K), chosen by Hultgren.¹ The melting point used has been recently determined by Cezairliyan² using pulse techniques.

Drops of the solid material were made in order to confirm the extrapolation. A helium atmosphere was necessary to resolidify the melt. The scatter in these data was higher than the liquid, presumably due to thermal gradients in the sample, even though the sample was solidified from a molten sample before the drop. The values derived for enthalpy do, however, agree in substance with the extrapolated values.

ZIRCONIUM

In undertaking the study of zirconium, several factors were important. From the experimental standpoint, the ease with which zirconium forms both nitrides and oxides required a purification train totally eliminating even trace amounts of nitrogen and oxygen. An initial purification train element consisting of a glass tube packed with clean lithium chips was not satisfactory and it was necessary to build a stainless steel trap holding calcium turnings heated to 600°C in order to scavenge nitrogen from the argon gas. Four to six hour flush times were necessary to purge the levitation chamber. Since both oxides and nitrides of zirconium are exceptionally refractory, and their vapor pressures are even lower than the metal, back-diffusion of atmospheric nitrogen through minute leakage paths forced the adoption of stringent measures to seal the system.

The coils used were of the solenoid design, with a very small (~0.20 in.) gap between upper and lower turns, which were both wound on the 0.375 in. form. The chemically interesting factor in this system was the extremely long temperature range which could be studied. Normally, if the power requirement can be met, the upper temperature limit is set by the vapor pressure of the metal. When the vapor pressure of the substance in question rises much above 1×10^{-4} atm., the physical quantity of material evolved per unit time is sufficient to obscure the pyrometer sighting path. Zirconium has a liquid range of more than 1000°C for which the vapor pressure is less than 1×10^{-4} .

The obvious trend in all the systems studied to this point was that, to experimental accuracy, the heat capacity of the liquid phase was constant. The longer base line provided by this large liquid range was expected to provide a good test of the linearity of the enthalpy function.

Table 2 presents the results obtained so far in this study.¹ The linearity of the fit to the data for the range 2128 K to 2839 K, a span of more than 700 degrees, is representative of the best work obtained from this apparatus. The inclusion of the three higher points increases the average error by more than a factor of 2. It seems quite unlikely that this is merely random error for two reasons. First, the minimum deviation of the higher points is nearly four times the average and the maximum deviation almost ten times the average. For purely statistical reasons, this seems unlikely. Second, all of the normally expected systematic errors tend to cause data points to fall below the line. For instance, oxidation effects lead to higher surface emissivities which would give a higher apparent brightness temperature based on the assumed emissivity for the pure metal; any loss of material, or any loss of energy during the drop would produce an anomalously low enthalpy. Only a failure of the assumption that emissivity is a constant for liquids, or the appearance of a temperature dependence in C_p would seem to explain the trend of the upper data points.

An attempt to treat the possible temperature dependence of C_p was made by fitting the trial functions of Table 3.

Equations 1, 2, and 3 of this fit are the standard forms used to treat solid data. These functions followed the data no better than could be expected from the addition of another degree of freedom. In an attempt to represent an interpolation formula, deviation plots following the form of equations 4 through 7 were investigated. None of these forms followed the data, with the exception of equation 7 with $n = 6$. Table 4 shows the effect of this interpolation formula. It should be emphasized that this formula behaves quite badly beyond the limits of the data, and is to be used only for interpolation purposes.

Table 5 shows the change in emissivity necessary to account for the apparent curvature. This effect, also, is not a simple relationship and the necessity for gathering more data to characterize the effects observed is obvious.

References

1. R. Hultgren, R. L. Orr, P. D. Anderson and K. K. Kelley, Supplement to Selected Values of Thermodynamic Properties of Metals and Alloys, Wiley and Sons, Inc., New York (1970).
2. A. J. Cezairliyan, NBS Report No. 10326, 133 (1971).

TABLE 1

ENTHALPY INCREMENTS FOR LIQUID NIOBIUM

Temperature K	EXPERIMENTAL	$(H_T^\circ - H_{298}^\circ)$ j/mole	CALCULATED	Dev. %
	Mass gm		$(H_T^\circ - H_{298}^\circ)$ j/mole	
2738	1.1363	108379	107740	0.59
2746	1.7781	107232	108064	-0.77
2759	1.6678	108520	108592	-0.07
2802	1.5317	109995	110337	-0.31
2812	1.6401	111865	110743	1.01
2830	1.6965	110690	111474	-0.70
2888	1.8309	113793	113828	0.03
2890	1.8451	114550	113909	0.56
2901	1.0195	114476	114356	-0.11
2986	1.7886	118525	117806	0.61
2989	1.5379	117659	117928	-0.23
3011	1.9678	117767	118821	-0.89
3101	1.7506	122031	122474	-0.36
3141	1.5926	124638	124097	0.44
3292	2.0276	130275	130226	0.04

$H_T^\circ - H_{298}^\circ = C_p (T - T_m) + H(l, T_m)$	Average Deviation	0.45%
$C_p = 40.6 \pm 1.0$ j/mole-K (9.70 \pm 0.24 cal/mole-K)	Standard Deviation	617 j/mole
$H(l, T_m) = 108227 \pm 240$ j/mole	Temperature Range	554 K
$T_m = 2750$ K	Drop Distance	26 cm
$\Delta H_f = 34570 \pm 240$ j/mole	Spectral Emissivity	0.317
$\Delta S_f = 12.57 \pm 0.09$ j/mole (3.00 cal/mole-K)	Total Emissivity	0.33
$\rho = 7.6$ gm/cm ³		
$M = 92.91$ gm/mole		

TABLE 2

ZIRCONIUM LIQUID ENTHALPY INCREMENTS
-LINEAR FIT-

Temperature K	EXPERIMENTAL		CALCULATED		Dev. %
	Mass gm	$(H_T^\circ - H_{298}^\circ)$ j/mole	$(H_T^\circ - H_{298}^\circ)$ j/mole		
X*	2233	0.6190	78416	79347	-1.17
Q	2353	0.3967	84070	84236	-0.20
H	2376	0.4872	85570	85173	0.47
Y*	2425	0.9536	87926	87170	0.87
L	2527	0.7188	91519	91326	0.21
R	2575	0.4290	93648	93281	0.39
O	2670	0.9553	97227	97152	0.08
M	2688	0.9034	97880	97885	-0.01
S	2720	0.8208	99078	99189	-0.11
K	2817	0.9776	102778	103141	-0.34
J	2839	1.0448	103816	104038	-0.21
Average Deviation					0.38%
U	2933	1.1672	109222	107867	1.24
T	2997	1.2750	112657	110475	1.94
V	3048	1.4742	116381	112552	3.29

*Drop distance 22.3 cm, all others 26 cm.

$$H_T^\circ - H_{298}^\circ = C_p (T - T_m) + H(l, T_m)$$

Standard Deviation 436 j/mole

$$C_p = 40.7 \pm 0.7 \text{ j/mole-K}$$

Temperature Range (fit) 606 K

$$(9.74 \pm 0.16 \text{ cal/mole-K})$$

Emissivity (6450 Å) 0.318

$$H(l, T_m) = 75069 \pm 320 \text{ j/mole}$$

Total Emissivity 0.30

$$T_m = 2128 \text{ K}$$

Density 5.84 gm/cm³

$$\Delta H_f = 14652 \pm 320 \text{ j/mole}$$

$$\Delta S_f = 6.89 \pm 0.15 \text{ j/mole-K}$$

(1.65 cal/mole-K)

$$M = 91.22$$

TABLE 3

TRIAL FUNCTIONS FOR NON-LINEAR FIT TO
ZIRCONIUM LIQUID ENTHALPY INCREMENTS

$$H_T - H_{298} = a + b\Delta T + c\Delta T^2 \quad (1)$$

$$H_T - H_{298} = a + b\Delta T + c\Delta T^2 + d\Delta T \quad (2)$$

$$H_T - H_{298} = a + b\Delta T + c\Delta T^2 + d\Delta T^{-1} \quad (3)$$

$$H_T - H_{298} = a' + b\Delta T + c\Delta T^2 \quad (4)$$

$$H_T - H_{298} = a' + b'\Delta T + ce^{k\Delta T} \quad (5)$$

$$H_T - H_{298} = a' + b'\Delta T + c\Delta T^2 \quad (6)$$

$$H_T - H_{298} = a' + b'\Delta T + c\Delta T^N; N = 4, 5, \underline{6}, 8 \quad (7)$$

$$\Delta T = T - T_m$$

Prime indicates value from linear fit.

TABLE 4

ZIRCONIUM LIQUID
INTERPOLATION FORMULA

	<u>T</u> <u>K</u>	<u>(T-T_m)</u>	<u>OBS</u> <u>H_T - H₂₉₈</u> <u>j/mole</u>	<u>CALC</u> <u>H_T - H₂₉₈</u> <u>j/mole</u>	<u>Dev.</u> <u>%</u>	<u>C_P</u> <u>j/m-K</u>
X	2233	105	78416	79347	-1.17	40.7
Q	2353	225	84070	84236	-0.20	40.7
H	2376	248	85570	85173	0.47	40.8
Y	2425	297	87926	87170	0.87	40.8
L	2527	399	91519	91350	0.19	41.1
R	2575	447	93648	93327	0.34	41.3
O	2670	542	97227	97298	0.08	42.6
M	2688	560	97880	98065	-0.19	42.7
S	2720	592	99078	99436	-0.36	43.5
K	2817	689	102778	103760	-0.95	46.4
J	2839	711	103816	104787	-0.93	47.3
U	2933	805	109222	109453	-0.21	52.7
T	2997	869	112657	112981	-0.29	58.2
V	3048	920	116381	116075	0.26	64.0

$$H_T - H_{298} = 75069 \pm 320 + (40.7 \pm 0.7) \times (T - T_m) + (5.82 \pm 0.47 \times 10^{-15}) \times (T - T_m)^2$$

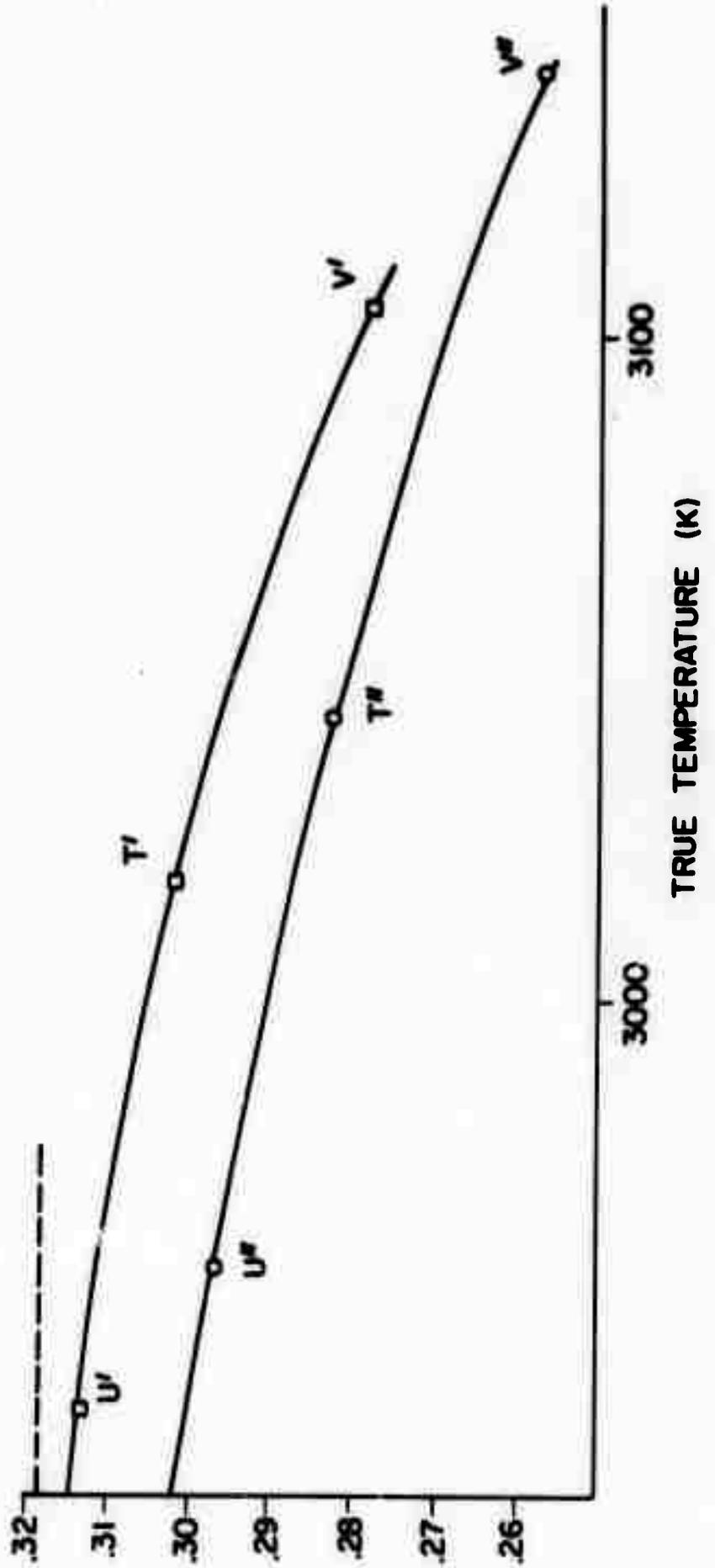
$$C_P = 40.7 (T - T_m) + 3.49 \times 10^{-14} (T - T_m)^2$$

Average Deviation 0.47%

TABLE 5

Emissivities of Zirconium near 3000 K if C_p is constant

POINTS	T^*B	$T_{\epsilon} = .318$	T To Fit Line Fit W/O T, U, V --Fit W/T, U, V	ϵ Calculated Est. Range
U	2529	2933	2960-U [#] 2938-U'	.297--.313
T	2576	2997	3042-T [#] 3018-T'	.284--.302
V	2614	3048	3140-V [#] 3105-V'	.257--.278



PYROLYSIS OF POLYMERS AND
SIMPLE ORGANIC MOLECULES*

J. L. Margrave

Abstract

The techniques of mass spectrometry and of matrix-isolation infrared spectroscopy may be combined to yield meaningful descriptions of the species which are formed in the primary mechanistic steps of pyrolysis. Decomposition of the hydrides of boron, carbon and silicon establish the presence of species like MH, MH₂ and M₂H as well as the expected MH₃, MH₄, M₂H₆, etc. Mass spectrometric characterization then facilitates the interpretation of infrared spectra of such pyrolysis intermediates when they are frozen out on surfaces at liquid helium temperatures.

Recent experimental work at Rice has identified species like Si(CH₃)₂, CCl₂, CCl₃, CCl₂Br, etc. Such studies are needed to allow reliable interpretation of pyrolysis data on hydrocarbon and other more complex polymers.

*Published, please see the following:

1. A. K. Maltsev, O. M. Nefedov, R. H. Hauge, J. L. Margrave and D. Seyferth, J. Phys. Chem. 75, 3984 (1971).
2. A. K. Maltsev, R. G. Mikaelian, O. M. Nefedov, R. H. Hauge and J. L. Margrave, Proc. Natl. Acad. USA 68, 3238 (1971).

Preceding page blank

THE EMISSIVITIES OF LIQUID METALS
AT THEIR FUSION TEMPERATURES

D. W. Bonnell, J. A. Treverton, A. J. Valerga
and J. L. Margrave

Abstract

A survey of the literature through 1969 shows an almost total lack of experimental emissivity data for metals in the liquid state. The emissivities for several transition metals and various other metals and compounds in the liquid state at their fusion temperatures have been determined in this laboratory. The technique used involves electromagnetic levitation-induction heating of the materials in an inert atmosphere. The brightness temperature of the liquid phase of the material is measured as the material is heated through fusion. Given a reliable value of the fusion temperature, which is available for most pure substances, one may readily calculate an emissivity for the liquid phase at the fusion temperature. Even in cases where melting points are poorly known, the brightness temperatures are unique parameters, independent of the temperature scale and measured for a chemically defined system at a fixed point. One may recalculate better emissivities as better melting point data become available.

Preceding page blank

THE EMISSIVITIES OF LIQUID METALS AT THEIR FUSION TEMPERATURES

D. W. Bonnell, J. A. Treverton, A. J. Valerga
and J. L. Margrave

Introduction

The use of brightness temperature as a measure of temperature is subject to an evaluation of the surface emissivity of the substance in question. The ability to generate blackbody conditions in most research applications has in the past made brightness temperature a secondary value, normally used only for monitoring purposes. However, when one deals with refractory metals, and wishes thermodynamic quality data concerning the liquid phase, blackbody conditions are no longer easily obtainable, especially if the container problem is also to be avoided. Application of other methods of non-contact temperature measurement, such as two color pyrometry are less convenient and of dubious advantage when good spectral emissivity data are available.

Interest in emissivity measurements has been sporadic, the primary literature sources of data being a review by Burgess and Waltenberg in 1915¹, and a collection of NBS measurements by Roeser and Wenzel².

The recent advent of direct applications of electromagnetic levitation to drop calorimetry³ and its immediate success have brought the temperature measurement problem to the fore.

The primary source of error for property measurements in the 2000 to 3000 K temperature range is now almost exclusively the lack of accuracy in temperature measurements.

For this research, a direct attempt to measure brightness temperatures at a standard wavelength for a variety of liquid metals was undertaken. If the brightness temperature is monitored as a function of time for a substance being heated during levitation, it is observed that at the point of fusion, the brightness temperature remains almost constant for a relatively long time. Experiments for several substances including copper were run at various heating rates, varying by nearly an order of magnitude, and the final value of emissivity showed no correlation with heating rate in any case. This corresponds well with other observations on substances subjected to constant heating, i.e., the phase change occurs at constant temperature. By using a recording pyrometer and noting the brightness temperature at the end of this plateau, the emissivity for the liquid at its fusion temperature can be calculated immediately from the Wien equation.

The advantages of levitation as a heating device have been pointed out³ previously. In this application the primary advantage is the ability to use a relatively massive sample, corresponding to sizes of surfaces normally under consideration in laboratory-scale experiments. With no container to interfere, the material is unaltered by its surroundings.

Apparatus

The levitation apparatus has been described in detail elsewhere³. The primary addition to the experimental arrangement is the Leeds and Northrup automatic recording pyrometer Model 8641-1, Serial #1720818. The automatic pyrometer operates at an effective wavelength of 6450 Angstroms with a band width of approximately 350 Å. Basic response time is about 1 sec. for 0.5°C resolution at 1063°C. The automatic pyrometer was calibrated by comparison against an L & N Model 8622-C manual pyrometer (Serial #1077349, calibrated by L & N by comparison with NBS Test No. 201571⁴ reference standard, report dated 9 Dec., 1970). Both pyrometers were sighted alternatively on a G.E. 20A pyrometer lamp. Ten calibration points across the medium range of the automatic machine were determined individually by each of two observers and a least squares line fitted to the deviations. The RMS error of the line is less than 2 degrees.

The technique used for measurement was generally the same for all samples. The flask surrounding the levitation coil was flushed for 5 to 10 minutes with pure argon which had been dried by passing through a magnesium perchlorate column and then deoxygenated by passing over a 19" long column of fine copper turnings heated to 350°C. In cases such as copper and nickel where surface coatings were noticed, the materials were cleaned with 1:1 diluted reagent hydrochloric acid, then rinsed with deionized water and dried with acetone. The sample was then immediately suspended in the coil and power applied. The auto-

matic pyrometer was used to follow the passage through fusion, a process which took more than 10 seconds in all cases. Each measurement was made with a fresh sample and every recovered sample showed a surface at least as bright as the material before levitation.

The samples used were supplied in massive form (either 1/4" rod, platelets, or shot). Table I lists sources and purity. In all cases where more than one source was used, there was no difference in results attributable to source.

Results

The results of this investigation are presented in Table I. The uncertainties given are in terms of precision of measurement. The accuracy of calibration of pyrometers in general can contribute an error of $\pm 5 - 7\%$. It is expected that the higher precision of automatic pyrometers will soon allow a much needed improvement in this figure. The fusion temperatures reported are either those of Hultgren, Orr, and Kelley⁵, modified to agree with the International Practical Temperature Scale of 1968⁶ or values which are secondary reference points of IPTS-68. The brightness temperature is related to the emissivity through the Wien equation⁷. At the temperatures involved in this work, the error introduced by this approximation to Plank's law is less than 0.1%. Wien's law gives, for a body of emissivity E_λ , the intensity of radiation at some wavelength λ ,

$$J_\lambda = E_\lambda C_1 \lambda^{-5} e^{-C_2/\lambda T} \quad (1)$$

TABLE I

Substance	Fusion Temp. (K)	Brightness Temp. of Liq. at Fusion (K)	Emissivity ($\lambda=0.645\mu$)	No. of Determinations
Ag ^a (99.999%)	1235 ^f	1085 ± 1	0.082 ± .001	12
Co ^a (99.95%)	1767 ^f	1626 ± 2	0.335 ± .006	6
Cr ^a (99.95%, 99.997%)	2133	1891 ± 3	0.262 ± .005	9
Cu ^b (99.9%)	1357.6 ^f	1216 ± 2	0.147 ± .010	7
Fe ^a (99.95%)	1811	1671 ± 1	0.357 ± .003	3
Mo ^a (99.9%)	2895	2510 ± 4	0.306 ± .004	8
Nb ^{a,c} (99.9%, 99.8%)	2744	2405 ± 2	0.317 ± .002	9
Ni ^d (99.8%)	1728 ^f	1597 ± 2	0.346 ± .005	11
Pd ^a (99.95%)	1827 ^f	1684 ± 2	0.354 ± .005	5
Ta ^a (99.9%)	3256	2779 ± 3	0.309 ± .003	2
Ti ^a (99.95%)	1946	1814 ± 1	0.434 ± .003	16
V ^e (99.9%)	2178	1973 ± 2	0.343 ± .004	10
Zr ^c (99.9%, 99.8%)	2128	1918 ± 2	0.318 ± .004	8

^aA. D. Mackay, Inc.

^bBaker and Adamson Reagent Chemicals

^cAlfa Inorganics, Inc.

^dJ. T. Baker Chemical Co.

^eResearch Organic/Inorganic Chemical Corp.

^fSecondary reference points on the IPTS-68. The other fusion temperatures are taken from reference 4 by Hultgren, Orr and Kelley and are adjusted to the IPTS-68.

The relation between brightness temperature, T_B , and true temperature T is then

$$\frac{1}{T} - \frac{1}{T_B} = \frac{\lambda \ln E_\lambda}{C_2} \quad (2)$$

where C_2 is 1.4388 cm K (IPTS-68) and λ is 6.45×10^{-5} cm. If (2) is solved for E_λ ,

$$E_\lambda = \text{EXP} \left(\frac{C_2}{\lambda} \left[\frac{1}{T} - \frac{1}{T_B} \right] \right) \quad (3)$$

is obtained. Taking differentials of both sides yields

$$dE_\lambda = \left\{ \text{EXP} \left(\frac{C_2}{\lambda} \left[\frac{1}{T} - \frac{1}{T_B} \right] \right) \right\} \left\{ -\frac{C_2}{\lambda} \right\} \left\{ -\frac{1}{T_B^2} \right\} dT_B \quad (4)$$

Dividing (4) by (3) gives

$$\frac{dE_\lambda}{E_\lambda} = \frac{C_2}{\lambda T_B} \left(\frac{dT_B}{T_B} \right) \quad (5)$$

where $C_2/\lambda = 22310$ K.

Use of equation (5) shows the extreme sensitivity of emissivity values to small errors in brightness temperatures. For example, in the neighborhood of 2000 K, if the brightness temperature is in error by 10 K (a typical value), and if the nominal emissivity is 0.30, $dE_\lambda = .017$, an error of more than 5%. Thus, the opposite side of the coin to temperature measurement being relatively insensitive to errors in emissivity is

that emissivity values derived from temperature measurements are very sensitive to inaccuracies in brightness temperature.

Discussion

These data were specifically produced at a known calibration point in order to provide stable reference values independent of the normal problems of temperature stability. For this reason, the reported brightness temperature is the fundamental measurement and is specific to the fusion temperature of the metal under one atmosphere of inert gas.

Reported errors in brightness temperature correspond to less than 0.2 mv deviation in the measured variable recorded on a L & N Speedomax H 6½" chart recorder equipped with Azar universal range and span device. The span used for all measurements was 2 mv full scale and selected points across all ranges at this span were calibrated against a L & N Type K-3 potentiometer. Calibration corrections were of the order of .07 - .10 mv in the ranges of interest and were applied.

The results of this research are compared with literature values in Table II. Agreement is reasonable with few exceptions. The most notable of these exceptions are titanium and chromium. Reference to oxide emissivity data of Roeser and Wenzel⁹ shows the emissivities of the oxides of these metals to be considerably higher than the clean surface. Titanium is very reactive at higher temperatures, showing an ability to oxidize even in the presence of trace amounts of carbon dioxide. Observations of

TABLE II

Substance	Emissivity	
	This Research	Literature
Ag	0.082 ± .001	0.072 ^a
Co	0.335 ± .006	0.37 ^a
Cr	0.262 ± .005	0.39 ^a
Cu	0.147 ± .010	0.150 ^a ; 0.148 ^b
Fe	0.357 ± .003	0.365 ^a
Mo	0.306 ± .004	0.40 ^a
Nb	0.317 ± .002	0.40 ^a
Ni	0.346 ± .005	0.37 ^a
Pd	0.354 ± .005	0.37 ^a
Ta	0.309 ± .003	0.398 ^c
Ti	0.434 ± .003	0.65 ^a
V	0.343 ± .004	0.32 ^a
Zr	0.318 ± .004	0.30 ^a

^aReference 2

^bReference 11

^cReference 10

chromium after levitation melting and dropping on a cold plate showed the bulk material to discolor easily in open air, presumably oxidation. The conclusion to be drawn is that probably the earlier measurements were of a surface slightly contaminated with oxide, resulting in a higher brightness temperature.

The emissivity value from Treverton and Margrave⁸ for molybdenum has changed because recalibration of the automatic pyrometer showed a slight but significant deviation from agreement with IPTS-68.

Measurements of niobium and tantalum are at temperatures high enough that questions of equilibrium for the melting point and actual true measured values at other temperatures must be a factor in other measurements (1, 2, 10). The levitation technique allows a reasonable time for equilibration at the melting point.

It is to be noted that for almost all the materials reported on here, no sign of change in brightness temperature was observed during the melting process. The single exception was copper. Even in the case of copper, however, the observed effect was not a change in emissivity from .10 to .15, as reported earlier. The current observations showed the surface to superheat slightly, settling back smoothly to a lower brightness temperature which coincided with the visual observation of the entire sample collapsing into the shape forced by the surrounding RF field. For this reason, the stated deviation in precision was reported as twice the calculated standard deviation.

Conclusions

Application of the levitation phenomenon to the measurement of surface brightness shows distinct advantages and monochromatic emissivities for several liquid metals have been determined at their respective melting points with a 450 KHz generator and levitation coil. The possible variation of emissivity as a function of phase is still subjected to question. Continuing work at various levitation frequencies where the skin depth is much larger is in progress.

Acknowledgements

This work was supported by the National Aeronautics and Space Administration.

References

1. G. K. Burgess and R. G. Waltenberg, Bull. Bur. Std. 11, 591.
2. W. F. Roeser and H. T. Wenzel, Temperature, Its Measurement and Control in Science and Industry, Appendix, Table 16, Reinhold Publ. (1941).
3. A. K. Chaudhuri, D. W. Bonnell, A. L. Ford and J. L. Margrave, High Temp. Sci. 2, 203 (1970).
4. Leeds and Northrup calibration on IPTS-48 was corrected to IPTS-68.
5. R. Maltgren, R. L. Orr, K. K. Kelley, Supplement to Selected Values of Thermodynamic Properties of Metals and Alloys (1968 to 1970).
6. The International Practical Temperature Scale of 1968. Comptes Rendus de la 13^{me} Conference Generale des Poids et Mesures 1967-1968, Annex 2. Metrologia 1969, 5, 35-44.
7. W. D. Kingery, Property Measurements at High Temperatures, John Wiley and Sons, Inc., N. Y. (1959) p. 12 ff.
8. J. A. Treverton and J. L. Margrave, "Thermodynamic Properties of Liquid Molybdenum by Levitation Calorimetry," ASTM Conf. on Thermophysical Properties, Boston, Mass., Oct., 1970.
9. W. F. Roeser and H. T. Wenzel, Temperature, Its Measurement and Control in Science and Industry, Appendix, Table 17, Reinhold Publ. (1941).
10. Ared Cezairliyan, NBS Report No. 10326, p. 133.
11. G. G. Gubareff, J. E. Janssen, and R. H. Horborg, Thermal Radiation Properties Survey (Minneapolis-Honeywell Regulator Co., Minneapolis, 1960).

HEATS OF COMBUSTION OF CARBONS AND GRAPHITES

J. L. Wood and J. L. Margrave

Abstract

A measure of the extent of conversion to "ideal" graphite can be obtained by high-precision combustion calorimetry on various samples of vitreous carbon, pyrographites, etc. It appears certain that a new standard reference state graphite has been identified and that the heat of formation of CO₂ (ideal gas) should be

$$\Delta H_f = -94,042.8 \pm 5.9 \text{ cal/mole}$$

Samples of various carbons are being sought from Professor E. Hücke and from other sources for combustion in a high-precision calorimeter, in order to place each on a relative energy scale.

THE Li/CFX BATTERY AND ITS CHARACTERISTICS

R. B. Badachhape, J. L. Wood, A. J. Valerga
and J. L. Margrave

Abstract

Current approaches to the construction, characteristics and other aspects of the Li/CFX battery system are being reviewed. Under a new contract with the U. S. Signal Corps at Fort Monmouth, New Jersey, Professor Margrave and his associates are determining basic thermodynamic parameters for CFX of various stoichiometries. New techniques for forming rugged electrodes are also being investigated.

Preceding page blank

STRESS WAVES DUE TO A SHORT DURATION PRESSURE PULSE
ON A SEMI-INFINITE BODY OF LAYERED COMPOSITE

J. A. Krumhansl and E. H. Lee

Abstract

The Fourier-Floquet theory for wave propagation in a continuum of composite material with a periodic structure has been developed to represent transient solutions. For a layered composite half-space with displacement and wave propagation normal to the layers, waves generated by impact, or pressure applied on the surface for a limited duration, can be treated by the same theory. The initial motion is analyzed by following the reflection and transmission of the resulting plane waves, for example, by the method of characteristics, until the force or contact ceases; and this solution then provides initial conditions for a Fourier-Floquet expansion for the subsequent motion. This form may be the most convenient for investigating the far field solution for "head of the wave" contributions. Waves generated by plate slap can be investigated in this way.

Preceding page blank

STRESS WAVES DUE TO A SHORT DURATION PRESSURE PULSE
ON A SEMI-INFINITE BODY OF LAYERED COMPOSITE

J. A. Krumhansl and E. H. Lee

Introduction

It has been pointed out by Krumhansl¹ that Floquet solutions² for waves in infinitely extended periodic elastic composites form a complete orthonormal function set, which can be utilized to evaluate the propagation of transients by means of Fourier-Floquet integrals. The case of an impulsive delta-function initial velocity distribution was presented¹, where the impulse occurred on a plane of symmetry of a one-dimensional lattice of equally spaced reinforcing plates, the displacement and propagation direction being normal to the plates. This formulation permits asymptotic evaluation of the far field comprising "head of the wave" components which yield the major contributions of the wave effect transmitted through the composite. Since transient loading of a composite commonly occurs on an external surface, we exploit the fact that waves generated by such a time dependent short duration pressure distribution on a surface can be analysed within the framework of the infinitely extended body theory¹ when the surface is a plane of symmetry of the lattice. This is achieved by first considering plane wave propagation with reflection and transmission at the matrix-

reinforcement interfaces for the half space, say, $x \geq 0$. This is easy to evaluate for the short duration of load application, but would be cumbersome for long times because of the multiplicity of reflections and transmissions which continue to occur. If a tensile variation of surface traction, the negative of the pressure pulse, is applied to the half-space $x \leq 0$, the surface velocity at any time will be the same for both half-spaces by symmetry and the linearity of the problem. Thus the two half-space surfaces will move with the same velocity and displacement. They can thus be considered to be in contact and welded together without modifying the stress wave distribution in either, so that the displacement and velocity at any instant after pressure cessation can be used as initial values for the full space, in order to evaluate the subsequent motion. Since surface forces on the half-spaces are then no longer acting, no body forces at the weld arise in the combined problem, so that the theory of¹ can be applied.

Because these "initial" distributions of stress and velocity are finite in magnitude and spread over one or more cells of the composite, the Fourier-Floquet integral is likely to exhibit practical convergence, the higher modes not being too highly stimulated. This form of the solution will permit study of the far field by asymptotic methods, with the determination of "head of the wave" contributions. Although these effects would eventually appear in the reflection - transmission procedure mentioned above, they are likely to lag far behind the precursor

and thus involve so many interactions that the Fourier-Floquet approach is likely to be advantageous in determining far field response. It is perhaps worth pointing out that use of the delta-function initial velocity solution to build up a transient solution by super-position, in the usual manner, is not applicable in its simplest form in the present case, since the basic solution depends on the location of the delta-function spike relative to the lattice configuration and is thus not translationally invariant.

Wave Transmission and Reflection Solution

Consider a pressure variation $p(t)$, $t \geq 0$, on the surface of a semi-infinite block, $x \geq 0$, of layered composite, as shown in Fig. 1. The matrix and reinforcement are considered each to be homogeneous linear elastic materials, and hence waves transmitted through them satisfy the wave equation:

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c_i^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad (1)$$

where c_i , the wave velocity, is equal to c_m in the matrix and c_f in the reinforcement, and is related to the elastic modulus for waves of one-dimensional strain: $\eta_i = \lambda_i + 2\mu_i$ and density ρ_i , by the relation:

$$\rho_i c_i^2 = \eta_i \quad (2)$$

for i equal to m or f in the two materials, respectively. The

λ_i and μ_i are the Lamé elastic constants. Perhaps the most convenient way of evaluating the waves generated by the applied pressure $p(t)$ in Fig. 1 is to utilize the theory of characteristics (see Ref. 3, p. 119, for example) which determines the relations

$$\rho_i c_i v \pm \sigma = \text{constant} \quad (3)$$

on

$$\frac{dx}{dt} = \pm c_i \quad (4)$$

where v is the particle velocity $\partial u / \partial t$ and σ is the stress. Thus for the initially undisturbed composite block, Fig. 1, with a layer of matrix material of thickness $(a - b)/2$ at the surface (since the surface is considered to be on the centerline of a matrix slab, which is of width $[a - b]$) the stress field for $t < (a - b)/2c_m$ is given by

$$\begin{aligned} \sigma(x, t) &= -p(t - x/c_m), & x < c_m t \\ &= 0, & x > c_m t \end{aligned} \quad (5)$$

since tensile stress is considered positive, and the velocity by

$$v(x, t) = -\sigma(x, t) / \rho_m c_m \quad (6)$$

These results follow from (3), since the constant for the lower sign is zero because the body is initially undisturbed before the pressure is applied, and thus both v and σ are zero throughout the half-space at $t = 0$.

For the half-space $x \leq 0$ and tensile surface traction of magnitude $p(t)$, the dependent variables are related to the above by evenness as a function of x for velocity v and displacement u , and oddness for the stress.

Suppose, for example, that $p(t)$ has the constant value p_0 , applied at $t = 0$ and removed at $t = t_1 < (a - b)/2c_m$. Then for $t = t_1$, the stress displacement and velocity in the full space are as illustrated in Fig. 2, and these could then be considered as "initial" conditions for the whole space, which will determine the motion of the half-space $x \geq 0$ with the pressure applied to its surface.

If the pressure application continues for $t > (a - b)/2c_m$, the stress wave reaches the first interface at $x = (a - b)/2$ while the pressure is still acting. Reflection and transmission at this interface must therefore be included in the analysis of wave propagation which must be extended through the time until the surface pressure ceases. We must therefore add to the characteristic relations (3) and (4) the requirement that the displacement and traction must be continuous across each interface, thus

$$u[(a/2 \pm b/2 \pm na)^+, t] = u[(a/2 \pm b/2 \pm na)^-, t] \quad (7)$$

$$\sigma[(a/2 \pm b/2 \pm na)^+, t] = \sigma[(a/2 \pm b/2 \pm na)^-, t] \quad (8)$$

where

$$\sigma(x, t) = \eta_1 \epsilon(x, t) = \eta_1 \partial u / \partial x \quad (9)$$

and n is an integer. Since the arguments in (7) and (8) are

interface positions (see Fig. 1), η_i takes on the values η_f and η_m on either the plus or minus sides of the interface. In (9), ϵ denotes the strain.

Combination of (3), (4), (7) and (8) permits the solution to be extended indefinitely, but the multiple reflections and transmissions at the interfaces associated with repeated application of (7) and (8) with $n = \pm 1, \pm 2, \text{etc.}$, can become cumbersome as t increases. For short time loading, no difficulty arises. A computer code based on such a procedure is mentioned by Lundergan and Drumheller⁴.

As described in the Introduction, as soon as the applied pressure has ceased, this solution can be used for initial values for displacement and velocity to express the subsequent motion in Floquet form¹. This formulation is presented briefly in the next section.

There is no need to limit the characteristic solution presented in this section to prescribed surface pressure. The solution for prescribed velocity can be obtained from (3) and (4) in a similar manner, but after some prescribed time the surface must be permitted to move freely without force applied, so that the problem for the two combined half spaces is free from body-force for the Floquet representation part of the analysis. Moreover, impact due to plate slap can also be handled in this way. The impact surface, $x = 0$, would become an additional interface if the impacting plate were not of matrix material, otherwise it would simply be a surface incapable of transmitting tensile stress.

The characteristic solution with free rear surface of the impacting plate separates from the surface of the half-space. This time could be used to provide initial conditions for the two half-space problem, the solution of which can be expressed in Floquet form for the subsequent motion. Thus this approach can be applied to a variety of loading situations.

Fourier-Floquet Representation

As shown¹, Floquet wave solutions form an ortho-normal set according to

$$\int_{-\infty}^{\infty} \rho(x) u_{\nu}^*(x, q) u_{\nu'}(x, q') dx = \delta_{\nu\nu'} \delta(q - q') \quad (10)$$

where q and q' are wave numbers and ν and ν' denote integer labels for the denumerable sequence of modes. The solution of (1) and the interface conditions (7) and (8) for prescribed initial values of displacement, $u(x, 0)$, and velocity, $\dot{u}(x, 0)$, for the extended composite $-\infty < x < \infty$, can be represented in the form:

$$u(x, t) = \sum_{\nu} \int_{-\pi/a}^{\pi/a} [c_{\nu}(+, q) e^{i\omega_{\nu}(q)t} + c_{\nu}(-, q) e^{-i\omega_{\nu}(q)t}] u_{\nu}(x; q) dq \quad (11)$$

The coefficients are given by the usual inverse relations:

$$\int_{-\infty}^{\infty} \rho(x) u_{\nu}^*(s; q) u(x, 0) dx = c_{\nu}(+, q) + c_{\nu}(-, q) \quad (12)$$

$$\int_{-\infty}^{\infty} \rho(x) u_{\nu}^*(x;q) \dot{u}(x,0) dx = i\omega_{\nu}(q) [c_{\nu}(+,q) - c_{\nu}(-,q)] \quad (13)$$

Note that for convenience a new time origin has been selected, and the "initial" displacement and velocity distributions correspond to $t = t_1$ of the characteristic solution, which is the time of cessation of surface pressure or a later time. Note also that $u(x,t)$ of (11) is complex, and the physical displacement is given by the real part. Similarly for other dependent variables. The functions $u_{\nu}(x,q)$ in (10) are the Floquet solutions, which have a real part even in x , and an imaginary part odd in x ⁵. The asterisk in (10) denotes the complex conjugate. Since $u(x,0)$ and $\dot{u}(x,0)$ are even in x , (12) and (13) determine that $c_{\nu}(+,q) + c_{\nu}(-,q)$ is real, and $c_{\nu}(+,q) - c_{\nu}(-,q)$ is pure imaginary, hence

$$c_{\nu}(-,q) = c_{\nu}^*(+,q) \quad (14)$$

Thus the two terms, in the square brackets in (11) are complex conjugates, hence their sum is real. Since the real part of $u_{\nu}(x;q)$ is even⁵, its derivative with respect to x , the corresponding strain, is odd, and hence will be zero at $x = 0$, since strain is continuous there. The stress given by (9) will thus also be zero for $x = 0$, as we expect it to be, since for $x \geq 0$, the solution is to apply for the half-plane after the surface pressure has ceased to act.

As shown¹, the form of solution (11) lends itself to asymptotic analysis by stationary phase or steepest descent

methods, and so is convenient for studying the response to the surface impact in the far field.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253.

References

1. Krumhansl, J. A., "Propagation of Transients in Periodic Composites", *Prelim. Reports, Memoranda and Technical Notes of ARPA Materials Summer Conference*, Dept. Chem. and Metall. Engrg., University of Michigan, pp. 176-195, July 1970.
2. Kohn, W, Krumhansl, J. A. and Lee, E. H., "Variational Methods for Dispersion Relations and Elastic Properties of Composite Materials", ASME Reprint No. 71-APMW-21, 1971, to appear in *J. Appl. Mech.*
3. Lee, E. H., "The Continuum Mechanics Aspect of Material Properties Determination", *Energetics*, Vol. III, pp. 85-122, 1967.
4. Lundergan, C. D. and Drumheller, D. S., "Propagation of Stress Waves in a Laminated Plate Composite", *J. Appl. Phys.*, Vol. 42, pp. 669-675, 1971.
5. Lee, E. H., "Stress Distributions in Plane Waves in a Periodic Layered Composite", *Prelim. Reports, Memoranda and Technical Notes of ARPA Materials Summer Conference*, Dept. Chem. and Metall. Engrg., University of Michigan, pp. 153-173, July 1970.

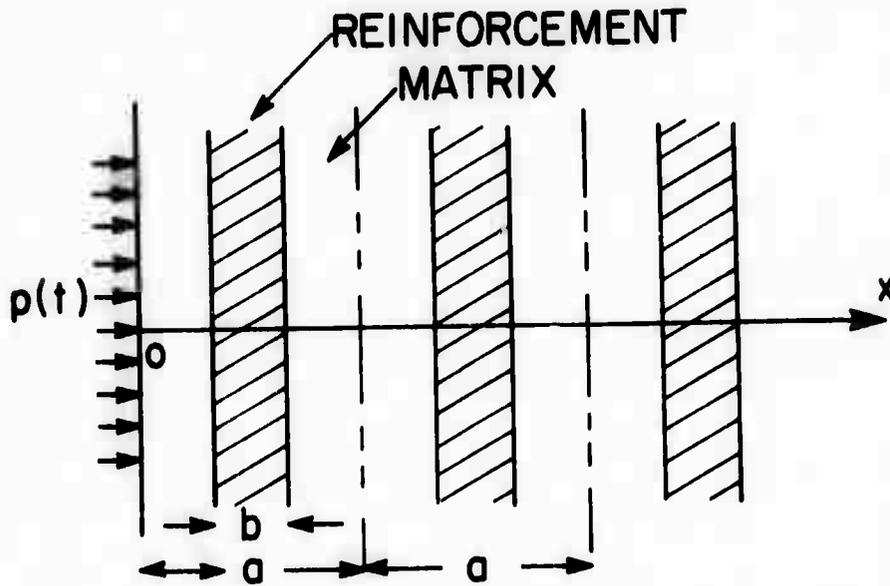


Figure 1. Composite Half-Space with Applied Surface Pressure

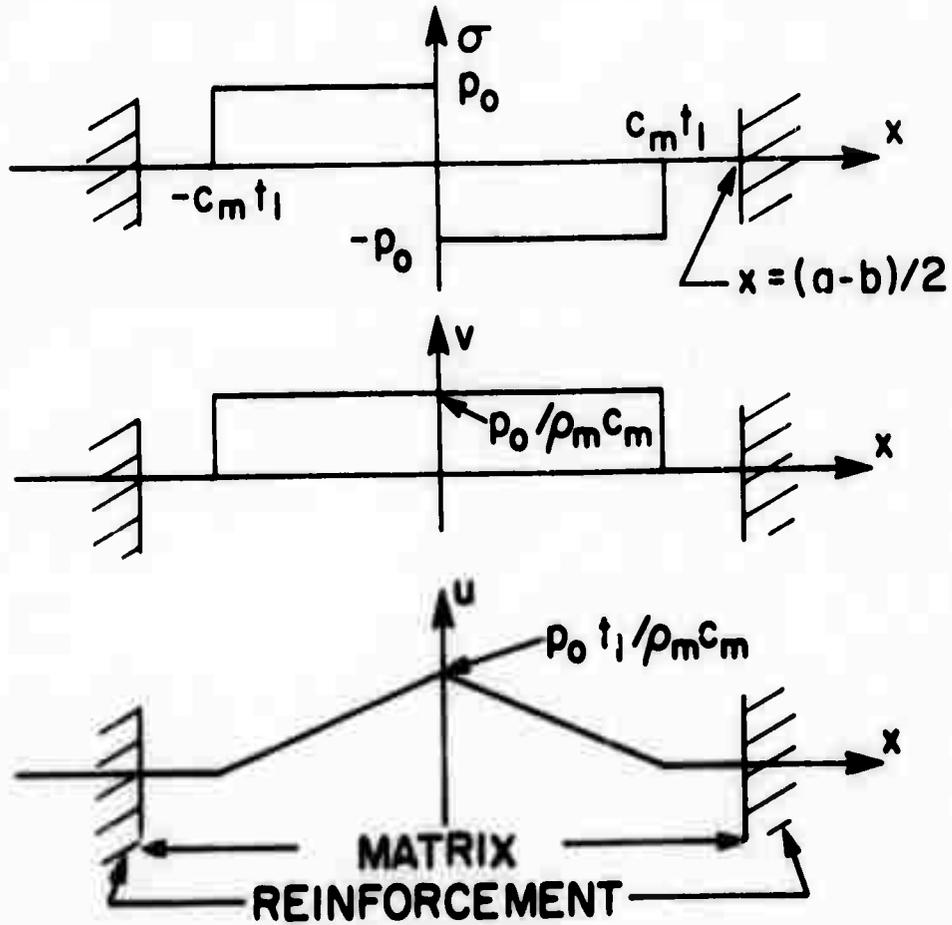


Figure 2. Stress, Velocity & Displacement for Constant Applied Pressure, p_0 , until $t = t_1 < (a-b)/2c_m$

DETERMINATION OF STRESS PROFILES FOR WAVES
IN PERIODIC COMPOSITES

L. Bevilacqua, W. Kohn, J. A. Krumhansl
and E. H. Lee

Abstract

Floquet or Bloch wave theory provides a convenient basic set of functions for representation of the propagation of transient elastic stress waves in periodic composites (Krumhansl, ARPA Mat. Res. Council Report, p. 175, 1970). Variational principles for computing dispersion relations and hence phase velocities generate a band structure of pass and no-pass frequency bands (Kohn, Krumhansl and Lee, ARPA Mat. Res. Council Report, Vol. I, Paper No. 2, 1969 and ASME Preprint 71-APMW-21, to appear in Jour. Appl. Mech.) Dispersion curves (frequency versus wave number) were accurately evaluated for laminar composites by using smooth Fourier series test functions for displacement in a Rayleigh-Ritz approximation procedure, but the corresponding stress profiles were unsatisfactory since the required continuity of stress at the inclusion-matrix interface was ruled out by the use of the smooth test functions for displacement and corresponding continuous strain profiles.

In this paper exact stress profiles are calculated for waves propagated normally to the laminae, and satisfactory approx-

Preceding page blank

imations to these are generated with the extended variational principle which permits independent test functions to be used in the matrix and inclusions. An augmented plane wave approach in the Rayleigh-Ritz procedure was adopted in which exact solutions of the wave equation were used as component test functions in the filament. It was found that with this procedure, accuracy was essentially independent of the ratio of elastic moduli of the inclusion and matrix. The variational approach is applicable to two and three-dimensional composite configurations, which are not amenable to exact evaluation.

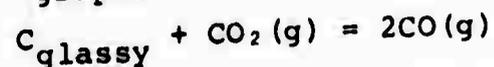
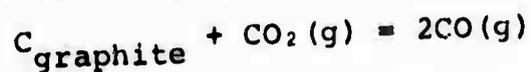
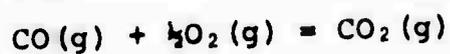
A PROPOSED METHOD FOR THE EVALUATION
OF THE THERMODYNAMIC PROPERTIES OF THE
GLASSY CARBON-GRAPHITE EQUILIBRIUM

E. E. Hucke and S. K. Das

Abstract

The available thermodynamic data for the graphite-carbon oxide equilibrium are critically reviewed, together with the meager data presently existing for heat capacity and heat of combustion of glassy carbon.

Five different, experimentally feasible, solid oxide cells and one molten salt cell are proposed to measure directly and more accurately the equilibrium thermo-chemical properties of the following reactions in the range of 600-1000°C.



The above data can be combined with low temperature heat capacity data for glassy carbon and graphite to yield configuration entropy and enthalpy values of glassy carbon relative to graphite. The configuration entropy value is a direct measure

of the degree of disorder and can be used to compare with structural models of glassy carbon deduced from other physical measurements such as X-ray and neutron diffraction.

I. INTRODUCTION

The most fundamental characteristic of any material is its structure. Once the structure is precisely established, all the known properties of that material can be explained and the unknown ones can be estimated. Whenever a new material is developed, it is first subjected to a series of physical and mechanical tests. These data are useful, but they do not give much insight into the basic structure of the material. The use of X-ray or other probes yield much structural information, but not enough to explain or estimate every property of that material. On the other hand, thermodynamic data can explain many of the gross properties of a material without utilizing detailed information. Precise thermodynamic data is, therefore, a basic tool in understanding, processing, and utilization of a material.

A new form of pure carbon has recently been prepared which, because of its similarity in appearance to glass, has been called "glassy" carbon. The glassy carbon is intermediate between glass and ceramic in many respects and can be adequately described as a conducting ceramic. At the time when glassy carbon was developed, its gas-impermeability at high temperature met a timely demand of the atomic energy industry. Now, however, it has a variety of new applications including outstanding promise as a biomaterial.

Only very sketchy and conflicting reports have been

published on the atomic structure of the glassy carbon. Although its heat of combustion and heat capacity at low temperatures have been measured, these data are not sufficient to calculate the thermodynamic properties of its transition into graphite. So far, no direct measurements of this transition have been published.

At present, there are no satisfactory criteria to characterize samples of glassy carbon. Activity of carbon in glassy carbon may be taken as one of the consistent criteria which makes one sample have different properties than others. The data on activity would, therefore, lead to a better processing, property control, and further systematic development of this material to suit wide-spread applications. Any attempt to measure the standard Gibb's free energy change of glassy carbon-graphite transition would be very significant and the outcome would be a contribution to this field. The derived entropy and enthalpy change would give a long awaited clue for a better understanding of the structure, hence, all other properties of this rather unusual material.

The objective of the proposed research will be the determination of activity of carbon in glassy carbon with respect to graphite as a standard state in the temperature range of 600-1000°C. If the high temperature heat capacity values of glassy carbon become available in the future, the entropy difference values to be deduced from the outcome of the proposed research could be used to compute the entropy of

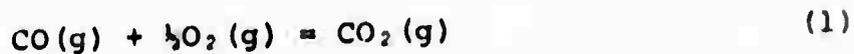
glassy carbon at absolute zero, and to tell whether glassy carbon is a perfect crystalline material in its internal equilibrium or not.

Probably the most direct method of measuring activity of carbon would be to accurately analyze the CO-CO₂ gas mixture in equilibrium with glassy carbon by a mass spectrometer. The only objection which outrules this method is the problem of getting the equilibrium gas samples out of the glassy carbon-CO-CO₂ system. It is anticipated that the gas samples would be subjected to thermal diffusion and convection phenomena which would ultimately result in unrepresentative gas samples.

In the proposed research, two electrochemical methods are outlined. One method involves solid electrolytes and the other a fused salt electrolyte. Both electrolytes are capable of giving accurate activity data. The solid electrolyte cells would measure the partial pressure of oxygen in equilibrium with glassy carbon and carbon oxides, which would then be related to activity of carbon in glassy carbon with respect to graphite as a standard for carbon. The fused salt electrolyte cells would directly yield activity data. A cell, once assembled, can be studied over the entire temperature range of interest. Furthermore, the attainment of equilibrium at each temperature is directly checked on each cell. The precision of the recorded emf and temperature can be related to the uncertainty in activity data.

It will become clear later in the text that the appli-

cation of the solid electrolyte method would need the standard free energy change of reactions 1 and 2.



The best available data on the standard free energy change of reactions 1 and 2 have been indirectly computed using calorimetric and spectroscopic data. The directly measured data by CO/CO₂ gas equilibration method are either less accurate than indirectly calculated values or have only historical importance. In view of the great importance of the equilibrium constants of reactions 1 and 2 in the problems of process metallurgy, it is a challenge to determine the values of the reaction constants with a more refined technique and to reach or even to exceed the accuracy deduced from calorimetric data. Solid electrolyte cells are capable of accurately measuring these equilibrium constants and it is anticipated that the proposed measurements will be a significant contribution to the thermodynamic literature.

II. LITERATURE REVIEW

A. Carbon Oxides, Graphite, and Glassy Carbon

1. Carbon Oxides

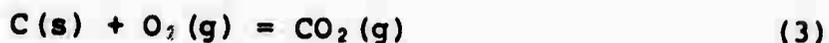
The first study on the thermodynamics of the producer gas reaction was done by Boudouard¹ in 1901. He measured the equilibrium CO-CO₂ gas composition of reaction 2 as a function of temperature at only atmospheric pressure. Shortly after Boudouard's classic study, Rhead and Wheeler^{2,3}, Jellinek and Diethelm⁴, and Dent and Cobb⁵ remeasured the equilibrium CO-CO₂ gas composition as functions of both temperature (700-1100°C) and pressure (0.5-4 atms.). They passed CO₂ or CO gas over purified wood charcoal powder and left the system at a constant temperature and pressure for a long time to achieve equilibrium. Samples of gases were drawn from the system at regular intervals of time and the constant gas composition over a long period of time was used to calculate the equilibrium constant of reaction 2. Their data are in accordance with the Le Chatelier's principle, e.g., a decrease in the total pressure and an increase in the temperature increases the proportion of carbon monoxide in the C-CO-CO₂ system in equilibrium at a constant temperature and pressure respectively.

These and certain indirect studies were reviewed by Chipman⁶ to obtain an equation representing the average experimental results. The position of the equilibrium line (Log K vs. 1/T) was determined in large measure by the 1911 results of

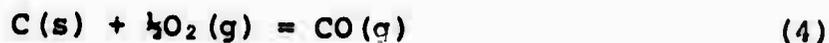
Rhead and Wheeler³, an unfortunate choice since there is reason to believe that their 1910 data are more dependable. The experimental results in the range 800-1000°C were plotted. Below 800°C the reaction is so slow that the attainment of equilibrium is doubtful; above 1000°C the concentration of CO₂ is below the range of accurate analysis. The principle sources of error include non-uniformity of temperature and inaccuracies in its measurement, and limitations in accuracy of gas analysis. Thermal diffusion, if present, would yield a too-low value of the equilibrium constant. Deposition of carbon from the effluent gas sample would likewise lower the value of the equilibrium constant.

In the 1910 series of Rhead and Wheeler² the thermocouple was embedded in the carbon specimen, whereas in the 1911 series it was outside the reaction vessel in contact with the furnace tube. They pointed out that in this second series the carbon surface may have been at lower temperature than measured. The furnace used in this study was too short to provide a satisfactory zone of constant temperature. The thermocouple measured the temperature at the hottest point and it seems altogether possible that parts of the reaction bulb were more than 20° lower in temperature. To sum up, even the best available data of Rhead and Wheeler^{2, 3} may be considered to be of historical importance only.

In 1940, the calorimetric data on the heat of combustion of graphites (Reaction 3) were measured by many investigators⁷.



These values are in excellent agreement with all recently published figures⁶. The heat of reactions 2 and 4 were calculated using directly measured values of heat of reactions 1⁷ and 3⁹.



Utilizing the available entropy and enthalpy values of oxygen, carbon monoxide, carbon dioxide (determined from statistical calculations, spectroscopic, molecular, and heat capacity data), and graphite (deduced from the calorimetric measurements of the heat capacity and the concept of the third law of thermodynamics), the standard Gibb's free energy change of reactions 1-4 have been carefully computed and compiled many times by many authors. However, every author lists his own estimate of the accuracy in the computed free energy values. The estimated accuracy varies from ± 20 cal/mole to ± 1 Kcal/mole. All the heat of combustion and free energy values have been critically examined in Appendix I. The estimate of ΔG_{1000}° values according to the present calculation are $-46,770 \pm 119$, $-1,089 \pm 338$, $-94,629 \pm 101$, $-47,859 \pm 220$ cal/mole for reactions 1-4, respectively.

Richardson and Dennis¹⁰, Smith¹⁷, and Chipman and Rist¹¹ performed many experiments to measure carbon activities in liquid and solid Fe-C alloys, which were then extended by Chipman and Fuwa¹² to numerous ternary alloys. In all the experiments

the same method of CO-CO₂ gas equilibrium was used. In calculating the carbon activities in iron alloy, with respect to graphite as a standard state, one requires the equilibrium constant of the producer gas reaction. When Richardson¹⁰ and Chipman^{11,12,16} were posed with this problem, they preferred to use the equilibrium constant available from indirect calculations rather than either using Rhead and Wheeler's^{2,3} value or re-measuring the equilibrium constant themselves. Probably they realized that the reason which prevented Rhead and Wheeler^{2,3} from getting accurate values of the equilibrium constant of the producer gas reaction was not only the lack of careful experimentation, but also the inherent limitation of the gas equilibrium techniques. In one of his recent articles, Wagner¹³ has clearly stressed the need of determining the equilibrium constant values of the producer gas reaction directly with a more refined technique and to reach or even to exceed the accuracy of the equilibrium constant values deduced from calorimetric and spectroscopic data.

2. Graphite

The hexagonal crystal structure of graphite, as first proposed by Bernal³¹ in 1924, is shown in Figure 1. It consists of carbon atoms arranged as hexagons in flat parallel sheets, the distance between adjacent sheets (which are held together by van der Waal's forces) is approximately 3.35 Å°, while that between atoms within a sheet (covalent bonded) is about 1.42 Å°. The layers are packed so that half of the carbon atoms in one

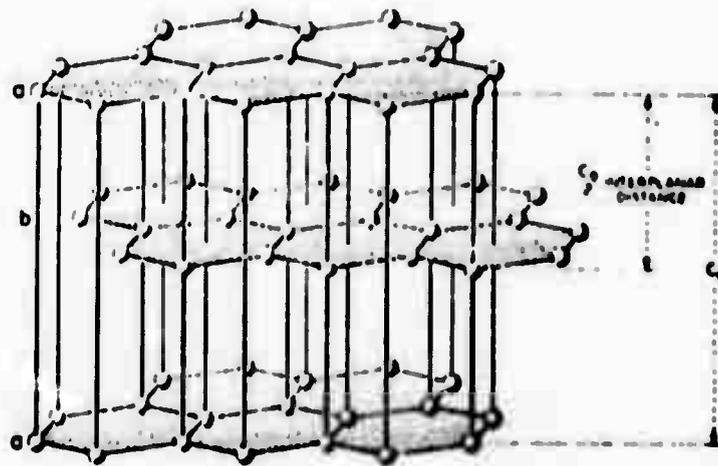


Figure 1. Structure of the hexagonal form of graphite

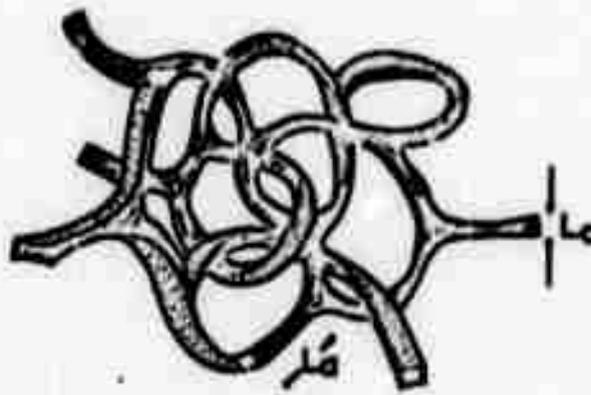


Figure 2. Structural Model for Glassy Carbon

layer lie over the mid-points of the hexagons in the next layer, the atoms in the third layer being situated directly above those in the first layer. This is the abab arrangement or hexagonal structure. Another structure, in which the position of the third layer with respect to the second layer is the same as that of the second with respect to the first, the abcabc arrangement or rhombohedral structure, occurs to the extent of a few percents in most samples of graphite. Graphite is noted for its highly anisotropic characteristics.

The physical, mechanical, and thermodynamic properties of graphite have been measured, calculated, and compiled many times by several authors.^{14,15,17} Its heat capacity in the range 1 to 20°K has recently been the object of intensive theoretical and experimental investigations. In recent JANAF Tables¹⁵, the low temperature C_p measurements of Keenan and Pearlman¹⁶ (1° to 4°K and 10° to 20°K) and of DeSorbo and Nichols¹⁷ (1° to 20°K) were joined smoothly with the C_p measurements of DeSorbo and Tyler¹⁵. C_p values above 100°K were taken from National Bureau of Standards Report 6928. Heat capacity value above 1000°K were adjusted to give smooth results. C_p values above 4000°K have been estimated.

3. Glassy Carbon

Carbon exists in many forms, but of these only two, diamond and graphite, can be completely characterized, and these are true allotropes; both occurring naturally. A wide range of artificial bulk carbons has been prepared for indus-

trial use but these materials are basically graphitic in structure or may be graphitized readily, and their properties (e.g., degree of graphitization) can be controlled to some extent to suit particular application.

In the past ten years, it has been found that several cross-linked polymers can be carefully decomposed so as to yield relatively massive shapes of a highly disordered non-graphitizing carbon that has been called "glassy", "glass-like", or "vitreous".^{12, 17}

It is a very pure form of carbon with impurity content of less than 200 ppm. Its name stems from physical appearance and not from a detailed knowledge of the structure. The physical appearance of these materials resembles black glass, particularly on fracture surfaces. It combines some of the properties of glass and silica with some of those normal industrial carbons. It has been generally noted that the resulting product depends on the original polymer composition (polyfur furyl alcohol and phenolic resin are most common), polymerized structure, the highest temperature of the heat treatment operation, and the rate and other conditions of thermal decomposition.

The fabrication of glassy carbon involves a thermosetting organic resin preparation, moulding the shape, and a pyrolysis operation. Shapes are produced to the net dimension by incorporating a uniform shrinkage factor into the mold design.

The properties distinguishing glassy carbon from the

trial use but these materials are basically graphitic in structure or may be graphitized readily, and their properties (e.g., degree of graphitization) can be controlled to some extent to suit particular application.

In the past ten years, it has been found that several cross-linked polymers can be carefully decomposed so as to yield relatively massive shapes of a highly disordered non-graphitizing carbon that has been called "glassy", "glass-like", or "vitreous".^{32, 37}

It is a very pure form of carbon with impurity content of less than 200 ppm. Its name stems from physical appearance and not from a detailed knowledge of the structure. The physical appearance of these materials resembles black glass, particularly on fracture surfaces. It combines some of the properties of glass and silica with some of those normal industrial carbons. It has been generally noted that the resulting product depends on the original polymer composition (polyfur furyl alcohol and phenolic resin are most common), polymerized structure, the highest temperature of the heat treatment operation, and the rate and other conditions of thermal decomposition.

The fabrication of glassy carbon involves a thermosetting organic resin preparation, moulding the shape, and a pyrolysis operation. Shapes are produced to the net dimension by incorporating a uniform shrinkage factor into the mold design.

The properties distinguishing glassy carbon from the

more common forms of the element includes high strength, low density, high hardness, resistance to corrosion and air oxidation, low impact strength, and above all, isotropic characteristics even in the fibrous form. The typical physical and mechanical properties of glassy carbon³³⁻³⁶ are given in Table I. One of the most striking differences between glassy carbon and conventional graphites is its extremely low permeability to gases, comparable to borosilicate glass. The internal pore structure of glassy carbon is non-communicating, quite small in size, and uniformly dispersed. These factors make glassy carbon attractive for high temperature uses. One class of use has been in the area of biomaterials where it has shown outstanding promise for compatibility in a wide range of uses in the human body such as heart valves, blood pumps, bone bridges, etc.⁴⁴

In spite of the extensive investigation conducted on this rather unusual material, only very sketchy and conflicting reports have been published on the atomic structure of glassy carbon. The present understanding is that glassy carbon would be made of small regions consisting of graphite-like atomic arrangement, stacked up in the c direction in a random manner. Based on the X-ray diffraction studies, Noda et al.^{38, 39} concluded that in addition to the trigonal bonding typical of the graphite structures, there were also tetrahedrally bonded carbon atoms, such as in diamonds, which constitute the main part of the criss-cross linkages which link graphite-like layers in a random way. Very recently, Jenkins and Kawamura⁴⁰ have

TABLE I: Typical Physical and Mechanical Properties of Glassy Carbon

<u>Property</u>	<u>Unit</u>	<u>Value</u>
Apparent density	grm/c.c.	1.43-1.50
Apparent porosity	%	0.2-0.4
Helium gas permeability	cm ² /sec	10 ⁻¹² -10 ⁻⁷
Hardness	moh	4-5
Tensile strength	psi	15,000-29,000
Compressive strength	psi	86,000-200,000
Young's modulus	psi	3.4×10 ⁶ -4.0×10 ⁶
Poisson's ratio	dimensionless	0.1
Izod impact strength	in.-lb/in.-notch	1.2
Thermal expansion	per °C	2×10 ⁻⁶ -5×10 ⁻⁶
Thermal conductivity	cal/cm/°C/sec	0.010-0.020
Electrical resistivity	ohm-cm	30×10 ⁻⁴ -50×10 ⁻⁴
Flexural strength	psi	13,000-22,000
Shear modulus	psi	1.4×10 ⁶ - 1.6×10 ⁶

proposed a structural model for the network of ribbon stacks in glassy carbon which is schematically shown in Figure 2.

The interlayer spacing decreases and the crystallite size increases with an increase in the heat treatment temperature. It can be said that the graphitization, that is to say, the development of a graphite-like layer structure, proceeds in glassy carbon with heat treatment. However, glassy carbon is a typical non-graphitizing carbon because the progress of its graphitization is very little in comparison with that of graphitizing carbons.

The heat of combustion of glassy carbon (Reaction 3) at 298.15°K has been measured by Lewis et al.⁴¹ The measurements were made in an aneroid bomb calorimeter and the reported value is 95,277 calories/grm-mole (94,054 cal/grm-mole for graphite) with a standard deviation of 19 cal/grm-mole. The mean bond energy in glassy carbon is clearly much less than in graphite, and this could well be attributed to the presence of an appreciable fraction of carbon-carbon bond in a strained state. Takahashi and Westrum⁴² have reported the heat capacity of glassy carbon in the temperature range of 5 to 350°K. Their results parallel that of pyrolytic graphite in showing an approximate T^2 dependence up to 30°K suggesting that the structure of this material may involve microdomains of lamellar graphite. Although a definitive conclusion regarding the structure of glassy carbon is still precluded at present, Takahashi and Westrum⁴² concluded that graphite type structure

seems favored, rather than an isotropic three dimensional random network configuration of carbon atoms such as that proposed by Furukawa.^{4 3}

No literature, either direct measurement or indirect computation using physical and thermal properties, on the thermodynamics of glassy carbon-graphite transition has been published. Utilizing the available thermodynamic data, a rough estimate of the thermodynamic properties of the glassy carbon-graphite and diamond-graphite transitions at 298.15°K have been shown in Appendix II.

B. Solid Oxide Electrolytes

Solid oxide electrolytes have experienced an intense level of activity since the rebirth of interest supplied by Kiukkola and Wagner's paper^{4 5} involving solid electrolytes. Within the last fifteen years solid oxide electrolytes have been used over the temperature range of 400-1600°C to measure oxygen activities in electrodes of solids, liquids, and gases.

Independent of any specific knowledge of the mechanism of ionic conduction in a crystalline electrolyte which exhibits exclusive ionic conduction (defined at $t_{ion} > 0.99$), a simple consideration of energy conversion provides the relationship between the reversible cell potential E and the standard Gibb's free energy change ΔG of the virtual chemical reaction of the cell I,

$$p_{O_2}^I / \text{solid oxide electrolyte} / p_{O_2}^{II} \quad (I)$$

$$\Delta G = -nEF = RT \ln \left(\frac{p_{O_2}^{II}}{p_{O_2}^I} \right)^{1/4} \quad (5)$$

where n is the equivalent of charge passed through the external circuit, F is the Faraday's constant, and $p_{O_2}^I$ and $p_{O_2}^{II}$ are the oxygen partial pressures of two electrodes. The thermodynamic quantities of the left and the right hand electrode of the cell under consideration are designated by the Roman superscripts and subscripts I and II, respectively. The application of the Gibbs-Helmholtz relation to the temperature dependence of the cell voltage provides values of ΔH and ΔS of the virtual cell reaction.

At present, the values of ΔG° deduced from emf measurements according to equation 5 is superior to values of ΔG° obtained with the help of other methods. The accuracy of this type of measurement is of the order of ± 1 mV corresponding to an uncertainty of the order of ± 50 calories/mole in ΔG° , whereas the uncertainty of ΔG° obtained with the help of calorimetric measurements is much higher in view of the uncertainty of $\Delta H_{298.15}^\circ$ and the heat capacity values.

While electronic conduction in aqueous electrolytes is always negligible, crystalline compounds exhibit partial electronic conductivities which are rarely negligible. A few binary and ternary compounds exhibit exclusive ionic conduction

and this conduction is limited to a specific range of metal or non-metal chemical potentials.

If a compound should contain in solid solution an ion of valence lower than those of the host matrix, then the equilibrium defect concentration may be altered greatly. This mechanism is often used intentionally to introduce or enhance predominant ionic conduction in a compound.

Both ZrO_2 -15 mole percent CaO and ThO_2 -15 mole percent $YO_{1.5}$, the most commonly used solid electrolytes, are known to have a CaF_2 -type structure where Ca^{2+} and Y^{3+} ions, respectively, substitute directly on Zr^{4+} and Th^{4+} sites giving rise to a substantial fraction of vacant oxygen sites.⁴⁷ This maintains the electrical neutrality in the compound.

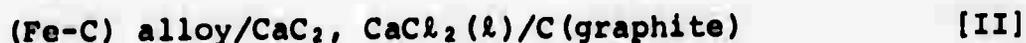
Of prime importance in the successful application of ZrO_2 -CaO and ThO_2 - $YO_{1.5}$ electrolytes is the knowledge of adequate information about the range of temperature and corresponding oxygen partial pressure over which each electrolyte exhibits predominantly ionic conduction. The criterion for the predominantly ionic conduction at a given temperature is a p_{O_2} -independent total conductivity at that temperature. This is established experimentally and is represented as a Kröger and Vink type of plot.

Patterson et al.⁴⁸ measured the electrical conductivity of both $Zr_{0.85}Ca_{0.15}O_{1.85}$ and $Th_{0.85}Y_{0.15}O_{1.925}$ electrolytes under various ambient conditions and suggested p_{O_2} lower limits for exclusive ionic conduction. Their data suggests p_{O_2} lower

limits of 10^{-29} , 10^{-28} , 10^{-27} , 10^{-26} , and 10^{-25} atmospheres at 800, 850, 900, 950, and 1000°C, respectively, for the ZrO_2 -CaO electrolyte. The corresponding p_{O_2} lower limits for the ThO_2 - $YO_{1.5}$ are considerably lower. Steele and Alcock⁴⁸, as well as Schmalzried⁴⁹, report higher oxygen partial pressures for the lower limits of ZrO_2 -CaO electrolyte (about 10^{-16} and 10^{-23} at 1000°C, respectively). Patterson⁵⁰ has recently presented the conduction domains for a number of known solid electrolytes, deduced from various conductivity and galvanic cell measurements, in the form of conduction domain maps in $\log p_{O_2}$ vs. $1/T$ space. His results for $t_{ion} > 0.99$ are shown in Table II.

C. Fused Salt Electrolyte

Hawkes and Morris⁵⁴ have recently employed the following galvanic cell in the temperature range of 800-1000°C to obtain carbon activities in Fe-C alloys:



The electrolyte used was a fused salt solution of 5-10% calcium carbide in calcium chloride. Their data reasonably matched with those obtained by gas equilibration methods using CO-CO₂ and CH₄-H₂ gas mixtures. Based upon X-ray examination of crystals of calcium carbide, carbon is believed to be present in the electrolyte as the so called acetylide ion, C_2^{2-} .

TABLE II. Oxygen Partial Pressure Ranges for Solid Oxide Electrolytes

Temperature °C	Lower P_{O_2} (limit atmospheres)			Higher P_{O_2} (limit atmospheres)		
	Zr _{0.85} Ca _{0.15} O _{1.85}	Th _{0.85} Y _{0.15} O _{1.85}	O _{1.825}	Zr _{0.85} Ca _{0.15} O _{1.85}	Th _{0.85} Y _{0.15} O _{1.825}	O _{1.825}
600	10^{-31}	---	---	10^{14}	---	---
700	10^{-25}	10^{-36}		10^{11}	10^{-8}	
800	10^{-22}	10^{-32}		10^9	10^{-6}	
900	10^{-19}	10^{-28}		10^7	10^{-6}	
1000	10^{-16}	10^{-23}		10^6	10^{-6}	
1100	10^{-14}	10^{-20}		10^5	10^{-6}	
1200	10^{-12}	---		10^3	---	
1300	10^{-10}	---		10^2	---	
1400	10^{-9}	---		10^2	---	
1500	10^{-8}	---		10	---	
1600	10^{-7}	---		1	---	
1700	10^{-6}	---		10^{-1}	---	

D. Thermodynamic and Kinetic Considerations

There are few specific criteria to be satisfied for a reliable and meaningful measurement employing electrochemical techniques:

1. Thermodynamic considerations

The electrolyte should be an exclusive ionic conductor of the ion of interest in the range of temperature and activity of the element prevailing in the experiments.

In the experiments using solid electrolytes (to be discussed later) the oxygen partial pressures of the systems would be about 10^{-10} to 10^{-18} atms. at 1000°C and about 10^{-24} atms. at 600°C . In this range of temperature and oxygen partial pressure, even zirconia-calcia electrolytes are exclusively ionic conductors of oxygen ions.⁴⁶ However, according to Steele and Alcock⁴⁸ and Patterson⁵⁰, an oxygen partial pressure of 10^{-16} atms. at 1000°C is almost the lowest partial pressure at which the zirconia-calcia electrolytes still behave as ionic conductors. In that case, thoria-yttria electrolytes can be used, because their lower limit of oxygen partial pressure is much lower than that of zirconia-calcia electrolytes.

The work of Hawkes and Morris⁵⁴ can be taken as a proof for the fact that calcium carbide is an exclusive conductor of C_2^{2-} ion in the temperature range of $800-1000^{\circ}\text{C}$.

There should be no side reactions; such as those between electrodes and electrolytes, electrodes and contact leads, electrolytes and electrodes and cell atmospheres, and any

significant ionic exchange reactions in the electrolytes.

All possible reactions between ZrO_2 , CaO , ThO_2 , and Y_2O_3 , and C , CO , and CO_2 , forming carbides or metals, have positive Gibb's free energy change even at $1000^\circ C$. There is no known way in which carbon electrodes can react with $CaCl_2$ and CaC_2 . Platinum and platinum-10% rhodium wires are inert to argon gas, solid carbon, and carbon oxides up to $1000^\circ C$. Argon gas is inert to calcium carbide, calcium chloride, and solid electrolytes. There is no known possibility of any exchange reaction with C_2^{2-} ion in the fused salt.

The elements or compounds fixing activity of non-metals at the electrodes should be stoichiometric. NiO , CoO , NbO_2 , Nb_2O_5 (to be used in calibration runs), CO , and CO_2 involved in fixing activity of oxygen are all stoichiometric compounds. Therefore, all the thermodynamic criteria are fulfilled for precise measurements.

2. Kinetic considerations

After satisfying the thermodynamic criteria, the reversible emf should be obtainable in a reasonable time. The kinetics of $Ni-NiO$ ⁵¹, $Co-CoO$ ⁵², $NbO_2-Nb_2O_5$ ⁵³, and $O_2-CO-CO_2$ ⁵² equilibria are reasonably fast, and the reversible emf can be obtained in about 9 hours even at $600^\circ C$. The sensitivity of the cells involving solid electrolytes, which depends on the mobility of oxygen ions in the electrolyte, is reasonably good (± 1 mV) even at $600^\circ C$. Hawkins and Morris⁵⁴ report that the mobility of C_2^{2-} ions in the fused salt (5% CaC_2 in $CaCl_2$) is

high and the reversible emf could be obtained in about 8 hours.

The kinetics of gassification of carbon by carbon dioxide has not completely been established. Kaftanov et al.⁵⁵ report that the rate of gassification depends mainly on the chemical reaction at the gas-solid interface and only to a negligible extent on the diffusion in the gaseous phase. Their results do not agree with the earlier investigations. Turkdogan and Vinters⁵⁶ have recently measured the kinetics of oxidation of graphite and charcoal in carbon dioxide. They found that the rate is proportional to the square root of the pressure of the carbon dioxide and the rate is controlled primarily by the formation of carbon monoxide from chemisorbed oxygen and carbon on the pore surface. Rhead and Wheeler^{2,3} equilibrated charcoal powder and CO₂ gas in connection with their thermodynamic study of the producer gas reaction. In their experiments, equilibrium was reached rapidly at higher temperatures, but at 800°C it was only attained after heating for 72 to 96 hours. At 700°C the reaction $C(s) + CO_2(g) \rightarrow 2CO(g)$ is extremely slow, and the equilibrium was attained only after five days.

Due to the extremely low gas permeability of glassy carbon, the producer gas reaction equilibrium with glassy carbon cylinders and pure carbon dioxide is expected to be very slow even at moderately high temperatures. The composition of CO-CO₂ gas mixtures of different temperatures in equilibrium with graphite and glassy carbon can be roughly calculated from the available thermodynamic data. It is proposed to equilibrate

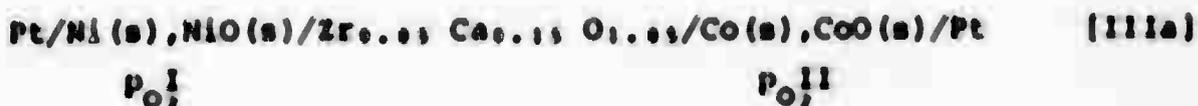
graphite and glassy carbon samples with calculated CO-CO₂ gas mixtures rather than with pure CO₂ or CO gas. The equilibrium with such gas mixtures is expected to be reasonably fast.

III. PROPOSED RESEARCH

A. Plan of Work

1. Calibration Runs

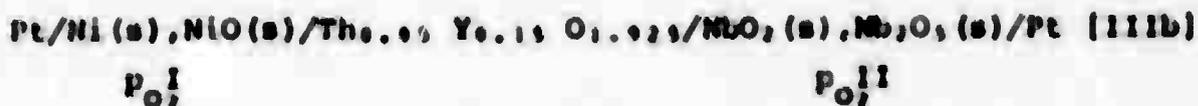
Cells IIIa and IIIb involving reference electrodes in the pertinent temperature and oxygen partial pressure range will be investigated to demonstrate the reliability of the experimental apparatus.



Temperature range: 600-1000°C

Oxygen partial pressure range: 10^{-21} (600°C) - 10^{-12}
(1000°C) atm.

Cell atmosphere: Purified argon



Temperature range: 600-1000°C

Oxygen partial pressure range: 10^{-21} (600°C) - 10^{-12}
(1000°C) atm.

Cell atmosphere: Purified argon

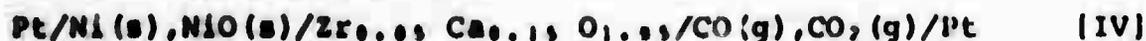
Relating the Gibb's free energy change of virtual cell reaction with the emf of the cells III, one gets

$$E(\text{III}) = \frac{1}{2F} (\Delta G_{\text{I}}^{\circ} - \Delta G_{\text{II}}^{\circ}) = \frac{RT}{2F} \ln \left(p_{\text{O}_2}^{\text{I}} / p_{\text{O}_2}^{\text{II}} \right) \quad (6)$$

$$\frac{dE(\text{III})}{dT} = \frac{1}{2F}(\Delta S_{\text{II}}^{\circ} - \Delta S_{\text{I}}^{\circ}) \quad (7)$$

The expected emf of cells IIIa and IIIb, calculated from the most accurate available values of $\Delta G_{\text{NiO}}^{\circ}$, $\Delta G_{\text{CoO}}^{\circ}$, $\Delta G_{\text{NbO}_2}^{\circ}$, and $\Delta G_{\text{Nb}_2\text{O}_5}^{\circ}$ are shown in Table III.

2. CO-O₂-CO₂ Equilibrium



Temperature range: 600-1000°C

Oxygen partial pressure range: 10^{-23} (600°C) - 10^{-12}
(1000°C) atms.

Cell atmosphere: Solid electrode side; Purified argon
Gaseous electrode side; CO-CO₂ gas
mixture of controlled oxygen
potentials

The Cell IV would be investigated to measure the standard free energy change of the reaction 1. The CO-CO₂ gas mixture of controlled oxygen potentials would be prepared by using the constant pressure head capillary flow meters. The emf of cell IV as a function of temperature would give the standard enthalpy and entropy changes of reaction 1.

$$\Delta G_{\text{II}}^{\circ} = \Delta G_{\text{I}}^{\circ} - 2E(\text{IV})F - RT \ln(p_{\text{CO}_2}/p_{\text{CO}})^{\text{II}} \quad (8)$$

$$\Delta H_{\text{II}}^{\circ} = \Delta H_{\text{I}}^{\circ} - 2(E(\text{IV}))F \quad (9)$$

T=0

$$\Delta S_{\text{II}}^{\circ} = \Delta S_{\text{I}}^{\circ} + 2\left(\frac{dE(\text{IV})}{dT}\right)F + R \ln(p_{\text{CO}_2}/p_{\text{CO}})^{\text{II}} \quad (10)$$

TABLE III. Expected Emf of Calibration Runs

Temp. °C	$P_{O_2}^I$ (Ni-NiO) atms	$P_{O_2}^{II}$ (Co-CoO) atms	$P_{O_2}^{III}$ (NbO ₂ -Nb ₂ O ₅) atms	E (IIIa) mV	E (IIIb) mV
600	$10^{-19.1}$	$10^{-20.8}$	$10^{-29.4}$	72	444
650	$10^{-17.6}$	$10^{-19.2}$	$10^{-27.3}$	76	446
700	$10^{-16.2}$	$10^{-17.9}$	$10^{-25.5}$	80	448
750	$10^{-15.0}$	$10^{-16.6}$	$10^{-23.9}$	84	450
800	$10^{-13.9}$	$10^{-15.5}$	$10^{-22.4}$	88 (±3)	452 (±4)
850	$10^{-12.9}$	$10^{-14.5}$	$10^{-21.0}$	92	454
900	$10^{-11.9}$	$10^{-13.5}$	$10^{-19.8}$	96	456
950	$10^{-11.1}$	$10^{-12.7}$	$10^{-18.6}$	100	458
1000	$10^{-10.3}$	$10^{-11.9}$	$10^{-17.6}$	104	460

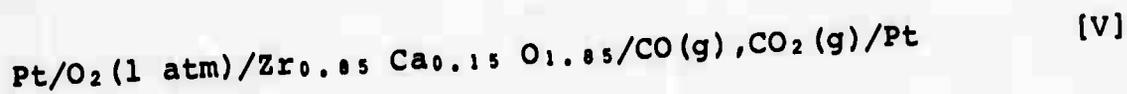
The success of calibration runs (reasonable match between the the expected and actual emf) would demonstrate that the solid electrolytes are indeed exclusive ionic conductors over the range of temperature and P_{O_2} prevailing in the experiments.

If equation 8 is differentiated,

$$\delta(\Delta G_{II}^{\circ}) = \delta(\Delta G_I^{\circ}) - 2\delta(E(IV))F - R(\delta T)\ln(p_{CO_2}/p_{CO})^{II} \quad (11)$$

Taking $\delta(\Delta G_I^{\circ})$ value of ± 50 calories from the work of Charette and Flengas,⁵¹ with one millivolt uncertainty in the measured cell voltage, 1° uncertainty in the furnace temperature, and for a fixed $(p_{CO_2}/p_{CO})^{II}$ ratio of 10, corresponds to 100 calories uncertainty in ΔG_{II}° . In the absence of any special problem, careful experimentation can achieve measurements of this accuracy.

Even more accurate values of ΔG_{II}° (± 50 calories) can be obtained from the emf of the cell V.



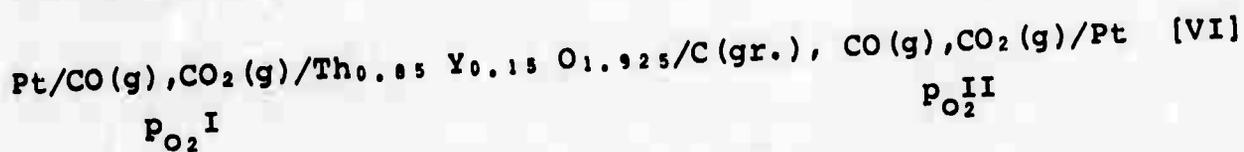
Temperature range: 600-1000°C

Oxygen partial pressure range: 10^{-23} (600°C) - 10^{-12} (1000°C) atms.

$$\Delta G_{II}^{\circ} = -2E(\text{V})F - RT \ln(p_{CO_2}/p_{CO})^{II} \quad (12)$$

3. C-CO-CO₂ Equilibrium

It would become clear later that the accuracy of carbon activities in glassy carbon would strongly depend on the accuracy of ΔG° of reaction 2. The cell VI would be investigated to accurately measure ΔG° of reaction 2.



Temperature range: 600-1000°C

Oxygen partial pressure range: 10^{-24} (600°C) - 10^{-18}
(1000°C) atms.

Cell atmosphere: Gaseous electrode; CO-CO₂ gas mixture
of controlled oxygen partial
pressure

Graphite electrode; CO-CO₂ gas mixture
in equilibrium with graphite

The use of CO-CO₂ gas mixture instead of Ni-NiO or Co-CoO as a
reference electrode has two distinct advantages:

- (a) Since the reaction $\text{CO(g)} + \frac{1}{2}\text{O}_2\text{(g)} = \text{CO}_2\text{(g)}$ is
common on both sides of the electrode, the un-
certainty in its free energy values would not
further increase the uncertainty in free energy
values of the reaction $\text{C(s)} + \text{CO}_2\text{(g)} = 2\text{CO(g)}$.
- (b) By increasing the $p_{\text{CO}_2}/p_{\text{CO}}$ ratio of CO-CO₂ electrode,
its oxygen partial pressure can be increased by a
few orders of magnitude.

The oxygen partial pressures on two electrode sides and the emf
of cell VI can be related as

$$p_{\text{O}_2}^{\text{I}} = A^2 e^{\frac{2\Delta G_{\text{I}}^{\circ}}{RT}} \quad (13)$$

$$p_{\text{O}_2}^{\text{II}} = B^2 e^{\frac{2\Delta G_{\text{I}}^{\circ}}{RT}} \quad (14)$$

$$p_{\text{O}_2}^{\text{I}} = p_{\text{O}_2}^{\text{II}} e^{\frac{4E(\text{VI})F}{RT}} \quad (15)$$

where ΔG_I° is the standard free energy change of reaction 1 and $A = (p_{CO_2}/p_{CO})^I$, $B = (p_{CO_2}/p_{CO})^{II}$. "A" would be maintained constant and at a total pressure of 1 atmosphere, and taking graphite as a standard state for carbon, "B" can be expressed in terms of ΔG_{II}° of reaction 2.

$$B = \frac{1 + \frac{e}{2} e^{-\Delta G_{II}^\circ/RT} - \left(\frac{e}{4} e^{-2\Delta G_{II}^\circ/RT} + e^{-\Delta G_{II}^\circ/RT} \right)^{1/2}}{-\frac{e}{2} e^{-\Delta G_{II}^\circ/RT} + \left(\frac{e}{4} e^{-2\Delta G_{II}^\circ/RT} + e^{-\Delta G_{II}^\circ/RT} \right)^{1/2}} \quad (16)$$

Combining equation 13 to 16, one gets an expression for ΔG_{II}° as

$$\Delta G_{II}^\circ = RT \ln \left(\frac{A}{1 + A e^{-2E(VI)F/RT}} \right) - 2E(VI)F \quad (17)$$

It is worthwhile to note that the emf of cell VI would not be a straight line when plotted against temperature (as in most of the electrochemical cells), because equation 18 is non-linear.

$$E(VI) = \frac{RT}{2F} \ln \left(\frac{2A}{-1 + \sqrt{1 + 4e^{\Delta G_{II}^\circ/RT}}} \right) \quad (18)$$

This behavior of $E(VI)$ vs. T curve becomes clear when $E(VI)$ is expressed in terms of p_{CO}^{II} . Combining equation 16 and 18,

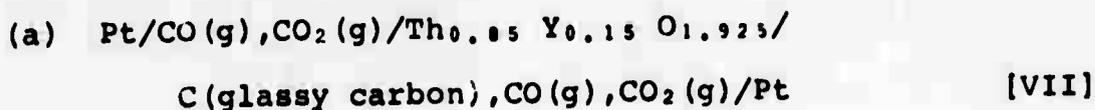
$$E(VI) = -\frac{\Delta H_{II}^\circ}{2F} + \frac{T}{2F} \left(\Delta S_{II}^\circ + R \ln \left(\frac{A}{p_{CO}^{II}} \right) \right) \quad (19)$$

where ΔG_{II}° , ΔH_{II}° , and ΔS_{II}° are the standard free energy, heat, and entropy change of reaction 2, respectively.

Since P_{CO}^{II} is itself a non-linear function of T, this makes E(VI) a non-linear function of T. The nature of the curves of P_{CO}^{II} and E(VI) against T has been discussed in Appendix III. The directly measured ΔG_{II}° values would be independent of ash content, ash composition, and even the kind of graphite used. Emf of cell VI would be measured as a function of temperature which would then give ΔG_{II}° from equation 17. ΔH_{II}° and ΔS_{II}° of reaction 2 can be readily obtained from the ΔG_{II}° vs. T curve. The accuracy of ΔG_{II}° would depend upon the standard deviation of the ΔG_{II}° vs. T curve.

4. Activity of Carbon in Glassy Carbon

Carbon activities in glassy carbon samples would be measured with reference to graphite as a standard state, using the following cell configurations:



Temperature range: 600-1000°C

Oxygen partial pressure range: 10^{-24} (at 600°C)

- 10^{-18} (at 1000°C) atms.

The emf of the above cell can be shown to be

$$E(\text{VII}) = \frac{RT}{2F} \ln \frac{2A}{-1 + \left[1 + \frac{4 e^{\Delta G_{II}^{\circ}/RT}}{a_c} \right]^{1/2}} \quad (20)$$

where a_c is carbon activity in glassy carbon and other symbols have their predefined meanings.

In this case also E(VII) vs. T curve would not be linear for the reasons discussed previously. However, carbon activity can be calculated explicitly as a function of temperature by using the following relation

$$a_c = \frac{e^{(\Delta G_{II}^{\circ} + 2E(VII)F)RT}}{A(1 + A e^{-2E(VII)F/RT})} \quad (21)$$

If equation 21 is differentiated, one gets

$$\delta(a_c) = \frac{a_c}{RT^2} \left\{ (T\delta(\Delta G_{II}^{\circ}) - \Delta G_{II}^{\circ} \delta T) + 2F(T\delta E(VII) - E(VII)\delta T) (1 + A a_c e^{\frac{-(\Delta G_{II}^{\circ} + 4E(VII)F)}{RT}}) \right\} \quad (22)$$

It is evident from equation 22 that the uncertainty in carbon activity would depend on the uncertainties in the emf, temperature, ΔG_{II}° , as well as on the values of these variables.

(b) Pt/C_I(glassy carbon)CO(g), CO₂(g)/Th_{0.95}Y_{0.15}O_{1.925}/

C_{II}(graphite, CO(g), CO₂(g)/Pt [VIII]

Temperature range: 600-1000°C

Oxygen partial pressure range: 10⁻²⁴ (at 600°C)

- 10⁻¹⁰ (at 1000°C) atms.

It can be readily shown that the electrode with lower carbon activity is at higher oxygen potential. The emf of cell VIII is related to $\left(\frac{P_{CO_2}}{P_{CO}}\right)^I$ and $\left(\frac{P_{CO_2}}{P_{CO}}\right)^{II}$ which are ultimately related to carbon activities in two different forms and ΔG_{II}° .

$$\left(\frac{P_{CO_2}}{P_{CO}}\right)^I = \left(\frac{P_{CO_2}}{P_{CO}}\right)^{II} e^{2E(VIII)F/RT} \quad (23)$$

Referring to graphite as a standard state for carbon, carbon activity in glassy carbon can be written as

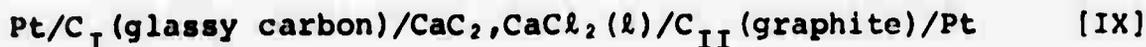
$$a_{C_I} = \left[\left(\frac{P_{CO_2}}{P_{CO}}\right)^{II} e^{\frac{2E(VIII)F - \Delta G_{II}^\circ}{RT}} \left\{ 1 + \left(\frac{P_{CO_2}}{P_{CO}}\right)^{II} e^{2E(VIII)F/RT} \right\}^{-1} \right] \quad (24)$$

where

$$\left(\frac{P_{CO_2}}{P_{CO}}\right)^{II} = \frac{1 + 2e^{\Delta G_{II}^\circ/RT} - \left(1 + 4e^{\Delta G_{II}^\circ/RT}\right)^{1/2}}{-1 + \left(1 + 4e^{\Delta G_{II}^\circ/RT}\right)^{1/2}} \quad (25)$$

Once the emf of cell VIII is measured as a function of temperature, free energy, entropy, and enthalpy change of the glassy carbon-graphite transition can be readily calculated.

(c) Fused salt electrolyte



Temperature range: 800-1000°C

Atmosphere: Purified argon

The carbon activity in glassy carbon is related to the emf of

cell IX as:

$$\Delta G^\circ = RT \ln a_{c_I} = -E(IX)F \quad (26)$$

5. Entropy of Glassy Carbon at Absolute Zero

The configuration entropy of a solid substance at 0°K is a direct measure of its disorder. Since the third law demands that a perfect crystalline substance have zero entropy at 0°K where no vibrational contributions are present, the failure of a substance to reach zero entropy at °K gives a measure of the lack of crystalline order. In studies of many imperfectly crystalline materials such as polymers accurate calorimetric methods have been used to deduce the degree of crystallinity. Necessarily, these methods can only directly yield values for configurational enthalpy differences. Since the proposed method can give a direct measure of the standard Gibb's free energy change, values for both the configurational entropy and enthalpy are obtainable.

The entropy change of the glassy carbon-graphite transition, proposed to be deduced from the results of cell VII or VIII or IX, can be written as

$$\Delta S_T^\circ = (S_{T,c,graphite}^\circ + S_{T,v,graphite}^\circ) - (S_{T,c,glassy\ carbon}^\circ + S_{T,v,glassy\ carbon}^\circ) \quad (27a)$$

The subscripts c and v refer to configurational and vibrational contributions, respectively. Similarly, ΔH_T° can be separated

into configurational and vibrational contributions.

$$\begin{aligned} \Delta H_T^\circ &= (H_{T,c,\text{graphite}}^\circ + H_{T,v,\text{graphite}}^\circ) \\ &- (H_{T,c,\text{glassy carbon}}^\circ + H_{T,v,\text{glassy carbon}}^\circ) \end{aligned} \quad (27b)$$

If high temperature heat capacity values for glassy carbon become available in the future, $S_{T,v,\text{glassy carbon}}^\circ$ values can be calculated. The heat capacity of graphite is available over a wide range of temperature.^{18, 28-30} Assuming graphite to be a perfect crystalline solid and utilizing above mentioned information, it would be possible to calculate configurational entropy of glassy carbon at absolute zero. Similarly, $(H_{T,c,\text{graphite}}^\circ - H_{T,c,\text{glassy carbon}}^\circ)$ values can be calculated using available heat capacity and heat of combustion data for graphite and glassy carbon. This provides a unique method of separating the total entropy and enthalpy into its configurational and vibrational contributions. Once the configurational entropy and enthalpy is known, different types of structural models of glassy carbon available in the literature can be tested for their validity and consistency.

B. Electrode, Electrolyte, and Atmosphere Preparation

1. Electrode preparation

Electrode pellets ($\frac{1}{4}$ " dia. \times $\frac{1}{4}$ " ht.) of Ni-NiO, Co-CoO, and NbO₂-Nb₂O₅ will consist of roughly equal amounts of metals and oxides.

Graphite would be machined to serve as the outer

electrode and the flat bottom solid electrolyte tube would slip into the cavity in the graphite electrode. This geometry would give a rigid support to the solid electrolyte tube as well as would provide more contact area between the electrode and the electrolyte.

It is almost impossible to machine baked glassy carbon therefore, unbaked glassy carbon would be machined in the form of small cylinders, and then baked pieces would slip inside the solid electrolyte tube and serve as the inner electrode. During the baking operation, the glassy carbon shrinks appreciably (20-22%), therefore, proper shrinkage allowances would be considered.

The electrodes of graphite and glassy carbon for fused salt experiments will be crucibles and long cylinders, respectively.

2. Electrolyte preparation

The solid electrolyte tubes would be purchased from Zirconium Corporation of America.

Fisher Scientific Company supplies technical grade calcium carbide. Hawkes and Morris⁵⁴ have described a laboratory procedure to prepare pure calcium carbide powder. Pure calcium and finely divided graphite (outgassed at 1000°C) mixture containing a small stoichiometric excess of calcium are heated in a mild steel crucible at 900°C for about 48 hours in a stream of pure argon.

3. Atmosphere preparation

(a) Purified Argon

The high grade argon would be purified of oxygen, moisture, carbon dioxide, and carbon monoxide and hydrogen by passing the gas through reduced BTS catalyst at 200°C, P₂O₅, and magnesium perchlorate, ascarite, and oxidized BTS catalyst at 250°C, respectively.

(b) CO/CO₂ Gas Mixture of Controlled Oxygen Potentials

Purified carbon monoxide (by passing it through dehydrating compounds, ascarite, and BTS catalyst in the reduced form) and carbon dioxide (by passing it through dehydrating compounds and BTS catalyst in both oxidized and reduced forms) gases would be mixed for a controlled oxygen potential by using the constant pressure head capillary flow meters which were first developed by Johnston and Walker³⁷ and later modified by Darken and Gurry³⁸ and Schwerdtfeger and Turkdogan³⁹. The basic principle is to maintain a constant pressure drop through each of two flowmeters by bleeder tubes. The pressure drop across the capillary is measured by the monometer M. The gases from both flow meters are mixed in column G packed with glass beads. The liquid heads in the bleeder tubes are adjusted to the required levels by transferring dibutyl phthalate from the reservoir R. The capillaries C₁ and C₂ are held in position by ground glass joints for easy exchange from one calibrated capillary to another.

The capillary tubes would be calibrated by measuring the flow rate of the rise of a soap bubble in a graduated glass

tube of uniform cross-section, e.g., a soap bubble flow meter.

Johnston and Walker¹⁷ report that the partial pressure ratio of gases as low as 10^{-5} (or as high as 10^5) could be maintained constant for days. The gas mixing system to be used in the proposed research is schematically shown in Figure 3.

C. Experimental Apparatus and Procedure

1. Experimental Apparatus

(a) Solid Oxide Electrolytes

The electrode isolation requirement is best fulfilled by the complete enclosure of one electrode in an impervious electrolyte tube. Therefore, the electrolyte of the tubular form, 1/2" OD, 3/8" ID, 24" long, would be used. A platinum lead will be passed through a quartz capillary tube which is pressed on the electrode by a spring to provide good contact with the electrolyte. The other platinum lead wire will be pushed against the outside flat end of the tube. A Pt/Pt-10% Rh thermocouple would be used for the temperature measurement. The whole assembly would be enclosed in a mullite tube. The cells would be slowly heated in a vertical resistance furnace capable of reaching 1000°C. An electrically grounded metallic tube (nichrome) surrounding the mullite tube would minimize induced voltages from the resistance furnace as well as broaden the constant temperature zone and increase its heat capacity (to reduce temperature fluctuations). The proposed cell assembly is shown in Figure 4.

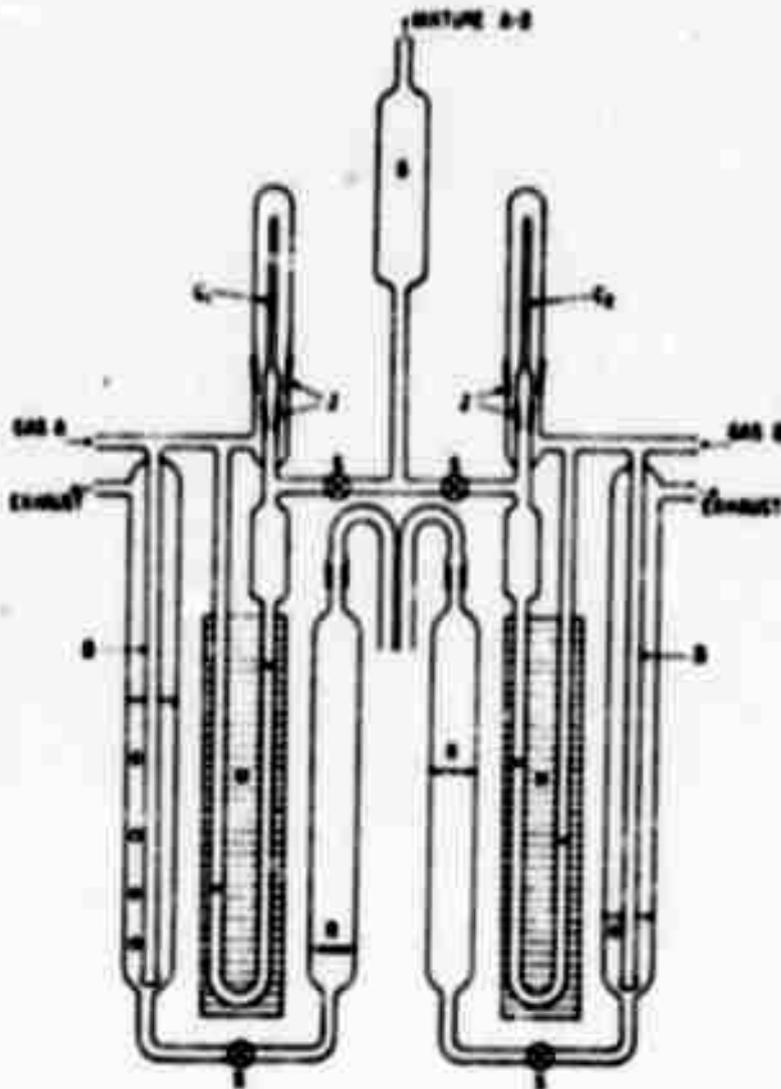


Figure 3. CO/CO₂ Gas Mixing Train for Controlled Oxygen Potentials

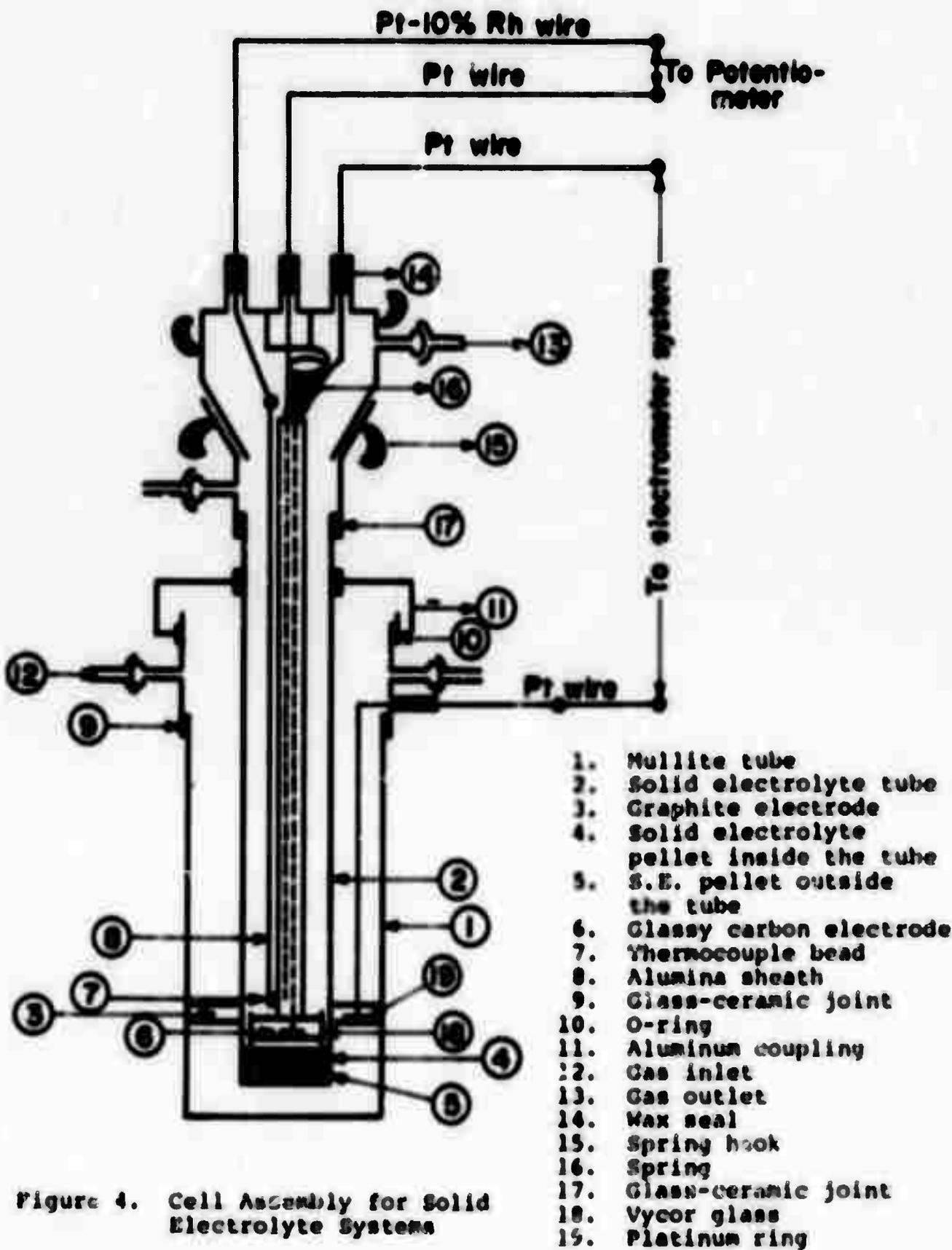


Figure 4. Cell Assembly for Solid Electrolyte Systems

(b) Fused Salt Electrolyte

The apparatus is shown in Figure 5. The cell would consist of a graphite crucible positioned inside a mullite tube. The glassy carbon electrode would rest on a zirconia disk covering the graphite crucible. The lower end of the glassy carbon electrode would be immersed inside the electrolyte (5% CaC_2 in CaCl_2) contained in the graphite crucible. The temperature would be controlled ($\pm 0.5^\circ\text{C}$) by a Wheelco proportionate controller.

The emf can be measured by a null-balance potentiometer. This method depends on the exact balancing of the cell emf by an emf from the potentiometer, however, if the balance is not exact, two problems arise:

- (1) the measured emf is in error by an amount depending upon the resistance of the circuit and the out-of-balance current, and
- (2) the passage of current through the cell causes the measured potential to deviate from that of the true equilibrium (polarization effect).

The second method depends on the use of electrometer amplifiers to read the emf directly. They have huge internal impedance (of the order of a few megohms) compared to potentiometers (500 ohms in L & N K-3 type potentiometer). Furthermore, the speed of response, convenience, continuity of reading, of the electrometer system are great advantages when an accuracy of about 0.1% is sufficient. Thus, for fused salt as well as

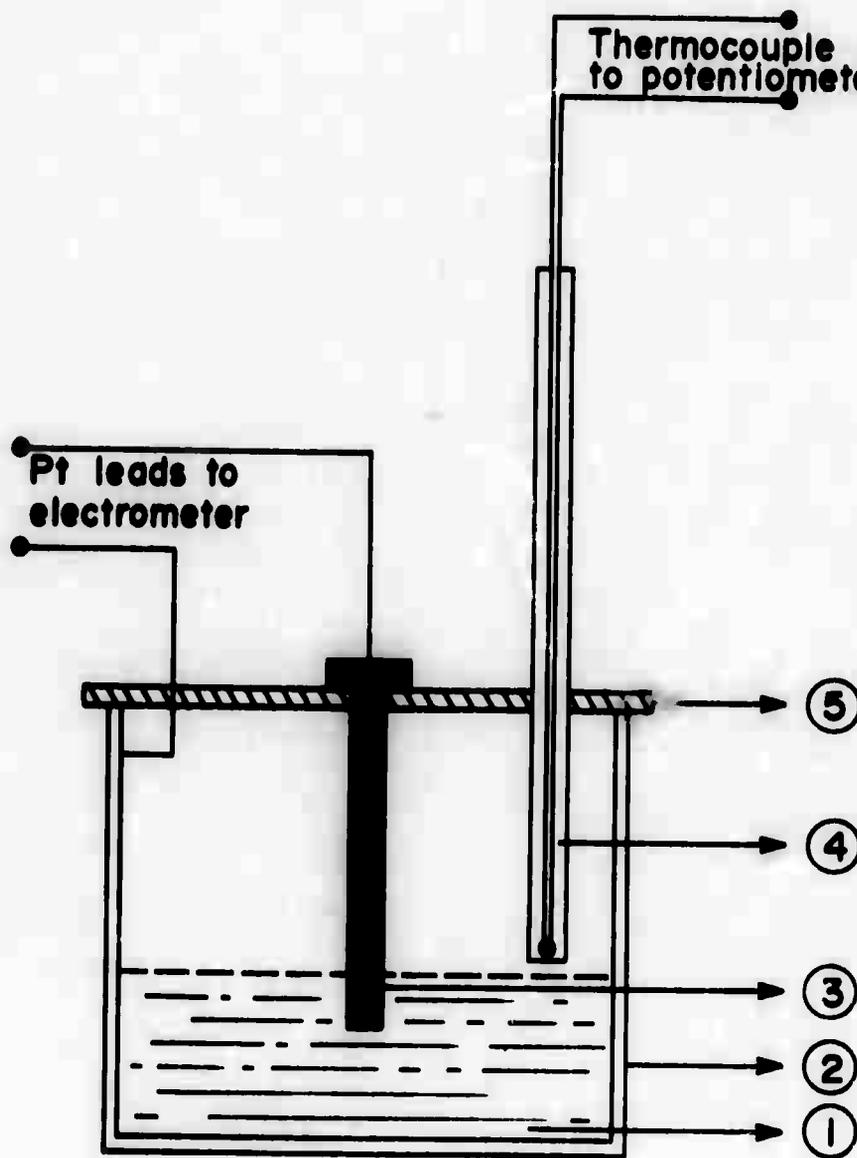


Figure 5. Cell Assembly for Fused Salt Systems

- (1) Fused salt electrolyte (5% CaC_2 in CaCl_2)
- (2) Graphite crucible
- (3) Glassy carbon cylinder
- (4) Alumina sheath
- (5) Zirconia disk

for solid electrolyte studies, the electrometer system would be used.

2. Procedure

Each cell would be assembled at room temperature. It will then be evacuated several times and flushed with purified argon or CO/CO₂ mixture. The cell would be heated slowly to 1000°C and left overnight for homogenization and equilibration. Emf would be recorded at intervals of 30 minutes and if the variations are within 0.1% over a period of 2-3 hours, the emf values would be considered to be steady. The rate of gas flow past the cell would be varied to check its influence on the cell voltage. The temperature would then be lowered by 50°C and measurements would be made until a steady emf is obtained. In experiments involving solid electrolytes, the cells would be evacuated at each temperature and then would be flushed with CO-CO₂ gas mixture, in equilibrium with graphite or glassy carbon (calculated from available data) at that temperature. This would avoid any deposition of carbon during cooling. This procedure would be repeated until the emf at lowest temperature is measured. The reversibility of cell voltage would be checked at each temperature by polarizing the cell in both directions. Some of the samples would be repeated and the temperature cycling would be done to check the reproducibility of the data.

V. REFERENCES

Carbon Oxides

1. O. Boudard, Ann. Chim. Phys., 24, 1(1901).
2. T. F. E. Rhead and R. V. Wheeler, J. Chem. Soc., 99, 1140(1910).
3. T. F. E. Rhead and R. V. Wheeler, Ibid, 1140(1911).
4. K. Jellinek and A. Diethelm, Z. anorg. allgem. Chem., 124, 203(1922).
5. F. J. Dent and J. W. Cobb, J. Chem. Soc., 1903(1929).
6. J. Chipman, J. Phys. Chem., 60, 1467(1956).
7. E. J. Prosen, R. S. Jessup, and F. D. Rossini, J. Res. Nat. Bur. Std., 33, 447(1944).
8. P. Hawtin, J. B. Lewis, N. Moul, and R. H. Phillips, Phil. Trans. Roy. Soc. Ser. A, 261, 67(1966).
9. F. D. Rossini, J. Res. Nat. Bur. Std., 22, 407(1939).
10. F. D. Richardson and W. E. Dennis, Trans. Faraday Soc., 49, 171(1953).
11. A. Rist and J. Chipman, Physical Chemistry of Steel Making, The Technology Press of M.I.T. and John Wiley & Sons, Inc., New York, 3(1958).
12. T. Fuwa and J. Chipman, Trans. Met. Soc. AIME, 215, 708(1959).
13. C. Wagner, Techniques of Metal Research, Volume IV, Part 1, 1, Interscience Publishers (1970).
14. D. R. Stull, JANAF Thermochemical Tables, Dow Chemical Co., Midland, Michigan (1965).
15. D. R. Stull and H. Prophet, JANAF Thermochemical Tables, Second Edition, U.S. Dept. Comm. (1971).
16. J. Chipman, Trans. Met. Soc. AIME, 239, 2(1967).
17. R. P. Smith, J. Am. Chem. Soc., 68, 1163(1946).
18. F. D. Richardson and J. H. E. Jeffes, J. Iron & Steel Institute, 160, 261(1948).

19. M. D. Thompson, The Total and Free Energies of Formation of the Oxides of Thirty-Two Elements, Electrochemical Soc. Inc., New York (1942).
20. O. Kubaschewski, E. L. Evans, and C. B. Alcock, Metallurgical Thermochemistry, Fourth Edition, Pergamon Press (1967).
21. R. G. Ward, An Introduction to the Physical Chemistry of Iron and Steel Making, Edward Arnold Ltd., London (1962).
22. F. D. Rossini, D. D. Wagman, and W. H. Evans, Selected Values of Chemical Thermodynamic Properties, Nat. Bur. Std. Circular No. 500, Ser. III (1952).
23. F. D. Rossini, Selected Values of the Properties of Hydrocarbons, Nat. Bur. Std., Circular No. C461 (1947).
24. J. P. Coughlin, US Bur. Mines, Bull. No. 542 (1954).
25. J. F. Elliott and M. Gleiser, Thermochemistry of Steelmaking, Addison Wesley Publ. Co., Cambridge, Mass. (1960).
26. C. E. Wicks and F. E. Block, US Bur. Mines, Bull. No. 605 (1963).

Graphite

27. R. E. Nightingale, Nuclear Graphite, Academic Press, 87 (1962).
28. P. H. Keesom and N. Pearlman, Phy. Rev., 99, 1119 (1955).
29. W. DeSorbo and G. E. Nichols, Phys. and Chem. Solids, 6, 352 (1958).
30. W. DeSorbo and W. W. Tyler, J. Chem. Phys., 21, 1660 (1953).
31. Bernal, Proc. Roy. Soc. A, 106, 749 (1924).

Glassy Carbon

32. P. E. Duwez and E. E. Hucke, to be published.
33. F. C. Cowlard and J. C. Lewis, J. Mat. Sci., 2, 507 (1967).
34. Impermeable Carbon, Vitreous Carbon Division, Beckwith Carbon Corporation (1969).
35. LMSC Glass-Like Carbon, Lockheed, Palo Alto Res. Lab. (1969).
36. S. Yamada, A Review of Glass-Like Carbons, Battelle Memorial Institute, Columbus, Ohio (1968).

37. P. E. Duwez, to be published.
38. T. Noda, M. Inagaki, and S. Yamada, J. Non-Crys. Solids, 1, 285 (1969).
39. T. Noda and M. Inagaki, Bull. Chem. Soc. Japan, 37, 1534 (1964).
40. G. M. Jenkins and K. Kawamura, Nature, 231 (1971).
41. J. B. Lewis, R. Murdock, and A. N. Moul, Nature, 221 (1969).
42. Y. Takahashi and E. F. Westrum, Jr., J. Chem. Thermodynamics, 2, 847 (1970).
43. K. Furukawa, Nippon Kessho Gakkaishi, 6, 101 (1964).
44. J. Benson, North American-Rockwell, Rocketdyne Report R-7855 (1969).

Solid Oxide Electrolytes

45. K. Kiukkola and C. Wagner, J. Electrochem. Soc., 104, 308, 379 (1957).
46. J. W. Patterson, E. C. Bogren, and R. A. Rapp, J. Electrochem. Soc., 114, 752 (1967).
47. R. A. Rapp and D. A. Shores, Techniques of Metal Research, Volume IV, Part 2, 123, Interscience Publishers (1970).
48. B. C. H. Steele and C. B. Alcock, Trans. Met. Soc. AIME, 233, 1359 (1965).
49. V. H. Schmalzried, Z. Electrochem., 66, 572 (1962).
50. J. W. Patterson, J. Electrochem. Soc., 118, 1033 (1971).
51. G. G. Charette and S. N. Flengas, J. Electrochem. Soc., 115, 796 (1968).
52. K. Schwerdtfeger, Trans. Met. Soc. AIME, 239, 1276 (1967).
53. W. L. Worrell, J. Phys. Chem., 68, 952 (1964).

Fused Salt Electrolyte

54. G. L. Hawkes and D. R. Morris, Trans. Met. Soc. AIME, 242, 1083 (1968).

Kinetics of Carbon Gassification

55. S. V. Kaftanov, S. D. Fedoseev, and L. V. Svirida, Tr. Mosk. Khim-Tekhnol. Inst., 57, 29(1968).
56. E. T. Turkdogan and J. V. Vinters, Carbon, 7, 101(1969).

CO/CO₂ Gas Mixer for Controlled Oxygen Potentials

57. J. Johnston and A. C. Walker, J. Am. Chem. Soc., 47, 1807 (1925).
58. L. S. Darken and R. W. Gurry, J. Am. Chem. Soc., 67, 1398 (1945).
59. K. Schwerdtfeger and E. T. Turkdogan, Techniques of Metal Research, Volume IV, Part 1, 321, Interscience Publishers (1970).

Diamond-Graphite Equilibrium

60. R. Berman and S. F. Simon, Zeit. Für Elektrochemie, 59, 333(1955).
61. F. P. Bundy, H. P. Bovenkerk, H. M. Strong, and R. H. Wentorf, Jr., J. Chem. Phys., 35, 383(1961).

APPENDIX I

Critical Evaluation of the Thermodynamic Data of Carbon Oxides

In one of their classical publications, Richardson and Jeffes¹⁸ compiled and graphically represented the standard Gibb's free energy of formation of all oxides pertinent to the iron and steel making. However, they assigned only an order of magnitude value to the accuracy of free energy data by classifying them as A (± 1 Kcal), B (± 3 Kcal), C (± 10 Kcal), and D ($> \pm 10$ Kcal). The standard free energy values of carbon oxides were taken from the compilation of Thompson¹⁹, who had not stated the accuracy of his calculated values. Then Kubaschewski and Evans²⁰ and Ward²¹ again compiled free energy data from the same source and quoted the same order of magnitude value of the accuracy. Therefore, all the later compilations derived from Thompson's¹⁹ calculations give only the order of magnitude of error and their accuracy is anything less than 1 Kilocalorie, but the exact number has never been stated.

In JANAF Thermochemical Tables, first published in 1965, Stull et al.¹⁴ compiled the thermodynamic data of carbon oxides. The $\Delta H^{\circ}_{f,298.15}$ value for CO_2 (Reaction 3) was taken from Prosen, et al.'s⁷ paper and was corrected for the change in the molecular weight of CO_2 from 44.010 to 44.011. This correction would change Prosen et al.'s⁷ value by a factor of 1.0000227. The $\Delta H^{\circ}_{f,298.15}$ value for CO (Reaction 4) was computed by using $\Delta H^{\circ}_{f,298.15}$ values of reactions 1 and 3 taken from the works of

Rossini⁹ and Prosen, et al.⁷, respectively. Table 12 lists the $\Delta H^\circ_{f,298.15}$ values from the original sources,^{7,9} according to the present calculations, and from original and recent JANAF Thermochemical Tables.^{14,15} The values of $\Delta H^\circ_{f,298.15}$ (according to the present calculations) match reasonably well with those listed in the original JANAF Thermochemical Tables,¹⁴ but the accuracy limits attached to the values are in complete disagreement. The tabulation of Schwerdtfeger and Turkdogan⁵⁹ is based on the original JANAF Thermochemical Tables¹⁴ and therefore they list highly exaggerated values of the uncertainty limits. However, the accuracy limits reported in recently published (June, 1971) edition of JANAF Thermochemical Tables¹⁵ have been corrected and they match extremely well with those obtained in the present calculations.

Rossini²² made an elaborate compilation of chemical thermodynamic properties which for carbon oxides are essentially the same as made by Rossini et al.²³ under American Petroleum Institute Research Project No. 44. Later on Coughlin²⁴, Elliott and Gleiser²⁵, Wicks and Block²⁶ reported the same values. They used $\Delta H^\circ_{f,298.15}$ values for reaction 1 and 3 with an accuracy of ± 10 cal/mole (which, according to the present calculations, should be ± 39.5 , and ± 10.8 cal/mole, respectively) and calculated the values of $\Delta G^\circ_{f,T}$ of reactions 3 and 4 at various temperatures. The listed accuracy of $\Delta G^\circ_{f,298.15}$ values of reactions 3 and 4 are 30 and 20 cal/mole, respectively. Using equation 5, and the most accurate values of $\Delta S^\circ_{f,298.15}$

TABLE 12: Critical Evaluation of Thermodynamic Data of Carbon Oxides

Reaction	$\Delta H^{\circ}_f, 298.15$ cal/mole		$\Delta G^{\circ}_f, 298.15$ cal/mole		$\Delta G^{\circ}_f, 1000$ cal/mole
	Source	Value	Source	Value	
C (gr.) + O ₂ (g) = CO ₂ (g)	23	-94,051.8 ± 10.8	*	-94,260.6 ± 55.8	-94,628.6 ± 100.8
	*	-94,053.9 ± 10.8	29	-94,260.0 ± 30.0	
	22	-94,054.0 ± 30.0	22	-94,265.2	
	25	-94,054.0 ± 11.0	25	-94,265.2	
CO (g) + ½ O ₂ (g) = CO ₂ (g)	24	-67,636.1 ± 28.7	*	-61,480.8 ± 73.7	-46,769.6 ± 118.7
	*	-67,637.6 ± 28.7			
	22	-67,636.5			
	25	-67,638.0 ± 30.0			
C (gr.) + ½ O ₂ (g) = CO (g)	24	-26,393.8 ± 30.8	*	-32,779.8 ± 130	-47,859.0 ± 219.5
	*	-26,416.3 ± 39.5	29	-32,808.0 ± 20.0	
	22	-26,416.5 ± 620	22	-32,786.0	
	25	-26,417.0 ± 40.0	25	-32,786.0	
C (gr.) + CO ₂ (g) = 2CO (g)	24	41,242.3 ± 58.4	*	28,701.0 ± 203	-1,089.4 ± 338.2
	*	41,221.3 ± 68.2			

*Present Calculation

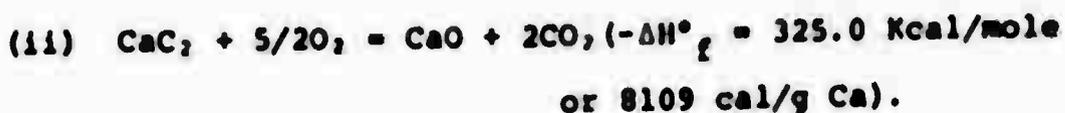
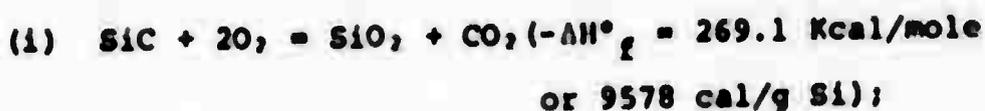
of reactions 1 and 3 (-20.650 ± 0.045 and 0.693 ± 0.045 cal/mole- $^{\circ}\text{K}$, respectively) the calculated accuracy of $\Delta G^{\circ}_{f,298.15}$ values of reactions 3 and 4 are ± 55.8 and ± 129.5 cal/mole, respectively. The accuracy of $\Delta S^{\circ}_{f,T}$ used in the present calculations incorporates errors in the experimental heat capacity values as well as their numerical integration. Error in $(\Delta H^{\circ}_{f,T} - \Delta H^{\circ}_{f,298.15})$ values are taken to be T times error in $\Delta S^{\circ}_{f,T}$ values.

$$\delta(\Delta G^{\circ}_{f,T}) = \delta(\Delta H^{\circ}_{f,298.15}) + \delta\left(\int_{298.15}^T \Delta C_p dT\right) + T\delta(\Delta S^{\circ}_{f,T}) \quad (5)$$

Table 12 shows $\Delta G^{\circ}_{f,T}$ values of reactions 1-4 at 298.15 and 1000 $^{\circ}\text{K}$ as obtained in the present calculations, listed by Coughlin²⁴, and original and recent JANAF Thermochemical Tables.^{14,15}

The thermochemical calculations and the extended discussion of their probable accuracy depend heavily on accuracy of the heat of combustion of graphite (reaction 3) as reported by Prosen et al.⁷ in 1944. When an accuracy of ± 10.8 cal/mole has been reported, one has to critically consider the effects of impurity content, impurity composition, pretreatment, and the kind of graphite. Experimentally it is almost impossible to determine the weight of ash present in the sample after combustion, and in calculating the results it is necessary to assume that any residue present in the crucible after combustion has the same composition as in the original graphite. This method of correcting for the weight of ash is not necessarily sound because the assumption is made that the impurities are

originally present in the graphite in the same form as they are found in the ash, namely, as oxides.⁶ It is more correct to assume that the impurities are present as carbides which oxidize during the combustion process. Thus, not only will the weight of ash differ from the original impurity content, but in addition, a significant quantity of heat will also be evolved by the oxidation of carbides. The effect of the impurity content on the heat liberated per gram of graphite burned can be demonstrated by considering the behavior of a graphite which has an ash content of 300 ppm; that is 1 gram of graphite produces 0.3 mg of ash. During the thermal pretreatment silicon and calcium present in the graphite are converted into carbides which in turn, during the combustion, undergo exothermic reactions:



If the concentrations of silicon and calcium impurities are both x mg/g carbon, then on combustion by stoichiometry, $2.14x$ mg of silica and $1.40x$ mg of calcium oxide will be produced for every gram of carbon. The weight of ash will be $3.54x$ mg. Equating this with the assumed figure of 0.3 mg gives x as 0.085. Thus the silicon and calcium content of the original graphite is 0.085 mg/g and consequently the SiC and CaC₂ contents are 0.121 and 0.136 mg/g, respectively. The amount of carbon which burns as graphite is depleted by the amounts combined as carbides and

these are 0.036 mg/g as SiC and 0.051 mg/g as CaC₂, leaving 0.999743 g/g of "free" carbon. The heat of combustion value (94,051.8 cal/mole) reported by Prosen et al.⁷ can be written as:

heat of combustion/g ash free graphite

$$= \frac{\text{observed heat release/g sample}}{1 - \text{ash content/g sample}}$$

Thus, the observed heat release = $7830.47(1-0.0003) = 7828.12$ cal/g. Therefore, 0.121 mg of SiC and 0.136 mg of CaC₂ would liberate 0.6g and 0.81 cal, respectively, and the heat liberated by the combustion of 0.999743 gm of "free" carbon in graphite is 7826.62 cal. This yields a value of 7828.63 cal/grm or 94029.7 cal/mole for the true heat of combustion of graphite. It can be seen that the apparent heat of combustion measured by Prosen et al.⁷ is 22.1 cal/mole (94,051.8 - 94029.7) higher than the true value for a graphite assumed to contain equal Si and Ca impurity and totaling 300 ppm of ash.

If the graphite sample was not prepared by heating to at least 2700°C in an inert atmosphere, the heat of combustion obtained was significantly higher (40 cal/mole).⁸ This result which was fully corrected for ash characteristics, shows the significant effect on heat of combustion due to difference in degree of graphitization. However, the value of heat combustion reported by Prosen et al.⁷ is an average of 17 measurements on Buckingham natural graphite and various artificial samples, with varying degrees of graphitization and ash characteristics.

APPENDIX II

Estimate of the Thermodynamic Properties of the Glassy Carbon-Graphite and Diamond-Graphite Equilibria at 298.15°K

Glassy Carbon-Graphite Equilibrium

The following data have been used in the present calculation:

- (1) The heat of combustion of glassy carbon at 298.15°K¹
- (2) The heat of combustion of graphite at 298.15°K⁰
- (3) The heat capacity of glassy carbon in the temperature range of 5-298.15°K²
- (4) The heat capacity of graphite in the temperature range of 1-298.15°K³

If both glassy carbon and graphite obey the third law, the following thermodynamic properties for the reaction C(glassy carbon) → C(graphite) can be obtained at 298.15°K.

$$\Delta H^\circ = -1,223 \pm 30 \text{ cal/mole}$$

$$\Delta S^\circ = -0.046 \pm 0.01 \text{ cal/mole-}^\circ\text{K}$$

$$\Delta G^\circ = -1,209 \pm 33 \text{ cal/mole}$$

$$(a_c)_{\text{glassy carbon}} = 0.13 \pm 0.006$$

However, it is anticipated that the configurational entropy of glassy carbon at absolute zero would be significantly high and hence, activity of carbon in glassy carbon would be close to unity.

Diamond-Graphite Equilibrium

For purposes of comparison, it is worthwhile to calculate

activity of carbon in diamond. The probable equilibrium line between diamond and graphite forms of carbon has been calculated thermodynamically from the physical properties and heats of formation of the two forms of carbon.^{60,61}

The following thermodynamic properties for the reaction $C(\text{diamond}) \rightarrow C(\text{graphite})$ can be obtained at 298.15°K:

$$\Delta H^\circ = -453 \text{ cal/mole}$$

$$\Delta S^\circ = 0.802 \text{ cal/mole-}^\circ\text{K}$$

$$\Delta G^\circ = -692 \text{ cal/mole}$$

$$(a_c)_{\text{diamond}} = 0.31$$

APPENDIX III

Nature of p_{CO}^{II} and E(VI) vs. Temperature Curves

Applying L'Hospital's rule, the limiting value of few functions which constitute the required equations 18 and 19 are:

$$\lim_{T \rightarrow 0} e^{\Delta G_{II}^{\circ}/RT} = \infty$$

$$\lim_{T \rightarrow 0} e^{-\Delta G_{II}^{\circ}/RT} = 0$$

$$\lim_{T \rightarrow \infty} e^{\Delta G_{II}^{\circ}/RT} = e^{-\Delta S_{II}^{\circ}/RT}$$

$$\lim_{T \rightarrow \infty} e^{-\Delta G_{II}^{\circ}/RT} = e^{\Delta S_{II}^{\circ}/RT}$$

$$\lim_{T \rightarrow 0} p_{CO}^{II} = 0$$

$$\lim_{T \rightarrow \infty} p_{CO}^{II} = 1$$

$$\lim_{T \rightarrow 0} \frac{d}{dT} (p_{CO}^{II}) = \infty$$

$$\lim_{T \rightarrow \infty} \frac{d}{dT} (p_{CO}^{II}) = 0$$

$$\lim_{T \rightarrow 0} E(VI) = -\Delta H_{II}^{\circ}/2F$$

$$\text{Lt } E(\text{VI}) = [-\Delta H_{\text{II}}^{\circ} + T(\Delta S_{\text{II}}^{\circ} + R \ln A)]/2F$$

$$T \rightarrow \infty$$

$$\text{Lt } \frac{d}{dT} (E(\text{VI})) = \infty$$

$$T \rightarrow 0$$

$$\text{Lt } \frac{d}{dT} (E(\text{VI})) = (\Delta S_{\text{II}}^{\circ} + R \ln A)/2F$$

$$T \rightarrow \infty$$

Figure 6 shows schematically the curves for both $p_{\text{CO}}^{\text{II}}$ and $E(\text{VI})$ vs. temperature.

$p_{\text{CO}}^{\text{II}}$ vs. T Curve

$p_{\text{CO}}^{\text{II}}$ starts with a value of zero and has an infinite slope at $T = 0$, and then asymptotically merges to $p_{\text{CO}}^{\text{II}} = 1$ line.

$E(\text{VI})$ vs. T Curve

The behavior of $p_{\text{CO}}^{\text{II}}$ vs. T curve is fully reflected on this curve. This curve also starts with an infinite slope, has a value of $-\Delta H_{\text{II}}^{\circ}/2F$ at $T=0$, and as $p_{\text{CO}}^{\text{II}}$ starts tending to unity, it is evident from equation 19 that this asymptotically merges to a straight line with a slope of $(\Delta S_{\text{II}}^{\circ} + R \ln A)/2F$. Even when $p_{\text{CO}}^{\text{II}}$ is close to unity but not unity, $E(\text{VI})$ almost merges to the straight line because in this region $\ln\left(A/p_{\text{CO}}^{\text{II}}\right)$ is very close to $\ln A$. It is worthwhile to note that the behavior of $E(\text{VI})$ vs. T curve resembles in many respects that of activity vs. mole fraction curve in a typical non-ideal metallic binary solution.

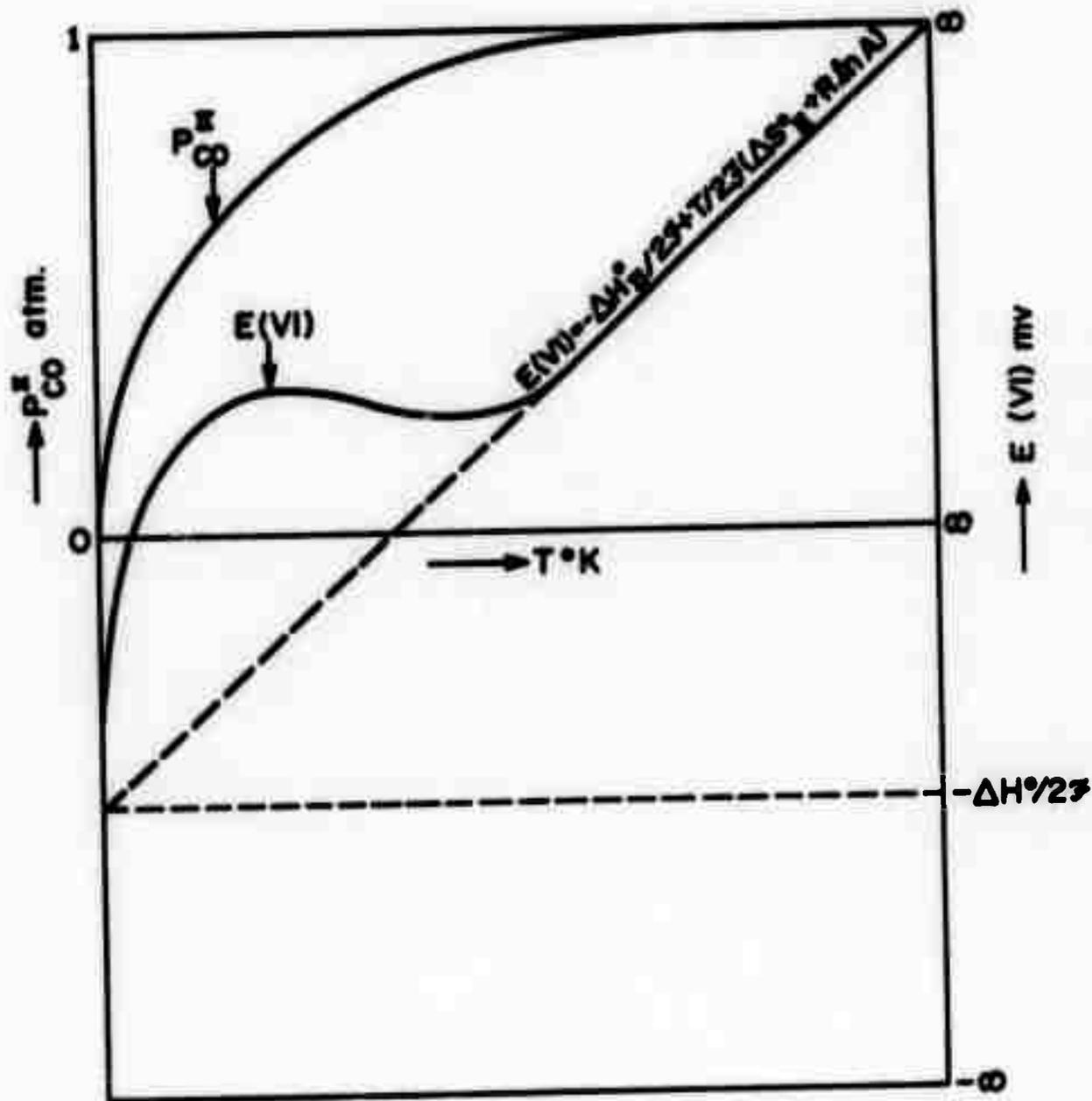


Figure 6. Nature of p_{CO}^{II} & $E(VI)$ vs. T Curves

FLOW VIA DISLOCATIONS IN IDEAL GLASSES

J. J. Gilman

Outline

- I. Introduction
- II. Nature of Dislocation Lines in Glasses
- III. Experimental Evidence for the Existence of Dislocation Lines
- IV. Kinematics of Flow Via Dislocation Motion
- V. Dynamics of Dislocation Motion
- VI. Implications of Dislocation Mechanics

Preceding page blank

FLOW VIA DISLOCATIONS IN IDEAL GLASSES

J. J. Gilman

I. Introduction

A few years ago the author (Gilman, 1968)¹ proposed that flow in noncrystalline solids often consists of correlated molecular shear events that can be described in terms of the motions of dislocation lines. In the meantime, experimental results have appeared that support this proposal so it is worthwhile to examine it and some of its consequences further. Previously it had been supposed that flow in glasses consisted of uncorrelated individual molecular shears as in a liquid of low viscosity (Eyring, et al., 1964)². Since the correlations between shears and, hence, the discrete dislocation lines will only persist for finite times, this mode of flow will only be important if the time duration of the plastic flow is short compared with the viscous relaxation time. Supposing that a flow event lasts 10 seconds and the applied stress is moderately high (say 10^{10} d/cm²) then a transition in the deformation mode will occur at a viscosity level of about $10 \times 10^{10} = 10^{11}$ poise. Flow in glasses more rigid than this will tend to be associated with dislocation lines. In glasses less rigid than this, the flow events will tend to be uncorrelated.

II. Nature of Dislocation Lines in Glasses

Figure 1 compares two dislocation lines, one in crystalline silica (quartz) and one in vitreous silica. In the crystalline case there is a constant Burgers vector along the line and the structure left behind as the dislocation moves is perfectly restored crystal (if the Burgers vector is a perfect translation vector of the crystal structure). In the noncrystalline case the local translations that are needed to restore the bonding vary along the line length so the Burgers vector fluctuates in magnitude and direction.

The drawings are projections onto the plane of the paper of the positions of the silicon atoms of a single sheet in the structure. The oxygen atoms are not shown, but each silicon atom is bonded to an oxygen atom that lies just above it plus three that lie below it, approximately parallel to the plane of the drawing. If parts of the upper oxygen layers are translated while the rest are not, dislocation lines can be formed at positions indicated by the dashed lines in Figure 1. The arrows represent the translations that move oxygen atoms from initial sites to equivalent final sites during an elementary motion of the dislocation line. The sheet is not necessarily flat but it will tend to lie parallel to the plane of maximum shear stress.

In addition to the fixed vs. fluctuating Burgers vectors, the two cases are different because the structure will not be perfectly restored behind a moving dislocation line in the noncrystalline case. This lack of complete restoration occurs because the distribution of Burgers vectors changes as the dislocation line moves. Consider a network with an average primitive translation vector, $\langle b \rangle$, which has a dispersion, β . Then start with a dislocation line at one position and having an average Burgers vector, $\langle b \rangle_1$, whose dispersion is β_1 . Move this line forward by an average amount, $\langle b \rangle$, to a new position with average Burgers vector $\langle b \rangle_2$ and dispersion β_2 . In general, $\langle b \rangle_1$ and $\langle b \rangle_2$ will not be the same so various defects will be left behind. The number of these defects will be proportional to the fractional change: $\frac{\langle b_2 \rangle - \langle b_1 \rangle}{\langle b \rangle}$. Or, if b is normally distributed the number of defects will be proportional to β .

The defects left behind can consist of various types that will depend on the class of glass, whether it be of the silicate, polymeric, or metallic type. In the silicate case, a glass is an open covalently-bonded network; and a metallic glass consists of a densely consolidated collection of approximately spherical atoms. In open or tangled chain networks, broken bonds (free radicals) can be left behind. In any network, density fluctuations can

be left behind, as well as dislocation loops. The deposition of these defects in the wake of a moving dislocation will produce a "structural drag" force in addition to whatever viscous drag force is acting.

The concept of dislocation lines with variable Burgers vectors is reasonable in this case because the increase in self-energy that occurs as a result of small fluctuations in b is relatively small since the energy is proportional to the square of the Burgers displacement. Suppose a unit length of dislocation has a Burgers displacement of $b + \delta$ along one-half of its length, and $b - \delta$ along the other half, where δ is a small increment. Then the ratio of its elastic energy to that of the same length without fluctuations is: $1 + \delta^2/b^2$. Thus, fluctuations as large as 30% cause only a 10% energy increase.

In the noncrystalline case illustrated by Figure 1b, the average Burgers displacement has a definite value that is determined by the network dimensions, but there are fluctuations in both its magnitude and its direction along the line. In order to minimize the energy of such a dislocation, it is necessary for the mean, b , to be conserved over long distances; thus, although the local b may fluctuate, there are long range correlations (occasional large energy densities may cause this condition to be relaxed). Furthermore, there will be little tendency for the line to lie on

a single plane, and its local structure will change as it moves. Nevertheless, it is expected that such dislocations will exist in noncrystalline solids, especially under flow conditions. When they are viewed with a somewhat fuzzy microscope (resolution of approximately 10A), their behavior should resemble that of dislocations in crystals.

Application of the concept of dislocation lines to glassy materials is not just a forced analogy with crystals, but is desirable because it allows the flow properties of these structures to be discussed in a more organized way than is otherwise possible. That is, it provides a simple means for describing the correlations that must exist between adjacent elementary shear processes in a rigid glass. It is clear that if an elementary plastic shear occurs at a particular point in a glass, then the probability that another will occur adjacent to this point is much greater than the probability that another shear will occur at some random point. The enhanced probability results from the local stress concentration and the favorable local configuration. This correlation of shear events leads directly to flow via the motions of dislocation lines.

The dislocation line description of flow in glasses becomes more useful for describing and interpreting the kinematics and dynamics of the flow process. It allows phenomena such as glide-line markings, abrupt yielding, non-linear viscous response, low-temperature internal friction,

stress relaxation, etc. to be given a microscopic interpretation. These topics will be discussed further as this paper progresses.

III. Experimental Evidence for the Existence of Dislocation Lines

Direct evidence for the existence of flow via the motion of dislocation lines in glasses has been obtained by Leamy, Chen and Wang (1971)³. They have observed distinct and relatively sharp shear-step markings on the surfaces of bent specimens of Pd₄Si glass. Such surface steps require inhomogeneous flow propagation within the material. Since the boundary of a locally sheared region is defined to be a dislocation line, this means that dislocations must move through these glass specimens.

The strain-stress curves measured by Masumoto and Maddin (1971)⁴, as well as those measured by the above authors, show definite evidence of plastic flow (as contrasted with viscous flow) followed rather immediately by fracture in glassy Pd₄Si.

Plastic flow in various silicate glasses has been observed directly by Marsh (1963)⁵ and indirectly by various authors. Although it may be inferred that such flow is inhomogeneous, this author is not aware of direct microscopic evidence.

For partially glassy polymers the existence of flow by means of dislocation motion may be inferred from the general character of the macroscopic response (Dey, 1967)⁶, and for microstructural observations that show inhomogeneous banded flow (Imada, et al., 1971⁷; Brady and Yeh, 1971⁸).

Since plastic flow via dislocation motion in glasses is both theoretically plausible and strongly supported by experimental observations, it is worthwhile to develop the kinematics and dynamics further. This is done in the next two sections.

IV. Kinematics of Flow Via Dislocation Motion

A general expression for flow via dislocation lines was derived by Gillis and Gilman (1965)⁹. The plastic strain rate in this case is given by:

$$\dot{\epsilon}_p = \langle b \rangle \oint (v \cdot n) dl \quad (1)$$

where b is the average Burgers displacement in the shear direction, v is the velocity of a line segment, n is the generalized outer normal vector of the segment, and dl is an infinitesimal line segment.

Since glasses are isotropic, the velocity vector is not expected to change much from one point to another so this equation may be simplified to read:

$$\dot{\epsilon}_p = \langle b \rangle v l \quad (2)$$

where v is the average velocity and l is the line length per unit volume.

Although the length of a dislocation line will meander in a complicated way, it can be represented as the net resultant of a collection of small loops of radius, r , and volume concentration, N . Then $L = 2\pi Nr$ and when these loops expand at velocity, v , the length per unit volume will increase:

$$\dot{L}_0 = 2\pi Nv \quad (3)$$

As the concentration of lines increases attrition will occur as a result of collisions. If the collision parameter is p , then the rate of attrition of L is:

$$\dot{L}_1 = pL^2v \quad (4)$$

so the overall rate of change is:

$$\dot{L} = 2\pi Nv - pL^2v = \alpha - \beta L^2 \quad (5)$$

Let $C^2 = \alpha/\beta$, then:

$$\dot{L} = \beta(C^2 - L^2) \quad (6)$$

Integration of this with $L = L_0$ at $t = 0$ yields:

$$\beta ct = \tanh^{-1}(L/C) - \tanh^{-1}(L_0/C) \quad (7)$$

and since $L_0 \ll C$ this may be inverted to read:

$$L = C \tanh(\beta Ct) \quad (8)$$

Thus, L increases with time to the steady-state value:

$$L_{ss} = C = (2\pi N/p)^{1/2} \quad (9)$$

and the initial slope is $\beta C = \sqrt{\alpha\beta}$. This is illustrated in Figure 2.

The behavior of Figure 2 combined with that of Equation (2) indicates that the plastic-strain-rate will increase with time at constant stress until it reaches a steady value that is determined by L_{ss} and the average velocity. Furthermore, since the collision parameter will depend on the strength of the interactions between dislocations relative to the applied stress, it will most likely vary inversely with the square of the applied stress. As a result, L_{ss} will tend to increase in direct proportion to the applied stress; and if the velocity is proportional to stress the strain-rate will increase with the square of the stress.

The discussion above assumes that the concentration of loops is fixed. However, since screw dislocations in glasses can cross-glide without difficulty it seems likely that the loop concentration will increase via multiple-crossglide. The N will be proportional to L ; making L initially proportional to L so first order multiplication occurs, followed by attrition proportional to $-L^2$. This behavior is well-known in crystals and is discussed in detail elsewhere (Gilman, 1969)¹⁰ so only the results will be given here. If k is the multiplication constant and k_2 is the attrition constant, the kinetic equation is:

$$\dot{L} = k_1 L - k_2 L^2 \quad (10)$$

This integrates to give the dependence of L on time:

$$L = \frac{k_1 L_0 e^{k_1 t}}{k_1 + k_2 L_0 (e^{k_1 t} - 1)} \quad (11)$$

According to this equation L increases exponentially at first starting from the initial value L_0 and eventually saturates at the steady-state value k_1/k_2 .

This completes this brief discussion of kinematics, and attention will be directed next to the velocity factor in the strain-rate equation.

V. Dynamics of Dislocation Motion

The forces acting on a dislocation line segment will include the following:

- a) driving force = $\tau \langle b \rangle$ where τ is the shear stress on the glide plane and $\langle b \rangle$ is the average Burgers displacement. The shear stress may derive from applied stress, internal, or both.
- b) inertial resistance = $m (dv/dt)$, where m = effective mass and (dv/dt) is the acceleration.
- c) line tension
- d) drag force
 - 1) caused by defects left in wake
 - 2) drag constant and v is the velocity

For most cases of interest only forces (a) and (d) are important. The inertial resistance is small because the effective mass is approximately ρb^2 (where ρ is the mass density) and this is a molecular mass per unit length. Thus, the time needed to overcome inertial effects is small and dislocation motions quickly approach steady-state.

Unless dislocation lines are sharply curved, line tension effects are small. In the present discussion small average curvatures will be assumed so these effects may be neglected.

Since a dislocation in a glass has no definite glide plane the appropriate shear stress is simply the maximum applied shear-stress which equals half the applied tension or compression in a uniaxial test. When multiaxial stresses are present, the root mean square or octahedral shear stress will be the appropriate driving stress.

At steady-state the driving force is balanced by equal but oppositely directed drag forces. The most fundamental of these are the viscous ones. These take many forms (Rosenfield et al., 1968)¹¹ but the most important ones act at the very core of a dislocation because this is where the maximum velocity gradient is.¹

In the case of a glass, an estimate of the viscosity coefficient at a dislocation core may be obtained by applying the free volume theory of viscosity. In order to do this an estimate of the local change in the structural free volume is needed. This can be obtained by using some results from soil mechanics as follows. This free volume is not the same as thermally induced free volume, but it does have an effect on the local viscosity and has its own temperature dependence.

The core of a dislocation is the place where the relative displacement across the glide plane changes from zero to b . In elastic theory the change is discontinuous, but in physical theory the change is more gradual and takes

place over a distance, w , that is called width of the dislocation. More precisely, the displacement along the glide plane near a dislocation line may be represented by:

$$\mu(x) = \frac{-b}{\pi} \tan^{-1} (2x/w) \quad (12)$$

As x varies from $-\infty$ to $+\infty$, $w(x)$ changes from $-b/2$ to $+b/2$ and the shear strain increases from zero far away to b/w at the center of the dislocation. The width parameter, w , may be as small as b or as large as many times b .

In an aggregate of random particles such as a glass, the dilatance that results from a given shear strain has been measured by students of soil mechanics and also predicted theoretically (Horne, 1965)¹². The dilatance is:

$$\frac{\Delta V}{V} \approx 1.4 \epsilon_{xy} \quad (\epsilon_{xy} \ll 1) \quad (13)$$

Therefore, if $w = b$, the dilatance at the center of a dislocation in a glass would be approximately 1.4. This dilatance or excess free volume will cause a reduction in the viscosity in the immediate vicinity of the dislocation core. The amount of this reduction can be estimated from the free volume theory of viscosity (Cohen and Turnbull, 1959).¹³

The Doolittle equation for the viscosity is:

$$\eta = \eta_0 e^{BV_0/V_f} \quad (14)$$

where η_0 and B are constants, V_0 is the molecular volume at $T = 0$ and V_f is the free volume. Differentiation and rearrangement of Equation (14) yields an expression for the fractional change in viscosity caused by a given fractional change in the free volume:

$$\frac{\Delta\eta}{\eta} = \frac{-BV_0}{V_f} \left(\frac{\Delta V_f}{V_f} \right) \quad (15)$$

since B is of order unity, and $V_0 \gg V_f$ this means that a small change in free volume can cause a large change in the local viscosity.

Taking $V_f = \alpha TV_0$, where α is the volume thermal expansion coefficient, plus Equation (13) with $b = w$, yields:

$$\frac{\Delta\eta}{\eta} = \frac{-1.4B}{\alpha T} \quad (16)$$

or, if η_c is the viscosity at a dislocation's core, then:

$$\eta_c = \left(\frac{1.4B}{\alpha T} - 1 \right) \eta \quad (17)$$

which gives a large reduction in the local viscosity at low temperatures and/or for small thermal expansion coefficients.

The above equations indicate one way in which flow in glasses tends to become increasingly localized at low temperatures. Dilatance left in the wake of a moving

dislocation increases the ease with which a subsequent dislocation can move through the same region. However, at high temperatures the tendency toward localized flow should disappear, as indeed it does. Equation (17) indicates that there is a critical temperature at which $\eta_c = \eta$ when the term in parentheses equals unity. Determination of this critical temperature, $T_c \approx 0.7B/\alpha$ might give a means for determining the value of B.

Dilatance will affect other physical properties in addition to the structural free volume change associated with it. For example, since the elastic stiffness arises from electrostatic forces it depends on the four-thirds power of the atomic density. If M is a particular stiffness modulus and n is the atomic density (atoms/volume) then a change in atomic density causes a stiffness change as follows:

$$\frac{\Delta M}{M} = \frac{4}{3} \left(\frac{\Delta n}{n} \right) \quad (18)$$

and a dilatance at a dislocation core of 1.4 would be expected to cause nearly a two-fold decrease in the elastic stiffness.

Another effect that needs to be kept in mind is that dilatance created by shearing a group of atoms is not distributed uniformly. It tends to be spread out on the glide plane in the glide direction. This causes anisotropy in a sheared soil and can be expected to do the same in a sheared glass.

The creation of dilatance during plastic shearing tends to soften the local structure as discussed above. Offsetting this effect to various extents is the creation of defects during the flow process. These defects cause drag on a moving dislocation and also tend to interfere with the motions of subsequent dislocations.

If the energies of the defects are known the drag stress can be readily calculated. Let δ be the density of defects per unit area left in the wake of a moving dislocation. Let U be the energy per defect; and σ be the local stress. The work done by the stress during a forward movement, x , is σbx , where b is the Burgers displacement. If $\Delta\sigma$ is the incremental stress caused by defect production then the work done by it must equal the energy needed to create the defects which is $U\delta x$. Therefore, the incremental stress is:

$$\Delta\sigma = \frac{U\delta}{b} \quad (19)$$

Suppose that $U = 1\text{ e.v./defect}$; $\delta = 10^{12}/\text{cm}^2$; and $b = 2A$. Then $\Delta\sigma \approx 10^8 \text{ d/cm}^2$. This is not a large stress, so defect drag will not have a large effect on the flow stresses of glasses unless δ is very large.

In the case of polymers it is often argued that the breakage of molecular chains (thereby leaving defects behind)

actually limits the rate of plastic flow. However, presently available data do not support this hypothesis. For example, data presented by DeVries (1971)¹⁴ show that in Nylon 6 flowing at 8.2×10^9 d/cm.² about 3.3×10^{16} free radicals are produced per cm.³ per unit strain (this being a lower limit). Now the energy needed to break a carbon-carbon bond is about 5.7×10^{-2} ergs., so the energy needed to form the observed spins is about 9×10^4 ergs/cm.³ per unit strain. But the plastic work done per unit strain is about 8×10^9 ergs/cm.³. Thus, the plastic work done is 10^5 times as large as the energy stored in measurable free radicals. The data presented by DeVries for polyethylene yield a similar discrepancy factor. Although the measurements may not observe many of the spins that are created, it is unlikely that this can account for such a large discrepancy.

VI. Implications of Dislocation Mechanics

The plastic behaviors of crystalline solids range widely from elastic-ideal plastic at one extreme to ideal visco-elastic at the other. A great variety of stress-strain, and strain-time, curves is exhibited by the multitude of solids that lie between these two extremes. Dislocation mechanics has succeeded in interpreting this range of behavior to a remarkable extent.¹ At times past it was thought that glasses are always nearly visco-elastic in their behavior,

but it is now known that they too have a wide range of behavior. Silicate glasses tend to be viscous, while metallic glasses tend to be plastic, and amorphous polymers lie somewhere in between.

Fortunately, just as for crystalline solids, dislocation mechanics is flexible enough to encompass the behavior range without straining its credibility. The reason for this is the heterogeneous nature of the processes that are involved. The macroscopic behavior is the resultant of sums and products of these heterogeneous elements. Since they are varied, the macroscopic behavior is even more varied. This is illustrated by Equation (1) or (2), which state that the macroscopic strain-rate depends on the product of the dislocation line length (and line shape) and the average velocity. But neither of these quantities is fixed. The velocity depends strongly on stress and the line length depends on the prior strain. Furthermore, the average velocity of a dislocation line represents a distribution of local velocities that can vary markedly in space and time. All of these factors lead to a great variety of net macroscopic behaviors.

Depending on the values of the constants in the appropriate equations, stress-strain curves with and without upper yield points are obtained; creep curves with and without incubation periods can be expected; and various

shapes for the transition regions of inhomogeneously drawn fibers can be predicted in much the same way as has been done for crystalline solids.

Because of its inhomogeneous nature, plastic flow leads to stress concentrations and broken chemical bonds. These in turn result in the fracturing of glasses as emphasized by Marsh.⁵ They also lead to enhanced chemical reactivity and stress-corrosion.

Serious studies of the mechanical behavior of silicate, metallic, and polymeric glasses have only just begun in terms of dislocation and disclination lines (Gilman and Li, 1970)¹⁵, but this is a most promising approach for developing an understanding of their varied behavior and for unifying this branch of the physical mechanics of solids with present knowledge of crystalline solids.

It may be possible to develop more incisive experimental knowledge by measuring internal friction during the uniaxial deformation of glasses. This might reveal the mean life times of dislocation loops and the variation of the line density with plastic strain. Studies of changes in deformation modes with temperature, stress level, and strain rate might also be revealing. It might also be possible to detect heterogeneous flow at the molecular scale by applying gold decoration techniques to the surface steps that

emerging dislocation lines must create. After the steps have been decorated, a replica of the surface can be made for study with an electron microscope.

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-71-C-0253 with The University of Michigan.

References

1. J. J. Gilman, in Dislocation Dynamics, Ed. by Rosenfield
2. et al. (McGraw-Hill, New York, 1968), pp. 3-25.
2. H. Eyring, D. Henderson, B. Stover, and E. Eyring,
Statistical Mechanics and Dynamics (J. Wiley & Sons,
New York, 1969), pp. 470-473.
3. H. J. Leamy, H. S. Chen, and T. T. Wang, Plastic Flow and
Fracture of Metallic Glass (Bell Laboratories Report, May
1971).
4. T. Masumoto and R. Maddin, *Acta Met.* 19 725 (1971).
5. D. M. Marsh, in Fracture in Solids, Ed. by Drucker and
Gilman (Interscience, New York, 1963), p. 143.
6. B. N. Dey, *J. Appl. Phys.* 38, 4144 (1967).
7. K. Imada, T. Yamamoto, K. Shigematsu, and M. Takayanagi,
J. Mat. Sci. 6, 537 (1971).
8. T.E. Brady and G. S.Y. Yeh, *J. Appl. Phys.* 42, 4622 (1971).
9. P. P. Gillis and J. J. Gilman, *J. Appl. Phys.* 36, 3370 (1965).
10. J. J. Gilman, The Micromechanics of Flow in Solids (McGraw-
Hill, New York, 1969).
11. Dislocation Dynamics, Ed. by Rosenfield et al. (McGraw-
Hill, New York, 1968).
12. M. R. Horne, *Proc. Roy. Soc. Lond.* 286A, 62 (1965), 286A, 79
(1965), 310A, 21 (1969).

13. M. H. Cohen and D. Turnbull, J. Chem. Phys. 31, 1164 (1959).
14. K. L. DeVries, J. Polymer Sci. C32 (Interscience, New York, 1971), 325-341.
15. J. J. Gilman and J. C. M. Li, J. Appl. Phys. 41, 4248 (1970).

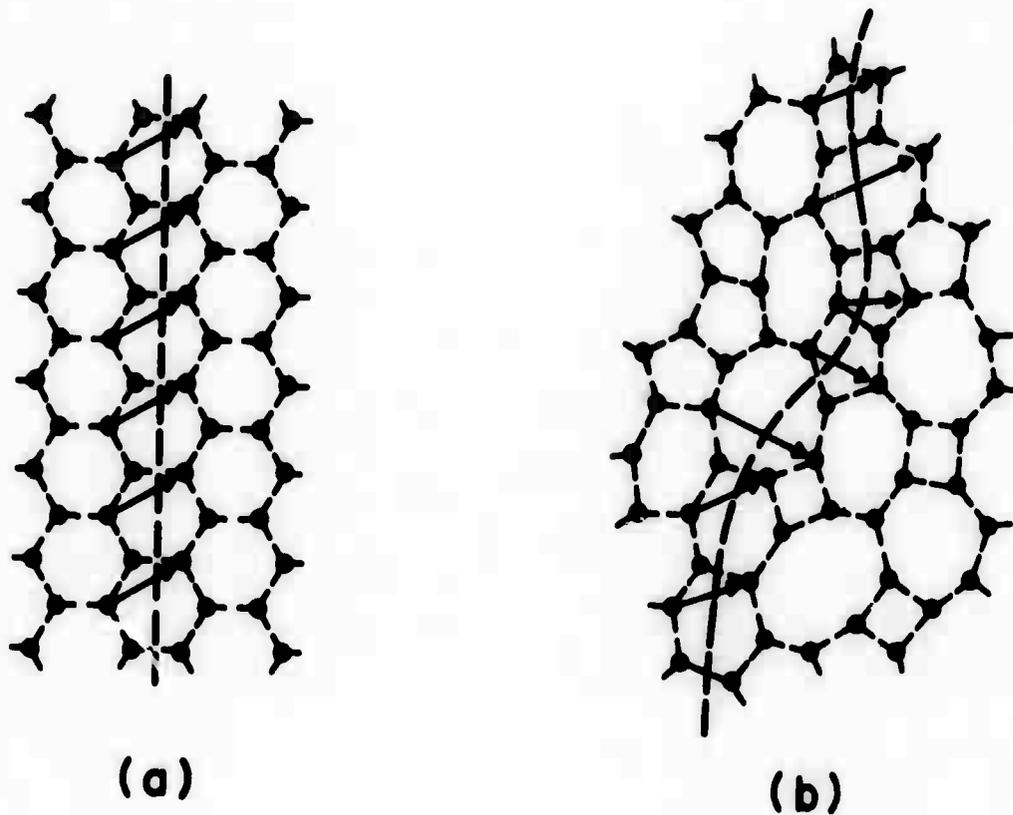


Figure 1. Dislocation lines in crystalline and vitreous silica compared. Only the silicon atoms (solid dots) as shown; each is surrounded by an oxygen tetrahedron. (a) Dislocation line has fixed Burgers vector. (b) Fluctuations in the length and direction of the Burgers vector occur in the glass structure.

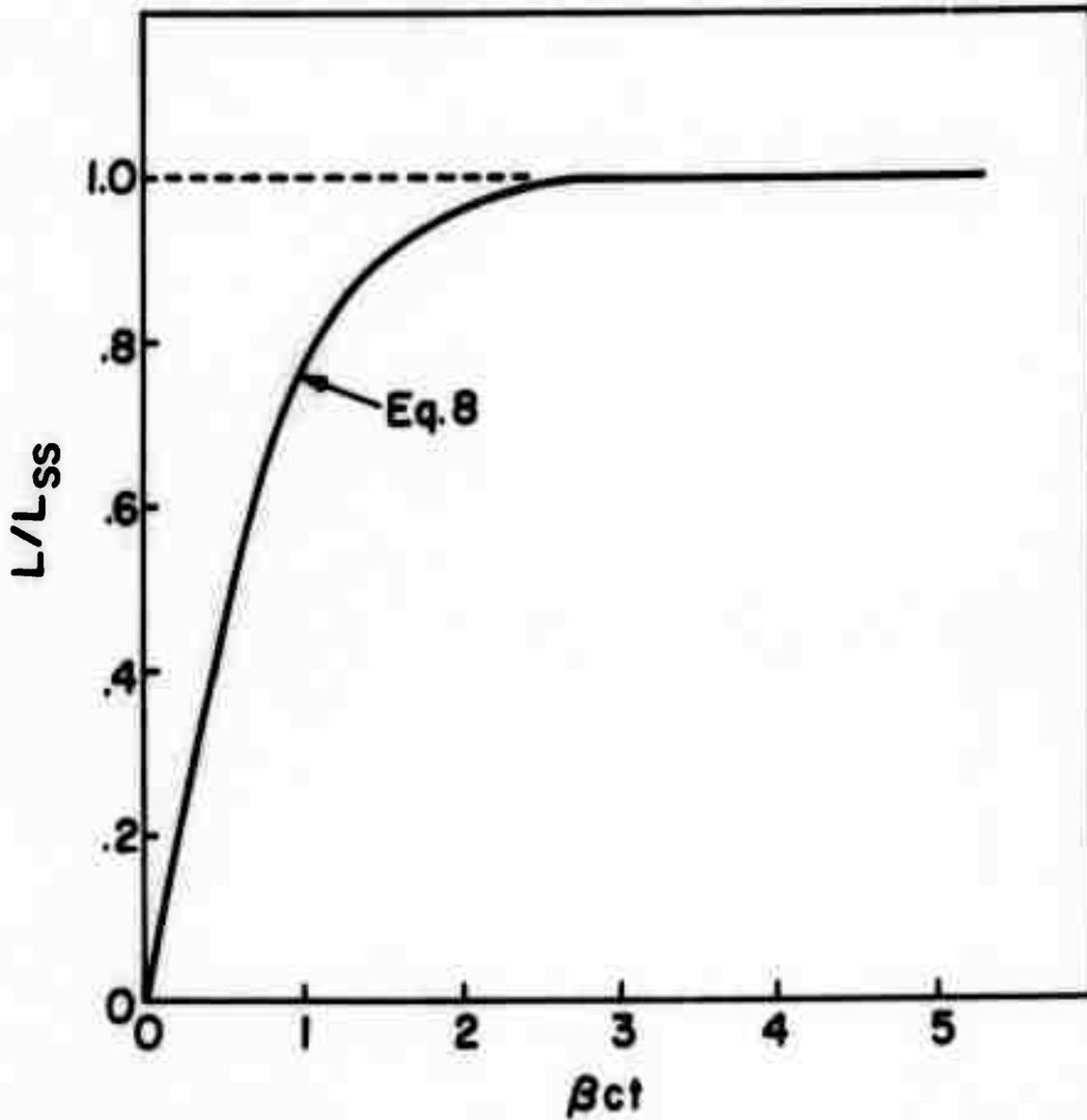


Figure 2. Dependence of dislocation line length on time assuming constant number of expanding loops and attrition through collisions.

HARDNESS - A STRENGTH MICROPROBE

J. J. Gilman

Outline

- I. Introduction
- II. Some Complexities of the Indentation Process
 - a. Friction
 - b. Pressure
 - c. Plastic anisotropy
- III. Fundamental Factors That Determine Hardness
 - a. Cohesive shear strengths
 - b. Strain-Hardening (high dislocation mobility)
 - c. Intrinsic plastic resistance (low dislocation mobility)
- IV. Hardness as a Research Tool (dynamic yielding)
- V. Suggestions for Future Work

Preceding page blank

HARDNESS - A STRENGTH MICROPROBE

J. J. Gilman

I. Introduction

Hardness measurements are at once among the most maligned and the most magnificent of physical measurements. Maligned because they are often misinterpreted by the uninitiated, and magnificent because they are so efficient in generating information for the skilled practitioner. They can quickly yield quantitative information about the elastic, anelastic, plastic, viscous, and fracture properties of a great variety of both isotropic and anisotropic solids. The tools that are used are simple and the sample sizes that are needed are typically small and sometimes sub-microscopic. This makes it unnecessary to have large specimens in order to measure strength properties and makes it possible to measure the properties of various microscopic particles within the matrix phase of a poly-phase metal, mineral, or ceramic material. This is why hardness may be considered to be a strength microprobe.

It will not be possible here to discuss the multitude of relationships between hardness measurements and strength properties. Only a few selected topics can be considered. First, some features of the test itself will

receive attention and then some of the correlations that exist between hardness and other fundamental properties.

In the first category are the effects of:

- a) friction
- b) pressure
- c) plastic anisotropy

In the second category are physical factors that determine hardness:

- a) cohesive shear strength
- b) strain-hardening (high dislocation mobility)
- c) intrinsic plastic resistance (low dislocation mobility)

Next, an example will be given of how hardness can be used as a research tool to solve a specific strength problem.

II. Some Complexities of the Indentation Process

For an isotropic, perfectly-plastic solid with frictionless surfaces, the theory of hardness indentations is relatively simple as reviewed by Tabor (1951).¹ A conclusion of the theory is that the hardness number, H , equals approximately $3Y$, where Y is the tensile yield stress which is in turn equal to $2S$, where S is the shear yield stress. Thus, $H \approx 6S$. However, many solids do not satisfy the initial assumptions. Crystals, for example, are not isotropic. Most solids exhibit strain-hardening and some transform under pressure into new solid or liquid phases. Also, the assumption

of low friction between the indenter and the specimen is doubtful at high pressure levels, especially for hard covalently-bonded solids.

a) Friction

No attempt will be made here to calculate the effects of friction on the forces involved in hardness tests. What will be done is to sketch the slip line fields for various cases in order to emphasize that the friction between an indenter and a specimen can markedly affect the plastic flow field and therefore the measured hardness number. In addition to difficult computational problems, a proper theoretical assessment of these effects requires that the friction coefficient be known at very high pressure levels. Since this is not the case, the theory is not yet completely satisfactory.

Consider two dimensional indentations as illustrated in Figure 1. Rigid blunt indenters are sketched at the top with the smooth, or frictionless, case on the left and the rough, or infinite friction, case on the right.² For a discussion of the methods used to construct these fields the reader is referred to the book by Hill (1951).³ The differences in the patterns arise because the slip lines lie at 45° to a shear-free surface and at 90° to a rigidly-clamped surface. The resulting flow fields are quite different both in form and extent. The field is more compact in the

smooth case with flow occurring immediately under the indenter except at its very center where a singular point exists at which fracture will tend to occur. In the rough case, no flow occurs immediately under the indenter so it pushes a wedge of undeformed specimen material ahead of itself.

Next, compare the slip line fields for indenters with 90° points on them as shown in the middle of the figure. Again, the field for the smooth case is more compact and it can be seen that the point of the indenter causes efficient lateral flow like the prow of a boat pushing through water. Direct lateral shear is not possible in the rough case because the slip lines adjacent to the indenter must be curved in order to satisfy the perpendicularity condition. This expands the size of the flow field.

Finally, the fields for indenters with the 135° angle of a standard Vickers diamond are illustrated in Figure 1c. Here, for the smooth indenter, the extent of the flow field is smaller than for the 90° case. This has the advantage of reducing the volume that becomes damaged during a test. In the rough case an undeformed zone is pushed ahead of the indenter so the flow field is the same as for the 90° rough indenter and considerably larger than for the smooth 135° case.

In all of the cases illustrated above the flow patterns are markedly different for the rough and smooth cases. Therefore, friction at the indenter-specimen interface can indeed have an effect on the measured hardness number. Its importance in any given case will depend on various factors, such as the strain-hardening rate of the specimen material, the depth of penetration, etc. In some cases the net result may not be significant, but the possible effects of friction should often be considered and the subject deserves further theoretical attention.

b) Pressure

Beneath some indenters the local pressure reaches quite high values. The pure pressure zones are the blank regions beneath the rough-blunt and rough-135° indenters on the right in Figure 1. The magnitude of the pressure is simply the hardness number. If this is 10,000 kg/mm² for diamond then the pressure may be as high as one megabar for this case. Such high pressures may cause phase changes to occur in some solids. These changes may consist of melting or of allotropic changes. It is well-known, for example, that the pressure created by the indentation of ice by the blade of a skate causes the ice to melt.

The indentation pressure is almost great enough to cause germanium to change phase. According to Bundy (1964), germanium transforms to its metallic form at about 115 kbar

at room temperature. This is only about twice the Vickers hardness of 600 kg/mm^2 ($\sim 60 \text{ kbar}$) so an indenter with a shape that increased the ratio of pressure to shear stress could easily cause the transformation.

In the case of silicon a standard Vickers indenter is likely to cause a transformation because hardness numbers of about 1100 kg/mm^2 are sometimes measured at room temperature, and the transformation pressure⁴ is just equal to this.

A pressure induced transformation that can be observed readily under an indenter occurs in CdS. The transformation occurs at about 24 kbar and is accompanied by a color change from orange to yellow. Therefore, if a plano-convex lens is used to indent the material and create the pressure, the color change can be observed with a microscope looking through the plane surface of the lens at the convex surface which is in contact with the specimen. The present author has done this.

c) Plastic Anisotropy

Many solids, especially crystals, are plastically anisotropic. For such a solid it may be easier to scratch a particular surface along one direction rather than another. Also, it may be easier to scratch a surface that is cut with one orientation than another. This is true even for cubic crystals because the critical shear stress for plastic flow varies from one plane to another.

The simple plastic theory applied by Tabor and mentioned previously here does not take anisotropy into account. When it is taken into account the ratio of H to Y becomes markedly changed. This will be discussed in more detail shortly.

Investigations of both the directionality effect and the surface orientation effect are often done using the Knoop type of indenter which makes a slender "diamond-shaped" impression with the ratio of the diagonals being 7:1. An excellent review and extension of the work on directionality effects has been published very recently by Brookes, O'Neill, and Redfern (1971).⁵ These authors have established a relationship between the "effective resolved-shear-stress" just below the indenter and the observed hardness number. Orientations of the long diagonal of the indenter which produce minimum values of "effective resolved-shear-stress" on the primary glide planes correspond to maximum observed hardness and conversely. The predictions of the theory of Brookes, O'Neill, and Redfern are in good agreement with observations.

Since other authors at this symposium are scheduled to discuss anisotropy in detail, I shall not dwell on the orientation dependence here; but I do wish to say more about the effect of anisotropy on the ratio between H and Y .

Nearly fifteen years ago, J. H. Westbrook (1958)⁶ did an unpublished study in which he compared measured uniaxial yield-stresses with Vickers hardness numbers. Two

classes of materials were involved: metals constituted one class, and compounds with the rocksalt crystal structure constituted the other. The results (Figure 2) show a dramatic difference in the H/Y ratios. For the metals the ratio is about 3 and for the compounds it is about 35. Westbrook did not offer a satisfactory explanation of this phenomenon, but further work (also unpublished) by Johnston and Nadeau (1964)⁷ did lead to a proper qualitative interpretation. These authors explained the difference in terms of the limited number of primary glide systems $\{110\}\langle\bar{1}\bar{1}0\rangle$ in rock-salt compounds compared with f.c.c. metals $\{111\}\langle\bar{1}\bar{1}0\rangle$. There are only 2 independent systems in the former case compared with 5 in the latter.

For an arbitrary plastic shape-change five independent glide systems are required. Thus the deformation caused by an indenter in a f.c.c. metal can occur entirely on primary systems; whereas it cannot for rock-salt type compounds. For these compounds, flow on secondary glide systems is required so the effects of plastic anisotropy are strongly manifested. The stress for flow on the secondary system in LiF (for example) may be as much as 100 times the stress for flow on the primary system (Gilman (1959)).⁸

I am indebted to Professor J. R. Rice of Brown University for elucidating the mechanics that are involved in a simplified model. This model is two-dimensional and

therefore is intended only to be illustrative of principles. It should not be subjected to quantitative comparisons with experiments in its present form.

Consider two-dimensional indentation by a blunt-smooth punch. The yield surface is sketched in Figure 3 and the slip-line field is outlined in Figure 4. In Figure 3 it is assumed that there are two glide systems operating; and only two. One glide system is (010)[100] with yield stress = b; and the other is (110)[$\bar{1}\bar{1}0$] with yield stress = a.

Figure 4 (top) outlines the slip line field which has seven regions. It also indicates (bottom) the stress states within each of the regions. The pressure on the indenter that is needed to just begin flow in the over-all configuration is:

$$P \text{ (limit)} = 4a + 2b$$

Since a is the critical flow stress of the primary glide system it is the major factor in this equation. If flow occurs with equal ease on the two systems, $a = b$ and $P = 6a = 3Y$ in accordance with the isotropic theory. However, if flow is difficult on the secondary system (say $b = 10a$; so $P = 24a$) then the ratio of H to Y increases to 12. In this way the ratio of 35 that Westbrook observed can be accounted for if $b = 66a$ which is consistent with the observations mentioned above for LiF crystals.

III. Fundamental Factors That Determine Hardness

If materials were perfectly homogeneous the size of the indenter used in a hardness test would not influence the measured hardness. In reality materials are not homogeneous, and plastic flow in them is localized at dislocation lines. Therefore, when a sufficiently small indenter makes contact with a solid, the region in which the stress is high may be so small that it does not contain any dislocations or other major defects. Therefore, flow cannot begin until dislocations have been nucleated under the indenter, and the factor that may determine the hardness for very small indenters is the cohesive shear strength of the material.

a) Cohesive Shear Strengths

The converse of the statement just above is that plastic indentations may be used to measure the cohesive shear strengths of materials. This was first attempted by the author (Gilman, 1959)⁹ and more recent studies have been made by Gane and Bowden (1968).¹⁰

The method used by Gilman consisted of touching crystals with spherical glass indenters in a carefully controlled fashion. Extremely clean spherical surfaces can be produced by drawing out Pyrex glass rods, breaking them, and then fusing the ends. Spherical surfaces that range in radius from 0.05 mm up can be produced this way. After such a rod is mounted suitably, a microhardness balance can be

used to accurately apply a load to it ranging from 1 g to 200 g. This allows the indenter to be moved smoothly into contact with a crystal at a rate of $\sim 10^{-3}$ cm/sec.

In order to avoid dislocation nucleation under a small spherical indenter Gilman showed that it is necessary to:

(a) apply loads less than some critical value, (b) contact dislocation-free areas of the crystal surfaces, (c) use freshly prepared indenters, (d) not re-use an indenter that had once produced dislocation nucleation, and (e) minimize the radius of the indenter to reduce the size of the stressed volume. With these precautions a load as high as 10 g can be applied to an indenter of 52μ radius in contact with LiF without causing dislocation nucleation. This should produce a maximum shear stress of about 7.4×10^9 d/cm² according to the Hertz theory of elastic contact. This stress is more than two orders of magnitude greater than the yield stress of the crystals ($\sim 5 \times 10^7$ g/mm²). Since the shear modulus for LiF is about 6.25×10^{11} d/cm, this maximum stress is about 1/85 of the shear modulus. It was suggested that this method, after some refinement, can be used to measure the stresses needed for homogeneous dislocation nucleation in crystals.

In the work of Gane and Bowden, small indenters of TiC were made by electropolishing rods down to fine tips of $\sim 500\text{A}$ radius. These were loaded electromagnetically inside

an electron microscope so the indentation process could be observed in situ. The results were similar to those described for LiF, but were made on crystals of gold, copper, and aluminum. No penetration occurred until a critical load was reached, and this critical load frequently corresponded to the cohesive shear strengths of the metals.

On the basis of both sets of experiments it may be concluded that solids which are ordinarily soft become highly resistant to penetration if mobile dislocations are not present. Penetration only occurs when the load reaches a value that is high enough to create the stresses needed for homogeneous dislocation nucleation.

b) Strain-Hardening (high dislocation mobility)

If a crystalline solid contains numerous dislocations its yield stress can vary over a wide range that may include several orders of magnitude. If it is a metal or a salt, impurity defects largely determine whether it is hard or soft. If it is a covalent compound, the intrinsic strengths of its chemical bonds determine its hardness.

Yield stresses are usually measured in a uniaxial stress test that does not constrain the flow pattern. Such a test does not model the situation under an indenter where a complex and highly constrained flow pattern exists. From the slip-line patterns of Figures 1 and 4 considerable intersecting glide must occur and therefore much strain-hardening.

As a result, two limiting cases arise. In one case dislocation mobility is high, so the flow rate becomes limited by strain-hardening. In the other dislocation mobility is low, and it is the factor that limits the flow rate.

The two limiting cases are illustrated in Figure 5. The lower line correlates data for pure face-centered-cubic metals. These metals are particularly suited to the type of correlation that is given here for two reasons. First, being noble metals, they are readily purified so that measurements on quite pure samples are available. Second, the dependence of their yield stresses on impurity content, and on temperature, is quite small and, hence, large errors due to these factors are unlikely.

Individual dislocations move readily in these metals at much lower stresses (a factor of 10^4 lower) than the stresses of Figure 5. Therefore, the flow that occurs during indentation is not being limited by the drag on isolated dislocations. Intersections between dislocations create jogs on them and these jogs create trails of dislocation dipoles behind the moving dislocations. Soon the structure becomes filled with interacting dislocations and further flow is limited by the strengths of the interactions. Since the interactions are elastic, for a given geometric configuration their strengths are proportional to the elastic moduli. In this way the linear proportionality of Figure 5 arises.

c) Intrinsic Plastic Resistance (low dislocation mobility)

Figure 5 illustrates the point that the behavior of covalently bonded crystals is qualitatively different from that of metals. First, the ratio of hardness to elastic modulus is much greater (being about 0.24 for diamond compared with 0.3×10^{-2} for rhodium). Second, the slopes of the correlation lines are different (being 1.5 for the covalent crystals compared with 1.0 for the f.c.c. metals).

The qualitative difference in behavior between these two classes of material arises because the flow-rate is limited in the metals primarily by interactions between dislocations, whereas it is limited in covalent crystals primarily by interactions between dislocations and the chemical bonds of element or compound. Another way of saying this is that short range (atomic) interactions are involved in the latter case; and longer range (elastic) interactions in the former. This manifests itself in the temperature dependence which is very weak for a f.c.c. metal and is very strong for a covalent crystal above a critical temperature.

The basic reason for the difference in behavior is simply that bonding in covalent compounds is highly stereospecific whereas in f.c.c. metals only a small fraction of the binding energy depends on stereographic factors. Therefore, the energy of a dislocation in a metal is nearly independent of its position. In contrast, a dislocation's energy depends strongly on its position in a covalent crystal

because an unpaired electron is associated with it when it is in its mid-glide position.

There is no space here to discuss details of dislocation structures in covalent crystals which involve (for the diamond or wurzite crystal structure) edge, screw, and 60° lines as well as positive and negative kinks on each type of line. Regardless of the details, the basic fact remains that in order to plastically shear such a crystal, electron-pair bonds must be first broken and then remade. Thus, two unpaired electrons exist each time an atomic shear process is half completed. Because of this it may be expected that the plastic characteristics will be closely connected with the electronic characteristics of covalent crystals. This is indeed the case as shown in Figure 6 where measured values of the activation energy for plastic glide (Chaudhuri, Patel, and Rubin, 1962)¹¹ are shown to equal twice the optical band gap. This is consistent with the statements above because breaking an electron-pair bond (inside a crystal) means that the two electrons become excited from the valence band to the conduction band so the energy required is the gap energy taken twice. This is also consistent with the tunneling model for dislocation mobility (Gilman, 1955)¹² in which low temperature motion is determined by the rate of electron tunneling across the gap.

Hardness correlates with glide activation energy as shown in Figure 7 so it would be expected that hardness should correlate with the energy gap. However, the units of hardness are energy/volume so it is the energy gap per unit volume that should be correlated and this is done in Figure 8. A similar correlation is shown in Figure 9 for III-IV compounds. In both cases the correlations are good but they have different slopes. This may result from the greater ionicity in the case of the III-IV compounds.

Since ionic bonding results from long range electrostatic forces it does not contribute directly to hardness. In this respect it is similar to metallic bonding. Therefore, it is appropriate to separate the homopolar and heteropolar components of the bonding in compounds in order to see whether hardness correlates better with the homopolar component than with the total bonding. One way of doing this is shown in Figure 10, where the data of Figure 9 are replotted. Here the homopolar energy gap as defined by Phillips (1970)¹³ is used instead of the optical band gap. The data correlate quite well and now the slope is $3/2$ which is consistent with Figures 5 and 7. This is also consistent with the well-known qualitative fact that for pure crystals the greater the homopolar component of bonding the harder the crystal.

Another correlation that demonstrates the close connection between hardness and bond strength is the

dependence on bond length since bond lengths and bond energies are closely related; especially for a homologous series of compounds. This type of correlation was discussed long ago by Goldschmidt (1926)¹⁴ and reviewed by Evans (1946)¹⁵ in his book. For group IV crystals and III-V compounds it has been demonstrated by Wolff et al. (1958).¹⁶ For II-VI compounds it has been discussed by Cline and Kahn (1963).¹⁷ Figure 11 shows the data for the III-V compounds.

Because of the clear connections between chemical bonding, hardness, and dislocation mobilities, an attempt has been made here to relate the mineralogist's Moh scale of hardness to physical properties. The first step was to relate indentation hardness with Moh hardness and the result of this is shown in Figure 12. This figure represents data averaged from various sources including: Taylor (1949),¹⁸ Khrushchov (1949),¹⁹ Winchell (1945),²⁰ and Knoop et al. (1939).²¹

The indentation hardness values for the Moh scale minerals are plotted in Figure 13 versus the bulk moduli of these minerals. The bulk moduli were obtained from Clark (1966).²² The correlation is only fair, but certainly good enough to indicate that the Moh scale is based on relative bond strengths. The data for quartz and orthoclase are anomalous. The reason for this is unknown at present.

The next class of materials whose hardness can be given a physical interpretation is the "hard metals." That

is, compounds consisting of a transition metal plus a metalloid such as B, C, or N. These compounds are metals in the sense that they conduct electricity readily, but they are covalent compounds in the sense that they are not ductile, but brittle. This mixed behavior is possible because the conduction mainly occurs in the s-band of energy, whereas much of the cohesion occurs in the d-band.

One large class of "hard metals" has the rock-salt structure. This class contains some 80-90 hydrides, borides, carbides, nitrides, and oxides. They tend to glide on the {111} plane and a cutaway schematic drawing of this is shown in Figure 14. According to the bonding theory worked out by Lye (1967)^{2 3} for TiC an important contribution to the bonding arises from overlapping of the d-orbitals of the Ti atoms at the octahedral holes in the f.c.c. titanium array. The carbon atoms at these sites have six titanium neighbors and, therefore, six overlapping d-orbitals. The overlap allows one or more electrons of the carbon atoms to reduce their kinetic energies by moving in the d-band of the Ti sublattice. Also, the overlapping d-orbitals are stabilized by the positive potentials of the carbon nuclei, so a mutually stabilizing cooperation occurs between the two sublattices.

It may be seen in Figure 14 that there are two types of site in the plane of the carbon atoms. The A-sites are occupied by carbon atoms and overlapping orbitals in the figure and the B-sites (at the center of the triad defined by the shaded circles) are unoccupied. Now in order for glide to occur in the structure, carbon atoms must glide from A-sites to B-sites and then back to A-sites. Since there are no overlapping orbitals at B-sites the carbon atoms must lose their excess binding energies (irreversibly) and then regain them when they return to A-sites. This means that strong forces resist the motions of dislocations in these compounds. The stress needed to overcome the chemical resisting force has been estimated by Gilman (1970)²⁴ to be: $\sigma = 2\Delta F/V$ where ΔF is the free energy of formation of the compound and V is the molecular volume.* The approximate validity of this

*An improved version of his argument is as follows. The barrier to be overcome is the energy per bond of carbon; or $\Delta F/6$, where 6 is the coordination number. The work done by the applied shear stress, S , in overcoming the barrier is $(Sb'w)b'$ where b' is the partial Burgers displacement which equals $b/\sqrt{3}$ in this structure, and w is the length of the dislocation segment. Equating these terms yields an expression for the local shear stress: $S = \Delta F/2b^3$. Because of the constraint factor the hardness, H , equals $6S$; or since $b^3 = \sqrt{2}V$ the hardness equals $2.14\Delta F/V$ which is nearly the same as the previous result.

relationship is demonstrated for carbides in Figure 15, and it is also approximately valid for the refractory diborides.

IV. Hardness as a Research Tool (dynamic yielding)

Hardness measurements are used in all aspects of structural materials research as this symposium attests. Its use in alloy development, measuring creep resistance, studying environmental effects, etc. will be discussed in detail by other authors. I shall take just one example of an engineering application in which it was useful because it could partially replace a much more complex measurement. My example comes from the recent program to develop improved ballistic armor.

The first step in stopping a high-speed armor-piercing projectile is to blunt its tip so that it cannot "knife" its way through a piece of armor. In order to cause such blunting the dynamic yield stress of the armor must be greater than about 40 kbar (corresponding to a hardness of about 1000 kg/mm^2).

The dynamic yield stress is measured by creating an intense plane wave (uniaxial strain) in the test material and then observing its profile as it propagates. The behavior is illustrated in Figure 16. At the left the initial state of an element of material is shown. In the middle the element has been compressed uniaxially by the wave. This changes its volume and also puts shear strain into it. On

the right the volume change remains, but the shear strains have relaxed. The stress level at which shear relaxation occurs is called the dynamic yield stress. For hard compounds it correlates well with static indentation hardness as shown in Figure 17 (the data are a courtesy of C. F. Cline). Thus, hardness can be used as an initial screen for candidate armor materials. Knowledge of the chemical factors that determine hardness can, in turn, be used to predict new materials such as Cline did for Be₃B.

V. Suggestions for Future Work

On the basis of the discussion presented here, future studies that use indentation hardness need to be supported by a careful study of the role that friction plays in the hardness test. Also, the quantitative effects of plastic anisotropy need study in order to predict the large effects it has on measured hardness values.

The chemistry of hardness needs systematic and comprehensive study in order to provide more understanding followed by predictive capability. For example, what role is played by the band gap energy? Is the homopolar component of the binding a dominant factor? Is electronegativity difference an important variable? How do hard covalent compounds differ from hard metals? Why do quartz and

orthoclase appear to behave anomalously? Can correlations between hardnesses and heats of formation be improved? What is the effect of pressure on hardness?

Acknowledgement

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAH15-71-C-0253 with The University of Michigan.

References

1. D. Tabor, The Hardness of Metals, Oxford, Clarendon Press (1951).
2. The author was aided in the construction of these slip-line fields by discussions with Professor E. H. Lee.
3. R. Hill, The Mathematical Theory of Plasticity, Oxford, Clarendon Press (1951).
4. F. P. Bundy, "Phase Diagrams of Si and Ge to 200 kbar, 1000°C," J. Chem. Phys. 41, 3809 (1964).
5. C. A. Brookes, J. B. O'Neill, and A. W. Redfern, "Anisotropy in the Hardness of Single Crystals," Proc. Roy. Soc. Lond. 322A, 73 (1971).
6. J. H. Westbrook, "Flow in Rock-Salt Structures," Report #58-RL-2033 of the G.E. Research Laboratory (August 1958).
7. W. G. Johnston, "Deformation Mechanisms of Refractory and Other Materials," Report #ARL 64-135 of the Aerospace Research Laboratories, OAR, U.S. Air Force.
8. J. J. Gilman, "Plastic Anisotropy of LiF and Other Rock-Salt-Type Crystals," Acta Met. 7, 608 (1959).
9. J. J. Gilman, "Dislocation Sources of Crystals," J. Appl. Phys. 30, 1584 (1959).
10. N. Gane and F. P. Bowden, "Microdeformation of Solids," J. Appl. Phys. 39, 1432 (1968).
11. A. R. Chaudhuri, J. R. Patel, and L. G. Rubin, "Velocities and Densities of Dislocations of Germanium and Other Semiconductor Crystals," J. Appl. Phys. 33, 2736 (1962).
12. J. J. Gilman, "Escape of Dislocations from Bound States by Tunneling," J. Metals 18, 1171 (1966); J. Appl. Phys. 39, 6068 (1968).
13. J. C. Phillips, "Ionicity of the Chemical Bond in Crystals," Rev. Mod. Phys. 42, 317 (1970).
14. V. M. Goldschmidt, Norske Vid. Akad. i Oslo Skr. Mat. Nat. Kl. 8, p. 102 (1926).

15. R. C. Evans, An Introduction to Crystal Chemistry, Cambridge Press (1946).
16. G. A. Wolff, L. Toman, N. I. Field, and J. C. Clark, Semiconductors and Phosphors, Ed. by Schoen and Welker, Wiley, Interscience, New York (1958), p. 463.
17. C. F. Cline and J. S. Kahn, "Microhardness of Single Crystals of BeO and Other Wurzite Compounds," J. Electrochem. Soc. 110, 773 (1963).
18. E. W. Taylor, Miner. Mag. 28, 718 (1949).
19. M. M. Khrushchov, Zavod. Labor. 15, 243 (1949).
20. H. Winchell, Amer. Min. 30, 583 (1945).
21. F. Knoop et al., Jour. Res. Bur. Stand. Wash. 23, 39 (1939).
22. S. P. Clark, Jr., Ed., Handbook of Physical Constants, Memoir #97, The Geological Society of America, Inc. (1966).
23. R. G. Lye, "Band Structure and Bonding in Titanium Carbide," Atomic and Electronic Structure of Metals, p. 99, ASM, Metals Park, Ohio (1967).
24. J. J. Gilman, "Hardness of Carbides and Other Refractory Hard Metals," J. Appl. Phys. 41, 1664 (1970).

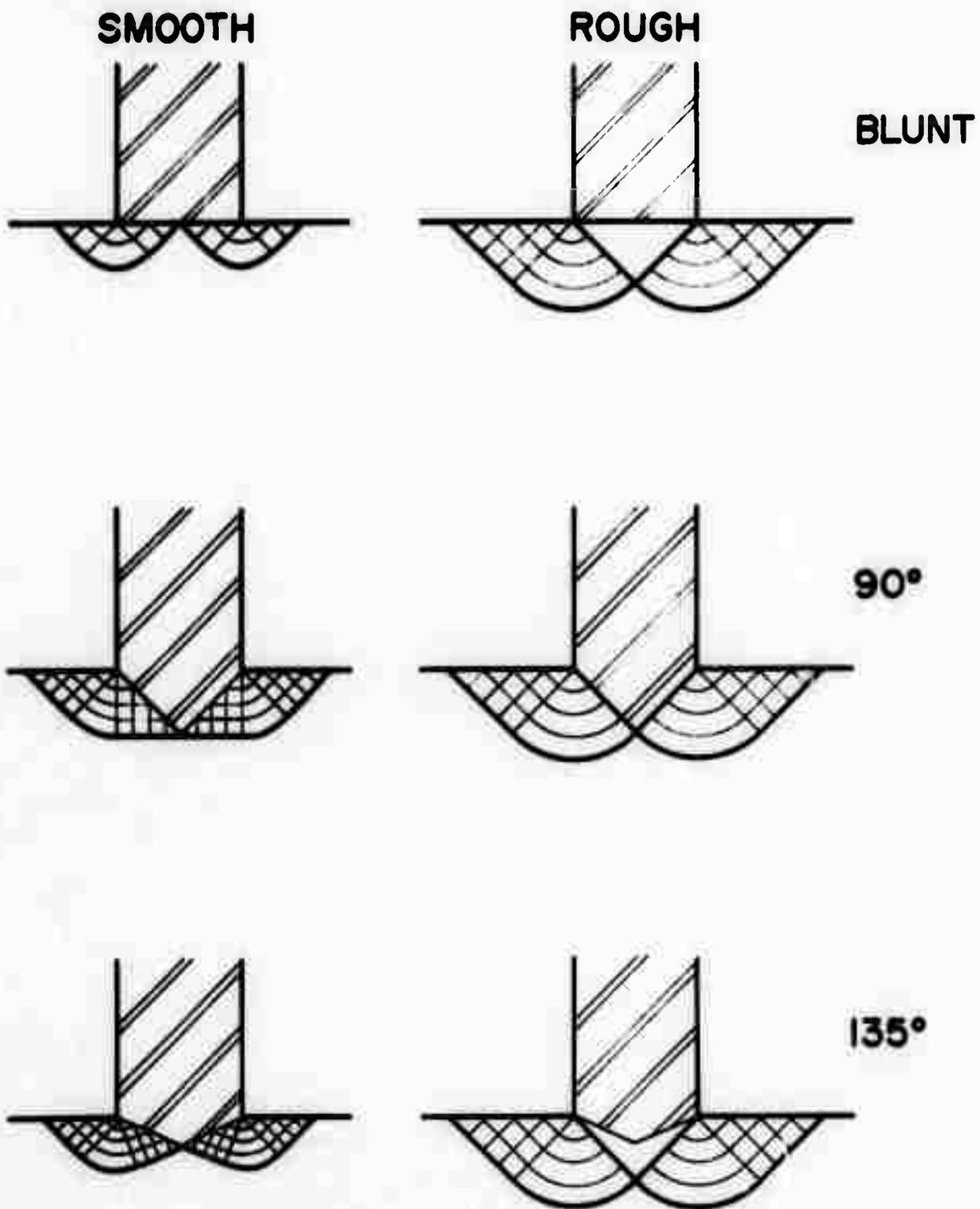


Figure 1. Effect of Friction on Plastic Flow Patterns under Rigid Indenters (two-dimensional).

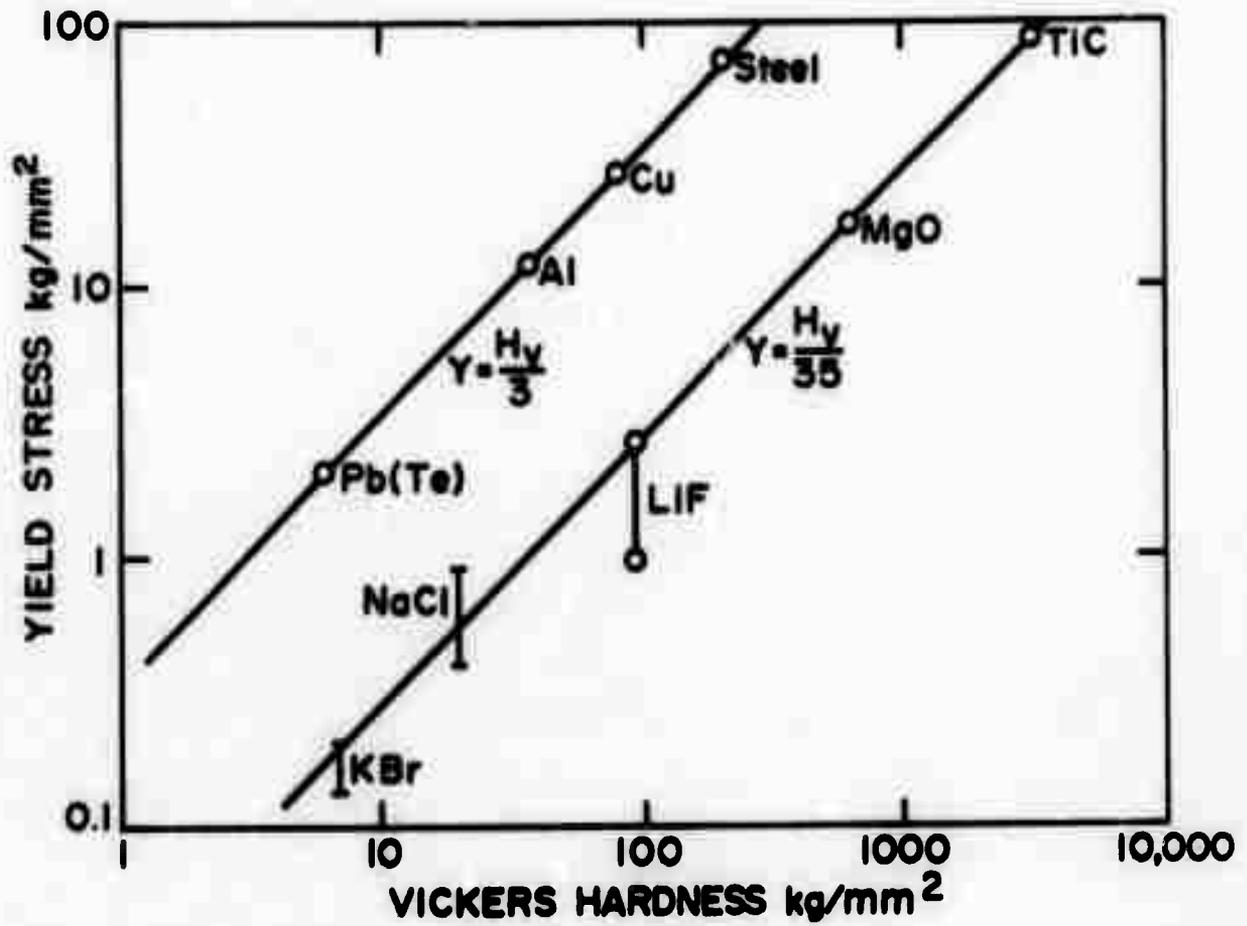


Figure 2. Correlation Lines for Metals and Ionic Crystals Relating Hardness Numbers and Yield Stresses (after J. H. Westbrook)

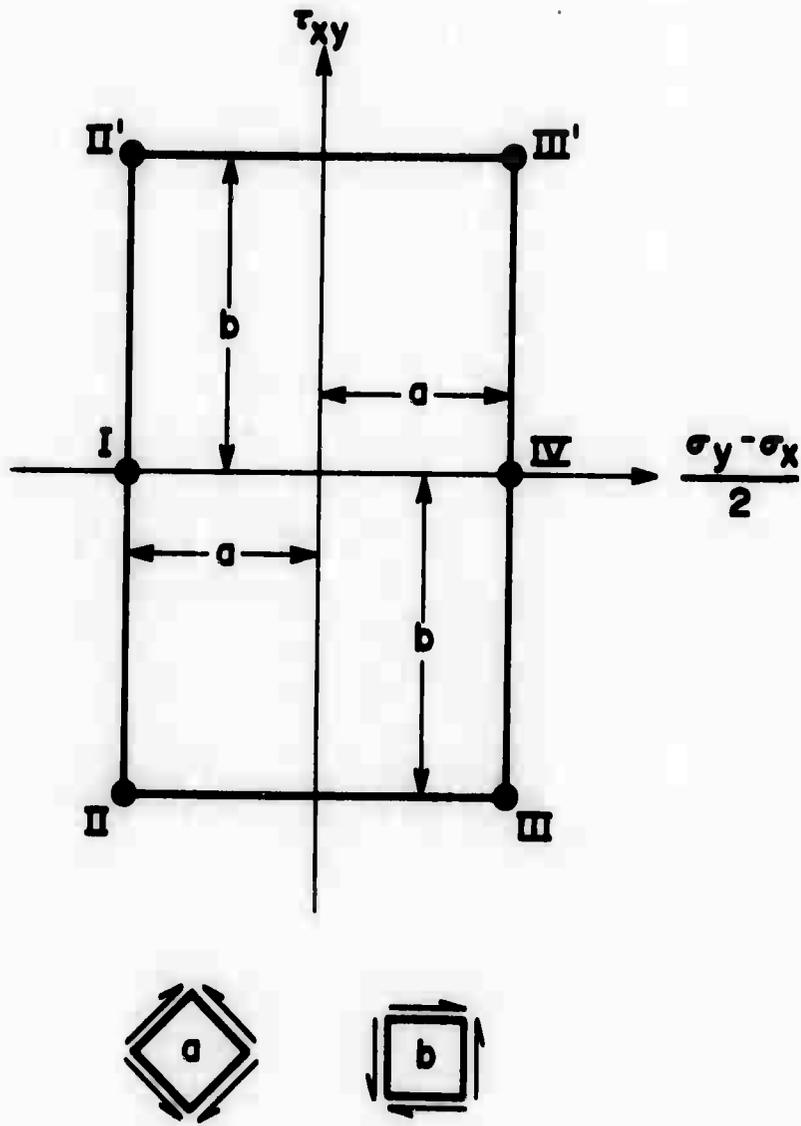


Figure 3. Two-dimensional Yield Surface for an Anisotropic Ideal Plastic Material

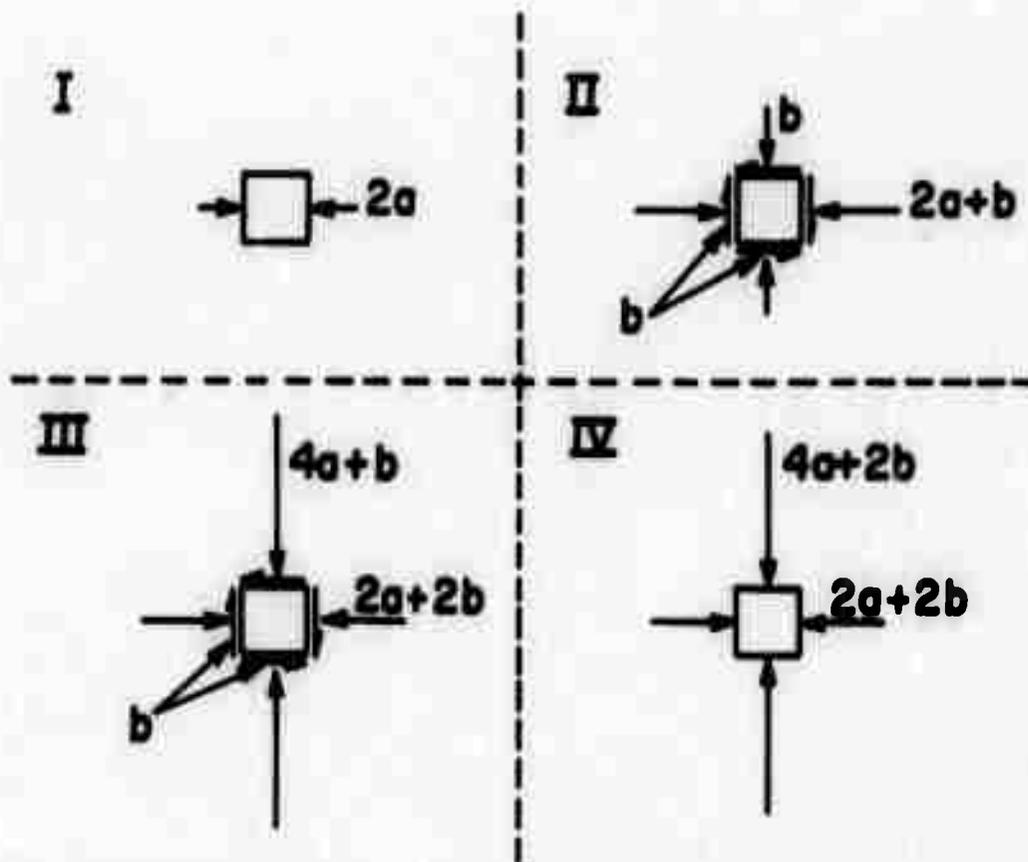
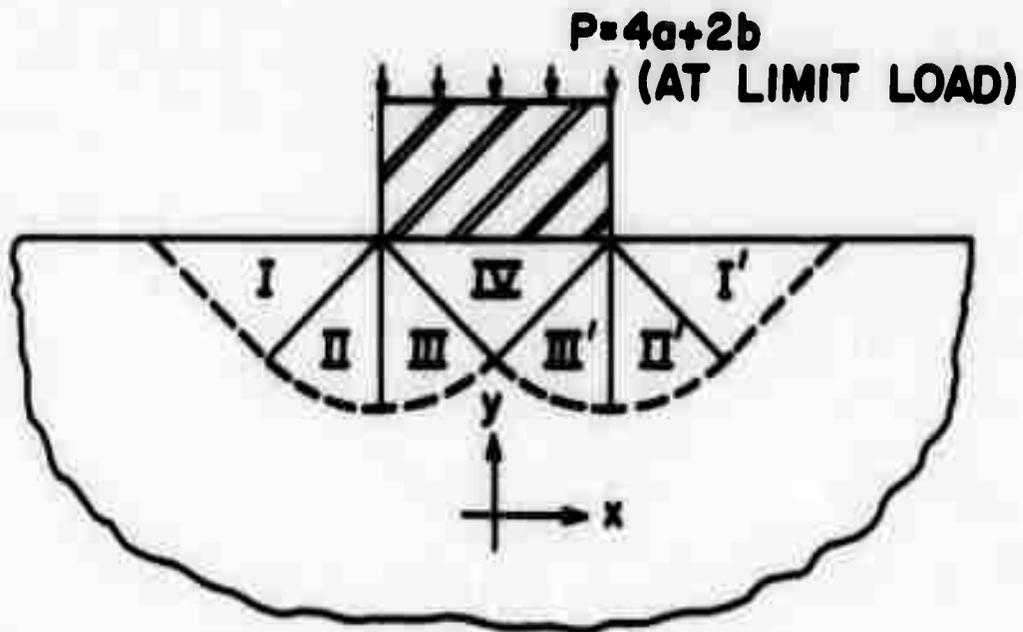


Figure 4. Stress Distributions Under a Two-dimensional Punch at the Limiting Load for General Plastic Flow (after J. R. Rice)

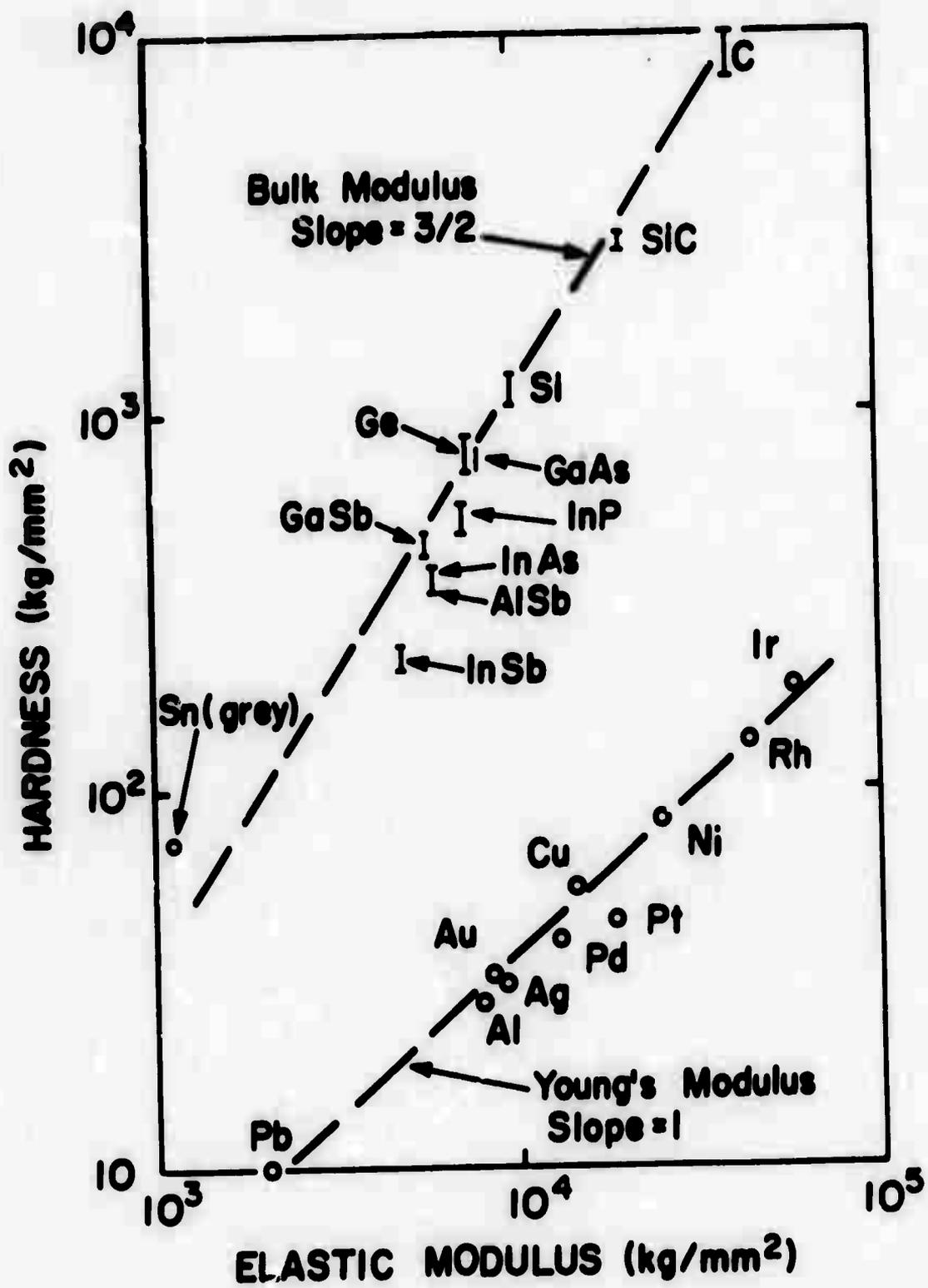


Figure 5. Correlation of Hardnesses and Elastic Moduli for Pure f.c.c. Metals and Covalent Crystals

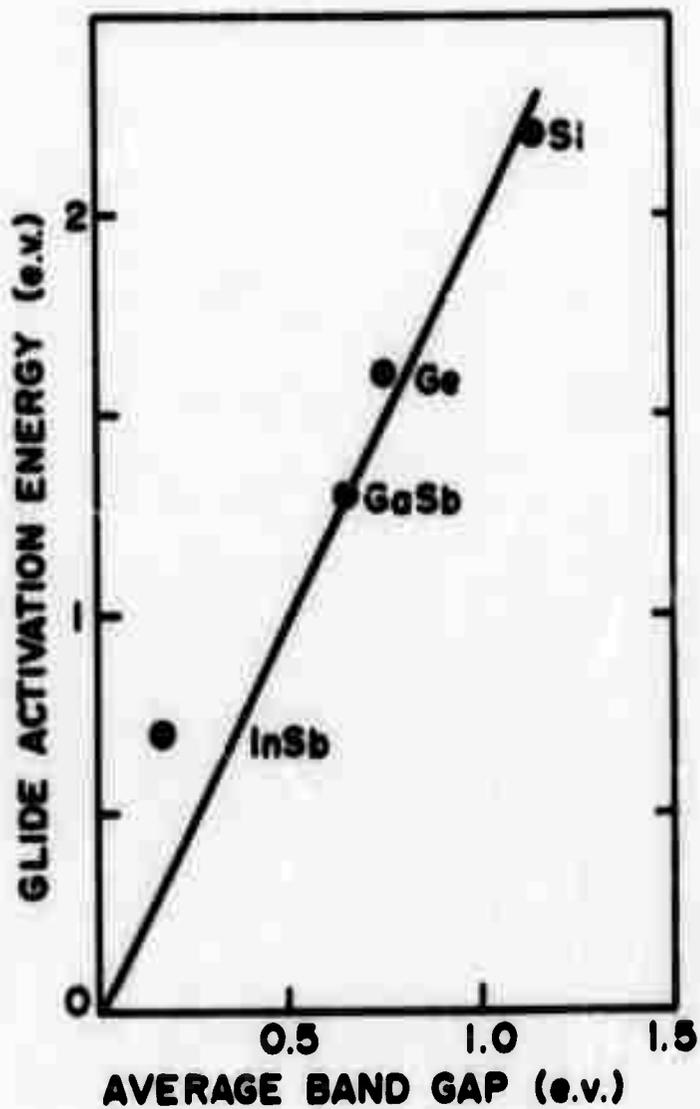


Figure 6. Relationship Between Glide Activation Energies and Average Band Gaps for Covalent Crystals

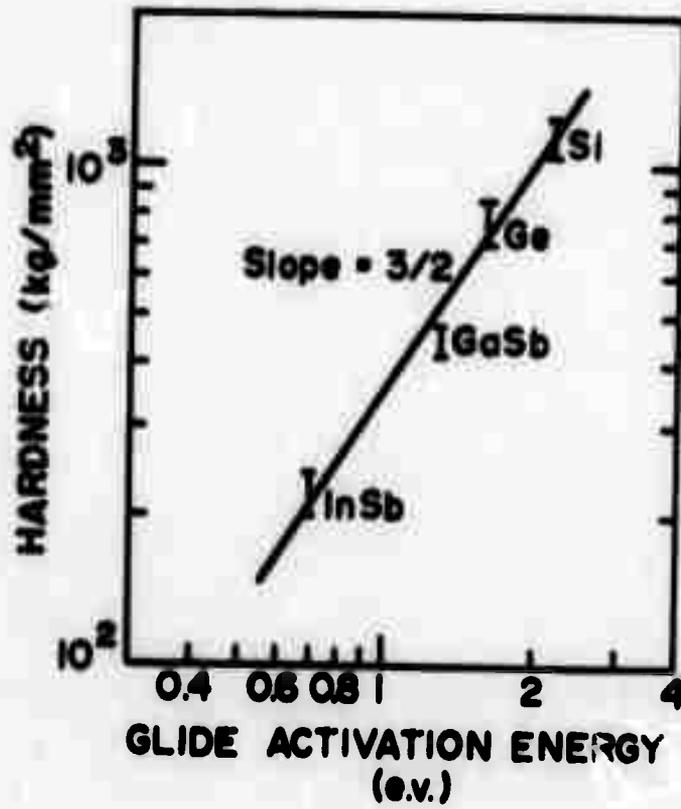


Figure 7. Correlation of Hardness and Glide Activation Energy

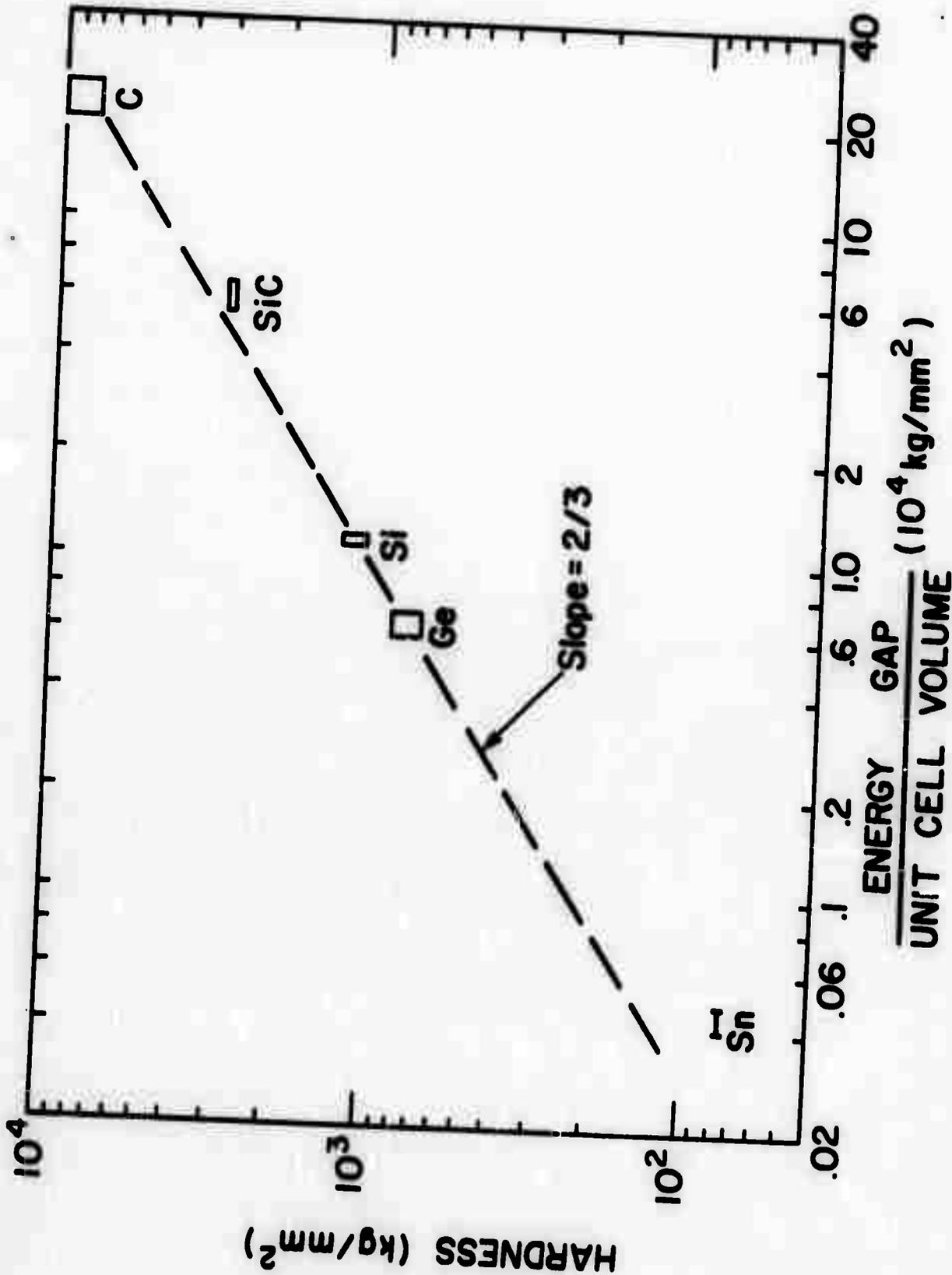


Figure 8. Correlation of Hardness and Energy Gap Density for Group IV Crystals

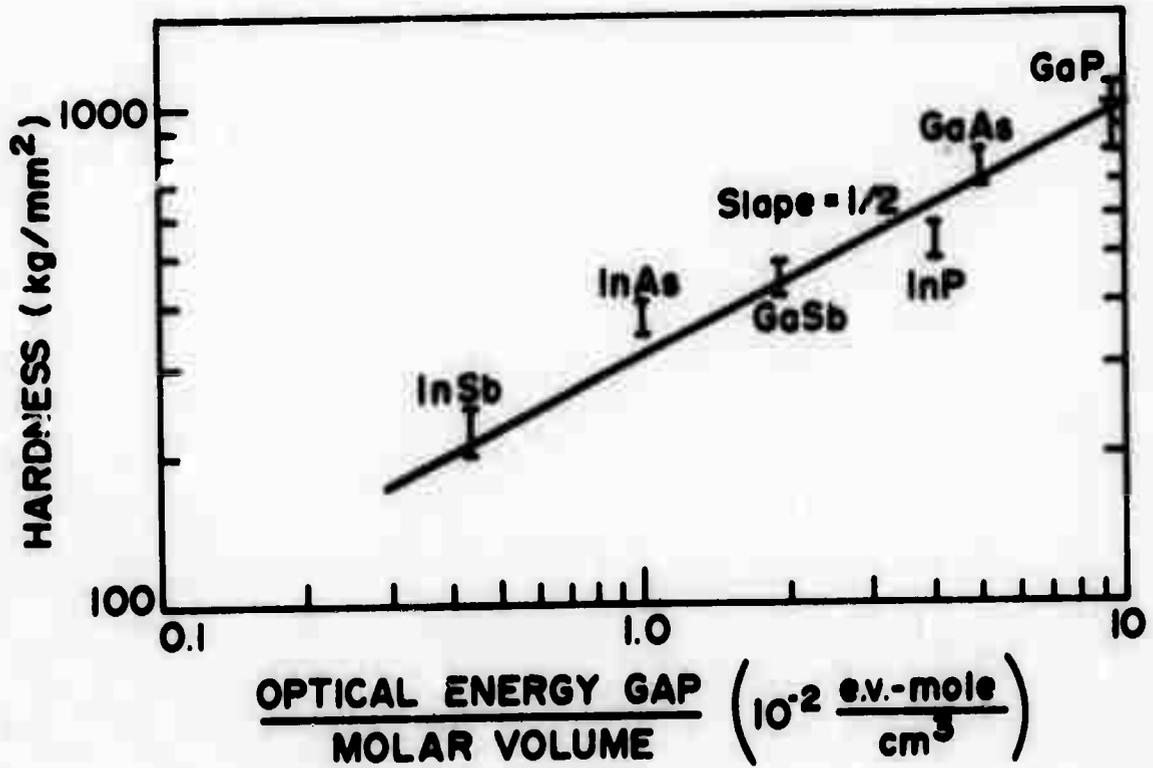


Figure 9. Correlation of Hardness and Energy Gap Density for Group III-V Compound Crystals

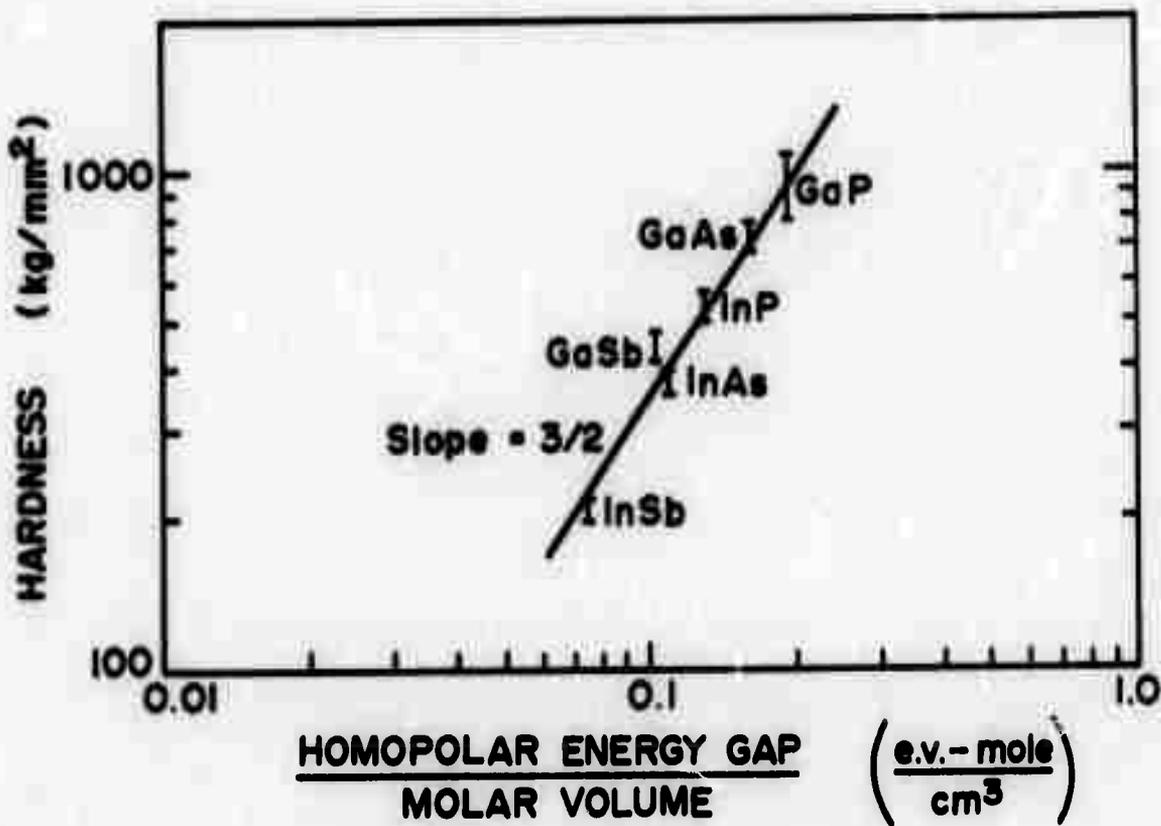


Figure 10. Correlation of Hardness and Energy Gap Density for Group III-V Compound Crystals Using the Homopolar Part of the Energy Gap

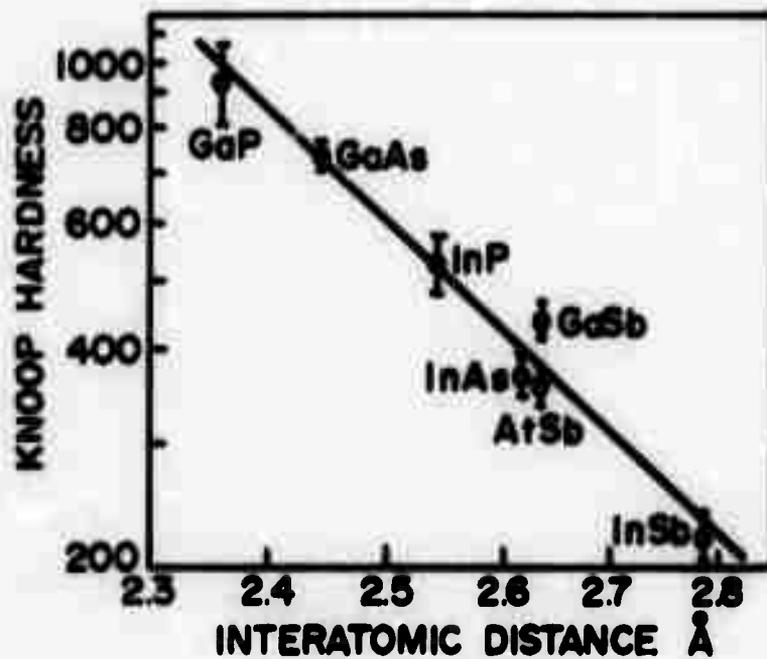


Figure 11. Correlation of Hardness and Bond Distance for III-IV Compound Crystals

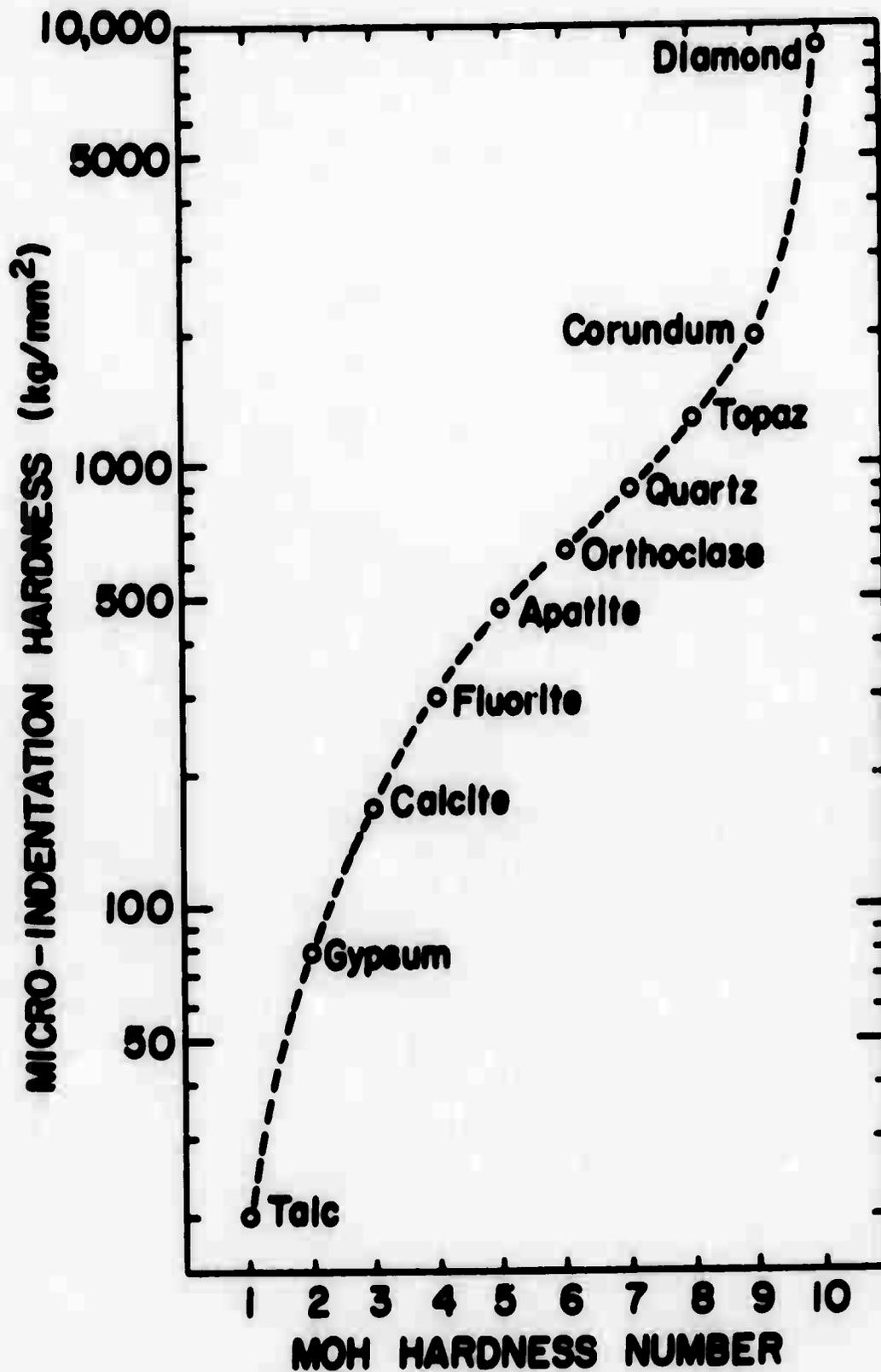


Figure 12. Correlation of Moh Scratch Hardness Numbers and Indentation Hardness Numbers

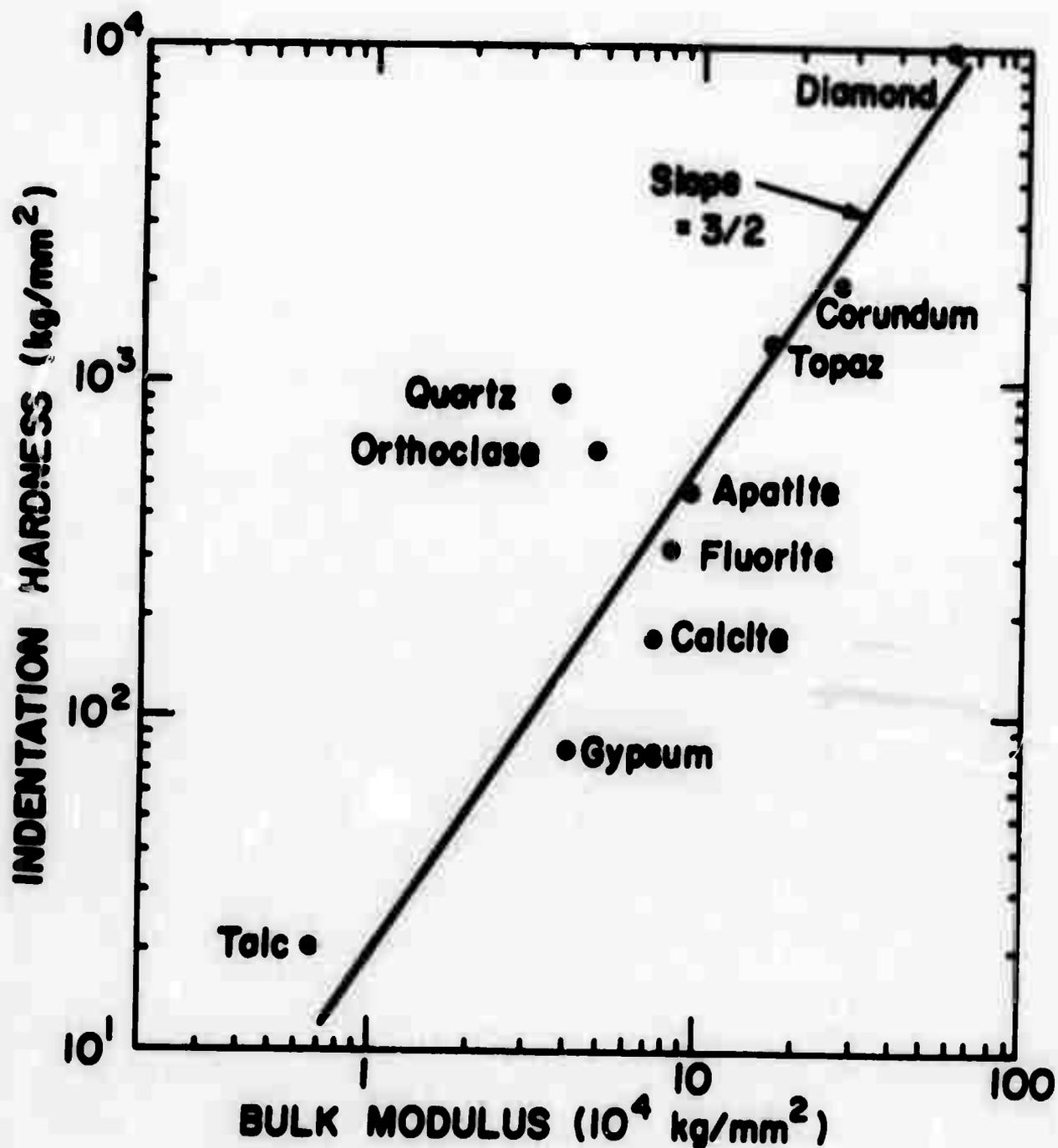


Figure 13. Result of Using Figure 12 to Relate Hardnesses of the Moh Scale Minerals with Their Bulk Elastic Moduli

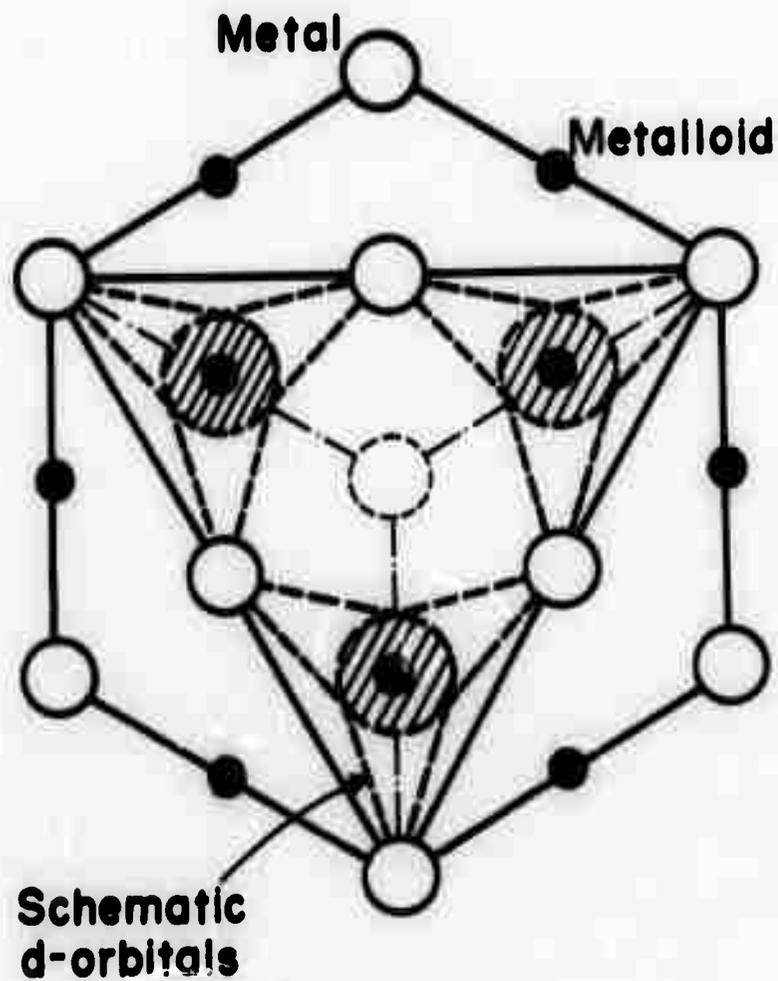


Figure 14. Schematic Bonding within Hard Metal Compounds that have the Rock-Salt Crystal Structure

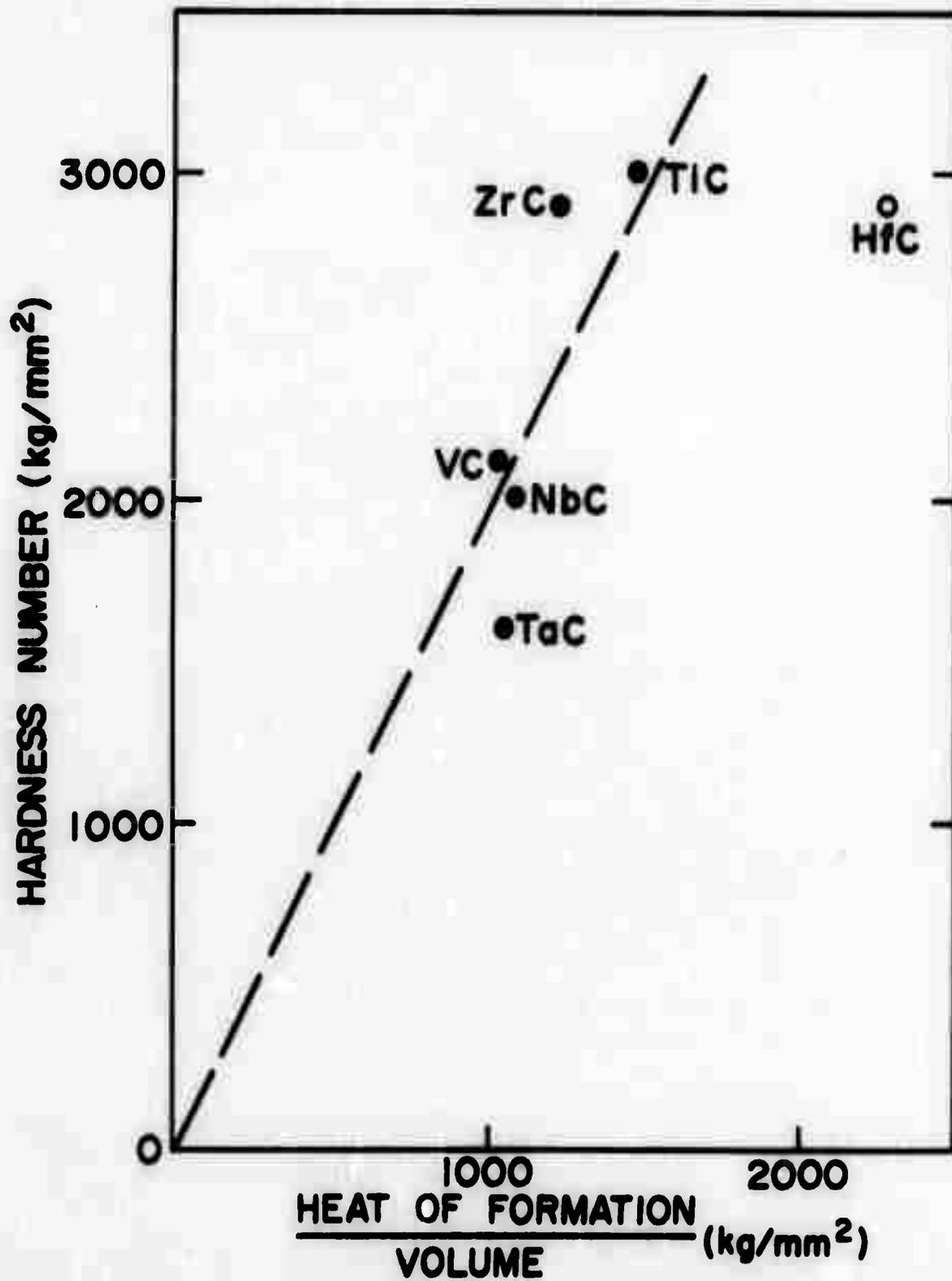
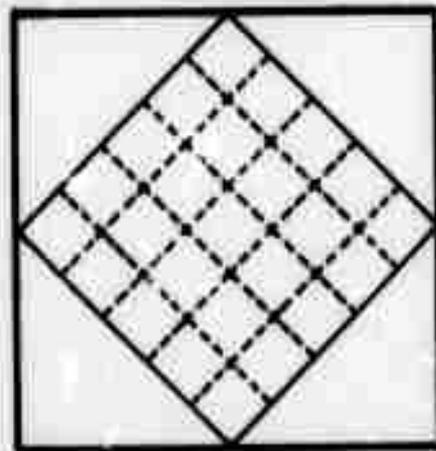
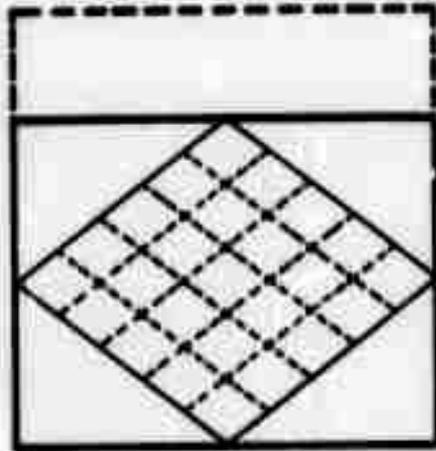
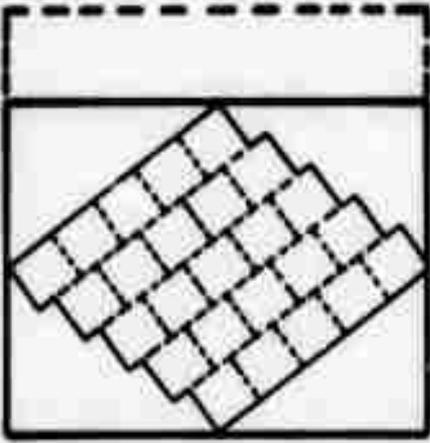


Figure 15. Correlation of the Hardnesses of Typical Carbides with their Formation Energy Densities as Predicted by Simple Theory



INITIAL STATE PLANE WAVE CAUSES SHEAR STRAINS RELAXED
COMPRESSION LEAVING ONLY VOLUME CHANGE

Figure 16. Schematic Diagram Illustrating the Plastic Relaxation of a Propagating Mechanical Shock Wave

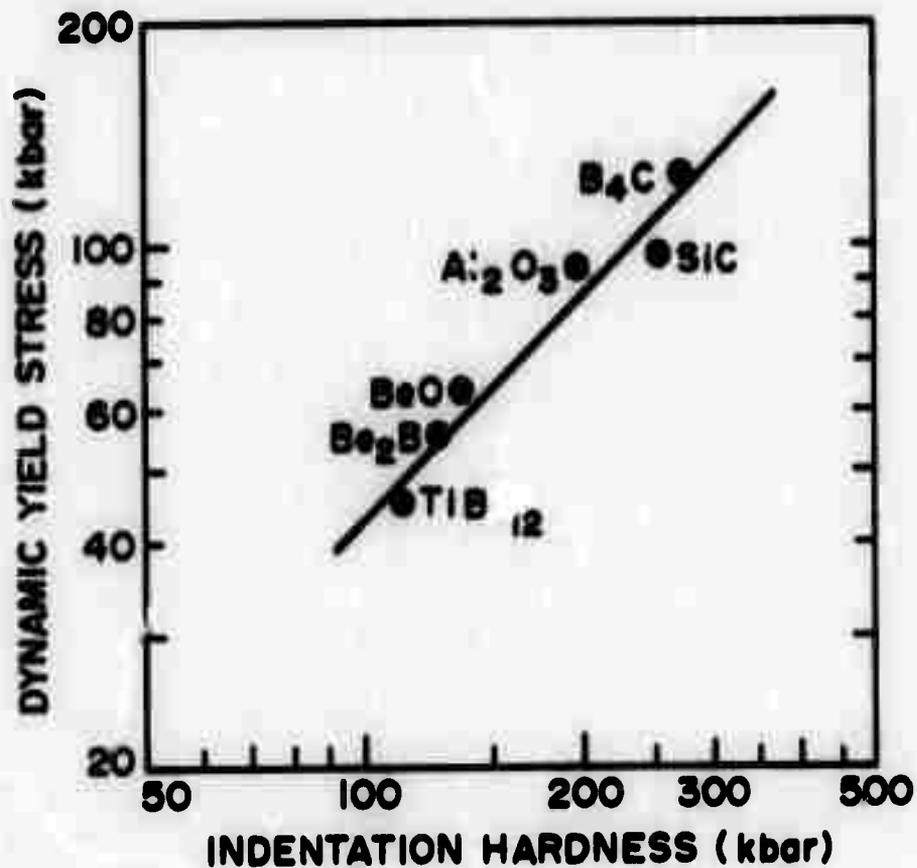


Figure 17. Correlation of Dynamic Yield Stresses Measured in Shock Experiments with the Indentation Hardnesses of Some Hard Compounds (after C. Cline)