THE ANNALS OF THE COMPUTATION LABORATORY OF HARVARD UNIVERSITY

VOLUME XXVI

THE ANNALS OF THE COMPUTATION LABORATORY OF HARVARD UNIVERSITY

I	A Manual of Operation for the Automatic Sequence Controlled Calculator	1946
II	Tables of the Modified Hankel Functions of Order One-Third and of Their Derivatives	1945
III	Tables of the Bessel Functions of the First Kind of Orders Zero and One	1947
IV	Tables of the Bessel Functions of the First Kind of Orders Two and Three	1947
V	Tables of the Bessel Functions of the First Kind of Orders Four, Five, and Six	1947
VI	Tables of the Bessel Functions of the First Kind of Orders Seven, Eight, and Nine	1947
VII	Tables of the Bessel Functions of the First Kind of Orders Ten, Eleven, and Twelve	1947
VIII	Tables of the Bessel Functions of the First Kind of Orders Thirteen, Fourteen, and Fifteen	1947
IX	Tables of the Bessel Functions of the First Kind of Orders Sixteen through Twenty-Seven	1948
X	Tables of the Bessel Functions of the First Kind of Orders Twenty-Eight through Thirty-Nine	1948
XI	Tables of the Bessel Functions of the First Kind of Orders Forty through Fifty-One	1948
XII	Tables of the Bessel Functions of the First Kind of Orders Fifty-Two through Sixty-Three	1949
XIII	Tables of the Bessel Functions of the First Kind of Orders Sixty-Four through Seventy-Eight	1949
XIV	Tables of the Bessel Functions of the First Kind of Orders Seventy-Nine through One Hundred Thirty-Five	1951
XV	(In preparation)	
XVI	Proceedings of a Symposium on Large-Scale Digital Calculating Machinery	1948
XVII	Tables for the Design of Missiles	1948
XVIII	Tables of Generalized Sine- and Cosine-Integral Functions: Part I	1949
XIX	Tables of Generalized Sine- and Cosine-Integral Functions: Part II	1949
XX	Tables of Inverse Hyperbolic Functions	1949
XXI	Tables of Generalized Exponential-Integral Functions	1949
XXII	Tables of the Function $\frac{\sin \phi}{\phi}$ and of its First Eleven Derivatives	1949
XXIII	(In preparation)	
XXIV	Description of a Relay Calculator	1949
XXV	(In preparation)	
XXVI	Proceedings of a Second Symposium on Large-Scale Digital Calculating Machinery	1951
XXVII	Synthesis of Electronic Computing and Control Circuits	1951

PROCEEDINGS OF A SECOND SYMPOSIUM ON LARGE-SCALE DIGITAL CALCULATING MACHINERY

Jointly Sponsored by The Navy Department
Bureau of Ordnance and Harvard University
at The Computation Laboratory
13–16 September 1949

CAMBRIDGE, MASSACHUSETTS
HARVARD UNIVERSITY PRESS
1951

LONDON : GEOFFREY CUMBERLEGE
OXFORD UNIVERSITY PRESS

The opinions or assertions contained herein are the private ones of the writers and are not to be construed as official or reflecting the views of the Navy Department or the naval service at large.

Composition by The Pitman Press, Bath, England

Printed by offset lithography by the Murray Printing Company, Wakefield, Massachusetts, U.S.A.

PREFACE

In January 1947 the Bureau of Ordnance of the United States Navy and Harvard University together sponsored a Symposium on Large-Scale Digital Calculating Machinery as a means of furthering interest in the design, construction, application, and operation of computing machinery. This meeting was attended by over three hundred people, nearly four times the originally expected attendance, and by popular demand the proceedings were published as Volume XVI of the Annals of the Computation Laboratory.

At the Oak Ridge meeting on computing machinery in April 1949, Mina Rees and John Mauchly, representing the Association for Computing Machinery, suggested that another symposium should be held at Harvard summarizing recent and current developments. The staff of the Computation Laboratory had already considered this possibility in connection with the announcement of the completion of Mark III Calculator, and were delighted with the suggestions of Dr. Rees and Dr. Mauchly. Accordingly, the Bureau of Ordnance was again invited to join Harvard University in sponsoring a second symposium with emphasis on the application of digital calculating machinery.

From experience with the first symposium, it was expected that perhaps three hundred people might attend. The response of more than seven hundred participants clearly indicated the rapidity with which the field of automatic computation is growing.

This volume, the twenty-sixth of the Annals of the Computation Laboratory, contains all the papers presented at the second symposium except one. Two of the speakers, Manuel S. Vallarta and Frederick V. Waugh, found at the last minute that they were unable to attend. However, their papers were received and were read by J. Curry Street and Leon Moses, respectively, both of Harvard University. Because of the tremendous editorial difficulties experienced with the proceedings of the first symposium, each speaker at the second was requested to supply his manuscript in advance, in order to avoid dependence upon transcription from sound recording. Thirty-nine papers are herein published essentially as submitted. Thus the work required to prepare this volume for publication was greatly reduced. However, it was necessary to redraw many of the illustrations for offset reproduction; this was done by Carmela M. Ciampa, assisted by Paul Donaldson, photographer of Cruft Laboratory, Harvard University.

Since the symposium was held in September, prior to the opening of the fall term, it was possible to make use of the dormitories in the Harvard Yard and the dining facilities of the Harvard Union. Arthur Trottenberg of Harvard University supervised arrangements for the use of these facilities and other accommodations. Preparation of the program and registration lists and the registration of the members of the symposium after their arrival were carried out by Betty Jennings, Jacquelin Sanborn, Jean Crawford, and Holly Wilkins. It is

a pleasure to acknowledge the coöperation of Edmund C. Berkeley, secretary of the Association for Computing Machinery, in this connection.

The staff of the Computation Laboratory wishes to express its appreciation to the members of the symposium for their attendance and for their participation in the discussions, to the chairmen of the several sessions for their assistance, and to the speakers not only for their addresses during the symposium but also for their coöperation in preparing the manuscripts of their papers.

The staff also wishes to express its gratitude to the Bureau of Ordnance and to its representatives, Captain G. T. Atkins and Mr. Albert Wertheimer, for many years of pleasant association throughout the building of Mark II and Mark III Calculators, for their continued interest and help, and for making possible both the Second Symposium on Large-Scale Digital Calculating Machinery and the publication of its proceedings.

HOWARD H. AIKEN

Cambridge, Massachusetts
May 1950

CONTENTS

PROGRAM

FIRST SESSION

Tuesday, September 13, 1949
10:30 A.M. to 12:00 P.M.

OPENING ADDRESSES

Presiding

Howard H. Aiken
Director of the Computation Laboratory

Edward Reynolds
Administrative Vice President of Harvard University

Rear Admiral F. I. Entwistle, USN
Director of Research, Bureau of Ordnance

SECOND SESSION

Tuesday, September 13, 1949
2:00 P.M. to 5:00 P.M.

RECENT DEVELOPMENTS IN COMPUTING MACHINERY

Presiding

Mina Rees, *Office of Naval Research*

1. The Mark III Calculator

   Benjamin L. Moore
   Harvard University

2. The Bell Computer, Model VI

   Ernest G. Andrews
   Bell Telephone Laboratories

3. An Electrostatic Memory System

   J. Presper Eckert, Jr.
   Eckert-Mauchly Computer Corporation

4. The Digital Computation Program at Massachusetts Institute of Technology

   Jay W. Forrester
   Massachusetts Institute of Technology

SECOND SESSION—CONTINUED

5. The Raytheon Electronic Digital Computer 50

Richard M. Bloch
Raytheon Manufacturing Company

6. A General Electric Engineering Digital Computer 65

Burton R. Lester
General Electric Company

BANQUET

Tuesday, September 13, 1949
7:00 P.M.

Toastmaster

Edward A. Weeks, Jr.
Editor of *The Atlantic Monthly*

Speaker

William S. Elliott
Research Laboratories of Elliott Brothers (London) Limited

"The Present Position of Computing-Machine Development in England" 74

THIRD SESSION

Wednesday, September 14, 1949
9:00 A.M. to 12:00 P.M.

RECENT DEVELOPMENTS IN COMPUTING MACHINERY

Presiding

E. Leon Chaffee, *Harvard University*

1. Semiautomatic Instruction on the Zephyr 43

H. D. Huskey
National Bureau of Standards, Institute for Numerical Analysis

2. Static Magnetic Delay Lines 91

Way Dong Woo
Harvard University

THIRD SESSION—CONTINUED

FOURTH SESSION

Wednesday, September 14, 1949
2:00 P.M. to 5:00 P.M.

NUMERICAL METHODS

Presiding
Raymond C. Archibald, *Brown University*

FOURTH SESSION—CONTINUED

5. An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differentia and Integral Operators 164

Cornelius Lanczos
Institute for Numerical Analysis, UCLA

6. The Monte Carlo Method 207

S. M. Ulam
Los Alamos Scientific Laboratory

FIFTH SESSION

Thursday, September 15, 1949
9:00 A.M. to 12:00 P.M.

COMPUTATIONAL PROBLEMS IN PHYSICS

Presiding

Karl K. Darrow, *Bell Telephone Laboratories*

1. The Place of Automatic Computing Machinery in Theoretical Physics 215

Wendell H. Furry
Harvard University

2. Double Refraction of Flow and the Dimensions of Large Asymmetric Molecules 219

Harold A. Scheraga, John T. Edsall, and J. Orten Gadd, Jr.
Cornell University, Harvard Medical School, and Computation Laboratory of Harvard University

3. L-Shell Internal Conversion 240

Morris E. Rose
Oak Ridge National Laboratory

4. The Use of Calculating Machines in the Theory of Primary Cosmic Radiation 244

Manuel S. Vallarta
University of Mexico
(read by J. C. Street, *Harvard University*)

5. Computational Problems in Nuclear Physics 250

Herman Feshbach
Massachusetts Institute of Technology

SIXTH SESSION

Thursday, September 15, 1949
2:00 P.M. to 5:00 P.M.

Aeronautics and Applied Mechanics

Presiding
Harald M. Westergaard, *Harvard University*

1. Computing Machines in Aeronautical Research 263

   R. D. O'Neal
   University of Michigan

2. Problem of Aircraft Dynamics 271

   Everett T. Welmers
   Bell Aircraft Corporation

3. A Statistical Method for Certain Nonlinear Dynamical Systems 281

   George R. Stibitz
   Consultant in Applied Mathematics, Burlington, Vermont

4. Combustion Aerodynamics 293

   Howard W. Emmons
   Harvard University

5. Application of Computing Machinery to Research of the Oil Industry 305

   Morris Muskat
   Gulf Research & Development Company

6. The 603-405 Computer 316

   William W. Woodbury
   Northrop Aircraft, Inc.

SEVENTH SESSION

Friday, September 16, 1949
9:00 A.M. to 12:00 P.M.

The Economic and Social Sciences

Presiding
Edwin B. Wilson, *Office of Naval Research*

1. Application of Computing Machinery to the Solution of Problems of the Social Sciences 323

   Frederick Mosteller
   Harvard University

SEVENTH SESSION—CONTINUED

EIGHTH SESSION

Friday, September 16, 1949
2:00 P.M. to 4:00 P.M.

DISCUSSION AND CONCLUSIONS

Presiding
Willard E. Bleick, *U.S. Naval Academy Post Graduate School*

* * *

OPEN DISCUSSION

MEMBERS OF THE SYMPOSIUM

MATTHEW C. ABBOTT, Engineer, W. S. MacDonald Company, Inc., Cambridge

MILTON ABRAMOWITZ, Numerical Mathematics Service, New York

CHARLES W. ADAMS, Massachusetts Institute of Technology, Cambridge

HOWARD H. AIKEN, Professor of Applied Mathematics and Director of the Computation Laboratory, Harvard University, Cambridge

MOE LAWRENCE AITEL, Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

J. CHARLES AJEMIAN, Massachusetts Institute of Technology, Cambridge

MILTON ALDEN, President, Alden Products Company, Brockton, Massachusetts

SAMUEL N. ALEXANDER, Chief, Electronic Computers Section, National Bureau of Standards, Washington, D.C.

WILLIAM R. ALLEN, Computation Laboratory, Harvard University, Cambridge

JAMES C. ALLER, LT., USN, Medford, Massachusetts

R. K. ALLERTON, JR., Public Relations, Underwood Corporation, New York

FRANZ L. ALT, National Bureau of Standards, Washington, D.C.

BIAGIO F. AMBROSIO, Engineer, National Bureau of Standards, Los Angeles, California

FREDERICK J. ANDERSON, Development Engineer, Sylvania Electric Products, Inc., Boston

LOWELL O. ANDERSON, Physicist, Naval Ordnance Test Station, China Lake, California

RUTH K. ANDERSON, Mathematician, Naval Ordnance Test Station, China Lake, California

ERNEST G. ANDREWS, Technical Staff, Bell Telephone Laboratories, New York

THOMAS B. ANDREWS, JR., Aeronautic Research Scientist, National Advisory Committee for Aeronautics, Langley Aeronautical Laboratory, Hampton, Virginia

FRANK H. ANDRIX, Engineer, Bell Aircraft Corporation, Buffalo, New York

LEONARD J. ANGUS, Consultant, Manhasset, Long Island, New York

RAYMOND C. ARCHIBALD, Professor, Brown University, Providence, Rhode Island

WALTER E. ARNOLDI, Project Engineer, Hamilton Standard Propellers, Wethersfield, Connecticut

ELEANOR ASMUTH, Badger, Wisconsin

JOHN ASMUTH, Instructor, University of Wisconsin, Madison, Wisconsin

ALBERT A. AUERBACH, Design Engineer, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

ISAAC L. AUERBACH, Senior Engineer, Burroughs Adding Machine Company, Philadelphia, Pennsylvania

DONALD E. BABCOCK, Republic Steel Corporation, Youngstown, Ohio

PAUL H. BACKUS, Lt.-Cdr., USN, Chief, Ballistics Research Section, Bureau of Ordnance, Washington, D.C.

GEORGE A. BALL, Research Assistant, Harvard University, Cambridge

STANLEY S. BALLARD, Professor of Physics, Tufts College, Medford, Massachusetts

MELVIN D. BALLER, Engineer, Air Force Cambridge Research Laboratories, Cambridge

ROBERT M. BARRETT, Electronic Engineer, Air Force Cambridge Research Laboratories, Cambridge

JEAN J. BARTIK, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

STEPHEN D. BATCHELOR, Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

DWIGHT W. BATTEAU, Research Assistant, Harvard University, Cambridge

D. T. BELL, Bell Telephone Laboratories, Inc., New York

ALBERT I. BELLIN, Assistant Professor, Harvard University, Cambridge

LAWRENCE W. BELOUNGIE, Raytheon Manufacturing Company, Waltham, Massachusetts

J. L. BELYEA, Lt., Royal Canadian Navy, Ottawa, Ontario, Canada

ALBERT A. BENNETT, Professor of Mathematics, Brown University, Providence, Rhode Island

ROBERT J. BERGEMANN, JR., Electronics Engineer, Office of Naval Research, Boston

STEFAN BERGMAN, Harvard University, Cambridge

EDMUND C. BERKELEY, President, E. C. Berkeley and Associates, New York

ERIC W. BETH, Physicist, Geophysical Research Directorate, Air Force Cambridge Research Laboratories, Cambridge

VICTOR E. BIEBER, JR., Aeronautical Engineer, Bureau of Aeronautics, Washington, D.C.

WALTER J. BINGEL, Mechanical Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

ROBERT W. BIRGE, Nuclear Laboratory, Harvard University, Cambridge

ERNEST W. BIVANS, Air Force Cambridge Research Laboratory, Cambridge

GERRIT A. BLAAUW, Computation Laboratory, Harvard University, Cambridge

PAUL B. BLACK, Head of Equipment Engineering, Sylvania Electric Products, Inc., Boston

BARTOLOME C. BLANCO, Harvard University, Cambridge

WILLARD E. BLEICK, Professor, U.S. Naval Postgraduate School, Annapolis, Maryland

ALAN BLOCH, Arma Corporation, Brooklyn, New York

RICHARD M. BLOCH, Manager, Analytical Section, Raytheon Manufacturing Company, Newton, Massachusetts

H. W. BODE, Bell Telephone Laboratories, New York

GEORGE A. W. BOEHM, Science Editor, *Newsweek*, New York

JOHN M. BOERMEESTER, John Hancock Mutual Life Insurance Company, Boston

MORTON BOISEN, Statistician, Bureau of the Census, Washington, D.C.

ROBERT N. BONNER, Junior Research Chemist, Carter Oil Company, Tulsa, Oklahoma

RICHARD C. BOOTON, JR., Research Assistant, Differential Analyzer Computer Laboratory, Massachusetts Institute of Technology, Cambridge

GUY F. BOUCHER, Computation Laboratory, Harvard University, Cambridge

JOHN R. BOWMAN, Head, Department of Research in Physical Chemistry, Mellon Institute, Pittsburgh, Pennsylvania

HUGH R. BOYD, Research Engineer, Massachusetts Institute of Technology, Cambridge

HENRY B. BRAINERD, Massachusetts Institute of Technology, Cambridge

C. CLINTON BRAMBLE, Director of Computation Ballistics, U.S. Naval Proving Ground, Dahlgren, Virginia

C. S. BRAND, Arnold Engineering Company, Chicago, Illinois

LEON BRILLOUIN, International Business Machines Corporation, New York

PAUL BROCK, Reeves Instrument Corporation, New York

DOUGLAS A. BROWN, Assistant Director, Harvard University News Office, Harvard University, Cambridge

GEORGE W. BROWN, Rand Corporation, Santa Monica, California

J. H. BROWN, Research Engineer, Massachusetts Institute of Technology, Cambridge

ROBERT G. BROWN, Operations Evaluation Group, CNO, Navy Department, Washington, D.C.

T. H. BROWN, Professor, Harvard Business School, Harvard University, Boston

T. WISTAR BROWN, Sales Manager, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

WILLIAM FULLER BROWN, JR., Research Physicist, Sun Oil Company, Newtown Square, Pennsylvania

JOSEPH A. BRUSTMAN, Chief Engineer, Remington Rand, Inc., South Norwalk, Connecticut

EDWARD C. BRYANT, Student, Boston University, Boston

RALPH W. BUMSTEAD, Patent Attorney, Westfield, New Jersey

RICHARD S. BURINGTON, Director, Evaluation and Analysis Group, Bureau of Ordnance, Washington, D.C.

WILLIAM BURKHART, Monroe Calculating Machine Company, Orange, New Jersey

ROBERT J. BURNS, Computation Laboratory, Harvard University, Cambridge

ROBERT R. BUSH, Research Fellow, Harvard University, Cambridge

SAMUEL H. CALDWELL, Professor of Electrical Engineering, Massachusetts Institute of Technology, Cambridge

FRANK J. CAMPAGNA, Computation Laboratory, Harvard University, Cambridge

ELIZABETH JEAN CAMPBELL, Computer, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge

R. E. CAMPBELL, Liaison Engineer, Signal Corps Engineering Laboratories, Massachusetts Institute of Technology, Cambridge

ROBERT V. D. CAMPBELL, Department Supervisor, Burroughs Adding Machine Company, Philadelphia, Pennsylvania.

E. W. CANNON, Mathematician, National Bureau of Standards, Washington, D.C.

JOEL CARROLL, Geodetic Engineer, U.S. Geological Survey, Washington, D.C.

JOHN B. CARROLL, Assistant Professor of Education, Harvard University, Cambridge

ELBERT P. CARTER, Senior Engineer, Transducer Corporation, Boston

CARL C. CHAMBERS, Acting Dean, University of Pennsylvania, Philadelphia, Pennsylvania

GEORGE C. CHASE, Research Engineer, Monroe Calculating Machine Company, Orange, New Jersey

JOSEPH CHEDAKER, Senior Engineer, Burroughs Adding Machine Company, Philadelphia, Pennsylvania

T. C. CHEN, Senior Engineer, Burroughs Adding Machine Company, Philadelphia, Pennsylvania

HERMAN CHERNOFF, Research Associate, Cowles Commission, University of Chicago, Chicago, Illinois

ALFRED C. CHEVERIE, Computation Laboratory, Harvard University, Cambridge

BENJAMIN F. CHEYDLEUR, Mathematician, U.S. Bureau of Ordnance, Washington, D.C.

HENRY W. F. CHIN, Assistant Mechanical Engineer, Raytheon Manufacturing Company, Newton, Massachusetts

MANUEL P. CHINITZ, Mathematician, U.S. Naval Proving Ground, Dahlgren, Virginia

ALLEN G. CHRISTENSEN, Computation Laboratory, Harvard University, Cambridge

THOMAS J. CHRISTMAN, Lt., USN, Post-graduate Student, Massachusetts Institute of Technology, Cambridge

J. C. CHU, Senior Scientist, Argonne National Laboratory, Chicago, Illinois

CARMELA M. CIAMPA, Computation Laboratory, Harvard University, Cambridge

A. G. CLAVIER, Assistant Technical Director, Federal Telecommunication Laboratories, Nutley, New Jersey

GERALD M. CLEMENCE, Director, Nautical Almanac, U.S. Naval Observatory, Washington, D.C.

R. F. CLIPPINGER, Computing Laboratory, Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland

A. M. CLOGSTON, Bell Telephone Laboratories, Murray Hill, New Jersey

RICHARD P. COATES, Computation Laboratory, Harvard University, Cambridge

TODD D. COCHRAN, JR., Development Engineer, Eastman Kodak Company, Rochester, New York

R. C. COILE, Operations Evaluation Group, Office of the Chief of Naval Operations, Navy Department, Washington, D.C.

CHARLES F. COIT, Senior Engineer, Raytheon Manufacturing Company, Newton, Massachusetts

ARNOLD A. COHEN, Senior Engineer, Engineering Research Associates, Inc., St. Paul, Minnesota

CHARLES J. COHEN, Mathematician, Naval Proving Ground, Dahlgren, Virginia

JOHN J. CONNOLLY, Engineering Supervisor, Teleregister Corporation, New York

CHARLES A. COOLIDGE, JR., Computation Laboratory, Harvard University, Cambridge

JOHN M. COOMBS, Director of Development, Engineering Research Associates, St. Paul, Minnesota

GERALD COOPER, Research Assistant, Massachusetts Institute of Technology, Cambridge

CHARLES L. CORDERMAN, Research Assistant, Massachusetts Institute of Technology, Cambridge

DOUGLAS S. CRAIG, Second Vice-President, Metropolitan Life Insurance Company, New York

JEAN CRAWFORD, Computation Laboratory, Harvard University, Cambridge

JOHN H. CREDE, Associate Director of Research, Allegheny Ludlum Steel Corporation, Brackenridge, Pennsylvania

L. P. CROSMAN, Director, Electronic Calculator Research, Remington Rand, Inc., South Norwalk, Connecticut

EDWARD D. CROSS, Director of Engineering, Alden Products Company, Brockton, Massachusetts

W. J. CROZIER, Professor, Harvard University, Cambridge

WILLIAM H. CUMMINS, Chief, Classification and Coding Branch, Federal Security Administration, Washington, D.C.

HASKELL B. CURRY, Professor of Mathematics, The Pennsylvania State College, State College, Pennsylvania

W. W. CURTIS, Office Supervisor, Aluminium Company of America, Boston

R. J. CYPSER, Instructor, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge

JOHN A. DAELHOUSEN, Staff Member, Massachusetts Institute of Technology, Cambridge

EVERETT J. DANIELS, Staff Member, Massachusetts Institute of Technology, Cambridge

M. DANILOFF, Senior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

GEORGE B. DANTZIG, Mathematician, US. Air Force Comptroller, Washington, D.C.

KARL K. DARROW, Physicist, Bell Telephone Laboratories, New York

GERALD W. DAVIS, Electronic Scientist, National Bureau of Standards, Washington, D.C.

MALVIN E. DAVIS, Actuary, Metropolitan Life Insurance Company, New York

PHILLIP DAVIS, Mathematician, Harvard University, Cambridge

CHRISTOPHER DEAN, Teaching Fellow, Harvard University, Cambridge

FRANKLIN R. DEAN, Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

J. T. DEBETTENCOURT, Section Manager, Raytheon Manufacturing Company, Waltham, Massachusetts

L. S. DEDERICK, Associate Director, Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland

GEORGE H. DEPINTO, Harvard University, Cambridge

HENRY DESTEFANO, Harvard University, Cambridge

JOHN E. DETURK, Senior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

M. L. DEUTSCH, Physicist, Socony-Vacuum Research Laboratory, Paulsboro, New Jersey

A. J. DEVAUD, Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

R. L. DEVEER, C. P. Clare and Company, Boston

GEORGE C. DEVOL, Manager, Magnetic Devices Department, Research Laboratory, Remington Rand, Inc., South Norwalk, Connecticut

WALTER L. DEVRIES, Actuarial Supervisor, Equitable Life Assurance Society, New York

C. B. DEWEY, Vice-President, Reeves Instrument Corporation, New York

ERNEST J. DIETERICH, Junior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

JOHN D. DILLON, Assistant Chief, Research Services, Air Force, Cambridge Research Laboratories, Cambridge

BERNARD DIMSDALE, Mathematician, Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland

L. P. DISNEY, Chief, Section of Predictions, U.S. Coast and Geodetic Survey, Washington, D.C.

NATHAN DIVINSKY, Research Associate, Cowles Commission, Chicago, Illinois

STEPHEN H. DODD, Research Engineer, Massachusetts Institute of Technology, Cambridge

CHARLES H. DOERSAM, JR., Mechanical Engineer, Office of Naval Research, Section of Development Contract, Port Washington, Long Island, New York

FRANCIS W. DRESCH, Assistant Director, Computation and Ballistics, U.S. Naval Proving Ground, Dahlgren, Virginia

L. B. DUMONT, Engineer, General Electric Company, Lynn, Massachusetts

ROBERT F. DUNCAN, Vice-President, John Price Jones Company, New York

S. W. DUNWELL, Future Demands Department, International Business Machines Corporation, New York

J. R. DYER, Colonel, USAF, Staff Officer, Munitions Board, Department of Defense, Washington, D.C.

E. C. EASTON, Dean of Engineering, Rutgers University, New Brunswick, New Jersey

J. PRESPER ECKERT, JR., Vice-President, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

W. J. ECKERT, Director, Pure Science, International Business Machines Corporation, New York

ROBERT P. EDDY, Mathematician, Naval Ordnance Laboratory, Silver Spring, Maryland

NIELS E. EDLEFSEN, Associate Technical Director, North American Aviation, Santa Monica, California

JOHN T. EDSALL, Associate Professor of Biological Chemistry, Harvard Medical School, Boston

MILTON EFFROS, Management Research, Metropolitan Life Insurance Company, New York

ROBERT D. ELBOURN, National Bureau of Standards, Washington, D.C.

PETER ELIAS, Teaching Fellow, Harvard University, Cambridge

WILLIAM S. ELLIOTT, Research Laboratories of Elliott Brothers (London), Ltd., Borehamwood, Hertfordshire, England

MURRAY ELLIS, Junior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

GEORGE V. ELTGROTH, Vice President, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

JOHN O. ELY, Research Associate, Massachusetts Institute of Technology, Cambridge

CLAUDE L. EMMERICH, Senior Physicist, Martin-Hubbard Corporation, Cambridge

HOWARD EMMONS, Associate Professor, Harvard University, Cambridge

HOWARD T. ENGSTROM, Vice President, Engineering Research Associates, Inc., St. Paul, Minnesota

F. I. ENTWISTLE, Rear Admiral, USN, Director of Research, Bureau of Ordnance, Navy Department, Washington, D.C.

HANS H. ESTIN, Brookline, Massachusetts

HELENTY ESTIN, Brookline, Massachusetts

ROBERT R. EVERETT, Research Engineer, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge

ROBERT G. EVERSEN, Computation Laboratory, Harvard University, Cambridge

HARRIS FAHNESTOCK, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge

R. M. FAIRBROTHER, Massachusetts Institute of Technology, Cambridge

R. S. FALLOWS, Engineer, Sylvania Electric Products, Inc., Boston

JOHN T. FARREN, Computation Laboratory, Harvard University, Cambridge

LOUIS FEIN, Senior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

J. H. FELKER, Member of Technical Staff, Bell Telephone Laboratories, Whippany, New Jersey

SAMUEL FELTMAN, Ordnance Engineer, Office of Chief of Ordnance, Washington, D.C.

FRED G. FENDER, Professor of Mathematics, Rutgers University, New Brunswick, New Jersey

DAVID T. FERRIER, Lawrence, Long Island, New York

HERMAN FESHBACH, Associate Professor, Massachusetts Institute of Technology, Cambridge

F. A. FICKEN, Research Associate, New York University, New York

L. R. FINK, Manager, Electronics Laboratory, General Electric Company, Syracuse, New York

HAROLD A. FINLEY, Manager, Management Research, Metropolitan Life Insurance Company, New York

LYMAN C. FISHER, Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

HARLAND W. FLAGG, Marchant Calculating Machine Company, Boston

DONALD A. FLANDERS, Senior Mathematician, Argonne National Laboratory, Chicago, Illinois

MARGARET I. FLORENCOURT, Research Engineer, Massachusetts Institute of Technology, Cambridge, Massachusetts

WILLIAM B. FLOYD, Sears Roebuck and Company, Chicago, Illinois

J. A. FLYNN, Mechanical Engineer, Cambridge Research Laboratory, Cambridge

JAMES W. FOLLIN, JR., Physicist, Applied Physics Laboratory, Johns Hopkins University, Silver Spring, Maryland

G. DONALD FORBES, Electronics Consultant, Sudbury, Massachusetts

RICHARD E. FORBES, Computation Laboratory, Harvard University, Cambridge

JAY W. FORRESTER, Associate Director, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge

GEORGE E. FORSYTHE, National Bureau of Standards, University of California at Los Angeles, Los Angeles, California

FRANKLIN H. FOWLER, JR., Associate Editor, *Product Engineering*, New York

PHILIP FRANKLIN, Professor, Massachusetts Institute of Technology, Cambridge

WALTER J. FRANTZ, Research Engineer, Boeing Aircraft, Seattle, Washington

WILLIAM H. FRATER, Director, Wage Analysis, General Motors Corporation, Detroit, Michigan

DAVID FRAZIER, Research Chemist, Standard Oil Company, Cleveland, Ohio

R. O. FREDETTE, Mathematician, Bureau of Ordnance, Navy Department, Washington, D.C.

F. N. FRENKIEL, Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

A. W. FRICK, Radio Corporation of America, Camden, New Jersey

W. BARKLEY FRITZ, Mathematician, Ballistics Research Laboratory, Aberdeen, Maryland

A. E. FROST, Assistant Equipment Research Engineer, Western Union Telegraph Company, New York

JOSEPH FUCARILE, Computation Laboratory, Harvard University, Cambridge

W. H. FURRY, Associate Professor of Physics, Harvard University, Cambridge

IRVING J. GABELMAN, Engineer, Watson Laboratories, USAF, Long Branch, New Jersey

EDWIN GABRIEL, Associate Engineer, Air Force Cambridge Research Laboratories, Cambridge

J. ORTEN GADD, JR., Computation Laboratory, Harvard University, Cambridge

JAMES A. GEAN, Chief Analytical Engineer, Parsons Corporation, Traverse City, Michigan

ARTHUR A. GENTILE, Computation Laboratory, Harvard University, Cambridge

EUGENE M. GETTEL, Research Engineer, United Aircraft Corporation, East Hartford, Connecticut

N. ELIOT GIBBS, Senior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

RICHARD C. GIBSON, Assistant Professor of Electrical Engineering, USAF Institute of Technology, Wright-Patterson Air Force Base, Dayton, Ohio

DONALD B. GILLIES, Computation Centre, Toronto, Canada

DOROTHY F. GILLETTE, Physicist, Air Force Cambridge Research Laboratories, Cambridge

H. F. GINGERICH, U.S. Navy Department, Office of Naval Research, Washington, D.C.

G. GLINSKI, Vice-President, Computing Devices of Canada, Ltd., Ottawa, Canada

SIMON E. Gluck, Research Associate, Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pennsylvania

THOMAS N. K. GODFREY, National Bureau of Standards, Massachusetts Institute of Technology, Cambridge

STANFORD GOLDMAN, Professor of Electrical Engineering, Syracuse University, Syracuse, New York

HARRY H. GOODE, Special Devices Center, Office of Naval Research, Sands Point, Long Island, New York

C. C. GOTLIEB, Assistant Professor, University of Toronto, Toronto, Ontario, Canada

R. S. GRAHAM, Bell Telephone Laboratories, Murray Hill, New Jersey

E. F. GRANT, Chief, Applied Mathematics Branch, Air Force Cambridge Research Laboratories, Cambridge

RANULF W. GRAS, Massachusetts Institute of Technology, Cambridge

HARRY J. GRAY, JR., Instructor of Electrical Engineering, University of Pennsylvania, Moore School of Electrical Engineering, Philadelphia, Pennsylvania

WALTER H. GRAY, JR., Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

BEN F. GREENE, Radio Engineer, Air Force Cambridge Research Laboratories, Cambridge

H. VOSE GREENOUGH, JR., Director, Technichord Records, Brookline, Massachusetts

DARRIN H. GRIDLEY, Naval Research Laboratories, Washington, D.C.

D. D. GRIEG, Division Head, Federal Telecommunication Laboratories, Nutley, New Jersey

B. A. GRIFFITH, Assistant Professor, University of Toronto, Toronto, Canada

IRVING I. GRINGORTEN, Meteorologist, Air Force Cambridge Research Laboratories, Cambridge

H. R. J. GROSCH, Watson Scientific Computing Laboratory, New York

E. V. GULDEN, Research Engineer, National Cash Register Company, Dayton, Ohio

WILLIAM F. GUNNING, Engineer, Rand Corporation, Santa Monica, California

DANIEL HAAGENS, Electronics Engineer, Underwood Corporation, Hartford, Connecticut

JAMES V. HAGGERTY, Procedural Consultant, Social Security Administration, Baltimore, Maryland

GILBERT O. HALL, Electronic Scientist, Air Force Cambridge Research Laboratories, Cambridge

W. K. HALSTEAD, Chief Engineer, W. S. MacDonald Company, Cambridge

F. E. HAMILTON, Engineer, International Business Machines Corporation, Endicott, New York

PRESTON C. HAMMER, Los Alamos Scientific Laboratory, Los Alamos, New Mexico

R. W. HAMMING, Bell Telephone Laboratory, Murray Hill, New Jersey

MRS. R. W. HAMMING, Morristown, New Jersey

GARNET HANES, Computation Centre, University of Toronto, Toronto, Canada

KENNETH C. HANNA, Computation Laboratory, Harvard University, Cambridge

GEORGE A. HARDENBERGH, Assistant Research Engineer, Engineering Research Associates, St. Paul, Minnesota

E. L. HARDER, Consulting Transmission Engineer, Westinghouse Electric Corporation, East Pittsburgh, Pennsylvania

JOHN A. HARR, Computation Laboratory, Harvard University, Cambridge

ROBERT W. HART, Electronics Engineer, Office of Naval Research, Boston
M. L. HASELTON, Vice President, Teleregister Corporation, New York
SETH HASTINGS, Mutual Life Insurance Company, New York
WILLARD D. HATCH, Ohio State University, Cleveland, Ohio
ROBERT L. HAWKINS, Computation Laboratory, Harvard University, Cambridge
MILES V. HAYES, Computation Laboratory, Harvard University, Cambridge
INEZ HAZEL, Assistant Engineer, Raytheon Manufacturing Company, Newton, Massachusetts
VERNON H. HEAD, National Advisory Committee for Aeronautics Project, Gordon McKay Laboratory, Harvard University, Cambridge
SAUL D. HEARN, Chief, Employee Statistics Section, Social Security Administration, Baltimore, Maryland
J. C. HEBARD, JR., Mechanical Engineer, Raytheon Manufacturing Company, Newton, Massachusetts
MAURICE H. HELLMAN, Engineer, Air Force Cambridge Research Laboratories, Cambridge, Massachusetts
L. L. HENKEL, Industrial College of Armed Forces, Fort Lesley McNair, Washington, D.C.
H. A. HENNING, Bell Telephone Laboratories, Inc., New York
PAUL HERGET, Director, University of Cincinnati Observatory, Cincinnati, Ohio
FRANK C. HERNE, Computation Laboratory, Harvard University, Cambridge
ROGER W. HICKMAN, Lecturer, Harvard University, Cambridge
JOHN L. HILL, Engineer, Engineering Research Associates, St. Paul, Minnesota
GEORGE W. HOBBS, Electronics Engineer, General Electric Company, Schenectady, New York
GEORGE G. HOBERG, Research Associate, Burroughs Adding Machine Company, Philadelphia, Pennsylvania
RICHARD HOFHEIMER, Computation Laboratory, Harvard University, Cambridge
MURRAY HOFFMAN, Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts
DOUGLAS L. HOGAN, Electronics Engineer, Navy Department, Washington, D.C.
JOHN V. HOLBERTON, Mathematician, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland
A. B. HOLLISTER, Engineer, Underwood Corporation, Hartford, Connecticut
MARVIN R. HOLTER, Research Assistant, University of Michigan, Ann Arbor, Michigan
LAWRENCE F. HOPE, Assistant Head, Department ME–6, Research Division, General Motors, Detroit, Michigan
RALPH HOPKINS, Special Representative, International Business Machines Corporation, Washington, D.C.
GRACE HOPPER, Chief of Application Department, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania
VIRGIL M. HORN, Assistant Section Head, Actuary Planning, Metropolitan Life Insurance Company, New York

JACOB HOROWITZ, Engineer, Raytheon Manufacturing Company, Newton, Massachusetts
JOHN H. HOWARD, Senior Project Engineer, Sperry Gyroscope Company, Great Neck, Long Island, New York
BRADFORD HOWLAND, Graduate Student, Department of Physics, Harvard University, Cambridge
JOHN F. HUBBARD, Treasurer, Martin-Hubbard Corporation, Boston
WILLIAM HULME, Computation Laboratory, Harvard University, Cambridge
J. STUART HUNTER, Statistician, Institute of Statistics, University of North Carolina, Raleigh, North Carolina
ALLEN HUNTINGTON, Electronics Engineer, U.S. Navy Electronics Laboratory, San Diego, California
C. C. HURD, International Business Machines Corporation, New York
HARRY D. HUSKEY, Mathematician, National Bureau of Standards, Los Angeles, California
W. R. HYDEMAN, Mathematician, U.S. Navy, Washington, D.C.
FRANK T. INNES, Research Engineer, The Franklin Institute, Philadelphia, Pennsylvania
EUGENE ISAACSON, Assistant Professor, Institute for Mathematics and Mechanics, New York University, New York
DAVID R. ISRAEL, Research Assistant, Massachusetts Institute of Technology, Cambridge
ARVID W. JACOBSON, Assistant Professor of Mathematics, Wayne University, Detroit, Michigan
BETTY JENNINGS, Computation Laboratory, Harvard University, Cambridge
PAUL V. JESTER, Vice President, A. C. Nielsen Company, Chicago, Illinois
G. D. JOHNSON, Bell Telephone Laboratories, Murray Hill, New Jersey
PAUL A. JOHNSON, Engineer, Boeing Airplane Company, Seattle, Washington
STANLEY A. JOHNSON, Computation Laboratory, Harvard University, Cambridge
R. F. JOHNSTON, Computation Centre, University of Toronto, Toronto, Canada
R. CLARK JONES, Senior Physicist, Polaroid Corporation, Cambridge
ROBERT HUDSON JONES, Head Engineer, Radar Design Branch, Navy Department, Bureau of Ships, Washington, D.C.
THOMAS F. JONES, JR., Assistant Professor, Differential Analyzer Computer Laboratory, Massachusetts Institute of Technology, Cambridge
WILLARD C. JONES, Assistant Chief Engineer, Underwood Corporation, Hartford, Connecticut
WILLIAM B. JORDAN, Engineer, General Electric Company, Schenectady, New York
THEODORE A. KALIN, Computation Laboratory, Harvard University, Cambridge
SIDNEY KAPLAN, Mathematician, Naval Ordnance Laboratory, Washington
MIDA KARAKASHIAN, Member, Joint Computing Group, Massachusetts Institute of Technology, Cambridge
ARTHUR A. KATZ, Mathematician, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania
J. KATZ, Engineer, University of Toronto, Canada
MARTIN KATZIN, Consultant, Naval Research Laboratory, Washington, D.C.

ALLEN KELLER, Assistant Division Engineer, Turbine Engineering Division, General Electric Company, Lynn, Massachusetts

E. G. KELLER, General Electric Company, Schenectady, New York

DANIEL W. KELLIHER, Computation Laboratory, Harvard University, Cambridge

JACQUELIN KELLY, Assistant Engineer, Raytheon Manufacturing Company, Newton, Massachusetts

JOHN P. KELLY, Head, Central Statistics Laboratory, Atomic Energy Commission, Oak Ridge, Tennessee

WENTWORTH KENNARD, Technical Editor, Raytheon Manufacturing Company, Waltham, Massachusetts

HARRY KENOSIAN, Research Engineer, Massachusetts Institute of Technology, Cambridge

CHESTER H. J. KEPPLER, Captain, USN, Retired, Counsellor for Foreign Students, Harvard University, Cambridge

J. H. KERFOOT, Civilian Technical Officer, Royal Canadian Navy, Ottawa, Ontario, Canada

MARSHALL KINCAID, Computation Laboratory, Harvard University, Cambridge

GILBERT W. KING, Arthur D. Little, Inc., Cambridge

R. W. P. KING, Professor, Harvard University, Cambridge

HENRY KINZLER, Supervisor, Procedure Planning, Metropolitan Life Insurance Company, New York

HANS KLEMPERER, Research Engineer, Massachusetts Institute of Technology, Cambridge

HERBERT M. KNIGHT, Electronic Engineer, Air Force Cambridge Research Laboratories, Cambridge

RALPH J. KOCHENBURGER, Instructor, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge

FLORENCE K. KOONS, Mathematician, National Bureau of Standards, Washington, D.C.

ZDENEK KOPAL, Massachusetts Institute of Technology, Cambridge

JOHN J. KORZDORFER, Project Engineer, Red Bank Division, Bendix Aviation Corporation, Red Bank, New Jersey

HANS KRAFT, Aerodynamicist, General Electric Company, Schenectady, New York

H. P. KUEHNI, Division Engineer, General Electric Company, Schenectady, New York

JOSEPH H. KUSNER, Munitions Board, Department of Defense, Washington, D.C.

N. L. KUSTERS, Research Engineer, National Research Council, Ottawa, Ontario, Canada

EDWARD LACEY, Electrical Engineer, National Bureau of Standards, Los Angeles, California

H. N. LADEN, LT., USN, Research and Development Engineer, Bureau of Ships, Washington, D.C.

LEON J. LADER, Radio Engineer, Watson Laboratories, Red Bank, New Jersey

CORNELIUS LANCZOS, Mathematician, Institute for Numerical Analysis, University of California at Los Angeles, California

CARNEY LANDIS, Professor of Psychology, Columbia University, New York

GERALD W. LAVIGNE, Computation Laboratory, Harvard University, Cambridge

TIMOTHY LEARY, Division of Industrial Coöperation, Massachusetts Institute of Technology, Cambridge

PHILIPPE E. LECORBEILLER, Professor of Applied Physics, Harvard University, Cambridge

HARRY S. LEE, Engineer, Massachusetts Institute of Technology, Cambridge

DERRICK H. LEHMER, Professor of Mathematics, University of California, Berkeley, California

R. A. LEIBLER, USN, Washington, D.C.

POLLY LEIGHTON, Supervisor, Computing Group, Massachusetts Institute of Technology, Cambridge

ALAN L. LEINER, Physicist, National Bureau of Standards, Electronic Computers Section, Washington, D.C.

MRS. HENRIETTA C. LEINER, Physicist, National Bureau of Standards, Electron Tube Laboratory, Washington, D.C.

J. PLUMER LEIPHART, Electronic Engineer, Naval Research Laboratory, Washington, D.C.

WASSILY W. LEONTIEF, Professor of Economics, Harvard University, Cambridge

BURTON R. LESTER, Section Engineer, Electronics Laboratory, General Electric Company, Syracuse, New York

JOSEPH H. LEVIN, Mathematician, Computation Laboratory, National Bureau of Standards, Washington, D.C.

ARNOLD M. LEVINE, Department Head, Federal Telecommunications Laboratories, Nutley, New Jersey

FRED C. LEWIS, Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

C. C. LIN, Massachusetts Institute of Technology, Cambridge

PETER L. LINDLEY, Computation Laboratory, Harvard University, Cambridge

HERBERT G. LINDNER, Radio Engineer, Coles Signal Corps Laboratory, Red Bank, New Jersey

S. B. LITTAUER, Associate Professor, Columbia University, New York

ELBERT P. LITTLE, Chief, Computation Section, Office of Air Research, USAF; Computation Laboratory, Harvard University, Cambridge

HUBERT M. LIVINGSTON, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

MRS. HUBERT M. LIVINGSTON, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

CHARLES J. LODA, Physicist, U.S. Navy Underwater Sound Laboratory, New London, Connecticut

SAMUEL LUBKIN, Electronic Scientist, National Bureau of Standards, Washington, D.C.

E. E. LUCCHINI, Computation Laboratory, Harvard University, Cambridge

DUNCAN LUCE, Research Center for Group Dynamics, University of Michigan, Ann Arbor, Michigan

W. F. MACDONALD, S. D. Leidesdorf and Company, New York

JOHN H. MACNEILL, Senior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

H. M. MacNeille, Atomic Energy Commission, Washington, D.C.
W. H. MacWilliams, Jr., Bell Telephone Laboratories, Murray Hill, New Jersey
P. J. Maginniss, Engineer, General Electric Company, Schenectady, New York
Alfred T. Magnell, Cdr., USN, Office of Naval Research, Washington, D.C.
Henry M. Mallon, Assistant Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts
Harland Manchester, Roving Editor, *Reader's Digest*, New York
Roland A. Mangini, John Hancock Life Insurance Company, Boston
Charles S. Manning, Electrical Engineer, Navy Electronics Laboratory, San Diego, California
Ethel Cox Marden, Mathematician, National Bureau of Standards, Washington, D.C.
Emil Marecki, Assistant Chief, Machine Tabulation Division, Bureau of the Census, Washington, D.C.
Arthur E. Mari, Mari & Rentel Company, Inc., Boston
Joseph Marksteiner, Computation Laboratory, Harvard University, Cambridge
Lawrence Markus, Harvard University, Cambridge
Elbert W. Marlowe, Engineer, Union Switch & Signal Company, Pittsburgh, Pennsylvania
H. W. Marsh, Consultant, USN Underwater Sound Laboratory, New London, Connecticut
Robert H. Marsh, Engineer, Massachusetts Institute of Technology, Cambridge
Byron O. Marshall, Physicist, Air Force Cambridge Research Laboratories, Cambridge
David W. Marshall, Procedures Analyst, Metropolitan Life Insurance Company, New York
W. T. Martin, Professor, Massachusetts Institute of Technology, Cambridge
Edward Massell, Engineer, Electronic Associates, Inc., Long Branch, New Jersey
Morton P. Matthew, Research Engineer, Friden Calculating Machine Company, San Leandro, California
J. W. Mauchly, President, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania
Harold F. May, Assistant Director, Teleregister Laboratories, New York
William E. May, Department of the Army, Washington, D.C.
J. P. Mayberry, Computation Centre, University of Toronto, Toronto, Ontario, Canada
H. C. Mayer, Research and Development, Radio Corporation of America, Camden, New Jersey
Rollin Powell Mayer, Research Engineer, Massachusetts Institute of Technology, Cambridge
Everett E. McCown, Electronics Engineer, U.S. Navy Electronics Laboratory, San Diego, California
James O. McDonough, Research Engineer, Massachusetts Institute of Technology, Cambridge
William D. McGuigan, Engineering Coördinator, Raytheon Manufacturing Company, Waltham, Massachusetts
E. C. McKay, Chief, Section of Tides, U.S. Coast and Geodetic Survey, Department of Commerce, Washington, D.C.

R. O. McMANUS, Mechanical Engineer, Air Force Cambridge Research Laboratory, Cambridge

JAMES L. McPHERSON, Machine Development Officer, Bureau of the Census, Washington, D.C.

JOHN C. McPHERSON, Vice President, International Business Machines, New York

KENNETH E. McVICAR, Research Assistant, Electrical Engineering Department, Massachusetts Institute of Technology, Cambridge

LEONARD C. MEAD, Research Coördinator, Tufts College, Medford, Massachusetts

RALPH I. MEADER, Vice President, Engineering Research Associates, St. Paul, Minnesota

R. E. MEAGHER, Assistant Professor, University of Illinois, Urbana, Illinois

EUGENE A. MECHLER, Research Engineer, Franklin Institute, Philadelphia, Pennsylvania

C. S. MERCER, Sales Engineer, Aluminium Company of America, Boston

DAVID MIDDLETON, Associate Professor of Applied Physics, Harvard University, Cambridge

JAMES G. MILES, Engineer, Engineering Research Associates, Inc., St. Paul, Minnesota

E. J. MILLER, Technical Officer, National Defence, Ottawa, Canada

FREDERICK G. MILLER, Electrical Engineer, Naval Proving Ground, Dahlgren, Virginia

HAROLD C. MILLER, Assistant Chairman, Physics Research, Armour Research Foundation, Chicago, Illinois

BURTON E. MILLS, Air Force Cambridge Research Laboratories, Cambridge

W. E. MILNE, Professor, Oregon State College, Corvallis, Oregon

HARRY R. MIMNO, Professor of Applied Physics, Harvard University, Cambridge

MILTON J. MINNEMAN, Electronic Group Engineer, Glenn L. Martin Company, Baltimore, Maryland

WILLIAM G. MINTY, Computation Laboratory, Harvard University, Cambridge

HERBERT F. MITCHELL, JR., Webb Institute of Naval Architecture, Glenn Cove, Long Island, New York

SAMARENDRA K. MITRA, UNESCO, Calcutta, India

WILLIAM MITTELMAN, Griffiss Air Force Base, Rome, New York

CARLTON A. MIZEN, Project Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

ELMER B. MODE, Professor of Mathematics, Boston University, Boston

BRUCE MONCREIFF, Systems Reviewer, Prudential Insurance Company of America, Newark, New Jersey

ROBERT J. MONROE, Institute of Statistics, North Carolina State College, Raleigh, North Carolina

N. F. MOODY, National Research Council, Ontario, Canada

CALVIN N. MOOERS, President, Zator Company, Boston

C. M. MOONEY, International Business Machines Corporation, New York

BENJAMIN L. MOORE, Assistant Director, Computation Laboratory, Harvard University, Cambridge

EDWARD F. MOORE, Graduate Student, Brown University, Providence, Rhode Island

JOHN MORRIS, Junior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

REEVES MORRISSON, Head, Analysis Section, Research Department, United Aircraft Corporation, East Hartford, Connecticut

PAUL L. MORTON, Associate Professor of Electrical Engineering, University of California, Berkeley, California

LEON MOSES, Graduate Student, Harvard University, Cambridge

Z. I. MOSESSON, Senior Actuarial Assistant, Prudential Insurance Company of America, Newark, New Jersey

ROBERT G. MOSS, Secretary, Electronic Calculator Committee, Metropolitan Life Insurance Company, New York

FREDERICK MOSTELLER, Assistant Professor of Mathematical Statistics, Harvard University, Cambridge

HAROLD M. MOTT-SMITH, Assistant Chief, Reactor Branch, Atomic Energy Commission, Washington, D.C.

CLIFTON F. MOUNTAIN, Director, Office of Statistical Research, Boston University, Boston

J. D. MOUNTAIN, International Telephone and Telegraph Company, New York

CARL F. MUCKENHOUPT, Head, Scientific Section, Office of Naval Research, Boston

RALPH E. MULLENDORE, Bureau of the Census, Washington, D.C.

G. G. MULLER, Bell Telephone Laboratories, New York

ROBERT E. MUMMA, Manager, Electrical Development Laboratory, National Cash Register Company, Dayton, Ohio

J. H. MUNCY, Electronic Engineer, Naval Air Development Station, Johnsville, Pennsylvania

MORRIS MUSKAT, Director of Physics Division, Gulf Oil Corporation, Pittsburgh, Pennsylvania

FRANKLIN G. MYERS, Design Specialist, Glenn L. Martin Company, Baltimore, Maryland

ROBERT A. NELSON, Research Engineer, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge

A. J. NEUMANN, Research Associate, University of Pennsylvania, Philadelphia, Pennsylvania

HERBERT B. NICHOLS, Science Editor, *Christian Science Monitor*, Boston

NATALIE N. NICHOLSON, Librarian, Harvard University, Cambridge

R. F. NICHOLSON, Data Utilization Laboratory, Griffiss Air Force Base, Rome, New York

EDWIN N. NILSON, Assistant Professor of Mathematics, Trinity College, Hartford, Connecticut

WILLIAM J. NOLAN, JR., Division of Industrial Coöperation, Massachusetts Institute of Technology, Cambridge

R. H. NOYES, Radio Engineer, Signal Corps Engineering Laboratories, Coles Laboratory, Red Bank, New Jersey

ALEXANDER NYMAN, Technical Advisor, Alden Products Company, Brockton, Massachusetts

H. NYQUIST, Bell Telephone Laboratories, New York

JAMES J. O'BEIRNE, Branch Chief, Control, Social Security Administration, Baltimore, Maryland

JOHN A. O'BRIEN, Research Engineer, Massachusetts Institute of Technology, Cambridge

L. A. OHLINGER, Project Engineer, Northrop Aircraft, Inc., Hawthorne, California
JOHN A. O'KEEFE, Mathematician, Army Map Service, Washington, D.C.
BRUCE OLDFIELD, Mathematician, U.S. Naval Ordnance Test Station, China Lake, California
THOMAS K. OLIVER, Acting Chief, Office of Air Research, Wright-Patterson Air Force Base, Ohio
R. D. O'NEAL, Aeronautical Research Center, University of Michigan, Ann Arbor, Michigan
ALLAN H. O'NEIL, Mathematician, U.S. Naval Proving Ground, Dahlgren, Virginia
ALEXANDER ORDEN, Research Associate, Massachusetts Institute of Technology, Cambridge
SIDNEY OVIATT, John Price Jones Company, Inc., New York
CHESTER H. PAGE, Electronics Consultant, National Bureau of Standards, Washington, D.C.
HENRY E. PAGE, Computation Laboratory, Harvard University, Cambridge
RALPH L. PALMER, Engineer, International Business Machines Corporation, Poughkeepsie, New York
WILLIAM N. PAPIAN, Research Assistant, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge
JOHN E. PARKER, President, Engineering Research Associates, Inc., St. Paul, Minnesota
G. B. PARKINSON, Senior Engineer, Raytheon Manufacturing Company, Newton, Massachusetts
JAMES PASTORIZA, Electronic Engineer, Air Force Cambridge Research Laboratory, Cambridge
GEORGE W. PATTERSON, Assistant Professor, Moore School, University of Pennsylvania, Philadelphia, Pennsylvania
ALFRED M. PEISER, Mathematician, Hydrocarbon Research, Inc., New York
C. L. PEKERIS, Institute for Advanced Study, Princeton, New Jersey
DAVID P. PERRY, Student, University of Pennsylvania, Philadelphia, Pennsylvania
C. G. PETERSON, District Agent, Marchant Calculating Machine Company, Boston
B. L. PFEFER, Project Manager, General Electric Company, Syracuse, New York
HARRY PFORZHEIMER, Head, Economic Section, Standard Oil Company (Ohio), Cleveland, Ohio
EDWARD W. PHARO, JR., Engineer, Philadelphia, Pennsylvania
J. R. PIERCE, Bell Telephone Laboratories, Murray Hill, New Jersey
ROBERT P. PINCKNEY, Student, Massachusetts Institute of Technology, Cambridge
SAMUEL PINES, Design Engineer, Republic Aviation, New York
CHARLES A. PIPER, Supervising Engineer, Bendix Aviation Research Laboratories, Detroit, Michigan
H. POLACHEK, Mathematician, Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland
ROBERT E. POPIEL, Computation Laboratory, Harvard University, Cambridge
H. PORITSKY, Consulting Engineer, General Electric Company, Schenectady, New York
WILLIAM A. PORTER, Computation Laboratory, Harvard University, Cambridge
JOHN T. POTTER, President, Potter Instrument Company, Inc., Flushing, New York

F. D. POWELL, Dynamics Engineer, Bell Aircraft Corporation, Buffalo, New York
BETTE PREER, Librarian, Boston Public Library, Boston
GEORGE G. PROULX, Computation Laboratory, Harvard University, Cambridge
C. C. PYNE, Harvard University, Cambridge
E. J. QUINBY, Director, Electronic Research and Development, Monroe Calculating Machine Company, Orange, New Jersey
JAN RAJCHMAN, Research Physicist, Radio Corporation of America, Princeton, New Jersey
JOHN H. RAMSER, Senior Physicist, Atlantic Refining Company, Philadelphia, Pennsylvania
EUGENE A. RASOR, Actuarial Mathematician, Social Security Administration, Washington, D.C.
ROBERT RATHBONE, Division of Industrial Coöperation, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge
A. G. RATZ, Research Engineer, University of Toronto, Toronto, Ontario, Canada
KONRAD RAUCH, Special Application Engineer, National Cash Register Company, Dayton, Ohio
RICHARD W. READ, Massachusetts Institute of Technology, Cambridge
DONALD REAM, Electronics Engineer, Navy Department, Washington, D.C.
MINA REES, Director, Mathematic Sciences Division, Office of Naval Research, Washington, D.C.
K. M. REHLER, Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts
GEORGE W. REITWIESNER, Mathematician, Ballistic Research Laboratories, Aberdeen Proving Ground, Aberdeen, Maryland
THEODORE V. RENTEL, Mari and Rentel Company, Inc., Boston
EDWARD REYNOLDS, Administrative Vice President, Harvard University, Cambridge
GEORGE E. REYNOLDS, Mathematician, Air Force Cambridge Research Laboratories, Cambridge
JAMES R. REYNOLDS, Office of Special Advisor to the President, Harvard University, Cambridge
ROBERT R. REYNOLDS, Assistant Professor, Oklahoma Agricultural and Mechanical College, Stillwater, Oklahoma
CHARLES RHEAMS, Senior Engineer, Raytheon Manufacturing Company, Boston
IDA RHODES, Mathematician, National Bureau of Standards, Washington, D.C.
HOPE RICE, University of North Carolina, Chapel Hill, North Carolina
OSCAR K. RICE, Professor of Chemistry, University of North Carolina, Chapel Hill, North Carolina
EDWIN S. RICH, Engineer, Massachusetts Institute of Technology, Cambridge
C. H. RICHARDS, Computation Laboratory, Harvard University, Cambridge
LOUIS N. RIDENOUR, Dean, Graduate College, University of Illinois, Urbana, Illinois
DONALD V. RIDER, Electronics Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

LEONARD D. RINALDI, Research Engineer, Cornell Aeronautical Laboratory, Buffalo, New York

E. K. RITTER, Research Engineer, University of Michigan, Ann Arbor, Michigan

GORDON A. ROBERTS, Manager, Future Demands Department, International Business Machines Corporation, New York

REX ROBERTS, Vice President, Transducer Corporation, Boston

JOHN W. ROCHE, Computation Laboratory, Harvard University, Cambridge

NATHANIEL ROCHESTER, Engineering Laboratory, International Business Machines Corporation, Poughkeepsie, New York

HAROLDO AZEVEDO RODRIGUES, Rio de Janeiro, Brazil

EDWARD ROGAL, President, Central Records, Inc., Boston

STANLEY ROGERS, Research Engineer, Convair, San Diego, California

THOMAS A. ROGERS, Associate Professor of Engineering, University of California at Los Angeles, Los Angeles, California

E. ROOT, Springfield, Vermont

MORRIS E. ROSE, Principal Physicist, Oak Ridge National Laboratory, Oak Ridge, Tennessee

JOSHUA H. ROSENBLOOM, Physicist, Naval Ordnance Laboratory, Washington, D.C.

CLARENCE ROSS, U.S. Naval Proving Ground, Dahlgren, Virginia

JOHN ROTHERY, Air Force Cambridge Research Laboratories, Cambridge

WALTER ROTMAN, Electronic Engineer, Air Force Cambridge Research Laboratories, Cambridge

B. G. H. ROWLEY, LT.-CDR., British Naval Staff, Electronics Liaison Officer, Washington, D.C.

D. M. RUBEL, LT.-CDR., USN, Office of the Secretary of Defense, Washington, D.C.

MILTON D. RUBIN, Research Director, G. A. Philbrick Researches, Inc., Boston

M. RUBINOFF, Institute for Advanced Study, Princeton, New Jersey

PHILIP RULON, Professor of Education and Acting Dean of the Faculty of Education, Harvard University, Cambridge

HEINZ RUTISHAUSER, Swiss Federal Institute of Technology, Zurich, Switzerland

DAVID RUTLAND, Engineer, National Bureau of Standards, Los Angeles, California

HERBERT E. SALZER, Mathematician, Computation Laboratory, National Bureau of Standards, Washington, D.C.

JOHN M. SALZER, Research Associate, Massachusetts Institute of Technology, Cambridge

ARTHUR L. SAMUEL, International Business Machines Corporation, Poughkeepsie, New York

PAUL A. SAMUELSON, Professor of Economics, Massachusetts Institute of Technology, Cambridge

JACKIE SANBORN, Computation Laboratory, Harvard University, Cambridge

BERNARD L. SARAHAN, Mathematician, Naval Research Laboratory, Washington, D.C.

EMIL D. SCHELL, Chief, Mathematic and Electronic Computer Branch, U.S. Air Force, Washington, D.C.

HAROLD A. SCHERAGA, Instructor, Cornell University, Ithaca, New York
MAX G. SCHERBERG, Chief of Applied Mathematics, Office of Air Research, Wright-Patterson Air Force Base, Dayton, Ohio
ROBERT SCHILLING, Department Head, General Motors Research Laboratory, Detroit, Michigan
KURT SCHNEIDER, Chemist-Metallurgist, Salem, Massachusetts
HENRY W. SCHRIMPF, Methods Analyst, Prudential Insurance Company of America, Newark, New Jersey
R. E. SCHUETTE, Project Engineer, Barber-Colman Company, Rockford, Illinois
WILLIAM T. SCOTT, Associate Scientist, Brookhaven National Laboratory, Upton, Long Island, New York
JOHN F. SCULLY, Radio Engineer, Air Force Cambridge Research Laboratories, Cambridge
RAYMOND J. SEEGER, Naval Ordnance Laboratories, White Oak, Silver Spring, Maryland
ROBERT R. SEEKER, JR., International Business Machines Corporation, New York
WARREN L. SEMON, Computation Laboratory, Harvard University, Cambridge
AMIR H. SEPAHBAN, Student, Harvard University, Cambridge
ROBERT SERRELL, Radio Corporation of America Laboratories, Princeton, New Jersey
EDGAR D. SEYMOUR, Development Engineer, Eastman Kodak Company, Rochester, New York
JACOB SHAPIRO, Physicist, Atomic Energy Commission, Elmhurst, New York
ROBERT F. SHAW, Engineer, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania
C. BRADFORD SHEPPARD, Engineer, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania
HERBERT SHERMAN, Chief, Systems Branch, Plans Office, AMC, Watson Laboratories, Red Bank, New Jersey
JACK SHERMAN, Mathematician, The Texas Company, Beacon, New York
CHARLES E. SHINN, Electronics Department, Burroughs Adding Machine Company, Philadelphia, Pennsylvania
BERNARD SHOOR, Research Engineer, Northrop Aircraft Company, Hawthorne, California
GEORGE SHORTLEY, Operations Research Office, Department of the Army, Ft. Leslie J. McNair, Washington, D.C.
ARNOLD SHOSTAK, Engineer, Office of Naval Research, Washington, D.C.
BYRON SHREINER, Assistant to Vice President, A. C. Nielsen Company, Chicago, Illinois
T. R. SILVERBERG, Engineer, Raytheon Manufacturing Company; Waltham, Massachusetts
THEODORE SINGER, Computation Laboratory, Harvard University, Cambridge
ROGER L. SISSON, Research Assistant, Massachusetts Institute of Technology, Cambridge
MRS. ROGER L. SISSON, Waltham, Massachusetts
C. J. SITTINGER, Consulting Engineer, Winchester, Massachusetts
MORTON L. SLATER, Senior Mathematician, Ordnance Research, Chicago, Illinois
DAVID SLEPIAN, Harvard University, Cambridge

ALBERT E. SMITH, Physicist, Office of Naval Research, Washington, D.C.

BRUCE K. SMITH, Junior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

CHARLES V. L. SMITH, Head, Computer Branch, Office of Naval Research, Washington, D.C.

DEXTER SMITH, Computation Laboratory, Harvard University, Cambridge

LEONARD W. SMITH, Computation Laboratory, Harvard University, Cambridge

H. P. SMITH, Research, Underwood Corporation, Hartford, Connecticut

V. G. SMITH, Professor of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada

R. M. SNOW, Senior Physicist, Martin-Hubbard Corporation, Boston

FRANCES E. SNYDER, Program Analyst, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

R. L. SNYDER, Branch Chief, Ballistic Research Laboratory, Aberdeen Proving Ground, Aberdeen, Maryland

SAMUEL S. SNYDER, Research Analyst, Department of the Army, Washington, D.C.

THOMAS G. SOFRIN, Engineer, Pratt & Whitney Aircraft Corporation, Hartford, Connecticut

AARON S. SOLTES, Electronics Engineer, Air Force Cambridge Research Laboratories, Cambridge

G. WALTER SPAHR, Manager of Engineering, Underwood Corporation, Hartford, Connecticut

AMBROSE P. SPEISER, Swiss Federal Institute of Technology, Zurich, Switzerland

JOHN W. H. SPENCER, Chief, Computing Branch, Geodetic Division, Army Map Service, Washington, D.C.

DANIEL SPILLANE, Computation Laboratory, Harvard University, Cambridge

BERNARD I. SPINRAD, Associate Physicist, Argonne National Laboratory, Chicago, Illinois

H. P. STABLER, Professor of Physics, Williams College, Williamstown, Massachusetts

JACK J. STALLER, Mechanical Engineer, Raytheon Manufacturing Company, Newton, Massachusetts

E. D. STANLEY, JR., CDR., USN, Munitions Board, Department of Defense, Washington, D.C.

S. W. STARK, Research Staff Member, Optical Research Laboratory, Boston University, Boston

CHARLES L. STEC, Civilian Assistant, Electrical Design Division, Bureau of Ships, USN, Washington, D.C.

HAROLD STEIN, Physicist, Signal Corps Engineering Laboratories, Long Branch, New Jersey

HAROLD H. STEIN, Engineer, University of Toronto, Toronto, Ontario, Canada

RICHARD STEPHENSON, Bell Telephone Laboratories, Murray Hill, New Jersey

D. L. STEVENS, Head of Computer Group, Sylvania Electric Products, Inc., Boston

GEORGE R. STIBITZ, Consultant, Burlington, Vermont

W. W. STIFLER, JR., Physicist, Engineering Research Associates, Washington, D.C.

E. E. ST. JOHN, Electronic Engineer, Nepa Project, Oak Ridge, Tennessee

MORTON J. STOLLER, Aeronautical Research Scientist, National Advisory Committee for Aeronautics, Langley Air Force Base, Virginia

JOSEPH J. STONE, JR., Associate Electrical Engineer, Fairchild Engine and Aircraft Corporation, Oak Ridge, Tennessee

JOHN STRAND, Research Engineer, University of Michigan, Ann Arbor, Michigan

H. L. STRAUS, Chairman of the Board, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania

J. C. STREET, Professor of Physics, Harvard University, Cambridge

PETER F. STRONG, Computation Laboratory, Harvard University, Cambridge

DANIEL R. STULL, Physical Research Laboratory, The Dow Chemical Company, Midland, Michigan

GEORGE C. SUMNER, Massachusetts Institute of Technology, Cambridge

ALFRED K. SUSSKIND, Research Assistant, Project Whirlwind, Massachusetts Institute of Technology, Cambridge

LOUIS SUTRO, Instructor, Department of Electrical Engineering, Tufts College, Medford, Massachusetts

RUTH W. SUTRO, West Medford, Massachusetts

HENRY W. SYER, Assistant Professor of Education, Boston University, Boston

A. H. TAUB, Professor, University of Illinois, Urbana, Illinois

NORMAN H. TAYLOR, Research Engineer, Massachusetts Institute of Technology, Cambridge

RICHARD TAYLOR, Assistant Professor, Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge

J. B. TEPE, Research Engineer, DuPont Experimental Station, Wilmington, Delaware

BENJAMIN J. TEPPING, Statistician, Bureau of the Census, Washington, D.C.

P. J. THEODORIDES, Visiting Lecturer on Aeronautical Engineering, Harvard University, Cambridge

HELENE THOMAN, Computation Laboratory, Harvard University, Cambridge

L. H. THOMAS, International Business Machines Corporation, New York

PAUL D. THOMAS, Mathematician, U.S. Coast and Geodetic Survey, Washington, D.C.

GEORGE W. THOMSON, Senior Chemical Mathematician, Ethyl Corporation, Detroit, Michigan

R. THORARENSEN, Electronic Laboratory, General Electric Company, Syracuse, New York

W. K. TRAUGER, Professor, State Teachers College, Potsdam, New York

IRVEN TRAVIS, Director of Research, Burroughs Adding Machine Company, Philadelphia, Pennsylvania

G. W. TRICHEL, Chrysler Corporation, Detroit, Michigan

ARTHUR D. TROTTENBERG, Assistant to Administrative Vice President, Harvard University, Cambridge

LEDYARD R. TUCKER, Director of Statistical Analysis, Educational Testing Service, Princeton, New Jersey

JOHN W. TUKEY, Bell Telephone Laboratories, Murray Hill, New Jersey

L. R. TURNER, Aeronautical Research Scientist, National Advisory Committee for Aeronautics, Cleveland, Ohio

MRS. L. R. TURNER, Cleveland, Ohio
ARTHUR W. TYLER, Physicist, Eastman Kodak Company, Rochester, New York
STANISLAUS M. ULAM, Los Alamos Scientific Laboratory, Los Alamos, New Mexico
JOHN H. VAN VLECK, Professor of Mathematical Physics, Harvard University, Cambridge
CLEMENT J. VAN VLIET, Statistician, U.S. Navy Electronics Laboratory, San Diego, California
FRANK L. VERDONCK, Computation Laboratory, Harvard University, Cambridge
FRANK M. VERZUH, Research Associate, Massachusetts Institute of Technology, Cambridge
HUGH WAINWRIGHT, Sales Engineer, Sylvania Electric Products, Inc., Boston
J. H. WAKELIN, Associate Director of Research, Textile Research Institute, Princeton, New Jersey
AN WANG, Computation Laboratory, Harvard University, Cambridge
CLIFFORD A. WARREN, Bell Telephone Laboratories, Murray Hill, New Jersey
CHAUNCEY W. WATT, Engineer, Massachusetts Institute of Technology, Cambridge
EDWARD A. WEEKS, JR., Editor, *The Atlantic Monthly*, Boston
JAMES R. WEINER, Chief Electrical Engineer, Eckert-Mauchly Computer Corporation, Philadelphia, Pennsylvania
HERBERT G. WEISS, Senior Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts
JOSEPH WEINSTEIN, Mathematician, Signal Corps Engineering Laboratories, Ft. Monmouth, New Jersey
W. GORDON WELCHMAN, Member of Research Staff, Massachusetts Institute of Technology, Cambridge
DAVID R. WELLER, Electronic Engineer, Data Utilization Laboratory, Griffiss Air Force Base, Rome, New York
EVERETT T. WELMERS, Chief of the Dynamics Group, Bell Aircraft Corporation, Buffalo, New York
INA W. WELMERS, Instructor in Mathematics, University of Buffalo, Buffalo, New York
ALBERT WERTHEIMER, Engineer, Bureau of Ordnance, Navy Department, Washington, D.C.
CHARLES F. WEST, Section Head, Raytheon Manufacturing Company, Waltham, Massachusetts
HARALD M. WESTERGAARD, Professor of Civil Engineering, Harvard University, Cambridge
L. D. WHITELOCK, Electronics Engineer, Bureau of Ships, Navy Department, Washington, D.C.
C. A. WHITTEN, Chief, Section of Triangulation, U.S. Coast and Geodetic Survey, Washington, D.C.
C. ROBERT WIESER, Research Engineer, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge
HOLLY B. WILKINS, Computation Laboratory, Harvard University, Cambridge
ROBERT E. WILKINS, Computation Laboratory, Harvard University, Cambridge
DAVID A. WILKINSON, Project Manager, General Electric Company, Syracuse, New York

SAMUEL B. WILLIAMS, Consulting Electrical Engineer, Brooklyn, New York

DEAN H. WILSON, Research Associate, Aeronautical Research Center, University of Michigan, Ann Arbor, Michigan

EDWIN B. WILSON, Consultant, Office of Naval Research, Boston

MORRIS WINICK, Chief Engineer, Transducer Corporation, Boston

WILLIAM WOLFSON, Electronic Engineer, Raytheon Manufacturing Company, Waltham, Massachusetts

RICHARD D. WOLTMAN, U.S. Naval Proving Ground, Dahlgren, Virginia

WAY DONG WOO, Assistant Professor, Computation Laboratory, Harvard University, Cambridge

H. A. WOOD, Supervisor of Dynamic Analysis, Chance Vought Aircraft, Dallas, Texas

MARSHALL K. WOOD, Assistant Director, Program Standards and Cost Control, U.S. Air Force, Washington, D.C.

RALPH V. WOOD, JR., Air Force Cambridge Research Laboratories, Cambridge

WILLIAM W. WOODBURY, Northrop Aircraft Company, Hawthorne, California

L. F. WOODRUFF, Chief Scientific Advisor to Director of Intelligence, Washington, D.C.

R. W. WOODWARD, Underwood Corporation, Hartford, Connecticut

HOWARD WRIGHT, Electronic Scientist, National Bureau of Standards, Washington, D.C.

ROBERT H. YOUDEN, Student, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge

DAVID M. YOUNG, JR., Graduate Student, Harvard University, Cambridge

PATRICK YOUTZ, Research Associate, Massachusetts Institute of Technology, Cambridge

H. I. ZAGOR, Chief Physicist, Reeves Instrument Corporation, New York

SERGE J. ZAROODNY, Engineer, Ballistics Research Laboratory, Aberdeen Proving Ground, Aberdeen, Maryland

G. K. ZIPF, University Lecturer, Harvard University, Cambridge

FIRST SESSION

Tuesday, September 13, 1949
10:30 A.M. to 12:00 P.M.

OPENING ADDRESSES

Presiding

Howard H. Aiken

Director of the Computation Laboratory

EDWARD REYNOLDS

HARVARD UNIVERSITY

It is a great privilege and gives me great personal pleasure to have the honor of bringing the greetings and warm welcome of the President and Fellows of Harvard College to this large group of distinguished guests visiting the University on the occasion of this symposium on large-scale digital calculating machinery at the Harvard Computation Laboratory. We are all sorry that President Conant could not personally welcome you here and present these greetings, but he is on the West Coast keeping engagements made more than a year ago. I regret this necessity for his absence, but express his hope that you will find your visit here interesting and productive, and our hospitality cordial.

This is the third such ceremony in connection with Harvard's Computation Laboratory. The first, about five years ago during wartime, dedicated Mark I, a highly significant development in this new field, which was then generously presented to Harvard by the International Business Machines Corporation and temporarily located in our Cruft Laboratory, representing the fruit of several years of collaboration by Professor Aiken of Harvard with the leading research men in the IBM organization. Mark I, although in some respects overshadowed by subsequent developments here and elsewhere, is still the reliable old workhorse of the Laboratory which has rendered extremely valuable service to the armed forces. From its initial operation until the end of last year, it has had the generous support of the United States Navy, which we gratefully acknowledge. More recently, the United States Air Force and the Atomic Energy Commission have shared this support. All three of these agencies of the Government have been most generous and understanding in helping us to broaden the scope of the problems to which it has been applied and thus to broaden the field of interest and usefulness of this type of machinery.

The second such ceremony, early in 1947, dedicated this new laboratory building and offered for inspection the Mark II, then being completed and tested for the Navy. Shortly thereafter, Mark II was delivered to the Navy and installed at Dahlgren Proving Ground in Virginia.

Even though we understand that, with its significant advances in speed and capacity over Mark I, Mark II has proved its usefulness, it has not satisfied our good friends in the Bureau of Ordnance of the United States Navy, who have continued their generous support of the research of this Laboratory and are now jointly with the Harvard Computation Laboratory sponsoring this third symposium at which we have the pleasure of unveiling Mark III, also destined for delivery in the near future to the Dahlgren Proving Grounds.

We feel that we may properly take some pride in the quality of research being carried on in this Laboratory. The distinguished character of the talent attending this symposium

supports us in this feeling. We feel that rather too much time has been devoted to the development of actual machinery growing out of this research, and have some hope that one of the by-products of this and other meetings may be to stimulate the interest of others in this phase of the application of our research and thus to eliminate the need for this activity on the part of our staff. Certainly the manufacture of parts and even the assembly thereof are not activities for which we are fitted or which we wish to pursue, except possibly in the production of a machine for our own use within the University.

As a layman participating actively in the administrative problems which arise out of the organized group research that has become such an active part of research at universities in recent years, I am tremendously pleased at the evidences of awakening interest in the usefulness of these newest appliances in widely scattered fields of research. The inescapable application of mathematics in practically every field of human endeavor makes it seem important to us that the understanding of the availability and usefulness of the developments being made here and in other mathematical research laboratories, as aids in all other fields of research, be spread as widely as possible; and we are therefore particularly pleased that the latter part of the program for this symposium is devoted to discussions of the relation of the work of this Laboratory to research in a broad range of subjects.

Increased awareness of the usefulness of these new tools in the field to activities in other branches of science inevitably increases the demand for the already inadequate number of men and women who are educated not only in the theories of design and operation of such machinery but in the understanding of their applicability. This emphasizes the other great responsibility of the staff of this Laboratory—meeting their obligations as teachers to provide the instruction required for the training and development of personnel interested in these lines. Here again we gratefully acknowledge the understanding support of our friends in the Bureau of Ordnance and in the Air Force. While we are endeavoring to develop support for this program from other sources and to obtain permanent endowment for the Laboratory, the understanding contractual support from these Government sources has been and continues to be invaluable.

REAR ADMIRAL F. I. ENTWISTLE, USN

NAVY DEPARTMENT, BUREAU OF ORDNANCE

It is with distinct pleasure that the Bureau of Ordnance joins hands with Harvard University in sponsoring this Symposium on Large-Scale Digital Calculating Machinery. On behalf of the Bureau of Ordnance I take great pleasure in welcoming you to these meetings. Your presence—the presence of so many distinguished scientists—assures us that much value will be gained by us all from the deliberations and discussions which will take place during these next few days.

During the past few years the relation between the Bureau of Ordnance and Harvard University has been unusually close and cordial. This happy relation has given the Navy the benefit of Harvard's talents and facilities and has led to the development of greatly improved computing machinery and methods. Harvard University is most fortunate in having on its staff Professor Howard Aiken, under whose unusual leadership, energy, and ability the Computation Laboratory has grown.

As we in the Bureau of Ordnance look back on the computing problems with which we were faced prior to the First World War, we find that large-scale computations arose chiefly in connection with the problems of ballistics—problems in which we were principally concerned with construction of range tables for seagoing gun systems with limited angles of elevation. In those days, one computer (and by computer I mean a man with a slide rule, log book, and a set of Engel's *Ballistic Tables*) handled all such computations. It probably took the impetus, the acceleration, and the foreboding of World War II to permit conception of the machine that was originated here and is known as the Mark I.

In the olden days when we were youngsters, and possibly a bit more impatient, that one man (that computer) with his slide rule and his tables gave us a series of curves or figures in a book, and we were to go out to put them to use. Frequently we discovered that we did not know how to do this or we found that the tables were incorrect.

The availability of accurate tables in time of war is very important indeed. When the recent war came along with its bombings, rocket firing, and use of heavier guns for antiaircraft and bombardment, we found our range tables insufficient. In fact, we were about 500 range tables behind. In the course of some years, that figure was decreased to 350; but still it was a problem of one man, one slide rule, and tables. Naturally, 500 tables would equal 500 men or 500 years; even with 500 men we would still be at least a year behind.

World War II indicated by great numbers—in tonnages, people, dollars—the magnitude of effort required to fight a modern war. I believe it showed us that we can no longer afford to fight wars of that magnitude. Many of us in the armed services have come to realize that our job is not to fight wars but to prevent them. If we had realized this in the period from

1925 to 1930, we might have dissuaded the Japanese in 1939 from exerting the effort that was subsequently shown. By keeping us prepared to carry through a war, these machines may help us to prevent wars.

This science of computation, which we have come here to discuss, has grown up from the association of people requiring such machines and their results with other people willing to incorporate themselves into the effort to design and to build such machines. The fact that we can collaborate and coördinate our efforts with a university such as Harvard, and can arrange for the services of the laboratory here for our mutual benefit, should assure us of the continuation of our so-called democracy. In the installation of computers, the time factor and accuracy of the machine are certainly important. But more important to my mind is the coördination of the university and the military. This is, in itself, a step forward in the university's aim of first teaching the individual and then going further to educate the country.

May I again express the continuing deep interest of the Bureau of Ordnance in this important subject of large-scale calculating machinery, and its applications, which may better equip the Bureau of Ordnance of the Navy Department to carry on its work in national defense. Both for myself and for the Chief of the Bureau of Ordnance, I wish to express our appreciation of the close and wholehearted coöperation on the part of Harvard University and the Computation Laboratory, and to acknowledge great and significant contributions in the development of computing machines and methods that they have made through their skills, their talents, and their facilities. May I add further that the Chief of the Bureau of Ordnance and I will carry on all we can undertake and accomplish to continue this type of collaboration resulting in this broader effort to prevent wars.

HOWARD H. AIKEN

HARVARD UNIVERSITY

As Admiral Entwistle has remarked, for five years the staff of the Computation Laboratory has been engaged in the construction of automatic computing machinery for the Bureau of Ordnance. Without the Bureau's constant support, our share of this research could never have been undertaken. We look upon the completion of Mark III as representing the end of a phase in the development of this subject.

I have often remarked that if all the computing machines under construction were to be completed, there would not be staff enough to operate them. Instruction in computing machinery represents one of the more aggravated aspects of a generally recognized problem in technical education. We feel that the further development of mathematical methods and the extended use of computing machines in the various fields represented by speakers here are those points at which levers should be placed to make the greatest possible advance in computer research. Only by completing computing machines and then operating them can the operating experience and experimental results be obtained that are so essential as a point of departure in passing from one design to another. Therefore, at our laboratory we have decided not to undertake the construction of any more large-scale computing machines with the exception of one, which we hope to build for our own use and keep at Harvard.

There is an ever-increasing number of industries interested in constructing computing machines outside the universities. In applying computing machinery to new and different fields, many proposals have been made, ranging all the way from devices for an automatic continuous audit, an automatic continuous inventory, down through an automatically operated insurance office, public-utility billing department, department-store accounting system, to more specific and less general accounting-machine components. Other proposals have included airline ticket-inventory systems, similar devices for railroad reservations, and automatic railroad ticket-vending machines. On the technical side, machines have been proposed involving automatic computers in connection with air-traffic control, airport control and almost every other manufacturing operation up to and including the automatic factory. But until our universities are able to offer well-rounded programs in numerical methods and the application of computing machinery to prepare men to operate these machines, the success of many of the proposed industrial programs will not be realized.

I should like to take this opportunity to express the appreciation of our staff to the Bureau of Ordnance for its support throughout these years and, more than that, for the privilege of pleasant associations which we have had with the representatives of that Bureau. It has been a great pleasure to work with them throughout the construction of both Mark II and Mark III Calculators, and we have built up an association which I have every reason to believe will continue.

SECOND SESSION

Tuesday, September 13, 1949
2:00 P.M. to 5:00 P.M.

RECENT DEVELOPMENTS IN COMPUTING MACHINERY

Presiding

Mina Rees
Office of Naval Research

The Mark III Calculator

BENJAMIN L. MOORE

Harvard University

The large-scale digital computing machine known as Mark III has been built by the Computation Laboratory of Harvard for the Bureau of Ordnance of the Navy Department. It is to be installed at the Naval Proving Ground, Dahlgren, Virginia. The construction work has been completed and the machine is now under test.

The decimal number system is used throughout the entire machine. Normal operations are carried out with 16 digits. Provision is made, however, to use 32, 48, or more digits when needed. The operating decimal point is manually set by the operator in one of six positions. In addition, under control of the sequencing unit, the operator may choose at any time one of three locations of the decimal point. In order to reduce the size of the memory, digits are stored in the coded form where four binary digits are used to represent a decimal digit. Figure 1 shows the system that has been adopted, where the weights of the four binary digits are 1, 2, 4, and 2, respectively. It should be noted that the sum of these weights is 9 and therefore the nine's complement of any digit may be obtained by changing zeros to ones and vice versa. This is quite a convenience electronically where a positive voltage may be used for a one and a negative voltage for a zero. The nine's complement can then be obtained merely by inverting the signal. These complements are used for subtraction in this machine.

Decimal Digit	Coded Decimal Notation 2* 4 2 1
0	0 0 0 0
1	0 0 0 1
2	0 0 1 0
3	0 0 1 1
4	0 1 0 0
5	1 0 1 1
6	1 1 0 0
7	1 1 0 1
8	1 1 1 0
9	1 1 1 1

FIG. 1. Coded decimal number system used in Mark III.

The number-storage system consists of eight rotating drums whose surfaces are coated with a thin layer of magnetic material. Information recorded in the form of a small magnetic dipole during one revolution may be played back on any succeeding revolution. A zero is represented by a magnetic dipole oriented in one direction and a one by the opposite orientation. New information is recorded directly over the old making erasure of the surface unnecessary. The played-back voltage signal is double ended, with the positive voltage first for one orientation and the negative first for the opposite orientation of the dipole. Figure 2 is a photograph of a typical cathode-ray oscillograph pattern having two negative-first pulses in a group of positive-first pulses. Figure 3 is similar except that the pulses have been reversed 4×10^8 times. This photograph was taken upon completion of a test to determine whether

the noise background increased or the pulses changed shape after many reversals. It is easy to see that the two pictures are almost identical.

Using a single magnetic head on a channel or track, the access time to any one number is the time for one revolution of the drum. At the expense of more heads the access time can

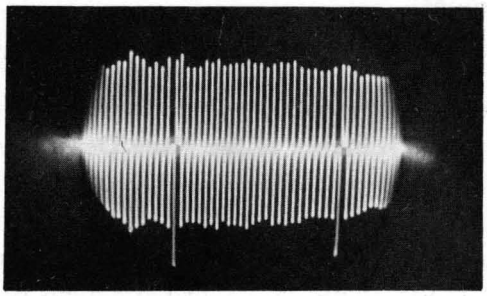


FIG. 2. Oscillograph trace of a typical playback-pulse pattern.

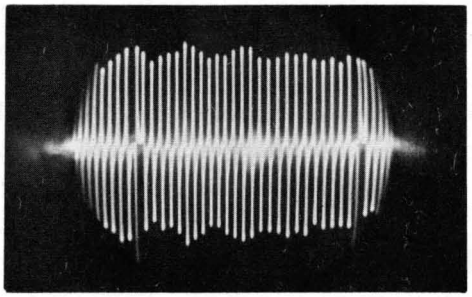


FIG. 3. Oscillograph trace of a typical playback-pulse pattern after 4×10^8 reversals.

be decreased. This machine uses two playback heads per track, so that the access time is reduced to the time of one half revolution. Since it is electrically more convenient, separate heads are used for recording and playback. Thus each binary track contains two record

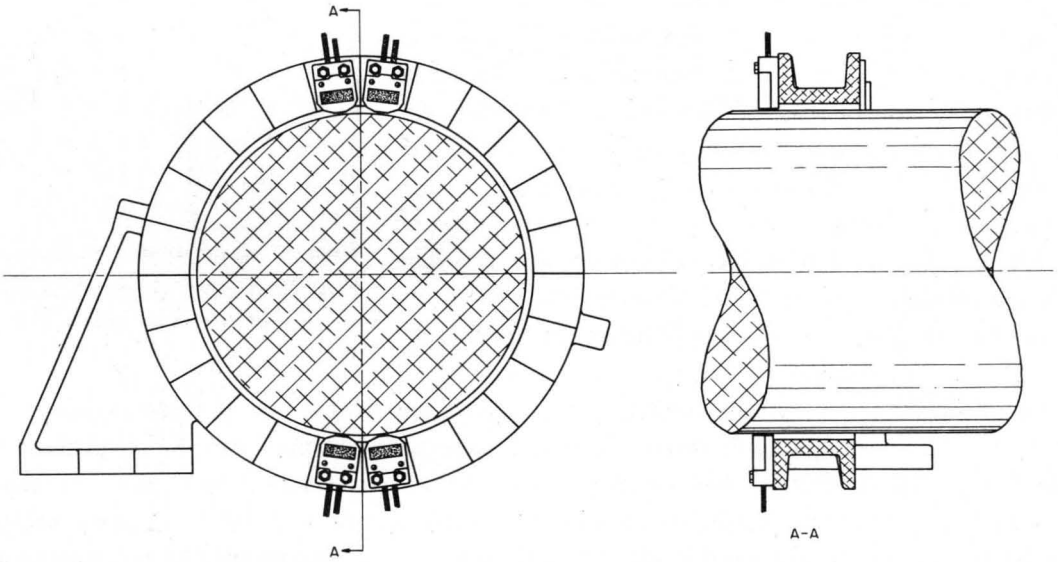


FIG. 4. Cross section of a binary magnetic storage channel.

heads as well as two playback heads. Figure 4 shows a typical cross section of a binary channel. Figure 5 shows the drum storage unit.

As it is convenient to have all components of a decimal digit available simultaneously, four parallel binary channels are used to represent one decimal channel where the binary

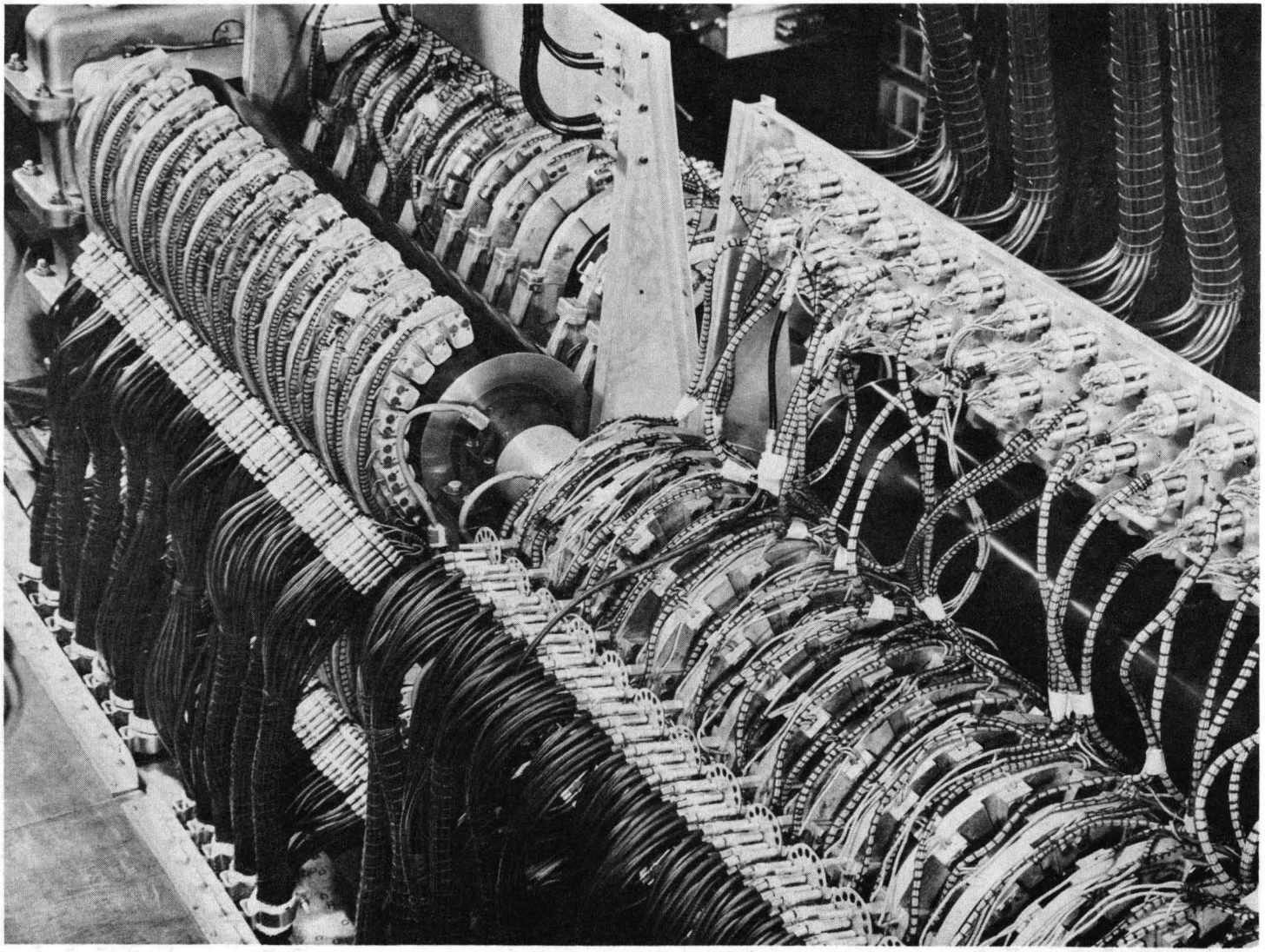


FIG. 5. View of the assembled magnetic-drum storage unit.

channels will have the weights 1, 2, 4, 2, respectively. The digits of a given number and the numbers themselves are stored serially around the periphery of the drum. The main storage system uses a pulse density of ten pulses per inch and stores ten 16-digit numbers in a decimal channel. To extend the storage-system capacity beyond ten numbers, parallel channels are used so that the selection of a given number involves the selection of a channel as well as a time selection, as the drum moves past the playback head.

The arithmetic unit of this machine is electronic and contains an adder, a multiplier, and certain sensing units. It contains no divider, since this operation is accomplished by an iteration process. The unit is serial in operation so that it is necessary to have the corresponding

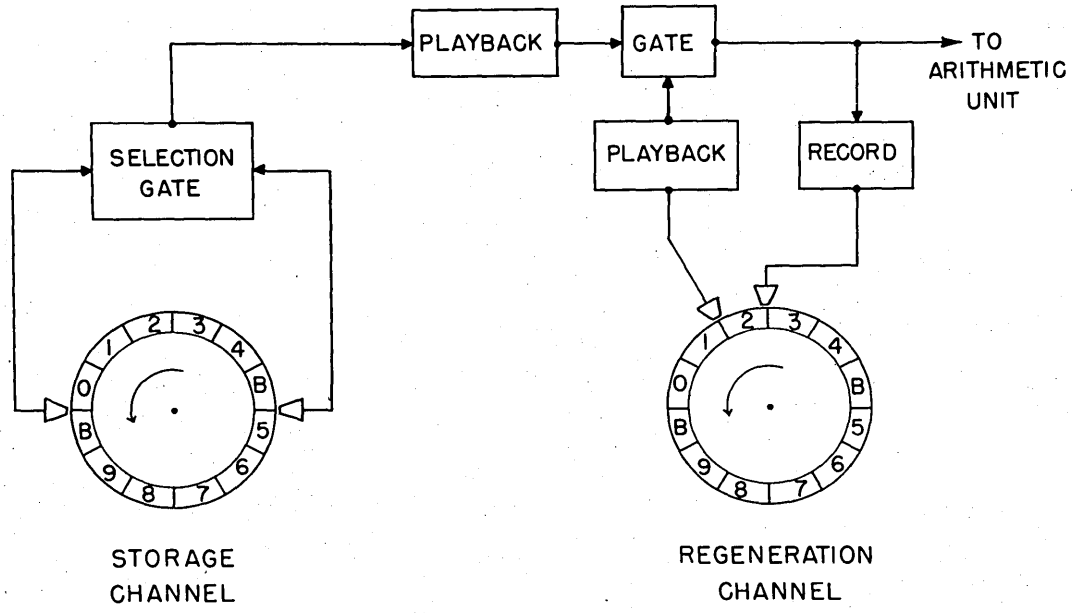


FIG. 6. Block diagram of a regeneration channel.

digits of two numbers to be added available simultaneously. Since, in general, two numbers will not necessarily be played back from the drum in the same time phase, some provision must be made to put them in phase. Figure 6 illustrates how this is done. On the left is the storage channel where a selection is made of the head to be used, depending on where the number is located and the phase of the drum. At the appropriate number time the gate opens and records the desired number on the regeneration channel. By the time the last digit is recorded the first digit is being played back by the regeneration playback so that the gate lets this number through to be recorded again. Thus the number is recorded 12 times around the channel and is available at any time thereafter. The two blank spaces on the storage channel provide time for switching operations and no numbers are stored in this interval.

With this brief picture of the magnetic storage system, let us examine the organization of the whole machine, a block diagram of which is shown in Fig. 7.

The storage system is divided into two parts. One, called slow storage, has a capacity of 4000 16-digit numbers and selection of channels is made by relays that are relatively slow. However, provision is made for transferring 20 numbers at a time from slow storage to the other section, which is called fast, where selection of numbers can be done at electronic

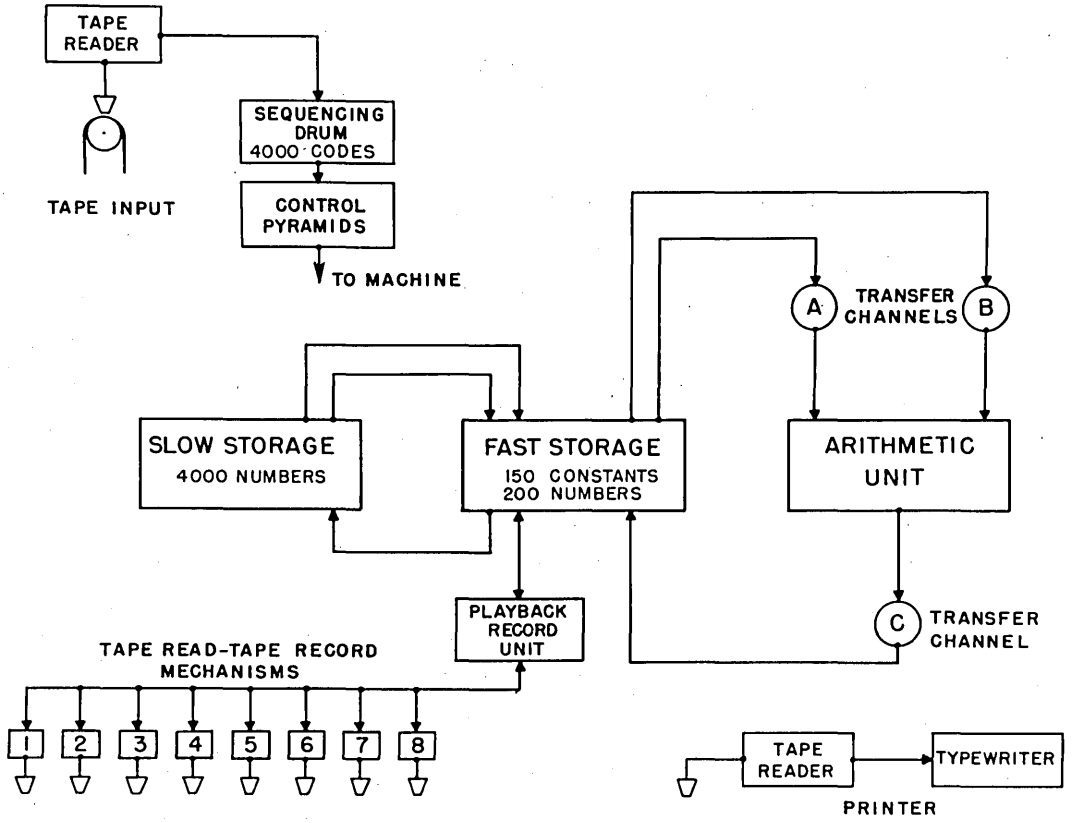


FIG. 7. Block diagram showing organization of the calculator.

speeds. Ten numbers at a time are transferred from the fast storage to slow. The slow section is mainly used for storage of functions.

The fast storage has a capacity of 200 numbers in addition to 150 permanent constants. The constants are used for computing functions such as $1/x$, $1/\sqrt{x}$, $\cos x$, $\log x$, antilog x, $\tan^{-1} x$.

The basic cycle of the machine is the access time to the storage, namely, one half revolution of the drum. Each cycle the machine delivers two numbers to the arithmetic unit over the two parallel busses A and B and returns the previous result to the storage. One addition can be performed each cycle, while multiplication requires 3 cycles. As the speed of the drums

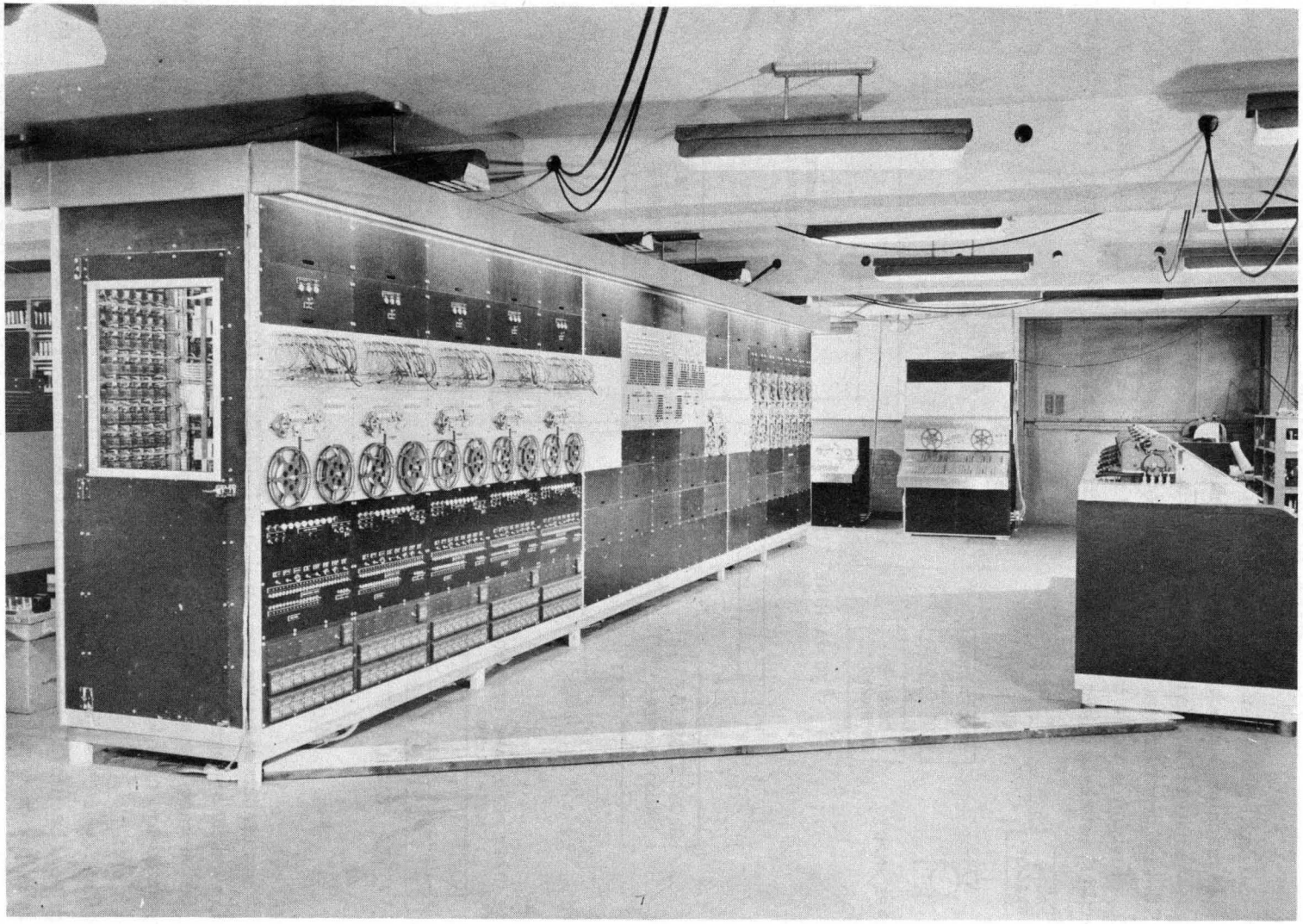


FIG. 8. View of the Mark III Calculator.

is slightly less than 7200 rev/min the cycle is about 4.2 msec, that is, addition requires 4.2 msec and multiplication 12.6 msec.

Numbers are fed into and recorded out of the machine by means of magnetic paper tape. There are eight mechanisms, any one or any combination of which can be set to read into or out of the machine. Recorded tapes are run through a tape reader which in turn operates

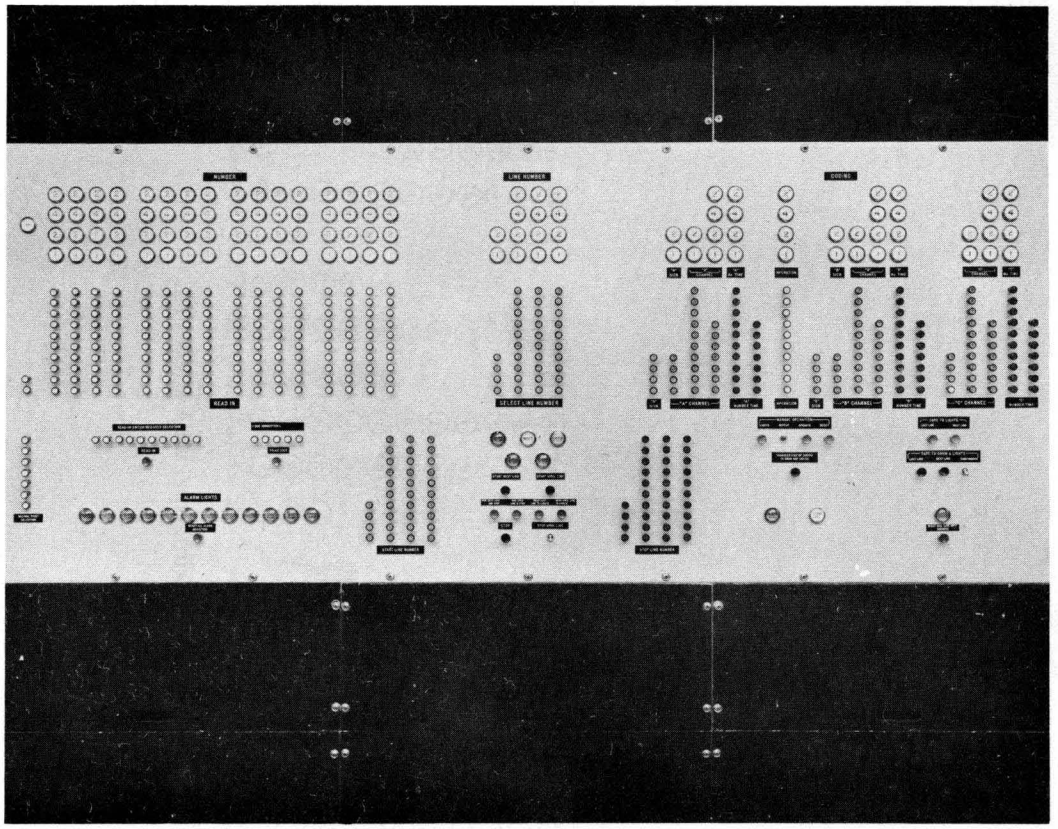


FIG. 9. Main control panel.

an electric typewriter to put the final results on the printed page. To handle the output the machine is equipped with five independent tape readers and typewriters.

There are essentially no checks built into the machine, dependence being put on mathematical checks. However, in transferring numbers from the machine to the printed page via the magnetic tape no mathematical checks are readily available. Therefore, to insure accuracy of this transfer, the output numbers from the machine are recorded into two separate channels on the tape by different sets of equipment. Before printing, the numbers from each channel of the tape are compared and if they are not identical the typewriter rings an alarm and stops.

Sequencing commands are stored on a separate drum turning at a much lower speed,

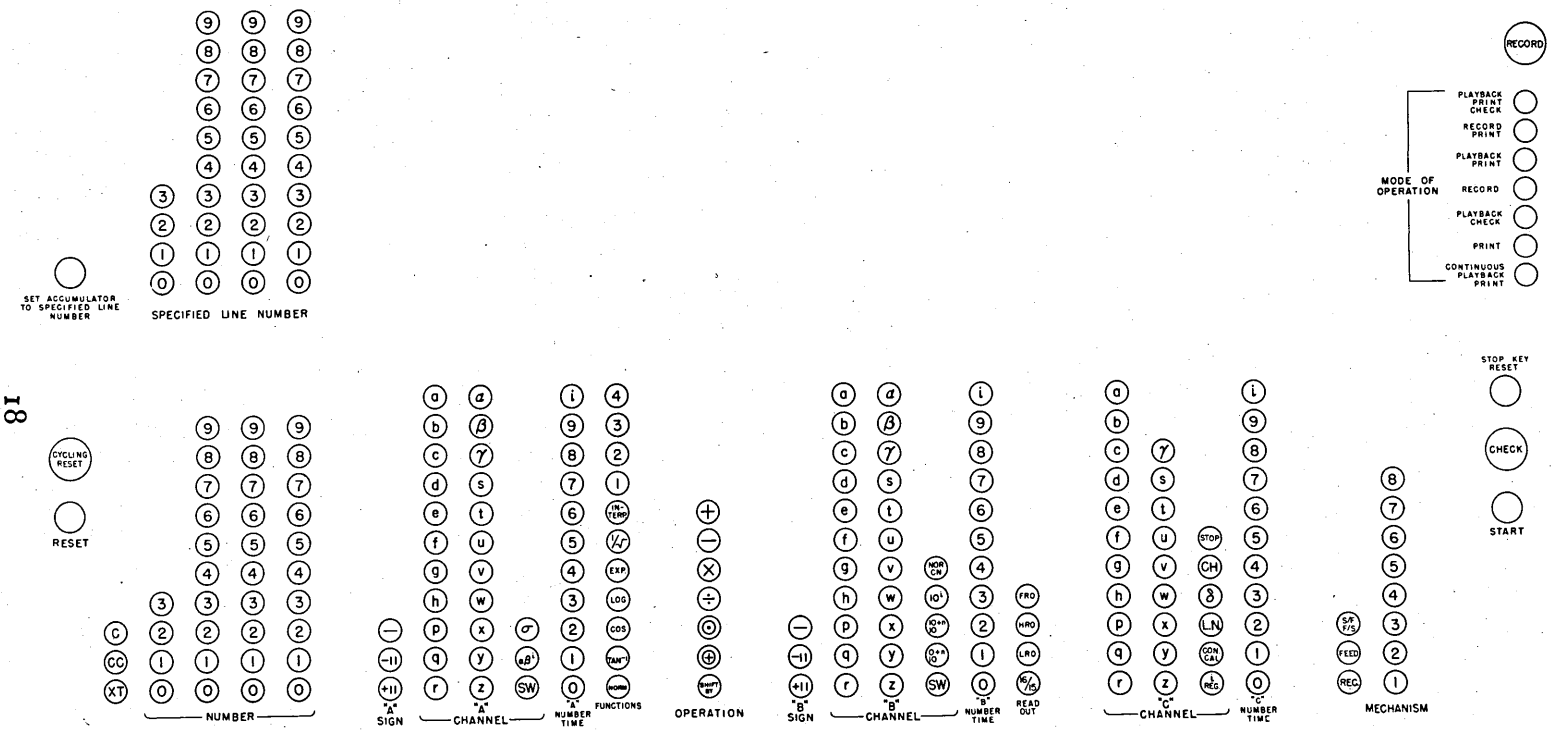


FIG. 10. Coding-machine keyboard.

approximately 1800 rev/min. Provision is made on this drum for storing 4000 lines of coding or sequencing commands. A line of coding will in general consist of the commands necessary to select two numbers from the storage, perform an operation, and return the resultant to storage. It is possible to jump from any line of coding to any other line with a maximum loss of a little more than 8 cycles and an average loss of about 4 cycles.

Figure 8 is a picture of the front of the machine. On the right facing the machine the five typewriters can be seen. The first five panels on the left are the tape readers that control the typewriters. In the center of the machine can be seen the main control panel. At the far end are the eight magnetic-tape input and output mechanisms. In the background can be seen the coding machine and the number-tape preparation unit.

Figure 9 is a view of the main control panel. On the left is a group of switches that enable the operator to feed numbers into the machine manually. Above these switches are a set of lights into which numbers, located anywhere in the machine, can be read. The center section contains controls for starting and stopping the machine. Controls are provided on the right to enable the operator to perform all operations manually in troubleshooting.

As is only too well known, the task of preparing a problem for machine solution is in many cases quite laborious. Even after the numerical analysis is completed, there still remains the work of translating the mathematical symbols and operations into a language the machine can understand. To reduce the work required in this part of the problem preparation, a special coding machine, whose keyboard is shown in Fig. 10, has been constructed. The storage registers in the fast storage are assigned letters and subscripts. If one uses these letters and subscripts in the numerical analysis, then it is a simple matter to operate the keyboard, thereby recording into a length of magnetic tape the necessary commands for carrying out the required operations. This tape may then be stored until the machine is available for solution of this problem, at which time the information on the tape is transferred to the sequencing drum. Provision is made for printing a copy of the coding commands for use of the operator in monitoring the problem.

Space does not permit a more detailed description of this coding machine. However, it should be pointed out that to operate the machine it is only necessary to know a few simple rules. In fact, many of the operations are obvious from the labels on the keyboard. It is the opinion of the staff of this Laboratory that this coding machine, which eliminates much of the labor in preparing a problem, represents a significant advance in the field of machine computation.

In conclusion, we would like to express our appreciation to the Bureau of Ordnance of the Navy Department whose interest and support have made this machine possible. It should also be pointed out that this machine is not the work of only a few individuals, but is the result of the combined effort of the entire staff of the Computation Laboratory.

The Bell Computer, Model VI

ERNEST G. ANDREWS

Bell Telephone Laboratories

It is customary in meetings on digital calculating machinery to listen to engineers extolling their latest creation with terms such as "a new giant brain" or "a machine that thinks."

I do not wish to criticize my colleagues in this fascinating field of engineering for using these metaphors because I have frequently indulged in the practice myself. However, a discordant note has been sounded in these meetings on occasion. This note, when translated into smooth, inoffensive English says, "Such claims by the engineers are not altogether in accordance with the facts." The thought behind these accusations has been aptly expressed by our own Dr. W. Bode, who, when speaking of our own Bell Laboratories Computers, said, "They are like very accurate but very dumb computresses." He was referring to the need for spelling out every elementary detail when programming a new problem. This criticism was also expressed by others, including Dr. George R. Stibitz when he addressed the first of these symposia. Partly as a consequence of all of these comments and partly because of the nature of the problems to be solved, the Bell Laboratories Model VI[1] relay computer has been endowed with more intelligence than its predecessors. But before delving into the details of this particular phase of the design let us take a broad look at the Model VI.

This computer has been placed in operation at the Bell Laboratories Murray Hill building. In many respects it resembles the Model V relay computers at the Ballistic Research Laboratory at Aberdeen, Maryland, and at the Laboratory of the National Advisory Committee on Aeronautics at Langley Field, Virginia. Complete descriptions of these computers have been presented by Dr. F. L. Alt and Mr. S. B. Williams.

The Model VI computer consists of two principal parts—the remote-control stations and the computing equipment. Figure 1 shows one of the remote-control stations. The three pieces of apparatus shown are types that are used extensively in Teletype printer telegraph systems, with minor changes to adapt them to computer operation. The printer on the movable table records the answers to the problems; the hand perforator is used for punching the data on the problem tapes and the tape reader transmits the data for the problems to the computing equipment.

Figure 2 shows part of the computing-room equipment. This part of the computer is made up of twelve bays of equipment consisting almost entirely of the heavy-duty-type relays used extensively in earlier computers and in telephone dial systems central-office equipment. The frames have light-gray enamel finish and other characteristics which make them resemble equipment that is found in the modern telephone office. There are about 4300 relays used, 86 cold-cathode tubes, and relatively small amounts of other miscellaneous apparatus.

An indicator and test panel with approximately 600 small lamps is provided for showing

the progress of various computing operations and for showing the numbers in various parts of the computer. It is illustrated in Fig. 3. This panel also has provisions for manually

FIG. 1. Remote-control station equipment.

inserting instructions into the computing program where necessary to make corrections in programming. These facilities are also used for making tests under closely controlled conditions.

Table 1. Comparison of number representations in Model VI and Model V computers.

Item of Comparison	Model VI	Model V
Number of digits for a number	3, 6, or 10	1, 2, 3, 4, 5, 6, or 7
Form of number: notation for π as an example	$+ 3.14159\ 2654 \times 10^{+00}$	$+ 0.3141\ 593 \times 10^{+01}$
Maximum number*	$\pm 9.99999\ 9998 \times 10^{+19}$	$\pm 0.9999\ 998 \times 10^{+19}$
Minimum number†	$\pm 1.00 \times 10^{-19}$	$\pm 0.1 \times 10^{-19}$

* The next higher number is the calculator's concept of infinity.

† The next lower number is the calculator's concept of zero. However, smaller numbers with the -19 exponent are possible with special problem coding.

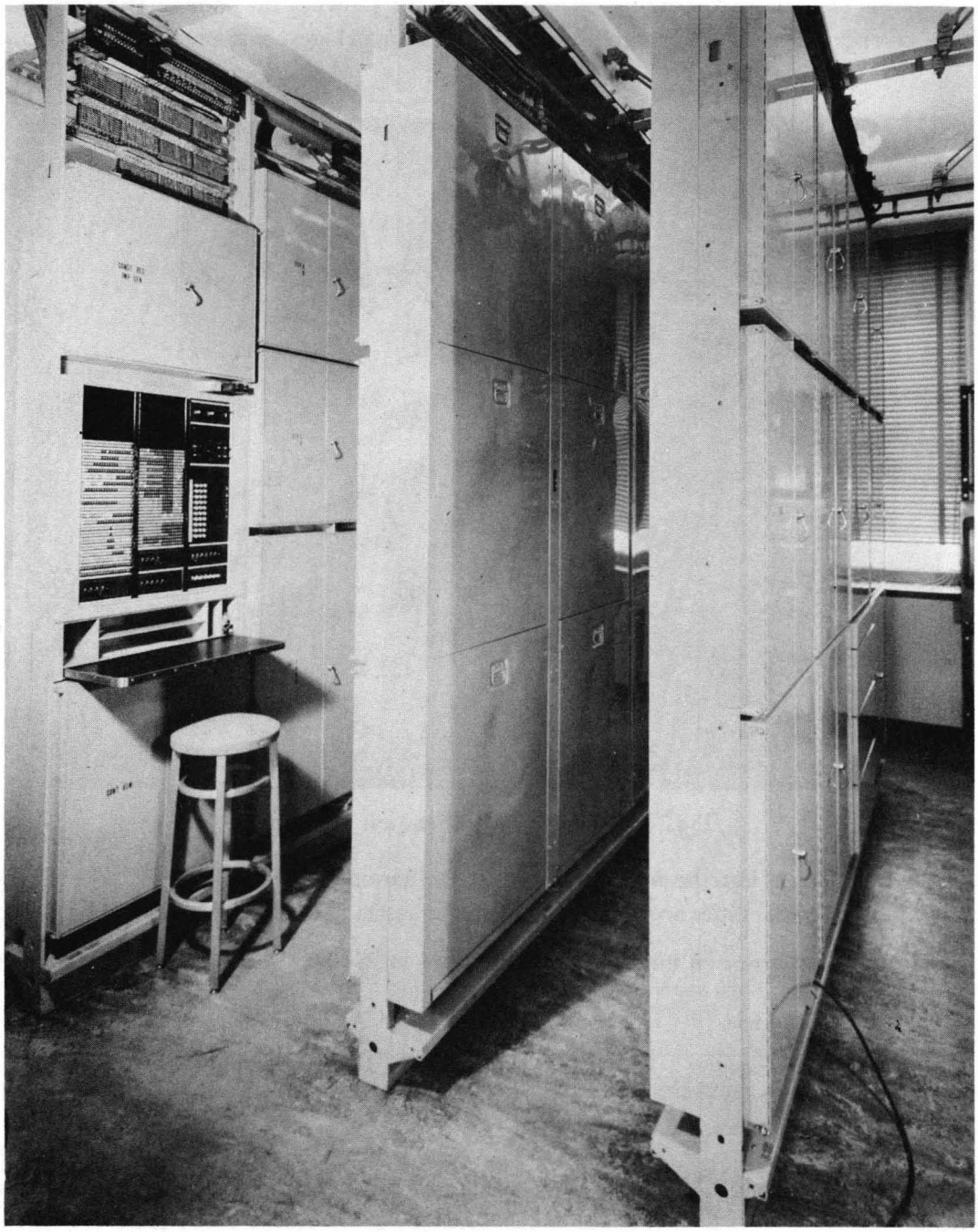


FIG. 2. Computing-room equipment.

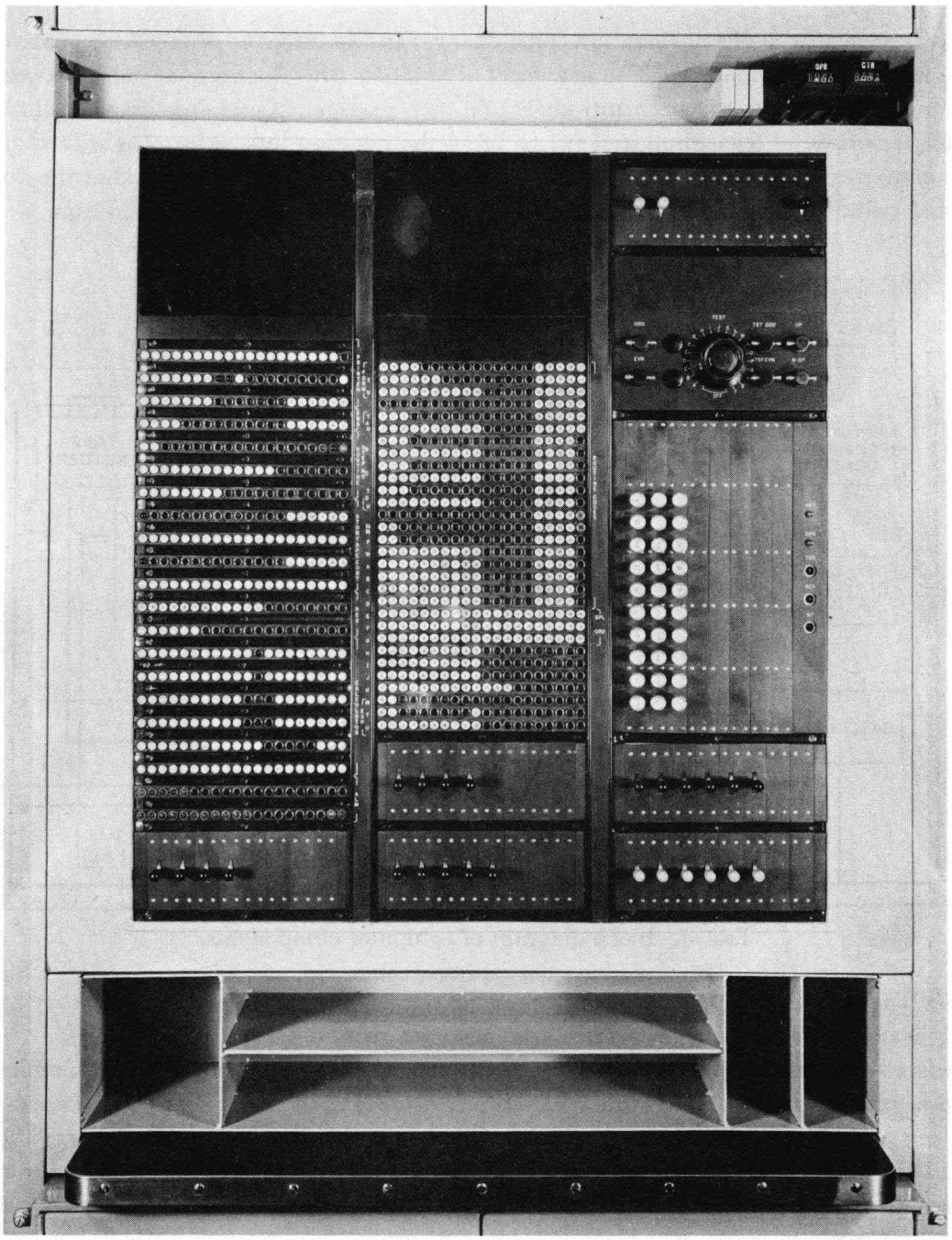


FIG. 3. Indicator and test panel.

As is shown in Table 1, there are some changes in the range of numbers compared with those in the earlier Model V. The new form of notation with the decimal point between the first and second digits has been introduced for two reasons: (i) to simplify the isolation of the characteristic of logarithms and (ii) to bring the notation into agreement with that now commonly used in expressing values in scientific literature. It will be noted that the floating decimal point has been retained along with the range of the exponent of ten from + 19 to

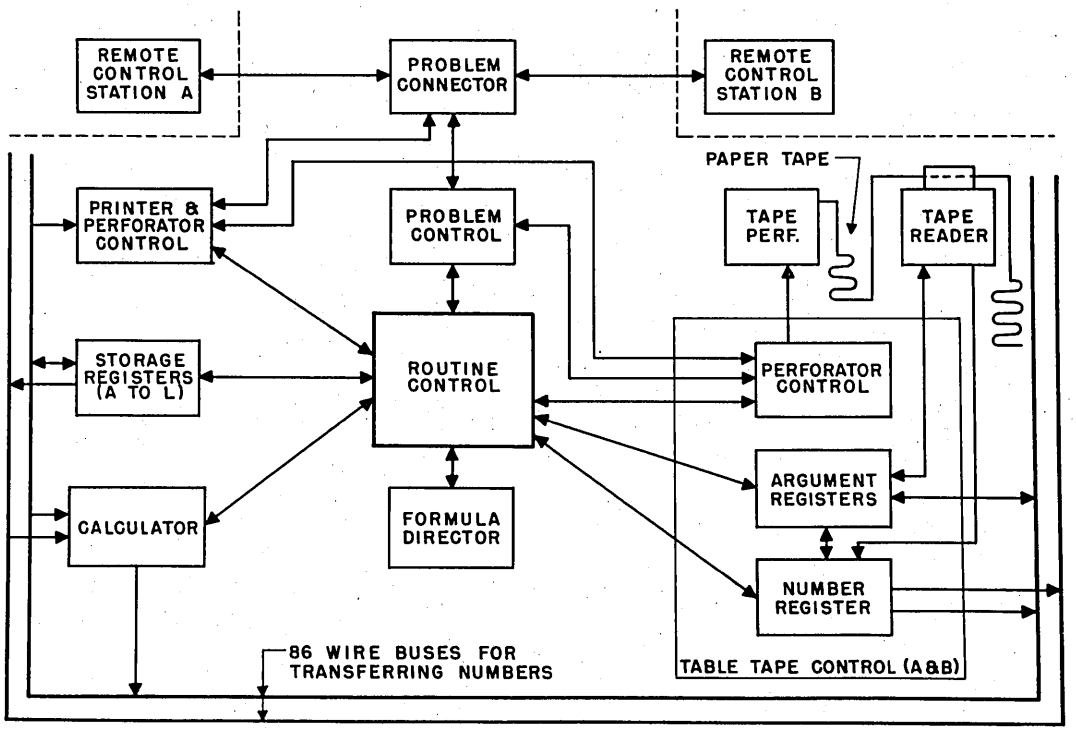


FIG. 4. Block diagram of computer components.

— 19. The calculator performs all arithmetic operations, namely, addition, subtraction, multiplication, division, and extraction of the square root.

Before returning to the new control-circuit features which provide for the higher intelligence, we wish to emphasize that the Model VI is as error-proof as its immediate predecessors. Our engineers with justifiable pride can still say, "Starting with the Model III delivered to the Armed Forces in 1944, not one of our customers has reported their computers giving out a wrong answer as a result of a machine error." To help understand the basic control system, reference is made to Fig. 4, which shows a block diagram of the relation between the routine-control circuit and the other computer components. To solve a problem, operators punch the problem data and computing instructions on a Teletype tape. This is loaded into the tape reader at a remote-control station and computing starts when the operator depresses a start key at this station.

Initially, the problem-control circuit carries out certain conditioning operations as instructed by the tape and then delegates control of computing operations to the central organ labelled "Routine Control." This component proceeds to control operations for computing or printing of answers, etc., as directed by instructions from the tape.

The instructions on the tape may be in either of two forms: (i) the specification of all of the detailed computing instructions, and (ii) the mere indication of which of the several sets of internal routines is to be used. The first form is usually used for types of problem that are encountered infrequently and the second for types of problem that are expected to occur often.

The routine-control circuit, in accordance with the instructions it receives, causes numbers to be passed from one register to another over the two 86-wire multiples or busses shown around the edge of Fig. 4. When an arithmetic operation is required the routine control designates the registers that hold the two numbers involved and directs the calculator to accept these numbers and perform the desired arithmetic operation on them. The control circuit then indicates what disposition is to be made of the result in the calculator by designating the place to which it is to be transferred; that is, to some particular register, the printer, or the perforator.

One of the new features is a "second trial" feature which is automatically brought into operation in almost all cases when the routine-control circuit fails to receive the usual OK signal indicating satisfactory execution of the instruction, or operation, called for.

The Model VI has an elaborate system of interrelated internal routines. It is this system with its own automatic seizure of various subroutines that gives this computer its higher intelligence.

The device used for this purpose is a combination relay and electronic device which is used in the Automatic Message Accounting System in a modern telephone office. As used in telephone switching it consists of a large group of code points, each corresponding to the location of the call-originating equipment of a subscriber. A signal on one of these code points denoting the origination of a call causes the circuit to ascertain and to record the subscriber's four-digit directory number.

As used in the computer, the code points correspond to the computing operations in a subroutine. A signal on one of these code points causes the computer routine-control circuit to set into operation the desired computer operation. A subroutine will use a train of from 6 to 20 of these code points in succession. Facilities are provided for 200 such subroutines, each being identified by the letter *A*, *B*, *C*, or *D* followed by a two-digit number.

The operation of a subroutine will be explained by an example which assumes that computations have reached the stage where the product of two complex numbers is required. Table 2 shows the formula used and the coding of the individual computer operations. Figure 5 shows how the operations are made a part of the computer and shows that there is an extremely close resemblance between the coding as it is written on paper in Table 2 and as it is memorized by the computer. The operation numbers 1 to 6 correspond exactly with code points 1 to 6. The letter designations correspond exactly with coil designations. The letters that form a particular computing operation are associated on the left of Fig. 5 by writing

Table 2. Coding of subroutine for complex-number multiplication.

$$(A + jB)(C + jD) = (AC - BD) + j(AD + BC)$$

Instruction		Explanation of Code
No.	Operation Code	
1	$B \times D\ E$	Multiply B by D and store result in E storage register
2	$A \times C\ S$	Multiply A by C and hold result in calculator
3	$S - E\ P$	Subtract BD from AC to obtain real part of product and print result
4	$A \times D\ E$	Multiply A by D and store result in E storage register
5	$B \times C\ S$	Multiply B by C and hold result in calculator
6	$S + E\ P$	Add AD and BC to obtain imaginary part of product and print result

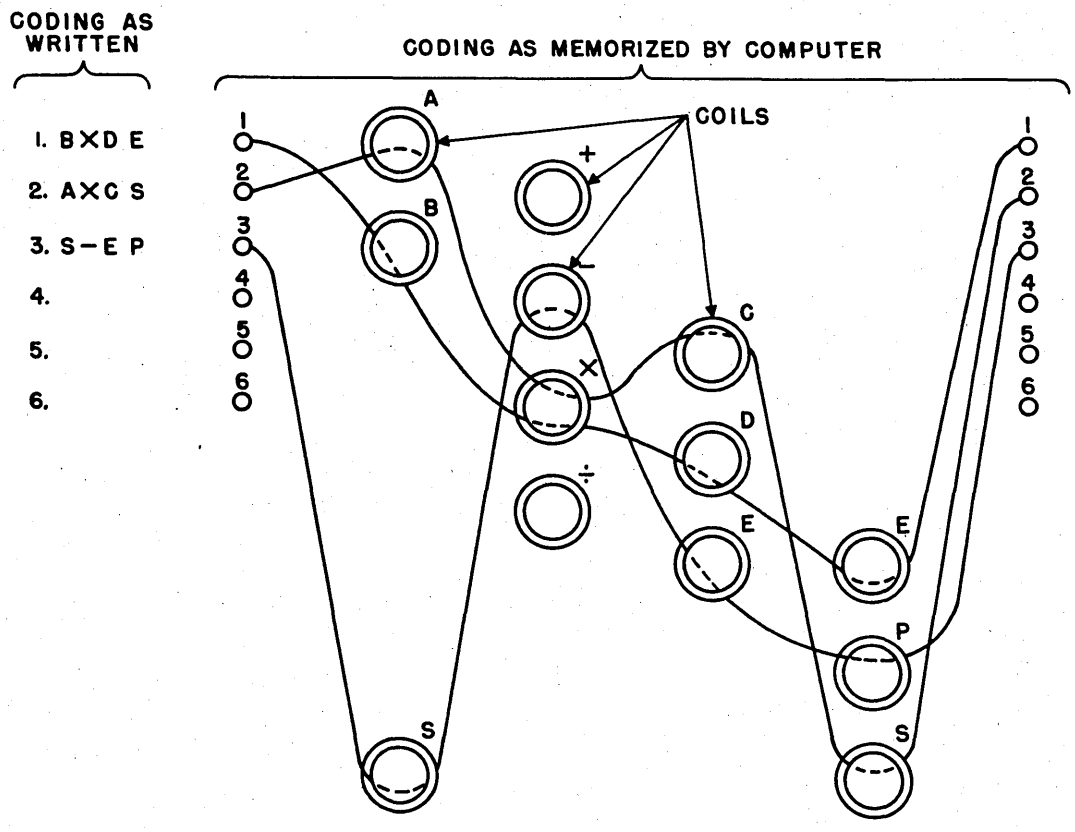


FIG. 5. Wiring of subroutine for complex-number multiplications.

them together on one line and on the right by running a single insulated cross-connection wire through coils having the same letter designations. This train of six operations would be identified by a subroutine number, such as D36 (not shown in Fig. 5). In practice these subroutine numbers are being called formula numbers.

When the 200 subroutines are fully employed there will be over 2000 insulated wires crisscrossed through the coils. These wires are not, however, part of the design of the computer but are placed in position by those responsible for operating the machine in accordance with the type of computing work they are engaged in. The actual physical operations of placing a wire can readily be performed in one or two minutes by the same personnel that operates the machine. No soldering of the wires is required. When a particular subroutine is no longer needed, the associated wires can be readily removed. The time required to set up a new subroutine, therefore, compares favourably with the time of setting up the same instruction on tapes.

Figure 6 shows a close-up of the coils with their associated cold-cathode tubes. Figure 7 shows the principles of the coil circuit.

A computer operation is initiated by causing a transient discharge from the resistance-capacitance-inductance network to be sent through the cross-connection wire. The part of the system consisting of the wire and the coil behaves like a transformer with the wire acting as a loosely coupled one-turn primary winding and the coil acting as the secondary. The transient through the primary induces enough voltage in the secondary to cause ionization between the control anode and the cathode of the type 313 cold-cathode tube. As

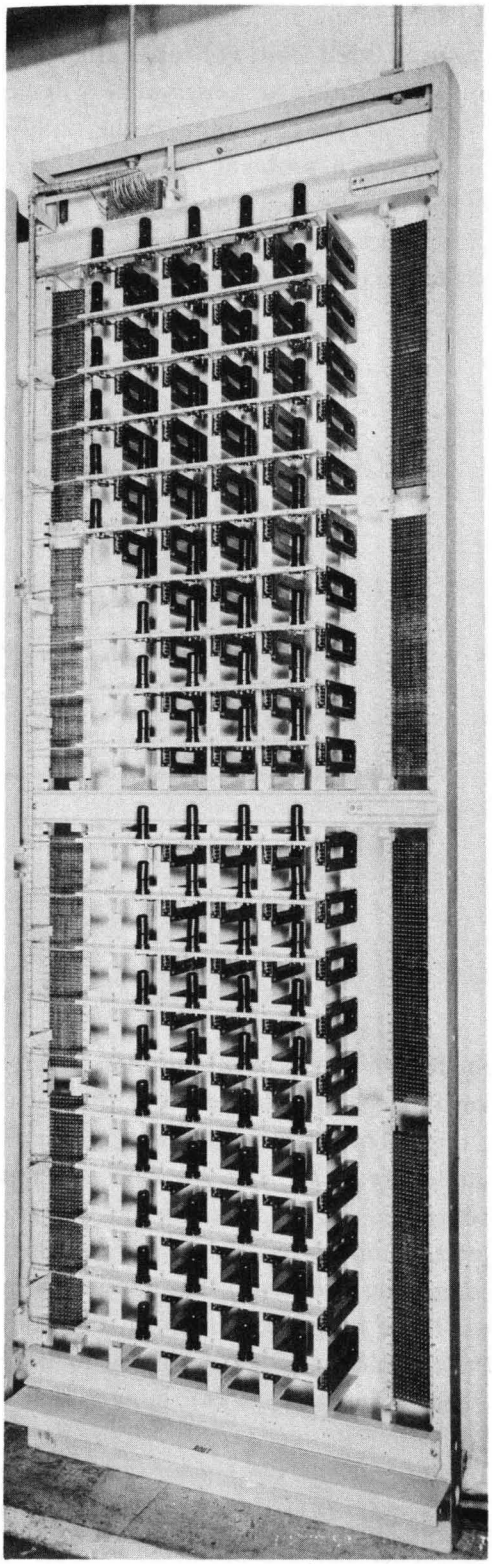


FIG. 6. Mounting of coils with their associated tubes.

soon as complete ionization takes place, the relay in the main anode circuit operates and disconnects itself from the tube. In this operation the tube is conducting for less than 0.01 sec and is operating at conservative voltage values. Very long tube life is, therefore, expected.

In the earlier computers the building blocks used for building up the complete program for solving a problem were single computing instructions. In the Model VI, the building blocks may be made much larger by making use of these subroutines. But these larger building blocks presented a challenge to make the most efficient use of them. Indiscriminate use would make for chaos. An efficiency expert would systematize the situation by grouping together

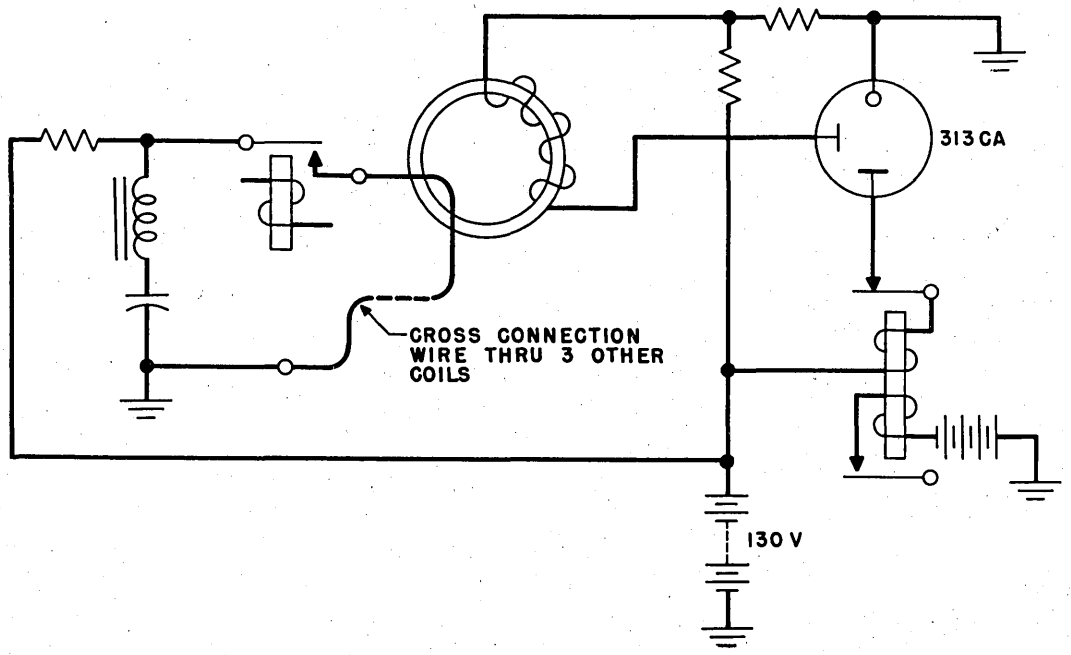


FIG. 7. Schematic diagram of coil circuits.

the simplest types of subroutines, which could be made complete in themselves, and by arranging other subroutines so that they would have various degrees of supervision over the first group. The degree of authority assigned has been designated for each subroutine with one of the letters *A*, *B*, *C*, or *D*. This is regimentation, but in a machine it is both acceptable and desirable. In fact, this regimentation is closely analogous to that which might exist in the strictest military school where an upperclassman would look with disdain upon any task (or computing operation) that a lowerclassman could perform. Consequently, the upperclassmen would be assigned to perform the more complicated tasks according to their own skill and they would be given authority to delegate lowerclassmen to do their more menial chores. These lowerclassmen can in turn delegate those parts of these chores beneath their skills to still lower classmen.

To continue the analogy, the Model VI computer has four levels of computing skill for

making use of the above-mentioned internal subroutines. They can be programmed to call in subroutines of a lower level and to regain control when the subordinate routine completes its task. The term "intelligence level,"[2] accurately describes the nature of these levels.

Figure 8 shows how this system operates in solving a rather complicated problem, called a ladder-network problem. One of the objectives of this problem is to determine the

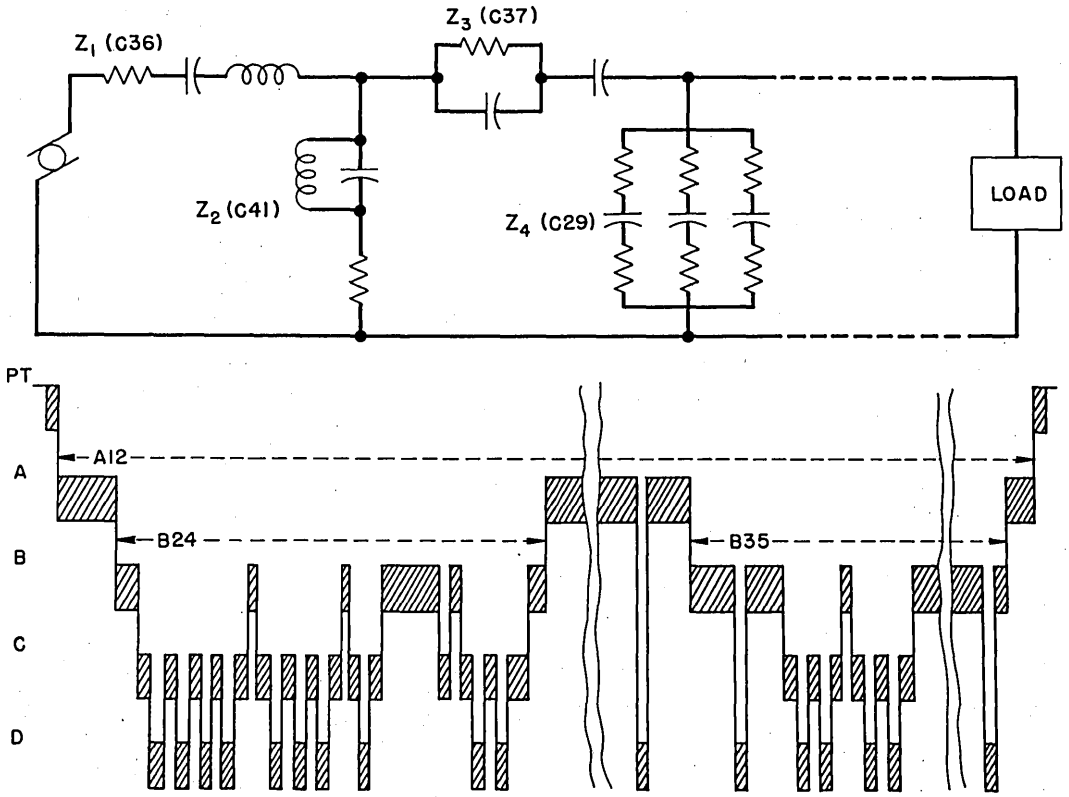


FIG. 8. Time-flow chart showing use of four intelligence levels to solve a ladder-network problem.

frequency response of the network. Mathematical analysis, therefore, is used in lieu of laboratory tests.

The top level in Fig. 8 is designated *PT*, denoting "Problem Tape." While this tape is primarily used for the introduction of the parameters of the problem, it may contain as much detail in computing instructions as required. In the case of the problem being described less than 1 in. of tape is required for the computing instructions because the Model VI is assumed to have already been taught how to solve the ladder-network problem. The problem tape then simply specifies that subroutine *A*12 be used and the tape thereby constitutes another intelligence level, higher than the four previously discussed.

In the interest of eliminating unnecessary detail, not all of the changes in level are shown

in Fig. 8, but there is still sufficient detail to show the principle of operation. By this systematic arrangement of internal routines, the labor involved in preparing the necessary problem data is near the irreducible minimum.

The solution of the network problem is carried out in the following principal steps:

The *A*12 routine assumes control at the start of the problem. It provides for organizing and setting up the initial conditions for computing the impedances for the various values of f (frequency). It then delegates control to subroutine *B*24.

Subroutine *B*24 controls the computation of the impedance of the various branches of the network. It also arranges to have the impedances recorded in block 400 on the storage tape for subsequent use by subroutine *B*35. After these preparations, *B*24 will instruct the computer to use the first number on the tape as the subroutine number to use in computing the impedance of the first branch.

Such a number might be *C*29. It would provide for reading off the parameters that follow and for combining the impedances for the individual elements, making use of *D* level subroutine for the complex-number arithmetic when required. The last instruction in *C*29 instructs the computer to use the next number on the tape as the subroutine to use. Additional subroutines are employed as required. After the last impedance has been computed, control is restored to subroutine *B*24.

On regaining control, *B*24 will instruct the computer to obtain the next value of frequency; it then arranges to have the impedances for the next value of frequency recorded in block 401 and then, using this new frequency value, repeats all of the operations just described. After completing the computing with the last value of frequency, control is restored to *A*12.

On regaining control, *A*12 will organize the computing of the complex values of voltage and current that the generator must supply. It then calls in *B*35 to control these calculations. Then *A*12, on regaining control, will provide for obtaining any additional information that may be required.

Table 3. List of intelligence levels.

Name	Designation or Symbol
1. Problem tape instructions	*SWR* (Switch to Routine)
2. *A* Subroutines	*A*12, etc.
3. *B* Subroutines	*B*24, etc.
4. *C* Subroutines	*C*29, etc.
5. *D* Subroutines	*D*59, etc.
6. Calculator instructions	$+, -, \times, \div, \sqrt{\ }$

From the above description it will be noted that the individual computing operations specified by a subroutine consist mostly of a collection of calculator and recording instructions.

It then follows that the calculator operations themselves comprise an intelligence level below all of the others. Instead of being identified and called for by a subroutine number, the usual arithmetic symbols $+, -, \times, \div$, and $\sqrt{\ }$ are employed. As in the other intelligence levels, the calculator returns control to the higher level that called it in as soon as the prescribed arithmetic operation is completed.

The Model VI then has at least six intelligence levels, as shown in Table 3, in descending order of authority or control.

A study of Table 3 shows that the Model VI will perform a series of computer operations in accordance with any of the last five levels by merely designating the three-element code or the symbol shown in the last column. In fact, the three-element code is so closely analogous to a symbol in this computer that it is proper to say that the Model VI responds to its own idea of a symbol for determining the logarithm of a number or of a symbol for determining the tangent of an angle, and so forth.

The Bell Computer, Model VI, has become an upperclassman. It can be taught how to solve a problem. It can retain this know-how for use whenever called upon in the future.

REFERENCES

1. Previous Bell computers have not carried model numbers, but the numbers that are now assigned to them in retrospect are: Model I, Complex-Number Computer; Model II, Relay Interpolator, *Bell Laboratories Record* (Dec. 1946), p. 457; Model III, Ballistic Computer at Fort Bliss, Texas, *Bell Laboratories Record* (May 1948), p. 208; Model IV, Ballistic Computer at Naval Research Laboratory; Model V, General-Purpose Computers at Langley Field and Aberdeen, *Bell Laboratories Record* (Feb. 1947), p. 49; and *Mathematical Tables and Other Aids to Computations* (Jan. 1948), p. 1, (April 1948), p. 69.

2. This term (ρ2ρ) was originated, it is believed, by G. R. Stibitz; see *Annals of the Computation Laboratory of Harvard University*, vol. 16 (Harvard University Press, Cambridge, 1948), p. 91.

An Electrostatic Memory System

J. PRESPER ECKERT, JR.

Eckert-Mauchly Computer Corporation

This paper is a progress report of the work done on a very high-speed memory constructed of ordinary cathode-ray tubes. The research has been performed in the laboratories of the Eckert-Mauchly Computer Corporation. Many persons have been engaged in this research, but particular credit should be given to Herman Lukoff, C. Bradford Sheppard, Gerald Smoliar, and Charles Michaels, all members of the Engineering Department.

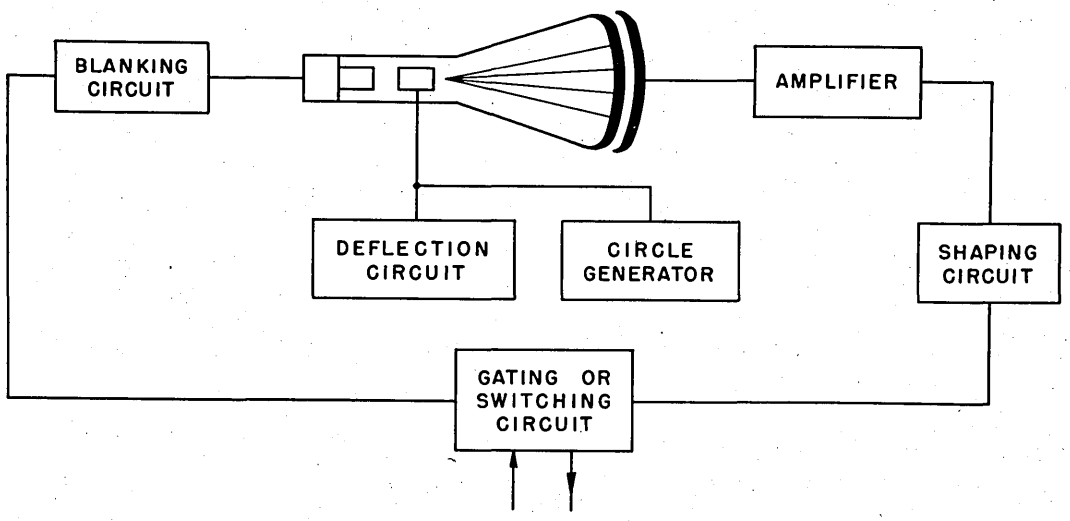


Fig. 1. Circuits used to operate a cathode-ray tube as an electrostatic memory.

A more complete report was submitted as a paper to the Institute of Radio Engineers in April 1949. It contains much more detailed information. In particular, it gives additional quantitative material.

The first part of this paper describes a memory system that is now under test. It is the second model of an electrostatic memory system to be constructed at the Eckert-Mauchly Computer Corporation. The second part of the paper describes a limited number of the tests performed on this system and gives some of the results. The final part of the paper gives a short glimpse into the research still to be performed on this memory system.

Work on a high-speed electrostatic memory system was originally begun by the author at the University of Pennsylvania. The tests at that time were preliminary, serving to indicate the large amount of research necessary to the developments described in this paper.

Figure 1 is a block diagram of the circuits used in operating an ordinary cathode-ray tube

as an electrostatic memory. A metallic electrode, actually a wire mesh, is attached to the base of the tube and is coupled to an amplifier having a gain of about 2000. Each time the beam strikes a charged area on the tube, a signal is developed on the electrode. This signal

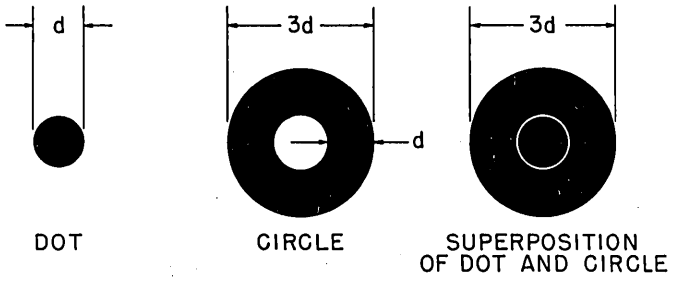


FIG. 2. Relative sizes of dot and circle.

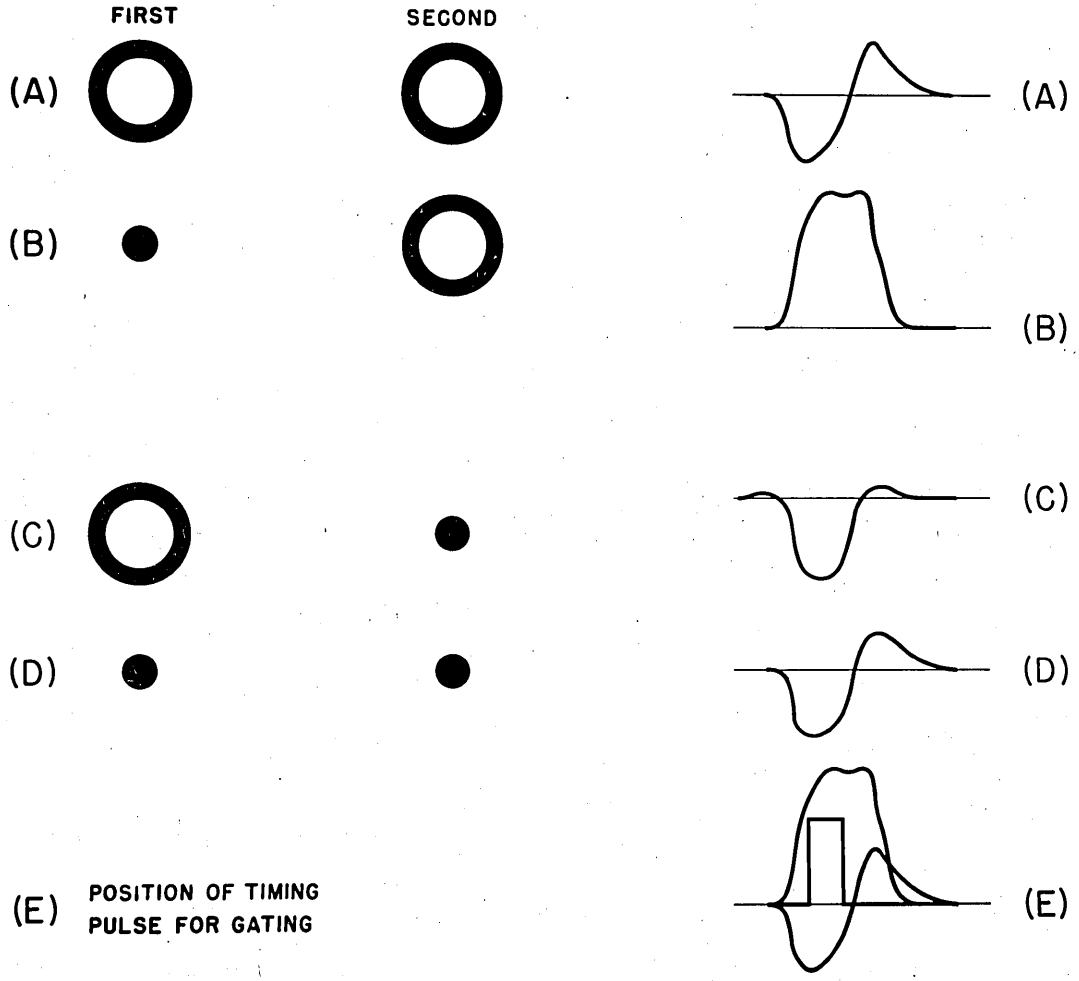


FIG. 3. Output signals.

is passed through a shaping circuit to the gating or switching circuits. These circuits turn the information in or out of the regenerating path, operating in a manner analogous to a mercury-delay memory. The blanking circuit turns the beam on or off for the writing or reading interval. The deflection circuit controls the position of the beam. The purpose of the circle generator is described later in the paper.

In operation, the screen of the tube is considered to be divided into many small elementary areas. These areas are approximately 0.1-in. squares. The patterns placed in these areas that gave the most satisfactory results consisted of dots and circles. The dots are formed by focusing the beam, as sharply as possible, in the center of the elementary area. The circle is

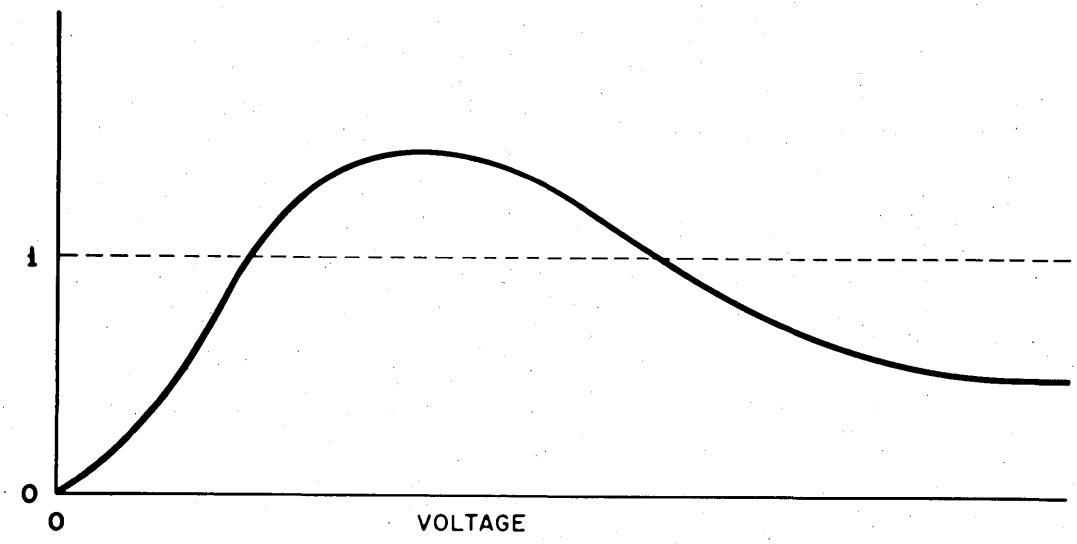


FIG. 4. Ratio of number of secondary to number of primary electrons as a function of voltage.

formed by superimposing two high-frequency sinusoidal emf's 90° out of phase with each other on the two deflection systems. As Fig. 2 shows, the diameter of the dot is about one-third to one-half the diameter of the circle. The two patterns can be considered as two different states, the dot representing a 1 and the circle representing a 0 in the binary system.

Reading the information stored on the screen of the tube is done by adjusting the potentials on the deflection plates so that the beam will fall directly on the desired elementary area. When the beam is turned on by the intensity grid, a potential is developed between the electrode and the collector which puts a signal into the amplifier.

Lines (A) and (B) of Fig. 3 show the output signals received by the amplifier during the reading operation. These signals are a result of both the previously stored pattern and the new reading pattern. While there are four types of signal, one of which has an initial positive rise, and the other three of which have a negative rise, only two of these signals are ordinarily used in the electrostatic memory system.

The high value of load resistance used with the pickup electrode tends to obscure certain factors important to an understanding of the problem. After a review of some of the pertinent properties of electrons striking insulated barriers, the output signals will be analyzed further.

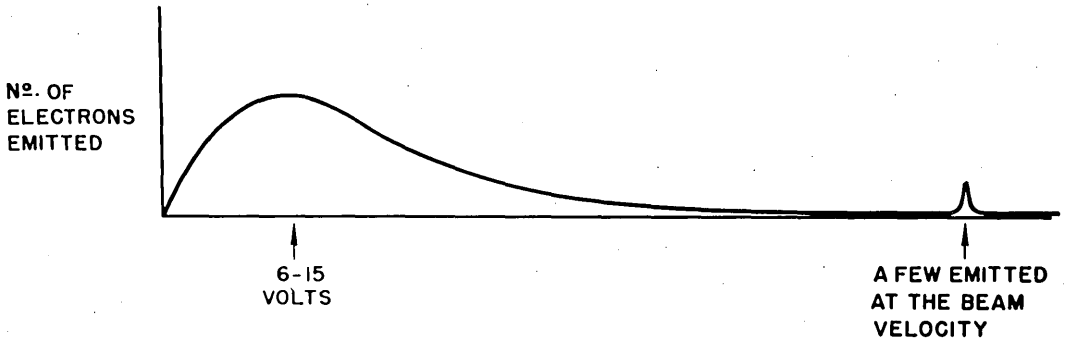


FIG. 5. Secondary-electron emission as a function of velocity of emission.

The phenomenon encountered in the electrostatic memory system involves both the primary electrons of the beam and the secondary electrons which are cast away from the surface of the phosphor by the beam. Figure 4 shows one of the well-known fundamental properties of secondary electrons. This curve shows the ratio of the number of secondary electrons to the number of primary electrons (which cause the primaries to be emitted) plotted against voltage. The curve is mainly of interest to this paper in the section where it is sub-

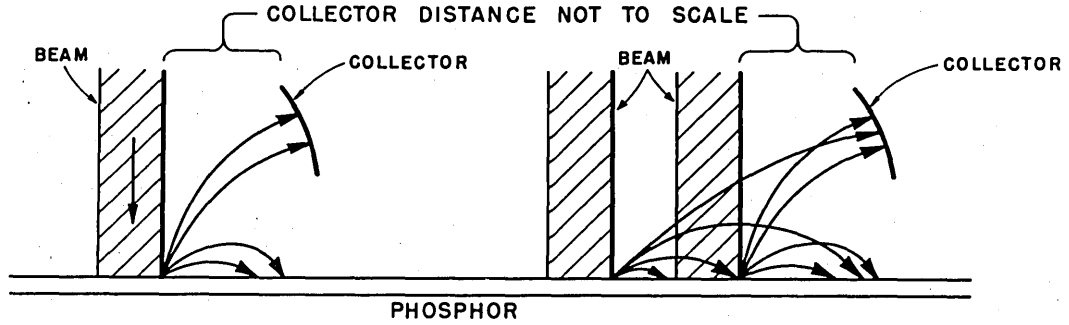


FIG. 6. Cross sections of electron beams.

stantially above 1. Most of the tests were made in this region with voltages between 1,500 and 4,000 v.

Figure 5 shows a second fundamental property of insulating surfaces and their effect upon the behaviour of secondary electrons. This curve shows the distribution of secondary electrons as a function of the velocity at which they are emitted from the surface. Except for a few electrons emitted at the beam velocity most of the electrons are emitted at velocities corresponding to between 3 and 15 v. These velocities are quite low compared to the velocity of the striking beam that causes the emission. Finally, few secondary electrons leave the surface

at right angles to the direction of the incident beam, the majority leaving at smaller angles to the incident beam.

Figure 6 shows the cross section of a beam of electrons striking the phosphor on the inner surface of the screen. Potential distributions are set up in each of the elementary areas.

POTENTIAL DISTRIBUTION

ALTERNATE POTENTIAL DISTRIBUTION

DOT CIRCLE

FIGS. 7 (upper) and 8 (lower). Possible potential distributions around dot and circle.

Secondary electrons are released in accordance with the principles just discussed. The secondary electrons travel to and are collected by the collector plate. This collector plate is usually the Aquadag coating on the inner walls of the tube. Other secondary electrons may fall back onto the surface of the screen. Since the beam arrives at the phosphor with a velocity

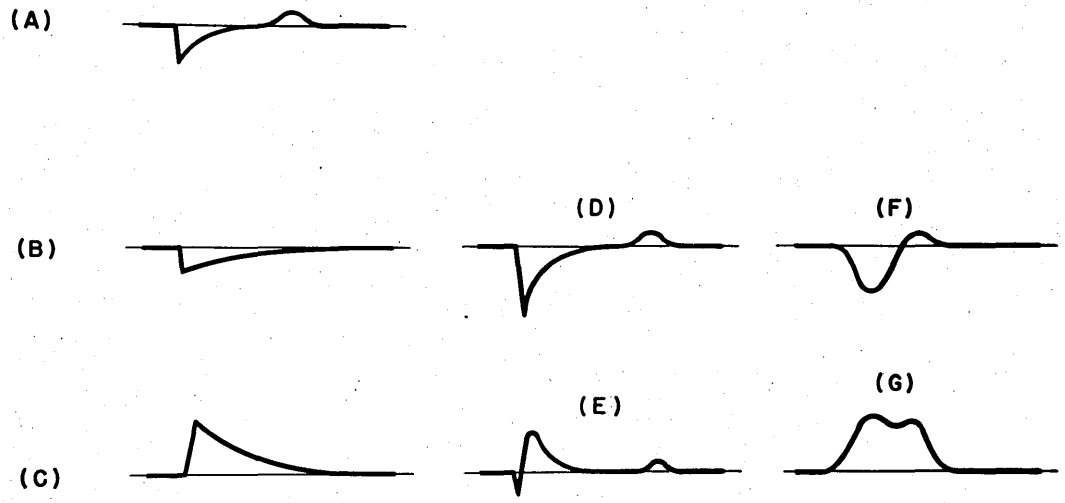


FIG. 9. Shapes of output signals.

corresponding to several thousand volts, the number of secondary electrons is greater than the number of primary electrons. Thus, the surface will not reach equilibrium at this point until the number of electrons that leave the surface and the number that arrive become equal.

The potential distribution around the dot may be as shown either in the left-hand section of Fig. 7 or in that of Fig. 8. Although much of the literature discusses the distribution shown

in Fig. 8, to the author's knowledge, no critical experiments have ever been made that would positively prove which curve describes the exact potential distribution.

A truer picture of the electrostatic memory phenomenon was obtained by lowering the load resistance connected across the input circuit consisting of the electrode and the grid of the first amplifier tube. By so doing, it was found that the signal obtained when similar patterns are placed on top of each other has the shape shown in curve *A* of Fig. 9. The initial negative kick is caused by the arrival of the beam from the gun after the intensity grid is turned on. The transit time for the beam is about 0.01 μsec. This negative kick then subsides toward zero since the electrons piled up on the end of the tube are drawn off through the emission of secondary electrons until equilibrium is reached. This equilibrium results in a potential plateau under the bombarding beam where there are as many electrons arriving as leaving. This is possible when the potential plateau becomes sufficiently positive that only some of the secondaries are knocked off the surface with sufficient velocity to reach the collector plate, allowing the remainder to fall back on the neighboring areas of the surface.

From the time equilibrium is reached until the beam is turned off, there is a steady inward and outward flow of electrons to the screen maintaining a space-charge cloud between the spot and the collector. When the beam is turned off, the space charge is rapidly taken up by the collector. Since this negative space charge leaves the screen, a positive kick is induced in the electrode. All other signals obtained contain curve *A* as a component. The components added to curve *A* to produce the other signals are sudden rises with simple exponential declines.

Curve *C* is obtained when a larger pattern is placed on a smaller one. The large positive kick occurs because most of the secondaries are drawn to the collector plate. Since the potential plateau of the dot is at a lower potential than the collector and small in area, only a few electrons are robbed from the secondary flow to the collector.

Curve *B* represents the exponential component obtained when a smaller pattern is placed on a larger pattern. Again, most of the secondaries go to the collector plate. But the desirable action would be for the secondaries to obliterate or cancel the circle as quickly as possible. Instead, the circle, in spite of its large area, collects electrons slowly since its potential plateau is lower than the potential of the collector. The secondaries that do fall on the circle do so mainly by virtue of the direction of their emission.

Thus, when a circle is put on a dot, there is a rapid net outward flow of electrons; while when a dot is put on a circle, there is a net inward but slower flow of electrons. The net inward flow occurs in spite of the influence of the higher potential on the collector owing to the large area of the plateau of the circle, which attracts the properly directioned electrons away from the collector. In either case the change in plateau area and, therefore, the number of electrons to be exchanged to reach equilibrium is the same. Therefore, the positive signal will be large since the output-voltage signal depends on the time rate of change of charge. The negative signal will be small since the time rate of change of charge is small.

Curve *D* is the sum of curves *A* and *B*, while curve *E* is the sum of curves *A* and *C*. Curves *F* and *G* are obtained by using a high load resistance on the amplifier input.

Two of the most interesting factors that cause destruction of the charge patterns are leakage and redistribution. Tests for leakage were conducted, and it was learned that the effective leakage may be considered negligible if the period between readings is less than 0.1 sec. Even after several seconds the signal has diminished by only a few percent. Redistribution is the spraying of secondary electrons from adjacent areas during the writing or reading process onto areas that have been previously charged. We have defined a "redistribution ratio" as the number of times the reading beam, with a certain duration, may operate adjacent to a particular spot at a certain distance before the signal that can be derived from the adjacent spot will have been degraded by more than by a certain percentage, say 10 percent. Experimentation has shown that the degradation is not proportional to the number of times of reading but is proportional to the total integrated reading time. For efficient operation, a minimum time sufficient to establish equilibrium in reading and regeneration can be chosen. The total allowable reading time on one spot can be divided by this minimum time to give the number of times an area may be read without appreciably affecting the adjacent areas. This is the redistribution ratio.

Since the process for reading the charged areas is a destructive one, immediate regeneration of the charges is necessary if it is desired to retain this information. As each spot is read out from the memory tube, it is temporarily held in a flip-flop or other simple form of memory for one binary digit and then if desired immediately read back into a cathode-ray tube. In this way, only one elementary memory or flip-flop need be used for each cathode-ray tube or group of tubes. In addition to this immediate regeneration, a systematic regeneration must be used.

Such a systematic regeneration pattern might divide each regeneration cycle into two intervals. During the first interval any arbitrary spot is read and regenerated; during the second interval, one of the other spots on the tube will be regenerated as part of a regular systematic regeneration procedure. In such a system, the condition of most interest would be that in which the same spot is read during all the arbitrary reading periods without losing the spot next to the arbitrary spot through redistribution. If there are 1000 spots on a tube, this requires a redistribution ratio of 1000 or better. Such a regeneration pattern utilizes 50 percent of the operating time for the purpose of regeneration.

If it is desirable to have less time in the memory for regeneration, a system of timing could be devised where two arbitrary spots are read in succession and then the systematic regeneration of a spot takes place. Such a system would cut down the time required for the systematic regeneration but would increase the intervals between regeneration of a particular spot and would require an improved redistribution ratio.

Figure 10 is a photograph of the second test model used for many of the tests conducted in the laboratories of the Eckert-Mauchly Computer Corporation.

In the experiments carried out in England, small imperfections in the phosphor of the

cathode-ray tubes, due either to a hole in the phosphor or to the inclusion of particles of carbon, would occasionally make it impossible to remember on certain parts of the tube. Difficulties of this type have not appeared so far in the work here described, although experiments on an extensive number of tubes have not yet been made.

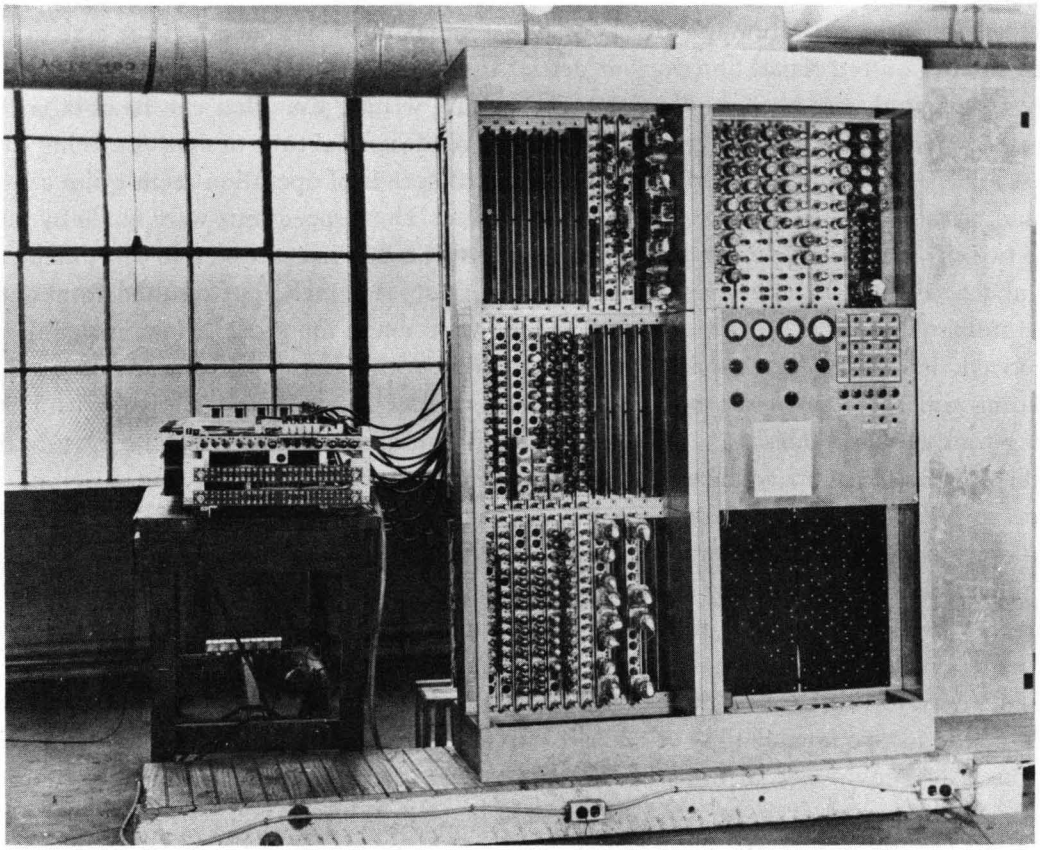


FIG. 10. Second test model of computer.

It is believed that the dot-circle system is the most insensitive to screen imperfections because, first, it gives the largest signal, and second, there is no sweeping action in which the edge of the beam may encounter a small discontinuity produced by phosphor imperfection. While a sweeping action is used to generate the circle, the frequency of sweeping around this circle is so great (approximately 20 million times a second) that the lag introduced by the finite charging time of the elementary area, combined with the finite transit time of the electrons, prevents the imperfections from having any effect on the shape of the output signal but simply changes its amplitude. Therefore, a system that has good output-signal amplitude used in connection with a cathode-ray tube in which the size of the imperfections is small compared to the size of an elemental area should be free of difficulty. According to Williams,

and others in this country, the size of an imperfection required to produce difficulty in those systems in which slow sweeping is used might be smaller than an elemental area. It would seem that as long as the imperfection were not smaller than the edge sharpness of the moving spot, this difficulty might be encountered in a slow- or fast-sweeping system. By the same reasoning, other nonsweeping or fast-sweeping systems should be fairly free of difficulties due to screen imperfection, although they would not be as good as the dot-circle system owing to the smaller output signal and various defects in its shape.

A careful study has been made of the reading and writing time that can be obtained from a dot-circle memory system using standard tubes. While special tube designs are being studied that would probably increase this speed, the present speeds of operation seem quite adequate for many uses and agree fairly well with expectations. These speed tests were made by putting down two circles in succession and then putting down two dots in succession in every elementary area and observing the effect on the shapes of the output signals that resulted from changes in the unblanking time. If this unblanking time were made too short, a loss in signal would be noticed, indicating that equilibrium had not been established. It was determined that a reading time of about 0.6 μsec was desirable in order to avoid loss of signal, and that a writing time of about 0.8 μsec was about the minimum allowable for adequate erasure of the old charge. Since each reference to the memory would require a reading and a writing time, a time to position the beam and time for the various switching operations, a total cycle of operation of about 2.5 to 3 μsec is indicated. Although not mandatory, a regeneration cycle of another 3 μsec would usually accompany the first cycle. Thus, a total time of about 6 μsec for an operation of reading or writing or both is indicated. A read signal might be sampled and made available to an arithmetic element about 1.5 μsec after the beginning of the cycle. Thus, 4.5 μsec of this time might be used for computing. This memory might, therefore, be considered to have a latency time of 1.5 μsec. In any case, the speeds involved are comparable with the fastest envisioned arithmetic elements. If this memory is used in a serial computing system, a pulse period of perhaps 2.5 μsec would be reasonable. About twice the speed could be obtained if the immediate regeneration were not interspersed but were separated into an individual reading and writing cycle.

A study of the effect of tube diameter, acceleration voltage, and the best focusing procedures was made. In addition, as many as half a dozen different types of phosphor were studied. A summary of some of the results obtained follows.

1. The acceleration voltage had a major effect upon the amount of storage in a single tube. The tests were made with a roster of 256 spots. The spacing between spots could be varied in such a way as to contract the entire pattern either vertically or horizontally on the face of the cathode-ray tube. This spacing was adjusted while observing a particular spot and coming back to the adjacent neighboring spot every other reading time a number of times equal to the redistribution time for which the test was made. Improvements greater than two in the number of spots stored on a particular tube were obtained for a 75-percent increase in accelerating voltage when a redistribution ratio of just two or three was required.

When a redistribution ratio of 1000 or more was required, as is common in practical application of the memory, an improvement of three or more was obtained.

2. Since increases in voltage produce such a large improvement in the amount of storage that can be obtained, a study of the phosphors was made to determine which phosphor would allow the highest operating voltage. Two things were considered important. The phosphor

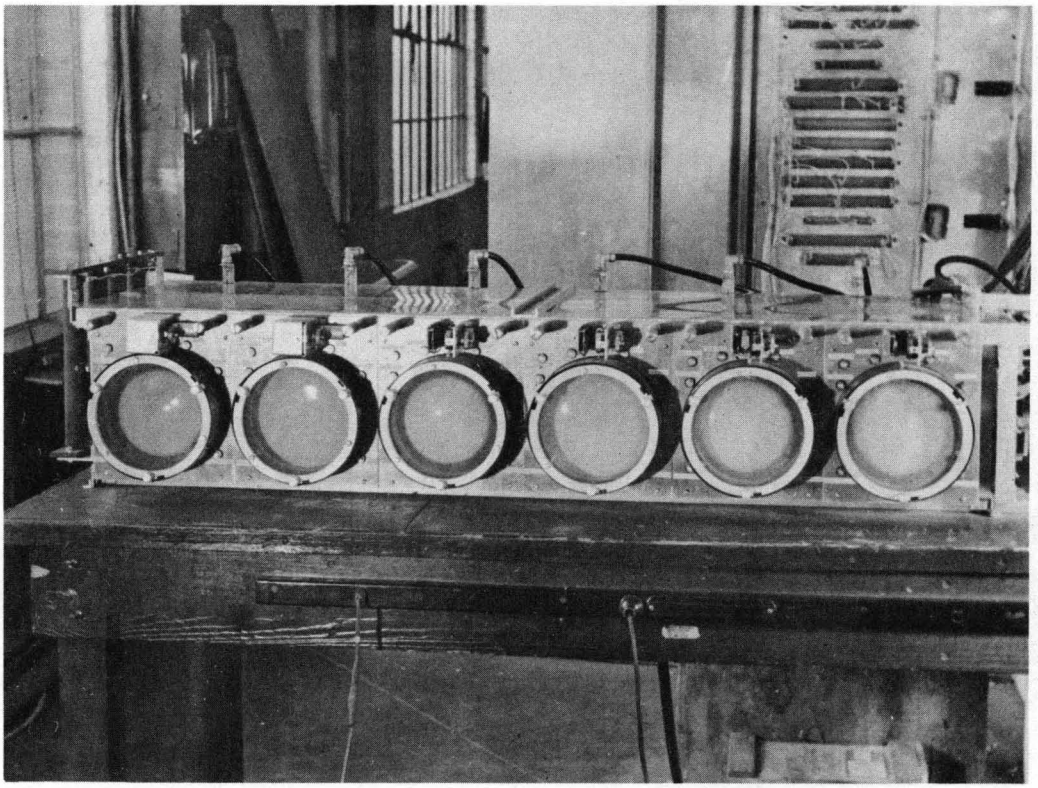


FIG. 11. Assembly of plug-in cathode-ray tube units.

should have a good secondary-emission ratio of the order of two or so at voltages safely in excess of the operating voltage in order that rapid functioning of the secondary cloud could be obtained. Secondly, the phosphor should be one that is easily made free of holes and that will not burn at high voltages. Tests of signal-reading and erasing ability showed that the P-1 phosphor, operating at 3000 to 4000 volts, nicely met all of the requirements.

3. Tubes containing almost similar guns of sizes 3, 5, and 7 in. using a P-1 phosphor and somewhat over 3000 volts for acceleration were tested to find their total storage capacity. The 3-in. tube would store over 2500 spots, on the assumption that a 5- to 10-percent decrease in output signal due to redistribution is tolerated, and that the area for 256 spots can be used as the basis for extrapolations to the number contained in the total roster area. On the same

assumption, the 5-in. tube would store over 3500 spots and the 7-in. tube would store about 5000 spots. Thus, while the larger tubes will store more information per tube, it is interesting to note that where space is a factor the smaller tube would be indicated. For example, the 7-in. tube stores only about one-third as many spots per unit volume as will the 3-in. tube, even though it will hold about twice as many spots.

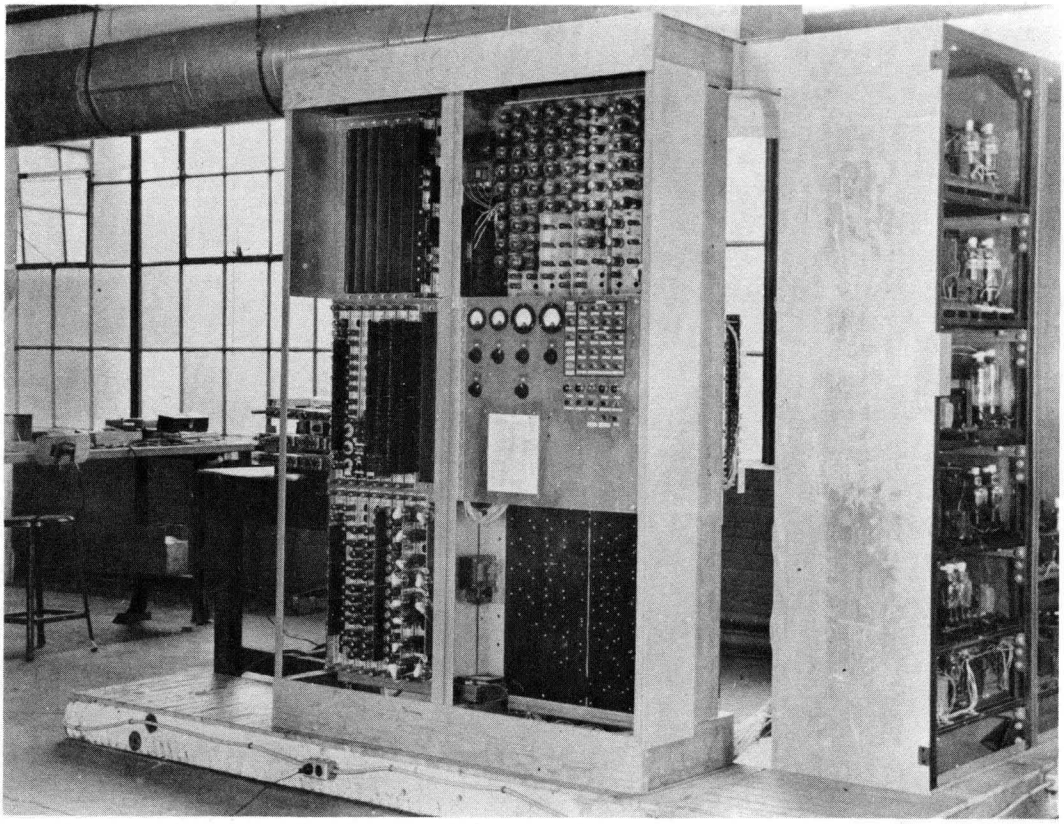


FIG. 12. Auxiliary equipment.

At the present time the Eckert-Mauchly Computer Corporation is engaged in setting up a complete memory involving more than 100 cathode-ray tubes suitable for use with a high-speed computer. This equipment includes several counters and other devices that allow for tests simulating the operation of what is essentially a complete memory system. The cathode-ray tubes are mounted in individual plug-in units. Each unit contains all the adjustments necessary for focusing and positioning the spots, and so forth. This plug-in assembly is rather important since one of the objectionable features of this type of memory is the multitude of adjustments required with each cathode-ray tube in order to allow for their rather wide manufacturing tolerances. However, since all of these adjustments are confined to the plug-in unit, a number of pretuned units can be held in reserve. These can be readily substituted for

an inoperative unit in about 1 minute, allowing any necessary readjustments of the defective unit to be made without serious interference to operation.

Figure 11 shows an assembly of six such units mounted in a framework with the necessary switching and deflection equipment. These frameworks can be stacked to relay-rack height, which allows for cabinets containing about 48 tubes to be assembled from these basic frames. A computer might use one or more such complete cabinets, depending upon the memory size required.

Figure 12 shows the complete auxiliary equipment, including power supply. The actual equipment for the memory requires less than half of the cabinet space shown. The additional space was left to allow this test equipment to be expanded into a laboratory model of a computer if desired. As it presently stands, this equipment will permit the study of tube life, maximum practical spot contents, maximum practical time between regenerations, and the degree to which readjustments may be required because of tube aging. Also, the effect of any imperfections in the phosphors may be studied on a really practical scale. Further, since regulators are included for all voltages affecting the cathode-ray tube operation, it will be possible to determine just which voltages must be regulated under practical operating conditions. In the present arrangement, regulated voltages appear in cases where calculations indicated a marginal requirement. Since the equipment is adjustable in pulse rate, experimental determinations of a reliable speed of operation can be made. Future reports will cover the findings of these further tests and also will describe several computing systems for which this memory is well adapted.

The Digital Computation Program at Massachusetts Institute of Technology

JAY W. FORRESTER

Massachusetts Institute of Technology

In this paper I wish to summarize the digital-computation activities at Massachusetts Institute of Technology. These will include the machine-development work on the Whirlwind I computer, the digital-computer educational program at M. I. T., and a few thoughts on future direction of work in digital computers.

The Whirlwind I computer is a prototype electronic computer which, following the precedent established by radio-frequency engineers, would probably be described as ultra high speed. We are aiming at the speed range of 10,000 to 20,000 complete arithmetic operations per second. Such speeds seem to be imperative for the application of digital computers to many of the more interesting control problems. Speed requirements dictate a parallel-type computer, and a sufficiently short storage-access time is provided thus far only by electrostatic tubes.

The computer is working in a new speed range and must be looked upon as a prototype design. As such, a short register length has been used to keep the first model as small as possible. The type of single-address instruction order used requires 16 binary digits of register length, and this was selected for the machine. Such a length is adequate for exploratory studies in control applications. In most mathematical work this short length would be a nuisance and double-length operations will often be employed until such time as the register length is expanded. Experience indicates that the choice of a short register was wise. Much has been learned since the design was frozen, and simplifications and improvements should be made before more equipment is built.

There has been no attempt to make a compact, small machine this first time. Flat panels on vertical racks permit complete access to both sides of electronic panels and is probably cutting to a third the time that would otherwise be required for installation and preliminary testing.

The design of Whirlwind I was begun two and one-half years ago at about the time of the first Harvard Symposium on computers. Prior to that time there had been a year of study of serial-type computers. A high-speed 5-digit parallel arithmetic element has been operating two years and giving valuable information on circuit performance and reliability.

The Whirlwind I computer might be divided into four parts: the arithmetic element, central control, storage, and terminal equipment for input-output. We have followed the design in that order. Most people in high-speed electronic computers have chosen to begin with the terminal equipment and work from there toward the central control of the machine.

We have followed the reverse order, designing the central control and arithmetic element first and leaving terminal equipment until the last. The input-output, it seems to us, is much more a function of the ultimate application of a computer like Whirlwind I than is any other part of the device. For some types of scientific work, page printing of results is sufficient. For many engineering jobs the easy, automatic plotting of curves is a necessity. In control applications the computer must have direct access to devices for converting to the analog quantities of the associated physical world. In many jobs an erasable external medium such as magnetic tape is required, and in others not. Therefore, it seems that terminal facilities may require continual adding of new equipment to the basic heart of the system which is the computer itself. Plans are not definitely formulated for the uses of Whirlwind I, and most of the terminal equipment will be fitted into those future plans. Initially we expect to have available the Eastman Kodak photographic-film units that were described at the first Harvard Symposium. These are being designed to read or write a thousand lines of information per second and, in addition, a duplicate checking channel. Each line contains one word length.

Turning now to the computer itself, which has received the principal attention of the laboratory, the arithmetic element and central control are both operating. The arithmetic element was installed in January 1949 and has been running since. At present it is being used as a tool for the preliminary testing of the central control. The central control for the execution of all the 32 machine orders was installed in June 1949 and is now being tested. No unusual difficulties have been encountered in obtaining desired performance. Storage will be the last part of the machine installed. Storage control circuits are now being connected and laboratory pilot quantities of tubes for 16-by-16 density are being built. These tubes still operate somewhat more slowly than desired.

A year ago last summer at the University of California Symposium, I estimated that Whirlwind I would be assembled by December 1949. It looks now as if this should be extended about 10 percent to February 1950. After assembly there will be a period of learning to use the equipment before one can really claim that it is in productive operation.

Figure 1 shows the switch and matrix section of Whirlwind I central control. Figure 2 is the test-equipment center used during installation.

The educational program in digital computers at M.I.T. is centered in four laboratories of the electrical engineering department. The differential analyzer is in the Center of Analysis directed by Professor S. H. Caldwell, who teaches a course in machine aids to computation. Professor Z. Kopal is in charge of the computation laboratory for the study of numerical processes and the operation of a hand computing center. A punched-card installation is operated by the Division of Industrial Coöperation under Mr. Frank Verzuh. The Whirlwind I digital computer is being constructed in the Servomechanisms Laboratory.

M.I.T. does not yet offer a packaged advanced study program in digital computation as does Harvard. However, available from the courses in the graduate school is a fairly complete master's degree level study selection. It is perhaps best to study numerical analysis and digital

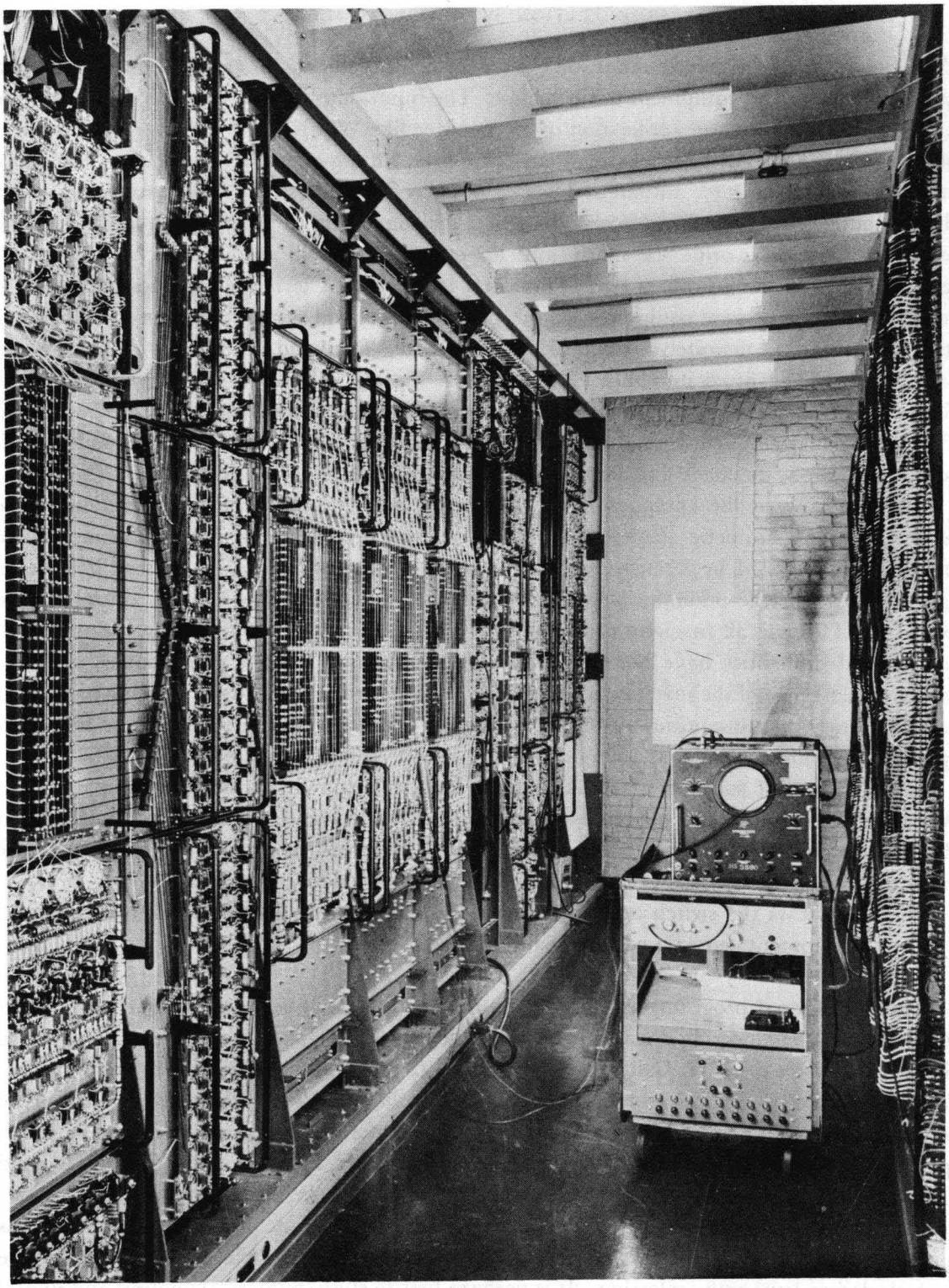


FIG. 1. Crystal switch and control-matrix section of Whirlwind I central control. The row of racks on the extreme right contains part of the arithmetic element.

computation with some preferred field of application in mind. The student of mathematics, physics, fluid flow, or statistics and operations analysis can add a study of digital computation techniques to his curriculum. The student in servomechanisms, meteorology, gas turbines, or aeronautics can add to his work the necessary mathematical analysis and machine-computation courses to allow the use of these tools in his special field.

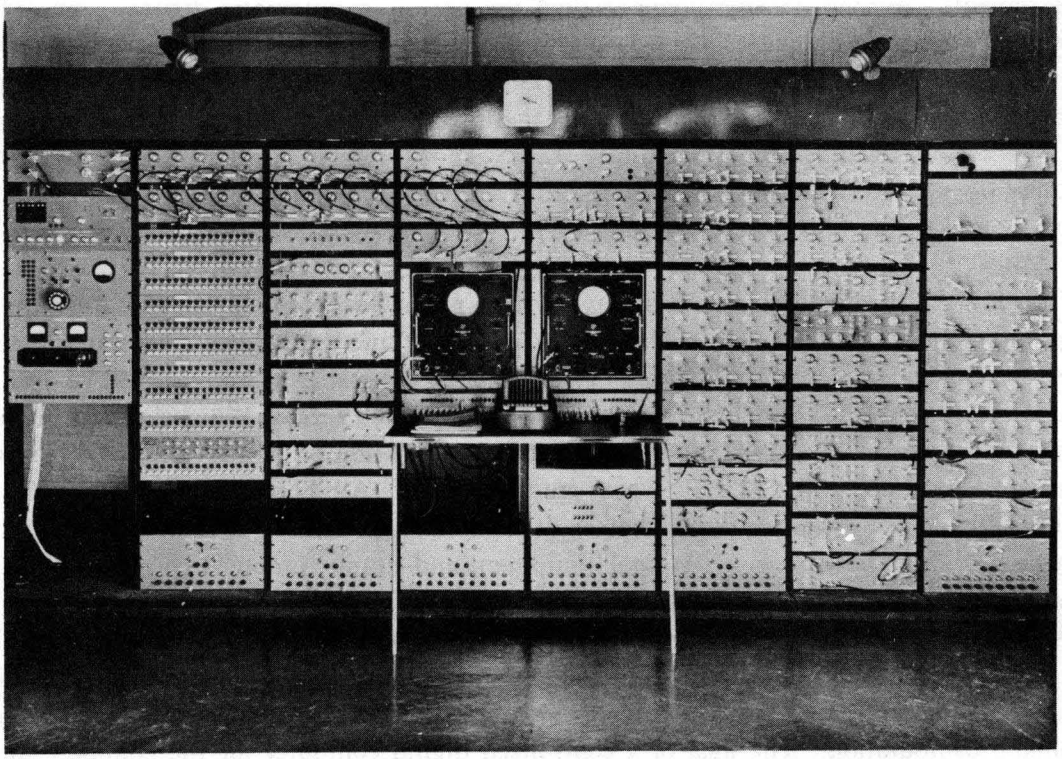


FIG. 2. Test center used during Whirlwind I installation. Power, marginal checking, and air-conditioning controls are on the right. Indicator lights in the center section connect to flip-flops in the computer. Both oscilloscopes have double-beam tubes connected to amplifiers, and remote probes for examining wave forms anywhere in the computer room. The remaining panels are computer test equipment used to generate any desired sequence of video testing pulses.

During part of the spring term, Mr. W. Gordon Welchman taught logic and coding for a digital computer and how to set up problems for automatic solution. This work will be expanded when Whirlwind I is operating and when arrangements are made to use it for student laboratory.

A major part of the M.I.T. training in digital-computer techniques is now made through the academic staff program. About a third of the Project Whirlwind staff is working toward advanced degrees. The men are on nominal full-time appointment which permits their taking two graduate courses. Fifteen to twenty research theses per year are related to the

digital-computer program. Last year these included electronics, such as studies of flip-flop circuits and secondary emission in vacuum; several were on trouble-location methods in digital computers; and others were in problem coding. One of the latter on the naval architecture procedure of *Intact Stability Study* of surface ships developed the digital-computer coding to go automatically from hull cross section to righting moments at various water lines and ship displacements. Another thesis student studied the use of an automatic digital computer in solving the alternating-current power-system problem for which the a-c network analyzer is commonly used. Now that machine construction is nearing completion and computer applications begin to occupy more of the staff time, thesis studies are less often on circuits and more often on computer coding and applications. Last year a doctorate thesis dealt with the theory of sampling servomechanisms where data from a digital computer are transmitted intermittently to control an external physical system.

Crystal balls are a little cloudy these days, but we might discuss the future digital-computer program at M.I.T. This is as diversified as the number of interested departments and laboratories. In the Servomechanisms Laboratory a principal interest is in using digital computers in automatic control. This includes simulation work for which Project Whirlwind was initially started by the Navy. Another long-range potential application of digital computers is in the control of air traffic, which M.I.T. is now studying for the Air Force. Digital computers in control will require extensive paper studies of methods and utility; and in the laboratory must be developed new types of terminal equipment and simple, practical conversion devices between digital and analog information.

A little closer at hand, I hope we can work on ways to make digital computers accessible to a wider group of users. There is an amazing number of technical people, from routine engineering offices to research scientists, who should be able to save time and money by using automatic computation. The idea of a centralized digital computer for the common use of many clients brings cries of anguish from those who hope to own machines for their private use. However, most potential users can have no hope of privately owning such facilities, or establishing training, and machine administrative procedures, which will make the central machine a success. We expect to approach this cautiously by beginning to work with other groups at M.I.T., first the other laboratories in the M.I.T. electrical engineering department, and to expand as conditions warrant.

Another untouched field is in the industrial applications of digital control. Thus far, most computer work has been sponsored by military research. The military uses are more obvious and urgent and, to date, few; but government groups have been able to invest in this long-range research. Already commercial concerns are actively working on the accounting and bookkeeping possibilities of automatic computers. Other areas are untouched, and I hope M.I.T. can extend its work of developing the theory of linear servomechanisms into nonlinear control using digital computers. Such things as the operation of chemical plants and calculating the plant balance of oil refineries are attractive possibilities. The Servomechanisms Laboratory

is now working with one company on digital control in a manufacturing process, although this does not involve a complete digital computer. Another indirectly related project is being sponsored by the Carnegie Foundation on the logic and coding of bibliographic information. Here methods of doing indexing by the association of idcas rather than in the elementary manner of card catalogs will at least require flexible computing facilities for the research, if not for the ultimate location of information. Thus far not all the time of Whirlwind I is scheduled for use, and I believe arrangements for any legitimate use can be made with the Navy. I say legitimate to exclude certain obvious statistical studies in connection with horse racing and the stock market. Whirlwind I should help to assess the value of digital computers in many proposed but as yet untried applications.

I expect that Whirlwind I will be available for exploring new ways of using digital computers. It will be most useful if it carries new applications to the point where success is demonstrated. Other computers designed and located elsewhere should then take over routine work as the need develops, in order that the M.I.T. laboratories may be free to continue explorations in new fields.

The Raytheon Electronic Digital Computer

RICHARD M. BLOCH

Raytheon Manufacturing Company

This paper describes the essential characteristics of the electronic digital computers now being designed and constructed by the Raytheon Manufacturing Company for the Special Devices Center of the Office of Naval Research; a machine having similar features is also being developed by Raytheon for the National Bureau of Standards.

I should first like to stress certain considerations that governed the design of these machines. It is clear that whenever the specifications for an electronic digital computer are set forth, certain minimal speed requirements are given, depending upon the application for which the computer is primarily intended. However, it behooves the machine designer to meet these requirements in such a way that the computer possesses a desirable speed balance among its several major components. Thus, there would appear to be little sense in designing an arithmetic unit capable of operating at very high speeds if the information *required* by the arithmetic unit cannot be obtained from the internal or external memory at a correspondingly high speed. As a matter of fact, the purchaser of a digital computer is not particularly interested in the speed of any single component of the machine. His interest obviously rests in the time required by the computer to complete successfully the solution of those problems that will most frequently be placed upon the machine.

A second consideration is that of reliability of operation. If any considerable time must be spent in repair and maintenance of a digital computer, the effective speed of the machine may easily be reduced by a factor of two or three, and the patience of the operating crew reduced by an even greater amount. Time spent by the designer in improving the reliability of the machine's components will be returned many fold when the computer is placed in operation. However, even though the error rate can be substantially reduced by stressing the reliability aspect in the design stage, errors will nevertheless occur. It is at this point that the diagnostic capabilities of the computer assume an important role. The locating of a machine fault may be a very serious and discouraging matter. There seems to be a popular misconception that when an error occurs, the failure can be traced to the arithmetic unit. This, unfortunately, is far too coöperative a spirit to expect from such a complex device. Whereas it is true that the primary task of the machine is that of performing the fundamental arithmetic operations, there are multitudes of operations of a *nonarithmetic* nature that are taking place within the computer. As an illustration, consider the case wherein an incorrect product is obtained by the machine, and suppose this fact is detected by some programmed or automatic checking device. Now, let us inspect a few of the possibilities. (1) The multiplication may have been performed incorrectly. (2) Although the multiplication was executed

properly, the product read-out circuit may have failed. (3) The transmission of either the multiplier or the multiplicand to the arithmetic unit may have been in error. (4) Either of the factors may have undergone a change while stored in the memory unit. (5) Indeed, a multiplication may never have been performed, but rather an addition or a division. It is of interest to note that only the first-mentioned failure involves an error in the process of arithmetic combination. The others are attributable to failures in the control and transmission networks. Clearly then, from the diagnostic point of view, if checking means are to be employed at all, it is important that the entire machine—not one particular unit—be under the surveillance of a comprehensive checking system. In the computer to be described shortly, an automatic self-checking system is employed throughout the machine to monitor all input, output, inter-unit, control, and arithmetic networks.

It should also be stated that from the mathematical viewpoint this same checking system which serves so vital a function for diagnostic purposes performs an invaluable service in halting the solution of a problem at the instant an error is detected. As problems of greater complexity are introduced to electronic digital computers, it becomes essential to stop the machine immediately lest several hours of erroneous computation ensue. Programming checks, such as performing a multiplication to check a division, or indeed solving the problem a second time utilizing an entirely new set of program orders, have distinct disadvantages which precluded the possibility of their use in the Raytheon computer. If such checks are to be applied conscientiously, then the time for solution is more than doubled and as many machine errors are to be anticipated in the execution of the checking operations as are to be expected from the original programming. Furthermore, from the diagnostic viewpoint programmed checks appear to be exceedingly weak.

A third design consideration that should be mentioned concerns programming and problem preparation. The computer to be described has been designed with a view toward reducing the time and labor involved in programming a problem for machine computation. Whereas theoretically it is possible logically to reduce all machine operations to a few basic processes, under such a scheme the program coder must resign himself to a task requiring an undue waste of time and mental effort—such gymnastics should be relegated to the sphere of the *machine's* operation. In general, the programming should closely correspond to the original mathematical formulation of the problem. As far as possible, the identity of what we understand as a single mathematical operation should be preserved in the programming to the extent that that operation is represented as a single order in the programming routine. It has been our thought that programming should be a straightforward and natural process—not one involving elaborate planning, numerous restrictions, and the continual use of mathematical ingenuity. It is in the *formulation*, in terms of machine processes, of the complex and as yet unsolved problems of mathematics and its kindred sciences that the mathematical talent of today might better spend its energy.

With the foregoing considerations forming the background for the machine's design, we

shall proceed with a description of the computer. The Raytheon machine operates in the true binary scale of notation, having a basic precision of 30 binary columns; all 30 columns are located to the right of the binary point. Information can, however, be accepted by or transmitted from the machine in decimal as well as binary notation. In the case of decimal numbers, each digit is represented by its four-column binary-coded-decimal equivalent. Conversion of decimal numbers to binary, as well as the inverse conversion from binary to decimal scale, is accomplished within the arithmetic unit of the computer. In standard

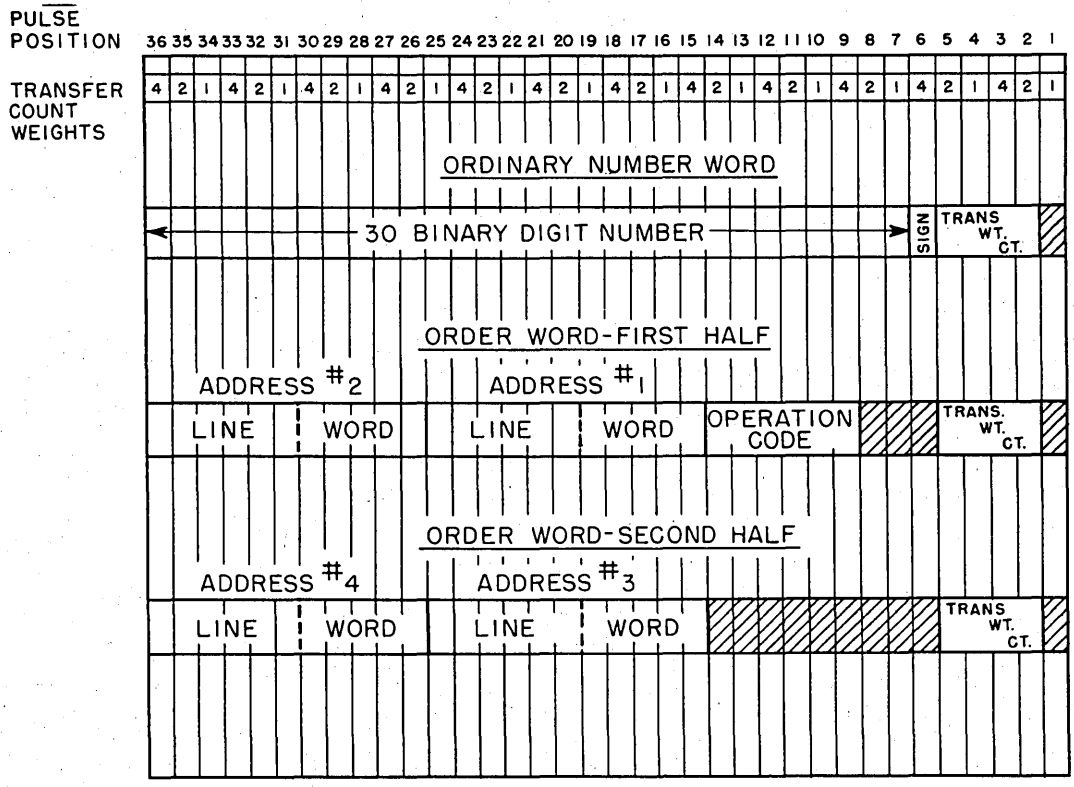


FIG. 1. Information allocation.

operation, a number is stored as a positive absolute value with proper algebraic sign and to a precision of 30 binary columns. Numbers of 60 binary digits are stored as a pair of standard numbers, and are processed in the arithmetic unit through the use of double-precision operations.

Floating-point operation is possible in this computer, and when this mode of operation is utilized numbers are stored with the first significant digit resting in the second binary place; the appropriate exponent on base 2 is stored in a separate memory position.

Each memory position of the computer has a capacity of 36 binary digits, and this sequence of digits is termed a word, which is the basic unit of information storage. Figure 1 shows the allocation of information in a number word. Pulse positions 7 through 36 hold the absolute

value of the number; the first binary column is located at position 36 and the thirtieth binary column lies at position 7. Pulse position 6 contains the algebraic sign—a 1 if the number is negative, a 0 if it is positive. Positions 2 through 5 contain the transfer weighted count of the number, which will be defined shortly; position 1 is blank.

Programming orders are stored in the internal or external memory of the computer in two parts, termed the first and second order words. The two half-orders are always stored in adjacent memory positions, and when the first half-order is called forth for control purposes, the second half-order automatically follows in sequence. Positions in the memory are numbered successively in binary notation, and these numbers will henceforth be referred to as addresses. A four-address programming system is employed in the computer. The first two addresses specify the positions in the memory where the operands for a given arithmetic operation are to be found. The third address indicates the position in the memory to which the result of the arithmetic operation is to be transmitted. Finally, the fourth address specifies the memory position wherein is located the next order to govern the machine's operation. The operation code, which indicates which of the 30 arithmetic operations is to be performed, also forms part of the information residing in each order. The details of the allocation of information in the first and second order words are also shown in Fig. 1. The line and word indications in each address in the diagram will be explained when the memory arrangement is discussed in more detail. All information in the order words is of course represented in binary notation. Other words may be combined in the arithmetic unit in the same fashion as number words. This feature, together with the fact that order words are stored in the regular memory units, permits a high degree of flexibility in programming which would not otherwise be possible; frequently, the number of orders required to program a problem may be reduced substantially. Furthermore, certain routines such as interpolation may be performed very rapidly, and without recourse to hunting techniques.

All storage of words in both the internal and external memory and all inter-unit transfers of words are checked by means of a weighted count, i.e., a weighted sum of the digits of the informational portion of the word (including the algebraic-sign digit if the word is a number word). As the diagram shows, this weighted count is stored with the number or half-order and is called the transfer weighted count. The weights chosen for the sum are the numbers 1, 2, 4, 1, 2, 4, 1, 2, 4, etc., which are assigned to the successive digital positions from right to left. This weighted binary sum is computed modulo 16 and is then modified by the addition of unity; thus a number and its weighted count cannot both be zero simultaneously. With this modification, a null word is not a valid word, and the complete failure of a gate or other device controlling the entire transmission channel will be detected. Since the sum is computed modulo 16, obviously only four digital positions are required to represent the transfer weighted count. The transfer weighted count is automatically constructed and checked when the numbers and orders are being prepared for machine entry by the problem preparation unit. Thenceforth, whenever the number word or order word is transferred from one machine unit to another, a new weighted count is constructed; failure of this new count to check with

the original count stops the machine, and an indication of this particular error is given to the operator. The transfer weighted count is not completely foolproof—no check is, for that matter. However, the extraordinary power of the count as a checking means rests in the very peculiar array of compensating and simultaneous errors that must occur in order to invalidate the check.

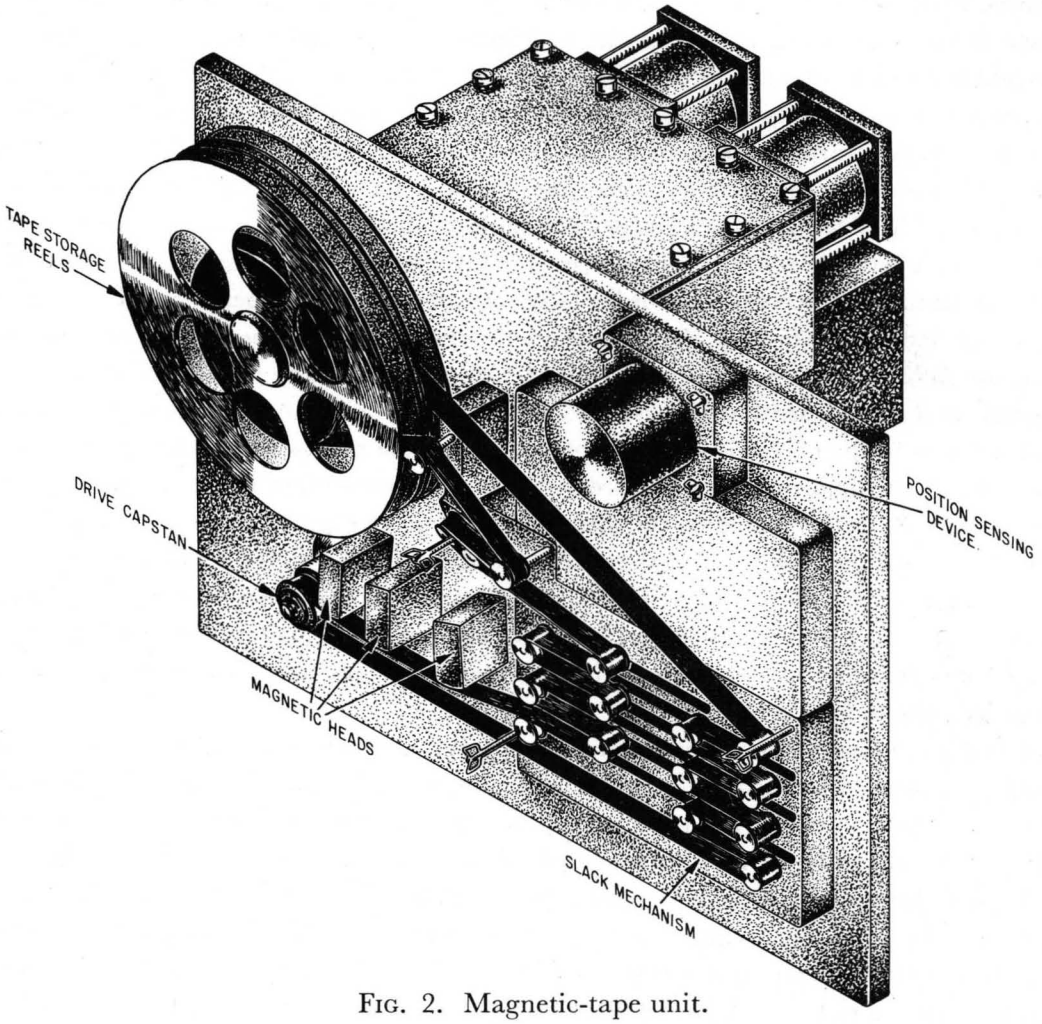


FIG. 2. Magnetic-tape unit.

The external memory consists of four magnetic-tape units, each having a storage capacity of approximately 100,000 words. A diagram of one of these units is shown in Fig. 2. These devices may be used as input units to supply numbers and orders to the machine, or as output units for the recording of intermediate and final results. Six-channel plastic tape coated with iron oxide is used as the magnetic medium, and tape having a width of approximately 0.5 in. and a thickness of 0.003 in. In each of the six channels, pulses are recorded with a density of 100 pulses to the inch, and the over-all reading and recording rate of each magnetic-tape

unit is roughly 400 words per second, corresponding to a tape speed of 30 in./sec. Recording upon or reading from the tape is performed in blocks of 32 words, each block occupying somewhat more than 2 linear inches on the tape. To each block of words on the tape is associated a 12-column binary number which is termed a "block number." The 12 binary indications are permanently placed on the back of the tape using a system of horizontal markings which are sensed photoelectrically during the hunting process; two sets of such markings are used—one when hunting for a block number in a forward direction, the other when hunting in a backward direction; in this way the number of reversals of tape motion

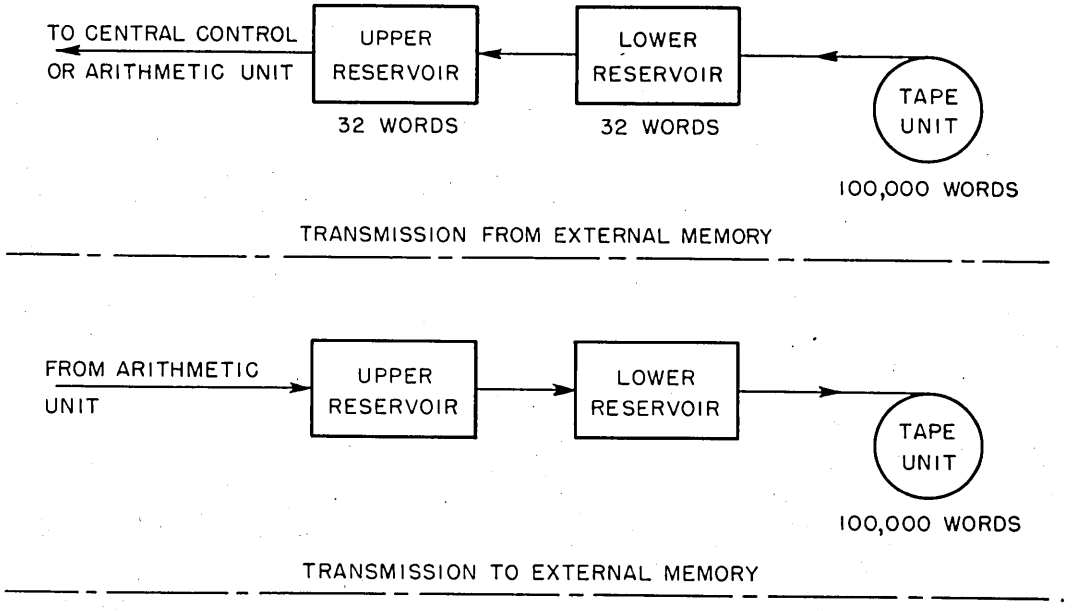


FIG. 3. Organization of external memory.

is held to a minimum. Separate markings are also used to indicate the beginning and end of each block of words. These permanent markings are not affected in the process of magnetic erasure, and they provide a means for visual inspection of the block numbers whenever this seems desirable.

As shown schematically in Fig. 3, there are two 32-word mercury delay lines or reservoirs associated with each tape unit; these reservoirs are used as buffers between the tape mechanism and the main electronic part of the machine. The external memory units respond to four distinct commands: (1) Tape Read, (2) Tape Record, (3) Hunt-Prepare to Read, (4) Hunt-Prepare to Record. In the Tape Read operation, the 32 words in the lower reservoir are transferred at high speed to the upper reservoir from which point the words are subject to call at any future time and in any sequence whatsoever by the central control. The addresses assigned to the upper reservoir are very similar to those that identify the internal memory positions; and, in fact, these upper reservoirs may be utilized as additional internal memory capacity. When the lower reservoir has been emptied in the course of the Tape Read operation,

the tape unit proceeds to fill this reservoir with the next block of 32 words. The computing routine of the machine, however, does not cease during this tape-to-reservoir transfer. If now a second Tape Read order occurs before the lower reservoir has been filled, then the machine stops and awaits the conclusion of this process. It can be seen, however, that if the mean rate of call for new words from a single external memory unit does not exceed 400 words per second, then no time is lost as a result of the tape-reading process; and the machine effectively possesses a high-speed internal memory capacity of several hundred thousand words.

In the case of recording, the words to be recorded are transmitted to the upper reservoir by the arithmetic unit in the course of the computation; here again, this reservoir receives words in the same fashion as any of the regular internal memory positions. Upon receipt of a Tape Record order, the contents of the upper reservoir are shifted at high speed to the lower reservoir, whereupon the computer is free to proceed with its computational routine. Meanwhile, the contents of the lower reservoir are recorded upon the tape at a rate of 400 words per second. If a second Tape Record order is called before the lower reservoir has been emptied, the machine is stopped awaiting the completion of this process, but, once again, this will occur only if the mean rate of words to be recorded on this one unit exceeds 400 words per second, and this situation should occur very infrequently.

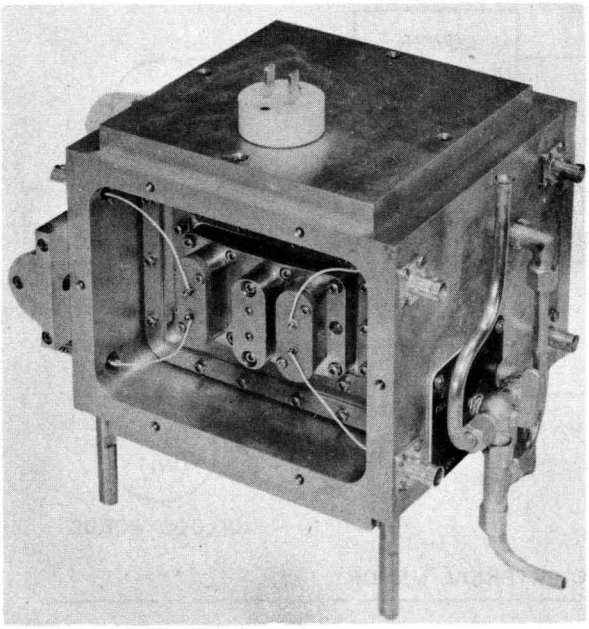


FIG. 4. Four-line mercury tank.

The Hunt-Prepare to Record order causes the particular external memory unit called to hunt for the block number resting in a special one-word external memory storage position known as the Hunt Register. When this block is found on the magnetic tape and verified through the use of the weighted count check, the tape unit stops and is now prepared to record at the proper block position.

The fourth external memory order, namely Hunt-Prepare to Read, is executed in substantially the same manner as the hunt operation just described. Here, however, when the proper block is located and verified, two successive blocks of 32 words are then read into the upper and lower reservoirs respectively, and the machine is prepared to use this information directly without further delay. Figure 4 shows a four-line temperature-controlled mercury tank having dimensions of approximately 6 in. on a side; this tank contains the upper and lower reservoirs associated with each of the four external memory units.

The internal memory of this computer consists of a set of 32 circulating mercury delay

lines operating at a pulse repetition rate of approximately 3.78 megacycles per second. Each line is capable of storing 32 words of 36 binary digits each. Thus, the internal memory has a total capacity of 1024 words. Words are stored serially and are transmitted in a pulse-by-pulse fashion to the other units of the machine. The time during which one word is made available for transmission is termed a minor cycle, having a time duration of about 9.5 μsec. One circulation of the delay line requires 32 minor cycles, and in this time interval each word stored in the line is available for transmission precisely once. The minor cycles are numbered in binary notation successively from 0 to 31, and each number identifies a word position in the line. Furthermore, the 32 delay lines, which are all in synchronism, are also numbered from 0 to 31 in binary notation. Thus, the number of a delay line and the number of a minor cycle together completely identify a word position and constitute what we have previously defined as an address. Although 10 binary digits suffice for the representation of each internal memory position, an eleventh digit is also used; this permits addresses to be assigned to the 128 additional positions located in the upper reservoirs of the external memory, as well as certain special one-word storage positions such as the Hunt Register.

Information is never erased from a memory position unless a new word is specifically transmitted to that position, at which time the erasure of the former contents occurs automatically.

The mercury lines have undergone extensive testing for the reliability of recirculation, read-in, and read-out; the results have shown that these lines will meet successfully the high reliability requirements that have been placed upon the computer.

The central-control unit may be described as the nerve center of the computer. It is the duty of this unit to extract programming orders from the memory at the proper time, interpret them accurately, perform the appropriate selections, and verify the fact that the order has been correctly executed.

Central control must select information from the memory in accordance with the line and the word numbers that form each address. The proper line is chosen under the control of a line-selection matrix. A check is performed to ensure that this selection was not in error. Whenever one of the memory lines is selected by the matrix, the binary identification number assigned to that line is generated automatically. This identification number or tag is compared with the portion of the address that governed the original selecting process, and any discrepancy will indicate a false selection; the machine stops at this point and the operator is given an indication of the cause of the failure.

For each address, another selection must be made under the jurisdiction of the central control—namely, the *word* selection determined by the word-number section of the governing address. At the beginning of the appropriate minor cycle, a pulse is transmitted to the memory gates permitting a word to be read from or into the appropriate word position of the line that has already been selected. To check that this temporal selection was performed correctly, an additional mercury line is employed which is in synchronism with the other lines of the memory unit. In each of the 32 word positions of this word-check line, as it is termed, is contained

the binary code which is associated with that particular word position. At the instant when a word selection is made from any of the information lines, the read-out gate of the word-check line is also activated, and the appropriate binary number is sent forth to the central control. If this number agrees with the word-number portion of the governing address, a powerful check is obtained which ensures that the desired word-time selection was properly performed. Although the line and word selections are known to be correct, it is still possible that the information within the word has undergone a change while stored in the memory. This possibility is guarded against by the application of a transfer-weighted-count check. In the case of order words, this check is performed by central control, while mutations in number words are detected by the arithmetic unit before any arithmetic process occurs.

The selection of the proper arithmetic operation in accordance with the operation code of the order is also, logically, a function of the central-control unit. This selection is performed by means of an operation code matrix; however, to preclude the possibility that the arithmetic unit will misinterpret the operation to be carried out, a check-back with central control is initiated.

In accordance with the tentative operation signal received by the arithmetic unit, the code corresponding to this operation is transmitted to central control for verification. Only if this new code corresponds to the operation code called for in the governing order, does the central control notify the arithmetic unit to proceed; otherwise, the appropriate error signal is flashed to the operator, and the computer's operation ceases. Generally speaking, any deviation from proper performance on the part of any of the machine's units comes under the cognizance of the central control.

This computer operates in a variable-cycle mode of operation whereby each new function to be performed by the machine is initiated by the successful completion of the previous function. As a result, the total time for a machine operation is not fixed, but will vary with the exigencies of the particular order being performed. Thus, in determining the speed of the machine, the *mean* time consumed in the performance of a complete order must form the basis of such calculations.

Since there is only one line-selection matrix in the computer, one might expect that the average time required for a memory selection would be approximately 16 or 17 minor cycles, there being 32 words in each line. However, by means of a system of anticipatory selection, the central-control unit is capable of reducing this average time substantially. Essentially, the four addresses are treated in pairs for selection purposes, and the address that is capable of being selected with the least time delay is chosen; thus address 2 may be selected before address 1, or address 4 before address 3, depending upon the time relation existing between the word-number parts of the addresses. Such possible inversions of the sequence of selections have no effect upon the proper execution of the programming order.

In some operations one or more of the addresses may be void; also, certain special addresses may occur that correspond to one-word storage positions. In the first case, the central control by-passes the selection completely; in the second case, advantage is taken of the fact that the

word is available for transmission during every minor cycle. In both cases, further substantial reductions in memory reference time are accomplished.

The arithmetic unit of the Raytheon computer consists of the necessary equipment for the execution and checking of the various arithmetic operations. The unit is designed for parallel operation, thus maintaining the desired speed balance among the various units of the machine. A small checking unit operates simultaneously with the main unit; this checking unit combines quantities known as arithmetic weighted counts, and the fulfilment of certain relations involving these counts provides a check upon the operation of the main arithmetic unit. The method by which the arithmetic weighted count of a binary number is derived

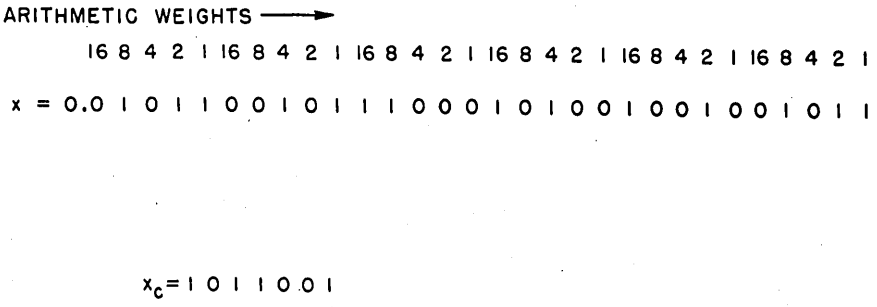


FIG. 5. Arithmetic weights.

is shown in Fig. 5. This count is defined as the weighted binary sum of the digits of a number, where the successive weights are 1, 2, 4, 8, 16, 1, 2, 4, 8, 16, etc., applied from the right to the left, beginning with the lowest-order column of the number. In Fig. 5, the count has been applied to the 30-column binary number x, and x_c indicates the resultant 7-column binary number which is the arithmetic weighted count of x. It should be noted that, unlike the transfer weighted count, no reduction modulo 16 is performed.

Figure 6 shows the checking identities that are applied to the four basic arithmetic operations. The subscript c denotes the arithmetic weighted count of the quantity to which it is appended. It is seen that in each case an invariant result is obtained which is independent of the values of X, Y, and Z; as indicated, the checking identity will produce the value 31, or in the binary system a sequence of five 1's, provided the result Z obtained in the main arithmetic unit is correct.

In Fig. 7 is given a list of the 30 operation codes that will be employed in this computer. The column headed "Binary" shows the actual sequence of digits as they occur within the machine in the operation-code section of an order. However, in order to simplify the

programming procedure, all orders are coded in the octal notation; and, as a matter of fact, the programmer need not concern himself with the equivalent binary notation, since all addresses and operation codes are prescribed beforehand in the octal system. In this system the binary columns are grouped three at a time, and to the eight possible configurations of the three binary digits there corresponds one of the eight possible octal digits—0 through 7. It should be noted that in terms of total digits required for number representation, the octal system closely approaches the decimal system in efficiency. As a further aid to the programmer, the operation codes are arranged so that the first octal digit denotes a certain family of related operations. Thus, where the first octal digit is 0, addition or subtraction is indicated; 1 denotes

1. ADDITION $X + Y = Z$

$$(|Z_c - (X_c + Y_c)| + 31)_c \equiv 31$$

2. SUBTRACTION $X - Y = Z$

$$(|Z_c - (X_c - Y_c)| + 31)_c \equiv 31$$

3. MULTIPLICATION $X \cdot Y = Z$

$$(|(X_c \cdot Y_c)_c - Z_c| + 31)_c \equiv 31$$

4. DIVISION $X/Y = Z + R/Y$

$$(|(Y_c \cdot Z_c)_c + R_c - X_c| + 31)_c \equiv 31$$

FIG. 6. Checking identities.

multiplication or division; 2, a transfer operation; 3, a shifting or extraction process; 4, the two substitution operations; 5, the branch operations; 6, the floating-point processes; and 7, the codes that pertain to the external memory units.

Space does not permit a complete description of each operation; however, Fig. 8 shows the manner in which each of six representative operations is programmed. In addition, the address of the addend is placed in the Address 1 position, that of the augend in the Address 2 position; the addition code 01 is inserted in the Operation Code position. The address to which the sum is to be transmitted is located at the Address 3 position. If the result is to be used immediately in the next operation, and if there is no need to transmit this result to the memory, then the third address position may be left void, and the result may be called forth in the next operation through the use of a special address. However, whether the third address is void or not, the result of the present operation remains available in a special one-word register of the arithmetic unit, and is subject to call by employing the above-mentioned special

OPERATION	BINARY	OCTAL
ADDITION	0 0 0 0 0 1	0 1
ADDITION (DOUBLE - PRECISION - PART 1)	0 0 0 0 1 0	0 2
ADDITION (DOUBLE - PRECISION - PART 2)	0 0 0 0 1 1	0 3
SUBTRACTION	0 0 0 1 0 0	0 4
SUBTRACTION (DOUBLE-PRECISION - PART 1)	0 0 0 1 0 1	0 5
SUBTRACTION (DOUBLE-PRECISION - PART 2)	0 0 0 1 1 0	0 6
MULTIPLICATION (TRANSMIT HIGH ORDER WITH ROUNDOFF)	0 0 1 0 0 0	1 0
MULTIPLICATION (TRANSMIT HIGH ORDER NO ROUNDOFF)	0 0 1 0 0 1	1 1
MULTIPLICATION (TRANSMIT LOW ORDER)	0 0 1 0 1 0	1 2
DIVISION (QUOTIENT NOT ROUNDED-OFF; REMAINDER AVAILABLE)	0 0 1 1 1 0	1 6
DIVISION (QUOTIENT ROUNDED-OFF)	0 0 1 1 1 1	1 7
TRANSFER (NORMAL)	0 1 0 0 0 0	2 0
TRANSFER (POSITIVE ABSOLUTE VALUE)	0 1 0 0 0 1	2 1
TRANSFER (NEGATIVE ABSOLUTE VALUE)	0 1 0 0 1 0	2 2
TRANSFER (SELECTIVE)	0 1 0 0 1 1	2 3
SHIFT (CONTROLLED)	0 1 1 0 0 0	3 0
SHIFT FACTOR (NORMAL)	0 1 1 0 0 1	3 1
SHIFT FACTOR (SQUARE ROOT)	0 1 1 0 1 0	3 2
EXTRACTION	0 1 1 1 1 0	3 6
SUBSTITUTION (ADDITIVE)	1 0 0 0 0 1	4 1
SUBSTITUTION (SUBTRACTIVE)	1 0 0 1 0 0	4 4
BRANCH (NORMAL)	1 0 1 0 0 0	5 0
BRANCH (EQUALITY SENSING)	1 0 1 0 0 1	5 1
ADDITION (FLOATING)	1 1 0 0 0 0	6 0
SUBTRACTION (FLOATING)	1 1 0 0 0 1	6 1
MULTIPLICATION (FLOATING)	1 1 0 0 1 0	6 2
TAPE RECORD	1 1 1 0 0 0	7 0
TAPE READ	1 1 1 0 0 1	7 1
HUNT – PREPARE TO RECORD	1 1 1 0 1 0	7 2
HUNT – PREPARE TO READ	1 1 1 0 1 1	7 3

FIG. 7. Operation codes.

address in the succeeding order. It is clear that by this means useless transmissions to or from the memory are completely avoided. It should be understood, of course, that the device just discussed may be applied to any desired operation and is not restricted to the addition process. To continue, Address 4 will contain the memory address from which the new order is to be obtained—or, more exactly, from which the first half of the new order is to be selected. The programming arrangements for subtraction, multiplication, and division, shown in Fig. 8, I believe, are self-explanatory in the light of the above discussion.

The selective transfer order directs the machine to multiply the number in memory position

OPERATION	ADDRESS #1	ADDRESS #2	OPERATION CODE	ADDRESS #3	ADDRESS #4
ADDITION	ADDEND	AUGEND	01	SUM	NEXT ORDER
SUBTRAC-TION	MINUEND	SUBTRA-HEND	04	DIFFER-ENCE	NEXT ORDER
MULTIPLI-CATION	MULTIPLI-CAND	MULTI-PLIER	10	PRODUCT	NEXT ORDER
DIVISION	DIVISOR	DIVIDEND	17	QUOTIENT	NEXT ORDER
TRANSFER (SELECTIVE)	A	C	23	B	NEXT ORDER
BRANCH (NORMAL)	A	B	50	C	D

FIG. 8. Construction of computer orders.

A by $+1$ or -1, according to whether the number in memory position C is positive or negative; the result of this process is then to be transmitted to memory position B.

The Branch-Normal operation is somewhat unusual both in its effect on the subsequent computation and in the treatment of the third address. If the number in memory position A is greater than or equal to the number located at memory position B, the machine is directed to obtain its next order from storage position C as indicated by Address 3; otherwise, the central control is to obtain the next order from memory position D as specified by the fourth address.

The problem-preparation unit is a manually operated device, independent of the main computer, that places the programming orders and numerical input information on the magnetic tape in preparation for entry into the machine. A first Teletype keyboard unit is used to prepare a standard five-hole Teletype paper tape. This tape is used in conjunction

with another keyboard unit to prepare a second paper tape. The operator of the second unit, reading from the same manuscript that was used in the preparation of the first tape, causes a second paper tape to be perforated. However, each key that is depressed must establish an identity with the Teletype code appearing on the preliminary tape; otherwise, the keyboard is automatically locked and the intended perforation of an erroneous code on the second tape is intercepted. A printer is also associated with the keyboard so that a printed copy of the programming is available to the operator in the course of the tape-preparation process. Programming orders are entered in the keyboard in octal notation, whereas numerical information will generally be entered in decimal notation; this conforms to the notation that is prescribed for the original manuscript. However, where desired, *numbers* may also be entered in the octal system if the information should happen to be available in this form; provision is made on the keyboard to indicate the particular notation being used.

After the second paper tape is prepared, it is transferred to a magnetic recording unit where the Teletype codes appearing on the tape are converted to binary or binary-coded decimal notation in accordance with a coded indication that accompanies each word on the paper tape; a transfer weighted count of each number and half-order is also automatically constructed. All of this information is then recorded on the magnetic tape in the required word-and-block form previously described; this magnetic tape is now transferred to one of the external memory units from which point the information is automatically available to the computer. These conversion and weight-counting processes, as well as the magnetic recording process itself, are all automatically checked.

The printing of the final results of a problem is performed by Teletype printers operating independently of the main machine. In the course of a computation, numbers to be printed are shuttled in binary-coded decimal form to one or more of the magnetic tapes associated with the external memory units. At a later time, these reels are transferred manually to the output printers where the numerical quantities are typed in final form. Directions to the printer involving considerations such as page format, location of the decimal point, etc., are supplied by auxiliary control devices.

Certain external memory tapes contain words in binary notation only, these quantities being intermediate values obtained in the computation and intended for direct feedback to the machine at a subsequent point in the solution of a problem; therefore, the computer obviously will not have been instructed to convert these numbers into binary-coded decimal notation. However, it may be desired on certain occasions to print the *binary* quantities contained in these intermediate tapes; for this reason, provisions have been made for the output printers to type numerical quantities in octal notation as well as decimal. Transfer weighted-count checks based upon the actuation of the printer code bars are employed to intercept printing errors.

A printer directly connected with the computer is provided so that the operator may monitor intermediate results of the computation while the machine is in operation. The transmittal of information to this printer is prearranged in the original programming of the problem.

Provisions for reading from, as well as into, the various high-speed memory positions of the machine under manual control have been made. Furthermore, the computer is designed in such a way that when an error occurs, causing the machine to stop, all numerical and control quantities involved in the execution of the order then governing the machine are available for immediate read-out by the operator; this feature should be an invaluable aid in the diagnosis of failures.

The Raytheon computer operates at a mean speed of 1600 complete operations per second. The machine has a complement of roughly 3500 vacuum tubes and 6500 crystal diodes. It is expected that construction of the computer will be completed at the end of September 1950

A General Electric Engineering Digital Computer

BURTON R. LESTER

General Electric Company

The General Electric Company feels very grateful for the opportunity to discuss our computer at this symposium. Our present effort is the design and construction of a computer suitable for the engineering problems that arise within the General Electric Company—a computer simple in design, accurate and reliable, easy to operate, and economical to maintain.

The General Electric Company has been interested in the computer field for many years. Our first computers were network analyzers. These were soon followed by the more complex differential analyzers. We built up small computation groups in our engineering divisions. In some instances, as the requirements for speed and accuracy increased, we rented IBM machinery.

Approximately six years ago, the first investigation of the possibility of constructing an automatic digital computer was made. In 1946 work was started on a small binary machine for control problems. Our Engineering Council also directed that we investigate the possibility of constructing another machine for internal use. A careful review was made of the various computer projects and a small group of engineers visited these projects to weigh their progress.

These efforts culminated in the decision to construct a computer. Our design and construction were based on the experience and ability of our Research Laboratory and our electronic accomplishments during the past war.

Our purpose in constructing this computer was threefold. The computer would enable more accurate and rapid computation of our engineering designs. It would provide our Research Laboratory with a long-needed facility. Last, we would gain extensive knowledge and position in this field by constructing and operating this machine.

The purpose of our computer set the major design considerations. Accurate and reliable operation was foremost. Consequently, only proven principles were utilized. A reasonable operating speed was set with the idea of increasing it gradually as we become more familiar with the capabilities of the computer. Operation and maintenance procedures were simplified. Unitized construction was employed to aid design, speed maintenance, and provide means for adding future improvements.

In discussing the features of any computer, it is well to break the design down into the following items for easy assimilation: number base, mode of operation, memory, arithmetic unit, control unit, input-output mechanism, tape-preparation unit, and printer.

Our computer operates in the decimal system. All numbers and instructions are expressed in decimal digits. The 2* coded decimal system is used within the computer. The basic

length of a number is 8 decimal digits using a fixed decimal point and 6 decimal digits with a floating point. The range of the latter system is from 10^{-9} to 10^{9}. The simplicity of the decimal operation outweighed the loss of capacity as compared with the equivalent binary machine. Actually, the construction of the machine is binary with the exception of three vacuum tubes per decimal digit in the accumulator. These tubes and associated circuits sense and correct forbidden combinations.

Serial and parallel operation are used to advantage. Serial operation occurs between all major units. Parallel operation is used within the arithmetic and control units. Numbers and instructions are stored serially in the memory in single binary channels. Serial read-in and read-out of the memory occurs at a 48-kc/sec repetition rate. Numbers and instructions are transferred between the arithmetic and control registers at a 200-kc/sec rate. Parallel operation is utilized for basic operations within the arithmetic and control registers at a 200-kc/sec rate.

Operation of the computer is as follows. An order stored in the memory is read serially to the control unit. This unit reads each part of the order in parallel by means of sensing circuits. First it directs the memory to transfer the two operands serially to the arithmetic unit. Then the control unit senses the order to determine the operation code and directs the arithmetic unit to perform this operation in parallel. The answer is transferred serially to the memory again under the direction of the control unit and to the address specified in the order. Finally, the control unit senses the present order to determine the address of the next order and transfers it to the control unit. Thus the basic cycle is repeated until the problem is solved.

The computer has a magnetic-drum memory that stores 4000 numbers and instructions in 100 tracks, 40 numbers per track. Pulses are spaced 20 to the inch in the tracks and the tracks are spaced 8 to the inch. Forty pulse spaces are required for each number—36 for the number and four guard spaces. The drum rotates at 1800 rev/min; consequently, the pulse repetition rate is approximately 48 kc/sec. The drum is constructed of aluminium with a magnetic coating and is 24 in. in diameter and 30 in. long.

Separate magnetic heads are provided for playback and recording. They are spaced 3 mils from the drum surface. Two playback amplifiers are used, one for information and one for clock pulses. A low-level crystal gating system connects the proper head to the information playback amplifier. The amplifiers have an automatic gain control to eliminate signal variation caused by eccentricity of the drum. A high-level gating system delivers recording pulses to the proper head. Recording and playback heads are spaced 180°. Consequently, a number recorded may be read back one half revolution later to verify memory operations. This check is performed after each recording.

Initially the design called for a serial arithmetic unit. However, it soon became apparent that the shifting registers, utilized to synchronize the numbers received from the memory, could easily be modified for parallel operation with a small increase in equipment. The arithmetic unit contains three basic registers: A, B, and C. Register A is an accumulator

and contains additional equipment to sense and correct forbidden binary numbers occurring as a result of addition. Register *B* is a shifting register. Register *C*, in addition to being a shifting register, can also be connected as a counter for use in multiplication and division.

The arithmetic unit performs the following basic operations: addition, subtraction, multiplication, division, and choice. Actually, the basic operation is the addition obtained in the accumulator. In subtraction, the nine's complement is used with end-around carry.

Multiplication is performed as a series of additions.

Division utilizes the oscillating overdraft method. Both the multiplication and division processes have been simplified to the extent that they are nominally equivalent to their binary counterparts.

Table 1. Operation times (μsec) for basic operations.

Operation	Fixed-Point Operation	Floating-Point Operation
Addition	15	350
Subtraction	15	350
Multiplication	450	490
Division	450	530

The times required for basic operations are listed in Table 1. These times do not include memory-access time. These times represent a small fraction of the time of one revolution of the drum. Under average conditions the time for completion of one operation including access time is equal to the time of one revolution of the memory.

The control unit is the telephone central of the computer. It controls the operations of the various units of the computer. The control signals are either a 40-v positive 2.5-μsec pulse or a d-c switching voltage of zero or + 40 v. The cyclic control rate is determined by the unit under control. Control signals to the memory are based on the 48-kc/sec clock pulse generated on the drum. Control signals to the tape input-output unit are based on synchronizing pulses on the tape. Control signals to the arithmetic unit are based on a 200-kc/sec oscillator in the control unit.

This type of control permits a great deal of freedom in introducing modification. Increasing memory operation only requires that the memory supply the correct clock pulse frequency to the control unit. No other modifications are needed. This flexibility is highly desirable to permit future improvements.

The four-address system is used in this computer. The first two addresses of the order are the locations of the operands. The third address is the location of where the answer is to be stored and the fourth address is that of the next order. A number consists of nine decimal digits, numbered from 0 to 8. The zero digit is the sign digit and the remaining digits comprise the actual number. Since there are 4000 memory positions, four decimal digits are required

to describe an address and sixteen decimal digits are required to describe four addresses. Consequently, each order consists of two numbers and is stored in two consecutive addresses in the memory. The remaining two digits of the order are used to denote the type of operation. The zero digits of the two numbers comprising an order contain this operation code. These digits are also used as the means of distinguishing between orders and numbers. Zero and nine are reserved for positive and negative number sign indication. The remaining values involving 64 possible combinations are operation codes. Thus a number must have a zero or a nine in the sign digit and an order may have any value but zero or nine in the sign digit.

Table 2. Operations and corresponding codes.

Operation Code	Operation
33	Read in
34	Read out
35	Transfer
36	Add using fixed decimal points
43	Add
44	Subtract
45	Multiply
46	Divide
53	Choice + or −
54	Choice zero or not zero
63	Move exponent to No. 2 address
65	Change address A
66	Change address B

In Table 2 are listed some of the more important operations and their corresponding codes. It will be noted that provision has been made to enable the machine to change its orders. In addition, it is possible to set up a problem in such a way that a single order can modify all the orders in the problem so as to call in new values of the variables.

Additional operations, such as raising to a power, extracting the root, finding the logarithm, etc., can be prepared as sub-sequences and stored in the memory for future use. With the addition of extra shifting registers, these operations may be performed entirely within the arithmetic unit.

The memory storage may be described as a huge matrix, the address being the key to

any element. The first two digits of the address control the row selection or the equivalent number time on the drum. The last two digits control column selection or the track designation.

The tape input-output component takes the tapes prepared by the tape-preparation unit and under direction of the control unit reads information into the computer memory. It also prepares magnetic tapes containing the answers to the problem (the output from the computer). Input and output tape data are the same; that is, data read out of the machine may be run back into the machine. A unit piece of information on the tape consists of a number plus its memory location. The tapes are operated at a speed such that approximately 25 pieces of information per second are read into or out of the machine. These tapes are utilized by the printer to type out the information. The reels of magnetic tape belonging to this unit are located in the center of the main console.

Tape preparation and printing are performed by units similar to those of the Harvard Mark III computer. The tape-preparation unit has a standard keyboard and can be used to prepare both number and sequence-control tapes.

The printer uses an electric typewriter, operating at 10 strokes per second. Provision has been made to vary the typography of the printed page at the discretion of the operator.

The main components of the computer are housed in five racks 24 in. wide and 8 ft tall. The Tape-Preparation Unit, Printer, Memory, and Power Supply are housed in separate units. The Magnetic Memory is contained in a cabinet 3 ft wide, 5 ft long and 6 ft high. It appears that the power supply will be contained in a similar cabinet. The Tape-Preparation and Printer Units will be housed in racks 24 in. wide and 6 ft tall.

Unitized construction is used throughout the computer. At present, there are 15 basic circuits. These circuits are constructed as plug-in assemblies such that the components are mounted on turrets between the tube socket and plug. These assemblies are approximately 1.5 in. in diameter and 2.5 in. long. The circuits plug into standard panels which contain the signal and power wiring. The panels are mounted vertically in the cabinets. A vertical sheet of Plexiglas mounts 2.5 in. in front of these panels. The circuits are plugged into the panels through the Plexiglas. Air, circulated in the channel formed by these two panels, is used to cool the circuits. The vacuum tubes mount directly in the sockets and project out of the Plexiglas panel.

One of the major achievements in the design of this computer has been the reduction of the tube complement to less than 1000 vacuum tubes. This reduction has been made possible through the use of germanium diodes and careful circuit design. Approximately 4000 diodes are used in the computer.

In summarizing, I would like to highlight the following points.

First: This computer has been designed to solve problems requiring engineering accuracy. For problems of this type, it must be reliable and accurate first; speed comes next.

Second: Design, operation, and maintenance have been simplified by the reduction of tube complement and unitized construction.

Third: The computer design is flexible; that is, individual units such as the Arithmetic Unit, Control Unit, Memory Unit, etc., stand by themselves. They can be readily modified with minor effect on the rest of the computer. The resulting building blocks which comprise this computer can be used to construct a computing machine for almost any purpose.

BANQUET

Tuesday, September 13, 1949

Toastmaster

Edward A. Weeks, Jr.
Editor of The Atlantic Monthly

TOAST BY D. H. LEHMER

There is a man among us here tonight who deserves our special vote of thanks and appreciation. He recognized the necessity for a medium of communication—"a standard source to which one might naturally turn for guidance in connection with all mathematical tables of importance in contemporary research." Through the National Research Council in 1943 he established the quarterly journal *Mathematical Tables and Other Aids to Computation.* Now after seven years of unflagging effort, Raymond C. Archibald is retiring as Editor of *MTAC.* I propose that we show him our appreciation for his excellent work.

TOAST BY SAMUEL H. CALDWELL

It is with deep regret that we note the absence from this banquet and from the Symposium of one of the world's great figures in the field of scientific computation. Some of you have known him as a teacher. Many of you met him and heard him at the Machine Computation Conference held at the Massachusetts Institute of Technology four years ago. All of us who have known Doctor L. J. Comrie have been stimulated by his appreciative response, impressed by his intellectual grasp, and conquered by his wit and charm.

As the founder of the Scientific Computing Service of London, and in his former connection with His Majesty's Naval Almanac Office, Doctor Comrie has been a prolific contributor to the literature of scientific computation. But history will name him also as one of the pioneers in the development and application of machine methods to computation problems.

Doctor Comrie is unable to be with us because of serious illness, and this I know is a matter of profound disappointment to him. It is proposed that we members of this Symposium stand at the side of Doctor Comrie in his fight for health and that we let him know it. I therefore ask that we request our Toastmaster to send to Doctor Comrie our prayers for quick and full recovery of his health and vigor, and our earnest hope that he can be with us at our next Symposium.

The Present Position of Automatic Computing Machine Development in England

W. S. ELLIOTT

Research Laboratories of Elliott Brothers (London) Limited

I have come from a place in England named Borehamwood. Borehamwood contributes both to the arts and to the sciences. A small part of its contribution to the sciences is work on what our popular press, unfortunately, in my view, calls "Electronic Brains." On the side of the arts a large part of the British Motion Picture Industry is at Borehamwood. Some of you may have heard that a certain William Shakespeare has been trying to earn dollars for his country by writing the screen plays "Henry V" and "Hamlet."

On my desk at Borehamwood, I have a volume which I prize very highly. It is a report of the proceedings of the first Symposium on large-scale digital calculating machines held here at Harvard in 1947. This is a book which, I think, contains much weighty and interesting material—material made no less significant by the advances of the last two and a half years. To me not the least interesting paper in this volume is that by Richard Babbage dealing with the work of his English grandfather, Charles Babbage, that first designer of computing machines. And when I read this paper, my attention focuses on one passage.

"Propose to any Englishman any principle or any instrument, however admirable, and you will observe that the whole effort of the English mind is directed to find a difficulty, a defect, or an impossibility in it. If you speak to him of a machine for peeling a potato, he will pronounce it impossible; if you peel a potato with it before his eyes, he will declare it useless because it will not slice a pineapple. Impart the same principle or show the same machine to an American . . . and you will observe that the whole effort of his mind is to find some new application of the principle, some new use for the instrument."

When Professor Aiken, just ten days ago, asked me to speak at this Symposium, my first thought was that I might take as a text the differences between English and American computing machines in the light of that passage. But when I came to think about it, I decided I could find no significant difference except perhaps that the groups developing our machines are a little smaller. Certainly the projects that we have are as diverse as those in this country, and the ways that the different groups go to work are similarly varied. For instance, the logical design of one machine was completed well before the team was set up to build it. Another machine grew as the ideas came. The first machine is more engineered, and the second machine is breadboard.

I shall mention seven groups in England working on computing-machine projects: three at Universities—the Universities of Cambridge, Manchester and London; three at Government establishments—the National Physical Laboratory (NPL), which I think corresponds

to your National Bureau of Standards, the Telecommunications Research Establishment (TRE) of the Ministry of Supply, and the Royal Aircraft Establishment; and I shall mention one industrial firm, Elliott Brothers (London) Limited, the Research Laboratories of which I represent. Of these groups, that at Cambridge has one machine fully operating. The Manchester group has a machine fully operating though with restricted input and output units. Other machines are in various stages of development. I shall describe the Cambridge

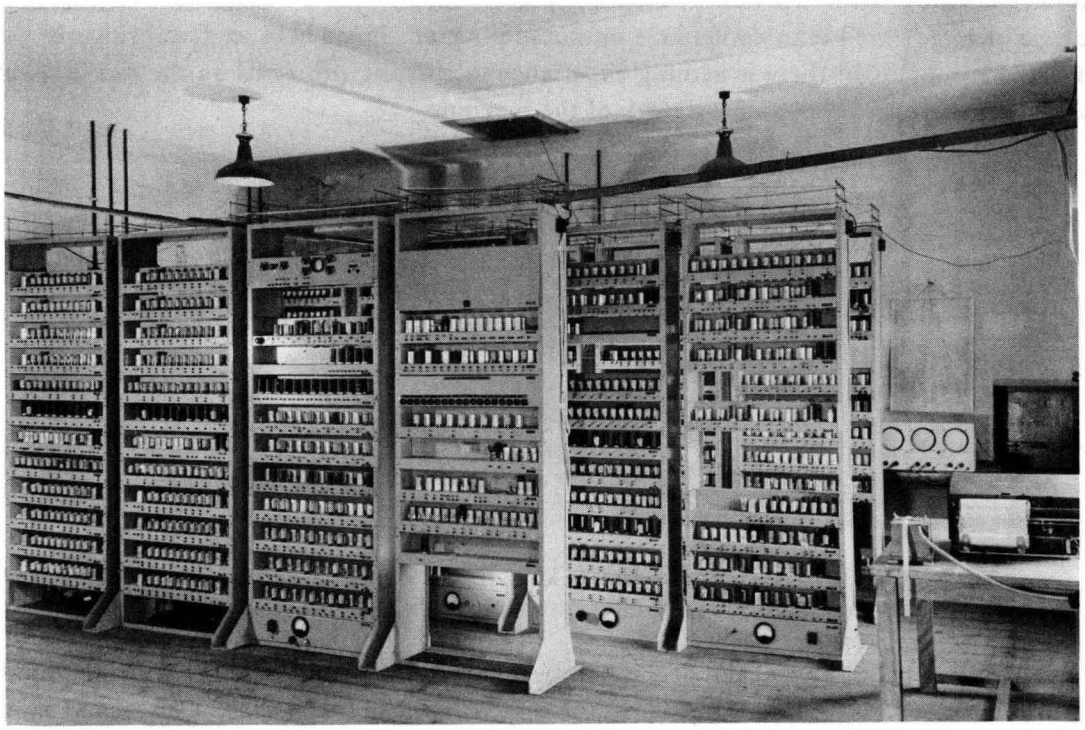


FIG. 1. The Electronic Storage Delay Automatic Calculator (EDSAC).

machine more fully and I shall compare other machines with it. I shall not give here precise figures for the memory capacity, speed, and so on.

The first electronic computer to run in England, the Electronic Delay Storage Automatic Calculator, or EDSAC, was designed and built by Mr. M. V. Wilkes, the Director of the Cambridge University Mathematical Laboratory, assisted by Mr. Renwick. Besides being a theoretical physicist Mr. Wilkes is a practical electronic engineer.

EDSAC (Fig. 1) was projected by Mr. Wilkes during a visit he made to the United States in 1946 when he attended part of a course on computing machines at the Moore School. The logical design of EDSAC was influenced by the ideas of Mauchly, Eckert, Goldstine, and Sharpless of the Moore School. At the outset Wilkes stated that he was not interested in building the best possible machine. He wanted to make a reliable machine and to make it quickly. He chose mercury delay-line storage as being the only principle which at that time

promised reliable storage. He chose a 500-kc/sec digit rate as being the fastest that, with the techniques then known, would give a reliable computer. The store capacity is 512 words (numbers or orders) of 32 binary digits. Input to the machine is by punched paper tapes prepared on a teleprinter keyboard, and output is directly printed on a typewriter. EDSAC uses a one-address code for instructions. The storage, control, and arithmetic units were designed in 1947, and they and the input and output units were built in 1948. Toward the and of that year parts of the machine were being tested, and the machine was fully operating and was demonstrated at a conference on computing machines held at Cambridge in June 1949. Today the team there is gaining experience in running problems on the machine, and I have with me two samples of the work of the machine.

FIG. 2. The Manchester machine. Control and input circuits are at the left, the memory in the right center, and the arithmetic units at the right.

The first sample is a list of the prime numbers up to 1021. The list starts with the number 5—Mr. Wilkes assured me that they know the prime numbers below that. The second sample is a tabulated solution of a second-order differential equation.

Before I leave the Cambridge group I should like to say that Wilkes is very active in holding fortnightly colloquia during University term time, and the different teams in England attend these colloquia very well and keep closely in touch.

In June of this year the colloquia culminated in a four-day conference. Descriptions were given at the conference of the various computing-machine projects, not only in England, but also in France, Holland and Sweden. A contribution from Doctor Huskey was read, giving an account of the present state in America. Discussion subjects included CRT storage, programming and coding, checking facilities, and permanent and semipermanent storage.

The second University group is at Manchester. The machine (Fig. 2) is being built by the Electrical Engineering Department under Professor F. C. Williams for the use of Doctor Newman and Doctor Turing of the Mathematical Department. There is close contact between the engineers and the mathematicians, but the machine is definitely being designed by the engineers. The machine uses the well-known CRT store of F. C. Williams and T. Kilburn.

This store features a standard cathode-ray tube and a physically simple mechanism. Experimental work on the store was completed about March 1948. Having built the store, Williams and Kilburn wanted to test it, so they added a second storage tube as an accumulator, and a third tube as control. They thus had a baby computing machine. The baby machine was of breadboard construction and today the machine at Manchester consists of these same breadboards and others that have been added. In fact, the machine has grown gradually as ideas came—unlike some machines, which have been built according to a master plan conceived at the outset. For this reason any description of the machine is liable to be outdated very quickly. At the time of the Cambridge conference, there were a fast multiplier CRT and a special tube for modifying instructions. The machine uses magnetic-drum auxiliary storage running at the rate of the working store. A feature of this is that the drum is synchronized from the machine's clock—the drum does not generate the clock. The drum stores true binary numbers and has access only to and from the main store.

Input to the Manchester machine is on a binary button board, and the output is a CRT display of the binary content of one of the CRT stores. The digit rate is limited by the CRT store to a quarter of that in EDSAC, but since the store is noncyclic the average access time in the two machines is similar. One CRT store has the capacity of four long tanks in EDSAC, that is to say, 32 words.

Because of the restricted input and output units on the Manchester machine the type of problem that can be run on it is rather limited but some interesting work has been done on the Mersenne numbers.

A second machine for the Manchester group is being built by Ferranti Limited. This is to be a more engineered version, and it will have 16 main CRT stores. The engineers consider four to eight to be the optimum number of main storage tubes in a computer of this type, having regard to transfer time from the auxiliary store to the main stores. The mathematicians would be content with eight tubes but in some cases would like 16, so to be on the safe side 16 main storage tubes have been decided on, in addition to an accumulating store, a store for control numbers and a "B" tube where instructions are modified. There is no proper name for the Manchester machine, though I understand from Professor Williams that it has a variety of improper names. The Manchester machine recently gave rise to some correspondence in the London *Times* on whether a machine could rival the brain of man. In an interview with the paper Doctor Turing said he did not exclude the possibility that the machine could write a sonnet. He added, however, that only another machine could appreciate the sonnet fully.

The third university group, at London, directed by Dr. A. D. Booth, is working on three machines. The first is called "Automatic Relay Calculator" or ARC. This is a binary relay machine which, in logical design, is somewhat similar to EDSAC and follows some of the ideas of the Moore School, in that, for example, numbers and orders are lumped together in the store, and orders can be modified. It is a parallel computer with a small magnetic-drum store. Input and output and semipermanent storage are all on punched paper tape. The machine was made by Doctor Booth for experiments in logical design. It has 800 relays and

cost about £2,500. The machine was being tested in June 1949. The magnetic drum is now being changed to another storage system, an electromechanical store, which Doctor Booth is developing and which is of some interest. In essence this is a very concentrated collection of small relays. It packs 256 numbers, each of 21 binary digits, into about 12 by 8 by 16 in.

Doctor Booth is also designing an electronic machine and is to make two models of it in parallel, for different uses. It is to have magnetic-drum storage, magnetic-tape input and output, and magnetic-tape auxiliary storage. It is to have multiplier and divider units, and Doctor Booth thinks it will have fewer than 1000 tubes. He gives no date for its completion. It may be one or two years.

Doctor Booth's third machine is a "Simple Electronic Computer" or SEC. This he proposes as the smallest electronic computer that will have all the main facilities of a general-purpose machine but will be as small as 181 tubes, and he hopes that University departments will be able to build it for themselves.

Turning now to the Government establishments, a considerable amount of work was done at the National Physical Laboratory in 1946 on the Automatic Computing Engine or ACE. This work was done under Doctor Turing and by the end of that year the quite complicated and sophisticated logical design was completed and several problems had been coded. In September 1947 an Electronics section was set up at the National Physical Laboratory to work on electronic computing machines, but before this team had started on the actual construction of the ACE Doctor Turing left. About the middle of 1948, it was decided that the theoretical team of the Mathematics Division, which was now under Mr. J. H. Wilkinson, should join the electronics group under Mr. Colebrook, and the two teams should work together on the construction, not of the full-scale machine, but of a Test Assembly. This Test Assembly represents an attempt to construct the smallest machine that will serve as an adequate testing ground for the concepts involved in the full-scale machine, but that will nevertheless be large enough to be a useful computer.

The TA is somewhat similar to EDSAC. It has, for example, delay-line storage. It works at twice the digit rate of the EDSAC. It has some logical orders other than those needed for arithmetic operations and uses the two-address code for instructions. Input and output are on Hollerith cards.

In EDSAC, instructions are stored serially in a long tank. This means that after obeying one instruction the machine has to wait for the remainder of a major cycle before the next instruction is available. In the TA this is overcome by facilities for putting instructions in nonserially and in such a way that when one is obeyed the next instruction is immediately available. The TA is being carefully engineered. About one-half or two-thirds of the chassis for it is completed, and Doctor Wilkinson hopes it will all be completed by the end of 1949 so that testing will start in 1950. It is not likely that the full machine of the 1946 design will be built now. Any further machine will probably have a much smaller number of mercury delay-line stores and auxiliary magnetic-drum storage.

At the Telecommunications Research Establishment of the Ministry of Supply, Dr. A. M.

Uttley is working on a parallel electronic machine for the use of mathematical physicists in the Ministry of Supply. His decision to make a parallel machine was taken after a visit to the United States in 1948 and was influenced by the fact that no one else in England at that time was building a parallel electronic machine. He uses storage tubes similar to those of F. C. Williams, working in the same way, and uses as many storage tubes as there are digits in his words. The tubes are scanned in parallel, and a word is represented by taking one digit from the corresponding position in each of the tubes. Like the Manchester machine, it uses magnetic-drum auxiliary storage, but unlike the Manchester machine the numbers on the drum are in binary-coded decimals, and there is direct access between the drum and the outside world. Doctor Uttley's idea is that the drum will be prepared at leisure by mathematicians; it will then be taken to the machine, and its contents transferred as a whole into the working store of the machine. In an attempt to make the machine completely self-checking, Doctor Uttley has developed a complete series of three-state circuits for the arithmetic and control units of the machine. The whole machine works in three states apart from the store, the states being nought, one, and fault. There is now a one-digit working model of the store and of all the three-state circuits, and the magnetic drum is completed, together with the tape puncher, and transfer from tape to drum and from drum to electromatic typewriter. Doctor Uttley thinks the whole machine will be working in one or two years.

Another small relay machine is being made at the Royal Aircraft Establishment by Doctor Hollingdale for the use of people in that establishment.

I come now to Elliott Brothers Research Laboratories. Our interest is in the development of reliable components such as storage, arithmetic, input and output units for high-speed machines. We are working on a CRT storage method similar to but not the same as that used by F. C. Williams at Manchester. In his paper F. C. Williams called the method of his that we use, "anticipation pulse storage." We find that we can use a higher writing speed than in the dot-dash method that is actually used in the Manchester machine though we have not decided the maximum number of digits that can be stored reliably on one tube. We are working on small logical units for serial operation at up to 1-Mc/sec digit rate and we have working a series-parallel multiplying unit, using these logical elements, that forms the rounded-off product of two numbers entering the unit simultaneously, the rounded-off product appearing in the following number time. In its final form this multiplier will feed the output straight back to one of the inputs so that n numbers can be multiplied together in n number times.

We are working on photographic methods of feeding input data and function tables into a high-speed computer and of recording the output from a computer. The input data and function tables are prepared by photographing a lamp display controlled from the register of a desk calculating machine working in binary scale, which we have made especially for this purpose. The film can be read at .1 megadigit per second into a computer.

No description of the English automatic computing machine projects is complete without mentioning the name of Professor Hartree, Plummer Professor of Mathematical Physics in

the University of Cambridge. The early work of Hartree and Porter on the differential analyzer at Manchester is well known, and today Professor Hartree plays a leading part in encouraging work in England on digital machines. He is a regular attendant at the Cambridge colloquia and is regarded as our chief contact with work in the United States.

In conclusion, I would say that the greatest diversity of opinion in England at the moment is on the best method of storage to use. In the Cambridge machine the component that gives the least trouble in the whole equipment is the mercury delay line. The F. C. Williams' store is welcomed enthusiastically by some groups in England, though others are unhappy about the noise level. Doctor Booth's electromechanical store is interesting in its simplicity and digit density, though it is limited in speed. There is general agreement in England on the use of magnetic-drum auxiliary storage. I think, however, that the greatest possibility of technical improvement or simplification is in storage systems.

The tendency in England at the moment is to gain experience with the small machines that have been built or are being built, and I think that after one or two years of gaining this experience, some further machines may be built. It is likely that when this happens the move will be in the direction of logically simpler rather than of larger machines.

Finally, I should like to return to the subject of human and mechanical brains. Professor Sir Geoffrey Jefferson of the department of neurosurgery of the University of Manchester gave the annual Lister Oration to the Royal College of Surgeons of London on this subject. He referred to the fact that some workers believe that by embodying in a machine the electrical principles underlying neural activity, light can be thrown on the way we think and remember. He did not think, however, that the day would dawn when the gracious rooms of the Royal Society of London would have to be converted into garages to house the new fellows.

THIRD SESSION

Wednesday, September 14, 1949
9:00 A.M. to 12:00 P.M.

RECENT DEVELOPMENTS IN COMPUTING MACHINERY

Presiding

E. Leon Chaffee
Harvard University

Semiautomatic Instruction on the Zephyr

HARRY D. HUSKEY

National Bureau of Standards, Institute for Numerical Analysis, UCLA

Presently designed calculators cannot be entirely automatic with respect to coding; they may do problem after problem automatically without human attention, but somebody must initially tell the machine what it is to do. We will develop in this paper a method of operation of such a computer in which the user need not tell in explicit detail everything that the calculator must do in the course of carrying out the computation. This concept of semiautomatic instruction has been called abbreviated-code instruction.[1]

To illustrate by example, assuming we wish to invert a matrix having m rows and n columns, the only essential information is (1) where the coefficients of the matrix are, (2) how many rows and columns there are, (3) what process is to be carried out, and (4) where the answers are to be placed. We expect to be able to obtain sheets of paper upon which appear in the appropriate order the coefficients of the inverted matrix without doing more than sending the initial coefficients into the calculator and a single coded instruction specifying the three items mentioned above.

The Zephyr, the electronic digital calculator under construction at the Institute for Numerical Analysis, will be used as the model in this paper to illustrate how these coded instructions will operate. Thus, before explaining the abbreviated code in detail a brief description of the Zephyr will be given.

The Zephyr consists of: (1) an arithmetic unit where the information is processed or modified; (2) a high-speed memory which remembers both the numerical and the instructional words needed during the computation; (3) a low-speed memory, which we shall refer to as the store,[2] inasmuch as it serves as a warehouse wherein numerical information, main routines of code words, and subroutines of commands or code words are stored; (4) a control unit which scans the memory for its commands, and executes them by sending out the appropriate signals to the other units; (5) input-output equipment which we will not discuss in this paper.

Information is stored and processed in the Zephyr in units that are 41 binary digits long. Such a unit is called a word. Words may be interpreted as numerical information or as instructions.

Numbers can be subclassified as follows. A word may represent a signed binary number lying somewhere between -2^{40} and $+2^{40}$. Or, it may represent a signed ten-decimal-digit number where each decimal digit is represented as a four-digit binary number. A floating binary representation may be used where one is dealing with numbers of the form $\pm a \times 2^b$. For example, the first digit represents the sign, the next ten binary digits may represent the exponent b, and the remaining 30 digits may represent the significant figures of the number

in binary form. In this manner, with some loss of relative accuracy, a word can represent numbers in the range between $\pm 2^{30} \times 2^{\pm 2^9}$.

In a similar manner instructions are subclassified into three classes. First, there are *command words* of which there are 13 in the Zephyr. A command word is a 41-binary-digit word, a portion of which determines one of the 13 operations, and the remainder of which, in general, specifies four addresses in the memory. A second class of instructions are the *control words*. Control words may serve as parameters that determine the number of repetitions of certain routines; they may be the bounds used to stop certain computational processes; or they may serve as factors in logical products or extraction operations. The third class of instructions are called *code words*. A code word is a compact representation in one word of several parameters that are needed to specify the operation of subroutines. Each subroutine extracts its various parameters from this one code word. Thus, one may specify a scalar multiplication with only one code word. The appropriate subroutine in the calculator extracts and properly places the various parameters from the code word. These parameters must specify the common factor (that is, specify its address in the high-speed memory), the location of the elements of the vector (say by specifying the address of the first element and the number of terms in the vector), and where the result is to be placed.

We can summarize the various types of instruction as follows. A *command word* is a 41-binary-digit word which the calculator explicitly understands and obeys. A *control word* is not directly obeyed by the calculator nor is it a direct part of the calculation; it in some way controls the course of the computation or enters into the arithmetic-like operations that are performed upon command words. A *code word* is an abbreviated instruction that specifies in one word a whole sequence of events for the calculator.

The high-speed memory will consist of a bank of cathode-ray tubes used in a manner devised by F. C. Williams, of Manchester University, England.

This memory will have a capacity of 512 41-binary-digit words and will be able to deliver any one of its words to the other units of the machine in about 20 μsec.

The high-speed memory will be divided into three parts: first, a part that stores the numerical information temporarily; second, a part that stores the subroutines which are to be used in the problem; third, a part that stores that portion of the main routine which must be stored in the high-speed memory. As the main routine is carried out new segments must be read in, and in the course of doing the problem numerical information must be transferred to and from the magnetic drum. If all the necessary subroutines cannot be stored in the high-speed memory at once these, too, must be read in and out during the course of the computation.

The store, or low-speed memory, will consist of a magnetic drum with a capacity of 10,000 words of standard 41-binary-digit length. It will have a multiplicity of reading and recording heads so arranged that all the 41 digits of a particular word will appear simultaneously at 41 different magnetic heads. Thus, the access time for a word on the drum depends upon the orientation of the drum when the number is called for, and will vary from a few

micro-seconds to a maximum of 16,000 μsec (the time it will take for the drum to make a complete revolution).

In similar fashion to the high-speed memory the magnetic-drum storage will be divided into three parts. One part will store the numerical information needed to do the respective problems. A second part will store all the standard routines, such as division, floating operations, etc. A third part will store the commands or coded instructions of the main routine. In our present experience the number of commands per routine seems to average around 30. Thus, we could store 100 different standard routines on the drum and only take up 3000 words of storage. Most problems should involve only a few hundred instructions, say not more than 1000. This leaves approximately 6000 words, which, for example, is ample room for storing all the numerical information involved in solving 70 simultaneous linear equations.

A command word may be represented in the form α, β, γ, δ, F; α, β, γ, and δ, generally, represent addresses or the position of words in the memory, while F determines which one of the 13 commands is involved.

In normal situations the next command is specified by a fifth address, called ε, which is remembered by a binary counter in the control unit. Each time a command is obeyed the number in ε is increased by unity; thus, the machine normally obeys a sequence of commands coming from successive addresses in the memory.

There are three special commands wherein the next command is determined by the fourth address, δ, of the present command. By the use of these special commands the machine may transfer source of control with each command. When operating in this manner the machine may obey any arbitrary pattern of commands in the memory. Naturally, the special commands may be interspersed in any desired manner among the other commands.

In addition and subtraction operations the capacity of the calculator may be exceeded. In case this happens the normal commands behave exactly like the special commands; that is, the next command is determined by the fourth address δ.

In order to explain efficiently the 13 commands, let us introduce the following notation. Let $w(\alpha)$ denote the word stored in address α. Let the symbol $\rightarrow$ be read as "replaces." Let NC $= w(\delta)$ mean that the next command the calculator is to obey is the word in address δ of the memory. The 13 commands, their symbols and effects, and the next command are given in Table 1.

Two principles have been followed in deciding upon the system of commands. The first is that there should be as few commands as possible so as to simplify the electronic circuitry. (Actually, the electronic function table which interprets these commands has only eight positions.) The second principle is that the commands should be as general as possible. For example, the Extract Command (logical product) allows the use of any factor whatsoever, and the elections in case of overflow are completely general.

One should notice that there is no Transfer of Control Command; the special commands do this automatically. Also, there is no Halt Command; the Input Command with δ specifying a nonexistent input device causes the machine to stop. Division is accomplished by a routine.

Table 1. Commands, symbols, effects, and next commands.

Command	Symbol	Effect	Next Command
Addition	$\alpha, \beta, \gamma, \delta, A$	$w(\alpha) + w(\beta) \rightarrow w(\gamma)$	$w(\varepsilon)$; $w(\delta)$ if overflow
Special Addition	$\alpha, \beta, \gamma, \delta, A_1$	$w(\alpha) + w(\beta) \rightarrow w(\gamma)$	$w(\delta)$
Subtraction	$\alpha, \beta, \gamma, \delta, S$	$w(\alpha) - w(\beta) \rightarrow w(\gamma)$	$w(\varepsilon)$; $w(\delta)$ if overflow
Special Subtraction	$\alpha, \beta, \gamma, \delta, S_1$	$w(\alpha) - w(\beta) \rightarrow w(\gamma)$	$w(\delta)$
Multiplication with Round-Off	$\alpha, \beta, \gamma, \delta, M$	$w(\alpha) \cdot w(\beta)$ rounded off to 40 digits and sign $\rightarrow w(\gamma)$	$w(\varepsilon)$
Special Multiplication with Round-off	$\alpha, \beta, \gamma, \delta, M_1$	$w(\alpha) \cdot w(\beta)$ rounded off to 40 digits and sign $\rightarrow w(\gamma)$	$w(\delta)$
Exact Multiplication	$\alpha, \beta, \gamma, \delta, P$	$w(\alpha) \cdot w(\beta) \rightarrow w(\gamma)$ and $w(\delta)$	$w(\varepsilon)$
Compare	$\alpha, \beta, \gamma, \delta, C$	Causes change in source of command	$w(\varepsilon)$ if $w(\alpha) < w(\beta)$; $w(\delta)$ if $w(\alpha) \geqslant w(\beta)$
Special Compare	$\alpha, \beta, \gamma, \delta, C_1$	Causes change in source of command	$w(\varepsilon)$ if $\lvert w(\alpha)\rvert < \lvert w(\beta)\rvert$; $w(\delta)$ if $\lvert w(\alpha)\rvert \geqslant \lvert w(\beta)\rvert$
Extract	$\alpha, \beta, \gamma, \delta, E$	$w(\beta)$ is blanked (made into zeros) wherever there are ones in $w(\alpha)$, the result is shifted right or left a certain amount as determined by δ, the final result $\rightarrow w(\gamma)$	$w(\varepsilon)$
Input	$\alpha, \beta, \gamma, \delta, I$	Information is transferred from an input device determined by δ to the address α in the memory	$w(\varepsilon)$
Special Input*	$\alpha, \beta, \gamma, \delta, I_1$	Incoming information goes to $w(\varepsilon)$ instead of $w(\alpha)$	$w(\varepsilon)$
Output	$\alpha, \beta, \gamma, \delta, 0$	Information is transferred from address α of the memory to the appropriate piece of output equipment as determined by δ	$w(\varepsilon)$

In the case of the input and output commands δ may specify that the transfer is between the memory and the magnetic drum. In this event γ and part of β determine the address on the drum.

* This command is particularly useful in the process of initial input (that is, the process of reading-in information when there are no commands in the high-speed memory).

It takes nine digits to specify an address in the memory. Thus, in the standard 41-binary-digit words there are five digits left after one accounts for the four addresses. One of the five

digits is used for checking purposes to hasten the detection of any error caused by the calculator trying to obey ordinary numbers as commands. Three of the digits define the eight distinct commands described earlier. The remaining digit defines modifications of five of the eight commands, referred to in the table as the special commands.

Operations more complicated or more elaborate than those described in the discussion of the 13 commands must be done by a sequence of commands called a routine or subroutine. For example, division can be done by repeated subtraction if the appropriate routine is used. The whole process of division, which may amount to 100 operations, can be completely determined by approximately 15 commands. Furthermore, various commands used in the routine are repeated over and over again, operating each time on different numbers.

We cannot go into details of routines at this time. Suffice it to say that the subject is a very interesting one and that there are many pitfalls for the unwary; for example, has "division by zero" been taken into account?

In a general-purpose computer there are many relatively simple operations that we want the calculator to carry out. For example, we want the calculator to perform division, floating addition[3] and subtraction, floating multiplication and division, store-to-memory sequence transfers, and many other routines. Each such routine can be represented by a single word.

Table 2. Storage of control words and interpretation routine.

Address	Number α β γ δ F	Remarks
1	0	(= 000...00)
2	1's, 1's, 1's, 0's, 1's	Used to extract the δ portion of a word
3	1's, 1's, 1's, 1's, 0's	Used to extract F portion
4	0's, 1, 0's, 0's, 0's	Used to increase β addresses by unity
5	0's, 0's, 0's, 1's, 1's	Used to extract α, β, and γ portions
6	1, 0's, 1, 0's, 0's	Used to simultaneously increase α and γ addresses by unity
200	——	Address 200 shall be used to store the present coded instruction upon which the calculator is operating. This address plays a role analogous to the control register which registers a command word while the machine is obeying it.
201 to 205	——	The five addresses following 200 contain an interpretation routine that keeps track of the coded instruction we are presently obeying, and provides a method of entering the proper subroutine. In this system there are no general "links" (transfer of control instructions) to tell the machine what to do when it finishes the present subroutine; when each subroutine is finished the control always returns to this interpretation routine

If one were to try to build in circuitry to enable the calculator to perform all these tasks it is clear that the machine would become so cumbersome from the circuitry point of view that it would be almost impossible to construct, and, very likely, impossible to maintain in operation. However, we have seen that a routine of standard instructions can be set up in the memory whose effect will be to carry out such operations as those described above.

Before considering in more detail an example of an abbreviated code instruction we will look into the storage of control words and examine the interpretation routine. We shall assume that control words and the interpretation routine are stored as indicated in Table 2.

Each abbreviated command will be stored in address 200 while it is being obeyed. A portion of it (analogous to the function in the command word) is extracted and added to a dummy command to arrange for an entry to the subroutine. The first step in the interpretation routine is to read the code word from a general place in the memory into a fixed place, address 200. Next the extraction and addition with the dummy command takes place. This dummy command must be carried along to allow the command to be used over and over (the old command which provided the last entry remains in the subroutine until such time as it is replaced).

The new entry shifts the source of commands into the subroutine. Each subroutine begins with certain extraction and addition commands that split the parameters off the code word in address 200 and add them into the appropriate blanks in the routine. The last command to be obeyed in the subroutine refers the control to address 201 for its next instruction. This address in turn specifies what coded instruction is to be transferred to address 200.

If we assume that fifty such abbreviated commands can be stored in the high-speed memory and still leave sufficient room for the arithmetic data and the appropriate subroutines, then either we must arrange that the interpretation routine counts and causes segments of the main routine to be read in, or every fiftieth command must read in fifty new coded commands from the main routine of the problem.

Consider the coded instruction

20, 18, 60, VC, 19.

This means that a constant in address 18 is to be multiplied by a vector of order 19 stored in addresses 20, 21, $\cdots$, 38 and the resultant vector is to be stored in addresses 60, 61, $\cdots$, 78. Let us also assume that the above vector-constant multiplication coded instruction is stored in address 225. The sequence of commands is given in Table 3.

The vector-constant multiplication routine is chosen as an example since it clearly can be considered as a unit in a higher-level program (for example, solution of simultaneous linear equations) and it may itself control subroutines. For instance, the numbers may be stored in floating-binary form and the command in address 306 would have to be replaced by a coded instruction calling for a floating-multiplication routine.

We can imagine much more elaborate situations in which the main routine is given as a sequence of coded instructions. Each of these coded instructions calls for a routine that is in

Table 3. Sequence of commands for multiplication of a vector by a constant.

Next Command is Determined By		The Command Is	Remarks
$w(201)$	Last command of preceding routine	1,225,200, 0,A	Takes (20,18,60,VC,19) to address 200
$w(202)$	ε	2,200,204, 0,E	The "VC" [= 300] is extracted into address 204 to use as an entry to the vector-constant routine
$w(203)$	ε	205,204,204, 0,A	A "dummy" command in 205 has the extracted VC added to it [$w(205)$ = 201,4,201,0,A_1]
$w(204)$	ε	201,4,201, 300,A_1	Adds "1" to the "225" in address b of $w(201)$ to provide for obtaining the next coded instruction
	(The interpretation routine is now finished and we are about to enter the vector-constant routine at address 300)		
$w(300)$	$\delta(204)$ [the δ address of $w(204)$]	5,200,204, 0,E	Extract the "20,18,60" of the VC instruction into address 204. Note that using 204 does not harm the interpretation routine
$w(301)$	ε	204,302,306, 303,A_1	Extractee is added to dummy in 302 to produce the first multiplication command [$w(302)$ = 0,0,0,0,M]
$w(303)$	$\delta(301)$	3,200,204,1–32,E	"19" extracted from $w(200)$ and shifted left into the α position in address 204
$w(304)$	ε	204,306,204, 0,A	A bound is produced in address 204 to tell when to stop this multiplication process [$w(204)$ = 39,18,60,0,M]
$w(305)$	ε	306,204, 0, 201,C	The process will be complete when $w(306)$ = 39,18,79,0,M and the source of command will shift to address 201
$w(306)$	ε	20, 18, 60, 0,M	First product is done
$w(307)$	ε	306,6,306, 305,A_1	Certain addresses of $w(306)$ are increased by unity and the calculator turns to the command in address 305

turn made up of coded instructions, and so forth, until finally one reaches subroutines whose elements are commands that the calculator explicitly obeys.

One approach to the problem of keeping track of position as one drops from one hierarchy of routine to another is by a process called *reversion storage*.[4] In this method a so-called queue

is established which stores in reverse order the addresses one must return to after completion of the respective subroutines in order to proceed with the problem.

Our approach to this problem has been to classify all routines into various orders. First-order routines are made of units which are explicit commands that the calculator obeys. There is no need for an interpretation routine for these routines since the ε counter keeps track of position here. Second-order routines are those whose elements are first-order routines. Third and higher orders are similarly defined. For each level a different interpretation routine must be used. Not only this, but one may need to use a certain coded instruction representing a particular first-order routine as part of, say, a second-order routine and a third-order routine. Therefore, a record must be kept of the level from which the entry was made to each subroutine. When that routine is finished the source of command will revert to the appropriate place in the correct level. Thus, we see that the interpretation routine divides into several parts (one for each order of routines used after the first) with a *record section* that stores the level from which the entry to subroutines is made.

In a sense we have made the situation two-dimensional. There are discrete levels on which we may operate with the record portion of the general interpretation routine controlling the choice of levels.

Success of a system like this simplifies coding by putting more of the responsibility for routine operations upon the calculator.

NOTES

1. The term "abbreviated-code instruction" was developed at the National Physical Laboratory, England, in a group headed by Dr. A. M. Turing of which the author was a member.

2. The term "store" was used by Charles Babbage of England, who is credited with being the first to design an automatic calculator. In England the term "store" is commonly used when referring to the "memory." Note that in this paper its use is restricted to the low-speed memory.

3. When two numbers of the form $a \times 2^b$ are to be added, one must be shifted until they have the same exponent.

4. This approach was developed at NPL. See note 1.

Static Magnetic Delay Lines

WAY DONG WOO

Harvard University

The magnetic delay line is a storage device which is built of rectifiers and transformers with cores made of ferromagnetic material that has nearly rectangular hysteresis curves. As shown in Fig. 1, a binary "1" is stored in a magnetic core as a residual magnetism in one direction, while a binary "0" is stored as a residual magnetism in the opposite direction. The difference

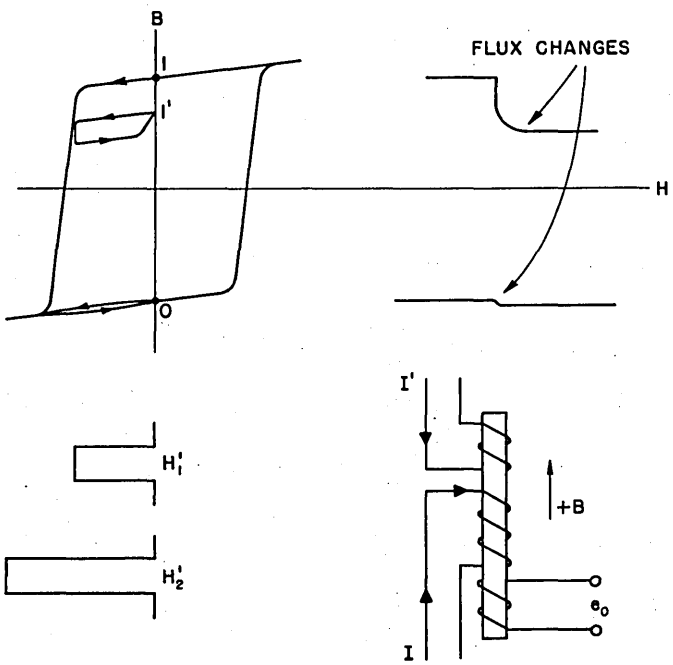


Fig. 1. Paths of operating point.

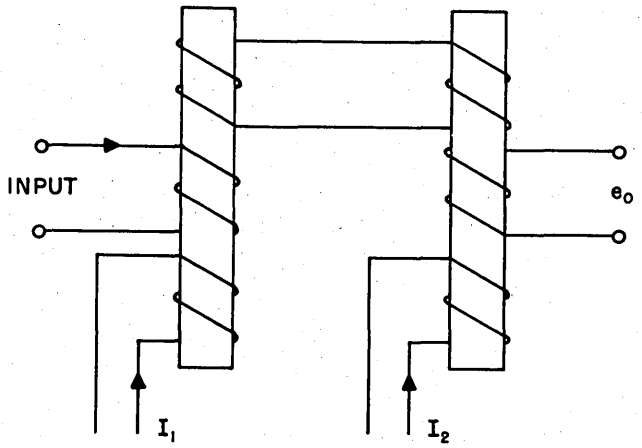
between this storage device and the conventional rotating magnetic drum or tape is that the storage medium is not moving. The information is recorded in discrete cores instead of on small spots in a continuous medium.

To record a binary "1" a positive magnetizing current is applied and to record a binary "0" a negative magnetizing current is applied. After the magnetizing-current pulse is over, the information will be preserved until another magnetizing force passes through the arc.

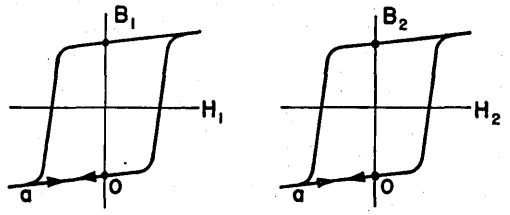
In order to read out the information without mechanical motion, it is necessary to apply a probing magnetizing force H', which is obtained when I' is applied. If the digit is a binary "1," then a large flux change occurs and a large induced voltage e_0 is obtained at the output

winding. If the digit is a binary "0," little voltage is induced. Thus the digit stored is indicated by the magnitude of the induced voltage when a probing current is applied.

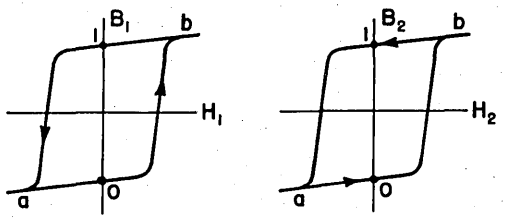
The residual magnetism remains essentially the same before and after a sufficiently small probing current. However, application of another probing current of the same magnitude



(a)



(b) PATHS OF OPERATION WHEN DIGIT STORED IS 0.



(c) PATHS OF OPERATION WHEN DIGIT STORED IS 1.

FIG. 2. Diagrams to show basic operation.

will produce a very small change of flux no matter whether the original state of the core is "1" or "0." After a small H', repeated application of H' will only describe the minor hysteresis loop shown at 1'. Thus, after the information is read out once, it can be considered destroyed unless one resorts to increasingly large probing currents.

If a probing current large enough to reverse the saturation is applied, a very large induced voltage results. It is so large as to be able to reverse saturation of another core of identical

construction. Referring to Fig. 2(*a*), if both cores were saturated in the negative direction originally, repeated application of I_1 and I_2 will not change the saturation of either core, as shown in Fig. 2(*b*), and little voltage is induced at the output winding. One can consider this as a "0" stored in this pair of cores. If, however, core number 1 is positively saturated originally, application of I_1 will cause flux Φ_1 to change from positive saturation to negative saturation. The voltage induced in the link winding will produce a current that opposes the effect of I_1. This current causes the flux in the second core to go to positive saturation even if it was originally at negative saturation. Now, if I_2 is applied to the second core, the flux in this core will go back to the state of "0" while that of the first core will go to the state of "1." Alternate application of I_1 and I_2 will result in an exchange of "1" from one core

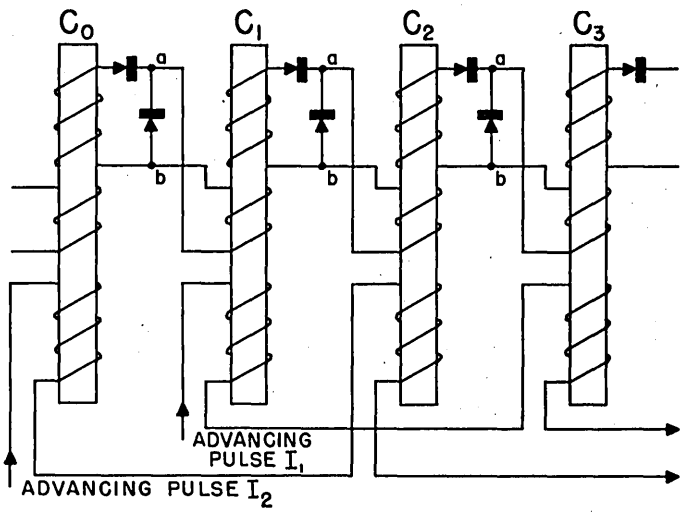


FIG. 3. Circuit diagram of magnetic delay line.

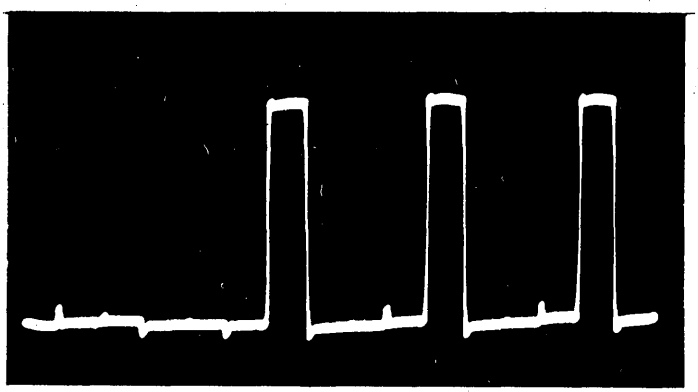
to the other, and there is an induced voltage at the output winding on every I_1 and I_2 pulse. Thus a digit "1" is stored in this pair of cores.

From the basic mode of operation, a number of cores are connected as shown in Fig. 3. The coils are wound so that the advancing current pulses produce a negative saturation corresponding to "0." The series rectifiers in the link winding are such as to stop any current in the link windings that would produce a negative flux. Consider now the case of the cores C_1 and C_2 having negative saturation. Then application of advancing current pulse I_1 will have no effect at all, and both cores retain their "0." One can also consider this as a "0" having been passed on from C_1 to C_2. On the other hand, if C_1 is positively saturated but C_2 is negatively saturated, when advancing current pulse I_1 is applied, C_1 will be saturated negatively, and the current in the circuit linking C_1 and C_2 will saturate the latter positively. Core C_1 returns to "0," while C_2 takes on the "1" that was originally in C_1.

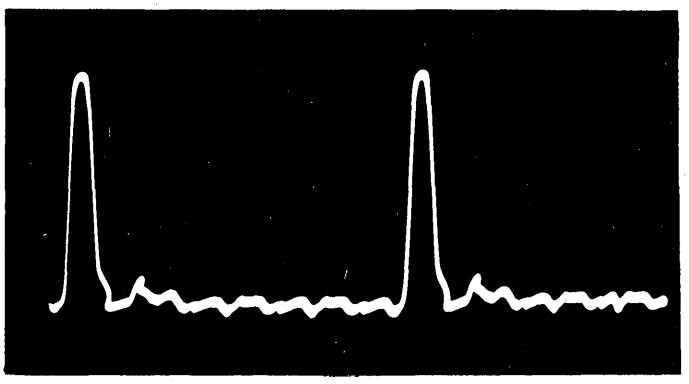
The rectifiers in the circuit prevent the effect of changes of flux in other cores on the two cores considered. The shunt rectifiers will prevent positive linking current in core C_0 when C_1

reverses saturation from "1" to "0" and produces the driving voltage so that the "1" does not go in the backward direction. However, it will have no effect when the driving voltage is from the C_0, because in this case the point a is at higher potential than b, while in the former case, a is at a lower potential than b.

The series rectifiers prevent the effect on C_3 when a "1" is advanced from C_1 into C_2. As



(*a*)



(*b*)

FIG. 4. Flux in a given core as a function of time:
(*a*) information rate, 3 kc/sec; information = 0111;
(*b*) information rate, 30 kc/sec; information = 1000.

the flux in C_2 goes positive, the voltage induced in the link winding to C_3 is such as to produce current causing negative flux in both C_2 and C_3. This current is prevented by the series rectifier. Aside from isolating C_3 from C_1 this rectifier also makes change of flux in C_2 from negative to positive easier.

Since each core when pulsed advances its stored digit only to the next core and has no effect on any other core, it is possible to advance a digit from every other core at the same time. Thus the advancing current windings of every other core are connected in series. The advancing current pulse I_1 will step the digits in all odd cores to the even cores, and the

advancing current pulse I_2 will step all the digits in the even cores to the odd cores. A pair of the alternate pulses will cause the digit to step two cores, which are considered as one unit of storage.

It is obvious that material having a nearly rectangular hysteresis loop, high retentivity, and low coercive force is required. The cores are made of wound strips of Deltamax (manu-

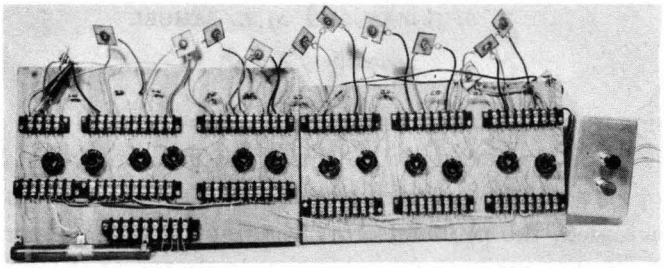


FIG. 5. High-speed magnetic delay line using selenium rectifiers.

factured by the Allegheny-Ludlum Steel Corporation) of about four convolutions. The diameter is $\frac{1}{2}$ in. and the strip is $\frac{1}{8}$ in. wide and 0.001 in. thick.

At present the maximum speed is 30,000 digits (i.e., 30,000 digits can be stepped through each unit of storage) per second. There is no lower limit of speed. The system acts like a system of trigger pairs, where digits are stepped from one trigger pair to the next. The fact that the speed is entirely controlled by the rate of advancing current pulses makes it a very useful intermediate storage system between two systems of widely different information rates.

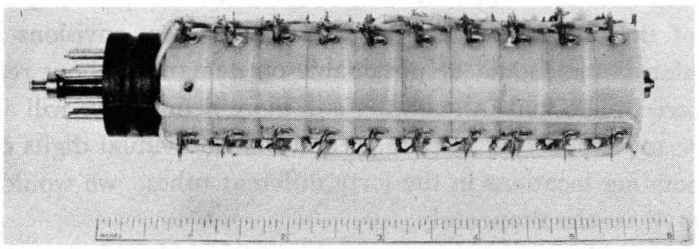


FIG. 6. Five-digit magnetic delay line.

Figure 4 shows the flux in a given core as a function of time when the information rate is (*a*) 3000 and (*b*) 30,000 digits per second. Figure 5 shows a ten-digit magnetic delay line on a breadboard. Figure 6 shows a five-digit line mounted on an octal plug.

Professor Howard H. Aiken, the director of the Harvard Computation Laboratory, proposed this form of storage device, and Dr. An Wang has done most of the work to make it successful. Special acknowledgment is due the Allegheny-Ludlum Steel Corporation, which has coöperated actively with the Computation Laboratory and has supplied all the core material.

Coördinate Tubes for use with Electrostatic Storage Tubes

R. S. Julian and A. L. Samuel

University of Illinois

One of the basic problems in connection with high-speed digital computers is that of storing the necessary amount of information which must be available at high speed as needed in the course of computation. As the speed of computing systems increases and as the amount of storage is also increased, the problem of locating any desired information becomes more acute. With this in mind, a research program was instituted at the University of Illinois to develop precise and rapid methods of locating stored information. Recent developments in storage systems in which continual memory refreshing is employed have somewhat reduced the long-term stability requirements, at least for these cases, so that the system to be described may not be needed. However, the system does possess a number of unique features which were thought to be of general interest and to warrant description. When and if storage systems progress to the point that a very much larger number of digits—say 10^6—can be stored in one electrostatic storage tube, then the need for precise locating equipment will again be urgent regardless of the type of refreshing used and the present scheme may warrant investigation.

We will assume that the information is stored in a binary code on the surface of target plates in tubes of the cathode-ray type and that the system envisions the requirement that any one bit of information should be obtainable on demand without regard to its particular location on the screen. To make the matter still more definite, we will assume that a bank of 40 such tubes are to be operated in parallel and that individual digits of a 40-digit code are stored in corresponding locations in the forty different tubes; we would then like to be able to locate these 40 digits simultaneously.

To make this possible we propose to combine these 40 cathode-ray tubes together with two special coördinate tubes into a master-slave relationship in which all of the tubes are connected to the same power supply with their corresponding deflection plates all tied together. There will then be a one-to-one correspondence between spot positions in the different tubes, although distortions may occur in the mapping from one tube to another as the result of minor differences and imperfections in the tube structures. If now some independent means is provided for precisely identifying specific spots on the screen of one tube, which then acts as the master, such that the beam of this tube can be returned to these spots with certainty when desired, the beams of all the other tubes will be returned to the corresponding spots in these tubes quite independently of any distortions that may occur in the different tubes. This will be true for a group of tubes that are structurally quite different as long as the other

voltages on the tubes are maintained constant. If the tubes are reasonably similar in their geometrical construction, it will still be true to a high degree of accuracy when these other voltages are allowed to vary within the usual engineering limits. It is only necessary, therefore, to introduce auxiliary beam-locating equipment into the master tube in order to control the motion of the spots in all the slave tubes to the desired accuracy. Furthermore, it is quite feasible to control the horizontal motions of the beams in the slave tubes with one master tube while at the same time controlling the vertical motion of the beams with a second master tube. The master tubes need not contain provisions for storage and they can be specialized to the necessary extent to perform their control functions, all the while preserving their essential similarity to the slave tubes in regard to one of their deflecting systems so as to retain the desired mapping characteristics.

Models of two quite different types of master tube have been constructed. Tests made on these tubes will be described later. Both types of tube are similar to the extent that they provide a definite number of stable beam positions (in this case 32) by means of mechanical positions on target plates contained within the master tubes. The beam position is maintained by means of servo amplifiers which obtain their input signals from the beam currents associated with the target plates.

The basic principle underlying the control system can be illustrated by reference to Fig. 1, which shows a system in which there is but one stable position. A single-stage amplifier is used in the illustration to simplify the discussion. Assume that the beam of the tube has been deflected so as to strike the top portion of the second plate. All the current of the beam will then be to this electrode. The voltage produced by this current in the grid resistor of the amplifier tube (augmented by a d.c. grid bias) will cause the control tube to be biased nearly to cutoff. As the result the plate current will be small, and the tap on the plate power supply will be set at a value that will cause the beam to be deflected downward. Alternatively, if the spot had been deflected downward so as to strike the interceptor plate only, there would be no current in the grid resistor, with the result that there would exist an appreciable plate current. With proper adjustments the resulting negative voltage across the plate resistor will exceed the positive bias on the vertical deflecting vane and the beam will be deflected upward. Obviously there exists but one stable equilibrium position in which the division of beam current between the target and the interceptor is such as to produce no net deflecting effect. If the amplifier circuit is properly designed to prevent hunting, any deviation of the beam from this equilibrium position will bring into play the necessary restoring forces to return the beam to the desired location. On the basis of this scheme, it is a relatively simple matter to visualize the interceptor electrode in any one of several forms such that there may exist a multiplicity of stable positions separated by regions of instability. Given a scheme for stepping the beam from one stable location to another, the necessary elements for the master tubes are evident.

While it would be possible to step the spot from a starting position to any desired ultimate

position by a series of equal steps, economy-of-time considerations suggest the desirability of utilizing steps of different sizes corresponding to the particular number system used in specifying the address (which in this case is binary). If information defining a desired memory location is supplied to the tubes in positional-notation form, then two possibilities exist: either this

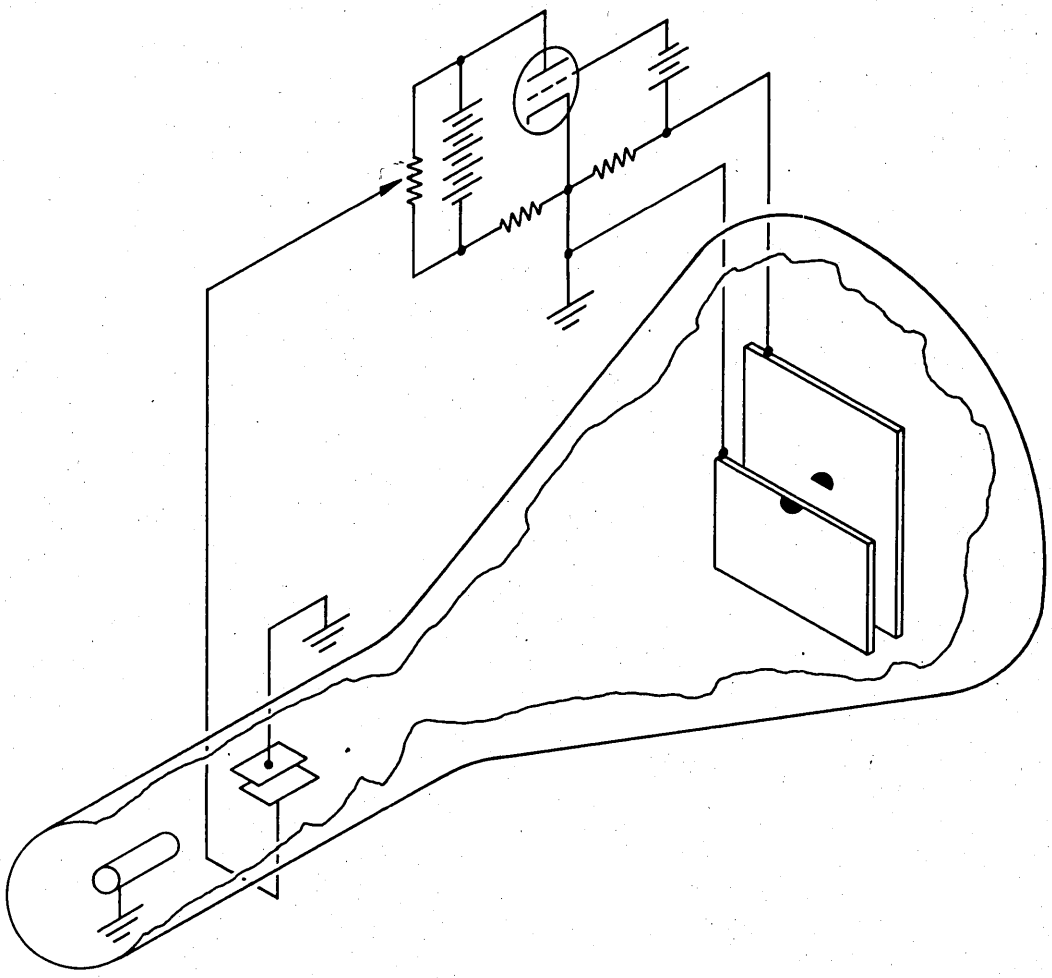


FIG. 1. Principle of stabilization.

information can be supplied in timed sequence or it can be supplied simultaneously. These two alternatives have resulted in the development of two quite different types of coördinate tube.

A simplified form of the serial coördinate tube is shown diagrammatically in Fig. 2 (in this case for only 8 stable positions). A comb-shaped interceptor is used in which the slots between the teeth are cut to different depths. The vertical position of the beam in this tube (which will be assumed to be acting as the master governing the vertical motion of the beams

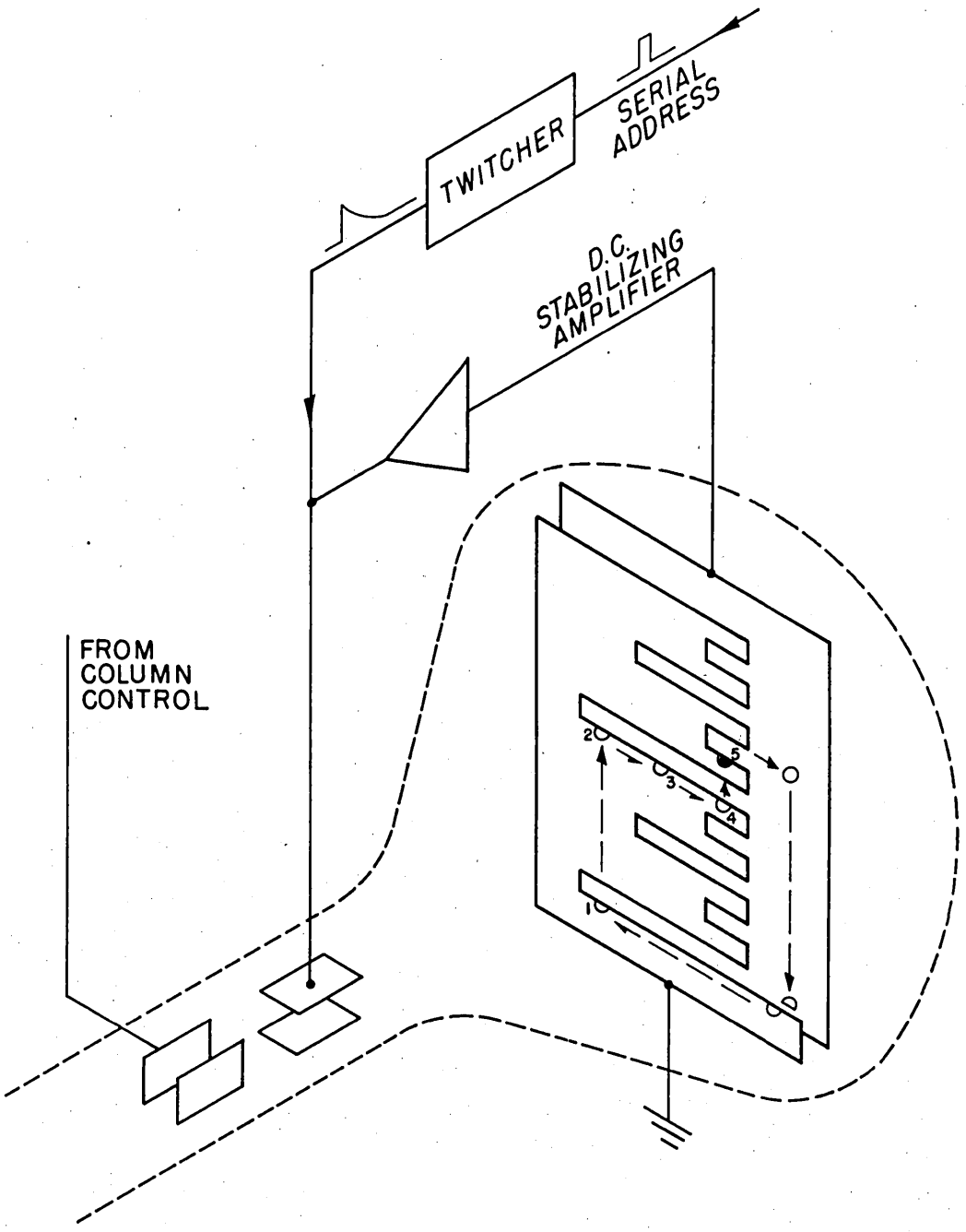


FIG. 2. The serial coördinate tube.

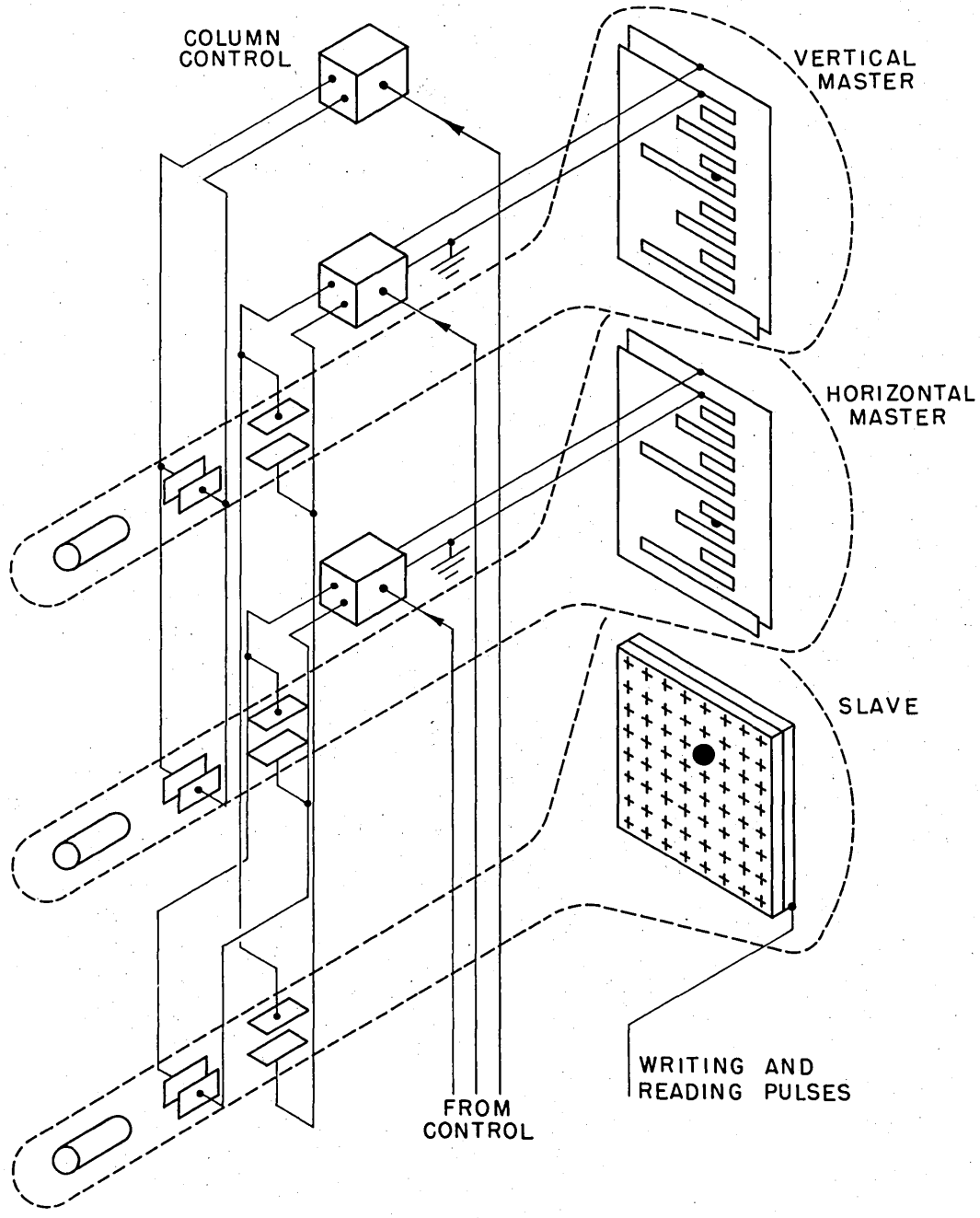


FIG. 3. The storage locating system shown with a single slave tube.

in the storage tubes) is controlled by the servo amplifier, the polarity being such that the beam is stable when partially intercepted by the top of any one tooth. The horizontal position of the beam in the master tube (i.e. the one controlling the vertical motion of the beam in the storage tubes) is, however, subject to independent control, there being three column positions corresponding to different columns in the binary-notational number system. The beam will be assumed to be initially deflected to the position labeled 1 in the figure. This corresponds to the starting position before an address has been located. We will assume that the desired address is the fifth slot, which in binary notation is 101, and that digits corresponding to this binary number are transmitted to the tube in timed sequence with provision for stepping the beam from column to column between these periods. With the beam in position 1, the first digit of the address is supplied through the circuit designated at the twitcher. This supplies

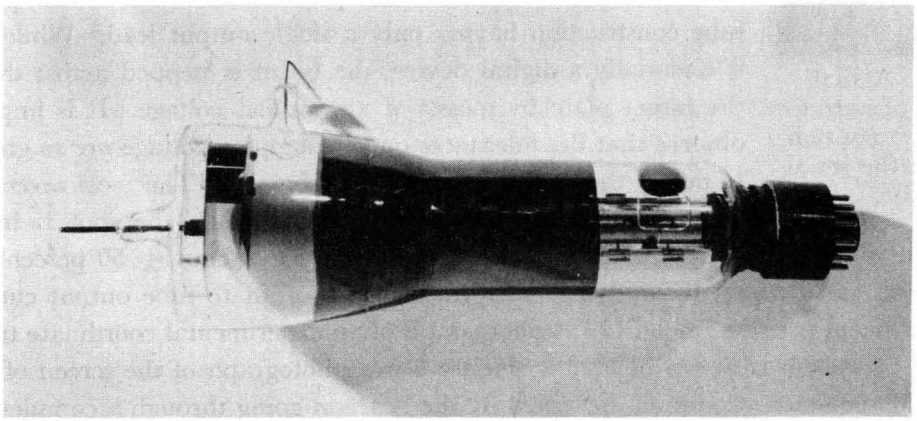


FIG. 4. Photograph of the serial tube.

to the deflecting plates a step voltage which transports the spot upward beyond the first slot so suddenly that the servo amplifier is practically not effective. The amplifier then continues to deflect the beam upward until it comes to rest at the next stable position, labeled 2 in the diagram. The column-control circuit then deflects the beam to position 3, where it is ready to receive the second digit of the address; in the present case this digit is zero, so that no signal is supplied from the twitcher circuit and the beam remains in position 3. The horizontal position of the beam is then moved to the next column and the final digit of the address is supplied, causing the beam to step to position 5. This then is the desired location.

If we assume that the tube just described was controlling the vertical position of the spots in the slave tubes, a similar master tube could be at the same time controlling the horizontal positions of the beams in these same slave tubes, with the result that the desired address would be located in a time required to transmit the three-digit address code for either the horizontal or vertical positions. This is shown in Fig. 3, where, for simplicity, only one slave tube is drawn.

A simple method of returning the beam to the starting position is also illustrated in Fig. 2. It is only necessary to move the beam to a restoring column on the right to cause it to be

returned to the desired horizontal position; it can then be deflected to the left at position 1 as shown. A number of quite different stepping arrangements have been proposed and investigated, but since the one shown in Fig. 2 proved to be the simplest and most reliable, these other schemes will not be discussed in detail.

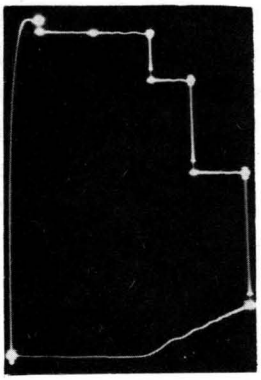


FIG. 5. Actual sequence of steps to locate the position 11101 on the serial tube.

Several characteristics of this scheme warrant special attention. It will be observed that the address must be supplied in time sequence and in the normal forward binary notation rather than in the reverse binary notation which is frequently employed in serial machines for the numbers on which arithmetic operations involving carry are performed. This must be carefully noted but should cause no trouble except in those cases where arithmetic operations are performed on addresses. The use of a timed-sequence address allows a simple master-tube construction having only a single output lead. While the tube is essentially a digital device, the beam is stepped across the slots in the target plate by means of an applied voltage. It is important to observe that the tolerances on this stepping voltage are so great as not to nullify the digital principle of operation. The most severe requirements on this stepping voltage occur when the spot is in the last column, in which position the amplitude of step must vary by about $\pm$ 50 percent to cause failure. This can be seen by studying the form of the coördinate tube output current as a function of beam position in Fig. 12. A photograph of an experimental coördinate tube of the serial-address system is shown in Fig. 4. Figure 5 is a photograph of the screen of this tube showing the sequence of positions occupied by the beam in going through a complete cycle to locate the address specified by the binary number 11101.

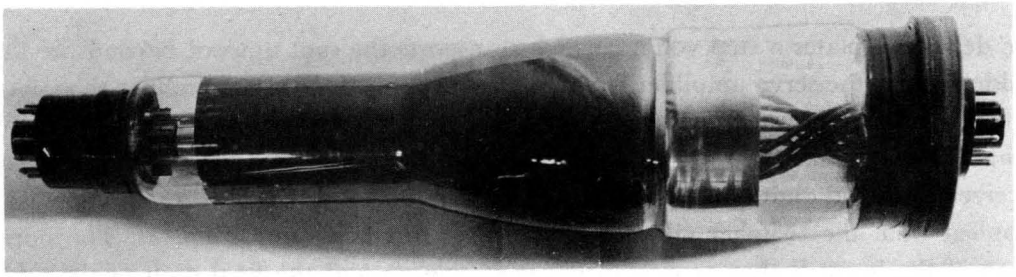


FIG. 6. Photograph of the parallel coördinate tube.

A distinction between the use of a serial or parallel address code and the operation of the complete computer on a parallel basis should be noted. In the system just described the address is supplied to the coördinate tubes in time sequence, but since the master tubes control 40 slave tubes each containing one digit, the stored information is available for use in a parallel adder if this is desired.

If a parallel system is envisioned, it would be more logical to supply the address to the

coördinate tubes in parallel rather than in timed sequence. For this reason a second type of beam-position tube, shown in Fig. 6, has been constructed. Structurally this tube is similar to the one just described in that it consists of a cathode-ray tube with a special target replacing the fluorescent screen. However, this target now consists of a metal plate containing vertical columns of windows, the vertical position of the beam being stabilized at one of eight locations determined by each of these windows.

The principle of operation is as follows. The electron beam is swept horizontally by a sine-wave generator at a speed that is high compared to the vertical operating speed. We have found it convenient to use a 30-Mc/sec oscillator for this purpose. The trace of the sweep spans all of the vertical columns of windows, as shown in Fig. 7. Wherever the beam encounters a window it enters and impinges upon one of the curved metal surfaces which may be seen behind each column. The secondary electrons ejected from a given one of these surfaces will either arrive or not arrive at a collector *C*, according as the bias upon the corresponding grid *G* is positive or negative. The current to the collectors produces a voltage drop which is then amplified by a high-frequency amplifier, and rectified, and the output voltage is supplied to the vertical deflecting plates of the tube in such a way that the vertical position of the beam rises as long as the secondaries in one or more columns reach a collector.

With this mechanism in mind we can now see how the beam finds the proper vertical position. The binary digits of the vertical address are applied as biases to the grids *G*, positive bias if a digit is one and negative if it is zero. The most significant digit is applied to *G*1, the next to *G*2, and so forth. As may be seen now by studying the positions of the windows in Fig. 7, the beam will rise to a unique position for each three-digit number applied to the grids if the initial position is near the bottom where the beam encounters windows in each column. For example, in the figure the beam is shown in the 010 position; only column 2 is open so the beam rises to the upper end of the lower window in this column and stops. Had the address been 110, the long window in column 1 would have bridged the gap between windows in column 2 and the trace would have risen to the top of the upper window in column 2. The zero position of the beam is established by cutting away the lower portions of the secondary emitting surfaces so that the primary beam can strike the collectors. This can be seen in columns 2 and 3 of the figure.

The over-all speed of this type of coördinate tube is somewhat better than that of a serially operated tube because the beam need stop only once in its search for an address. The tube also avoids the need of a direct-coupled external amplifier and column-stepping equipment. On the other hand, the high-frequency amplifier required by the parallel tube needs more gain than does the direct-coupled amplifier of the serial tube. The parallel tube itself is fairly complicated, and the over-all complexity of the two systems seems to be about the same.

In any application of coördinate tubes, the speed of operation is likely to be a matter of primary interest. While the over-all speed may depend upon the computer as a whole, certain limitations inherent in the coördinate tube itself may appropriately be discussed here. These

limitations are essentially those that are encountered in any feedback amplifier because of parasitic capacitance and finite tube transconductance, and are associated with feedback stability.

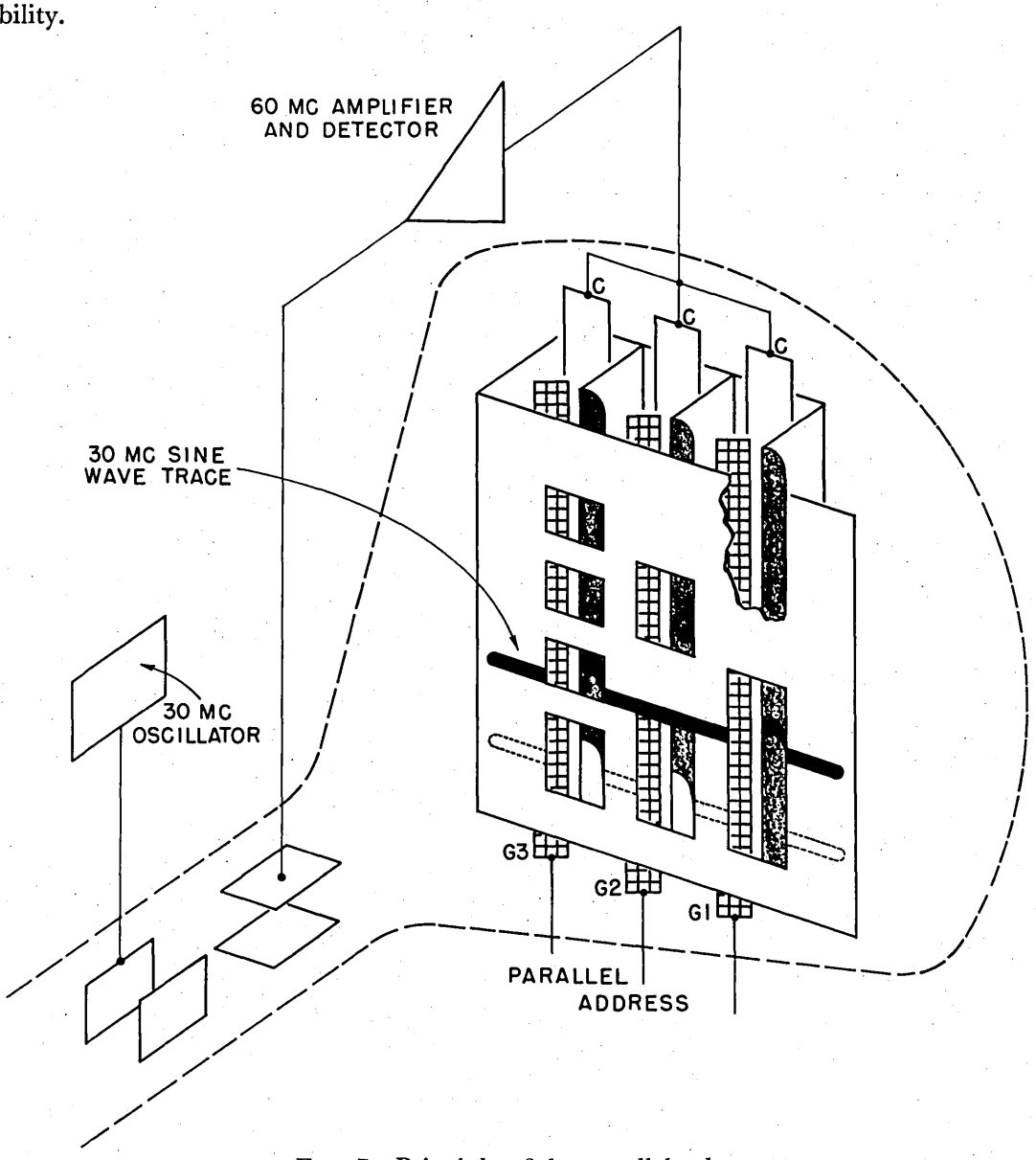


FIG. 7. Principle of the parallel tube.

In either of the two types of tube, the beam is expected to rise during switching until some edge of the locating comb is encountered and there to remain. For example, referring to Fig. 2, suppose that the beam is given a vertical twitch so that the feedback system causes it to rise from position 1 toward position 2. When the spot reaches the open slot it must not cross

the slot because the feedback system would then cause it to rise still farther to some undesired position. This places a restriction upon the speed with which the spot may be allowed to move relative to the speed of response of the stabilizing amplifier. Moreover, when the spot is resting in some position the feedback system should cause it to remain quiescent; that is, the feedback loop should be stable. These two requirements are somewhat similar, the second being ordinarily the more stringent.

The basic factors upon which these two types of stability depend may be understood by analyzing in detail a specific typical system, first for overshoot and second for static stability. The system considered is that shown in Fig. 8.

The following nomenclature will be used:

C_0, total amplifier output capacitance including the collector electrode of the coördinate tube;
d, equivalent spot diameter;
D, deflection sensitivity of coördinate tube;
g_m, mutual conductance of each amplifier tube;
G, zero-frequency gain per stage [$= g_m R$];
i_0, net beam current to collector;
$L(\omega)$, total gain ratio of feedback loop;
N, number of amplifier stages;
R, interstage shunt resistance;
S, vertical position of spot;
t, time;
v_n, output voltage of nth amplifier stage;
W, slot width;
ω, angular frequency.

With the electron beam in the position shown in Fig. 8, the beam current i_0 charges C_0 at a certain rate. The rising voltage across C_0, when amplified, causes the beam position to change at a rate given by

$$\frac{ds}{dt} = \frac{i_0}{C_0} G^n D. \tag{1}$$

When the spot position reaches the lower edge of the slot, the current in C_0 abruptly stops (in this part of the analysis the beam focus is assumed to be infinitely sharp). The requirement we wish to impose, then, is that the spot cease rising before it has travelled one additional slot width.

To calculate this we must know the response of the amplifier to an abruptly starting or stopping linearly rising voltage. For the amplifier shown in Fig. 5 we find that

$$V_{n+1} = CGe^{-x} \int_0^x V_n e^x dx, \quad x > 0 \tag{2}$$

if all v_n's are zero for $x < 0$. In this expression $x = t/RC$. If we take

$$v_0 = 0 \text{ for } t < 0$$

and

$$v_0 = at \text{ for } t > 0, \tag{3}$$

where a is constant, then Eq. (2) gives

$$v_n = ARCG^n \left[x - n + e^{-x} \sum_{0}^{n-1} \frac{n-k}{k} x^k \right]. \tag{4}$$

If the deflecting voltage of the coördinate tube is the output v_n of the last amplifier stage, Eq. (4) says that the ultimate ($t \gg RC$) motion of the spot is given by

$$S = atG^N D - aNRCG^N D. \tag{5}$$

The second term of this expression is the ultimate lag of the spot behind where it would have been with a perfect amplifier and the input given by Eq. (3). Since this is a linear analysis, this lag is the same as the overshoot when the beam reaches an edge after long travel. Comparing the two terms in the right-hand member of Eq. (5), we see that the overshoot is

$$NRC \frac{ds}{dt},$$

where ds/dt is the speed with which the spot enters the slot. The condition that the overshoot be less than one slot width, therefore, is

$$\frac{ds}{dt} < \frac{W}{NRC},$$

which may also be written

$$\frac{ds}{dt} < \frac{Wg_m}{NGC}. \tag{6}$$

The condition that the spot in a coördinate tube shall not overshoot its mark too far must certainly be met if the tube is to operate at all. Beyond this one might demand that the spot ultimately come to complete equilibrium and not dance about on the edge of the slit. Whether or not the spot will do this evidently depends upon the behavior of the complete feedback path, including the effect of finite spot focus. The feedback loop may be characterized by means of the complex loop gain, which for Fig. 8 is

$$L(\omega) = \frac{WDi_0R_0G^N}{d(1 + j\omega R_0C_0)(1 + j\omega RC)^N} \tag{7}$$

when expressed as a ratio. In this expression d is an equivalent spot diameter defined as the diameter the spot would have if the rate of variation of collector current with the elevation of the spot on the slot edge were constant over the spot diameter and equal to the actual rate at the equilibrium position. This gives a dimension that is proportional to diameter for similar spots and is of the same order as the visual spot diameter on a fluorescent screen.

The condition that the spot come to equilibrium can now be obtained from Nyquist's criterion for the stability of a feedback loop. This criterion is that the complex variable L

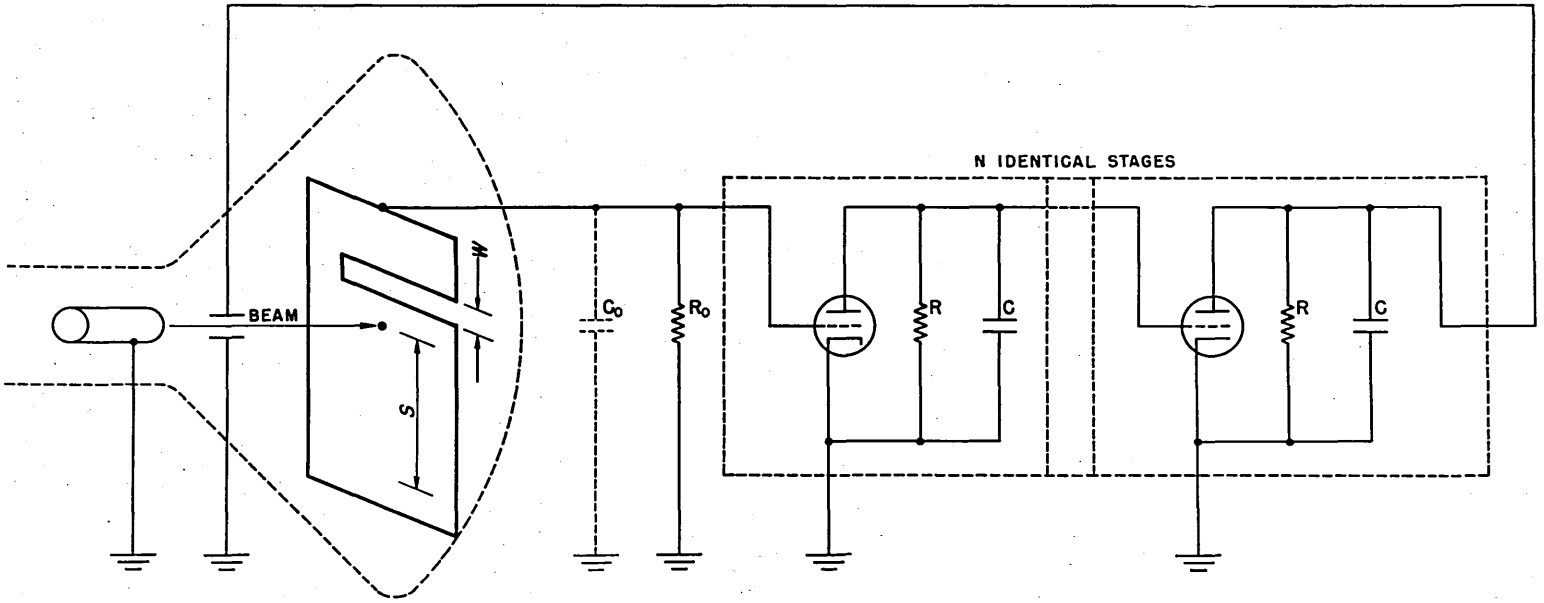


FIG. 8. System used to analyze stability.

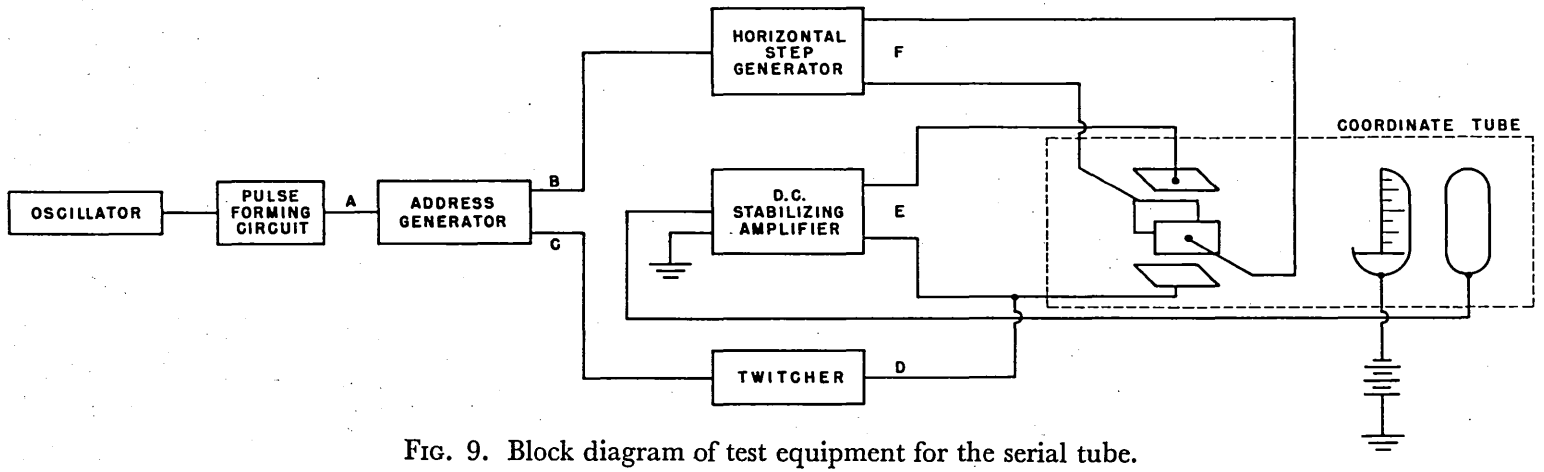


FIG. 9. Block diagram of test equipment for the serial tube.

not encircle the point -1 as ω traverses the real axis from $-\infty$ to $+\infty$. Examination of Eq. (7) shows that this requires that

$$\frac{i_0}{C_0} G^N D < \frac{\sin \frac{\pi}{2N}}{\cos^2 \frac{\pi}{2N}} \frac{dg_m}{GC},$$

since $L(0) \gg 1$. Making use of Eq. (1) this inequality becomes

$$\frac{ds}{dt} < K \frac{dg_m}{NGC}, \tag{8}$$

where K is a number of the order of unity.

Comparing (8) with (6) we see that the allowable spot speed for absolute stability is less than that for which the spot simply will not overshoot a slot in about the ratio of the spot diameter to the slot width. It is interesting to note that all factors describing properties of the amplifier enter both (6) and (8) in the same way.

The facts that must be considered in choosing the parameters of the amplifier system to operate a coördinate tube are relations (1) and (8), and a statement (9) of the maximum deflection which the amplifier must produce. Collected, these are

$$\frac{ds}{dt} = \frac{i_0}{C_0} G^N D, \tag{1}$$

$$\frac{ds}{dt} < K \frac{dg_m}{NGC}, \tag{8}$$

and

$$S_{\max} = i_0 R_0 G^N D. \tag{9}$$

When a specific coördinate tube is in mind the quantities D, W, i_0, d, and C_0 may be considered known. For fast operation, one would first of all choose an amplifier tube with as high a ratio g_m/C as possible so as to make the allowable speed high as given by (8). Furthermore, for a given total amplifier gain G^N the right-hand member of (8) is greatest when the stage gain G is 2.718. A logical design would be to choose this value of stage gain and then increase the number of stages so long as the spot speed given by (1) remains consistent with (8). Since this may require N to be as small as two or three for present tubes, some readjustment in the value of stage gain may be desirable. The resistor R_0 is then chosen so that the maximum deflection as given by (9) is 50 percent or so greater than the actual excursion over which the spot must stabilize. There are, of course, the usual matters of drift, dynamic range, allowable grid resistance, etc., to be considered in fixing the design.

The general considerations just given apply to the high-frequency amplifier of the parallel-type tube as well as to the low-pass amplifier. In the high-frequency case, however, the speed obtainable with a given type of amplifier tube is less than for the low-pass amplifier for two reasons: first, the speed is reduced by a factor of two because of the double-sideband operation;

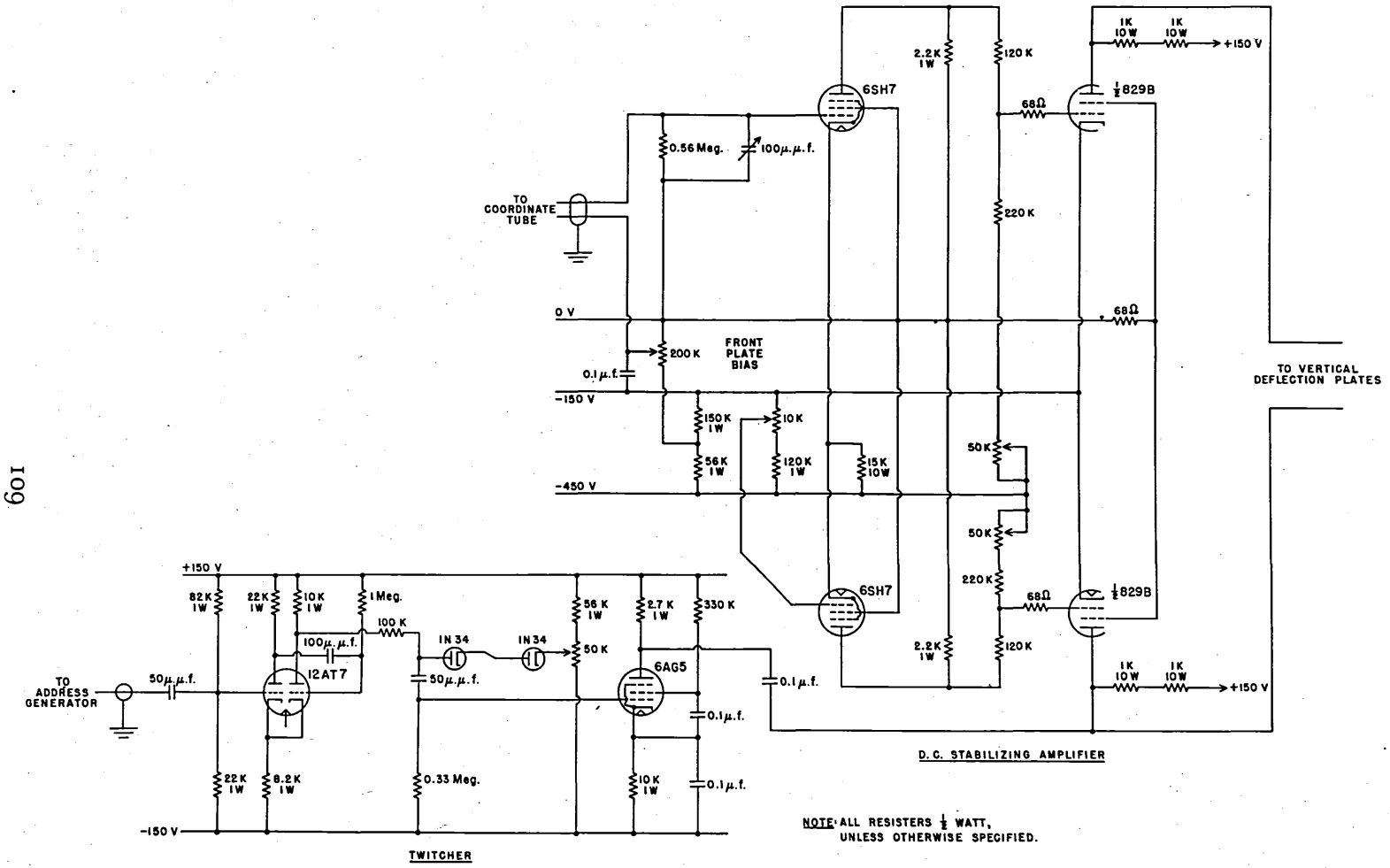


FIG. 10. Vertical-deflection circuits used to test the serial tube.

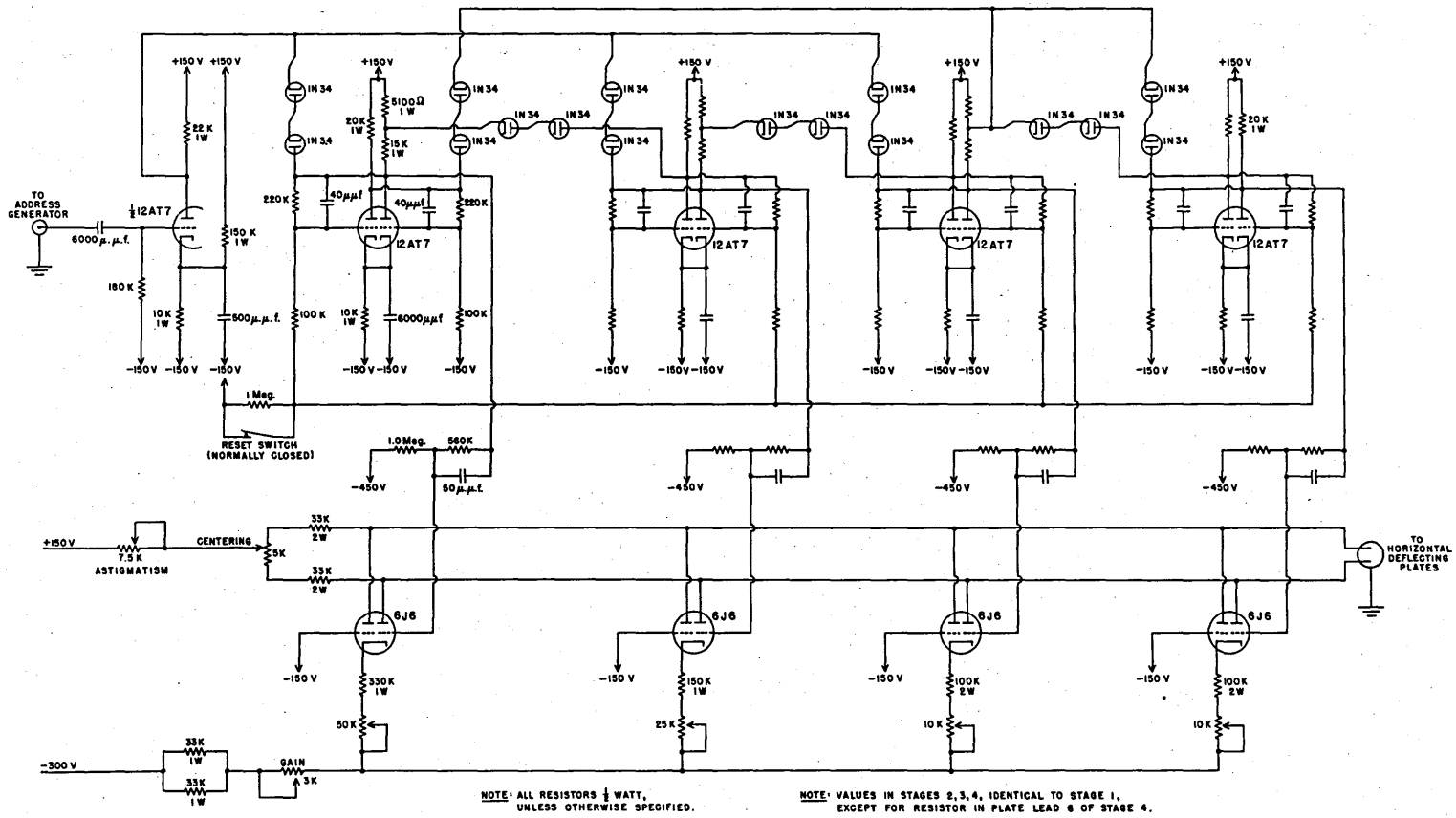


FIG. 11. Horizontal-step generator circuits.

and second, the value of high-frequency resistance R_0 required by the design procedure outlined above cannot be realized, and more than the optimum number of stages must therefore be used.

A number of experimental tubes have been built in the tube-construction laboratory of the University of Illinois. Two of these tubes are shown in Figs. 4 and 6. Except for the special target structures these tubes were made from standard cathode-ray tube parts. They have 32 possible spot positions determined by a five-binary-digit address. The operation of

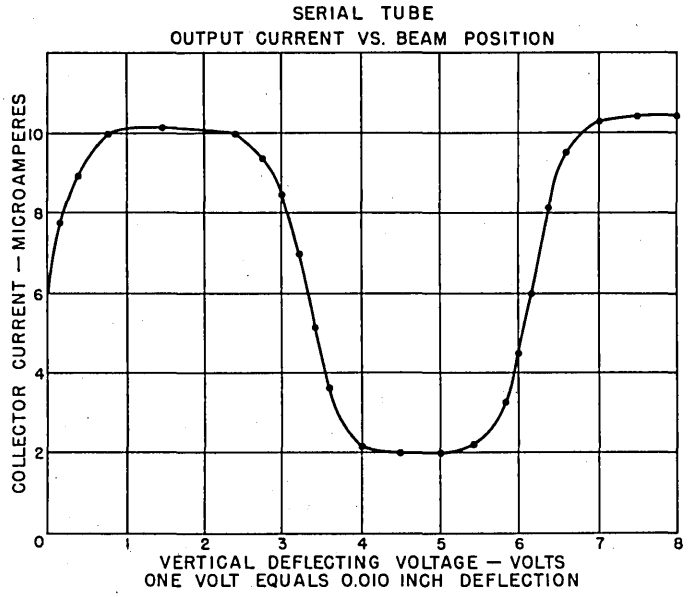


FIG. 12. The output current from the serial tube as a function of beam position when the beam is swept across the smallest steps.

these experimental tubes has been studied by means of special control-signal generators. A block diagram of the equipment for the serial tube is given in Fig. 9 and detailed schematic, in Figs. 10 and 11. This test equipment supplies the necessary signals for any address, either singly or repetitively, at a pulse rate up to about 200,000 pulses per second. A type 5UP1 oscilloscope tube, all of whose electrodes are in parallel with those of the coördinate tubes allows one to observe the path of the electron beam under various conditions.

By means of this equipment the effect of the various parameters of the tube and stabilizing amplifier upon the speed and stability of the system have been observed. The principle characteristics of a typical experimental tube are as follows: bulb diameter, 3 in.; number of teeth, 32; slot width, 0.032 in.; electron gun, 5U type; beam current, 10 μa; spot diameter, 0.010 in.; deflection sensitivity, 0.010 in./v; capacitance C_0, 50 $\mu\mu$f. This tube, when used with a two-stage stabilizing amplifier for which $g_m = 4000$ μmhos, $C = 20$ $\mu\mu$f, and $G = 10$, should be limited by overshoot stability (6) to a spot speed of about 5 teeth per microsecond.

This speed at which the spot will no longer lock in has been verified experimentally. It has also been observed that when the feedback-loop parameters are such that the speed is more than about one-third this great, the spot does not sit still but oscillates up and down on the edge of a slot. This observation checks (8) for the focus condition ordinarily used. The compromise between beam intensity and focus found most satisfactory is a spot diameter of about one-third of a slot width. Figure 12 shows the variation in collector current as the spot is moved over the smallest teeth with this focus condition.

Serial tubes also have been constructed in which the form of the comb was merely painted onto a piece of aluminum with India ink, rather than being a cutaway structure. This arrangement makes use of the difference in secondary emitting properties of aluminum and carbon. This method of constructing targets has the advantage that the target may have any outline

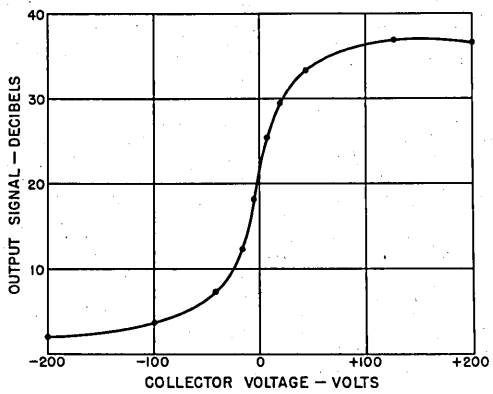


FIG. 13. The output from the parallel tube.

without regard for mechanical continuity of its parts. The two-piece cutaway target of Fig. 2 is, of course, more uniform in output than is the secondary-emission target.

Two types of parallel tube have been built: that shown in Fig. 6 and an earlier design without grids. In this simpler design a column is shut off by biasing the collector in the column negative so that the secondary electrons cannot reach it. However, the signal induced in the collector circuit by the secondary-electron space-charge cloud requires fairly large collector biases for its suppression. Figure 13 shows that an "open-closed" signal variation of 35 db is obtainable by the use of 200 v negative bias. The grids were placed in the later design so that smaller bias voltages would be effective. It was also expected that these grids would increase the "open-closed" discrimination by shielding the collectors from space charge since these grids were held at ground-high-frequency potential by built-in by-pass condensers. Unfortunately, owing to the difficulty of effectively grounding the entire target structure to the outside equipment, this improvement was not realized. Poor triode geometry also contributed to making this design less effective than the simpler arrangement.

The support given to this study by the Navy Department, Office of Naval Research, is

gratefully acknowledged. Most of the experimental work, including the construction of the experimental tubes, was done by Messrs. Robertson, Peiffer, and Haynes, graduate students in Electrical Engineering at the University of Illinois.

APPENDIX

HARMONIC OUTPUT OF PARALLEL-TYPE TUBE

As the electron beam scans sinusoidally across the windows in the parallel-type coördinate tube of Fig. 6, it produces current pulses to the collectors C of any columns that are not shut

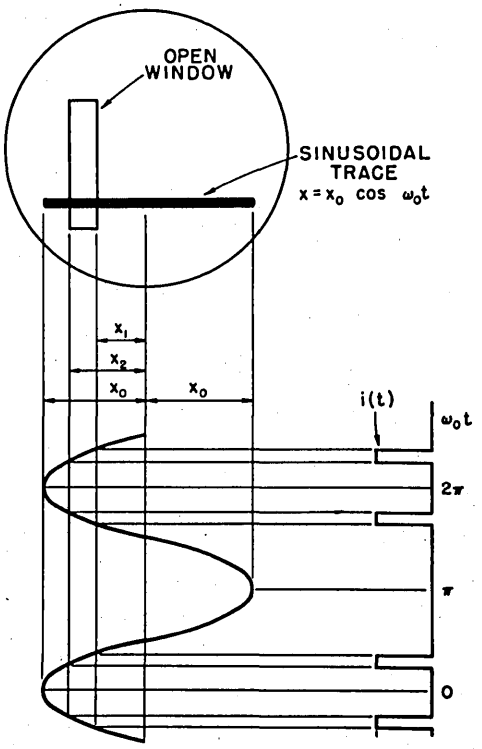


FIG. 14. Analysis of the output from the parallel tube.

off by the grids G. In order that the tube may function properly, these various pulses must all add in phase, and each pulse should contribute about equally to the output signal. The following analysis discloses the conditions under which this will occur.

The complex output current may be taken as the Fourier series in $\omega_0 t$, that is,

$$i(t) = \sum_{-\infty}^{\infty} a_n e^{jn\omega_0 t}, \tag{1A}$$

where the complex constants a_n are given by

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} i(t) e^{-jn\omega_0 t} \, d(\omega_0 t). \tag{2B}$$

Putting into Eq. (2A) the output current $i(t)$ due to a single open column located as shown in Fig. 14, we get

$$a_n = \frac{I_0}{n}\left[\sin\left(n\cos^{-1}\frac{x_1}{x_0}\right) - \sin\left(n\cos^{-1}\frac{x_2}{x_0}\right)\right].$$

Hence, the rms collector current in the nth harmonic may be written

$$i_n = \frac{\sqrt{2}I_0}{n}\,\triangle\left[\sin\left(n\cos^{-1}\frac{x}{x_0}\right)\right], \tag{3A}$$

where $\triangle$ [ ] means the value of the function in brackets at x_1 minus its value at x_2.

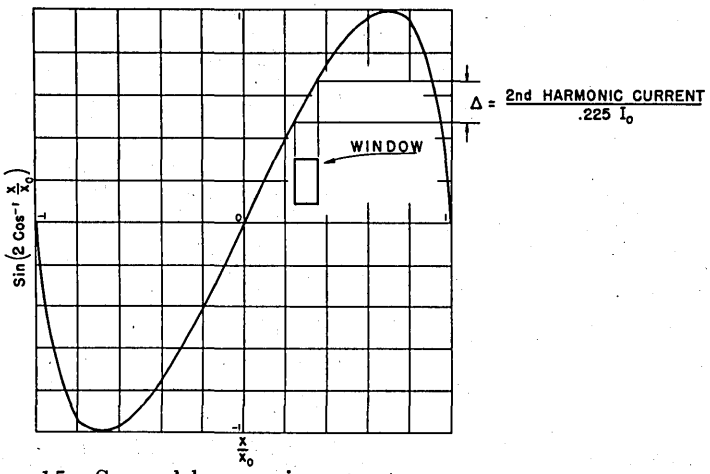


FIG. 15. Second-harmonic output curve;
$\triangle$ = (second-harmonic current)/$0.225I_0$.

Since a_n comes out real regardless of the values of x_1 and x_2, we see that the currents from any number of open slots will be either in phase or 180° out of phase. The total current in a given harmonic may readily be obtained for any placement of open windows from a graph of the function in brackets in Eq. (3A). This is illustrated in Fig. 15, where the function is plotted for $n = 2$. From Fig. 15 it is apparent that if the peak-to-peak sine-wave sweep is about twice the total width occupied by the windows then all pulses will add in phase and will contribute about equally.

The curve corresponding to Fig. 15 for the fundamental ($n = 1$) is a semicircle centered at zero. From this or direct physical reasoning, one sees that a window at the center of the sweep produces no fundamental output, and windows on opposite sides of the center tend to cancel. For this reason and because of shielding problems, the second harmonic rather than the fundamental output is used.

Basic Aspects of Special Computational Problems

HOWARD T. ENGSTROM

Engineering Research Associates, Inc.

This Symposium is impressive, both because of the large number of members and guests present and also because of the fearful and wonderful developments which have been and will be revealed in the papers presented here. The preceding papers on this program have been concerned largely with specific engineering developments; I should like to digress briefly and discuss some important factors of the general problem of procuring effective computing machinery.

I should like to emphasize, first, that the objective of all work on calculating machinery is to produce computational results. In spite of the large amount of activity in connection with the development of digital calculating machinery, I think it is a fact that nearly all of the computational results so important to our national defense and industrial economy are still produced by traditional methods. Most of the computational results are still obtained by machines of the desk calculator type, supplemented by the excellent machinery of the International Business Machines Corporation, Remington Rand, Burroughs, and others, including, particularly, such notable individual contributors as Professor Aiken, who has been responsible for the development of both numerical techniques and machinery for carrying them out.

Not very long ago, a war for which this country was unprepared found us equally unprepared to carry out many necessary computational problems. The procurement of adequate equipment to carry out these problems was a matter of great difficulty. In the light of the critical international situation in these postwar days, deep consideration should be given to the basic problem of procurement of this computing equipment.

In any discussion of a computational problem an immediate question is "Shall we use general purpose equipment, or design special equipment for this particular need?" It is axiomatic that any given computational process can be carried out more efficiently (i.e., either more rapidly or with less extensive machinery, or both) by equipment designed especially for the purpose. However, the decision between special-purpose and general-purpose equipment is difficult to make. It depends upon a number of factors related to each other in a complicated way. It depends not only upon the technical character of the problem but also on economic factors, the work load, and so on.

If solution time is not the most important factor, it is quite possible that general-purpose equipment may be preferable because of its versatility. If general-purpose equipment is obtained it can be applied later to other problems, as they arise. One activity equipped with general-purpose machinery which comes to mind immediately is, of course, the Computation

Laboratory at Harvard. Although this laboratory is devoted primarily to research into methods of computation, rather than to the actual performance of computing services, it has done much of the latter. Other general computation facilities are centered at the National Bureau of Standards. Some such services, notably in England, have been set up under private enterprise. The International Business Machines Corporation provides services of this character, particularly in connection with their large-scale computer in New York, although this service is again primarily scientific in objective. Some bureaus and divisions of the government have found the volume of individual computational problems sufficiently great to warrant setting up laboratories of general-purpose computing equipment to carry out services of this character which arise within their particular divisions.

At the other extreme are the computational problems requiring a large volume of specialized work of a repetitious nature where the load is kept constant. In these situations special-purpose equipment is the obvious choice. I shall sketch briefly some examples of specialized problems requiring extensive repetitive computation at nearly constant work load.

The airport facilities and the airways of this country are being subjected to increasing congestion, particularly under adverse weather conditions. The problems of air-traffic control and airport time utilization are essentially computational in character. They are problems of automatic continuous inventory. With respect to the airport time-utilization problem, the basic preliminary design plans for equipment that will solve it have already been prepared. This equipment will store information on airport runway assignments by hour and minute, classified as to class of aircraft and arrival and departure times. The proposed equipment will supply this information upon inquiry and will change the stored information to conform to changing situations occasioned by weather conditions. Supplementary information, such as the identity of the plane and its route, likewise will be made part of the record.

Present control equipment in general use performs no computations, and even the most routine decisions are presently made by human controllers. Eventually it is planned that the input and output to airport time-utilization equipment will come from communication channels, and that the proposed equipment, which I have described very briefly, will be used at all large airports.

This is a typical problem in which the use of special computational equipment is necessary. The development of such equipment must be pursued strenuously, and its installation encouraged. Operational control of large numbers of aircraft is of vital importance; the nation may be faced with a need for a practical solution of this problem on short notice.

Although of limited sophistication, the problem of reservation control also is one of importance. Those of you who spent valuable time during the war sitting in airports in far-off places waiting for air transportation, or in railroad stations attempting to get railroad transportation, realize this only too well. Technical methods of a computational nature for the solution of these reservation problems have been proposed. The basic reasons why this type of service is not yet available are nontechnical in character. They depend upon operational and financial conditions. For example, these are the questions which arise: Is it better for

each airline to maintain a single reservation control, centrally tied in by communication lines to its outlying offices, or to maintain separate centers in the major cities from which it operates? Is it preferable for the airlines to combine their reservation control on an intercompany basis in each major center? To what types of transaction must computational equipment provide the reply? Answers to these questions are being sought by the Air Transport Association and committees consisting of representatives of commercial airlines.

Another field in which large-scale computing is required and in which the arithmetic is straightforward can be designated under the heading of inventory control. Many important problems in this field are being handled adequately now, but the earlier years of the last war may be characterized by the statement "too little and too late," largely because of inadequate inventory control. The later years of the war were marked by the rise of priority systems and the resulting controversies. One basic assumption which may be made is that any future wars of these United States will be fought in the economy of limited scarcity. This means that improvement in the methods of the control of inventories must continually be carried out and that plans should be made to speed up even those methods that are satisfactory now. Applications to these problems of techniques such as those discussed at this Symposium are seriously lagging.

There are numerous other fields in which the application of special-purpose computing equipment is obvious. These are situations in which specific data-reduction problems exist; problems of control in which the required degree of precision is so high that digital rather than analogue techniques must be used. In all these fields the question remains: Why have not the successful results of researches been brought to bear on these problems? I believe the basic answer to this question lies essentially in nontechnical fields. The following reasons I believe are basic:

(1) *Lack of reliability.* The reliability of electronic equipment involving large numbers of vacuum tubes is still questionable. Reliability is of paramount importance in connection with any problems involving automatic control or inventory. In putting together a digital computer, whether special or general purpose, a great deal of time is spent in removing the bugs. Although components operate well individually, the interconnecting and matching problems assume large proportions. The maintenance of special computing installations, however soundly engineered, is a problem of the first magnitude.

(2) *Economic factors.* The economy of this nation is such that sources for procurement of computational equipment must be found in private industry. The researches on digital computing equipment have been carried out to a large extent at universities under government sponsorship. Large-scale computing devices are expensive. Private enterprise, which must make a profit, is naturally reluctant to invest the large sums necessary to establish procurement sources on an industrial basis. Rapid advances in the art are, paradoxically enough, a hindrance to industrial development because no one wants to spend money on equipment that may shortly become obsolete. Also, the industrialist requires some competitive protection in the form of patents or exclusive rights to equipment and techniques. The patent structure

with respect to the large-scale computing devices is complicated by the fact that so much of the work has been carried out either within the government itself or in nonprofit institutions. Moreover, there are no accurate data on the cost of producing such equipment because the methods of accounting employed in universities, government laboratories, and industry are so different. Hence, the question "What is a reasonable price for a computer?" is difficult to answer.

I have devoted much of this paper to generalities. In order to come within the compass of the title of this session, "Recent Developments in Computing Machinery," I must mention the contribution of the company which I represent. I believe our basic contribution to the practical solution of many special computational problems is our work with magnetic-drum storage and the ancillary electronic techniques. We have placed considerable emphasis on the perfection of the magnetic drum, both as a scientific instrument and as a competitive commercial component. I am happy to report that we have had magnetic-drum equipment operating satisfactorily for a period of two years. Our efforts to develop and design components and to evolve manufacturing techniques and processes have attained a degree of success such that magnetic-drum storage can be considered an industrial component. We have developed reliable magnetic heads, drum-surface materials and techniques, and mechanical-design principles. These will be the subject of papers presented elsewhere. I wish to point out, simply, that the magnetic drum, as a component of special digital computing machinery, is now available.

In closing, I should like to state again that the needs for special computing equipment in many aspects of our national defense have not been met. Large gaps exist in the fields of operational control and in highly specialized computing. Components to solve many of these problems have been developed, but are not industrially available. Increased attention must be given to these problems or the program on large-scale digital calculating machinery may be given the label "too many words, too few numbers."

Electrochemical Computing Elements

JOHN R. BOWMAN

Mellon Institute

Several fundamental and general electrochemical effects are potentially useful in the design of digital computing-machinery components. These include chemical deposition, electrolyte polarization, hydrogen-electrode polarization, anodic-film polarization, and alteration of surface tension. These can be combined in various types of cell to provide the functions of storage and selection. Such components have the advantages, over most of their equivalents, of small size and low cost. In speed, they fall in the millisecond, or more rarely in the microsecond, range. Their main disadvantage is that they are essentially low-voltage direct-current units, and hence are not particularly well suited to electronic coupling.

Electrochemical devices have found little application in communication engineering. This is largely because the effects are essentially qualitative and not reproducible to better than several percent. In digital networks, however, such reproducibility is not required, and electrochemical devices can be designed that will give good dependability in discriminating between two discrete states. This applies to all of the cells to be described. All of the effects discussed are reversible, not in the thermodynamic sense, but in the sense that input of a suitable signal will bring the device back to an original state after having received an intermediate signal.

From the qualitative character of these effects, detailed development of any unit must be closely associated with the development of the entire network. For this reason, the information presented here is essentially a theoretical discussion of principles supported by a minimum amount of experimental results. Further experimentation would be useless without a definite object of computer design as a whole.

As is well known in the electroplating arts, the passage of charge through a cell containing metal ions may cause deposition of metal on an electrode. This effect is a reversible one, and the presence of the metal film on the electrode gives the cell an output voltage. When the circuit is closed, a current is established in it and the metal returns to the electrolyte as ions.

This effect can, potentially at least, be utilized to make a storage device. Consider, for example, a cell composed of similar electrodes and an electrolyte containing ions of a metal that plates out well. An electric impulse to such a cell will cause it to have an emf of sign opposing the input pulse. Application of a second input pulse of opposite polarity to this charged cell will cause anodic removal of the metal originally plated out and simultaneous deposition of metal on the other electrode, the emf of the cell thereby being reversed in polarity. Such a cell has properties similar to those of a capacitor.

Selection of the electrode and electrolyte materials for a memory device depending on this principle requires certain obvious considerations. Perhaps the most important is that the metal to be deposited be not subject to corrosion by the electrolyte. Further, it is desirable that the cell potential be as high as practical. These requirements are essentially contradictory because the most active metals present the largest electrode potentials. The best compromise appears to be silver. This metal is unique in that it is the most active material that is unaffected by aqueous solutions of its own salts. It plates out well and develops usefully high potentials.

The electrodes to be used should be inert chemically and should polarize readily with respect to hydrogen. These requirements lead almost uniquely to gold as the electrode material.

The electrolyte should be stabilized to constant concentration of metal ions, a condition most readily met by providing for use of a saturated solution and supplying an excess of the solid-phase salt. For the silver-gold system, silver sulfate fulfills these conditions conveniently, being stable and soluble to a useful extent in water.

Numerous experiments have been conducted on the cell

$$Au/Ag_2SO_4, \text{ sat. aq.}/Au.$$

It is readily reversible, stable on standing for several months, and gives a steady output emf when charged to 0.1 to 0.2 v.

The actual value of the emf in the charged state is not reproducible, and appears to depend greatly on the nature of the gold surface on which the silver was plated out during the charging cycle, the rate at which the silver was deposited, and the amount of silver deposited. In general, high voltages are obtained for rough electrodes where small amounts of silver are deposited rapidly. In no case, however, was there ambiguity as to the sign of the polarity.

As will be discussed under hydrogen polarization, the emf of this type of cell may be high, i.e., 1 v or more immediately after charging, but this value decreases to that of a normal silver electrode in a few minutes.

This simple cell has the disadvantage that successive charging pulses or a long-continued one will deposit additional silver linearly with it, and reversal may require a large charge. This may be overcome by introduction of acid in the electrolyte to cause concentration polarization.

Since the mobilities of the ions in an electrolyte are in general different, a current in it gives rise to concentration gradients. In particular, the hydrogen ion is highly mobile and will carry a large part of the current relative to its concentration. If the electrolyte bearing silver ions is initially acidified and uniform, a substantial part of the current will initially be carried by the silver ion, but as the action proceeds the ratio of hydrogen- to silver-ion concentration near the cathode will decrease sharply. Continued current will then deposit relatively small amounts of metallic silver, and a saturation effect exists. The charge-retention characteristic of a typical cell is illustrated in Fig. 1. Charging curves have been obtained

on numerous cells of this type. The initial portion of the curve is nearly linear in deposition of silver. As the charging proceeds, the electrolyte becomes exhausted of silver ion in the vicinity of the cathode, and the principal reaction at that electrode is release of hydrogen, which is quickly lost and does not contribute to the charge retained by the cell.

On applying a charging pulse of reversed polarity to an already charged cell, the limited amount of silver originally deposited is promptly removed because there is an abundance of sulfate ion in the neighborhood of the electrode originally serving as cathode. As a consequence, a cell subject to electrolyte polarization can be reversed with a single pulse of sufficient size, even though it had previously received several successive charging pulses of equal size and opposite polarity. In a computer storage device this effect eliminates the need for erasure before reading in new signals.

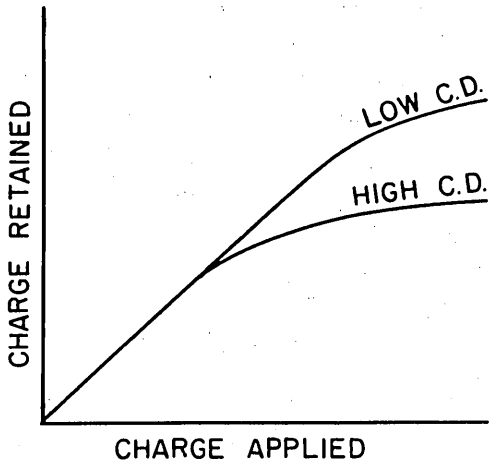


FIG. 1. Charge-retention characteristic of a typical silver sulfate cell.

Mathematically, a quantitative statement of this effect can be formulated as follows, neglecting acceleration and diffusion, which are negligible under all normal conditions.

The absence of space charge can be analytically stated in the form $\sum \rho_i = 0$, where the ρ_i's are charge densities corresponding to the different ionic species.

Letting the quantities i_k designate the components of current carried by the different ionic species we have

$$\sum i_k = i,$$

where i is the total current carried by the electrolyte.

The partial currents associated with the ionic species are proportional to the respective ionic specific velocities and to the concentrations; hence

$$\frac{i_1}{u_1\rho_1} = \frac{i_2}{u_2\rho_2} = \frac{i_3}{u_3\rho_3} = \cdots.$$

The partial currents and densities are related by the equation of continuity, which in one dimension reduces to

$$\frac{\partial i_k}{\partial x} = \frac{\partial \rho_k}{\partial t}.$$

The differential equation governing the over-all effect can be set up from these relations in terms of the ionic charges σ_k where

$$\frac{\partial \sigma_k}{\partial t} = i_k \text{ and } \frac{\partial \sigma_k}{\partial x} = \rho_k.$$

The result is that

$$\sum \frac{\partial \sigma_k}{\partial t} = i, \quad \sum \frac{\partial \sigma_k}{\partial x} = 0,$$

and

$$\frac{\dfrac{\partial \sigma_1}{\partial t}}{u_1 \dfrac{\partial \sigma_1}{\partial x}} = \frac{\dfrac{\partial \sigma_2}{\partial t}}{u_2 \dfrac{\partial \sigma_2}{\partial x}} = \frac{\dfrac{\partial \sigma_3}{\partial t}}{u_3 \dfrac{\partial \sigma_3}{\partial x}} = \cdots .$$

No analytical solution has yet been obtained for this nonlinear system, but it would be amenable to numerical treatment with a digital computer.

For high speed and sensitivity the cells should be constructed with minute dimensions. A useful form employs gold-wire electrodes of diameter 0.001 in. imbedded in a bead of acidified silver sulfate paste or gel about 0.5 mm in diameter. From experiment there is indication that the minimum stable silver film must be about 100 atom diameters thick. The basic constants for such a cell are given in Table 1. The constants associated with the

Table 1. Basic constants of silver.

Atomic weight	108
Density	10.5 g/ml
Molal volume	10.3 ml
Avogadro number	6.0×10^{23} atoms/mol
Volume of atom	1.71×10^{-22} ml
Diameter of atom	5.55×10^{-8} cm

Table 2. Characteristics of a silver film, of dimensions $(5 \times 10^{-2}) \times (7.5 \times 10^{-3}) \times (5.5 \times 10^{-6})$ cm.

Volume	2.1×10^{-9} ml
Molal equivalent	2.0×10^{-10} mol
Faraday constant	9.6×10^{4} coulombs/mol
Electrochemical equivalent	1.9×10^{-5} coulomb

Table 3. Characteristics of the electrolyte: a sphere of saturated Ag_2SO_4,Ag, 5×10^{-2} cm in diameter.

Volume	6.4×10^{-5} ml
Solubility, Ag_2SO_4 in water	6.0×10^{-3} g/ml
Mass of Ag_2SO_4 in solution	3.8×10^{-7} g
Molecular weight, Ag_2SO_4	312
Silver in solution	2.4×10^{-9} mol
Equivalent number of films	12

capacity of the cell are listed in Table 2, and the constants concerning its electrical characteristics in Table 3. Its resistance is of the order of 0.1 ohm. A suitable charging pulse is 200 ma for 1 msec; about one tenth of this is retained.

A basically different type of accumulative device is provided by similar inert electrodes in an electrolyte consisting of a dilute acid alone. Application of a charging pulse will, to a certain point, produce adsorbed hydrogen on the cathode. This hydrogen gives a relatively high emf, but one of relatively short duration; its half-life with 0.001-in. gold-wire electrodes is of the order of a few minutes, but its value may be between 2 and 3 v. Physically, it can be constructed much like the silver cell described.

The more detailed characteristics of such a cell are of practical as well as theoretical interest. At low voltage, i.e., below about 0.6 v, the cell behaves as a linear capacitor with a capacitance of about 10 μf per square centimeter of cathode area. As the voltage increases above this region, the cell accepts and retains a considerably larger charge, and finally at a still higher voltage hydrogen gas is released as bubbles, and current for the first time becomes steady. In the conducting range the back voltage of the cell increases logarithmically with the current. These phenomena are illustrated schematically in Fig. 2.

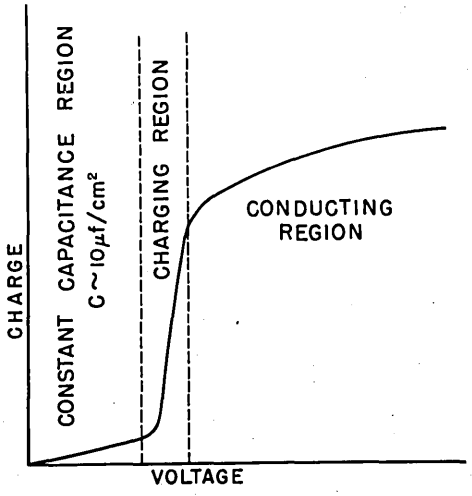


FIG. 2. Charge-voltage characteristic of an electrode-polarization cell.

Cells of this type are to be compared with the metal-disposition type in several ways. They supply potentials an order of magnitude higher, have far greater speed and sensitivity, but have half-lives several orders of magnitude less. They may, however, find application in operating organs of computers, such as adders, where short-time storage only is required. They could, of course, be used for long-time storage if periodically read and regenerated.

An extreme type of electrode-polarization cell can be constructed using tantalum-wire electrodes in an acidic electrolyte. The anion should preferably be of high valence, such as borate and phosphate. The anode of such a cell develops a high-resistance film which does not permit conduction until about 50 v and will retain available charge at voltages of this order of magnitude. The life of the charge, however, is short. The read-out operation can be measurement of either emf or resistivity.

The formation of high-resistivity films on anodes can be used for rectification as storage. By provision of one tantalum electrode and one inert one, preferably gold, a rectification unit having much the characteristics of a germanium-crystal rectifier is produced. Its main disadvantage is high capacitance, which precludes its use for extremely high-speed operation,

but this can be largely overcome by making the unit physically very small, as recommended for the other cells discussed here.

As is well known, the interfacial tension between mercury and an electrolyte is strongly dependent upon the potential difference between them. Extensive use has been made of the phenomenon in capillary electrometers. An extension of this device provides the function of a relay. The general arrangement is shown in Fig. 3. When the mercury droplet is negative with respect to the electrolyte by about 0.25 v, its surface tension is low and it flattens out

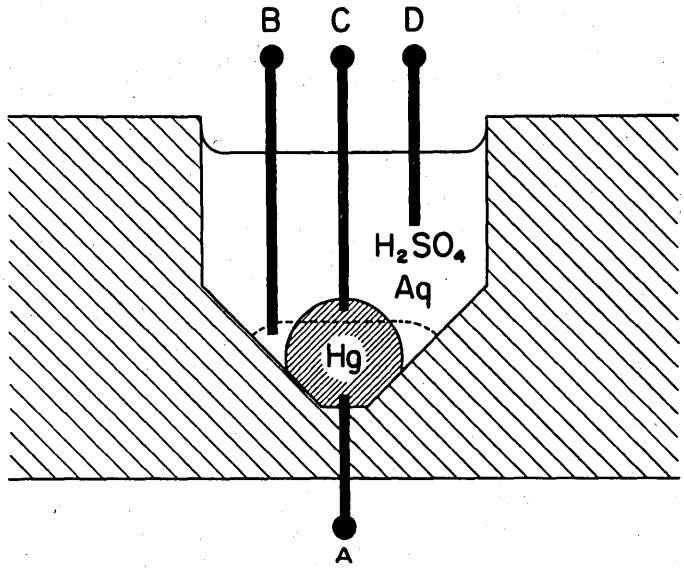


FIG. 3. Surface-tension cell.

to make contact with electrode *B*. When the applied control potential is removed, the surface tension very quickly resumes its normal value and the droplet returns to nearly spherical form, breaking connection with contact *B* and making connection with contact *C*. Contact *A* provides a common input and return for the control voltage, which is supplied to contact *D*, which at no time makes contact with the mercury droplet.

The device serves as a single-pole, double-throw, voltage-actuated relay. Its time of response is less than 1 msec.

Care must be used in applying voltages to open circuits of this relay which would take over the controlling function. Such errors can be wholly eliminated by establishing gating paths for signals before passing the signals through them, and keeping the signal pulses sufficiently short that they cannot assume control. Similar circuitry is good practice with mechanical relays, where the contacts are never required to make or break a current. Devices of this kind are potentially useful for large pyramid selectors.

Logical Syntax and Transformation Rules

GEORGE W. PATTERSON

University of Pennsylvania

This paper is a progress report on attempts to develop a unified theory of calculating machines. First of all it should be made clear what devices are under discussion. In the present state of the development it does not seem likely that analogue computers will be included; on the other hand, many devices not ordinarily considered to be calculators could fit into the analysis. As examples of such apparatus, the following are mentioned: teleprinter equipment, voting machines, cryptographic mechanisms, "tote" boards, type-casting machines, and at least certain portions of systems for railway signaling, centralized train control, automatic telephone switching, and pulse code modulation. It will be noted that all these devices are concerned with the handling of data or intelligence, generally in a coded form.

These mechanisms are obviously not of central interest to us, and are mentioned merely to indicate the possible broad scope of a unified theory. The principal motivation for a theory comes from the problems associated with the design and use of desk calculators, punch-card equipment, automatic message-accounting equipment, and, most important of all, large-scale digital calculators. The totality of devices already mentioned will be called syntactical machines, for reasons that will appear presently. Calculating machines constitute an important subclass of syntactical machines. Some of the problems whose solution could be aided by a unified approach are:

(1) To determine and describe the exact relation between operations built into the machine and the operations of mathematics.

(2) Given the constructional details of a machine, to determine its exact operational characteristics; this is the problem of machine analysis, which is important to the design engineer.

(3) Given the exact operational characteristics, and a family of components, to determine how they must be interconnected to produce the desired results; this is the problem of machine synthesis, even more important and difficult than the previous one.

(4) Given a set of basic machine operations, to construct by iteration and combination new and more complex operations; this is the problem faced by the coder.

(5) To determine a set of basic machine operations that are capable of the extension just referred to, with a reasonable amount of coding effort, and that can be physically realized with a reasonable amount of equipment; this is the main problem of logical design.

It would be rashly optimistic to predict that the solution of these problems will ever be reduced to a routine process; on the other hand, the development of theoretical tools that will assist in any way in their solution is a worthwhile project.

The building blocks of the development under discussion are logical syntax, symbolic logic, and conventional mathematical anaylsis. All that needs to be said here about symbolic logic is that it provides a useful symbolic apparatus for manipulating such words as "not," "or," "and," "implies," "is equivalent to," "there exists," "for every," in any arbitrary context.

Conventional mathematical analysis as the only tool for analyzing the operation of a calculating machine has several shortcomings:

(1) The machine operations are frequently not equivalent to ordinary mathematical operations but only approximations thereto.

(2) Mathematical analysis is ill-adapted to consideration of such operations as sorting, collating, selecting, extraction of fragments of words, and the like.

(3) Description of machine operations in purely mathematical terms obscures the possibility of nonmathematical interpretations of the data, such as would be involved in programming a machine to play chess, for example.

Where does logical syntax fit in? This question brings us to the basic characteristics of syntactical machines. They all accept input data or information, and produce output data. They are, in a sense, linguistic transducers, although the languages involved are usually extremely artificial and symbolic. But languages may be said to have three aspects: the structural, the meaningful, and the motivative, otherwise known as syntax, semantics, and pragmatics.[1]

It should be emphasized that this division of the theory of language into three parts has been made by logicians and philosophers, principally Professor Rudolf Carnap of the University of Chicago, rather than by students of natural spoken languages, and it has been principally applied to artificial symbolic languages that have been designed by mathematics, logicians, and engineers to be as accurate and unambiguous as possible. In the present state of the art, these are the only languages manipulated by machines.

Pragmatics, the motivative aspect of language, deals with the relation between the expressions of a language and the actions they produce in the hearer or the consumer. In the theory of natural languages we could ascribe the principles of rhetoric, propaganda and advertising copy writing to this subject. In mechanisms it would consider such things as the behavior of digital equipment as a link in a servomechanism; there is a close relation here to cybernetics. In calculators designed for scientific and accounting purposes the feedback link lies in the human consumer of the computational results. This aspect of language lies outside the scope of the present discussion.

Semantics, the meaningful aspect of language, deals with the relations between the expressions of a language and the objects or events that they designate. The designer of a scientific calculator selects and builds into it those operations that make the ascription of numerical meanings to the machine language as simple as possible, but it should always be borne in mind that the user is free to assign any meaning he wishes, and in machines for commercial purposes a nonnumerical meaning or interpretation may well be of equal importance. Felix Klein was one of the first to clearly recognize this fact; as he said, "*the rules of*

operation alone, and not the meaning of the numbers themselves, are of importance in calculating, for it is only these that the machine can follow; it is constructed to do just that; it could not possibly have an intuitive appreciation of the *meaning* of the numbers."[2]

The expression "the rules of operation alone" characterizes syntax. The following is quoted from Rudolf Carnap. "By the logical syntax of a language, we mean the formal theory of the linguistic forms of that language—the systematic statement of the formal rules which govern it together with the development of the consequences which follow from these rules. A theory, a rule, a definition or the like is to be called *formal* when no reference is made in it either to the meaning of the symbols (for example, the words) or to the sense of the expressions (e.g., the sentences), but simply and solely to the kinds and order of the symbols from which the expressions are constructed."[3]

What strings of characters are numerical expressions? This is a syntactical question, and its answer is a syntactical *formation* rule, since it explains how to form a numerical expression.

Given two numerical expressions, how do we form a third, which we may call their sum? This is a syntactical question, and its answer is a syntactical *transformation* rule, since it explains how to transform given expressions into new expressions. These rules are established with the meanings of the expressions as a basis, but this is in the background in syntactical investigations. The fact that symbols can be manipulated by formal rules, without any reference to their meanings, is what makes digital calculating machines possible.

In any investigation of a language, we require a language to state the results of our study. The language under investigation is the object language, the medium for expressing the results is the metalanguage. The metalanguage used is a matter of choice, but a judicious combination of symbolic logic, ordinary mathematics, and English is recommended.

So much for the abstract principles of logical syntax. Its development has been largely carried out by logicians in investigating the structure of mathematical proofs. These aspects of the subject do not bear directly on the theory of computing machines, and consequently the information available in the literature serves only as a starting point, and much further research is needed.

The application of syntax to computing machinery will be illustrated by a few specific examples. First of all we will consider the description of algorithmic number systems from a syntactical point of view. These number systems are those designed to provide names for all nonnegative integers. The ordinary denary, or decimal, system and the binary system are the most common examples. This description will be general and not restricted to a particular base or radix.

We assume the existence of β distinct kinds of symbols or *characters*, where $\beta \geq 2$ and is an integer. The exact nature of these characters is immaterial; they may be holes in paper tape, marks on paper, electrical pulses, distinct identifiable positions of rotating elements, light signals. All we require is that we can recognize any character, and unambiguously determine to which of the β classes it belongs.

It is further assumed that these characters can be arranged in strings with a definite beginning and end, and that each character except those at the ends has a unique immediate predecessor and a unique immediate successor. This constitutes a linear syntactical system or language. Each string is an *expression* of the language. Multidimensional expressions are also used, for example, in matrices and punch cards, but they will not be considered here. An expression, a string of characters, can be symbolized by $\overset{n}{\underset{i=m}{\frown}} x_i$. The variables of the meta-language are restricted to range over $\beta \geq 2$ possible values, i.e., to the characters. The integer β is a parameter that constitutes the radix or base of the system. If $\beta = 10$, and we are dealing with the ordinary written numerals, $\overset{2}{\underset{i=0}{\frown}} x_i$ = '365' would signify that x_0 = '5', x_1 = '6', x_2 = '3'. Note the numbering of characters from "right to left." The single quotation marks are used to denote the fact that we are considering the marks themselves and not the numbers they denote.

Certain other properties are assigned to the β distinct kinds of characters. First of all we require that a discrete cyclic order be established among the β characters. The red characters on the telephone dial exemplify this for the ordinary written system, $\beta = 10$. The cyclic order progresses counterclockwise around the dial. The cyclic successor of x will be denoted by $\sigma(x)$. Thus, on the dial, σ('0') = '1', σ('9') = '0', etc. Iterations of the cyclic successor operator will be denoted by exponents; for example, in the ordinary system, σ^5('0') = '5', $\sigma^{10}(x) = \sigma^0(x) = x$. In general, $\sigma^r(x) = \sigma^s(x)$ if and only if $r \equiv s \pmod{\beta}$. In addition to the cyclic order, we single out a particular character and call it 'Nu'. On the telephone dial, Nu appears immediately below the hook, i.e., in the ordinary system Nu = '0'. In modern written Arabic, Nu = ' . '; as transmitted by the telephone dial, Nu is a closely spaced time sequence of ten pulses. The telephone dial is a simple, inexpensive, syntactical machine; when properly manipulated, it transforms the '0' appearing in the directory into the requisite pulses.

The idea of the $\beta \geq 2$ characters, the cyclic order imposed on them, and the fiducial Nu, are the elements of the development. A numerical expression is defined as any expression with at least one character having the property that the first character is not Nu, i.e., $\overset{n}{\underset{i=m}{\frown}} x_i$ is a numerical expression provided $m \leq n$, and $x_n \neq$ Nu. It will be useful to have a symbol for the expression with no characters at all, and 'Λ' is selected for this purpose. The arch '$\frown$' means 'is followed by,' and capital letters will be used for expression variables, since small letters are reserved for character variables. Note that a single character is always an expression, but not conversely. To illustrate this notation:

$$\Lambda \frown X = X = X \frown \Lambda.$$

We have defined the numerical expressions as the totality of expressions that begin with a character distinct from Nu. This is a formation rule. In order to proceed further we define a transformation rule, which enables the determination of the successor of a numerical expression, in the sense of Peano's axioms, that is to say, the operation of counting. We do this

with the aid of the auxiliary notion of quasi successor which is defined for any arbitrary expression.

Definition 1. Quasi successor (quasinachfolger), qnf (X).

(1) $\text{qnf}(\Lambda) = \sigma(\text{Nu})$.

(2a) If $\sigma(z) \neq \text{Nu}$, then $\text{qnf}(Y \frown z) = Y \frown \sigma(z)$;

(2b) If $\sigma(z) = \text{Nu}$, then $\text{qnf}(Y \frown z) = \text{qnf}(Y) \frown \sigma(z)$.

This is the syntactical formulation of the operation performed by a counter. The last digit continually progresses through the cyclic order—note that (2a) and (2b) both terminate in $\sigma(z)$—but if the last digit becomes Nu, it is necessary to perform the counting operation on the expression formed by discarding the last digit; this is the carry. If when the last digit is discarded nothing remains, a new $\sigma(\text{Nu})$ (corresponding to '1') is prefixed. This assumes, of course, that the counter has unlimited capacity. This transformation rule is a recursive operation across the digits, from right to left.

On the basis of the above definition, we can prove

Theorem 1.

$$\text{If } \overset{n}{\underset{i=m}{\frown}} x_i \neq \overset{n}{\underset{i=m}{\frown}} \sigma^{-1}(\text{Nu}), \qquad n \geq m$$

then $\overset{n}{\underset{i=m}{\frown}} y_i = \text{qnf} \overset{n}{\underset{i=m}{\frown}} x_i$ if and only if

(1) $[y_i = \sigma(x_i)] \mathrel{<=>} (k)\,[(m \leq k < i) \Rightarrow (x_k = \sigma^{-1}(\text{Nu}))]$,

(2) $[y_i = x_i] \mathrel{<=>} (\exists k)\,[(m \leq k < i)\,.\,(x_k \neq \sigma^{-1}(\text{Nu}))]$.

If $\beta = 2$, then $x_i \neq \sigma^{-1}(\text{Nu})$ if and only if $x_k = \text{Nu}$, and $x_k = \sigma^{-1}(\text{Nu})$ if and only if $x_k = \sigma(\text{Nu})$ and hence lines (1) and (2) of Theorem (1) specialize to:

(1a) $[y_i = \sigma(x_i)] \mathrel{<=>} (k)\,[(m \leq k < i) \Rightarrow (x_k = \sigma(\text{Nu}))]$,

(2a) $[y_i = x_i] \mathrel{<=>} (\exists k)\,[(m \leq k < i)\,.\,(x_k = \text{Nu})]$.

The successor of a numerical expression X is simply $\text{qnf}(X)$, and it can be shown that the qnf operation, thus restricted, satisfies all of Peano's axioms. The theory of operations on natural numbers can be constructed from this transformation rule.

So far we have not considered the physical nature of the characters, nor how they are physically strung together to form expressions; we have been considering questions of axiomatic syntax. Suppose we have a language in which $\beta = 2$, and the characters are represented by two conditions of potential at a point ξ in an electric circuit. Suppose $\xi(t)$ is the proposition: the point ξ is at the higher of the two possible potentials at time t. We define $x_t = \text{Nu}$ and $x_t = \sigma(\text{Nu})$ (there are only two characters) as follows:

Definition 2.

$$[x_t = \text{Nu}] \mathrel{<=>} \sim \xi(t),$$
$$[x_t = \sigma(\text{Nu})] \mathrel{<=>} \xi(t).$$

We are now dealing with physical syntax, since the physical nature of the characters comes into the picture. Since the characters must have unique immediate predecessors and successors, it will be necessary to quantize our time scale. In a synchronous machine, we assume that t increases by constant increments. Suppose we have a black box with two inputs and two outputs (Fig. 1). The box has the property that η assumes the high potential if and only if ξ and α are at different potentials, and λ assumes the high potential if and only if ξ and α are both at the high potential. Such a device can be synthesized by methods developed by Burkhart and Kalin, to be described in a forthcoming Harvard Computation Laboratory publication.[4] As soon as we identify the states of the calculating mechanism with the characters of the object language, analysis and synthesis of its behavior merges with the discipline of physical syntax.

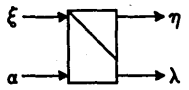


FIG. 1. A half adder.

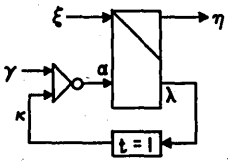


FIG. 2. Half adder adapted to perform the qnf transformation.

To return to the "black box," which is known as a half adder, its physical behavior is described by:

$$\eta(t) \Longleftrightarrow \sim [\xi(t) \Longleftrightarrow \alpha(t)], \tag{A}$$

$$\lambda(t) \Longleftrightarrow [\xi(t) \,.\, \alpha(t)]. \tag{B}$$

We define the character y_t by

Definition 3.

$$[y_t = \mathrm{Nu}] \Longleftrightarrow \sim \eta(t),$$

$$[y_t = \sigma(\mathrm{Nu})] \Longleftrightarrow \eta(t).$$

Definitions 2 and 3 and statements (A) and (B) give:

$$\alpha(t) \Longleftrightarrow [y_t = \sigma(x_t)], \tag{C}$$

$$\sim \alpha(t) \Longleftrightarrow [y_t = x_t], \tag{D}$$

$$\lambda(t) \Longleftrightarrow [(x_t = \sigma(\mathrm{Nu})) \,.\, \alpha(t)]. \tag{E}$$

In other words, if α is at the high potential (usually a positive pulse), then the box carries out the cyclic-successor transformation on each character, but if α is at the low potential, then the box performs the identity transformation.

A delay is now inserted in the circuit (Fig. 2) and the inputs to α are so connected that $\alpha(t) \Longleftrightarrow [\gamma(t) \vee \kappa(t)] \,.\, \gamma(t) \Longleftrightarrow t = (0)$, and $\kappa(t) \Longleftrightarrow \lambda(t-1)$. Now,

if $$\overset{n}{\underset{i=0}{\frown}} x_i \neq \overset{n}{\underset{i=0}{\frown}} \sigma(\mathrm{Nu}),$$

then $$\overset{n}{\underset{i=0}{\frown}} y_i = \mathrm{qnf} \overset{n}{\underset{i=0}{\frown}} x_i.$$

First, we note that for $t \geq 0$, $\lambda(t) <==> (t_k)\,[(0 \leq t_k \leq t) ==> \xi(t_k)]$.
This is demonstrable by an inductive argument. Hence:

$$\alpha(t) <==> (t_k)\,[(0 \leq t_k < t) ==> (x_{t_k} = \sigma(\mathrm{Nu}))]$$

and

$$\sim \alpha(t) <==> (\exists t_k)\,[(0 \leq t_k < t)\,.\,(x_{t_k} = \mathrm{Nu})].$$

Replacing $\alpha(t)$ and $\sim \alpha(t)$ in (C) and (D) by their equivalents given above, we see by Theorem 1 (for $\beta = 2$) that this circuit performs the qnf transformation.

Another transformation rule will be described. All systems for expressing integers, even Roman numerals, must have a qnf transformation rule, but complementation is characteristic of algorithmic systems, and has more syntactical than mathematical significance.

Definition 4.

$$\beta \text{ comp } (\sigma^r(\mathrm{Nu})) = \sigma^{(\beta-1)-r}(\mathrm{Nu}).$$

This defines the complement for a single digit. This is analogous (for $\beta = 10$) to the "nine's complement"; e.g., 10 comp ('0') = '9', 10 comp ('7') = '2'.

$$\beta \text{ comp } \left(\overset{n}{\underset{i=m}{\frown}} x_i\right) = \overset{n}{\underset{i=m}{\frown}} \beta \text{ comp } (x_i).$$

This definition extends β comp to expressions; β comp is a "linear operator" with respect to "$\frown$" and is thus more simply mechanized than the so-called ten's complement; there is no interaction between characters.

The black box just described can also perform the complement transformation. The peculiar property of binary systems that makes this possible is given by

Theorem 2.

If $\beta = 2$, $\sigma^{-1}(x) = \sigma(x) = 2 \text{ comp } (x)$.

From Theorem 2 and statements (C) and (D) we obtain:

$$\alpha(t) <==> [y_t = 2 \text{ comp } (x_t)], \qquad (F)$$

$$\sim \alpha(t) <==> [y_t = x_t]. \qquad (G)$$

If we supply clock pulses to α, the circuit complements; if not, then it gives the identity transformation.

Now consider the problem of forming the "ten's" complement. No special notation is needed, since it is simply qnf $(\beta \text{ comp } (x))$. The qnf indicates the necessity of carry mechanisms when a ten's complement is formed; qnf is not a "linear operator." From the previous analysis a complementer can be constructed by connecting two half adders in tandem (Fig. 3). All that is required is to supply clock pulses at δ and a starting pulse at $\gamma(t)$ simultaneous with the appearance of the least significant digit at ξ. There is a superior method, and the basis for it will now be derived.

The following theorem is easily proved from the preceding definitions and properties of cyclic order:

Theorem 3.

If $n \geq m$, then

(1) If $x_m \neq \text{Nu}$, $\text{qnf}\left(\beta \text{ comp } \left(\overset{n}{\underset{i=m}{\frown}} x_i\right)\right) = \beta \text{ comp } \left(\overset{n}{\underset{i=m-1}{\frown}} x_i\right) \frown \sigma(\beta \text{ comp } (x_m))$,

(2) If $x_m = \text{Nu}$, $\text{qnf}\left(\beta \text{ comp } \left(\overset{n}{\underset{i=m}{\frown}} x_i\right)\right) = \text{qnf}\left(\beta \text{ comp } \left(\overset{n}{\underset{i=m-1}{\frown}} x_i\right)\right) \frown x_m$.

From Theorem 1 follows

Theorem 4.

If $\overset{n}{\underset{i=m}{\frown}} x_i \neq \overset{n}{\underset{i=m}{\frown}} \text{Nu}$, $n \geq m$, then $\overset{n}{\underset{i=m}{\frown}} y_i = \text{qnf}\left(\beta \text{ comp } \left(\overset{n}{\underset{i=m}{\frown}} x_i\right)\right)$ if and only if:

(1) $[y_i = \sigma(\beta \text{ comp } x_i)] \langle\Longrightarrow\rangle (k)\,[(m \leq k < i) \Longrightarrow (x_k = \text{Nu})]$,

(2) $[y_i = \beta \text{ comp } x_i] \langle\Longrightarrow\rangle (\exists k)\,[(m \leq k < i) \Longrightarrow (x_k \neq \text{Nu})]$.

If $\beta = 2$ these two lines specialize to:

(1*a*) $[y_i = x_i] \langle\Longrightarrow\rangle (k)\,[(m \leq k < i) \Longrightarrow (x_k = \text{Nu})]$,

(2*a*) $[y_i = 2 \text{ comp } x_i] \langle\Longrightarrow\rangle (\exists k)\,[(m \leq k < i) \,.\, (x_k = \sigma(\text{Nu}))]$.

This theorem justifies the following ingenious circuit, invented by T. C. Chen.

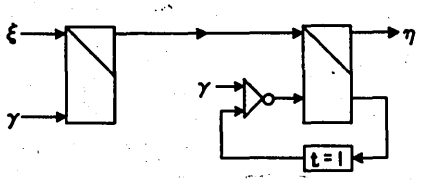


FIG. 3. A complementer.

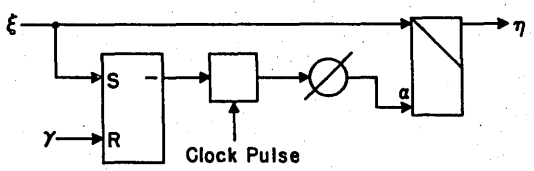


FIG. 4. Circuit for transformation corresponding to the two's complement.

The half adder is connected to a flip-flop as in Fig. 4. The gate-inverter combination serves to transform the static output of the flip-flop to a string of positive pulses, in order to permit a.c. coupling in the half adder. It will be noted that the positive pulses appear at α if and only if the flip-flop is set. The flip-flop is assumed to require one unit of time to change its state, and the condition of the circuit immediately preceding the arrival of the character x_0 is $\sim \xi(-1)$ and $\gamma(-1)$ and we require that $\gamma(t) \langle\Longrightarrow\rangle t = -1$. The circuit condition at α is then, for $t \geq 0$:

$$\alpha(t) \langle\Longrightarrow\rangle (\exists t_k)\,[(0 \leq t_k < t) \,.\, \xi(t_k)],$$
$$\dot{\sim} \alpha(t) \langle\Longrightarrow\rangle (t_k)\,[(0 \leq t_k < t) \Longrightarrow \sim \xi(t_k)].$$

Replacing $\xi(t)$ by its syntactical equivalent from Definition 2:

$$\alpha(t) \langle\Longrightarrow\rangle (\exists t_k)\,[(0 \leq t_k < t) \,.\, x_{t_k} = \sigma(\text{Nu})],$$
$$\sim \alpha(t) \langle\Longrightarrow\rangle (t_k)\,[(0 \leq t_k < t) \Longrightarrow (x_{t_k} = \text{Nu})].$$

Combining this with statements (*F*) and (*G*) we obtain

$$[y_t = 2 \text{ comp } (x_t)] \langle\Longrightarrow\rangle (\exists t_k)\,[(0 \leq t_k < t) \,.\, (x_{t_k} = \sigma(\text{Nu}))],$$
$$[y_t = x_t] \langle\Longrightarrow\rangle (t_k)\,[(0 \leq t_k < t) \Longrightarrow (x_{t_k} = \text{Nu})].$$

Thus Theorem 4 applies and we have proved that the circuit carries out the transformation rule (for $\beta = 2$) corresponding to the two's complement, provided that $\overset{n}{\underset{i=m}{\frown}} x_i \neq \overset{n}{\underset{i=m}{\frown}} \mathrm{Nu}$.

The circuits analyzed are rather elementary, but these methods provide a link between the physical properties of the equipment components and the object language, as well as a method of describing and analyzing the interrelations between the transformation rules of the object language. This method is capable of being extended to cover more complex situations which are at present difficult to investigate except by our sometimes fallible intuition.

Special Notation

$X^\frown Y$, the expression formed by adjoining the expression Y to the expression X.

$\overset{n}{\underset{i=m}{\frown}} x_i$, the expression formed by adjoining the characters $x_n, x_{n-1}, \cdots, x_{m+1}, x_m$ together (in that order).

$\sigma(x)$, the cyclic successor of the character x.

$\sigma^r(x)$, the rth iteration of the cyclic successor operator.

$<==>$, if and only if.

(k), for every (integer) $k_1 \cdots$.

$==>$, if, then.

$(\exists k)$, for some (integer) $k_1 \cdots$.

$\sim$ not $\cdots$.

. $\cdots$ and $\cdots$.

$\vee$ $\cdots$ or $\cdots$ (or both).

Lower-case letters are character variables; upper-case letters are expression variables; lower-case Greek letters are statement variables referring to voltage levels.

REFERENCES

1. R. Carnap, *Introduction to semantics* (Harvard University Press, 1942), p. 1.

2. F. Klein, *Elementary mathematics from an advanced standpoint*, vol. 1, *Arithmetic, algebra, analysis*, tr. by E. R. Hedrick and C. A. Noble, pp. 21–22.

3. R. Carnap, *The logical syntax of language* (Harcourt, Brace, 1937), p. 1.

4. Staff of the Computation Laboratory, *Synthesis of electronic computing and control circuits* (Harvard University Press, 1951).

FOURTH SESSION

Wednesday, September 14, 1949
2:00 P.M. to 5:00 P.M.

NUMERICAL METHODS

Presiding

Raymond C. Archibald
Brown University

Notes on the Solution of Linear Systems involving Inequalities

GEORGE W. BROWN

Rand Corporation

Consider the problem of minimizing a linear function $\Sigma b_j x_j$ subject to the conditions[1]

$$\sum_j A_{ij} x_j \geq c_i, \qquad i = 1, 2, \cdots, m_1$$
$$x_j \geq 0. \qquad j = 1, 2, \cdots, m_2$$

Notice at the outset that equalities may be admitted in this form by writing each equality as two inequalities with reversal of signs. Furthermore, the problem may be reformulated so that only inequalities of the form $x_j \geq 0$ are present, by defining appropriate new variables. Thus it is evident that the above form is simply one standard version of a general problem involving both inequalities and equalities.

In principle the solution of the problem stated is trivial. Observe that the set of inequalities defines in m_2-space a convex polyhedron (possibly empty) with at most $m_1 + m_2$ faces of dimension $m_2 - 1$, and that the minimum problem is that of finding an extreme point of the polyhedron in some direction. In general, the extremum will be taken on at a vertex, so the problem is that of evaluating $\Sigma b_j x_j$ at the vertices and choosing that vertex which yields the smallest value. A vertex is of course a point at which a subsystem (of rank m_2) of the inequalities is satisfied exactly as equalities, with the remaining inequalities satisfied. In principle, then, one could invert all subsystems of rank m_2, throwing out those whose solutions fail to satisfy the remaining inequalities, and then evaluate $\Sigma b_j x_j$. It is clear that this is not a practical method beyond the smallest values of m_1 and m_2. The practical difficulties stem from the fact that the convex polyhedron is specified by its faces, whereas the vertices are at the root of the problem.

In passing, it should be noted that the problem stated above has a very simple dual problem, obtained by transposing the matrix A, and making a few other obvious changes. The dual is the problem of maximizing $\Sigma c_i y_i$ subject to the conditions

$$\sum_i y_i A_{ij} \leq b_j, \qquad i = 1, 2, \cdots, m_1$$
$$y_i \geq 0. \qquad j = 1, 2, \cdots, m_2$$

The two dual problems have the property that if either problem has a solution so has the other, and the minimum value in one is the maximum value in the other. In certain economic applications the solutions of both problems are required.

Consider now the problem of maximizing $\min_j \Sigma \xi_i A_{ij}$ subject to $\xi_i \geq 0$, $\Sigma \xi_i = 1$, $i = 1, 2, \cdots, m_1$; and the dual problem of minimizing $\max_i \Sigma A_{ij} \eta_j$ subject to $\eta_j \geq 0$, $\Sigma \eta_j = 1$,

$j = 1, \cdots, m_2$. This problem provides optimum mixed strategies for the zero-sum game with matrix A, where A_{ij} represents the payment from player 1 to player 2, if player 1 plays his ith strategy and player 2 plays his jth strategy. The celebrated minimax theorem of von Neumann says that under the conditions stated

$$\operatorname*{Max}_{\xi} \operatorname*{Min}_{j} \Sigma \xi_i A_{ij} = \operatorname*{Min}_{\eta} \operatorname*{Max}_{i} \Sigma A_{ij} \eta_j.$$

The common value is referred to as the value of the game and the $\{\xi_i\}$ and $\{\eta_j\}$ of the solutions are the optimum mixtures for players 1 and 2, respectively. As in the first problem stated in this paper, geometrical considerations of convex bodies contribute to an understanding of the problem, and it turns out that in general the problem is practically solved if it is known which submatrix of A to invert.

There is of course an intimate relation between the theory-of-games problem and the problem first stated, although they are not quite identical problems, since the game problem always has a solution, while the first problem does not necessarily. To summarize briefly, the game problem is directly a special case of the first problem, while the first problem can always be embedded in a game problem, whose solutions yield solutions to the original problem if it has a solution. Thus, if problems of one type can be solved, so can problems of the other type.

Various iterative methods for solution of one or the other of these problems have been given by von Neumann, Dantzig, and others. While some of these methods may be practical over a certain range of problems, all of them have an apparent dependence, in required number of steps, of higher order than the first power of the linear dimensions of the problem. For very large matrices not possessing simplifying special properties, such a dependence can be a very serious obstacle in the way of getting numerical solutions. We will describe briefly, for the game solution, an iterative scheme which is quite different from those previously suggested, in that the amount of calculation required at each iterative step is directly proportional to the linear dimensions of the problem, so that the method has, a priori, some chance of beating the high-order dependence.

The procedure to be described can most easily be comprehended by considering the psychology of, let us say, a statistician unfamiliar with the theory of games. Such a person, faced with repeated choices of play of a certain game, might reasonably be expected to play, at each opportunity, that one of his strategies which is best against past history, that is, against the mixture constituted by his opponent's plays to date. Such a decision utilizes information of the past in the most obvious manner. The iterative scheme referred to here is based on a picture of two such statisticians playing repeatedly together. For purposes of calculation a slight modification is introduced which has the effect that the two players choose alternately, rather than simultaneously.

Restating the method algebraically, let A be the game matrix, let i_n and j_n be the nth choices of strategy for the two sides, and let $\xi_i^{(n)}$ and $\eta_j^{(n)}$ be the relative frequencies of strategies i and j in $(i_1, i_2, \cdots, i_n)$ and $(j_1, j_2, \cdots, j_n)$, respectively; then j_n minimizes $\Sigma_i \xi_i^{(n)} A_{ij}$ and

i \ j	1	2	3
1	3	1.1	1.2
2	1.3	2	0
3	0	1	3.1
4	2	1.5	1.1

FIG. 1. A 4 × 3 matrix.

n	i_n	j=1	j=2	j=3
1	2	1.3	2	0
2	3	1.3	3	3.1
3	1	4.3	4.1	4.3
4	1	7.3	5.2	5.5
5	1	10.3	6.3	6.7
6	4	12.3	7.8	7.8
7	2	13.6	9.8	7.8
8	3	13.6	10.8	10.9
9	4	15.6	12.3	12.0
10	3	15.6	13.3	15.1
11	3	15.6	14.3	18.2
12	3	15.6	15.3	21.3
13	2	16.9	17.3	21.3
14	4	18.9	18.8	22.4
15	4	20.9	20.3	23.5
16	2	22.2	22.3	23.5
17	4	24.2	23.8	24.6
18	2	25.5	25.8	24.6
19	4	27.5	27.3	25.7
20	4	29.5	28.8	26.8
21	3	29.5	29.8	29.9
22	1	32.5	30.9	31.1
23	4	34.5	32.4	32.2
24	3	34.5	33.4	35.3
25	4	36.5	34.9	36.4

$\underline{V}_n$	V_n
0	3.1
.65	2.1
1.37	1.77
1.30	1.60
1.26	1.52
1.30	1.55
1.11	1.46
1.35	1.46
1.33	1.59
1.33	1.53
1.30	1.48
1.28	1.44
1.30	1.48
1.34	1.49
1.35	1.51
1.39	1.52
1.40	1.52
1.37	1.49
1.35	1.47
1.34	1.48
1.40	1.49
1.40	1.48
1.40	1.47
1.39	1.47
1.40	1.47

j_n	i=1	i=2	i=3	i=4
3	1.2	0	3.1	1.1
1	4.2	1.3	3.1	3.1
2	5.3	3.3	4.1	4.6
2	6.4	5.3	5.1	6.1
2	7.5	7.3	6.1	7.6
2	8.6	9.3	7.1	9.1
3	9.8	9.3	10.2	10.2
2	10.9	11.3	11.2	11.7
3	12.1	11.3	14.3	12.8
2	13.2	13.3	15.3	14.3
2	14.3	15.3	16.3	15.8
2	15.4	17.3	17.3	17.3
1	18.4	18.6	17.3	19.3
2	19.5	20.6	18.3	20.8
2	20.6	22.6	19.3	22.3
1	23.6	23.9	19.3	24.3
2	24.7	25.9	20.3	25.8
3	25.9	25.9	23.4	26.9
3	27.1	25.9	26.5	28.0
3	28.3	25.9	29.6	29.1
1	31.3	27.2	29.6	31.1
2	32.4	29.2	30.6	32.6
3	33.6	29.2	33.7	33.7
2	34.7	31.2	34.7	35.2
2	35.8	33.2	35.7	36.7

FIG. 2. Cumulative payoffs.

i_{n+1} maximizes $\Sigma_j A_{ij}\eta_j^{(n)}$. This process defines a sequence $i_1, j_1, i_2, j_2, \cdots$, once i_1 is chosen (perhaps arbitrarily), except for possible ambiguities of the extrema. Any convenient rule will do for handling ambiguities. If $\underline{V}_n = \min_j \Sigma\xi_i^{(n)}A_{ij}$ and $\overline{V}_n = \max_i \Sigma A_{ij}\eta_j^{(n)}$, it is easily seen that $\underline{V}_n \leq V \leq \overline{V}_n$, where V is the value of the game. The mixtures $\{\xi_i^{(n)}\}$ and $\{\eta_j^{(n)}\}$ are mixed strategies, and the corresponding $\underline{V}_n$ and $\overline{V}_n$ are the most favorable outcomes ensured to each player if he uses the corresponding mixture.

At this moment not much is rigorously established about the properties of this iteration, except that if it converges at all it converges to a solution of the game for each side. Of course it would be sufficient if lim sup $\underline{V}_n$ = lim inf $\overline{V}_n$. There is considerable support, however, based on experience with the method, and also on the study of a related system of differential equations, for the conjecture that convergence is of the order of $1/n$ and does not depend essentially on the size of the matrix. If this is so, it is extremely important for the solution of large matrices, by virtue of the fact that each iterative steps requires only a number of operations proportional to the linear size of the matrix. Convergence of order $1/n$ is of course painful if high accuracy is needed. In such cases it may be possible, however, to use a method like this to get close to the solution, finishing with one step of another iteration.

Figure 2 is a worksheet showing 25 steps carried out for the 4×3 matrix given in Fig. 1. Note that each line is obtained by adding to the previous line, component by component, the corresponding row or column of the matrix, without troubling to divide by n. The $\underline{V}_n$ and $\overline{V}_n$ were calculated at each step, by division of the extrema by n, to show the progress of the calculation. In case of ties the lowest index was taken. Note particularly that $\overline{V}_n - \underline{V}_n$ is decreasing just about like $1/n$, in spite of the excursions which $\overline{V}_n$ and $\underline{V}_n$ make. The initial choice of $i_1 = 2$ was made deliberately as an unfavorable choice, with respect to minimum guaranteed payoff.

It is appropriate to report to this Symposium that preliminary discussions with Messrs. Harr and Singer, of the staff of the Harvard Computation Laboratory, indicate that Mark III could carry out 1000 of these iterative lines for a 40×40 matrix in comfortably under an hour. Of course the problem has not been completely programmed, but the estimate is believed to be conservative.

REFERENCE

1. The theoretical background of this paper is based on work of H. Weyl, von Neumann, Ville, Tucker, G. Dantzig, and others, on convex polyhedra and on the theory of games.

Mathematical Methods in Large-scale Computing Units

D. H. LEHMER

University of California

The title of this paper covers such a vast subject that it will be impossible to do it justice. In fact, this title might well have been chosen as that of the whole session. My aim is merely to discuss in a general way certain features of the mathematics that is characteristic of the large-scale computing unit. In pointing up these general remarks I shall discuss in considerable detail only one problem. Further illustrations will be contained, no doubt, in the other papers of this session.

The mathematical methods available to a computing unit depend of course on the versatility of the unit. Nearly all units can perform addition, subtraction, multiplication, and division. The advent of large-scale digital computers has added a fifth operation of considerable importance, namely, discrimination. This, in general terms, is the operation of making a choice of one of several branches of a program (or course of procedure), depending on the outcome of a previous calculation. This operation is peculiar to discrete-variable machines, since its outcome is not continuous. The purely analogue machine cannot distinguish the larger of two sufficiently small numbers, or determine the sign of either. This fact was recognized early in the construction of roulette wheels. By converting the wheel into a discrete-variable device countless arguments were avoided.

The ability of a discrete-variable machine to discriminate and thus to decide for itself what course of action to take has led to the popular misconception that such machines think or even have the ability to learn from experience.

Various criteria are employed in discrimination. Decisions are made according to whether:

(1) A given number is $\geqslant 0$ or < 0 (Harvard Mark I);
(2) A given number is 0 or not (ENIAC);
(3) A given number is odd or even (ENIAC);
(4) A given sequence of numbers has one sign pattern or another (Bell Telephone);
(5) The sum of two numbers exceeds the capacity of the machine or not (Zephyr);
(6) A given number belongs to one of a set of residue classes with respect to a given modulus (Electronic sieve).

These criteria are not independent and others can be constructed from them. All digital machines are capable of some form of discrimination and those named above are given only as examples.

The mathematical methods that call for much discrimination are very frequently iterative

ones. Here discrimination is used to decide whether or not to continue to iterate. Another simple use of discrimination is in forming the nonanalytic function $|x|$. More elaborate uses arise in the step-by-step solution of differential equations and of course still more in problems of combinatorial analysis and number theory. Incidentally, the electronic sieve is designed to make 10 million discriminations per second.

Another feature of mathematical methods that are being used in large-scale computing is that they tend to eliminate the elaborate formulas and to introduce instead what might be called combinatorial complexities. This is due for the most part to the high speed of operation. For instance, in using a quadrature formula for numerical integration it does not pay to use the accurate Weddle's rule; it is often simpler and even faster to employ the crude trapezoidal rule. As far as I know the superb method of Gauss for mechanical quadrature has never been used in large-scale work. The method of Heun is used much more frequently than the more accurate and complex Runge-Kutta method for the step-by-step solution of ordinary differential equations. Minima of functions are found by extensive numerical trial-and-error methods, rather than by the somewhat more sophisticated and traditional method of setting derivatives equal to zero and solving. Systems of many first-order differential equations are solved in lieu of single differential equations of high order. A large number of trial solutions of differential equations with one-point boundary conditions may be made in order to obtain a single solution of a two-point boundary problem. Solving problems in terms of special functions is passé; finite-difference methods are used instead. The power-series expansions of analytic functions are being used to a large number of terms and to a great accuracy in order to avoid the use of alternative asymptotic expansions.

All these examples show how mathematical subtleties are being replaced by stepped-up numerical activities. To make this replacement possible the operator naturally must surrender much of his control to the machine itself. He simply cannot follow the course of the numerical work with sufficient rapidity to make on-the-spot decisions as to what to do next. This means that the programmer may have to incorporate a large number of discriminations or branches in the program of the problem. Much has been said, but little written, about the logic or even the topology of programming. Logicians and topologists are not coming to the rescue of the desperate programmer. General rules for programming have been discovered. Most of them have been used in the Kansas City freight yards for a long time. This is the combinatorial complexity to which I have referred. Flow diagrams showing the routines, subroutines, and other wheels within wheels are hardly distinguishable from the block diagrams of the machine itself; the latter, however, are made once and for all. This then is the white man's burden of large-scale computing.

The third characteristic feature of discrete-variable methods is the possibility of introducing number theory into what at the outset appears to be a problem in continuous functions. By way of illustration, let me call attention to a method which is the subject of the last paper

of this session—the Monte Carlo method. In this method it is necessary to produce random variables. The problem here is not one of producing a table of random digits to be published and used by others. On the contrary, one can think ideally of a perfect stream of these random numbers produced at high speed by the machine and passing by a "gate." Whenever the computer needs a number it opens the gate and takes one. More explicitly, we might list the following desiderata:

(1) An unlimited sequence of randomly arranged eight-digit numbers;

(2) A simple process by which the machine may produce the sequence:

$$u_n = f(u_{n-1}, u_{n-2}, \cdots, u_{n-k});$$

(3) Immediate access by the machine to the current number of the sequence whenever necessary; of course, the whole sequence need not be retained in the machine.

If we examine these desiderata we see at once that they are inconsistent. In the first place, the number of eight-digit numbers is not unlimited. There are in fact only 100 million of them. Secondly, condition (2) forces the sequence to be ultimately periodic and therefore not random. We therefore scale down our demands and modify (1) and (2) to read:

(1) Millions of pseudo-random numbers u_n;

(2) $u_n = f(u_{n-1})$, f a simple function.

A pseudo-random sequence is a vague notion embodying the idea of a sequence in which each term is unpredictable to the uninitiated and whose digits pass a certain number of tests traditional with statisticians and depending somewhat on the uses to which the sequence is to be put. The worst possible departure from randomness is to have the period of the sequence small or equal to one. In constructing the function f, therefore, it is of the utmost importance to obtain one that produces a guaranteed proper period of immense length.

A method already in use on the ENIAC, due to von Neumann and Metropolis, is the following. Let u_0 be an arbitrary initial eight-digit number. Then u_1 is defined as the central block of eight digits in the square of u_0, and u_2 is defined as the same function of u_1 that u_1 is of u_0, etc. At first sight this would appear to give an ideal source of random numbers. Certainly is produces an unpredictable sequence of numbers. However, as has been pointed out already by several writers, this process cannot be expected to give random numbers. In fact, one must expect to obtain numbers u_n of the form $xyzw0000$ before many more than 10,000 numbers u_n are generated. When this happens, either $w = 0$ and $u_{n+1} = 00000000$ and all succeeding u's vanish, or $w \neq 0$ and all succeeding u's are of the form $x'y'z'w'0000$, where $w' = 1, 5$, or 6. Hence periodicity will set in in fewer than 3000 more steps. Also, one must expect to obtain numbers of the form $u_n = 0000xyzw$, in which case $u_{n+17}, u_{n+18}, \cdots$, all vanish. Thus it is seen that this process cannot be recommended as a source of random digits. It has an additional drawback in that it ties up the multiplier, which is a fairly busy component of any machine.

If we look at the problem from the standpoint of the theory of numbers, it is not difficult

to find a more satisfactory solution. We may proceed as follows. We begin as before with an arbitrary nonzero initial eight-digit number u_0. Next we compute $23u_0$. In general this will be a ten-digit number. The ninth and tenth digits (counting from the right) are now removed and subtracted from the remaining eight-digit number. This produces u_1. The next number u_2 is produced from u_1 in the same way. To illustrate in detail, suppose that the initial u_0 is 47594118 (chosen at random from a wastepaper basket of punched cards). Then

$$\begin{aligned} 20u_0 &= \ 9\ 51882360 \\ 3u_0 &= \ 1\ 42782354 \\ \hline 23u_0 &= 10\ 94664714 \\ \text{subtract } 10 &\qquad\quad -\ 10 \\ \hline u_1 &= \qquad 94664704 \end{aligned}$$

As in the first method, this process is necessarily ultimately periodic. In fact, it is actually periodic of period 5882352. This fact makes all the difference; the reason for it is simple. In computing u_n from u_{n-1} we are computing the remainder of $23u_{n-1}$ on division by $10^8 + 1$. Hence, in congruence notation

$$u_n \equiv u_0\, 23^n \pmod{10^8 + 1}.$$

By the theory of the binomial congruence, u_n is periodic of period 5882352 since $10^8 + 1 = 17 \cdot 5882353$. The number 23 is the best possible choice in the sense that no other number produces a longer period, and no smaller number produces a period more than half as long.

As set up for the ENIAC, for example, the process would tie up only two accumulators and would produce 5000 pseudo-random digits per second. The process would have a period of 2 hr 36 min 52 sec.

Whether such a set of digits or a reasonable subset satisfies the statisticians' tests for randomness is of course another question. To investigate this matter I have secured the kind coöperation of Professor L. E. Cunningham of the Astronomy Department of the University of California, who set up the calculation on the IBM calculating punch 602A. This produced the first 5000 u's (that is, 40,000 digits in all) in about 4 hr (the ENIAC would be faster by a factor of 1800). One of the secondary reasons for making the calculation was to test the accuracy of the 602A. It may distress some and surprise others to know that any isolated u_n can be computed on a desk calculator in 3 min. Thus u_{5000} was known in advance. The fact that this value agreed with the result obtained in 5000 steps is a rigorous check of the arithmetic unit of the 602A.

Once produced, the results were subjected to four standard tests with the assistance of Dr. Evelyn Fix and other members of the Statistical Laboratory of the University of California. All four tests were passed successfully. In case anyone is convinced that the numbers u_n are really random, I should like to call his attention to the fact that each number u_n is a multiple of 17.

For a binary machine a similar process can be set up with respect to a modulus of the form

$2^n \pm 1$. For example, with the Mersenne prime $2^{31} - 1$ as modulus, more than 66 billion pseudo-random binary digits can be generated.

The method is based on the function $f(x) = ax$. A very little extra complication would be produced by using the general linear function $f(x) = ax + b$. However, nothing is gained by this generalization, since the period is independent of b, as one can see from the theory of difference equations. With machines capable of parallel operation like the ENIAC and SSEC, the above process is especially advantageous for two reasons: (*a*) it can be incorporated into the program with very little expense, and (*b*) the "gating" of this routine at irregular time intervals serves further to randomize the sequence u_n. The serial-type machine would simply use the numbers u_m one after another.

I have gone into such great detail on this problem just to indicate how a problem in, let us say, nuclear physics, when attacked by a large-scale computing unit, can involve a mathematical method taken from the impractical theory of numbers.

It is only fair to point out that, conversely, large-scale digital equipment can be used to study certain problems in the theory of numbers. In fact, it is sometimes a little exasperating for the number theorist to assist the applied mathematician in juggling round-off errors, truncating errors and a flitting decimal point in order to adapt a problem in fluid mechanics to a discrete-variable machine when all the time the machine, being digital, is all ready to work on clean-cut problems involving whole numbers. However, I realize that this exasperation is shared by very few present. Most of you will be relieved to know that, to the best of my knowledge, very little valuable time on large-scale computing units has been spent on such unprofitable problems.

In fact, to date, only one small problem of this sort has been solved and published, and another is making slow progress. However, I hear that the University of Manchester's new computing machine is being used on such problems and doubtless there will be some interesting results published before long.

It may not be out of place to mention certain kinds of problems for which no mathematical method would seem to be available in order to apply large-scale computing units in a practical way. By an impractical application we mean one that produces results no more rapidly than a few hand computers using desk calculators. Since the large-scale machines are based on the four rational operations they have good control over functions that are defined by algebraic expressions. However, mathematics abounds with functions that are defined verbally, often in some negative way. Such functions are apt to give trouble if they cannot be expressed directly in terms of operations with which the machine is familiar. Simple examples of such functions occur in the theory of numbers, algebra, topology, statistics, organic chemistry, genetics, and elsewhere. Often these functions are of the enumerative sort. For example, one can ask for the number of nonequivalent maps of 135 countries, or the number of different ways that each map can be colored in five colors. If ten permutations of the digits 0, 1, · · ·, 9

are selected at random, what is the probability that they form a Latin square? If $a(n)$ denotes the number of prime factors of n, is the sum $\sum_{n=1}^{N} (-1)^{a(n)}$ negative for $1000 < N < 10^8$? It is perhaps best not to think of such questions, but to return with our all-purpose computing unit to our differential equations.

In conclusion a few words should be said about the possible future influence of large-scale computing units on mathematics and mathematicians. To quote an eminent physicist: "With the developments of greater capacity and speed it is almost certain that new methods will have to be developed in order to make the fullest use of the capabilities of this new equipment. It is necessary not only to design machines for the mathematics, but also to develop a new mathematics for the machines."

In one sense a new mathematics is arising from the development of these machines. I refer to the theory of programming and coding in all its aspects. Here we have developed a nomenclature and a mass of symbolism which belong properly to a small corner of symbolic logic. This is not the new mathematics of the quotation, however. Personally, I do not look for much really new mathematics as a result of the machine development. Of course there will be new interest in and new emphasis on old and recent mathematics. Processes which the mathematicians have been writing about but not carrying out will become realities. However, few mathematicians have ever been stopped by the fact that they could not carry out the operations that they contemplate. Long ago the mathematician broke through the restricting boundary of things that were practical. In this respect the mathematician is far ahead of the existing machines and will doubtless continue to be so.

There is no doubt that these new machines are creating new service jobs for mathematicians, young and old. However, it seems to me, the most important influence of the machines on mathematics and mathematicians should lie on the opportunities that exist for applying the experimental method to mathematics. Much of modern mathematics is being developed in terms of what can be proved by general methods rather than in terms of what really exists in the universe of discourse. Many a young Ph.D. in mathematics has written his dissertation about a class of objects without ever having seen one of the objects at close range. There exists a distinct possibility that the new machines will be used in some cases to explore the terrain that has been staked out so freely and that something worth proving will be discovered in the rapidly expanding universe of mathematics.

Empirical Study of Effects of Rounding Errors

C. CLINTON BRAMBLE

U.S. Naval Proving Ground, Dahlgren, Virginia

Among the tacit assumptions usually made in computation are, first, that the arithmetical processes involved are considerably diversified, and second, that the arithmetical processes produce results in such a way that the occurrence of the different digits is equally likely. In this paper I wish to present evidence that these conditions are not, in general, realized and to exhibit some results of actual computation which indicate that, after all, the situation is not too bad.

If it is assumed that the digits 0 to 9 occur with the same frequency when they appear as units digits of numbers that enter into computation—that is, that their distribution function is rectangular—then it follows that the distribution function of the units digits of the sum of two numbers is also rectangular and the digits 0 to 9 are again equally likely. As a consequence of this, we can say that the units digit of an extended sum is as likely to be one digit as another. This premises also the idea that the numbers that enter into the sum are uncorrelated. It is clear that we can violate this premise if we add numbers identical or correlated, such as $x + x$ or $x + x^2$, or if we subtract the same number from a given number repetitively. For instance, if we add numbers ending in 5 to a number x we will also get as successive sums numbers ending in x and $x + 5$; or if we subtract in succession a series of even numbers from an odd number, we will always get odd differences and it is clear that the different digits are under these circumstances not equally likely.

If we form a 0 to 9 multiplication table, we will find the distribution of the units digits of the 100 possible products of integers as follows:

0	1	2	3	4	5	6	7	8	9
27	4	12	4	12	9	12	4	12	4

This frequency table indicates a high frequency of 0's and a low frequency of odd integers as the units digits of products of integers. It is also noted that 75 percent of the units digits produced by multiplication are even and 25 percent are odd; also, that 32 percent of the digits are greater than 5 and 32 percent are less than 5 exclusive of 0, so that if we round numbers on 5 as a base, we expect the average of the positive errors to be the same as that of the negative ones.

If, however, a combination of integers of the form $ab + c$ is made from integers selected at random, an inspection of the results shows that the distribution function of units digits is again rectangular and all the digits from 0 to 9 have the same frequency. Addition is therefore a leveling process, while multiplication disturbs the uniformity of digit-distribution functions.

Suppose next we consider the product of two numbers, say $10a_1 + b_1$ and $10a_2 + b_2$, namely, $100a_1a_2 + 10(a_1b_2 + a_2b_1) + b_1b_2$. We note that, exclusive of the "carry" from b_1b_2, the tens digit is made up of a sum of two products, $a_1b_2 + a_2b_1$. An examination of this product shows that we will obtain an even number in those cases in which both products are even or in which both products are odd. As stated above, the first product will be even 3/4 of the time and the second product will be even 3/4 of the time; therefore the sum will be even under this contingency 9/16 of the time. In the same way we see that the two products will be odd $1/4 \times 1/4$ or 1/16 of the time. Hence the expression $a_1b_2 + a_2b_1$ in which a's and b's are the integers 0 to 9 with equal probability will be even 5/8 of the time. If we add another term to the product, as occurs in the hundreds digit of a product of two numbers of three or more digits, we have an expression of the form $a_1c_2 + a_2c_1 + b_1b_2$ and we find that this sum will be an even number in 9/16 of the possible cases. Extending this process, we find that the sum of n products, $\sum_{i=1}^{n} a_ib_i$, $i = 1, \cdots, n$, in which the numbers are not correlated, will be an even number in $1/2 + (1/2)^{n+1}$ of the cases. Thus it appears that there are components of the various digits, as we pass farther to the left in the formation of products, that tend toward the equality of the numbers of odd and even digits.

Next, let us look for the number of even digits in the tens place of all possible products of integers. Only the tens and units digits of the factors will affect this number and it will consist of the units digit of $(a_1b_2 + a_2b_1)$ plus the tens digit of b_1b_2. This will be referred to as the "tens digit without carry." To pursue this, let us form a table of the tens digits of all possible products of integers so that we will have available the numbers that are carried in forming products, together with their frequencies.

	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1	1	1	1	1
3	0	0	0	0	1	1	1	2	2	2
4	0	0	0	1	1	2	2	2	3	3
5	0	0	1	1	2	2	3	3	4	4
6	0	0	1	1	2	3	3	4	4	5
7	0	0	1	2	2	3	4	4	5	6
8	0	0	1	2	3	4	4	5	6	7
9	0	0	1	2	3	4	5	6	7	8

The frequencies of the integers are as follows:

0	1	2	3	4	5	6	7	8	9
42	17	13	9	9	4	3	2	1	0

We see, therefore, that an even number is carried 68 percent of the time and an odd number 32 percent of the time. An examination of individual cases shows that, considering the cases in which an even number is produced, there is no correlation between the oddness or evenness of the units digit of $a_1b_2 + a_2b_1$ and that of the carry from the product b_1b_2. Therefore, we can say that even tens digits will be produced in those cases in which the units digits of $a_1b_2 + a_2b_1$ is even and the carry is even and in those cases in which both of these are odd. Thus we have $5/8 \times 68$ percent plus $3/8 \times 32$ percent, or 0.545, as the fraction of cases in which the tens digit is expected to be an even number, and correspondingly in 0.455 of the cases an odd number. This shows again that even numbers tend to dominate in multiplication but that as we pass from right to left in products there appears to be a mixing process that tends toward a reduction in the excess of even numbers.

Consider an array of all possible products of the integers from 0 to 99. The distribution of the tens digits is given by the table:

0	1	2	3	4	5	6	7	8	9	even	odd
1210	900	1060	900	1060	950	1060	900	1060	900	5450	4550

We will find in this array 5450 numbers in which the tens digit is even. By means of an inspection based on enumerating all possible cases, it is seen that the number of even digits followed by 6, 7, 8, 9 as units digits is precisely the same as the number of odd digits so followed. Further, the number of even tens digits preceding a final 5 is the same as the number of odd ones. Thus it is clear that a rounding process based on 5 will create as many even numbers as odd numbers, since the even numbers followed by numbers greater than 5 will become odd and the odd numbers will become even. In this array there are 900 cases in which the tens digit is followed by a 5. A process of enumeration shows that 450 of these tens digits are even and 450 are odd. Thus we find that after rounding, if we round the numbers ending in 5 to an even digit we create 450 more even numbers and the resulting numbers will consist of 5900 even and 4100 odd. On the other hand, if we should agree that in the dubious case ending in 5 we should round to an odd number we will have 5000 even and 5000 odd. It may be of interest to point out that the complete frequency distribution for the tens digit of all possible products "without carry" is the following:

0	1	2	3	4	5	6	7	8	9	even	odd
1450	720	1200	720	1200	870	1200	720	1200	720	6250	3750

The distribution with carry in which the numbers ending in 5 are rounded to even numbers is:

0	1	2	3	4	5	6	7	8	9	even	odd
1320	800	1170	800	1120	900	1120	800	1170	800	5900	4100

It may be noted in particular in this distribution that the deviations from 1000 have become smaller, that is, each digit now has a frequency that is more nearly 10 percent than in the previous case. While in rounding in either case mentioned above it is expected that as large an amount will be dropped as is carried, rounding numbers ending in 5 to the even number tends toward a domination of even digits while the rounding to the odd number tends to restoring a balance between odds and evens. From this point of view it would appear to be preferable to round numbers ending in 5 to the odd digit. Furthermore, rounding to the odd digit never affects more than the next digit. The individual frequencies when we round the 5's to the odd tens digit are as follows:

0	1	2	3	4	5	6	7	8	9
1160	960	960	1010	960	1060	960	1010	960	960

The greater uniformity of occurrences of the different digits as a consequence of this rounding procedure is apparent.

EXAMPLES OF ERRORS IN ACTUAL COMPUTATION

(1) Suppose we wish to compute the successive values of $\pi/2 - n\pi/144$ for $n = 0, \cdots, 26$, using ten significant figures. These values may be found by subtracting $\pi/144 = 0.02181661565$ (ten significant figures) successively from $\pi/2 = 1.570796326$. After 26 steps of this calculation the result is 1.003564311, whose approximate error is 8×10^{-9}. Or we may find

$$\left(\frac{\pi}{2} - 26\,\frac{\pi}{144}\right) = 46\,\frac{\pi}{144} = 46 \times 0.02181661565$$

and get 1.003564319, whose approximate error is 1×10^{-10}.

The source of this discrepancy is seen to lie in the fact that in using the value of $\pi/144$ in subtraction only nine places of decimals were carried.

(2) A step-by-step computation of sine and cosine from the equations

$$\sin k(x+2) - \sin kx = 2 \sin k \cos k(x+1),$$
$$\cos k(x+2) - \cos kx = 2 \sin k \sin k(x+1)$$

exhibited certain periodic errors. That of the sine, after increasing, returned to zero at 90° in a computation in which k was taken as $\pi/360$.

Let ξ be the disturbance in the value of $\sin kx$ and η that of $\cos kx$ if a discrepancy of ξ_1 occurs in the value of $\sin k$ as used in the computation. Then ξ and η satisfy the following linear difference equations:

$$(E^2 - 1)\xi - 2aE\eta = 2\xi_1 \cos k(x+1),$$
$$2aE\xi + (E^2 - 1)\eta = -2\xi_1 \sin k(x+1).$$

In solving these equations it is advantageous to note that

$$Z = \xi + i\eta$$

satisfies

$$(E^2 - 1)Z + 2aiEZ = 2\xi_1 e^{-k(x+1)}.$$

The starting values of these quantities ξ and η are

$$\xi_0 = \eta_0 = 0, \quad \xi_1 = \xi_1, \text{ and } \eta_1 = 0.$$

Subject to these conditions the solutions may be written

$$\xi = \left\{\frac{\xi_1}{2}\frac{\sin k}{\cos^2 k} + \frac{\eta_1}{2\cos k}\right\}[\sin (k+\pi)x + \sin kx] + \xi_1 \frac{x\cos kx}{\cos k},$$

$$\eta = \left\{\frac{\xi_1}{2}\frac{\sin k}{\cos^2 k} + \frac{\eta_1}{2\cos k}\right\}[-\cos (k+\pi)x + \cos kx] - \xi_1 \frac{x\sin kx}{\cos k}.$$

These equations show that an error in the value of sin k in our first equations will lead to the propagation of errors in the calculated sines and cosines which vanish periodically but which have, in general, an amplitude that increases with x.

(3) The errors occurring in the inversion of matrices when digital calculators are used may be inspected by inverting matrices whose inverses are known. A method of constructing a matrix whose elements are, to within an integral factor, integers, whose inverse is known and the elements of whose inverse likewise are, to within an integral factor, integers has been devised.

A few of these matrices have been inverted by the Aiken Relay Calculator. The number of significant figures carried was ten. In these calculations certain tenth-order matrices lost two significant figures, while in one sixth-order case as many as four were lost. This investigation will be continued as opportunity permits.

(4) Numerical solutions were made by the Aiken Relay Calculator of differential equations which were of the form

$$\ddot{x} = -H(y)G(v)\dot{x},$$

$$\ddot{y} = -H(y)G(v)\dot{y} - g,$$

in which $v^2 = \dot{x}^2 + \dot{y}^2$ and G and H are given as tabular functions.

The basic solution was made with initial conditions with

$$v_0 = 866.6666667.$$

Other solutions were made from the following values of v_0

$$v_0 = 866.6566667,$$

and

$$v_0 = 866.6766667.$$

At $t = 40$ the values of x and y agreed to five significant figures.

Again, trajectories were calculated with a disturbance in an initial element of 1×10^{-8} in the initial values of the velocity components $\dot{x}_0$ and $\dot{y}_0$. The resultant disturbances, which were due largely to different sequences of rounding errors, were of the same order of magnitude, with the largest disturbance noted as 2×10^{-6}.

Numerical Methods Associated with Laplace's Equation

W. E. MILNE

Institute for Numerical Analysis, UCLA, and Oregon State College

For simplicity, the following discussion is limited to Laplace's equation. Actually, the ideas and methods presented are applicable, with suitable modifications, to more general linear partial differential equations of elliptic type.

The paper is further limited to methods based on replacing the partial differential equation with an appropriate partial difference equation, and takes no account of the vast amount of research that has been devoted to obtaining approximate solutions by analytical means, such as Bergman's method of orthogonal polynomials, to mention only one.

A third limitation is the restriction of the problem to two dimensions. The theory for three dimensions is not essentially different but the numerical labor is so great that little has so far been done with three-dimensional problems.

We consider Laplace's equation in a plane. The first step is to cover the plane with a net and to set up difference equations involving the values of the unknown function $U(x, y)$ at the nodes of the net. Theoretically, the meshes of the net need not be uniform in size, but may vary continuously over the plane, as in the problem of conformal mapping. Some work has in fact been done with such nets. But this leads to variable coefficients in our difference equation and seriously complicates the programming of the problem for automatic computing machinery. Restricting ourselves to nets with uniform mesh we find three possible cases,[1] where the meshes are (i) regular hexagons, (ii) squares, (iii) equilateral triangles.

The next problem is to select the nodes for which the difference equation is to be set up. The greater the number of nodes chosen, the more closely in general will the difference equation represent the differential equation. But in order that the resultant formula be both symmetrical and at the same time applicable to points adjacent to the boundary, it is evident that we can use only a central point together with the immediately adjacent points of the net. There are, with all these limitations, only four practical possibilities:

(1) Hexagonal mesh (Fig. 1), four-point formula,[1]

$$u_0 = \tfrac{1}{3}(u_1 + u_2 + u_3) + O(h^3); \tag{1}$$

(2) Square mesh (Fig. 2), five-point formula,[1, 2, 3]

$$u_0 = \tfrac{1}{4}(u_1 + u_2 + u_3 + u_4) + O(h^4); \tag{2}$$

(3) Triangular mesh (Fig. 3), seven-point formula,[1]

$$u_0 = \tfrac{1}{6}(u_1 + u_2 + u_3 + u_4 + u_5 + u_6) + O(h^6); \tag{3}$$

(4) Square mesh (Fig. 4), nine-point formula,[2]

$$u_0 = \tfrac{1}{20}[4(u_1 + u_2 + u_3 + u_4) + u_5 + u_6 + u_7 + u_8] + O(h^8). \tag{4}$$

The term $O(h^n)$ means that $|O(h^n)| < kh^n$ as $h \to 0$, where k is some constant. Each formula is exact if u is a harmonic polynomial of degree $n - 1$. When h is made sufficiently small the formulas obviously increase in accuracy from the first to the last.

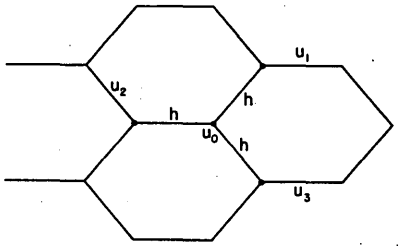


FIG. 1. Hexagonal mesh, four-point formula.

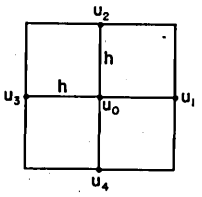


FIG. 2. Square mesh, five-point formula.

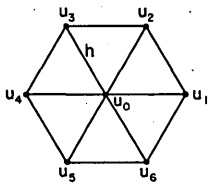


FIG. 3. Triangular mesh, seven-point formula.

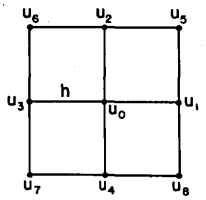


FIG. 4. Square mesh, nine-point formula.

Formula (1) is the least accurate and seems to have little to recommend it.

Formula (2) is the best known and has been much used, perhaps more than any of the others.

Formula (3) has been extensively used by Southwell and his followers. However, I gathered from conversation with Professor Southwell that he is now inclined to favor the square mesh rather than the triangular.

It is my own opinion that formula (4) is, all factors considered, the most useful finite-difference equation with which to replace Laplace's equation. In tests on a variety of harmonic functions formula (4) gave notably better accuracy than any of the other formulas, the gain in accuracy more than offsetting the slight increase in numerical computation. For the remainder of this paper we shall assume a square mesh and shall replace Laplace's differential equation by the difference equation (4) which we commonly write in the symbolic form[2]

$$\#u = \begin{array}{|c|c|c|} \hline 1 & 4 & 1 \\ \hline 4 & -20 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array} u = 0. \tag{4'}$$

This diagram is a convenient way of exhibiting the coefficients of the formula in their proper positions relative to the corresponding nodal values.

Dirichlet's problem is that of finding $U(x, y)$ harmonic in the interior of a plane region R and taking on assigned values on the boundary B of R. Let us lay a net with square mesh of side h over the region R, and for simplicity let us initially suppose that the net intersects the boundary only at nodal points. For each interior point of R we set up Eq. (4). Since the values at the interior nodes are unknown while those at the boundary nodes are known, we have a set of N linear equations in N unknowns, N being the number of interior nodes. It is readily shown that these equations always possess a unique solution. This solution then provides the nodal values of a function $W(x, y)$ which satisfies Eq. (4) at all interior nodal points and takes on the prescribed values at the boundary nodes.

The next question is, how closely does the numerical function $W(x, y)$ approximate the desired harmonic function $U(x, y)$? In spite of the importance of this question we shall not take the space here to investigate it, since it belongs more in the domain of pure analysis than in that of numerical methods. Suffice it to say that reasonably satisfactory bounds for the error can be obtained and that in most practical problems a net can be taken so that the solution of the difference equation is a satisfactory approximation to the desired harmonic function.

Attention will be centered on the problem of solving the N simultaneous linear equations. These equations are individually simple in form, but in problems of practical importance their number is so great that direct solution is at present a task too formidable to undertake. Perhaps in the future machines will be developed which can handle this large number of equations, but it seems unlikely that any of the machines contemplated at present will be able to tackle it.

We are driven, therefore, to consider methods of successive approximation. Let V be an approximation to the desired harmonic function U; let $V = U$ on the boundary B, and $V = U + E$ at interior points of R, so that E represents the error of the approximation. In order to see how the process of successive approximations works we shall consider a pair of physical problems, one in which the error E is interpreted as the displacement of a vibrating membrane, the other in which E is interpreted as temperature in a cooling conducting slab.

Consider the region R as a membrane clamped on the boundary B, and let E denote the normal displacement of the membrane at an interior point. Then E satisfies the differential system

$$\begin{aligned} &\frac{\partial^2 E}{\partial t^2} = a^2 \nabla^2 E; \\ &E = 0 \text{ on } B, \\ &E = E_0 \text{ when } t = 0, \\ &\frac{\partial E}{\partial t} = 0 \text{ when } t = 0. \end{aligned} \tag{5}$$

If λ_i are the characteristic numbers for the region R, we may assume a solution in the form

$$E = \sum_{i=1}^{\infty} \varphi_i \cos a\lambda_i t, \tag{6}$$

where the φ_i are characteristic functions satisfying the set of equations

$$\begin{aligned} \nabla^2 \varphi_i + \lambda_i^2 \varphi_i &= 0 \text{ in } R, \\ \varphi_i &= 0 \text{ on } B. \end{aligned} \tag{7}$$

Again, consider a conducting sheet shaped like the region R, with temperature E at interior points and zero on B. Then E satisfies the set of equations

$$\begin{aligned} \frac{\partial E}{\partial t} &= a^2 \nabla^2 E \text{ in } R; \\ E &= 0 \text{ on } B, \\ E &= E_0 \text{ when } t = 0. \end{aligned} \tag{8}$$

Here we may assume a solution of the form

$$E = \sum_{i=1}^{\infty} \varphi_i e^{-a^2\lambda_i^2 t}, \tag{9}$$

where the φ_i and λ_i are exactly the same as in the case of the vibrating membrane.

Assume now that the function V is given in $R + B$. Let U satisfy the equation $\nabla^2 U = 0$ in R. We have $V = U$ on B, and $E = V - U$ in R.

Evidently E so defined satisfies the system (8), and since U is independent of t and $\nabla^2 U = 0$ it follows that V satisfies the set of equations

$$\begin{aligned} \frac{\partial V}{\partial t} &= a^2 \nabla^2 V \text{ in } R; \\ V &= U \text{ on } B, \\ V &= V_0 \text{ when } t = 0. \end{aligned} \tag{10}$$

From Eq. (9) it is evident that

$$V = U + \sum_{i=1}^{\infty} \varphi_i e^{-a^2\lambda_i^2 t}. \tag{11}$$

Equation (11) shows that as t increases to infinity the terms of the series on the right approach zero and hence V approaches U, no matter what value V_0 the function V assumed initially. It is also of interest to observe that the larger the λ_i the more rapidly does the corresponding term in Eq. (11) die out. Referring to Eq. (6) we see that in the case of the vibrating membrane the terms with large λ_i correspond to terms having a high frequency of vibration, for which the corresponding characteristic functions have many nodal lines and many changes of sign.

Hence as t increases the error term E rapidly loses its highly variable components and tends toward a smooth function with few, if any, nodal lines in R.

In order to carry out numerically the limiting process indicated in the preceding section

we replace Eqs. (10) by the corresponding set of difference equations, as follows. In place of $\nabla^2 V$ we have

$$\frac{1}{6h^2}\begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & -20 & 4 \\ \hline 1 & 4 & 1 \\ \hline\end{array}\, V,$$

and in place of $\partial V/\partial t$ we set $(V_{n+1} - V_n)/\triangle t$, where V_n is the value of V at the time $t = n\triangle t$. Then

$$V_{n+1} = V_n + \frac{1}{\theta}\begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & -20 & 4 \\ \hline 1 & 4 & 1 \\ \hline\end{array}\, V_n, \tag{12}$$

or, with the notation defined in Eq. (4′),

$$V_{n+1} = V_n + \frac{1}{\theta} \# V_n,$$

in which

$$\frac{1}{\theta} = \frac{a^2 \triangle t}{6h^2}. \tag{13}$$

Equation (12) furnishes the desired recurrence relation for computing the successive approximations $V_1, V_2, \cdots, V_n$.

In order to investigate the effect of successive applications of Eq. (12), and in particular to determine whether or not the process converges, we need to know something about the characteristic numbers associated with the homogeneous partial difference equation

$$\begin{aligned} \# \varphi + \lambda^2 \varphi &= 0 \text{ in } R, \\ \varphi &= 0 \text{ on } B. \end{aligned} \tag{14}$$

Without taking time to consider details of proof and limitations on the region R we shall merely state loosely the results that we want.

1. If R contains N interior points, there are N real characteristic values of λ^2 for which Eq. (14) possesses nonzero solutions.
2. These values of λ^2 are not necessarily all distinct.
3. All characteristic values of λ^2 lie in the interval

$$0 < \lambda^2 < 32.$$

With this information we can proceed to the investigation of Eq. (12).

We assume a solution of Eq. (12) in the form

$$V_n = U + \sum_{i=1}^{N} \varphi_i \rho_i{}^n \tag{15}$$

in which the φ_i are point functions defined at each interior point of R and zero on B, while the ρ_i are constants. Substitution in Eq. (12) gives

$$\sum[\rho_i - 1]\varphi_i\rho_i^n = \sum\frac{1}{\theta}\#\varphi_i\rho_i^n,$$

and from Eq. (14) we see that this can be satisfied if

$$\theta[1 - \rho_i] = \lambda_i^2$$

Table 1. Values of λ_i^2, and corresponding values of ρ_i for various choices of θ.

λ_i^2	$\theta = 16$ ρ_i	$\theta = 20$ ρ_i	$\theta = 24$ ρ_i	$\theta = 36$ ρ_i
3.80	.76	.81	.84	.89
7.34	.54	.63	.69	.80
9.53	.40	.52	.60	.74
12.91	.19	.35	.46	.64
13.07	.18	.35	.46	.64
16.61	-.04	.17	.31	.54
17.53	-.10	.12	.27	.51
18.80	-.17	.06	.22	.48
19.09	-.19	.05	.20	.47
22.15	-.38	-.11	.08	.38
22.35	-.40	-.12	.07	.38
22.47	-.40	-.12	.06	.38
23.00	-.44	-.15	.04	.36
24.09	-.51	-.20	-.00	.33
25.53	-.60	-.28	-.06	.29
25.85	-.62	-.29	-.08	.28
26.47	-.65	-.32	-.10	.26
28.33	-.77	-.42	-.18	.21
28.85	-.80	-.44	-.20	.20
30.60	-.91	-.53	-.28	.15

where λ_i^2 is one of the characteristic numbers belonging to Eq. (14). We have then

$$\rho_i = 1 - \frac{\lambda_i^2}{\theta}.$$

In order to secure convergence[4] we obviously required that $|\rho_i| < 1$ for all i. Since λ_i^2 is greater than zero and less than 32, this requires that we take $\theta \geqq 16$. It follows that $\triangle t$ in Eq. (13) must be taken less than $6h^2/16a^2$.

For simplicity in computation it is desirable to choose θ as an integer. The effect of several choices of θ will be illustrated by a numerical example for which the characteristic numbers

can be readily computed. We choose the case where R is a rectangle with $N = 4 \times 5 = 20$ interior points. The values of the λ_i^2, and the corresponding values of ρ_i for $\theta = 16, 20, 24, 36$ are given in Table 1. The relation of ρ and λ for these different choices of θ may be readily seen from Fig. 5.

We now examine and compare the specific formulas obtained from Eq. (12) by setting $\theta = 16, 20, 24, 36$ in turn. These choices are sufficient to make clear the behavior in general, and these specific values were selected because in each case the corresponding formula has some special distinction.

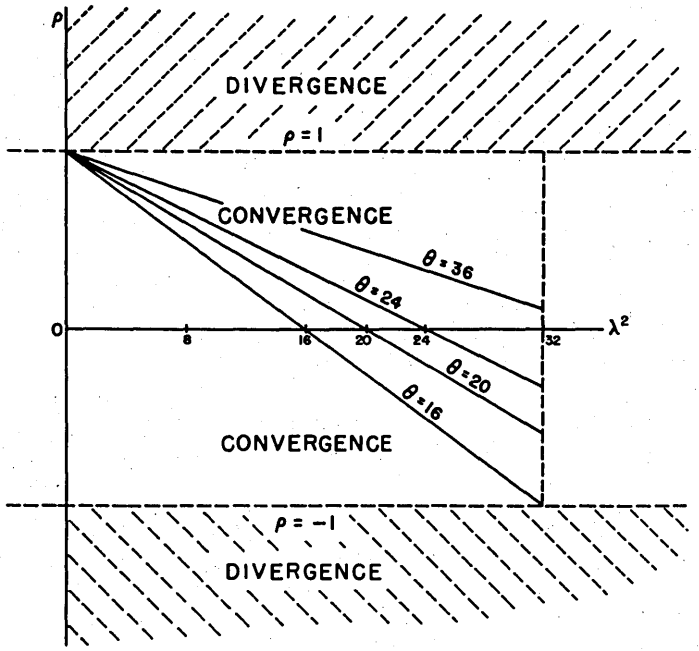


FIG. 5. Relation of ρ and λ for various choices of θ.

For $\theta = 16$, formula (12) becomes

$$\#u = \begin{array}{|c|c|c|} \hline 1 & 4 & 1 \\ \hline 4 & -20 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array} \quad u = 0 \tag{16}$$

Referring to Fig. 5, we see from Eq. (15) that the effect of repeated applications of Eq. (16) is to damp out rather slowly those components of the error term belonging to large λ_i^2 and to small λ_i^2, while those belonging to intermediate values are much more rapidly damped out. For brevity we say that Eq. (16) secures rapid liquidation of error components of intermediate frequencies, and slow liquidation of error components with low or high frequencies. This is the best formula of all for the lowest frequency.

For $\theta = 20$, formula (12) becomes

$$V_{n+1} = \tfrac{1}{20} \begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & 0 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array} V_n. \tag{17}$$

Formula (17) is less effective than (16) for low-frequency components but is much more effective for high-frequency components. Because of the zero term and the simple divisor, Eq. (17) is probably the simplest formula to compute.

For $\theta = 24$, formula (12) becomes

$$V_{n+1} = \tfrac{1}{24} \begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & 4 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array} V_n. \tag{18}$$

This is still less effective for low frequencies but is even better for high frequencies. Formula (18) is worth noting for one interesting characteristic. The largest coefficient by which any value of V_n is multiplied in the calculation of V_{n+1} is seen to be 1/6, and this is smaller than the largest coefficient in any other formula obtained from Eq. (12). This means that numerical errors, round-off errors, etc., are more rapidly damped out by Eq. (18) than by any of the other formulas for successive approximation.

For $\theta = 36$, formula (12) becomes

$$V_{n+1} = \tfrac{1}{36} \begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & 16 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array} V_n. \tag{19}$$

From the graph of ρ_i we see that for almost all frequencies this formula is not as rapidly convergent as Eq. (18). But Eq. (19) possesses the unique distinction of being factorable into the product of two operators, as may be indicated symbolically in the form

$$\tfrac{1}{36} \begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & 16 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array} = \left\{ \tfrac{1}{6} \begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline \end{array} \right\} \left\{ \tfrac{1}{6} \begin{array}{|c|}\hline 1 \\ \hline 4 \\ \hline 1 \\ \hline \end{array} \right\}$$

For programming the computation on certain types of machine this fact outweighs all other considerations. As an example, Dr. Yowell of the Institute for Numerical Analysis readily

set up a computation on the IBM 604 in which the V_n cards were run through the machine by columns and the values of

$$W_n = \tfrac{1}{6}\begin{array}{|c|}\hline 1 \\ \hline 4 \\ \hline 1 \\ \hline \end{array}\, V_n$$

were computed. Then these were run through by rows and the required values of

$$V_{n+1} = \tfrac{1}{6}\begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline \end{array}\, W_n$$

were obtained by the same program. Because of the serial nature of any machine using stacks of cards, this factoring of the operator into two operations of serial type makes possible the use of such machines in situations where they ordinarily could not be employed.

We now have a choice of formulas by which successive approximations can be carried out, together with some indication, through the size of the ρ_i, of the character of the convergence to be expected. There is also great latitude in the way in which such formulas can be used. For example, instead of calculating each value of V_{n+1} from the old values V_n one may at each step utilize such new values as are available. It has been shown[4] that this process tends to improve convergence. If this is done it may make considerable difference in what order one proceeds over the points of the region R. Conceivably it might be best to take first all points adjacent to the boundary, then all those adjacent to these, etc. So far as I am aware, this has not been investigated.

It is possible to obtain more rapid convergence by using the formula

$$V_{n+1} - V_n = \frac{1}{\theta} \# V_n + \alpha(V_n - V_{n-1})$$

with appropriate choices of the *two* parameters θ and α. Space does not permit a detailed explanation.[4]

When the number of interior points of R is large, the lowest characteristic number is small and the corresponding ρ is but slightly less than unity. For example, for a square with 101 units on a side, containing 10,000 interior points, the smallest characteristic number is about 0.01152 and for $\theta = 24$ the largest ρ is 0.99952. To reduce the error to 1 percent of its original value would take about 10,000 repetitions of the formula. For 10,000 points this means that we must apply the formula 100,000,000 times to reduce the error to 1 percent of its original value. For the computer with a desk calculator this is a dismaying prospect, and even the best electronic computers would require considerable time, since none now contemplated could store 10,000 numbers in the high-speed memory.

Hence some device is needed for chopping off errors in big chunks instead of gently polishing them down one at a time. It is here that the relaxation methods of Southwell becomes important.[1, 5] Unfortunately, from the standpoint of automatic computing machinery, applying these methods is something of an art. Considerable study will probably be required to formulate this art in such a way that it can be programmed for automatic computing machinery. I can only indicate roughly how this conceivably may be done.

Suppose that we are using formula (18). After a suitable number of repetitions the high-frequency terms will be pretty well eliminated, and the remaining errors, instead of being helter-skelter, will be collected in rather large heaps over the region R. Consider one of these heaps, the error being approximately zero on the boundary of the heap. By summing the residuals around the boundary of the heap we can calculate the average thickness of what we may call the bottom layer of the heap. Repeating the process for the next set of points inside the set adjacent to the boundary we get the thickness of the next layer, and so on. Having computed the whole pile in this fashion we remove the whole block of errors.

The process is not at all bad for a desk computer, but I do not know how it would be programmed for an automatic machine. In the examples so far tested it proved very effective. This device used two or three times and formula (18) about eight times gave better results than formula (19) applied 65 times.

This whole problem of liquidating large blocks of errors deserves additional study, especially with a view to devising processes adapted to automatic computing machines.

So far in this discussion we have assumed that the boundary B of the region R coincides with lines of the net. In most practical problems this simple situation is not realized. To meet the difficulties arising when the boundary points are not nodes of the net we may adopt any one of several expedients.

(*a*) In the vicinity of a curved boundary we may, as the computation proceeds, employ successively finer and finer nets until the actual curved boundary is closely enough represented by nearby nodal points. As far as the writer is aware, this is the method so far most commonly used. Yet it has the obvious objection of greatly increasing the number of points, often far beyond the number required to secure an adequate solution for Laplace's equation. Moreover, it is apparent that this procedure greatly complicates the programming of the problem for automatic machinery.

(*b*) Another line of attack is to devise special formulas with which to replace our standard nine-point formula for the case where the boundary cuts through the standard nine-point pattern. I have in fact derived a set of formulas for this purpose, but must admit that they also fall somewhat short of what we desire. They require a set of auxiliary interpolation tables, and also require calculation of the coördinates of each intersection of the boundary with the lines of the net.

(*c*) In certain simple cases, at least, it may be possible to handle curved boundaries by introducing curvilinear coördinates. This, however, is usually a very difficult job in itself,

and has the further objection of introducing a lot of analysis into what should be a strictly numerical procedure.

All in all, the treatment of curved boundaries stands out as one of the least satisfactory features of our whole procedure.

In the vicinity of a boundary point where the assigned boundary values suffer a discontinuity, the process of successive approximations not only converges with painful sluggishness but the limiting result is likely to be a poor approximation to the desired harmonic function. The accuracy can of course be improved by using a finer net, at the expense of additional labor.

Here, however, it appears possible to remove the difficulty before the computation is begun. Consider, for example, the case of a straight-line boundary (taken as the x-axis) with a finite jump in the boundary values at O (taken as the origin) of magnitude M, going from left to right (Fig. 6). Then the function (M/π) arctan (y/x) is harmonic in R, and has the

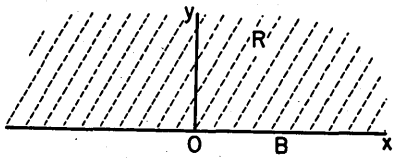


FIG. 6. Straight-line boundary with a finite jump in boundary values at O.

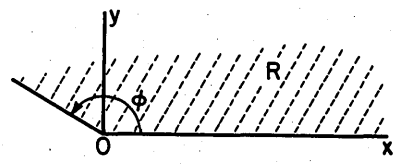


FIG. 7. Boundary with a corner at O.

value M to the left of the origin and the value zero to the right. If U is the required harmonic function, then

$$W = U + \frac{M}{\pi} \arctan \frac{y}{x}$$

is harmonic in R and has no discontinuity at O. We may therefore add the boundary values of (M/π) arctan (y/x) to the given boundary values, solve the new problem for W, and at the end obtain U from the equation

$$U = W - \frac{M}{\pi} \arctan \frac{y}{x}.$$

If the boundary has a corner at O, making an interior angle φ, the foregoing procedure is modified by using (M/φ) arctan (y/x). This device (Fig. 7) has been used with gratifying results in trial examples.

The process of successive approximations starts from some assumed value V_0. Theoretically, the process converges to the desired value V no matter what V_0 is selected, but in practice we want the original choice V_0 to be as good as possible in order to hasten the convergence. An experienced computer, familiar with the behavior of harmonic functions, can make a surprisingly good guess to start with. What we really need, however, is some definite procedure, adaptable to automatic computing machines which will provide a satisfactory initial function V_0. Some formulas have been proposed,[6] but the best of them leaves a good deal to be desired.

So far we have restricted ourselves to the simple Dirichlet problem with known values on the boundary. Actually we must also handle cases where the normal derivative is assigned on part or all of the boundary,[7] as well as cases where some combination of boundary values and normal derivatives is assigned. For straight-line boundaries coinciding with lines of the net, formulas to handle such cases are readily obtained. To judge by a few trial examples, the convergence of successive approximations will prove even slower than for the Dirichlet problem. And when curved boundaries are involved additional complexities occur. Considerably more research is needed in connection with such problems.

This incomplete and somewhat rambling paper may be brought to a close with the hope that it has served to emphasize two points:

(1) Numerical methods have been brought to the point where, with the aid of high-speed automatic computing machines, linear partial differential equations of second order can be successfully tamed and domesticated for the use of mankind; and

(2) Intensive research is still required to improve and polish the actual technique.

I am happy to acknowledge my obligation to the National Bureau of Standards and the Office of Naval Research, which have provided much of the material for my remarks and have made it possible for me to attend these meetings.

REFERENCES

1. R. V. Southwell, *Relaxation methods in theoretical physics* (Oxford: Clarendon Press, 1946).

2. W. G. Bickley, "Finite difference formulae for the square lattice," *Quart. J. Mech. Appl. Math.* **1,** 35–42 (1948).

3. G. H. Shortley, R. Weller, and B. Fried, *Numerical solution of Laplace's and Poisson's equations* (Ohio State University Studies, Engineering Experiment Station Bulletin No. 107).

4. S. Frankel, *The rate of convergence of relaxation method calculations* (mimeographed).

5. L. Fox, "Some improvements in the use of relaxation methods for the solution of ordinary and partial differential equations," *Proc. Roy. Soc. London* [A] **190,** 31–59 (1947).

6. M. M. Frocht and M. M. Leven, "A rational approach to the numerical solution of Laplace's equation," *J. Appl. Phys.* **12,** 596–604 (1941).

7. R. V. Southwell and G. Vaisey,' 'Plane potential problems involving specified normal gradients,'' *Proc. Roy. Soc. London* [A] **182,** 129–151 (1943).

An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators*

CORNELIUS LANCZOS

National Bureau of Standards, Institute for Numerical Analysis, UCLA

The eigenvalue problem of linear operators is of central importance for all vibration problems of physics and engineering. The vibrations of elastic structures, the flutter problems of aerodynamics, the stability problem of electric networks, the atomic and molecular vibrations of particle physics, are all diverse aspects of the same fundamental problem, viz. the principal-axis problem of quadratic forms.

In view of the central importance of the eigenvalue problem for so many fields of pure and applied mathematics, much thought has been devoted to the designing of efficient methods by which the eigenvalues of a given linear operator may be found. That linear operator may be of the algebraic or of the continuous type; that is, a matrix, a differential operator, or a Fredholm kernel function. Iteration methods play a prominent part in these designs, and the literature on the iteration of matrices is very extensive.[1] In the English literature of recent years the work of A. Hotelling[2] and A. C. Aitken[3] deserve attention. H. Wayland[4] surveys the field in its historical development, up to recent years. W. U. Kincaid[5] obtained additional results by improving the convergence of some of the classical procedures.

The present investigation, while starting out along classical lines, proceeds nevertheless in a different direction. The advantages of the method here developed[6] can be summarized as follows.

1. The iterations are used in the most economical fashion, obtaining an arbitrary number of eigenvalues and eigensolutions by one single set of iterations, without reducing the order of the matrix.

2. The rapid accumulation of fatal rounding errors, common to all iteration processes if applied to matrices of high dispersion (large "spread" of the eigenvalues), is effectively counteracted by the method of "minimized iterations."

3. The method is directly translatable into analytic terms, by replacing summation by integration. We then get a rapidly convergent analytic iteration process by which the eigenvalues and eigensolutions of linear differential and integral equations may be obtained.

The two classical solutions of Fredholm's problem. Since Fredholm's fundamental essay on integral equations[7] we can replace the solution of linear differential and integral equations by the solution of a set of simultaneous ordinary linear equations of infinite order. The problem

* The preparation of this paper was sponsored (in part) by the Office of Naval Research.

of Fredholm, if formulated in the language of matrices, can be stated as follows: Find a solution of the equation

$$y - \lambda A y = b, \tag{1}$$

where b is a given vector, λ a given scalar parameter, and A a given matrix (whose order eventually increases to infinity), while y is the unknown vector. This problem includes the inversion of a matrix ($\lambda = \infty$) and the problem of the characteristic solutions, also called "eigensolutions" ($b = 0$), as special cases.

Two fundamentally different classical solutions of this problem are known. The first solution is known as the "Liouville-Neumann expansion."[8] We consider A as an algebraic operator and obtain formally the infinite geometric series

$$y = \frac{1}{1 - \lambda A} b = (1 + \lambda A + \lambda^2 A^2 + \cdots) b. \tag{2}$$

This series converges for sufficiently small values of $|\lambda|$ but diverges beyond a certain $|\lambda| = |\lambda_1|$. The solution is obtained by a series of successive "iterations";[9] we construct in succession the set of vectors

$$\begin{aligned} b_0 &= b, \\ b_1 &= A b_0, \\ b_2 &= A b_1, \\ &\cdots \\ b_{n+1} &= A b_n, \end{aligned} \tag{3}$$

and then form the sum

$$y = b_0 + \lambda b_1 + \lambda^2 b_2 + \cdots \tag{4}$$

The merit of this solution is that it requires nothing but a sequence of iterations. The drawback of the solution is that its convergence is limited to sufficiently small values of λ.

The second classical solution is known as the Schmidt series.[10] We assume that the matrix A is "nondefective," i.e., that all its elementary divisors are linear. We furthermore assume that we possess all the eigenvalues[11] μ_i and eigenvectors u_i of the matrix A, defined by the equations

$$A u_i = \mu_i u_i. \quad i = 1, 2, \ldots, n \tag{5}$$

If A is nonsymmetric, we need also the "adjoint" eigenvectors u_i^*, defined with the help of the transposed matrix A^* by the equations

$$A^* u_i^* = \mu_i u_i^*. \quad i = 1, 2, \ldots, n \tag{6}$$

We now form the scalars

$$\gamma_i = \frac{b \cdot u_i^*}{u_i \cdot u_i^*} \tag{7}$$

and obtain y in the form of the expansion

$$y = \frac{\gamma_1 u_1}{1 - \lambda \mu_1} + \frac{\gamma_2 u_2}{1 - \lambda \mu_2} + \cdots + \frac{\gamma_n u_n}{1 - \lambda \mu_n}. \tag{8}$$

This series offers no convergence difficulties since it is a finite expansion in the case of matrices of finite order and yields a convergent expansion in the case of the infinite matrices associated with the kernels of linear differential and integral operators.

The drawback of this solution is—apart from the exclusion of defective matrices[12]—that it presupposes the complete solution of the eigenvalue problem associated with the matrix A.

Solution of the Fredholm problem by the S-expansion. We now develop a new expansion which solves the Fredholm problem much as the Liouville-Neumann series does but avoids the convergence difficulty of that solution.

We first notice that the iterated vectors $b_0, b_1, b_2, \ldots$ cannot be linearly independent of each other beyond a certain definite b_k. All these vectors find their place within the n-dimensional space of the matrix A; hence not more than n of them can be linearly independent. We thus know in advance that there must exist between the successive iterations a linear identity of the form

$$b_m + g_1 b_{m-1} + g_2 b_{m-2} + \cdots + g_m b_0 = 0. \tag{9}$$

We cannot tell in advance what m will be, except for the lower and upper bounds:

$$1 \leqq m \leqq n. \tag{10}$$

How to establish the relation (10) by a systematic algorithm will be shown presently; for the time being we assume that the relation is already established. We now define the polynomial

$$G(x) = x^m + g_1 x^{m-1} + \cdots + g_m \tag{11}$$

together with the "inverted polynomial" (the coefficients of which follow the opposite sequence)

$$S_m(\lambda) = 1 + g_1\lambda + g_2\lambda^2 + \cdots + g_m\lambda^m. \tag{12}$$

Furthermore, we introduce the partial sums of the latter polynomial:

$$\begin{aligned} S_0 &= 1, \\ S_1(\lambda) &= 1 + g_1\lambda, \\ S_2(\lambda) &= 1 + g_1\lambda + g_2\lambda^2, \\ &\cdots \\ S_{m-1}(\lambda) &= 1 + g_1\lambda + \cdots + g_{m-1}\lambda^{m-1}. \end{aligned} \tag{13}$$

We now refer to a formula which can be proved by straightforward algebra:[13]

$$\frac{S_m(\lambda) - \lambda^m G(x)}{1 - \lambda x} = S_{m-1}(\lambda) + S_{m-2}(\lambda)\cdot\lambda x + S_{m-3}(\lambda)\cdot\lambda^2 x^2 + \cdots + S_0\cdot\lambda^{m-1}x^{m-1}. \tag{14}$$

Let us apply this formula operationally, replacing x by the matrix A, and operating on the vector b_0. In view of the definition of the vectors b_i, the relation (10) gives

$$G(A)\cdot b_0 = 0, \tag{15}$$

and thus we obtain

$$\frac{S_m(\lambda)}{1-\lambda A}\, b_0 = S_{m-1}(\lambda)b_0 + S_{m-2}(\lambda)\lambda b_1 + \cdots + S_0\lambda^{m-1}b_{m-1} \tag{16}$$

and hence

$$y = \frac{1}{1-\lambda A}\, b_0 = \frac{S_{m-1}(\lambda)b_0 + S_{m-2}(\lambda)\lambda b_1 + \cdots S_0\lambda^{m-1}b_{m-1}}{S_m(\lambda)}. \tag{17}$$

If we compare this solution with the earlier solution (4) we notice that the expansion (17) may be conceived as a modified form of the Liouville-Neumann series because it is composed of the same kind of terms, the difference being only that we *weight* the terms $\lambda^k b_k$ by the weight factors

$$w_k = \frac{S_{m-k-1}(\lambda)}{S_m(\lambda)} \tag{18}$$

instead of taking them all with the uniform weight factor 1. This weighting has the beneficial effect that the series *terminates* after m terms, instead of going on endlessly. The weight factors w_i are very near to 1 for small λ but become more and more important as λ increases. The weighting makes the series convergent for *all* values of λ.

The remarkable feature of the expansion (17) is its *complete generality*. No matter how defective the matrix A may be, and no matter how the vector b_0 was chosen, the expansion (17) is *always* valid, provided only that we interpret it properly. In particular we have to bear in mind that there will always be m polynomials $S_k(\lambda)$, even though every $S_k(\lambda)$ may not be of degree k, owing to the vanishing of the higher coefficients. For example, it could happen that

$$G(x) = x^m, \tag{19}$$

so that

$$S_m(\lambda) = 1 + 0\lambda + 0\lambda^2 + \cdots + 0\lambda^m, \tag{20}$$

$$S_k(\lambda) = 1 + 0\lambda + 0\lambda^2 + \cdots + 0\lambda^k, \tag{21}$$

and formula (17) gives:

$$y = b_0 + \lambda b_1 + \cdots + \lambda^{m-1}b_{m-1}. \tag{22}$$

Solution of the eigenvalue problem. The Liouville-Neumann series cannot give the solution of the eigenvalue problem since the expansion becomes divergent as soon as the parameter λ reaches the lowest characteristic number λ_1. The Schmidt series cannot give the solution of the eigenvalue problem since it presupposes the knowledge of all the eigenvalues and eigenvectors of the matrix A. On the other hand, the expansion (17), which is based purely on iterations and yet remains valid for all λ, must contain implicitly the solution of the principal-axis problem. Indeed, let us write the right-hand member of Eq. (1) in the form

$$b = S_m(\lambda)\bar{b}. \tag{23}$$

Then the expansion (17) loses its denominator and becomes

$$y = S_{m-1}(\lambda)\bar{b}_0 + S_{m-2}(\lambda)\lambda\bar{b}_1 + \cdots S_0\lambda^{m-1}\bar{b}_{m-1}. \tag{24}$$

We can now answer the question whether a solution of the homogeneous equation

$$y - \lambda A y = 0 \tag{25}$$

is possible without the identical vanishing of y. The expression (23) shows that b can vanish only under two circumstances; either the vector $\bar{b}$ or the scalar $S_m(\lambda)$ must vanish. Since the former possibility leads to an identically vanishing y, only the latter possibility is of interest. This gives for the parameter λ the condition

$$S_m(\lambda) = 0. \tag{26}$$

The roots of this equation give us the characteristic values $\lambda = \lambda_i$, while the solution (24) yields the characteristic solutions, or eigenvalues, or principal axes of the matrix A:

$$u_i = S_{m-1}(\lambda_i)b_0 + S_{m-2}(\lambda_i)\lambda_i b_1 + \cdots + S_0 \lambda_i^{m-1} b_{m-1}. \tag{27}$$

It is a remarkable fact that although the vector b_0 was chosen entirely freely, the particular linear combination (27) of the iterated vectors has *invariant significance*, except for an undetermined factor of proportionality which remains free, in view of the linearity of the defining equation (25). That undetermined factor may even come out to be zero, i.e., a certain axis may not be represented in the trial vector b_0 at all. This explains why the order of the polynomial $S_m(\lambda)$ need not be necessarily equal to n. The trial vector b_0 may not give us *all* the principal axes of A. What we can say with assurance, however, is that all the roots of $S_m(\lambda)$ are true characteristic values of A, and all the u_i obtained by the formula (24) are true characteristic vectors, even if we did not obtain the *complete* solution of the eigenvalue problem. The discrepancy between the order m of the polynomial $G(\mu)$ and the order n of the characteristic equation

$$F(\mu) \begin{vmatrix} a_{11} - \mu & \cdots & a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & \cdots & a_{nn} - \mu \end{vmatrix} = 0 \tag{28}$$

will be the subject of the discussions of the next section.

Instead of substituting in formula (27) we can also obtain the principal axes u_i by a numerically simpler process, applying synthetic division. By synthetic division we generate the polynomials

$$\frac{G(x)}{x - \mu_i} = x^{m-1} + g_1{}^i x^{m-2} + \cdots + g_{m-1}{}^i. \tag{29}$$

We then replace x^j by b_j and obtain

$$u_i = b_{m-1} + g_1{}^i b_{m-2} + \cdots + g_{m-1}{}^i b_0. \tag{30}$$

The proof follows immediately from the equation

$$(A - \mu_i)u_i = G(A) \cdot b_0 = 0. \tag{31}$$

The problem of missing axes. Let us assume that we start with an arbitrary "trial vector" b_0 and obtain by successive iterations the sequence

$$b_0, b_1, b_2, \ldots, b_n. \tag{32}$$

Similarly we start with the trial vector b_0^* and obtain by iterating with the transposed matrix A^* the adjoint sequence

$$b_0^*, b_1^*, b_2^*, \ldots, b_n^*. \tag{33}$$

Let us now form the following set of "basic scalars"

$$c_{i+k} = b_i \cdot b_k^* = b_k \cdot b_i^*. \tag{34}$$

It is a remarkable fact that these scalars depend only on the *sum* of the two subscripts i and k; for example,

$$b_{k-1}b_{k+1}^* = b_{k-1}^* b_{k+1} = b_k b_k^*. \tag{35}$$

This gives a powerful numerical check of the iteration scheme, since a discrepancy between the two members of Eq. (35) (beyond the limits of the rounding errors) would indicate an error in the calculation of b_{k+1} or b_{k+1}^*, if the sequence up to b_k and b_k^* had been checked before.

Let us assume for the sake of the present argument that A is a nondefective matrix, and let us analyze the vector b_0 in terms of the eigenvectors u_i, while b_0^* will be analyzed in terms of the adjoint vectors u_i^*; thus,

$$b_0 = \bar{\beta}_1 u_1 + \bar{\beta}_2 u_2 + \cdots + \bar{\beta}_n u_n, \tag{36}$$

$$b_0^* = \bar{\beta}_1^* u_1^* + \bar{\beta}_2^* u_2^* + \cdots + \bar{\beta}_n^* u_n^*. \tag{37}$$

Then the scalars c_i become

$$c_i = \rho_1 \mu_1^i + \rho_2 \mu_2^i + \cdots + \rho_n \mu_n^i, \tag{38}$$

with

$$\rho_k = \bar{\beta}_k \bar{\beta}_k^*. \tag{39}$$

The problem of obtaining the μ_i from the c_i is the problem of "weighted moments," which can be solved as follows. Assuming that none of the ρ_k vanish and that all the λ_i are distinct, we establish a linear relation between $n+1$ consecutive c_i of the following form:

$$\begin{aligned}
c_0\eta_0 + c_1\eta_1 \quad &+ \cdots + c_{n-1}\eta_{n-1} + c_n = 0,\\
c_1\eta_0 + c_2\eta_1 \quad &+ \cdots + c_n\eta_{n-1} + c_{n+1} = 0,\\
\cdot \quad \cdot \quad & \quad \cdot \quad \cdot \quad \cdot\\
\cdot \quad \cdot \quad & \quad \cdot \quad \cdot \quad \cdot\\
\cdot \quad \cdot \quad & \quad \cdot \quad \cdot \quad \cdot\\
c_n\eta_0 + c_{n+1}\eta_1 &+ \cdots + c_{2n-1}\eta_{n-1} + c_{2n} = 0.
\end{aligned} \tag{40}$$

Then the definition of the c_i shows directly that the set (40) demands that

$$F(\mu_i) = 0, \tag{41}$$

where

$$F(x) = \eta_0 + \eta_1 x + \cdots + \eta_{n-1}x^{n-1} + x^n. \tag{42}$$

Hence, by solving the recurrent set (40) with the help of a "progressive algorithm," displayed in the next section, we can obtain the coefficients of the characteristic polynomial (42), whose roots give the eigenvalues μ_1.

Under the given restricting conditions none of the μ_i roots have been lost and we could actually establish the full characteristic equation (41). It can happen, however, that b_0 is orthogonal to some axis u_i^* and it is equally possible that b_0^* is orthogonal to some axis u_k. In that case $\overline{\beta}_i$ and $\overline{\beta}_k^*$ drop out of the expansions (36) and (37) and consequently the expansion (38) lacks both ρ_i and ρ_k. This means that the scalars c_j are unable to provide all the μ_j, since μ_i and μ_k are missing. The characteristic equation (41) cannot be fully established under these circumstances.

The deficiency was here caused by an unsuitable choice of the vectors b_0 and b_0^*; it is removable by a better choice of the trial vectors. However, we can have another situation where the deficiency goes deeper and is not removable by *any* choice of the trial vectors. This happens if the μ_i roots of the characteristic equation are not all distinct. The expansion (38) shows that two equal roots λ_i and λ_k cannot be separated since they behave exactly like one single root with a double amplitude. Generally, the weighted moments c_i can never show whether or not there are multiple roots, because the multiple roots behave like single roots. Consequently, in the case of multiple eigenvalues the linear relation between the c_i will be not of the nth but of a lower order. If the number of distinct roots is m, then the relations (40) will appear in the form

$$\begin{aligned}
c_0\eta_0 + c_1\eta_1 &+ \cdots + c_{m-1}\eta_{m-1} + c_m &&= 0,\\
c_1\eta_0 + c_2\eta_1 &+ \cdots + c_m\eta_{m-1} + c_{m+1} &&= 0,\\
\cdot \quad \cdot &\qquad \cdot \quad \cdot \quad \cdot &&\\
\cdot \quad \cdot &\qquad \cdot \quad \cdot \quad \cdot &&\\
\cdot \quad \cdot &\qquad \cdot \quad \cdot \quad \cdot &&\\
c_m\eta_0 + c_{m+1}\eta_1 &+ \cdots + c_{2m-1}\eta_{m-1} + c_{2m} &&= 0.
\end{aligned} \tag{43}$$

Once more we can establish the polynomial

$$G(x) = \eta_0 + \eta_1 x + \cdots + \eta_{m-1}x^{m-1} + x^m, \tag{44}$$

but this polynomial is now of only mth order and factors into the m root factors

$$(x - \mu_1)(x - \mu_2) \cdots (x - \mu_m), \tag{45}$$

where all the μ_i are distinct. After obtaining all the roots of the polynomial (44) we can now construct by synthetic division the polynomials

$$\frac{G(x)}{x - \mu_k} = x^{m-1} + g_1{}^k x^{m-2} + \cdots + g_{m-1}{}^k, \tag{46}$$

and replacing x^j by b_j we obtain the principal axes of both A and A^*:

$$\begin{aligned}
u_k &= b_{m-1} + g_1{}^k b_{m-2} + \cdots + g_{m-1}{}^k b_0,\\
u_k^* &= b_{m-1}^* + g_1{}^k b_{m-2}^* + \cdots + g_{m-1}{}^k b_0^*.
\end{aligned} \tag{47}$$

This gives a partial solution of the principal-axis problem, inasmuch as each multiple root contributed only one axis. Moreover, we cannot tell from our solution which one of the roots is single and which one multiple, nor can the degree of multiplicity be established. In order

to get further information, we have to change our trial vectors and go through the iteration scheme once more. We now substitute in the formulas (47) again and can immediately localize all the single roots by the fact that the vectors u_j associated with these roots do not change (apart from a proportionality factor), while the u_k belonging to double roots will generally change their direction. A proper linear combination of the new u_k' and the previous u_k establishes the second axis associated with the double eigenvalue μ_k; we put

$$u_k^1 = u_k, \qquad u_k^2 = u_k' + \gamma u_k,$$
$$u_k^{1*} = u_k^*, \qquad u_k^{2*} = u_k'^* + \gamma^* u_k^*.$$

The factors γ and γ^* are determined by the condition that the vectors u_k^1 and u_k^2 have to be biorthogonal to the vectors u_k^{1*} and u_k^{2*}.

In the case of triple roots a third trial is demanded, and so on.

An interesting contrast to this behavior of multiple roots associated with nondefective matrices is provided by the behavior of multiple roots associated with defective matrices. A defective eigenvalue is always a multiple eigenvalue, but here the multiplicity is caused not by the collapse of two very near eigenvalues, but by the multiplicity of the elementary divisor. This comes into evidence in the polynomial $G(x)$ by giving a root factor of order higher than the first. Whenever the polynomial $G(x)$ reveals a multiple root, we can tell in advance that the matrix A is defective in these roots, and the multiplicity of the root establishes the degree of deficiency.

It will be revealing to demonstrate these conditions with the help of a matrix which combines all the different types of irregularities that may be encountered in working with arbitrary matrices. Let us analyze the following matrix of sixth order:

$$\begin{pmatrix} 1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The eigenvalue 2 is the only regular eigenvalue of this matrix. The matrix is "singular" because the determinant of the coefficients is zero. This, however, is irrelevant from the viewpoint of the eigenvalue problem since the eigenvalue "zero" is just as good as any other eigenvalue. More important is the fact that the eigenvalue zero is a *double root* of the characteristic equation. The remaining three roots of the characteristic equation are all 1. This 1 is thus a triple root of the characteristic equation; at the same time the matrix has a double deficiency in this root because the elementary divisor associated with this root is cubic. The matrix possesses only 4 independent principal axes.

What will the polynomial $G(x)$ become in the case of this matrix? The regular eigenvalue 2 must give the root factor $x - 2$. The regular eigenvalue 0 has the multiplicity 2 but is reduced to the single eigenvalue 0 and thus contributes the factor x. The deficient eigenvalue

1 has the multiplicity 3 but also double defectiveness. Hence, it must contribute the root factor $(x-1)^3$. We can thus predict that the polynomial $G(x)$ will come out as follows:

$$G(x) = x(x-2)(x-1)^3 = x^5 - 5x^4 + 9x^3 - 7x^2 + 2x.$$

Let us verify this numerically. As a trial vector we choose

$$b_0 = b_0^* = 1, 1, 1, 1, 1, 1.$$

The successive iterations yield the following table:

$$\begin{aligned}
b_0 &= 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1, \\
b_1 &= 6 \quad 5 \quad 1 \quad 2 \quad 0 \quad 0, \\
b_2 &= 19 \quad 9 \quad 1 \quad 4 \quad 0 \quad 0, \\
b_3 &= 40 \quad 13 \quad 1 \quad 8 \quad 0 \quad 0, \\
b_4 &= 69 \quad 17 \quad 1 \quad 16 \quad 0 \quad 0, \\
b_5 &= 106 \quad 21 \quad 1 \quad 32 \quad 0 \quad 0, \\
b_6 &= 151 \quad 25 \quad 1 \quad 64 \quad 0 \quad 0;
\end{aligned}$$

$$\begin{aligned}
b_0^* &= 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1, \\
b_1^* &= 1 \quad 3 \quad 8 \quad 2 \quad 0 \quad 0, \\
b_2^* &= 1 \quad 5 \quad 23 \quad 4 \quad 0 \quad 0, \\
b_3^* &= 1 \quad 7 \quad 46 \quad 8 \quad 0 \quad 0, \\
b_4^* &= 1 \quad 9 \quad 77 \quad 16 \quad 0 \quad 0, \\
b_5^* &= 1 \quad 11 \quad 116 \quad 32 \quad 0 \quad 0, \\
b_6^* &= 1 \quad 13 \quad 163 \quad 64 \quad 0 \quad 0.
\end{aligned}$$

We now construct the c_i by dotting b_0 with the b_i^* (or b_0^* with the b_i); we continue by dotting b_6 with $b_1^*, \ldots, b_6^*$ (or b_6^* with $b_1, \ldots, b_6$). This gives the following string of $2n+1 = 13$ basic scalars:

$$c_i = 6, 14, 33, 62, 103, 160, 241, 362, 555, 884, 1477, 2590, 4735.$$

The application of the progressive algorithm of the next section to these b_i yields $G(x)$ in the predicted form. We now obtain by synthetic divisions

$$\frac{G(x)}{x-1} = x^4 - 4x^3 + 5x^2 - 2x,$$

$$\frac{G(x)}{x-2} = x^4 - 3x^3 + 3x^2 - x,$$

$$\frac{G(x)}{x} = x^4 - 5x^4 + 9x^2 - 7x + 2.$$

Inverting these polynomials we obtain the matrix

$$\begin{pmatrix} 0 & -2 & 5 & -4 & 1 & 0 \\ 0 & -1 & 3 & -3 & 1 & 0 \\ 2 & -7 & 9 & -5 & 1 & 0 \end{pmatrix}.$$

The product of this matrix with the iteration matrix B (omitting the last row b_6) yields three principal axes u_i; similarly, the product of the same matrix with the iteration matrix B^* yields the three adjoint axes u_i^*:

$$\begin{aligned} u(1) &= -8 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0, \\ u(2) &= \quad 0 \quad 0 \quad 0 \quad 2 \quad 0 \quad 0, \\ u(0) &= \quad 0 \quad 0 \quad 0 \quad 0 \quad 2 \quad 2; \end{aligned}$$

$$\begin{aligned} u^*(1) &= \quad 0 \quad 0 \quad -8 \quad 0 \quad 0 \quad 0, \\ u^*(2) &= \quad 0 \quad 0 \quad 0 \quad 2 \quad 0 \quad 0, \\ u^*(0) &= \quad 0 \quad 0 \quad 0 \quad 0 \quad 2 \quad 2. \end{aligned}$$

Since $G(x)$ is of only fifth order, while the order of the characteristic equation is 6, we know that one of the axes is still missing. We cannot decide a priori whether the missing axis is caused by the duplicity of the eigenvalue 0, 1, or 2.[14] However, a repetition of the iteration with the trial vectors

$$b_0 = b_0^* = 1, 1, 1, 1, 1, 0$$

causes a change in the rows $u(0)$ and $u^*(0)$ only. This designates the eigenvalue $u = 0$ as the double root. The process of biorthogonalization finally yields[15]

$$\begin{aligned} u_1(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1, \\ u_2(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad -1; \end{aligned}$$

$$\begin{aligned} u_1^*(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1, \\ u_2^*(0) &= 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad -1. \end{aligned}$$

The progressive algorithm for the construction of the characteristic polynomial $G(x)$. The crucial point in our discussions was the establishment of a linear relation between a certain b_m and the previous iterated vectors. This relation leads to the characteristic polynomial $G(x)$, whose roots $G(\mu_i) = 0$ yield the eigenvalues μ_1. Then by synthetic division we can immediately obtain that particular linear combination of the iterated vectors b_i which give us the eigenvectors (principal axes) of the matrix A.

We do not know in advance in what relation the order m of the polynomial $G(x)$ will be to the order n of the matrix A. Accidental deficiencies of the trial vectors b_0, b_0^*, and the presence of multiple eigenvalues in A can diminish m to any value between 1 and n. For this reason we will follow a systematic procedure that generates $G(x)$ gradually, going through all degrees from 1 to m. The procedure comes automatically to a halt when the proper m has been reached.

Our final goal is to solve the recurrent set of equations

$$\begin{aligned} c_0\eta_0 + c_1\eta_1 \quad &+ \cdots + c_m = 0, \\ \cdot \quad \cdot \quad &\cdot \quad \cdot \\ \cdot \quad \cdot \quad &\cdot \quad \cdot \\ \cdot \quad \cdot \quad &\cdot \quad \cdot \\ c_m\eta_0 + c_{m+1}\eta_1 &+ \cdots + c_{2m} = 0. \end{aligned} \tag{48}$$

This is possible only if the determinant of this homogeneous set vanishes:

$$\begin{vmatrix} c_0 c_1 & \cdots & c_m \\ c_1 c_2 & \cdots & c_{m+1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ c_m c_{m+1} & \cdots & c_{2m} \end{vmatrix} = 0. \tag{49}$$

Before reaching this goal, however, we can certainly solve for any $k < m$ the following inhomogeneous set (the upper k is meant as a superscript):

$$\begin{aligned} c_0\eta_0^k + c_1\eta_1^k \quad &+ \cdots + c_k \quad = 0, \\ c_1\eta_0^k + c_2\eta_1^k \quad &+ \cdots + c_{k+1} = 0, \\ \cdot \qquad \cdot \qquad & \qquad \cdot \quad \cdot \\ \cdot \qquad \cdot \qquad & \qquad \cdot \quad \cdot \\ \cdot \qquad \cdot \qquad & \qquad \cdot \quad \cdot \\ c_k\eta_0^k + c_{k+1}\eta_1^k &+ \cdots + c_{2k} \quad = h_k. \end{aligned} \tag{50}$$

The freedom of h_k in the last equation removes the overdetermination of the set (48). The proper m will be reached as soon as h_m turns out to be zero.

Now a recurrent set of equations has certain algebraic properties that are not shared by other linear systems. In particular, there exists a *recursion relation* between the solutions of three consecutive sets of the type (50). This greatly facilitates the method of solution, through the application of a systematic recursion scheme that will now be developed.

We consider the system (50) and assume that we possess the solution up to a definite k. Then we will show how this solution may be utilized for the construction of the next solution which belongs to the order $k + 1$.

Our scheme becomes greatly simplified if we pay attention to an additional set of equations which omits the first of Eqs. (50) but adds one more equation at the end:

$$\begin{aligned} c_1\bar{\eta}_1^k \quad + c_2\bar{\eta}_2^k \quad &+ \cdots + c_{k+1} \quad = 0, \\ \cdot \qquad \cdot \qquad & \qquad \cdot \quad \cdot \\ \cdot \qquad \cdot \qquad & \qquad \cdot \quad \cdot \\ \cdot \qquad \cdot \qquad & \qquad \cdot \quad \cdot \\ c_{k+1}\bar{\eta}_1^k + c_{k+2}\bar{\eta}_2^k &+ \cdots + c_{2k+1} = \bar{h}_{k+1}. \end{aligned} \tag{51}$$

Let us now multiply the set (50) by the factor

$$q_k = -\frac{\bar{h}_{k+1}}{h_k} \tag{52}$$

and add the set (51). We get a new set of equations which can be written down in the form:

$$\begin{aligned} c_0\eta_0^{k+1} + c_1\eta_1^{k+1} \quad &+ \cdots + c_{k+1} = 0, \\ \cdot \quad &\cdot \quad \cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \quad \cdot \\ c_k\eta_0^{k+1} + c_{k+1}\eta_1^{k+1} &+ \cdots + c_{2k+1} = 0, \end{aligned} \tag{53}$$

provided we put

$$\begin{aligned} \eta_0^{k+1} &= q_k\eta_0^k, \\ \eta_1^{k+1} &= q_k\eta_1^k + \bar{\eta}_1^k, \\ \cdot \quad &\cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \\ \eta_k^{k+1} &= q_k \cdot 1 + \bar{\eta}_k^k. \end{aligned} \tag{54}$$

We now evaluate the scalar

$$c_{k+1}\eta_0^{k+1} + \cdots + c_{2k+1}\eta_k^{k+1} + c_{2k+2} = h_{k+1}, \tag{55}$$

which is added to the set (53) as the last equation.

What we have accomplished is that a proper linear combination of the solutions η_i^k and $\bar{\eta}_i^k$ provided us with the next solution η_i^{k+1}. But now exactly the same procedure can be utilized to obtain $\bar{\eta}_i^{k+1}$ on the basis of $\bar{\eta}_i^k$ and η_i^{k+1}.

For this purpose we multiply the set (51) by

$$\bar{q}_k = -\frac{h_{k+1}}{\bar{h}_{k+1}} \tag{56}$$

and add the set (53), completed by (55) but omitting the first equation. This gives

$$\begin{aligned} c_1\bar{\eta}_1^{k+1} \quad &+ c_2\bar{\eta}_2^{k+1} \quad + \cdots + c_{k+2} = 0, \\ \cdot \quad &\cdot \quad \cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \quad \cdot \\ c_{k+1}\bar{\eta}_1^{k+1} &+ c_{k+2}\bar{\eta}_2^{k+1} + \cdots + c_{2k+2} = 0, \end{aligned} \tag{57}$$

provided we put

$$\begin{aligned} \bar{\eta}_1^{k+1} &= \bar{q}_k\bar{\eta}_1^k + \eta_0^{k+1}, \\ \bar{\eta}_2^{k+1} &= \bar{q}_k\bar{\eta}_2^k + \eta_1^{k+1}, \\ \cdot \quad &\cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \\ \cdot \quad &\cdot \quad \cdot \\ \bar{\eta}_{k+1}^{k+1} &= \bar{q}_k \cdot 1 + \eta_k^{k+1}. \end{aligned} \tag{58}$$

Once more we evaluate the scalar

$$c_{k+2}\bar{\eta}_1^{k+1} + \cdots + c_{2k+3} = \bar{h}_{k+2}, \tag{59}$$

which is added to (57) as the last equation.

This analytic procedure can be translated into an elegant geometric arrangement which generates the successive solutions η_i^k and $\bar{\eta}_i^k$ in successive columns. The resulting algorithm is best explained with the help of a numerical example.

For this purpose we choose the eigenvalue problem of an intentionally oversimplified matrix since our aim is to show not the power of the method but the nature of the algorithm that leads to the establishment of the characteristic equation. The limitation of the method due to the accumulation of rounding errors will be discussed in the next section.

Let the given matrix be

$$\begin{pmatrix} 13 & 5 & -23 \\ 4 & 0 & -4 \\ 7 & 3 & -13 \end{pmatrix}$$

We iterate with the trial vector $b_0 = 1, 0, 0$, and obtain

$$\begin{matrix} 1 & 0 & 0 \\ 13 & 4 & 7 \\ 28 & 24 & 12 \\ 208 & 64 & 112 \end{matrix}$$

We transpose the matrix and iterate with the trial vector $b_0^* = 1, 0, 0$, obtaining

$$\begin{matrix} 1 & 0 & 0 \\ 13 & 5 & -23 \\ 28 & -4 & -20 \\ 208 & 80 & -368 \end{matrix}$$

We dot the first row and the last row with the opposing matrix and obtain the basic scalars c_i as follows:[16]

1, 13, 28, 208, 448, 3328, 7168.

Table 1.

	0	0.5	1	1.5	2	2.5	3
h_1 =	1	13	-141	147.6923075	-163.4042569	0	
q_1 =	-13	10.84615384	1.047463173	1.106382990	0		0
1	1						
13		1	-13				
28			1	-2.15384616	-13.617021249		
208				1	-1.106382987	-16.000000004	0
448					1	0.000000003	-16
3328						1	0
7168							1

These numbers are written down in a column and the scheme shown in Table 1 comes into operation.

Instead of distinguishing between the η_i and $\bar{\eta}_i$ solutions we use a uniform procedure but mark the successive columns alternately as "full" and "half columns"; thus we number the successive columns as zero, one-half, one, ... The scheme has to end at a *full* column, and the end is marked by the vanishing of the corresponding "head number" h_i. In our scheme the head number is zero already at the half-column 2.5, but here the scheme cannot end, and thus we continue to the column 3, whose head number becomes once more 0, and then the scheme is finished. The last column gives the polynomial $G(x)$, starting with the diagonal term and proceeding upward:

$$G(x) = 1 \cdot x^3 + 0 \cdot x^2 - 16x + 0$$
$$= x^3 - 16x.$$

The head numbers h_i are always obtained by dotting the column below with the basic column c_i; for example, at the head of column 2 we find the number — 163.4042569. This number was obtained by the following cumulative multiplication:

$$448 \cdot 1 + 208 \cdot (- 1.106382987) + 28 \cdot (- 13.617021249).$$

The numbers q_i represent the negative ratio of two consecutive h_i numbers:

$$q_i = - \frac{h_{i+\frac{1}{2}}}{h_i};$$

for example, $q_{1.5} = 1.106382990$ was obtained by the division

$$\frac{- (- 163.4042569)}{147.6923075}.$$

The scheme grows as follows. As soon as a certain column C_i is completed, we evaluate the associated head number h_i; this provides us with the previous q-number $q_{i-\frac{1}{2}}$. To construct the next column $C_{i+\frac{1}{2}}$, we multiply the column $C_{i-\frac{1}{2}}$ by the constant $q_{i-\frac{1}{2}}$ and add the column C_i; thus

$$C_{i+\frac{1}{2}} = q_{i-\frac{1}{2}} \cdot C_{i-\frac{1}{2}} + C_i.$$

However, the result of this operation is shifted down by one element; for example, in constructing column 2.5 the result of the operation

$$1.10638299 \cdot (- 2.15384616) + (- 13.617021249) = - 16$$

is not put in the row where the operation occurred, but shifted down to the next row below.

The unfilled spaces of the scheme mean automatically "zero."

The outstanding feature of this algorithm is that *it can never come to premature grief*, provided only that the first two c-numbers c_0 and c_1 are different from zero. Division by zero cannot occur since the scheme comes to an end anyway as soon as the head number zero appears in one of the full columns.

Of interest is also the fact that the products of the head-numbers associated with the

full columns give us the successive recurrent determinants of the c_i; for example, the determinants

$$1, \quad \begin{vmatrix} 1 & 13 \\ 13 & 28 \end{vmatrix}, \quad \begin{vmatrix} 1 & 13 & 28 \\ 13 & 28 & 208 \\ 28 & 208 & 448 \end{vmatrix},$$

and

$$\begin{vmatrix} 1 & 13 & 28 & 208 \\ 13 & 28 & 208 & 448 \\ 28 & 208 & 448 & 3328 \\ 208 & 448 & 3328 & 7168 \end{vmatrix}$$

are given by the successive products

$$1,\ 1 \cdot (-141) = -141,\ (-141) \cdot (-163.4042569) = 23040, \text{ and } 23040 \cdot 0 = 0.$$

Similarly, the products of the head numbers of the half-columns give us similar determinants, but omitting c_0 from the sequence of c-numbers. In the example above the determinants

$$13, \quad \begin{vmatrix} 13 & 28 \\ 28 & 208 \end{vmatrix}, \quad \begin{vmatrix} 13 & 28 & 208 \\ 28 & 208 & 448 \\ 208 & 448 & 3328 \end{vmatrix}$$

are given by the products

$$13, \qquad 13 \cdot 147.6923075 = 1920, \qquad 1920 \cdot 0 = 0.$$

The purpose of the algorithm of Table 1 was to generate the coefficients of the basic identity that exists between the iterated vectors b_i. This identity finds expression in the vanishing of the polynomial $G(x)$:

$$G(x) = 0. \tag{60}$$

The roots of this algebraic equation give us the eigenvalues of the matrix A. In our example we get the cubic equation

$$x^3 - 16x = 0,$$

which has three roots

$$\mu_1 = 0, \qquad \mu_2 = 4, \qquad \mu_3 = -4. \tag{61}$$

These are the eigenvalues of our matrix. In order to obtain the associated eigenvectors, we divide $G(x)$ by the root factors

$$\frac{G(x)}{x} = x^2 - 16, \qquad \frac{G(x)}{x-4} = x^2 + 4x, \qquad \frac{G(x)}{x+4} = x^2 - 4x.$$

This gives, replacing x^k by b_k:

$$u(0) = -16b_0 + b_2,$$
$$u(4) = 4b_1 + b_2,$$
$$u(-4) = -4b_1 + b_2.$$

Consequently, if the matrix

$$\begin{pmatrix} -16 & 0 & 1 \\ 0 & 4 & 1 \\ 0 & -4 & 1 \end{pmatrix}$$

is multiplied by the matrix of the b_i (omitting b_3), we obtain the three eigenvectors u_i:

$$\begin{pmatrix} -16 & 0 & 1 \\ 0 & 4 & 1 \\ 0 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 13 & 4 & 7 \\ 28 & 24 & 12 \end{pmatrix} = \begin{matrix} 12 & 24 & 12 = u(0), \\ 80 & 40 & 40 = u(4), \\ -24 & 8 & -16 = u(-4). \end{matrix}$$

If the same matrix is multiplied by the matrix of the transposed iterations b_i^* (omitting b_3^*), we obtain the three adjoint eigenvectors u_i^*:

$$\begin{pmatrix} -16 & 0 & 1 \\ 0 & 4 & 1 \\ 0 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 13 & 5 & -23 \\ 28 & -4 & -20 \end{pmatrix} = \begin{matrix} 12 & -4 & -20 = u^*(0), \\ 80 & 16 & -112 = u^*(4), \\ -24 & -24 & 72 = u^*(-4). \end{matrix}$$

The solution of the entire eigenvalue problem is thus accomplished.

The method of minimized iterations. In principle, the previous discussions give a complete solution of the eigenvalue problem. We have found a systematic algorithm for the generation of the characteristic polynomial $G(\mu)$. The roots of this polynomial gave the eigenvalues of the matrix A. Then the process of synthetic division established the associated eigenvectors. Accidental deficiencies were possible but could be eliminated by additional trials.

As a matter of fact, however, the "progressive algorithm" of the last section has its serious limitations if large matrices are involved. Let us assume that there is considerable "dispersion" among the eigenvalues, which means that the ratio of the largest to the smallest eigenvalue is fairly large. Then the successive iterations will grossly increase the gap and after a few iterations the small eigenvalues will be practically drowned out. Let us assume, for example, that we have a 12-by-12 matrix which requires 12 iterations for the generation of the characteristic equation. The relatively mild ratio of 10:1 as the "spread" of the eigenvalues is after 12 iterations increased to the ratio 10^{12}:1, which means that we can never get through with the iteration scheme because the rounding errors make all iterations beyond the eighth entirely valueless.

As an actual example, taken from a physical situation, let us consider four eigenvalues which are distributed as follows:

$$1, \quad 5, \quad 50, \quad 2000.$$

Let us assume, furthermore, that we start with a trial vector which contains the four eigenvectors in the ratio of the eigenvalues, that is, the eigenvalue 2000 dominates with the amplitude 2000, compared with the amplitude of the eigenvalue 1. After one iteration the amplitude ratio is increased to $4 \cdot 10^6$, after two iterations to $8 \cdot 10^9$. The later iterations can give us no new information since they practically repeat the second iteration, multiplied every time

by the factor 2000. The small eigenvalues 1 and 5 are practically obliterated and cannot be rescued, except by an excessive accuracy which is far beyond the limitations of the customary digital machines.

We will now develop a modification of the customary iteration technique which obviates this difficulty. The modified scheme eliminates the rapid accumulation of rounding errors which under ordinary circumstances destroys the value of high order iterations. The new technique prevents the large eigenvalues from monopolizing the scene. It protects the small eigenvalues by constantly balancing the distribution of amplitudes in the most equitable fashion.

As an illustrative example, let us apply this method of "minimized iterations" to the above-mentioned dispersion problem. If the largest amplitude is normalized to 1, then the initial distribution of amplitudes is characterized as follows:

$$0.0005, \quad 0.0025, \quad 0.025, \quad 1.$$

Now, while an ordinary iteration would make this distribution still more extreme, the method of minimized iterations changes the distribution of amplitudes as follows:

$$0.0205, \quad 0.1023, \quad 1, \quad -0.0253.$$

We see that it is now the *third* eigenvector which gets a large weight factor, while the fourth eigenvector is almost completely in the background.

A repetition of the scheme brings about the following new distribution:

$$0.2184, \quad 1, \quad -0.1068, \quad 0.0000.$$

It is now the *second* eigenvector which gets the strongest emphasis.

The next repetition yields:

$$1, \quad -0.2181, \quad 0.0018, \quad 0.0000.$$

and we see that the weight is shifted over to the *smallest* eigenvalue.

After giving a chance to each eigenvalue, the scheme is exhausted, since we have all the information we need. Consequently the next minimized iteration yields an identical *vanishing* of the next vector, thus bringing the scheme to its natural conclusion.

In order to expose the principle of minimized iterations, let us first consider the case of *symmetric* matrices:

$$A^* = A. \tag{62}$$

Moreover, let us agree that the multiplication of a vector b by the matrix A shall be denoted by a prime:

$$Ab = b'. \tag{63}$$

Now our aim is to establish a linear identity between the iterated vectors. We cannot expect that this identity shall come into being right from the beginning. Yet, we can *approach* this identity right from the beginning by choosing such a linear combination of the iterated

vector b_0' and b_0 as makes the amplitude of the new vector as small as possible. Hence, we want to choose as our new vector b_1 the combination

$$b_1 = b_0' - \alpha_0 b_0, \tag{64}$$

where α_0 is determined by the condition that

$$(b_0' - \alpha_0 b_0)^2 = \text{minimum}. \tag{65}$$

This gives

$$\alpha_0 = \frac{b_0' b_0}{b_0^2}. \tag{66}$$

Notice that

$$b_1 \cdot b_0 = 0; \tag{67}$$

that is, the new vector b_1 is orthogonal to the original vector b_0.

We now continue our process. From b_1 we proceed to b_2 by choosing the linear combination

$$b_2 = b_1' - \alpha_1 b_1 - \beta_0 b_0, \tag{68}$$

and once more α_1 and β_0 are determined by the condition that b_2^2 shall become as small as possible. This gives

$$\alpha_1 = \frac{b_1' b_1}{b_1^2}, \qquad \beta_0 = \frac{b_1' b_0}{b_0^2}. \tag{69}$$

A good check of the iteration b_1' is provided by the condition

$$b_1' b_0 = b_1 b_0' = b_1^2. \tag{70}$$

Hence, the numerator of β_0 has to agree with the denominator of α_1.

The new vector b_2 is orthogonal to both b_0 and b_1.

This scheme can obviously be continued. The most remarkable feature of this successive minimization process is, however, that the best linear combination *never includes more than three terms*. If we form b_3, we would think that we should put

$$b_3 = b_2' - \alpha_2 b_2 - \beta_1 b_1 - \gamma_0 b_0. \tag{71}$$

But actually, in view of the orthogonality of b_2 to the previous vectors, we get

$$\gamma_0 = \frac{b_2' b_0}{b_0^2} = \frac{b_2 b_0'}{b_0^2} = 0. \tag{72}$$

Hence, every new step of the minimization process requires only two correction terms.

By this process a succession of orthogonal vectors is generated:[17]

$$b_0, b_1, b_2, \cdots, b_{m-1}, \tag{73}$$

until the identity relation becomes exact, which means that

$$b_m = 0. \tag{74}$$

If the matrix A is not symmetric, then we modify our procedure as follows. We operate simultaneously with A and A^*. The operations are the same as before, with the only difference

that the dot products are always formed between two *opposing* vectors. The scheme is indicated as follows:

$$
\begin{array}{cc}
b_0 & b_0^* \\
b_1 = b_0' - \alpha_0 b_0 & b_1^* = b_0^{*\prime} - \alpha_0 b_0^* \\
\alpha_0 = \dfrac{b_0' b_0^*}{b_0 b_0^*} = \dfrac{b_0^{*\prime} b_0}{b_0^* b_0} & \\
b_2 = b_1' - \alpha_1 b_1 - \beta_0 b_0 & b_2^* = b_1^{*\prime} - \alpha_1 b_1^* - \beta_0 b_0^* \\
\alpha_1 = \dfrac{b_1' b_1^*}{b_1 b_1^*} = \dfrac{b_1^{*\prime} b_1}{b_1^* b_1} & \\
\beta_0 = \dfrac{b_1' b_0^*}{b_0 b_0^*} = \dfrac{b_1^{*\prime} b_0}{b_0^* b_0} & \\
b_3 = b_2' - \alpha_2 b_2 - \beta_1 b_1 & b_3^* = b_2^{*\prime} - \alpha_2 b_2^* - \beta_1 b_1^* \\
\text{etc.} &
\end{array}
\tag{75}
$$

Operationally, the prime indicates multiplication by the matrix A. Hence, the succession of b_i vectors represents in fact a successive *set of polynomials*. Replacing A by the more familiar letter x, we have:

$$
\begin{aligned}
b_0 &= 1 \cdot b_0, \\
b_1 &= (x - \alpha_0) b_0, \\
b_2 &= (x - \alpha_1) b_1 - \beta_0 b_0, \\
b_3 &= (x - \alpha_2) b_2 - \beta_1 b_1, \\
&\cdots \\
b_m &= (x - \alpha_{m-1}) b_{m-1} - \beta_{m-2} b_{m-2} = 0.
\end{aligned}
\tag{76}
$$

This gradual generation of the characteristic polynomial $G(x)$ is in complete harmony with the procedure of the "progressive algorithm," discussed in the preceding section. In fact, *the successive polynomials of the set* (76) *are identical with the polynomials found in the full columns of the progressive algorithm in Table* 1. This explains the existence of the recursion relation

$$
p_{m+1}(x) = (x - \alpha_n) p_n(x) - \beta_{n-1} p_{n-1}(x) \tag{77}
$$

without additional γ, δ, . . . terms. The existence of such a relation is a characteristic feature of the *recurrent* set of equations that are at the basis of the entire development.

While the new scheme goes basically through the same steps as the previously discussed "progressive algorithm," it is in an incomparably stronger position concerning rounding errors. Apart from the fact that the rounding errors do not accumulate, we can effectively counteract their influence by constantly checking the mutual orthogonality of the gradually evolving vectors b_i and b_i^*. Any lack of orthogonality, caused by rounding errors, can immediately be corrected by the addition of a small correction term [Eq. (98)]. By this procedure the orthogonality of the generated vector system does not come gradually out of gear.

However, quite apart from the numerical advantages, the biorthogonality of the vectors

b_i and b_i^* has further appeal because *it imitates the behavior of the principal axes.* This is an analytical eminently valuable fact which makes the transition from the iterated vectors to the principal axes a simple and strongly convergent process.

In order to see the method in actual operation, let us apply it to the simple example of the preceding section. Here the matrix A is of third order, and thus we have to construct the vectors b_0, b_1, b_2, b_3, and the corresponding adjoint vectors. We obtain the following results:

$$b_0 = 1 \quad 0 \quad 0 \qquad b_0^* = 1 \quad 0 \quad 0$$

$$b_0' = 13 \quad 4 \quad 7 \qquad b_0^{*\prime} = 13 \quad 5 \quad -23$$

$$\alpha_0 = \frac{1 \cdot 13}{1} = 13$$

$$b_1 = 0 \quad 4 \quad 7 \qquad b_1^* = 0 \quad 5 \quad -23$$

$$b_1' = -141 \quad -28 \quad -79 \qquad b_1^{*\prime} = -141 \quad -69 \quad 279$$

$$\alpha_1 = \frac{1677}{-141} = -11.89361702 \qquad \beta_0 = \frac{-141}{1} = -141$$

$$b_2 = 0,\ 19.57446808,\ 4.25531914 \qquad b_2^* = 0,\ -9.53191490,\ 5.44680854$$

$$b_2' = 0,\ -17.02127656,\ 3.40425542 \qquad b_2^{*\prime} = 0,\ 16.34042562,\ -32.68085142$$

$$\alpha_2 = \frac{180.78768715}{-163.4042553} = -1.106382981 \qquad \beta_1 = \frac{-163.40425746}{-141} = 1.158895443$$

$$b_3 = 0, \quad 0, \quad 0 \qquad b_3^* = 0, \quad 0, \quad 0$$

The associated polynomials become:

$$p_0 = 1,$$

$$p_1(x) = x - 13,$$

$$p_2(x) = (x + 11.89361702)(x - 13) + 141$$

$$= x^2 - 1.10638298x - 13.61702126,$$

$$p_3(x) = (x + 1.106382981)(x^2 - 1.10638298x - 13.61702126) - 1.158895443(x - 13)$$

$$= x^3 - 16x.$$

Comparison with Table 1 shows that the coefficients of these very same polynomials appear in the full columns 0, 1, 2, 3 of the progressive algorithm.

Solution of the eigenvalue problem by the method of minimized iterations. The biorthogonal property of the vector system b_i, b_i^* leads to an explicit solution of the eigenvalue problem, in terms of the vectors b_i. Let us first assume that the matrix A is of the nondefective type and let us analyze the vectors b_i in terms of the eigenvectors u_i. The method by which the vectors b_i were generated yields directly the relation

$$b_i = p_i(\mu_1)u_1 + p_i(\mu_2)u_2 + \cdots + p_i(\mu_m)u_m. \tag{78}$$

If this relation is dotted with u_k^*, we obtain, in view of the mutual orthogonality of the two sets of axes,

$$b_i \cdot u_k^* = p_i(\mu_k) u_k \cdot u_k^*. \tag{79}$$

Let us now reverse the process and expand the u_i in terms of the b_i:

$$u_1 = \alpha_{i0} b_0 + \alpha_{i1} b_1 + \cdots + \alpha_{im-1} b_{m-1}. \tag{80}$$

The dotting by b_k^* yields

$$\alpha_{ik} = \frac{u_i \cdot b_k^*}{b_k \cdot b_k^*}. \tag{81}$$

Let us denote the "norm" of b_k by σ_k:

$$\sigma_k = b_k \cdot b_k^* \tag{82}$$

while the norm of u_k will be left arbitrary. Then the expansion (80) becomes:

$$u_i = \frac{b_0}{\sigma_0} + p_1(\mu_i) \frac{b_1}{\sigma_1} + p_2(\mu_i) \frac{b_2}{\sigma_2} + \cdots + p_{m-1}(\mu_i) \frac{b_{m-1}}{\sigma_{m-1}}. \tag{83}$$

This expansion contains the solution of the principal-axis problem. The eigenvectors u_i are generated in terms of the vectors b_i, which are the successive vectors of the process of minimized iterations. The expansion (83) takes the place of the previous "S-expansion" (27) which solved the eigenvector problem in terms of the customary process of iteration.

The adjoint axes are obtained in identical fashion:

$$u_i^* = \frac{b_0^*}{\sigma_0} + p_1(\mu_i) \frac{b_1^*}{\sigma_1} + p_2(\mu_1) \frac{b_2^*}{\sigma_2} + \cdots + p_{m-1}(\mu_i) \frac{b_{m-1}^*}{\sigma_{m-1}}. \tag{84}$$

The expansion (83) remains valid even in the case of defective matrices. The only difference is that the number of principal axes becomes less than n since a multiple root μ_j, if substituted into (83) and (84) cannot contribute more than one principal axis.[18] However, a defective matrix actually possesses less than n pairs of principal axes, and the above expansions give the general solution of the problem.

An interesting alternative of the expansion (83) arises if we go back to the original Fredholm problem and request a solution in terms of the minimized vectors b_i, rather than the simply iterated vectors of the expansion (17). One method would be to make use of the Schmidt series (8), expressing the u_i of that series in terms of the b_i, according to the expansion (83). However, the Schmidt series holds for nondefective matrices only, while we know that a solution must exist for any kind of matrix.

Hence, we prefer to proceed in a somewhat different fashion. We expand y directly in terms of the vectors b_i:

$$y = y_0 b_0 + y_1 b_1 + \cdots + y_{m-1} b_{m-1}. \tag{85}$$

We substitute this expansion into Eq. (1), replacing b_k' by

$$b_k' = b_{k+1} + \alpha_k b_k + \beta_{k-1} b_{k-1}. \tag{86}$$

Then we compare coefficients on both sides of the equation. The result can be described as follows.

Let us reverse the sequence of the α_i-coefficients and let us do the same with the β_i-coefficients. Hence, we define

$$\begin{aligned} \bar{\alpha}_0 &= \alpha_{m-1}, \\ \bar{\alpha}_1 &= \alpha_{m-2}, \ \bar{\beta}_0 = \beta_{m-2}, \\ &\cdots \\ \alpha_{m-1} &= \alpha_0, \quad \bar{\beta}_{m-2} = \beta_0. \end{aligned} \tag{87}$$

We now construct the following "reversed" set of polynomials:

$$\begin{aligned} \bar{p}_0 &= 1, \\ \bar{p}_1(x) &= x - \bar{\alpha}_0, \\ p_2(x) &= (x - \bar{\alpha}_1)\bar{p}_1(x) - \bar{\beta}_0, \\ &\cdots \\ \bar{p}_m(x) &= (x - \bar{\alpha}_{m-1})\bar{p}_{m-1}(x) - \bar{\beta}_{m-2}\bar{p}_{m-2}(x) \\ &= G(x). \end{aligned} \tag{88}$$

Then the solution of the Fredholm problem (1) is given by the expansion

$$y = \frac{\mu}{G(\mu)}[b_{m-1} + \bar{p}_1(\mu)b_{m-2} + \cdots + \bar{p}_{m-1}(\mu)b_0], \tag{89}$$

where we have put

$$\mu = \frac{1}{\lambda}. \tag{90}$$

The expansion (89) is *completely general* and remains valid, no matter how the vector b_0 of the right-hand member was given, and how regular or irregular the matrix A may be. The only condition to be satisfied is that the vector b_0*—while otherwise chosen arbitrarily—shall be free of accidental deficiencies, that is, b_0* shall not be orthogonal to some u_k if b_0 is not simultaneously orthogonal to u_k*.

The expansion (89) leads once more to a solution of the eigenvector problem, this time obtained with the help of the "reversed" polynomials $\bar{p}_i(x)$:

$$u_i = b_{m-1} + \bar{p}_1(\mu_i)b_{m-2} + \cdots + \bar{p}_{m-1}(\mu_i)b_0. \tag{91}$$

The expansions (91) and (83) actually coincide—except for a factor of proportionality—for algebraic reasons.

In order to see a numerical example for this solution of the eigenvalue problem let us return once more to the simple problem previously discussed. The minimized b_i and b_i* vectors associated with this matrix were given at the end of the preceding section, together

with the associated polynomials $p_i(x)$. We now construct the *reversed* polynomials $\bar{p}_i(x)$. For this purpose we tabulate the α_i and β_i:

13	— 141
— 11.89361702	1.158895443
— 1.106382981	

We reverse the sequence of this tabulation:

— 1.106382981	
— 11.89361702	1.58895443
13	— 141

and construct in succession

$$\begin{aligned}
\bar{p}_0 &= 1,\\
\bar{p}_1(x) &= x + 1.106382981,\\
\bar{p}_2(x) &= (x + 11.8361702)p_1(x) - 1.58895443.\\
&= x^2 + 13x + 12,\\
\bar{p}_3(x) &= (x - 13)\,\bar{p}_2(x) + 141\\
&= x^3 - 16x.
\end{aligned}$$

The last polynomial is identical with $p_3(x) = G(x)$. The zeros of this polynomial are

$$\mu_1 = 0, \qquad \mu_2 = 4, \qquad \mu_3 = -4;$$

substituting these values into $\bar{p}_2(\mu)$, $\bar{p}_1(\mu)$, $\bar{p}_0$ we obtain the matrix

$$\begin{pmatrix} 12 & 1.06382981 & 1\\ 80 & 5.106382981 & 1\\ -24 & -2.89361702 & 1 \end{pmatrix}.$$

The product of this matrix with the matrix of the b_i vectors gives the three principal axes u_i:

$$\begin{pmatrix} 12 & 1.06382981 & 1\\ 80 & 5.106382981 & 1\\ -24 & -2.89361702 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0\\ 0 & 4 & 7\\ 0 & 19.57446808 & 4.25531914 \end{pmatrix} = \begin{matrix} 12 & 24 & 12 = u(0),\\ 80 & 40 & 40 = u(4),\\ -24 & 8 & -16 = u(-4), \end{matrix}$$

in complete agreement with the previous result, but now obtained by an entirely different method. If the b-matrix is replaced by the b^*-matrix, the adjoint axes $u^*(0)$, $u^*(4)$, $u^*(-4)$ are obtained.

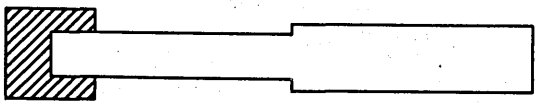


FIG. 1. Vibrating bar.

The lateral vibrations of a bar. In order to study the power of the method in connection with a vibration problem of large dispersion, the elastic vibrations of a bar were investigated. The bar was clamped at one end and free at the other. Moreover, the bar changed its cross section suddenly in the middle (Fig. 1). The change of the cross section was such that the

moment of inertia jumped from the value 1 to 2. The differential equation that describes the vibrations of such a bar is the fourth order equation

$$\frac{d^2}{dx^2}\left[k(x)\frac{dy^2}{dx^2}\right] = \lambda y, \tag{92}$$

with the boundary conditions

$$\begin{aligned} y(0) &= 0, \qquad y''(l) = 0, \\ y'(0) &= 0, \qquad y'''(l) = 0, \end{aligned} \tag{93}$$

and

$$k(x) = 1\left(0 \leq x < \frac{l}{2}\right), \quad k(x) = 2\left(\frac{l}{2} < x \leq l\right). \tag{94}$$

The differential operator d/dx was replaced by the difference operator $\Delta/\Delta x$, with $\Delta x = 1$. The length of the bar was chosen as $l = 13$, thus leading to a 12-by-12 matrix, since $y(0) = y(1) = 0$.

The first step was the inversion of the matrix. This was easily accomplished since a matrix that is composed of a narrow band around the diagonal can be inverted with little labor. The eigenvalues μ_i of the inverted matrix are the *reciprocals* of the original λ_i:

$$\mu_i = \frac{1}{\lambda_i}. \tag{95}$$

The general theory has shown that the iteration scheme applied to an arbitrary matrix automatically yields a biorthogonal set of vectors b_i and b_i^*; they can be conceived as the building blocks from which the entire set of principal axes may be generated. In the present problem, dissipative forces are absent, which makes the matrix A symmetric and the problem self-adjoint. Hence,

$$b_i = b_i^*, \tag{96}$$

and we get through with a single set of iterations.

Now the general procedure would demand that we go through 12 minimized iterations before the stage $b_{12} = 0$ is attained. However, the study of a system with high dispersion has shown that in such a system the method of minimized iterations practically separates the various vibrational moves, starting with the highest eigenvalue and descending systematically to the lower eigenvalues, provided that we employ a trial vector b_0 which weights the eigenvectors according to the associated eigenvalues, or even more strongly. In the present problem the trial vector 1, 0, 0, . . . was not used directly but iterated with the matrix A, and then iterated again. The vector b_0'' thus obtained was employed as the b_0 of the minimized iteration scheme.

The strong grading of the successive eigenvectors has the consequence that in k minimized iterations essentially only the highest k vibrational modes will come into evidence. This is of eminent practical value since it allows us to dispense with the calculation of the very low eigenvalues (that is, very high frequencies, since we speak of the eigenvalues of the *inverted* matrix), which are often of little physical interest, and also of little mathematical interest in

view of the fact that the replacing of the d operator by the Δ operator becomes in the realm of high frequencies more and more damaging.

Whether the isolation actually takes place or not can be tested with the help of the $p_i(x)$ polynomials that accompany the iteration scheme. The order of these polynomials constantly increases by 1. The correct eigenvalues of the matrix A are obtained by evaluating the zeros of the *last* polynomial $p_m(x) = 0$. What actually happens, however, is that the zeros of the polynomials $p_i(x)$ do not change much from the beginning. If the dispersion is strong, then each new polynomial basically adds one more root but corrects the higher roots by only small amounts. It is thus quite possible that the series of largest roots in which we are primarily interested is practically established with sufficient accuracy after a few iterations. Then we can stop, since the later iterations will change the values obtained by negligible amounts. The same can be said about the vibrational modes associated with these roots.

This consideration suggests the following successive procedure for the approximate determination of the eigenvalues and eigenvectors (vibrational modes) of a matrix. As the minimization scheme proceeds and we constantly obtain newer and newer polynomials $p_i(x)$, we handle the last polynomial obtained as if it were the final polynomial $p_m(x)$. We evaluate the roots of this polynomial and compare them with the previous roots. Those roots which change by negligible amounts are already in their final form.

A similar procedure holds for the evaluation of the eigenvectors u_i. Here the biorthogonality of the vectors b_i and b_i^*—which is reduced to simple orthogonality in the case of a symmetric matrix—is of very great help. Let us assume that the lengths of the vectors b_i are normalized to 1, by replacing b_i by $b_i/\sqrt{\sigma_i}$. Then the expansions (83) and (84) show that the following matrix must be an orthogonal—although in the diagonal terms not normalized—matrix:

$$\left\{\begin{array}{cccc} \frac{1}{\sqrt{\sigma_0}} & \frac{p_1(\mu_1)}{\sqrt{\sigma_1}} & \frac{p_2(\mu_1)}{\sqrt{\sigma_2}} & \cdots \frac{p_{m-1}(\mu_1)}{\sqrt{\sigma_{m-1}}} \\ \frac{1}{\sqrt{\sigma_0}} & \frac{p_1(\mu_2)}{\sqrt{\sigma_1}} & \frac{p_2(\mu_2)}{\sqrt{\sigma_2}} & \cdots \frac{p_{m-1}(\mu_2)}{\sqrt{\sigma_{m-1}}} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{1}{\sqrt{\sigma_0}} & \frac{p_1(\mu_m)}{\sqrt{\sigma_1}} & \frac{p_2(\mu_m)}{\sqrt{\sigma_2}} & \cdots \frac{p_{m-1}(\mu_m)}{\sqrt{\sigma_{m-1}}} \end{array}\right\} \quad . \quad . \quad . \quad . \quad (97)$$

The dot-product of any two rows of this matrix must come out as zero—thus providing us with a powerful check on the construction of the p_i polynomials and the correctness of the roots μ_1, which are the roots of the equation $p_m(\mu) = 0$.[19] In the case of strong dispersion, the transformation matrix (97) is essentially reduced to the diagonal terms and one term to the right and to the left of the diagonal; that is, the eigenvector u_k is essentially a linear combination of three b-vectors only, namely, b_{k-2}, b_{k-1}, and b_k.

These general conditions are well demonstrated by the tabulation of the final results of the above-mentioned bar problem. The minimized iterations were carried through up to $m = 6$. On the basis of these iterations, the first six eigenvalues and the first five vibrational modes of the clamped-free bar were evaluated. The iterations were constantly watched for orthogonality. After obtaining a certain b_i, this b_i was immediately dotted with all the previous b_j. If a certain dot product $b_i \cdot b_j$ came out as noticeably different from zero, the correction term

$$\varepsilon_{ij} = -\frac{b_i \cdot b_j}{b_j^2} b_j \tag{98}$$

was added to b_i, thus compensating for the influence of rounding errors. By this procedure the ten-significant-figure accuracy of the calculations was constantly maintained.[20]

The roots of the successive polynomials $p_i(x)$ are tabulated in Table 2.

Table 2.

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
2256.926071					
.943939	48.1610705				
.943939	.2037755	5.272311428			
.943939	.2037825	.355958260	1.513923859		
.943939	.2037825	.356269794	.582259337	0.546327303	
.943939	.2037825	.356269980	.5829955952	.591117817	0.2498132719

The successive orthogonal transformation matrices (97) likewise show strong convergence. We tabulate here only the last computed transformation matrix (rounded off to four decimal places), which expresses the first six eigenvectors $u_1, \ldots, u_6$ in terms of the first six normalized $b_i/\sqrt{\sigma_i}$ vectors, making use of the roots of $p_6(u) = 0$. The diagonal elements are normalized to 1:

$$\begin{pmatrix} 1 & 0.0028 & 0 & 0 & 0 & 0 \\ -0.0028 & 1 & 0.0316 & 0.0004 & 0 & 0 \\ 0 & -0.0316 & 1 & 0.1497 & 0.0081 & 0 \\ 0 & 0.0044 & -0.1520 & 1 & 0.2693 & 0.0249 \\ 0 & -0.0010 & 0.0335 & -0.2793 & 1 & 0.4033 \\ 0 & 0.0002 & -0.0087 & 0.0779 & -0.3816 & 1 \end{pmatrix}.$$

We notice how quickly the elements fall off to zero as soon as we are beyond one element to the right and one to the left of the main diagonal. The orthogonal reference system of the b_i and the orthogonal reference system of the u_i are thus in close proximity to each other.

The five vibrational modes $u_1, \ldots, u_5$ thus obtained (u_6 being omitted since the lack of the neighbor on the right side of the diagonal makes the approximation unreliable) are plotted in Fig. 2.

The eigenvalue problem of linear integral operators. The methods and results of the past sections can now be applied to the realm of continuous operators. The kernel of an integral equation can be conceived as a matrix of infinite order that may be approximated to any degree of accuracy by a matrix of high but finite order. One method of treating an integral equation is to replace it by an ordinary matrix equation of sufficiently high order. This procedure is from the numerical standpoint frequently the most satisfactory one. However, we can design methods for the solution of integral equations that obtain the solution by purely analytical tools, on the basis of an infinite convergent expansion, such as the Schmidt series, for example. The method we are going to discuss belongs to the latter type. We will find an expansion that is based on the same kind of iterative integrations as the Liouville-Neumann series, but avoiding the convergence difficulties of that expansion. The expansion we are going to develop *converges under all circumstances* and gives the solution of any Fredholm type of integral equation, no matter how defective the kernel of that integral equation may be.[21]

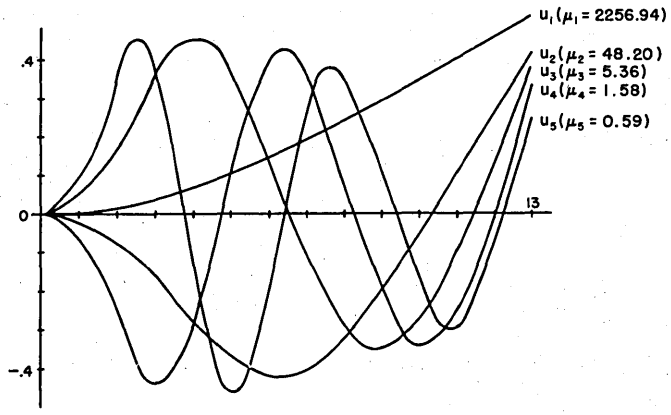


FIG. 2. Five vibrational modes of the bar of Fig. 1.

Let us first go back to our earlier method of solving the Fredholm problem. The solution was obtained as the S-expansion (17). The difficulty with this solution is that it is based on the linear identity that can be established between the iterated vectors b_i. That identity is generally of the order n; if n grows to infinity, we have to obtain an identity of infinite order before our solution can be constructed. That, however, cannot be done without the proper adjustments.

The later attempt, based on the method of minimized iterations, employs more adequate principles. We have seen that for any matrix A a biorthogonal set of vectors b_i and b_i^* can be constructed by successive minimizations. The set is uniquely determined as soon as the first trial vectors b_0 and b_0^* are given. In the case of the inhomogeneous equation (1) the right-hand member b may be chosen as the trial vector b_0 while b_0^* is still arbitrary.

The construction of these two sets of vectors is quite independent of the order n of the matrix. If the matrix becomes an integral operator, the b_i and b_i^* vectors are transformed into a biorthogonal set of functions

$$\begin{aligned} &\varphi_0(x), \quad \varphi_1(x), \quad \varphi_2(x), \ \ldots \\ &\varphi_0^*(x), \quad \varphi_1^*(x), \quad \varphi_2^*(x), \ \ldots \end{aligned} \tag{99}$$

which are generally present in infinite number. The process of minimized iterations assigns to *any* integral operator such a set, after $\varphi_0(x)$ and $\varphi_0^*(x)$ have been chosen.

Another important feature of the process of minimized iterations was the appearance of a successive set of polynomials $p_i(\mu)$, tied together by the recursion relation

$$p_{i+1}(\mu) = (\mu - \alpha_i)p_i(\mu) - \beta_{i-1}p_{i-1}(\mu). \tag{100}$$

This is again entirely independent of the order n of the matrix A and remains true even if the matrix A is replaced by a Fredholm kernel $K(x, \xi)$.

We can now proceed as follows. We stop at an arbitrary $p_m(x)$ and form the reversed set of polynomials $\bar{p}_i(\mu)$, defined by the process (87). Then we construct the expansion

$$y_m(x) = \frac{\mu}{p_m(\mu)} [\varphi_{m-1}(x) + \bar{p}_1(\mu)\varphi_{m-2}(x) + \cdots + \bar{p}_{m-1}(\mu)\varphi_0(x)]. \tag{101}$$

This gives a successive-approximation process that converges well to the solution of the Fredholm integral equation

$$y(x) - \lambda K y(x) = \varphi_0(x). \tag{102}$$

In other words,

$$y(x) = \lim_{m\to\infty} y_m(x). \tag{103}$$

By the same token we can obtain all the eigenvalues and eigensolutions of the kernel $K(x, \xi)$, if such solutions exist. For this purpose we obtain the roots μ_i of the polynomial $_m(\mu)p$ by solving the algebraic equation

$$p_m(\mu) = 0. \tag{104}$$

The exact eigenvalues μ_i of the integral operator $K(x, \xi)$ are obtained by the limit process:

$$\lim_{m\to\infty} p_m(\mu_1) = 0, \tag{105}$$

where the largest root is called μ_1 and the subsequent roots are arranged according to their absolute values. The corresponding eigenfunctions are given by the infinite expansion

$$u_i(x) = \lim_{m\to\infty} \left[\frac{\varphi_0(x)}{\sigma_0} + p_i(\mu_i)\frac{\varphi_1(x)}{\sigma_1} + \cdots + p_{m-1}(\mu_i)\frac{\varphi_{m-1}(x)}{\sigma_{m-1}}\right], \tag{106}$$

where μ_i is the ith root of the polynomial $p_m(\mu)$.[22]

As a trial function $\varphi_0(x)$ we may choose, for example,

$$\varphi_0 = \text{const.} = 1. \tag{107}$$

However, the convergence is greatly speeded up if we first apply the operator K to this function, and possibly iterate even once more. In other words, we should choose $\varphi_0 = K \cdot 1$, or even $\varphi_0 = K^2 \cdot 1$ as the basic trial function of the expansion (106).

We consider two particularly interesting examples which are well able to illustrate the nature of the successive-approximation process here discussed.

The vibrating plate. In the problem of the vibrating plate we encounter the self-adjoint differential operator

$$-\frac{d}{dx}(xy'). \tag{108}$$

This leads to the eigenvalue problem

$$-\frac{d}{dx}(xy') = \lambda y, \qquad (0 \leq x \leq 1) \qquad (109)$$

with the boundary condition

$$y(1) = 0. \qquad (110)$$

The solution of the differential equation (109) is

$$y = J_0(2\sqrt{\lambda x}), \qquad (111)$$

where $J_0(x)$ is the Bessel function of order zero. The boundary condition (110) requires that λ shall be chosen as follows:

$$\lambda = \frac{\xi_i^2}{4}, \qquad (112)$$

where ξ_i are the zeros of $J_0(x) = 0$.

Now the Green's function of the differential equation (109) changes the differential operator (108) to the inverse operator which is an integral operator of the nature of a symmetric Fredholm kernel function $K(x, \xi)$. Our problem will be to obtain the eigenvalues and eigenfunctions of this kernel.

If we start out with the function $\varphi_0(x) = 1$, the operation $K\varphi_0$ gives

$$1 - x,$$

and repeating the operation we obtain

$$\frac{x^2}{4} - x + \frac{3}{4},$$

and so on. The successive iterations will be *polynomials* in x. Now the minimized iterations are merely some linear combinations of the ordinary iterations. Hence the orthogonal sequence $\varphi_i(x)$ will become a sequence of polynomials of constantly increasing order, starting with the constant $\varphi_0 = 1$. This singles out the $\varphi_k(x)$ as the *Legendre polynomials* $P_k(x)$, but normalized to the range 0 to 1, instead of the customary range -1 to $+1$. The renormalization of the range transforms the polynomials $P_k(x)$ into Jacobi polynomials $G_k(p, q; x)$, with $p = q = 1$,[23] which again are special cases of the Gaussian hypergeometric series $F(\alpha, \beta, \gamma; x)$, in the sense of $F(k + 1, -k, 1; x)$; hence, we get:

$$\begin{aligned} \varphi_0 &= 1, \\ \varphi_1(x) &= 1 - 2x, \\ \varphi_2(x) &= 1 - 6x + 6x^2, \\ \varphi_3(x) &= 1 - 12x + 30x^2 - 20x^3, \\ &\cdots \end{aligned} \qquad (113)$$

The associated polynomials $p_i(x)$ can be obtained on the basis of the relation

$$K\varphi_m = \varphi_{n+1} + \alpha_n\varphi_n + \beta_{n-1}\varphi_{n-1}. \qquad (114)$$

This gives

$$\begin{aligned} p_0 &= 1, \\ p_1(x) &= 2x - 1, \\ p_2(x) &= 24x^2 - 18x + 1, \\ p_3(x) &= 720x^3 - 600x^2 + 72x - 1. \end{aligned} \tag{115}$$

In order to obtain the general recursion relation for these polynomials it is preferable to follow the example of the algorithm of Table 1 and introduce the "half-columns" in addition to the full columns. Hence, we define a second set of polynomials $q_k(x)$ and set up the recursion relations

$$\begin{aligned} p_n(x) &= nxq_{n-1}(x) - p_{n-1}(x), \\ q_n(x) &= 2(2n + 1)p_n(x) - q_{n-1}(x). \end{aligned} \tag{116}$$

We thus obtain, starting with $p_0 = 1$ and $q_0 = 2$, and using successive recursions:

$$\begin{aligned} p_0 &= 1, & q_0 &= 2 \\ p_1(x) &= 2x - 1, & q_1(x) &= 12x - 8, \\ p_2(x) &= 24x^2 - 18x + 1, & q_2(x) &= 240x^2 - 192x + 18, \\ p_3(x) &= 720x^3 - 600x^2 + 72x - 1, & q_3(x) &= 10080x^3 - 8640x^2 + 1200x - 32, \\ &\vdots \quad . & &. \quad . \quad . \quad . \end{aligned} \tag{117}$$

The zeros of the $p_m(x)$ polynomials converge to the eigenvalues of our problem, but the convergence is somewhat slow since the original function $\varphi_0 = 1$ does not satisfy the boundary conditions and thus does not suppress sufficiently the eigenfunctions of high order. The zeros of the q_i polynomials give quicker convergence. They are given in Table 3, going up to $q_5(x)$.

Table 3.

0.6677				
.69155	0.1084			
.69166016	.130242	.035241		
.6916602716	.1312564	.051130	.014842	
.6916602760	.13127115	.0532914	.025582	.00729
.6916602761	.13127123	.0534138	.028769	.01794

The last row contains the correct values of the eigenvalues, computed on the basis of Eq. (112),

$$\mu_i = \frac{4}{\xi_i^2}. \tag{118}$$

We notice the eminent convergence of the scheme.

The question of the eigenfunctions of our problem will not be discussed here.

The vibrating string: even modes. Another interesting example is provided by the vibrating string. The differential operator here is

$$-\frac{d^2}{dx^2} \tag{119}$$

with the boundary conditions

$$y(\pm 1) = 0. \tag{120}$$

The solution of the differential equation

$$-\frac{d^2}{dx^2}y = \lambda y \tag{121}$$

under the given boundary conditions is

$$y_i = \cos (2i + 1)\frac{\pi}{2}x, \qquad \text{(even modes)} \tag{122}$$

$$y_j = \sin j\pi x. \qquad \text{(odd modes)} \tag{123}$$

This gives the eigenvalues

$$\lambda_i = \left(\frac{2i + 1}{2}\pi\right)^2 \qquad \text{(even modes)} \tag{124}$$

and

$$\lambda_j = (j\pi)^2. \qquad \text{(odd modes)} \tag{125}$$

If we start with the trial function $\varphi_0 = 1$, we will get all the even vibrational modes of the string, while $\varphi_0 = x$ will give all the odd vibrational modes. We start with the first alternative.

Successive iterations give

$$\begin{aligned} K \cdot 1 &= \frac{x^2}{2} - \frac{1}{2}, \\ K^2 \cdot 1 &= \frac{x^4}{24} - \frac{x^2}{4} + \frac{5}{24}, \end{aligned} \tag{126}$$

and we notice that the minimized iterations will now become a sequence of *even* polynomials. The transformation $x^2 = \xi$ shows that these polynomials are again Jacobi polynomials $G_k(p, q; x^2)$, but now $p = q = \frac{1}{2}$, and we obtain the hypergeometric functions $F(k + \frac{1}{2}, -k, \frac{1}{2}; x^2)$:

$$\begin{aligned} \varphi_0 &= 1, \\ \varphi_1(x) &= 1 - 3x^2, \\ \varphi_2(x) &= 3 - 30x^2 + 35x^4, \\ \varphi_3(x) &= 5 - 105x^2 + 315x^4 - 231x^6, \\ &\cdots \end{aligned} \tag{127}$$

Once more we can establish the associated polynomials $p_i(x)$, and the recursion relation by which they can be generated. In the present case the recursion relations come out to be

$$\begin{aligned} p_n(x) &= (4n - 1)xq_{n-1}(x) - p_{n-1}(x), \\ q_n(x) &= (4n + 1)p_n(x) - q_{n-1}(x), \end{aligned} \tag{128}$$

starting with $p_0 = 1$, $q_0 = 1$. This yields

$$\begin{aligned} p_0 &= 1, & q_0 &= 1, \\ p_1(x) &= 3x - 1, & q_1(x) &= 15x - 6, \\ p_2(x) &= 105x^2 - 45x + 1, & q_2(x) &= 945x^2 - 420x + 15, \\ p_3(x) &= 10395x^3 - 4725x^2 + 210x - 1, & q_3(x) &= 135135x^3 - 62370x^2 + 3150x - 28, \\ &\cdots & &\cdots \end{aligned} \tag{129}$$

Table 4.

μ_1	μ_2	μ_3	μ_4	μ_5
0.40000				
.405059	0.02351			
.405284733	.044856	0.011397		
.4052847346	.04503010	.015722	0.00455	
.4052847346	.0450316322	.016192	.00752	0.00216
.4052847346	.0450316371	.016211	.00827	.00500

The successive zeros of the $q_i(x)$ polynomials, up to $q_5(x)$, are given in Table 4. The last row contains the correct eigenvalues, calculated from the formula

$$\mu_i = \left(\frac{2}{2i+1}\frac{1}{\pi}\right)^2. \tag{130}$$

The convergence is again very conspicuous.

The vibrating string: odd modes. In the case of the odd modes of the vibrating string the orthogonal functions of the minimized iterations are again related to the Jacobi polynomials $G_k(p, q; x)$, but now with $p = q = \frac{3}{2}$. Expressed in terms of the hypergeometric series we now get the polynomials of odd orders $xF(k + \frac{3}{2}, -k, \frac{3}{2}; x)$:

$$\begin{aligned} \varphi_0 &= x, \\ \varphi_1(x) &= 3x - 5x^3, \\ \varphi_2(x) &= 15x - 70x^3 + 63x^5, \\ \varphi_3(x) &= 35x - 315x^3 + 693x^5 - 429x^7, \end{aligned} \tag{131}$$

The associated $p_i(x)$ polynomials are generated by the recursion relations

$$\begin{aligned} p_n(x) &= (4n+1)xq_{n-1}(x) - p_{n-1}(x), \\ q_n(x) &= (4n+3)p_n(x) - q_{n-1}(x), \end{aligned} \tag{132}$$

starting with $p_0 = 1$, $q_0 = 3$. We thus get

$$\begin{aligned} p_0 &= 1, & q_0 &= 3, \\ p_1(x) &= 15x - 1, & q_1(x) &= 105x - 10, \\ p_2(x) &= 945x^2 - 105x + 1, & q_2(x) &= 10395x^2 - 1260x + 21, \\ p_3(x) &= 135135x^3 - 17325x^2 + 378x - 1, & q_3(x) &= 2027025x^3 - 270270x^2 + 6930x - 36. \end{aligned} \tag{133}$$

The zeros of $q_i(x)$, up to $q_5(x)$, are given in Table 5. The last row contains the correct eigenvalues calculated on the basis of the formula

$$\mu_i = \frac{1}{i^2\pi^2}. \tag{134}$$

Table 5.

μ_1	μ_2	μ_3	μ_4	μ_5
0.0952				
.10126	0.01995			
.10132106	.02500	0.00701		
.1013211836	.025323	.01068	0.00307	
.1013211836	.02533024	.011215	.00550	0.00156
.1013211836	.025330296	.011258	.00633	.00405

The eigenvalue problem of linear differential operators. Let $Dy(x)$ be a given linear differential operator, with given homogeneous boundary conditions of sufficient number to establish an eigenvalue problem. The problem of finding the eigenvalues and eigenfunctions of this operator is equivalent to the problem of the previous section in which the eigenvalue problem of linear integral operators was investigated. Let us assume that we know the Green's function $K(x, \xi)$ of the differential equation

$$Dy = \rho. \tag{135}$$

Then K is the reciprocal operator of D which possesses the same eigenfunctions (principal axes) as the operator D, while the eigenvalues of K are the reciprocals of the eigenvalues of D.

Hence, in principle, the eigenvalue problem of differential operators needs no special investigation. Actually, however, the situation in most cases is far less simple. The assumption that we are in possession of the Green's function associated with the differential equation (135) is often of only purely theoretical significance. Even very simple differential operators have Green's functions that are outside the limits of our analytical possibilities. Moreover, even if we do possess the integral operator K in closed form, it is still possible that the successive integrations needed for the construction of the successive orthogonal functions $\varphi_1(x)$, $\varphi_2(x)$, $\varphi_3(x)$, . . . go beyond our analytical facilities.

In view of this situation we ask the question whether we could not relax some of the practically too stringent demands of the general theory. We may lose somewhat in accuracy, but we may gain tremendously in analytic operations if we can replace some of the demands of the general theory by more simplified demands. The present section will show how that may actually be accomplished.

Leaving aside the method of minimized iterations, which was merely an additional tool in our general program, the basic principle of our entire investigation, if shaped to the realm of integral operators, may be formulated as follows.

We start out with a function $f_0(x)$ which may be chosen as $f_0(x) = 1$. We then form by iterated integrations a set of new functions

$$f_1(x) = Kf_0(x), f_2(x) = Kf_1(x), \ldots, f_m(x) = Kf_{m-1}(x). \tag{136}$$

Then we try to establish an approximate linear relation between these functions, as accurately as possible. For this purpose we make use of the method of least squares.

We notice that the general principle involves two processes: (*a*) the construction of the iterated set (136); (*b*) the establishment of a close linear relation between them. It is in the first process that the knowledge of the integral operator $K = D^{-1}$ is demanded. But let us observe that the relation between the successive f_i-functions can be stated in reverse order. We then get

$$f_m(x), f_{m-1}(x) = Df_m(x), \ldots, f_0(x) = Df_1(x). \tag{137}$$

If we start with the function $f_m(x)$, then the successive functions of lower order can be formed with the help of the given D operator and we can completely dispense with the use of the Green's functions.

Now the freedom of choosing $f_0(x)$ makes also $f_m(x)$ to some extent a free function. Yet the successive functions $f_i(x)$ do not have the same degree of freedom. While $f_0(x)$ need not satisfy the given boundary conditions, $f_1(x)$ of necessity satisfies these conditions, while $f_2(x)$ satisfies them even more strongly, since not only $f_2(x)$ but even $Df_2(x)$ satisfies the given boundary conditions. Generally, we can say that an arbitrary $f_n(x)$ need not satisfy any definite differential or integral equation but it is very restricted in the matter of boundary conditions; it has to satisfy the boundary conditions "in nth order." This means that not only $f_n(x)$ itself, but the whole sequence of functions

$$f_n(x), Df_n(x), D^2f_n(x), \ldots, D^{n-1}f_n(x) \tag{138}$$

must satisfy the given boundary conditions.

To construct a function $f_n(x)$ of this property is not too difficult. We expand $f_n(x)$ into a linear set of powers, or periodic functions, or any other kind of function we may find adequate to the given problem. The coefficients of this expansion will be determined by the boundary conditions that will be satisfied by $f_n(x)$ and the iterated functions (138). This leads to the solution of linear equations. In fact, this process can be systematized to a regular recursion scheme that avoids the accumulation of simultaneous linear equations, replacing them by a set of separated equations, each one involving but one unknown.

We have thus constructed our set (136), although in reverse order. We did not use any integrations, only the repeated application of the given differential operator D. The first phase of our problem is accomplished.

We now turn to the *second* phase of our program, namely, the establishment of an approximate linear relation between the iterated functions $f_i(x)$. The method of least squares is once more at our disposal. However, here again we might encounter the difficulty that the definite integrals demanded for the evaluation of the α_i and β_i are practically beyond our means. Once more we can simplify our task. The situation is similar to that of evaluating the

coefficients of a Fourier series. The "best" coefficients, obtained by the method of least squares, demand the evaluation of a set of definite integrals. Yet we can get a practically equally close approximation of a function $f(x)$ by a finite trigonometric series $\overline{f(x)}$, if we use the method of "trigonometric interpolation." Instead of minimizing the mean square of the error $f(x) - \overline{f(x)}$, we make $\overline{f(x)}$ equal to $f(x)$ in a sufficient number of equidistant points. This leads to no integrations but to simple summations.

The present situation is quite analogous. To establish a linear relation between the $f_i(x)$ means that the last function $f_m(x)$ shall be approximated by a linear combination of the previous functions. Instead of using the method of least squares for this approximation we can use the much simpler method of *interpolation*, by establishing a linear relation between the successive $f_i(x)$ in as many equidistant points as we have coefficients at our disposal. For the sake of better convergence, it is preferable to omit $f_0(x)$—which does not satisfy the boundary conditions and thus contains the high vibrational modes too pronouncedly—and establish the linear relation only from $f_1(x)$ on. For example, if we constructed a trial function $f_3(x)$ which, together with the iterated $f_2(x) = Df_3(x)$ and doubly iterated $f_1(x) = D^2f(x)$ satisfies the given boundary conditions, then we can choose two points of the region, say the two endpoints, where a linear relation of the form

$$f_3(x) + \alpha f_2(x) + \beta f_1(x) = 0 \tag{139}$$

shall hold. This gives the characteristic polynomial $G(x)$ in the form

$$G(x) = x^2 + \alpha x + \beta. \tag{140}$$

The two roots of this polynomial give us an approximate evaluation of the two highest μ_i (or the two lowest λ_i), that is, $\mu_1 = 1/\lambda_1$ and $\mu_2 = 1/\lambda_2$, while the corresponding eigensolutions are obtained by synthetic division:

$$\begin{aligned} \frac{G(x)}{x - \mu_1} &= g_1'x + g_2', \\ \frac{G(x)}{x - \mu_2} &= g_1''x + g_2'', \end{aligned} \tag{141}$$

which gives

$$\begin{aligned} u_1(x) &= g_1'f_2(x) + g_i'f_1(x), \\ u_2(x) &= g_1''f_2(x) + g_2''f_1(x). \end{aligned} \tag{142}$$

(The last root and its eigenfunction are always considerably in error, and give only rough indications.)

The remarkable feature of this method is that it *completely avoids any integrations*, requiring only the solution of a relatively small number of linear equations.

The following application of the method demonstrates its practical usefulness. The method was applied to obtain the first three eigenvalues of the lateral vibrations of a uniform bar, clamped at both ends. The given differential operator is here

$$Dy = \frac{d^4y}{dx^4},$$

with the boundary conditions

$$y(\pm 1) = 0, \qquad y'(\pm 1) = 0.$$

Only the even modes were considered, expanding $y(x)$ into even powers of x. The approximations were carried out in the first, second, and third orders. The eigenvalues obtained are given in Table 6, the last row containing the correct eigenvalues given by Rayleigh.[24]

Table 6.

μ_1	μ_2	μ_3
0.0323413		
.0319686	0.0007932	
.031963958	.0010875	0.0000788
0.031963996	0.0010946	0.0001795

We notice that the general convergence behavior of this method is exactly the same as that of the analytically more advanced, but practically much more cumbersome, method of minimized iterations.

Differential equations of second order: Milne's method. If a linear differential equation of second order with two-end boundary conditions is changed into a difference equation and then handled as a matrix problem, singularly favorable conditions exist for the solution of the eigenvalue problem. The matrix of the corresponding difference equation contains only diagonal terms plus one term to the right and one to the left. If we now start to iterate with the trial vector

$$b_0 = 1, 0, 0, \ldots, 0, \tag{143}$$

we observe that the successive iterations grow by one element only, as indicated in the following scheme, where the dots stand for the nonvanishing components:

$$\begin{aligned}
b_0 &= \cdot, \\
b_1 &= \cdot\ \cdot, \\
b_2 &= \cdot\ \cdot\ \cdot, \\
&\ \ \vdots \\
b_{n-1} &= \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot, \\
b_n &= \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot.
\end{aligned}$$

Under these conditions the establishment of the linear identity between the iterated vectors is greatly simplified since it is available by a successive recursion scheme. The coefficients of the equation

$$b_n + g_1 b_{n-1} + g_2 b_{n-2} + \cdots + g_n b_0 = 0 \tag{145}$$

are directly at our disposal, since the last column of the last two vectors gives g_1, then the

previous column gives g_2, . . ., until finally the first column gives g_n. The construction of the basic polynomial $G(x)$ is thus accomplished and the eigenvalues[25] λ_i are directly available by finding the roots of the equation $G(\lambda) = 0$.

Professor W. E. Milne, of the Oregon State College and the Institute for Numerical Analysis, applied the general theory to this problem, but with the following modification. Instead of iterating with the given matrix A, Milne considers the regular vibration problem

$$\frac{\partial^2 u}{\partial t^2} + Du = 0, \tag{146}$$

where the operator D has the following significance:[26]

$$Du = -\left[\frac{d^2}{dx^2} + p(x)\frac{d}{dx} + g(x)\right] u(x, t). \tag{147}$$

The differential equation (146) is now converted into a difference equation, with $\Delta x = \Delta t = h$. Then the values of $u(ih, jh)$ are determined by successive recursions, starting from the initial conditions

$$u(ih, 0) = 1, 0, 0, 0, \ldots, 0 \tag{148}$$

and

$$u(ih, h) = u(ih, -h). \tag{149}$$

The linear identity between the $n + 1$ vectors

$$u(ih, 0), u(ih, h), \ldots, u(ih, nh) \tag{150}$$

leads to a trigonometric equation for the characteristic frequencies ν_i, of the form

$$\cos n\nu_i h + A_{n-1}\cos(n-1)\nu_i h + \cdots + A_0 = 0. \tag{151}$$

We then put

$$\lambda_i = \nu_i^2. \tag{152}$$

On the other hand, the regular iteration method gives the eigenvalues $\bar{\lambda}_i$ of the operator Δu, defined in harmony with the operator Du but with the modification that the operation d/dx is replaced by the operation $\Delta/\Delta x$. The $\bar{\lambda}_i$ are in the following relation to the ν_i of Eqs. (151) and (152):

$$\bar{\lambda}_i = \bar{\nu}_i^2 = \left(\frac{\sin \frac{1}{2}h\nu_i}{\frac{1}{2}h}\right)^2. \tag{153}$$

It is of interest to see that the values (152) of Milne are much closer to the true eigenvalues than are the values obtained by iterations. The values of Milne remain good even for high frequencies while the iteration method gives gradually worse results; this is to be expected since the error committed by changing the differential equation to a difference equation must come into evidence with ever-increasing force, as we proceed to the vibrational modes of higher order.

Table 7 illustrates the situation. It contains the results of one of Milne's examples ("Example 1"). Here

$$D = -\left(\frac{d^2}{dx^2} + 2\frac{d}{dx}\right), \tag{154}$$

with the boundary conditions

$$u(0) = u(1) = 0. \tag{155}$$

Moreover, h was chosen as $\frac{1}{8}$ and $n = 7$. The column $\sqrt{\lambda_k}$ gives the correct frequencies, the column $\sqrt{\lambda_k}^*$ gives the frequencies obtained by Milne's method, while the column $\sqrt{\bar{\lambda}_k}$ gives the frequencies obtained by the iteration method.

Table 7.

k	$\sqrt{\lambda_k}$	$\sqrt{\lambda_k^*}$	$\sqrt{\bar{\lambda}_k}$
1	3.2969	3.2898	3.2667
2	6.3623	6.3457	6.1806
3	9.4777	9.4507	8.9107
4	12.6061	12.5664	11.3138
5	15.7398	15.6820	13.2891
6	18.8761	18.7870	14.7581
7	22.0139	21.8430	15.6629

Actually, it is purely a matter of computational preference whether we follow the one or the other scheme since there is a rigid relation between them. The frequencies ν_i obtained by Milne's method are in the following relation to the frequencies $\bar{\nu}_i$ obtained by the matrix-iteration method:

$$\nu_i \frac{\sin \frac{1}{2}h\nu_i}{\frac{1}{2}h\nu_i} = \bar{\nu}_i. \tag{156}$$

Hence, the results obtained by the one scheme can be translated into the results of the other scheme, and vice versa.

This raises the question why it is so beneficial to transform the frequencies $\bar{\nu}_i$ of the Δu operator to the frequencies ν_i by the condition

$$\sin \tfrac{1}{2}h\nu_i = \tfrac{1}{2}h\nu_i. \tag{157}$$

The answer is contained in the fact that the correction factor

$$\frac{\sin \frac{1}{2}h\nu_i}{\frac{1}{2}h\nu_i} \tag{158}$$

is exactly the factor that compensates for the transition from du/dx to $\Delta u/\Delta x$, if $u(x)$ is of the form

$$u(x) = C_i \sin(\nu_i x + \theta_i) \tag{159}$$

where the constants C_i and θ_i are arbitrary.

Now it so happens that for high frequencies ν_i the first term of the operator (156) strongly overshadows the other terms. The differential equation of the eigenvalue problem for large ν_i thus becomes asymptotically

$$\frac{d^2u}{dx^2} + \nu_i^2 u_i = 0, \tag{160}$$

the solution of which is given by (159). This asymptotic behavior of the solution for high frequencies makes it possible to counteract the damaging influence of the error caused by the initial transition to the difference equation. The correction is implicitly included in Milne's solution, while the results of the matrix-iteration scheme can be corrected by solving equations (157) for the ν_i.[27]

Multidimensional problems. The present investigation was devoted to differential and integral operators that belonged to a definite finite range of the variable x. This variable covered a one-dimensional manifold of points. However, in many problems of physics and engineering the domain of the independent variable is more than one-dimensional. A few general remarks may be in order as to the possibility of extending the principles and methods of the present investigation to manifolds of higher dimensions.

While the general theory of integral equations reveals that the fundamental properties of an integral equation are essentially independent of the dimensionality of the variable x, yet from the practical viewpoint the eigenvalue problem of multidimensional manifolds does lead to difficulties that are not encountered in manifolds of one single dimension. The basic difference is that an essentially multidimensional manifold of eigenvalues is projected on a one-dimensional manifold, thus causing a strong overlapping of basically different vibrational modes. A good example is provided by the vibrational modes of a rectangular membrane. The eigenvalues are here given by the equation

$$\lambda = \alpha_1^2 m_1^2 + \alpha_2^2 m_2^2,$$

where m_1 and m_2 are two independent integers, while α_1 and α_2 are two constants determined by the length and width of the membrane.

As another illustration, consider the bewildering variety of spectral terms that can be found within a very narrow band of frequencies, if the vibrational modes of an atom or a molecule are studied. To separate all these vibrational modes from one another poses a difficult problem which has no analogue in systems of one degree of freedom where the different vibrational states usually belong to well-separated frequencies.

It is practically impossible that one single trial function will be sufficient for the separation of all these vibrational states. Nor does such an expectation correspond to the actual physical situation. The tremendous variety of atomic states is not excited by one single exciting function but by a rapid succession of an infinite variety of exciting functions, distributed according to some statistical probability laws. To imitate this situation mathematically means that we have to operate with a great variety of trial functions before we can hope to untangle the very dense family of vibrational states associated with a more than one-dimensional manifold.

In this connection it seems appropriate to say a word about the physical significance of the "trial function" $\varphi_0(x)$ that we have employed for the generation of an entire system of eigenfunctions. At first sight this trial function may appear as a purely mathematical quantity that has no analogue in the physical world. The homogeneous integral equation that defines the eigenvalues, and the eigenfunctions, of a given integral operator, does not come physically into evidence since in the domain of physical reality there is always a "driving force" that

provides the right-hand member of the integral equation; it is thus the inhomogeneous and not the homogeneous equation that has direct physical significance.

If we carefully analyze the method of successive approximations by which the eigenvalues and the eigenfunctions of a given integral operator were obtained, we cannot fail to observe that we have basically operated with the *inhomogeneous* equation (102) and our trial function $\varphi_0(x)$ serves merely as the "exciting function" or "driving force." Indeed, the solution (106) for the eigenfunctions is nothing but a special case of the general solution (101), but applied to such values of the parameter λ as make the denominator zero. This means that we artificially generate the state of "resonance" which singles out one definite eigenvalue λ_i and its associated eigenfunction $\varphi_i(x)$.

From this point of view we can say that, while the separation of all the eigenfunctions of a multidimensional operator might be a practically insuperable task—except if the technique of "separation" is applicable, which reduces the multidimensional problem to a succession of one-dimensional problems—yet it might not be too difficult to obtain the solution of a given multidimensional integral equation if the right-hand member (that is, physically, the "driving force") is given as a sufficiently smooth function that does not contain a too large variety of eigenfunctions. Then the convergence of the method may still suffice for a solution that gives the output function with a practically satisfactory accuracy. This is the situation in many antenna and wave-guide problems which are actually input-output problems, rather than strict resonance problems. In other words, what we want to get is a certain mixture of weighted eigenfunctions, which appear physically together, on account of the exciting mechanism, while the isolation of each eigenfunction for itself is not demanded. Problems of this type are much more amenable to a solution than problems that demand a strict separation of the infinite variety of eigenfunctions associated with a multidimensional differential or integral operator. To show the applicability of the method to problems of this nature will be the task of a future investigation.

Summary. The present investigation establishes a systematic procedure for the evaluation of the latent roots and principal axes of a matrix, without constant reductions of the order of the matrix. A systematic algorithm (called the "progressive algorithm") is developed which obtains the linear identity between the iterated vectors in successive steps by means of recursions. The accuracy of the relation obtained increases constantly, until in the end full accuracy is obtained.

This procedure is then modified to the method of "minimized iterations," in order to avoid the accumulation of rounding errors. Great accuracy is thus obtainable even in the case of matrices that exhibit a large dispersion of the eigenvalues. Moreover, the good convergence of the method in the case of large dispersion makes it possible to operate with a small number of iterations, obtaining m successive eigenvalues and principal axes by only $m + 1$ iterations.

These results are independent of the order of the matrix and can thus be immediately applied to the realm of differential and integral operators. This results in a well-convergent

approximation method by which the solution of an integral equation of the Fredholm type is obtained by successive iterations. The same procedure obtains the eigenvalues and eigensolutions of the given integral operator, if these eigensolutions exist.

In the case of differential operators the too-stringent demands of the least-squares method may be relaxed. The approximate linear identity between the iterated functions may be established by interpolation, thus dispensing with the evaluation of definite integrals. Moreover, the iterations may be carried out with the given differential operator itself, instead of reverting to the Green's function, which is frequently not available in closed form. The entire procedure is then free of integrations and requires only the solution of linear equations.

Acknowledgments. The present investigation contains the results of years of research in the fields of network analysis, flutter problems, vibration of antennas, solution of systems of linear equations, encountered by the author in his consulting and research work for the Boeing Airplane Company, Seattle, Washington. The final conclusions were reached since the author's stay with the Institute for Numerical Analysis of the National Bureau of Standards. The author expresses his heartfelt thanks to Dr. C. K. Stedman, head of the Physical Research Unit of the Boeing Airplane Company, and to Dr. J. H. Curtiss, Acting Director of the Institute for Numerical Analysis, for their generous support of his scientific endeavors.

NOTES AND REFERENCES

1. The basic principles of the various iteration methods are exhaustively treated in the well-known book by R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary matrices* (Cambridge University Press, 1938; Macmillan, 1947).

2. A. Hotelling, *Psychometrika* **1,** 27–35 (1936).

3. A. C. Aitken, *Proc. Roy. Soc.* (*Edinburgh*) **57,** 269–304 (1937).

4. H. Wayland, *Quart. Appl. Math.* **2,** 277–306 (1945).

5. W. U. Kincaid, *Quart. Appl. Math.* **5,** 320–345 (1947).

6. The literature available to the author showed no evidence that the methods and results of the present investigation have been found before. However, Prof. A. W. Ostrowski of the University of Basle and the Institute for Numerical Analysis informed the author that his method parallels the earlier work of some Russian scientists; the references given by Ostrowski are: A. Krylov, *Izv. Akad. Nauk S.S.S.R.* **7,** 491–539 (1931); N. Luzin, *ibid.* **7,** 903–958 (1931). On the basis of the reviews of these papers in the *Zentralblatt,* the author believes that the two methods coincide only in the point of departure. The author has not, however, read these Russian papers.

7. Fredholm, *Acta Math.* **27,** 365–390 (1903).

8. See, for example, E. T. Whittaker and G. N. Watson, *Modern analysis* (Cambridge University Press, ed. 4, 1935), p. 221.

9. Throughout this paper the term "iteration" refers to the application of the given matrix A to a given vector b, by forming the product Ab.

10. Whittaker and Watson, reference 8, p. 228; Courant and Hilbert, *Methoden der mathematische Physik* (J. Springer, Berlin, 1931), vol. 1, p. 116.

11. We shall use the term "eigenvalue" for the numbers μ_i defined by Eq. (5) while the reciprocals of the eigenvalues λ_i [$= 1/\mu_i$] will be called "characteristic numbers."

12. The characteristic solutions of defective matrices—i.e., matrices whose elementary divisors are not throughout linear—do not include the entire n-dimensional space since such matrices possess fewer than n independent principal axes.

13. In order to prevent this paper from becoming too lengthy, the analytic details of the present investigation are kept to a minimum and in a few places the reader is requested to interpolate the missing steps.

14. We have in mind the general case and overlook the fact that the oversimplified nature of the example makes the decision trivial.

15. The reader is urged to carry through a similar analysis with the same matrix, but changing the 0, 0 diagonal elements of rows 5 and 6 to 1, 0 and 1, 1.

16. Instead of iterating with A and A^* n times, we can also iterate with A alone $2n$ times. Any of the columns of the iteration matrix can now be chosen as c_i numbers, since these columns correspond to a dotting of the iteration matrix with $b_0^* = 1, 0, 0, \ldots$; $b_0^* = 0, 1, 0, 0, \ldots$; $b_0^* = 0, 0, 1, 0, 0, \ldots$; and so on. The transposed matrix is here not used at all. Dr. E. C. Bower of the Douglas Aircraft Company pointed out to the author that from the machine viewpoint a uniform iteration scheme of $2n$ iterations is preferable to a divided scheme of $n + n$ iterations. The divided scheme has the advantage of less accumulation of rounding errors and more powerful checks on the successive iterations. The uniform scheme has the advantage that more than one column is at our disposal. Accidental deficiencies of the b_0^* vector can thus be eliminated, by repeating the algorithm with a different column. (For this purpose it is of advantage to start with the trial vector $b_0 = 1, 1, 1, \ldots 1$.) In the case of a *symmetric* matrix it is evident that after n iterations the basic scalars should be formed, instead of continuing with n more iterations.

17. The idea of the successive orthogonalization of a set of vectors was probably first employed by O. Szász, in connection with a determinant theorem of Hadamard; see *Math. és phys. lapok* **19,** 221–227 (1910) (in Hungarian). The method found later numerous important applications.

18. The reader is urged to carry through the process of minimized iterations and evaluation of the principal axes for the defective matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

which has only one pair of principal axes. (Choose the trial vector in the form $b_0 = b_0^* = 1, 1, 1$.)

19. For algebraic reasons, the orthogonality of the matrix (97) holds not only for the final m but for *any value* of m.

20. In a control experiment that imitated the conditions of the vibrating bar, but with a more regular matrix, the results were analytically predictable and the computational results open to an exact check. This example vividly demonstrated the astonishing degree of noninterference of orthogonal vectors. The spread of the eigenvalues was 1 : 3200. The trial vector b_0 strongly overemphasized the largest eigenvalue, containing the lowest and the highest eigenvectors with an amplitude ratio of $1 : 10^8$ (this means that if the vector of the smallest eigenvalue were drawn with the length of 1 in., the vector of the largest eigenvalue, perpendicular to the previous one, would span the distance from Los Angeles to Chicago). The slightest inclination between the two vectors would fatally injure the

chances of the smallest eigenvalue. When the entire computation scheme was finished, the analytically required eigenvalues were computed and the comparison made. *The entire string of eigenvalues, including the last, agreed with a maximum error of two units in the ninth significant figure,* thus demonstrating that the method is completely free of the cumulation of rounding errors. The author is indebted to Miss Fannie M. Gordon, former computing-staff member of the Mathematical Tables Project, New York, now of the Institute for Numerical Analysis of the National Bureau of Standards, for the eminently careful and skilled performance of the computing operations.

21. The Volterra type of integral equations which have no eigenvalues and eigensolutions are thus included in our general considerations.

22. This expansion is not of the nature of the Neumann series because the coefficients of the expansion are not rigid but constantly changing with the number m of approximating terms.

23. See Courant and Hilbert, reference 10, p. 76.

24. Lord Rayleigh, *The Theory of Sound* (Dover Publications, New York, 1945), vol. I, pp. 272-278 (reprint edition).

25. We call the eigenvalues here λ_i since the operator D is not inverted into an integral operator K.

26. See the forthcoming publication in the *Journal of Research of the National Bureau of Standards*; the variable s of Milne is changed to x and his λ^2 to λ, to avoid conflicts with the notations of the present paper.

27. This experience is valuable since in many eigenvalue problems similar conditions hold; the eigenfunctions of large order can often be asymptotically estimated, in which case the error of the Δ-process may be effectively corrected. For example, the values μ_i found in a previous section for the lateral vibrations of an inhomogeneous bar may be corrected as follows:

μ_i uncorrected: 2256.944, 48.2038, 5.3563, 1.5830, 0.59, (0.25);
μ_i corrected: 2258.924, 48.4977, 5.4577, 1.6407, 0.62, (0.28).

On the Monte Carlo Method

S. ULAM

Los Alamos Scientific Laboratory

The advent of fast computing machines is greeted as a way to "settle" a great many questions existing in applied mathematics and physics. It is hoped that the formulation of many problems in physical sciences is now correct, but the obtaining of solutions through methods that are perfectly well understood in principle is much too laborious and tedious to be attempted by paper and pencil. The mere number of arithmetical steps necessary to perform the procedures for numerical solution is such as to require, in some cases, years, even hundreds of years, of computer's time. The machines, especially the electronic calculators, can shorten this time very greatly, since the elementary steps, e.g., the multiplications, may take less than 1/1000 of the time required by a person, even if provided with a desk multiplier.

It seems that in addition to this program of testing and verification of existing mathematical formulations of physical problems, some other possibilities will be open. One—a rather general program—will be to use the calculations of electronic calculators for heuristic purposes. This may be equally possible in exploring new physical models and in pure mathematics itself. The latter possibility, perhaps less evident a priori, should be clear if one remembers the role of *examples* in abstract mathematics. It is sufficient to point out their importance in such parts of mathematics as geometry, topology, theory of functions, and abstract algebra. The field of combinatorial analysis, which is so hard to define just because it consists of a variety of special problems or examples not yet embraced by a simple general theory, seems the clearest case. The combinatorial problems studied (many of them problems of *enumeration*), drawn, as it were, from outside of mathematics, often have their origin in problems about configurations of physically existing objects and relations. In a field like this it is clear that the ability of machines to survey *all* the possibilities of specified arrangements will provide the material suggesting future theories.

The theory of probabilities, which from one point of view is a branch of combinatorial analysis, is a case in point. The so-called Monte Carlo method may be said to consist of a "physical" production of models of combinatorial situations.

A simple example would be this: The problem consists of estimating the proportion of the 52! permutations of objects (cards) possessing a given—in practice always a complicated—property. One should then consider all these permutations, counting the number of those among them that possess this property. This would of course be impossible, even granting a continuous development of the speed in computing for the next hundred years. A way to get the proportion with good probability is to produce a "large" number of permutations, say 10,000, *at random* and count the *proportion* of the permutations possessing the given property.

It is perhaps surprising how many mathematical problems have in practice a structure logically similar to the one of this example. The evaluation of a definite integral may be thought of as a task somewhat analogous to the one of the previous example. The problem consists in finding the value, on a region S of the unit cube defined by inequalities, of the number

$$\iint_{(S)} \cdots \int f(x_1 \cdots x_n) dx_1 \cdots dx_n,$$

f being a given function. This will be reduced to finding, say,

$$\iint_{(R)} \cdots \int dx_1 \cdots dx_{n+1}$$

where the region R is defined by a class of inequalities

$$\varphi_1(x_1 \cdots x_{n+1}) < 0,\ \varphi_2(\cdots) < 0,\ \varphi_k(\cdots) < 0.$$

The procedure of elementary calculus will consist in counting the lattice points of a subdivision in a space of $n + 1$ dimensions, and ascertaining the proportion of these points that satisfy the given inequalities. The Monte Carlo procedure would be to take a large number of lattice points *at random* and examine these only. This number need not be of the order of the total of all lattice points.

The problem of estimating very "small" volumes requires special tricks. To illustrate the nature of possible devices, let us again take a more purely combinatorial problem. In a solitaire (game of cards for one), one desires to estimate the probability of a successful outcome. (We assume that skill plays no role, so that it is purely a game of chance.) In cases where the game has a very small probability of success most actual plays will end in failure and only an upper limit will be obtainable for this probability. How can one get some idea of a lower limit > 0? Suppose that one obtains, still obeying the rules of the game, in a noticeable proportion α of tries, a situation A where only, say, ten cards are left uncovered; after that, however, we meet with "failure." It might be justifiable to restore the ten cards to their positions *in a different* permutation and try from the situation A again. By examining a large number of the 10! permutations we might obtain the number β expressing the chance that starting with A we "win" B. A reasonable guess for the chance of success from the beginning without "cheating" would then be greater than $\alpha\beta$. Of course A should be really a class of positions, not a possible or a very special one. It seems, however, that if the playing of the whole game is decomposed into two or more stages, there will be a saving in the number of experiments compared with the number necessary to play to the end each time and beginning anew after each failure from the start, that is, a new permutation of the 52 cards.

The validity of such a procedure can be established in some cases. One has to prove independence, or estimate from above the correlation between the classes of events A and success B.

It is of course obvious that one can study "experimentally" the behavior of solutions of equations which themselves describe a random process, by using the digital computer as an analogy machine, as it were.[1] This experimental—that is, statistical—approach by Monte

Carlo techniques has been applied by various authors to linear partial differential equations.[2] In the case of equations that are quadratic or of higher order in the unknown functions and their derivatives, the obvious Monte Carlo procedure would be much more cumbersome, but may still have heuristic value. As an example, let us take a bilinear system of two partial differential equations

$$\left.\begin{aligned}\frac{\partial u_1}{\partial t} &= \alpha_1 \Delta u_1 + \beta_1(u_2)u_1 \\ \frac{\partial u_2}{\partial t} &= \alpha_2 \Delta u_2 + \beta_2(u_1)u_2\end{aligned}\right\},$$

where u_1 and u_2 are unknown functions of coördinates x, y, z, and t; α_1 and α_2 are given constants; β_1 and β_2 are given functions, for simplicity linear in u_1 and u_2 and also involving the independent variables x, y, z. One would like to know the asymptotic form of u_1 and u_2 (for large values of t). This problem may be looked upon as a straightforward generalization of the diffusion model (Fermi) of the Schrödinger equation.[1] It would correspond to a model of a system of two particles with the *potential* function for u_i replaced by the corresponding u function of the other particle. This linked system, treated then somewhat in the spirit of a field theory, is nonlinear. There will not be in general eigenfunctions—the separation into a time-independent equation will not be possible; yet for large values of the parameter t the space part of u may have an almost periodic or *summable* (by the first mean) behavior. A numerical approach to the study of such systems could again be a Monte Carlo procedure. One would diffuse and multiply the (fictitious) particles corresponding to u_1 and u_2 according to their numbers, instead of a given function V of coördinates. Since these numbers change in t, it will be necessary to make frequent censuses—as it were, to interrupt the calculation periodically—in order to ascertain the values to be used for "potentials."

The problem of transforming first purely formally, an equation not of a diffusion or Boltzmann type into one of the above type thus becomes of practical importance. Let us indicate some possibilities in this direction. The equation of Hamilton-Jacobi in one dimension has the form

$$\left(\frac{\partial S}{\partial x}\right)^2 = \frac{1}{v^2(x)}. \tag{1}$$

On the other hand, consider the equation

$$\frac{\partial W}{\partial t} = \frac{\partial}{\partial x}\left[v\frac{\partial (vW)}{\partial x}\right]. \tag{2}$$

This latter equation will describe the probability behavior of a particle starting, say, from the origin, and performing a random walk on the line, steps being equally probable to the right or to the left. However, the length of the steps in the position x is proportional to the value of $v(x)$. If we perform the passage to the limit with the length of the step tending to zero the resulting continuous process gives a distribution of position in time t obeying Eq. (2). It can be proved[3] that the crest of the distribution, that is, the place x where $\partial W/\partial t = 0$, will satisfy a relation $S(x) = t^{\frac{1}{2}}$, where S is the solution of Eq. (1).

It is of course quite unnecessary to take recourse to such methods for a one-dimensional equation that is easily solved explicitly by quadratures. The example here given is meant merely to indicate the possibility of relations between two seemingly very different processes. One is a strictly deterministic one, described by the equation of geometric optics (or the equations of mechanics), the characteristic equation of Hamilton. The other is a continuous random-walk process with the *length* of the elementary step a given function of position. It turns out that at least in one dimension the locus of the points where the first derivative with respect to time of the probability distribution is equal to zero coincides with the locus of the points where the value of the Hamilton function $S = \sqrt{t}$. In two or more dimensions, the two loci are probably at least asymptotically equal, that is, for large values of t.

In the first examples of application of the Monte Carlo method to empirical evaluation of properties of solutions of differential equations, one studied the *density* of the diffusing and branching, that is, multiplying and transmuting, particles. This density as a function of the independent variables obeyed a linear partial differential equation of a parabolic or elliptic type. It is clear that for nonlinear equations one will have to examine, not this density directly, but appropriately chosen *functionals* of this function.

The diffusion process can be described, of course, as a Markoff chain, and this in turn by a study of the interaction of matrices with nonnegative coefficients. Let us indicate a way to study "experimentally" the behavior of powers of matrices with arbitrary real terms. This possibility rests on the fact that real numbers can be considered as matrices, with positive terms; for example, -2 corresponds to $\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. This correspondence obviously preserves both addition and multiplication. Any system described by any n-by-n matrix giving the transition *moments* as real numbers can be interpreted probabilistically by using $2n$-by-$2n$ matrices with nonnegative terms. The diffusion and branching or multiplication are performed by two kinds of particles—black and red—with the transformation rates given by the matrices above.

In having four kinds of particles one can then realize stochastic models for matrices with complex terms; more generally, with an appropriate number of *kinds* of particles, one can realize stochastic models for more general algebras over real numbers.[4]

The possibility of a statistical or probabilistic evaluation of definite integrals in n-dimensional space affords merely one example of an attempt to gain insight into a situation involving a system of n particles. Let us think here of n as having a value of the order of 10 or 20. The "appearance" of a set of points in a euclidean space of this dimension, if the set is defined as above by many inequalities, cannot of course be studied on graphs directly, or very well by projections of the set into three-dimensional component spaces. Now, in physical chemistry, for instance, the occurrence of this situation and its importance are well known. The properties of a molecule with a large number of atoms depends on characteristics of configurations of certain n-dimensional sets. The evaluation of various functionals of these configurations can, probably, be done best by a Monte Carlo procedure, that is, by testing a large number of

n-tuples, chosen at random with appropriate distribution, for the values of these given functionals.

It is rather curious that one meets with an analogous situation in pure mathematics itself. Let us describe it very briefly: a *formal* system in mathematics involves in addition to the Boolean operations of elementary logic or set theory (the addition and intersection of sets of points), the so-called quantifiers, the two symbols $\sum_x \varphi(x)$, meaning that there *exists* an x for which $\varphi(x)$, a propositional function, is true, and $\prod_x \psi(x)$, meaning that $\psi(x)$ holds for *all* x.

A large part of the study of mathematics as a formal system involves the study of classes of sets on which one performs these operations. One knows that a "geometric" interpretation of these operators is particularly simple.[5] The existence quantifier corresponds to taking an orthogonal projection parallel to one or more axes of a given set of points in n dimensions on a space of fewer dimensions. The other quantifiers can be expressed by means of the first and the Boolean operations.

Even the simplest mathematical definitions lead to sets defined in a higher number of dimensions. The problem presents this appearance: there are given in a space of n dimensions several "primitive" sets of points. Starting with these sets, one obtains new ones by adding them, intersecting them with one another; these are the Boolean operations. One also takes projections of the sets obtained and conversely, having sets in spaces of fewer dimensions, erects cylinder sets in the full n-dimensional space.

One can in this fashion, starting from two given sets, obtain an infinity of new sets.

The mutual relations of these sets form the object of the logical or metamathematical study of the system.

It is possible that, for heuristic purposes alone, it would be useful to study these constructions on a large number of examples.

A mathematical theorem can be formulated in this language as stating that a certain set of the class obtained is vacuous. In cases where a proof would appear very difficult it might be of value to, so to say, try to construct points of it by random choices of the starting sets or values of "free variables" in the n-dimensional space. The failure to obtain any after a great number of choices would then lead to the belief that if the set is not vacuous it is small. It is clear that a proof will never be obtained in this fashion. However, the heuristic value of such a procedure might not be negligible.

REFERENCES

1. N. Metropolis and S. Ulam, "The Monte Carlo method," *J. Am. Statist. Assoc.* **44,** 335–341 (1949).

2. M. D. Donsker and M. Kac, "A sampling method for determining the lowest eigenvalue and the principle eigenfunction of Schrödinger's equation," *J. Research Nat. Bur. Standards* **44,** 551–557 (1950).

3. C. J. Everett and S. Ulam, "Random walk and the Hamilton-Jacobi equation," *Bull. Am. Math. Soc.* (1950), abstract (to be published).

4. C. J. Everett and S. Ulam, "On an application of a correspondence between matrices over real algebras and matrices of positive real numbers," *Bull. Am. Math. Soc.* (1950), abstract (to be published).

5. C. J. Everett and S. Ulam, "Projective algebra. I," *Am. J. Math.* **68,** 77–88 (1946).

FIFTH SESSION

Thursday, September 15, 1949
9:00 A.M. to 12:00 P.M.

COMPUTATIONAL PROBLEMS IN PHYSICS

Presiding

Karl K. Darrow
Bell Telephone Laboratories

The Place of Automatic Computing Machinery in Theoretical Physics

WENDELL H. FURRY

Harvard University

We could scarcely expect to discuss the place of automatic computation in theoretical physics without taking some account of one extreme attitude, which, we may hope, is of decreasing prevalence, and which, at any rate, is surely not held by most of the members of this symposium. This is the opinion that, bluntly, it has no place at all.

A strong human tendency to resist and attack the introduction of new methods has shown itself repeatedly in mathematics and natural philosophy. As two examples, we might consider the cases of infinite series and of complex numbers. Hogben has pointed out that the keen sense of paradox that the Greeks got from Zeno's fable of Achilles and the tortoise was due simply to their lack of the concept of an infinite series with a finite sum. It was many centuries before this concept was finally cleared up to the point that absurd results were no longer derived occasionally by even the ablest mathematicians. The related idea of a limit, as used in the calculus, was the object of decades of bitter controversy.

The use of complex numbers aroused if anything even more violent opposition. At the beginning of the nineteenth century a Cambridge mathematician devoted thirty pages of the *Philosophical Transactions* of the Royal Society to a carefully reasoned plea for the acceptance of the use of complex arithmetic in a few simple arguments. He encountered the searing condemnation of a learned but anonymous writer in the *Edinburgh Review*, who launched upon his concluding page as follows:

> We shall be spared the task of examining further into a mode of explanation which, at the outset, is liable to so great objections. Operations deduced from such principles are undeserving the name of reasoning; and they cannot afford one particle of evidence either of truth or of falsehood.

We know how this case turned out, also, with the rigorous justification of complex analysis and its magnificently powerful development during the next few decades.

Actually, of course, it would be beside the point to argue the case on this basis. Surely it must be generally accepted that automatic computation can provide rigorously certain results, of a required degree of approximation. The whole program of digital computing machines, indeed, falls directly in the tradition of the arithmetization of mathematics, which played a great part in the development of modern standards of rigor. Many problems of the uniqueness of solutions and the limits of error in automatic computation call for further mathematical study, but it is certain that suitable standards can be established and met, in all but very exceptional cases.

A physicist denying automatic computation a place in theoretical physics would soon, if not immediately, come to the assertion, "Even if it is correct, it isn't physics." This is certainly not a logical statement, but a judgment of values. We cannot expect to deal conclusively—or perhaps the word should be exhaustively—with this sort of question, but a few remarks can be made.

To begin with, it is easy to think of extreme cases of both kinds—cases in which no one would expect automatic computation to play a part, and other cases in which only a definitely quaint degree of prejudice could refuse to welcome its help. An example of the first kind of case is the problem of divergences in quantum electrodynamics. The difficulty here is that the existing formulation of the theory is inadequate, and indeed mathematically meaningless. The solution must come from the discovery of suitable new physical ideas, and their proper incorporation into the theory. This has fundamentally nothing to do with computation, and only after the theoretical advance has been made can it become possible to see the part automatic computation might play in further developments.

The other extreme case is that in which a physicist succeeds in obtaining a formula, or perhaps an equation, containing, or contributing to, the solution of a problem. The natural next step is to use values of tabulated functions to evaluate the formula or solve the equation. If they are very simple functions, say trigonometric functions, comparatively ancient tables are available. If, on the other hand, something like the confluent hypergeometric function is involved, the physicist is not likely to be able to take the next step unless a computing project, nowadays probably a project using an automatic machine, has provided the necessary tables. Simple and obvious as this part of the place of automatic computation in physics is, it is of great importance. There has hardly been a real beginning on the production of the great variety of new tables that can be of inestimable value in scientific work.

Coming back again to the problem of our Tory who feels that automatic computation just isn't physics, we may note that one main ingredient of this feeling is the idea that physical explanations should be fundamentally simple. The demand for simple explanations is a basic part of a physicist's attitude. The final achieving of a simple synthesis out of previous complications is a source of the greatest satisfaction to the workers who experience it. It establishes one of the landmarks that adorn the history of the subject, and it provides an addition to the solid framework supporting the structure of the subject itself. Without a reliable framework of simple basic principles, the complicated structure could not grow far without collapsing.

Although his reputation among laymen may not be exactly that of a man preferring simple ideas, the physicist does succeed in satisfying his desire for simple explanations in many cases. Sometimes considerable time elapses before such an explanation is forthcoming. For a quarter century after the discovery of thermal diffusion in gases, any student who asked a professor for a simple explanation of this phenomenon was told that such an explanation is impossible, and that only the detailed mathematical theory could account for the effect. That was that, and the student had to subside. But about ten years ago a professor asked a student to include

in a coming colloquium talk an explanation of thermal diffusion. The demand was an insistent one, and the student, whose name was Sidney Frankel, was on the spot. Accordingly, for a bit less than ten years we have had a good simple explanation of thermal diffusion in gases.

It is clear that a simple explanation, or a simple test of a new theory, cannot be based on the most general sort of case. A happy choice of a special case can be of the greatest importance; a famous example is the spectrum of the hydrogen atom. The choice of suitable cases for theoretical explanation is a matter requiring the highest type of judgment, and perhaps inspiration. Often it may be the skill and insight of the experimental physicist that show the right way.

But one simple test, or a few simple tests, are often not enough for a new theory. A theory as radically new as quantum mechanics could not be accepted on the evidence of the Balmer spectrum alone, even if it had no rivals in explaining that case. Corroboration was found from many sources. Some of these arguments were very impressive although almost purely qualitative, but lengthy computations made by Hylleraas and by Coolidge and James, among others, were very important. I can testify that the fact that the prolonged labor required could not then be saved by automatic computation is a source of sincere regret to at least one veteran of that period.

Extensive computations are often needed, not only in finding theoretical results to check with experiment, but also in finding out what interpretation to give to the experimental results themselves. Elaborate calculations on the functioning of an instrument may be required to obtain data that throw light on important questions of fundamental theory. Professor Vallarta's paper discusses the interpretation of observations from a huge instrument, the earth itself acting as a magnetic spectrometer.

The subject of nuclear physics, on which Professor Feshbach reports, is one in which the wish of the physicist for simple explanation has suffered repeated rebuffs. Nuclear structure differs from atomic structure in such ways that much less can be expected from semiqualitative arguments. Some of the high-energy scattering data indicate that the basic phenomenon of the so-called saturation of nuclear forces will have to be explained in a rather complicated way. In this subject the number of different hypotheses that may need to be tested, as well as the number of separate problems, is so great that only automatic computation seems capable of progressing fast enough.

Besides the problems of testing theories, interpreting observations, and deciding between various hypotheses, there are questions how far existing theories are capable of accounting for certain kinds of phenomena. For example, it seems to be generally agreed that nonrelativistic quantum mechanics accounts for atomic and molecular structure and for many facts in the structure of solids, but the question may be raised how far it suffices to cover *all* of this last field. Are the striking phenomena of superconductivity to be accounted for as statistical effects of the ordinary electrical and quantum laws, or do they require the introduction of some perhaps radically new though pleasingly simple assumption? Probably authorities in this field have definite opinions on such questions, but they can scarcely be really certain

about them. The difficulty of such statistical calculations is so great that automatic calculation, for finite but rather extensive lattice structures, may eventually be called into use.

In summary it can be said that there are questions in theoretical physics with which automatic computation has nothing to do, and on the other hand there is a potentially very great service in the provision of function tables, which would be universally welcomed. There are also many questions in the testing of theories, in the interpretation of observations, in the choice between hypotheses, and in establishing the range of adequacy of theories, for which automatic computation could be extremely useful. The settling of such questions would generally be helpful, and in some cases probably indispensable, in the advancement of theoretical physics.

Double Refraction of Flow and the Dimensions of Large Asymmetric Molecules

HAROLD A. SCHERAGA

Cornell University

JOHN T. EDSALL

Harvard Medical School

and

J. ORTEN GADD, JR.

Computation Laboratory of Harvard University

Optical measurements of the double refraction produced when a solution of large asymmetric molecules is subject to a shearing force can be used to determine the dimensions of the dissolved molecules.[1–4] The method has already been extensively applied,[5–10] but many of the data obtained or obtainable could not heretofore be interpreted because the theory had been developed to give numerical values only under certain limiting conditions. The present work was undertaken in order to extend the applicability of the theory to a much wider range of experimental conditions, thus greatly increasing the usefulness of flow-birefringence measurements as a tool in the determination of particle sizes and the characterization of polydisperse systems of macromolecules.

Double refraction is produced, in a liquid containing large asymmetric molecules or colloidal particles, when a velocity gradient is set up in the liquid. This is most readily achieved by forcing the liquid through a capillary tube, or by subjecting it to shear between two concentric cylinders, one of which rotates while the other is held fixed. The latter procedure is best for quantitative measurements, and was employed in 1870 by J. Clerk Maxwell, who was apparently the first to describe the phenomenon, using Canada balsam as the liquid for study. This is also the method that has been adopted in most studies on double refraction of flow.[1–16]

In the concentric-cylinder type of system, the liquid is placed in the annular space between the cylinders, the suspended particles assuming random orientation when both cylinders are at rest, as shown in Fig. 1 (*a*). When one of the cylinders, say the outer one, is set in rotation, laminar flow is produced in the liquid and a velocity gradient is set up across the gap.[17] The resulting shearing forces produce an orientation of the suspended particles, which are here assumed to be rigid ellipsoids of revolution.[18] This orientation is represented schematically in Fig. 1 (*b*). If the cylinders are mounted between crossed Nicol prisms, where *AA* and *PP* represent the planes of transmission of the analyzing and polarizing Nicol, respectively, then

the field appears dark when the cylinders are at rest and, when one cylinder is rotating, becomes light in all regions except for a dark cross [Fig. 1(*b*)], the "cross of isocline."

To characterize the observed phenomena, there are two quantities that must be measured: (1) the extinction angle χ, the smaller of the two angles between the cross of isocline and the planes of transmission of the Nicols (this is also the angle between the optic axis in the flowing liquid and the direction of the streamlines); (2) the magnitude of the double refraction Δn, that is, the difference in refractive index between light transmitted with its electric vector parallel, and light with its electric vector perpendicular, to the optic axis. The problem is to measure χ and Δn as functions of the velocity gradient and relate them to the dimensions of the suspended particles.

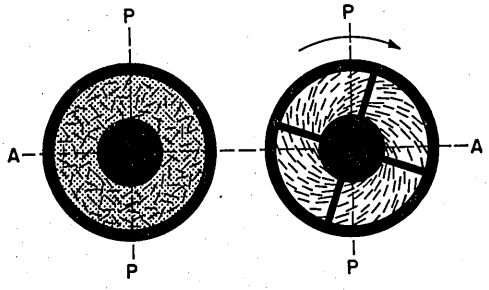


FIG. 1. Orientation of particles, each schematically represented by a line indicating its optic axis, in a doubly refracting liquid between concentric cylinders, when the outer cylinder is (*a*) at rest, (*b*) in motion. The lines *AA* and *PP* are the axes of crossed Nicol prisms. [Adapted from von Muralt and Edsall, *J. Biol. Chem.* **89**, 315 (1930).]

Empirically it is found, for solutions containing a single type of large molecule, that χ approaches 45° as the velocity gradient G approaches zero, and approaches 0° asymptotically as G increases to very large values, provided the flow is laminar; Δn is zero when $G = 0$, is a linear function of G at low values of G, and gradually approaches a constant limiting saturation value at very high values of G. Very elongated particles, like those of myosin (the structural protein of muscle) or tobacco-mosaic virus, which are several thousand Angstrom units in length, give low values of χ and high values of Δn even at low velocity gradients ($G = 10$ to $300\ \text{sec}^{-1}$ in water which has a viscosity of 0.01 poise at 20° C). Molecules, like those of human serum albumin and gamma globulin, which are near 200 A in length, require much higher velocity gradients and solvent media of high viscosity as well to attain a significant degree of orientation ($G = 1000$ to $10{,}000\ \text{sec}^{-1}$, or more; viscosity 50 to 100 times that of water).

Colloidal particles or large asymmetric molecules in the flowing solution are subject to shearing forces due to the velocity gradient G, which tends to orient their major axes.[19] In addition to the hydrodynamic forces, the particles are subject to rotary Brownian movement which causes a random fluctuation of the orientation. The Brownian movement is characterized by a rotary diffusion constant Θ. The relation of χ and Δn to the molecular dimensions has been developed chiefly by Boeder,[20] Peterlin and Stuart,[3, 21, 22] and Snellman and Bjornstahl.[4] This is expressed as a function of the parameter α (or σ in the notation of Peterlin and Stuart), which is equal to G/Θ. If Θ is known, the length of the semimajor axis a of the molecule can be evaluated.[1–4, 23, 24] The crux of the problem, therefore, is the determination of Θ from the experimental measurements of χ and G.

The double refraction Δn is the product of an optical factor which is evaluated independently[22] and an orientation factor f. Like the extinction angle χ, f is a function of α and the axial ratio p [$= a/b$].

We shall further define the quantity[25] R,

$$R = \frac{p^2 - 1}{p^2 + 1}; \tag{1}$$

R is thus equal to unity for an infinitely thin rod ($a/b = \infty$); to zero for a sphere ($a/b = 1$); and to -1 for a flat disk without thickness ($a/b = 0$).

Peterlin and Stuart obtained expressions for χ and f in terms of slowly converging infinite series in α and p. At very low values of α (< 1.5), corresponding to χ values between 45° and 38°, these series converge sufficiently rapidly to enable one to evaluate the rotary diffusion constant from a simple limiting equation. However, the errors in the experimental data are generally greatest at low velocity gradients, that is, at low values of α. The data are more accurate at somewhat higher gradients, but it has not been possible hitherto to evaluate from theory the numerical relation between χ and α under these conditions. Moreover, it is very important experimentally to determine whether a given solution under study contains only one or more than one constituent capable of orientation by the velocity gradients employed. The only way to be sure of this is to make measurements over a wide range of velocity gradients and compare the measured χ values with values calculated from the appropriate theoretical curve.[26] However, since only a small portion of the theoretical curve is given by the limiting formulas of Peterlin and Stuart, this method of analysis could not be satisfactorily carried out. A semiempirical method has been tried[9] but it was considered essential to have the complete theoretical curves using a rigid ellipsoid of revolution as a molecular model.[18] If these were available it would be possible not only to infer whether only a single type of elongated molecule is present but also, if several such components are present, to draw some important inferences concerning their relative sizes and concentrations in the solution. An observed χ value, in such a multicomponent system, is a function of all the values of both χ and f that would be found for each of the components, if it were present in the solution alone, at the same velocity gradient. The ability to analyze such complex systems would greatly increase the range and power of the method of double refraction of flow.

We shall, therefore, present the Peterlin and Stuart theory[3, 21, 22] wherein we have evaluated the quantities required to obtain χ and f values over a wide range of α values by the use of the Mark I computer of the Harvard Computation Laboratory.

If rigid ellipsoidal particles are suspended in a continuous medium under conditions of laminar flow, a steady-state distribution will be established very rapidly. This distribution will depend on α and R and is characterized by a distribution function F which, in the steady state, is given by the differential equation[27]

$$\frac{\Delta F}{\alpha} = \frac{1 + R\cos 2\varphi}{2}\frac{\partial F}{\partial \varphi} + R\frac{\sin\theta\cos\theta\sin 2\varphi}{2}\frac{\partial F}{\partial \theta} - \frac{3R\sin^2\theta\sin 2\varphi}{2}F. \tag{2}$$

The meaning of θ and φ may be understood by reference to Fig. 2, which is a section of the gap between the concentric cylinders, the inner one rotating. Here X is the direction of the streamlines at O; Z is the direction of the velocity gradient between the concentric cylinders; Y is parallel to the cylinder axis and is normal to the streaming plane; x, y, z are the directions of the principal axes of the index-of-refraction ellipsoid of the birefringent system; and χ is the angle between the z- and X-axes, x and z being coplanar with X and Z. An *individual* particle at O has its major axis in the ξ direction, where ξ, η, ζ are a set of axes fixed in the particle. Then Θ is the angle between Y and ξ, while φ is the angle between the YZ- and $Y\xi$-planes. This is the usual notation of spherical coördinates with volume element $d\Omega = \sin\theta d\theta d\varphi$.

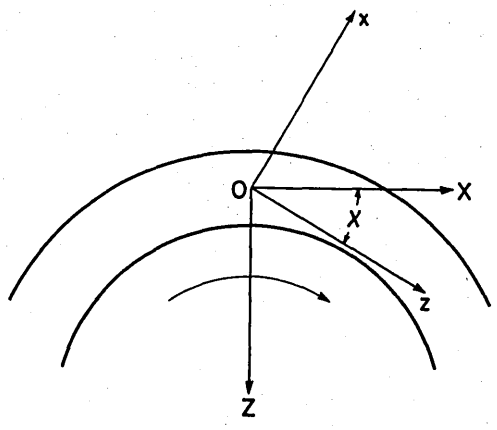


Fig. 2. Part of the coördinate system in the Couette cylinder apparatus. The X, Y, Z axes are fixed in the fluid and the x, y, z axes are the principal axes of the index-of-refraction ellipsoid in the birefringent system. [From Peterlin and Stuart, *Z. Physik* **112**, 1 (1939).]

As will be shown later, the determination of χ and f involves the evaluation of certain mean values. The distribution function F is required for this purpose and is evaluated as follows.

Express F as a power series in R,

$$F = \sum_{j=0}^{\infty} R^j F_j. \tag{3}$$

Each F_j then satisfies an inhomogeneous equation of the type

$$\frac{1}{\alpha}\Delta F_j - \tfrac{1}{2}\frac{\partial F_j}{\partial \varphi} = \tfrac{1}{2}\left[\cos 2\varphi \frac{\partial F_{j-1}}{\partial \varphi} + \sin\theta\cos\theta\sin 2\varphi \frac{\partial F_{j-1}}{\partial \theta} - 3\sin^2\theta\sin 2\varphi \cdot F_{j-1}\right]. \tag{4}$$

Now F_j may be expressed in terms of series of spherical harmonics as

$$F_j = \tfrac{1}{2}\sum_{n=0}^{\infty} a_{n0,j} P_{2n} + \sum_{n=1}^{\infty}\sum_{m=1}^{n}(a_{nm,j}\cos 2m\varphi + b_{nm,j}\sin 2m\varphi) P_{2n}^{2m}, \tag{5}$$

where P_{2n} is a Legendre polynomial of the first kind, and

$$P_{2n}{}^{2m} = \sin^{2m}\theta \cdot \frac{d^{2m}P_{2n}}{(d\cos\theta)^{2m}}$$

Since F_j is a function of α and is independent of R, the $a_{nm,j}$ and $b_{nm,j}$ coefficients will also have this dependency.

Substituting from Eq. (5) in Eq. (4) and making use of the orthogonality and recurrence relations for these polynomials,[28] one obtains the following recurrence formulas for $a_{nm,j}$ and $b_{nm,j}$:

$$\frac{n(2n+1)}{\alpha}\, a_{n0,j} = -\tfrac{1}{4}\left[-\frac{(2n-3)(2n-2)(2n-1)2n(2n+1)}{(4n-3)(4n-1)}\, b_{n-1,1;j-1}\right.$$
$$+\frac{3(2n-1)2n(2n+1)(2n+2)}{(4n-1)(4n+3)}\, b_{n,1;j-1}$$
$$\left.+\frac{2n(2n+1)(2n+2)(2n+3)(2n+4)}{(4n+3)(4n+5)}\, b_{n+1,1;j-1}\right], \tag{6}$$

$$b_{n0,j} = 0, \tag{7}$$

$$\frac{2n(2n+1)}{\alpha}\, a_{nm,j} + mb_{nm,j} = -\tfrac{1}{4}\left[\frac{2n+1}{(4n-3)(4n-1)}\, b_{n-1,m-1;j-1}\right.$$
$$-\frac{3}{(4n-1)(4n+3)}\, b_{n,m-1;j-1} - \frac{2n}{(4n+3)(4n+5)}\, b_{n+1,m-1;j-1}$$
$$-\frac{(2n-2m-3)(2n-2m-2)(2n-2m-1)(2n-2m)(2n+1)}{(4n-3)(4n-1)}\, b_{n-1,m+1;j-1}$$
$$+\frac{3(2n-2m-1)(2n-2m)(2n+2m+1)(2n+2m+2)}{(4n-1)(4n+3)}\, b_{n,m+1;j-1}$$
$$\left.+\frac{2n(2n+2m+1)(2n+2m+2)(2n+2m+3)(2n+2m+4)}{(4n+3)(4n+5)}\, b_{n+1,m+1;j-1}\right]$$
$$(m \neq 0), \tag{8}$$

$$-ma_{nm,j} + \frac{2n(2n+1)}{\alpha}\, b_{nm,j} = \tfrac{1}{4}\left[\frac{2n+1}{(4n-3)(4n-1)}\, a_{n-1,m-1;j-1}\right.$$
$$-\frac{3}{(4n-1)(4n+3)}\, a_{n,m-1;j-1} - \frac{2n}{(4n+3)(4n+5)}\, a_{n+1,m-1;j-1}$$
$$-\frac{(2n-2m-3)(2n-2m-2)(2n-2m-1)(2n-2m)(2n+1)}{(4n-3)(4n-1)}\, a_{n-1,m+1;j-1}$$
$$+\frac{3(2n-2m-1)(2n-2m)(2n+2m+1)(2n+2m+2)}{(4n-1)(4n+3)}\, a_{n,m+1;j-1}$$
$$\left.+\frac{2n(2n+2m+1)(2n+2m+2)(2n+2m+3)(2n+2m+4)}{(4n+3)(4n+5)}\, a_{n+1,m+1;j-1}\right]$$
$$(m \neq 0). \tag{9}$$

Normalization, that is, putting $\int F d\Omega = 1$, gives $a_{00,0} = 1/2\pi$ and all other $a_{00,j} = 0$. The complete distribution function is given by Eq. (3) after the F_j's are thus evaluated.

The evaluation of the coefficients $a_{nm,j}$ and $b_{nm,j}$ as solutions of the simultaneous Eqs. (6–9) is a formidable task and would have been hopeless without the aid of the Mark I computer. It should be pointed out that not all the coefficients are required for the problem of double refraction of flow, but, as has been shown by Peterlin and Stuart, only the $a_{11,j}$ and $b_{11,j}$ terms. However, many other terms are required in order to evaluate these particular ones. The task is somewhat eased by the vanishing of many of these terms for certain values of the indices.

Making use of the distribution function F of the particles it is possible to calculate the effect of the interaction of this oriented system with a beam of polarized light, that is, the double refraction. Results of such a computation are

$$\tan 2\chi = \frac{\overline{2 \cos(\xi X) \cos(\xi Z)}}{\overline{\cos^2(\xi X)} - \overline{\cos^2(\xi Z)}} \tag{10}$$

and

$$\Delta n = \frac{2\pi c}{n}(g_1 - g_2)\sqrt{[\overline{\cos^2(\xi Z)} - \overline{\cos^2(\xi X)}]^2 + [\overline{4 \cos(\xi Z) \cos(\xi X)}]^2}$$

$$= c\frac{2\pi(g_1 - g_2)}{n} f(\alpha, p) \tag{11}$$

where c is the concentration of the particles, n is the index of refraction of the isotropic solution at rest, and $(g_1 - g_2)$ is an optical factor depending on the axial ratio and indices of refraction of the particles.[22] Since $\cos(\xi X) = \sin\theta \sin\varphi$, and $\cos(\xi Z) = \sin\theta \cos\varphi$, it follows that

$$2 \cos(\xi X) \cos(\xi Z) = \tfrac{1}{3} \sin 2\varphi \cdot P_2^2$$

and $\tag{12}$

$$\cos^2(\xi Z) - \cos^2(\xi X) = \tfrac{1}{3} \cos 2\varphi \cdot P_2^2.$$

The mean values of these functions are evaluated by multiplying by F and integrating, giving[29]

$$-\tan 2\chi = \frac{\sum_{j=1}^{\infty} R^{j-1} b_{11,j}}{\sum_{j=1}^{\infty} R^{j-1} a_{11,j}} \tag{13}$$

and

$$f(\alpha,R) = \frac{16\pi R}{5}\sqrt{(\sum_{j=1}^{\infty} R^{j-1} a_{11,j})^2 + (\sum_{j=1}^{\infty} R^{j-1} b_{11,j})^2}. \tag{14}$$

Heretofore, these equations have been useful for valid computation only in the following limiting forms, which hold for $\alpha < 1.5$,

$$\chi = \frac{\pi}{4} - \frac{\alpha}{12}\left[1 - \frac{\alpha^2}{108}\left(1 + \frac{24R^2}{35}\right) + \cdots\right], \tag{15}$$

$$f(\alpha,R) = \frac{\alpha R}{15}\left[1 - \frac{\alpha^2}{72}\left(1 + \frac{6R^2}{35}\right) + \cdots\right]. \tag{16}$$

By machine computation of the $a_{11,j}$ and $b_{11,j}$ terms, the sums that appear in Eqs. (13)

and (14) have been evaluated for values of α up to $\alpha = 200$. At these high α values χ has fallen practically to zero from the initial value of 45° at $\alpha = 0$. However, f is still significantly far from its saturation value f_∞ which it would have at $\alpha = \infty$. The convergence of these series is much more dependent upon the rate of decrease of the values of the $a_{11,j}$ and $b_{11,j}$ terms, as j increases, than upon the decreasing values of R^{j-1}, especially since R approaches

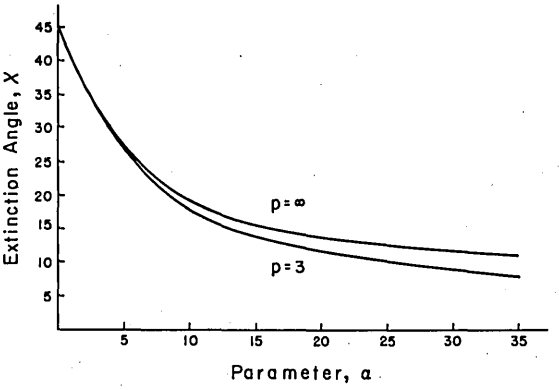


FIG. 3. Extinction angle χ as a function of the parameter α, together with its dependence on the axial ratio p.

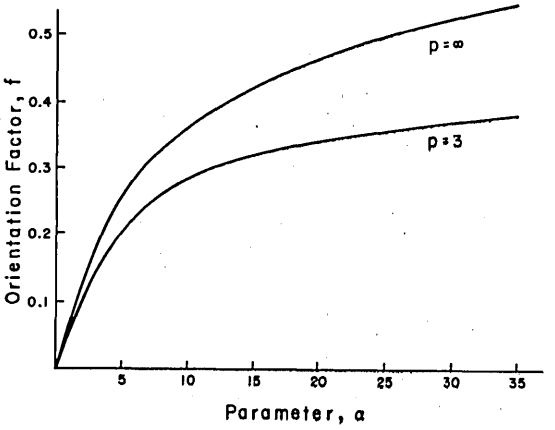


FIG. 4. Orientation factor f as a function of the parameter α, together with its dependence on the axial ratio p.

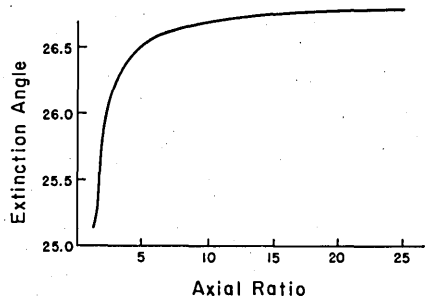


FIG. 5. Dependence of extinction angle on axial ratio for $\alpha = 5$.

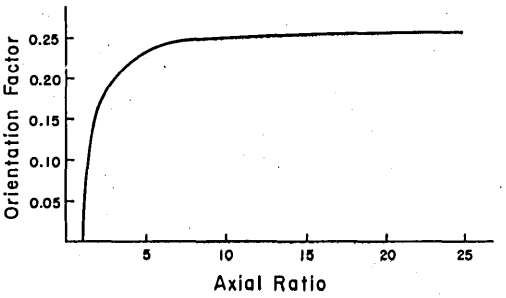


FIG. 6. Dependence of orientation factor on axial ratio for $\alpha = 5$.

unity very rapidly as p increases. The details of the computation problem are discussed in the next section. These series were evaluated for various values of the axial ratio p, and the data are summarized in Tables 1 and 2. The greater dependence on p at low p values is immediately apparent and is, of course, a consequence of the rapidity with which R approaches unity as p increases. The functions χ and f become insensitive to p at about $p = 16$, where they differ by very little from their values for $p = \infty$. By far the greatest dependence on p occurs below $p = 10$. These results are illustrated graphically in Figs. 3 to 6.

The values in Tables 1 and 2 are given for prolate ellipsoids ($p > 1$). However, the curves for an oblate ellipsoid of axial ratio $1/p$ are identical with those of a prolate ellipsoid of axial

Table 1. Extinction angle χ as a function of α for various axial ratios p.

α \ p	1.00	1.25	1.50	1.75	2.00	2.25	2.50	3.00	3.50	4.00	4.50
0.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00
0.25	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81
0.50	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62
0.75	41.44	41.44	41.44	41.44	41.44	41.44	41.44	41.44	41.44	41.44	41.45
1.00	40.27	40.27	40.27	40.28	40.28	40.28	40.28	40.29	40.29	40.29	40.29
1.25	39.12	39.12	39.12	39.13	39.14	39.14	39.14	39.15	39.16	39.16	39.16
1.50	37.98	37.99	38.00	38.01	38.02	38.02	38.03	38.04	38.05	38.05	38.06
1.75	36.87	36.88	36.89	36.91	36.92	36.93	36.94	36.96	36.97	36.98	36.99
2.00	35.78	35.79	35.81	35.84	35.86	35.88	35.89	35.91	35.93	35.94	35.95
2.25	34.72	34.74	34.76	34.80	34.82	34.85	34.87	34.90	34.92	34.94	34.95
2.50	33.69	33.71	33.74	33.79	33.82	33.86	33.88	33.93	33.96	33.98	33.99
3.00	31.72	31.74	31.80	31.87	31.93	31.98	32.03	32.09	32.14	32.17	32.19
3.50	29.87	29.91	30.00	30.09	30.18	30.25	30.31	30.41	30.47	30.52	30.55
4.00	28.16	28.21	28.32	28.45	28.56	28.66	28.75	28.87	28.95	29.02	29.06
4.50	26.57	26.64	26.78	26.94	27.08	27.21	27.31	27.47	27.58	27.66	27.71
5.00	25.10	25.18	25.36	25.55	25.73	25.88	26.01	26.20	26.33	26.42	26.49
6.00	22.50	22.61	22.85	23.11	23.35	23.55	23.72	23.98	24.16	24.29	24.38
7.00	20.30	20.44	20.74	21.05	21.35	21.60	21.81	22.13	22.36	22.51	22.63
8.00	18.43	18.60	18.94	19.31	19.65	19.94	20.20	20.57	20.84	21.02	21.16
9.00	16.84	17.03	17.40	17.81	18.20	18.53	18.81	19.25	19.54	19.75	19.90
10.00	15.48	15.67	16.08	16.53	16.95	17.31	17.62	18.09	18.42	18.66	18.83
12.50	12.82	13.03	13.47	13.97	14.45	14.87	15.23	15.80	16.19	16.48	16.68
15.00	10.90	11.11	11.57	12.09	12.60	13.05	13.45	14.07	14.52	14.84	15.07
17.50	9.46	9.67	10.12	10.64	11.16	11.63	12.05	12.72	13.20	13.55	13.81
20.00	8.35	8.55	8.98	9.49	10.02	10.50	10.93	11.62	12.13	12.51	12.79
22.50	7.46	7.65	8.06	8.57	9.08	9.56	10.00	10.71	11.25	11.64	11.94
25.00	6.75	6.92	7.32	7.80	8.30	8.78	9.22	9.95	10.50	10.91	11.22
30.00	5.66	5.81	6.17	6.61	7.08	7.54	7.98	8.71	9.28	9.72	10.06
35.00	4.86	5.00	5.32	5.73	6.17	6.61	7.03	7.76	8.34	8.80	9.15
40.00	4.27	4.39	4.68	5.05	5.46	5.88	6.28	7.00	7.58	8.04	8.41
45.00	3.80	3.91	4.17	4.51	4.90	5.29	5.67	6.37	6.95	7.41	7.78
50.00	3.42	3.53	3.76	4.08	4.43	4.80	5.17	5.84	6.41	6.87	7.25
60.00	2.86	2.94	3.15	3.42	3.73	4.06	4.39	5.00	5.54	5.99	6.36
80.00	2.14	2.21	2.37	2.58	2.82	3.09	3.36	3.88	4.34	4.75	5.08
100.00	1.72	1.77	1.90	2.07	2.27	2.49	2.71	3.15	3.56	3.90	4.20
200.00	0.86	0.89	0.95	1.04	1.14	1.26	1.38	1.62	1.84	2.04	2.21

a \ p	5.00	6.00	7.00	8.00	9.00	10.00	12.00	16.00	25.00	50.00	∞
0.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00
0.25	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81	43.81
0.50	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62	42.62
0.75	41.45	41.45	41.45	41.45	41.45	41.45	41.45	41.45	41.45	41.45	41.45
1.00	40.29	40.29	40.30	40.30	40.30	40.30	40.30	40.30	40.30	40.30	40.30
1.25	39.16	39.16	39.17	39.17	39.17	39.17	39.17	39.17	39.17	39.17	39.17
1.50	38.06	38.06	38.07	38.07	38.07	38.07	38.07	38.07	38.07	38.07	38.08
1.75	36.99	37.00	37.00	37.00	37.00	37.01	37.01	37.01	37.01	37.01	37.01
2.00	35.96	35.97	35.97	35.98	35.98	35.98	35.98	35.98	35.99	35.99	35.99
2.25	34.96	34.97	34.98	34.98	34.98	34.99	34.99	35.00	35.00	35.00	35.00
2.50	34.01	34.02	34.03	34.04	34.04	34.04	34.05	34.05	34.05	34.06	34.06
3.00	32.21	32.23	32.25	32.26	32.27	32.27	32.28	32.28	32.29	32.29	32.29
3.50	30.58	30.61	30.63	30.64	30.66	30.66	30.67	30.68	30.69	30.69	30.69
4.00	29.09	29.14	29.17	29.19	29.20	29.21	29.22	29.23	29.24	29.25	29.25
4.50	27.75	27.81	27.85	27.87	27.88	27.89	27.91	27.93	27.94	27.95	27.95
5.00	26.54	26.61	26.65	26.68	26.70	26.71	26.73	26.75	26.77	26.77	26.78
6.00	24.45	24.54	24.60	24.63	24.66	24.68	24.70	24.73	24.75	24.76	24.76
7.00	22.71	22.83	22.90	22.94	22.98	23.00	23.03	23.06	23.09	23.10	23.11
8.00	21.26	21.39	21.48	21.53	21.57	21.60	21.64	21.68	21.70	21.72	21.73
9.00	20.02	20.18	20.27	20.34	20.38	20.41	20.46	20.50	20.53	20.55	20.55
10.00	18.96	19.13	19.24	19.31	19.36	19.39	19.44	19.49	19.53	19.54	19.55
12.50	16.84	17.05	17.18	17.27	17.33	17.37	17.43	17.49	17.54	17.56	17.56
15.00	15.25	15.49	15.64	15.74	15.81	15.86	15.93	16.00	16.05	16.08	16.09
17.50	14.00	14.27	14.44	14.55	14.63	14.68	14.76	14.84	14.89	14.92	14.93
20.00	13.00	13.29	13.47	13.59	13.68	13.74	13.82	13.90	13.97	14.00	14.01
22.50	12.16	12.48	12.67	12.80	12.90	12.97	13.05	13.14	13.21	13.24	13.26
25.00	11.46	11.79	12.00	12.14	12.24	12.31	12.41	12.50	12.58	12.62	12.63
30.00	10.32	10.68	10.91	11.07	11.18	11.26	11.37	11.48	11.57	11.61	11.62
35.00	9.43	9.81	10.07	10.24	10.36	10.45	10.57	10.69	10.78	10.83	10.85
40.00	8.70	9.11	9.38	9.57	9.70	9.79	9.92	10.06	10.16	10.21	10.23
45.00	8.08	8.50	8.79	8.98	9.12	9.23	9.36	9.51	9.62	9.68	9.69
50.00	7.54	7.98	8.28	8.48	8.63	8.74	8.88	9.03	9.15	9.22	9.23
60.00	6.66	7.10	7.41	7.62	7.78	7.89	8.05	8.20	8.33	8.40	8.42
80.00	5.36	5.78	6.08	6.28	6.43	6.55	6.70	6.86	6.98	7.05	7.08
100.00	4.45	4.83	5.09	5.28	5.41	5.52	5.66	5.80	5.92	5.98	6.00
200.00	2.35	2.58	2.74	2.85	2.93	2.99	3.08	3.16	3.23	3.27	3.28

Table 2. Orientation factor f as a function of α for various axial ratios p.

α \ p	1.00	1.25	1.50	1.75	2.00	2.25	2.50	3.00	3.50	4.00	4.50
0.00	0.00000	0.00000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	0.00000	0.00366	0.0064	0.0085	0.0100	0.0112	0.0121	0.0133	0.0141	0.0147	0.0151
0.50	0.00000	0.00729	0.0128	0.0169	0.0199	0.0223	0.0241	0.0266	0.0282	0.0293	0.0301
0.75	0.00000	0.01089	0.0191	0.0252	0.0298	0.0332	0.0359	0.0397	0.0421	0.0437	0.0449
1.00	0.00000	0.01443	0.0253	0.0334	0.0394	0.0440	0.0476	0.0525	0.0557	0.0579	0.0595
1.25	0.00000	0.01791	0.0314	0.0414	0.0489	0.0546	0.0590	0.0651	0.0691	0.0718	0.0737
1.50	0.00000	0.02129	0.0373	0.0492	0.0581	0.0649	0.0701	0.0774	0.0821	0.0853	0.0876
1.75	0.00000	0.02458	0.0430	0.0568	0.0671	0.0749	0.0809	0.0893	0.0947	0.0984	0.1010
2.00	0.00000	0.02776	0.0486	0.0641	0.0757	0.0845	0.0913	0.1007	0.1069	0.1110	0.1139
2.25	0.00000	0.03082	0.0540	0.0712	0.0840	0.0938	0.1013	0.1118	0.1185	0.1231	0.1264
2.50	0.00000	0.03376	0.0591	0.0779	0.0920	0.1027	0.1109	0.1223	0.1297	0.1348	0.1383
3.00	0.00000	0.03925	0.0687	0.0906	0.1069	0.1193	0.1288	0.1421	0.1507	0.1565	0.1606
3.50	0.00000	0.04422	0.0774	0.1020	0.1204	0.1344	0.1451	0.1601	0.1697	0.1762	0.1809
4.00	0.00000	0.04868	0.0852	0.1123	0.1326	0.1480	0.1598	0.1763	0.1869	0.1941	0.1992
4.50	0.00000	0.05266	0.0922	0.1216	0.1436	0.1602	0.1730	0.1909	0.2024	0.2103	0.2158
5.00	0.00000	0.05620	0.0984	0.1299	0.1534	0.1712	0.1849	0.2041	0.2165	0.2249	0.2308
6.00	0.00000	0.06211	0.1089	0.1438	0.1700	0.1900	0.2053	0.2268	0.2407	0.2502	0.2568
7.00	0.00000	0.06674	0.1172	0.1550	0.1835	0.2052	0.2220	0.2456	0.2609	0.2712	0.2786
8.00	0.00000	0.07038	0.1238	0.1640	0.1945	0.2178	0.2358	0.2613	0.2778	0.2891	0.2970
9.00	0.00000	0.07326	0.1291	0.1714	0.2035	0.2282	0.2474	0.2746	0.2923	0.3044	0.3129
10.00	0.00000	0.07557	0.1334	0.1774	0.2111	0.2370	0.2573	0.2860	0.3048	0.3176	0.3267
12.50	0.00000	0.07961	0.1411	0.1886	0.2253	0.2539	0.2764	0.3086	0.3299	0.3444	0.3548
15.00	0.00000	0.08213	0.1461	0.1960	0.2351	0.2658	0.2902	0.3254	0.3487	0.3649	0.3764
17.50	0.00000	0.08378	0.1495	0.2012	0.2421	0.2746	0.3005	0.3383	0.3635	0.3810	0.3936
20.00	0.00000	0.08492	0.1518	0.2050	0.2473	0.2812	0.3085	0.3485	0.3755	0.3942	0.4077
22.50	0.00000	0.08573	0.1536	0.2078	0.2513	0.2864	0.3148	0.3568	0.3853	0.4052	0.4196
25.00	0.00000	0.08634	0.1549	0.2099	0.2544	0.2905	0.3199	0.3637	0.3936	0.4147	0.4299
30.00	0.00000	0.08714	0.1567	0.2130	0.2589	0.2966	0.3276	0.3744	0.4069	0.4299	0.4467
35.00	0.00000	0.08764	0.1578	0.2149	0.2619	0.3008	0.3331	0.3823	0.4169	0.4418	0.4599
40.00	0.00000	0.08797	0.1585	0.2163	0.2640	0.3038	0.3371	0.3883	0.4248	0.4513	0.4707
45.00	0.00000	0.08820	0.1591	0.2172	0.2656	0.3060	0.3400	0.3930	0.4311	0.4589	0.4796
50.00	0.00000	0.08836	0.1595	0.2179	0.2667	0.3077	0.3424	0.3967	0.4362	0.4653	0.4871
60.00	0.00000	0.08858	0.1600	0.2189	0.2683	0.3100	0.3456	0.4021	0.4438	0.4750	0.4987
80.00	0.00000	0.08880	0.1605	0.2199	0.2699	0.3125	0.3492	0.4082	0.4528	0.4868	0.5132
100.00	0.00000	0.08891	0.1607	0.2203	0.2707	0.3138	0.3509	0.4114	0.4576	0.4933	0.5213
200.00	0.00000	0.08904	0.1611	0.2209	0.2718	0.3155	0.3535	0.4161	0.4649	0.5034	0.5342

α \ p	5.00	6.00	7.00	8.00	9.00	10.00	12.00	16.00	25.00	50.00	∞
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	0.0154	0.0158	0.0160	0.0161	0.0162	0.0163	0.0164	0.0165	0.0166	0.0166	0.0167
0.50	0.0307	0.0314	0.0319	0.0322	0.0324	0.0325	0.0327	0.0329	0.0331	0.0332	0.0332
0.75	0.0458	0.0469	0.0476	0.0480	0.0483	0.0486	0.0489	0.0492	0.0494	0.0495	0.0496
1.00	0.0606	0.0621	0.0630	0.0636	0.0640	0.0643	0.0647	0.0651	0.0654	0.0656	0.0656
1.25	0.0751	0.0769	0.0781	0.0788	0.0793	0.0797	0.0802	0.0807	0.0811	0.0813	0.0813
1.50	0.0892	0.0914	0.0927	0.0936	0.0942	0.0947	0.0953	0.0958	0.0963	0.0965	0.0966
1.75	0.1029	0.1053	0.1069	0.1080	0.1087	0.1092	0.1098	0.1105	0.1110	0.1113	0.1114
2.00	0.1161	0.1189	0.1206	0.1218	0.1226	0.1231	0.1239	0.1246	0.1252	0.1255	0.1256
2.25	0.1287	0.1319	0.1338	0.1351	0.1359	0.1366	0.1374	0.1382	0.1388	0.1392	0.1393
2.50	0.1409	0.1443	0.1464	0.1478	0.1488	0.1494	0.1503	0.1512	0.1519	0.1523	0.1524
3.00	0.1636	0.1676	0.1700	0.1716	0.1727	0.1735	0.1745	0.1756	0.1764	0.1768	0.1769
3.50	0.1842	0.1887	0.1914	0.1932	0.1945	0.1954	0.1965	0.1977	0.1986	0.1991	0.1992
4.00	0.2029	0.2078	0.2108	0.2128	0.2142	0.2152	0.2165	0.2177	0.2187	0.2192	0.2194
4.50	0.2198	0.2251	0.2284	0.2306	0.2321	0.2331	0.2345	0.2359	0.2370	0.2376	0.2377
5.00	0.2351	0.2408	0.2444	0.2467	0.2483	0.2494	0.2509	0.2524	0.2536	0.2542	0.2544
6.00	0.2617	0.2681	0.2721	0.2747	0.2765	0.2778	0.2795	0.2812	0.2825	0.2832	0.2834
7.00	0.2839	0.2910	0.2954	0.2983	0.3003	0.3017	0.3036	0.3054	0.3069	0.3076	0.3079
8.00	0.3028	0.3106	0.3153	0.3185	0.3206	0.3222	0.3242	0.3262	0.3278	0.3286	0.3289
9.00	0.3191	0.3275	0.3326	0.3360	0.3383	0.3399	0.3421	0.3443	0.3460	0.3469	0.3472
10.00	0.3334	0.3423	0.3477	0.3513	0.3538	0.3556	0.3579	0.3603	0.3621	0.3630	0.3633
12.50	0.3624	0.3726	0.3788	0.3830	0.3858	0.3879	0.3906	0.3933	0.3953	0.3964	0.3968
15.00	0.3848	0.3962	0.4032	0.4078	0.4110	0.4133	0.4163	0.4193	0.4216	0.4228	0.4232
17.50	0.4028	0.4152	0.4229	0.4279	0.4314	0.4340	0.4373	0.4406	0.4431	0.4444	0.4449
20.00	0.4177	0.4311	0.4393	0.4448	0.4486	0.4513	0.4549	0.4585	0.4612	0.4626	0.4631
22.50	0.4302	0.4445	0.4533	0.4591	0.4632	0.4661	0.4699	0.4737	0.4766	0.4782	0.4787
25.00	0.4412	0.4564	0.4659	0.4721	0.4765	0.4796	0.4837	0.4878	0.4910	0.4926	0.4932
30.00	0.4592	0.4761	0.4867	0.4937	0.4985	0.5020	0.5066	0.5113	0.5148	0.5166	0.5173
35.00	0.4736	0.4921	0.5037	0.5114	0.5167	0.5206	0.5257	0.5308	0.5347	0.5367	0.5374
40.00	0.4854	0.5055	0.5181	0.5265	0.5324	0.5366	0.5422	0.5478	0.5521	0.5543	0.5551
45.00	0.4952	0.5168	0.5304	0.5395	0.5459	0.5505	0.5565	0.5626	0.5673	0.5698	0.5706
50.00	0.5037	0.5267	0.5413	0.5511	0.5580	0.5630	0.5696	0.5763	0.5814	0.5841	0.5850
60.00	0.5169	0.5425	0.5590	0.5701	0.5780	0.5838	0.5914	0.5991	0.6051	0.6083	0.6094
80.00	0.5338	0.5632	0.5826	0.5960	0.6055	0.6125	0.6218	0.6314	0.6389	0.6429	0.6442
100.00	0.5434	0.5753	0.5966	0.6113	0.6219	0.6298	0.6403	0.6511	0.6596	0.6642	0.6657
200.00	0.5588	0.5952	0.6199	0.6372	0.6498	0.6592	0.6718	0.6850	0.6954	0.7010	0.7029

ratio p. This ratio enters into Eqs. (13) and (14) only through the function R. From Eq. (1) it is clear that when p changes to $1/p$, R changes sign: $R(p) = -R(1/p)$. However, in the summations $\Sigma R^{j-1}a_{11,j}$ and $\Sigma R^{j-1}b_{11,j}$, only the a and b coefficients for which j is odd have values different from zero. Hence, only even powers of R enter into the summations, from which it follows that $\chi(\alpha,p) = \chi(\alpha,1/p)$ and $f(\alpha,p) = -f(\alpha,1/p)$.[30] These relations were clearly recognized by Peterlin and Stuart[22] (see their figures for χ and f as functions of α), but were not explicitly stated by them. From Eq. (13),[31]

$$\lim_{R\to 0} \tan 2\chi = \frac{-b_{11,1}}{a_{11,1}} = \frac{6}{\alpha}.$$

The equation $\lim_{\alpha\to 0} \tan 2\chi = 6/\alpha$, which had already been derived by Boeder[20] for the case of thin rods ($R = 1$), is the same as Eq. (15) as α approaches zero.

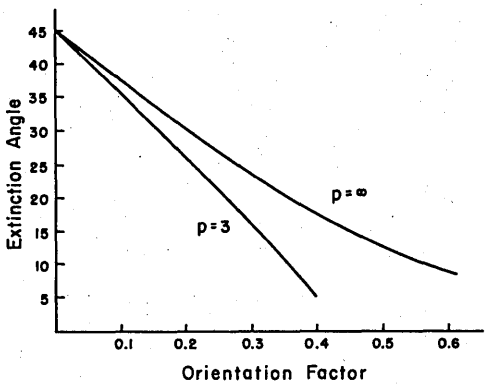


FIG. 7. Extinction angle as a function of orientation factor, together with its dependence on the axial ratio p.

Thus in all cases, when α approaches zero and the degree of orientation becomes very small, f approaches zero and $\tan 2\chi = 6/\alpha$. This holds for all p values and, in the limiting case where p approaches 1—that is, the ellipsoid approaches a sphere—these relations hold also for all values of α.

If experimental values of χ are plotted as a function of Δn, the data may be fitted by a theoretical curve of χ as a function of kf, where k is an adjustable constant used to fit the data to the curve. When determined, it gives the optical factor $(g_1 - g_2)$, since

$$k = \frac{2\pi c(g_1 - g_2)}{n}.$$

If this procedure is adopted, it is unnecessary to extrapolate to low values of α where the experimental errors are large.[9,10] The nature of the curve for χ as a function of f and its dependence on p are shown in Fig. 7.

Some question may arise as to the validity of the Peterlin and Stuart solution, especially since the viscosity problem, treated by Peterlin with the same distribution function, is at variance with the results of Simha,[32,33] whose treatment is considered to be the valid one.[34] The Simha treatment of viscosity is also identical with that of Kuhn and Kuhn[35] for low gradient. However, the disagreement in the viscosity theories does not arise from the use of an incorrect distribution function,[36] but rather from Peterlin's omission of certain terms in the hydrodynamic equations, which were taken into account by Simha. Thus, the inadequacy of the viscosity theory in no way affects the valid use of this distribution function for the treatment of double refraction of flow.

If the omitted terms were taken into account, Peterlin's viscosity treatment would presumably be valid. During the course of the computation in connection with the present problem

numerous other coefficients besides the $a_{11,j}$ and $b_{11,j}$ terms were evaluated. The availability of these additional terms may be of use for a viscosity theory formulated on a similar basis.

The first and crucial step in the organization of this problem for machine computation was an analysis of the recurrence formulas (6–9) and the following generalizations applying to them:

1. All terms with negative indices are zero;
2. $a_{nm,0} = 0$, and $a_{00,j} = 0$, except $a_{00,0} = 1/2\pi$;
3. $b_{nm,0} = b_{n0,j} = b_{0m,j} = 0$.

A rigorous analytical study seemed to be out of the question. Therefore several decisions were made arbitrarily. Since the limited internal storage capacity of the Mark I computer was one of the chief obstacles to be overcome, it was decided to consider each 24-column storage counter as two counters of 12 columns each so that the a and b terms for any nm,j could be contained in one counter, all terms being carried to ten places of decimals.

It may be observed that the values of all terms with an index of j_i depend entirely upon terms with the index j_{i-1}. Thus the decision was made to assign a storage counter to each nm combination, to put in these counters the $a_{nm,0}$ and $b_{nm,0}$ terms, to compute from these the $a_{nm,1}$ and $b_{nm,1}$ terms and transfer these results into the same set of counters, then to compute the $a_{nm,2}$ and $b_{nm,2}$ terms, and so on. This led to the adoption of the viewpoint that values of points in the nm-plane (Fig. 8) were being computed at time "zero," "one," . . ., that is, $j = 0, 1, \ldots$ There is, of course, a pair of values at each point, one value for the a term and one for the b term.

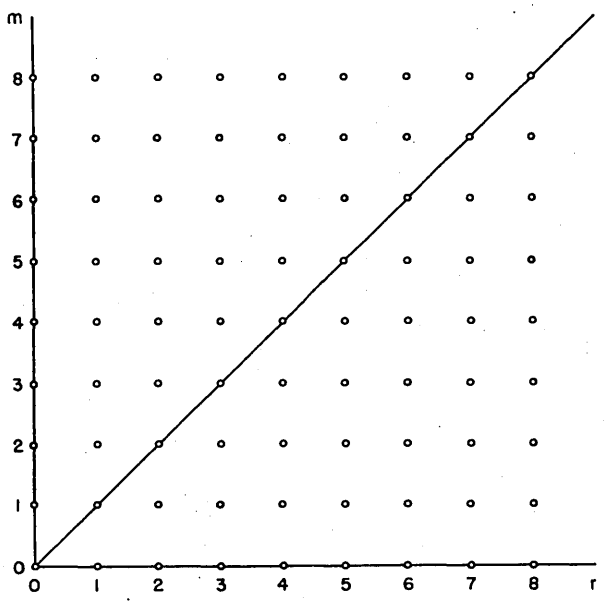


FIG. 8. Array of points in nm-plane.

It was hoped that a newly computed pair of values $a_{nm,j}$ and $b_{nm,j}$ might immediately replace the pair $a_{nm,j-1}$ and $b_{nm,j-1}$ in the storage counter assigned to that particular nm combination. Unfortunately, this seemed impossible because the previous values are needed to compute values at neighboring points, and it appeared that a second set of storage counters would be required. However, further study of the recurrence formulas shows that the values of all terms with the index m_i depend only upon terms with m_{i-1} or m_{i+1} indices. Since it is known that, for $j = 0$, the only point at which either the a or the b term has a value is the point 0,0 in the nm-plane, it is obvious that at time "one" ($j = 1$) values occur only in the

row $m = 1$ and at time "two" only in the rows $m = 0$ and $m = 2$. Thus it is seen that for even j only the rows of even m, which depend only upon the rows of odd m, need be calculated, and that for odd j it is necessary to compute only the odd rows in the m direction. This meant that only one set of storage counters was needed, since a newly computed pair of values for a point nm,j could immediately replace the values in storage for that point in the nm-plane, which were values for $nm,j-2$.

An inspection of the n indices shows that values can extend only one place farther in the n direction with each increase of 1 in j. Since 0,0 is the only point with a value for $j = 0$, it is clear that all $a_{nm,j}$ and $b_{nm,j}$ are zero for $n > j$.

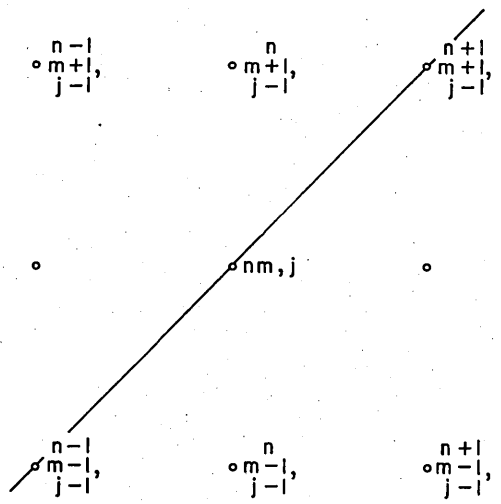


FIG. 9. Illustration of recurrence relations for a given term represented in the nm-plane.

A parallel assertion holds for $m > j$, but this becomes inconsequential in view of the fact that, in the present problem, it has been found unnecessary to compute the value of any point where $m > n$. This is the same as saying that the points to the left of the diagonal (Fig. 8) have no effect upon the values at the point 1,1 in the nm-plane, which is the only point at which results are required. By the recurrence relations, the a and b terms at any point in the nm-plane depend upon the previous time's terms at the three points centered directly above it and the three centered directly below it (Fig. 9). But, for any point on the diagonal ($n = m$), the coefficients applied to the first two points in the upper row vanish. Likewise, for any point just off the diagonal to the right ($n = m + 1$), the coefficient of the first term in the upper row becomes zero. Therefore, values to the left of the diagonal have no effect at any time upon any point on, or to the right of, the diagonal.

As a summary of the results of this analysis, the following may be added to the initial generalizations:

4. If m is even and j is odd, $a_{nm,j} = 0$, and $b_{nm,j} = 0$;
5. If m is odd and j is even, $a_{nm,j} = 0$, and $b_{nm,j} = 0$;
6. If $n > j$, or if $m > j$, $a_{nm,j} = 0$, and $b_{nm,j} = 0$;
7. In this problem, terms where $m > n$ need not be computed as they do not affect the values of $a_{11,j}$ or $b_{11,j}$.

Considering these findings as well as the storage capacity of Mark I, it was decided to compute the a and b terms at all the points within the block $n \leqslant 8$, $m \leqslant 8$, $j \leqslant 15$ that contributed anything to the final results. This meant that 40 points in the nm-plane were to be computed and that 40 storage counters would be required. Since terms with n or m indices of 9 were not being computed when $j = 9$, terms with n or m indices of 8 would be in error when $j = 10$, so these terms were not computed. Likewise, terms having n or m indices of 7

would be incorrect when $j = 11$ and so were not computed. Thus, when $j = 15$, the terms $a_{1,1;15}$ and $b_{1,1;15}$ were the only terms calculated. It remained then for the computer to determine whether 15 terms, or 8 nonzero terms, would suffice for the convergence of the series, or whether some change in organization would have to be made. All in all, there were now the following 136 $n,m;j$ points to be computed:

1,1;1	2,1;5	5,2;6	7,3;7	6,2;8	6,1;9	2,2;10	5,5;11
1,0;2	3,1;5	6,2;6	5,5;7	7,2;8	7,1;9	3,2;10	1,0;12
2,0;2	4,1;5	4,4;6	6,5;7	8,2;8	3,3;9	4,2;10	2,0;12
2,2;2	5,1;5	5,4;6	7,5;7	4,4;8	4,3;9	5,2;10	3,0;12
1,1;3	3,3;5	6,4;6	7,7;7	5,4;8	5,3;9	6,2;10	4,0;12
2,1;3	4,3;5	6,6;6	1,0;8	6,4;8	6,3;9	4,4;10	2,2;12
3,1;3	5,3;5	1,1;7	2,0;8	7,4;8	7,3;9	5,4;10	3,2;12
3,3;3	5,5;5	2,1;7	3,0;8	8,4;8	5,5;9	6,4;10	4,2;12
1,0;4	1,0;6	3,1;7	4,0;8	6,6;8	6,5;9	6,6;10	4,4;12
2,0;4	2,0;6	4,1;7	5,0;8	7,6;8	7,5;9	1,1;11	1,1;13
3,0;4	3,0;6	5,1;7	6,0;8	8,6;8	7,7;9	2,1;11	2,1;13
4,0;4	4,0;6	6,1;7	7,0;8	8,8;8	1,0;10	3,1;11	3,1;13
2,2;4	5,0;6	7,1;7	8,0;8	1,1;9	2,0;10	4,1;11	3,3;13
3,2;4	6,0;6	3,3;7	2,2;8	2,1;9	3,0;10	5,1;11	1,0;14
4,2;4	2,2;6	4,3;7	3,2;8	3,1;9	4,0;10	3,3;11	2,0;14
4,4;4	3,2;6	5,3;7	4,2;8	4,1;9	5,0;10	4,3;11	2,2;14
1,1;5	4,2;6	6,3;7	5,2;8	5,1;9	6,0;10	5,3;11	1,1;15

Next, Eqs. (8) and (9) were solved and written in the form

$$a_{nm,j} = -\frac{\alpha m A + kB}{\alpha m^2 + k^2/\alpha}, \quad b_{nm,j} = \frac{kA - \alpha m B}{\alpha m^2 + k^2/\alpha}, \tag{17}$$

where $k = 2n(2n + 1)$,

$$A = a_1c_1 + a_2c_2 + a_3c_3 + a_4c_4 + a_5c_5 + a_6c_6,$$

$$B = b_1c_1 + b_2c_2 + b_3c_3 + b_4c_4 + b_5c_5 + b_6c_6,$$

$$c_1 = \frac{2n + 1}{4(4n - 3)(4n - 1)},$$

$$c_2 = -\frac{3}{4(4n - 1)(4n + 3)},$$

$$c_3 = -\frac{n}{2(4n + 3)(4n + 5)},$$

$$c_4 = -(2n - 2m - 3)(2n - 2m - 2)(2n - 2m - 1)(2n - 2m)c_1,$$

$$c_5 = -(2n - 2m - 1)(2n - 2m)(2n + 2m + 1)(2n + 2m + 2)c_2,$$

$$c_6 = -(2n + 2m + 1)(2n + 2m + 2)(2n + 2m + 3)(2n + 2m + 4)c_3,$$

and

$$a_1 = a_{n-1,m-1;j-1}, \quad a_2 = a_{n,m-1;j-1}, \quad a_3 = a_{n+1,m-1;j-1},$$

$$a_4 = a_{n-1,m+1;j-1}, \quad a_5 = a_{n,m+1;j-1}, \quad a_6 = a_{n+1,m+1;j-1},$$

$$b_1 = b_{n-1,m-1;j-1}, \quad b_2 = b_{n,m-1;j-1}, \quad b_3 = b_{n+1,m-1;j-1},$$

$$b_4 = b_{n-1,m+1;j-1}, \quad b_5 = b_{n,m+1;j-1}, \quad b_6 = b_{n+1,m+1;j-1}.$$

In order that one computation routine might be used throughout, it was decided to compute the terms for $m = 0$ exactly like the others and then to wipe out the $b_{n0,j}$ term and double the $a_{n0,j}$ term, in accordance with Eqs. (6) and (7).

The first control tape placed upon the calculator simply computed c_1, c_2, c_3, c_4, c_5, c_6, k, k^2, m^2, and the sum of these nine quantities for each of the 136 points, and punched the results on cards. These calculations were not checked internally. Next was run another simple tape which fed these punched cards into the machine and printed all quantities. These printed sheets were checked for accuracy. Since all these quantities are functions of n and m only, most of them were computed more than once. In fact, the case of $n = m = 1$ occurs eight times. A check that the results in these duplicate cases agreed was a part of the visual check. However, the real check at this point was in the fact that all these quantities had been hastily hand computed in advance, and the tape was run on Mark I simply to verify these results and to produce them on punched cards. This portion of the job consumed 6 hours of machine time.

The next step was the calculation of the terms $a_{nm,j}$ and $b_{nm,j}$. This was by far the major part of the problem and consumed some two weeks of machine time. A separate run was made for each value of α, the quantities α and $1/\alpha$ being placed in constant switches. The main control tape read into six working counters the values in the six counters containing the particular a_1 and b_1, a_2 and b_2, . . ., a_6 and b_6 applying to the point being computed. A subsequence routine then computed $a_{nm,j}$ and $b_{nm,j}$, whereupon the main control tape directed these results into the counter assigned to that nm combination and set up the six working counters for the next point. The main tape thus comprised 136 of these small sections. When it had finished its run, which required about $6\frac{1}{2}$ hours, it was started over again with new values of α and $1/\alpha$.

The subsequence routine, which consumed about 3 minutes of running time for the computation of the a and b terms at each point, comprised the following operations. First, the cards containing the nine constants for that point were fed into the calculator, and the sum of these constants was checked against the punched card containing their sum. Then the A and B of Eqs. (17) were computed by 12 multiplications and checked by six additional multiplications, using the relation

$$A + B = (a_1 + b_1)c_1 + \cdots + (a_6 + b_6)c_6.$$

Then $a_{nm,j}$ and $b_{nm,j}$ were computed directly by means of Eqs. (17). Since these equations involve both m and m^2, both k and k^2, and both α and $1/\alpha$, no intermediate checks were necessary. The only check at this point was the substitution of the computed $a_{nm,j}$ and $b_{nm,j}$ into Eqs. (8) and (9). Next was applied the test whether m was zero, in which case the a term was doubled and the b term erased, in accordance with Eqs. (6) and (7). Then $a_{nm,j}$ and $b_{nm,j}$ were printed and, if they were for the point $1,1;j$, they were punched on cards.

For very small values of α the convergence was extremely rapid, the terms $a_{1,1;5}$ and $b_{1,1;5}$ being zero to ten decimal places. Not until α reached 6 did the terms $a_{1,1;15}$ and $b_{1,1;15}$ exceed 10^{-10}, and for all α under 25 it was felt that the error being committed in the dropping of the

$a_{1,1;17}$ and $b_{1,1;17}$ terms was insignificant. However, it could be seen that more terms would be needed for accuracy when α was large. So the case $\alpha = 100$ was then run in order that the results might indicate that many unnecessary terms were being computed. It was found that, even in this case of large α, the values of terms with an m index of 5 or more never exceeded 10^{-10}, and that there were several other points where this was the case. Thus it was possible to code a new control tape which would compute the terms through $a_{1,1;23}$ and $b_{1,1;23}$ with the use of only 36 counters for the storage of 36 points in the nm-plane. This required the computation of the following 244 $n,m;j$ points:

1,1;1	2,2;6	6,2;8	5,2;10	11,0;12	6,0;14	5,0;16	2,2;18
1,0;2	3,2;6	7,2;8	6,2;10	12,0;12	7,0;14	6,0;16	3,2;18
2,0;2	4,2;6	8,2;8	7,2;10	2,2;12	8,0;14	7,0;16	4,2;18
2,2;2	5,2;6	4,4;8	8,2;10	3,2;12	9,0;14	8,0;16	5,2;18
1,1;3	6,2;6	1,1;9	4,4;10	4,2;12	10,0;14	2,2;16	6,2;18
2,1;3	4,4;6	2,1;9	1,1;11	5,2;12	2,2;14	3,2;16	4,4;18
3,1;3	1,1;7	3,1;9	2,1;11	6,2;12	3,2;14	4,2;16	1,1;19
3,3;3	2,1;7	4,1;9	3,1;11	7,2;12	4,2;14	5,2;16	2,1;19
1,0;4	3,1;7	5,1;9	4,1;11	8,2;12	5,2;14	6,2;16	3,1;19
2,0;4	4,1;7	6,1;9	5,1;11	4,4;12	6,2;14	7,2;16	4,1;19
3,0;4	5,1;7	7,1;9	6,1;11	1,1;13	7,2;14	8,2;16	5,1;19
4,0;4	6,1;7	8,1;9	7,1;11	2,1;13	8,2;14	4,4;16	3,3;19
2,2;4	7,1;7	9,1;9	8,1;11	3,1;13	4,4;14	1,1;17	4,3;19
3,2;4	3,3;7	3,3;9	9,1;11	4,1;13	1,1;15	2,1;17	5,3;19
4,2;4	4,3;7	4,3;9	10,1;11	5,1;13	2,1;15	3,1;17	1,0;20
4,4;4	5,3;7	5,3;9	11,1;11	6,1;13	3,1;15	4,1;17	2,0;20
1,1;5	6,3;7	6,3;9	3,3;11	7,1;13	4,1;15	5,1;17	3,0;20
2,1;5	7,3;7	7,3;9	4,3;11	8,1;13	5,1;15	6,1;17	4,0;20
3,1;5	1,0;8	1,0;10	5,3;11	9,1;13	6,1;15	7,1;17	2,2;20
4,1;5	2,0;8	2,0;10	6,3;11	10,1;13	7,1;15	3,3;17	3,2;20
5,1;5	3,0;8	3,0;10	7,3;11	11,1;13	8,1;15	4,3;17	4,2;20
3,3;5	4,0;8	4,0;10	1,0;12	3,3;13	9,1;15	5,3;17	4,4;20
4,3;5	5,0;8	5,0;10	2,0;12	4,3;13	3,3;15	6,3;17	1,1;21
5,3;5	6,0;8	6,0;10	3,0;12	5,3;13	4,3;15	7,3;17	2,1;21
1,0;6	7,0;8	7,0;10	4,0;12	6,3;13	5,3;15	1,0;18	3,1;21
2,0;6	8,0;8	8,0;10	5,0;12	7,3;13	6,3;15	2,0;18	3,3;21
3,0;6	2,2;8	9,0;10	6,0;12	1,0;14	7,3;15	3,0;18	1,0;22
4,0;6	3,2;8	10,0;10	7,0;12	2,0;14	1,0;16	4,0;18	2,0;22
5,0;6	4,2;8	2,2;10	8,0;12	3,0;14	2,0;16	5,0;18	2,2;22
6,0;6	5,2;8	3,2;10	9,0;12	4,0;14	3,0;16	6,0;18	1,1;23
		4,2;10	10,0;12	5,0;14	4,0;16		

This tape was run for all values of α from 25 up, each run consuming 12 hours of machine time. The four extra terms in the series improved the convergence considerably. Although no accurate estimate of the size of the remainder term could be made, some idea of the rate of convergence of these series at high α values may be obtained from the data of Tables 3

Table 3. Values of $-\Sigma a_{11,j}$ as a function of j for various values of α.

j \ α	25	40	60	100	200
1	0.037622	0.038913	0.039395	0.039646	0.039753
3	0.045133	0.048785	0.050268	0.051071	0.051419
5	0.046291	0.052465	0.055325	0.056967	0.057700
7	0.045537	0.053452	0.057790	0.060498	0.061765
9	0.044715	0.053261	0.058821	0.062630	0.064509
11	0.044315	0.052705	0.059032	0.063820	0.066313
13	0.044244	0.052205	0.058871	0.064424	0.067487
15	0.044290	0.051896	0.058608	0.064699	0.068260
17	0.044338	0.051754	0.058368	0.064802	0.068783
19	0.044361	0.051713	0.058191	0.064821	0.069127
21	0.044367	0.051715	0.058079	0.064804	0.069338
23	0.044366	0.051729	0.058013	0.064771	0.069460

Table 4. Values of $\Sigma b_{11,j}$ as a function of j for various values of α.

j \ α	25	40	60	100	200
1	0.009029	0.005837	0.003939	0.002379	0.001193
3	0.015774	0.010806	0.007460	0.004560	0.002298
5	0.019813	0.014767	0.010583	0.006605	0.003359
7	0.021399	0.017518	0.013207	0.008500	0.004385
9	0.021545	0.019029	0.015137	0.010092	0.005295
11	0.021256	0.019628	0.016389	0.011325	0.006051
13	0.020927	0.019723	0.017107	0.012225	0.006651
15	0.020828	0.019616	0.017461	0.012848	0.007110
17	0.020818	0.019482	0.017597	0.013258	0.007451
19	0.020833	0.019383	0.017620	0.013519	0.007699
21	0.020844	0.019328	0.017595	0.013677	0.007874
23	0.020849	0.019303	0.017563	0.013768	0.007997

and 4, where the values of the $-\Sigma a_{11,j}$ and $\Sigma b_{11,j}$ series are given for several values of α. These data are also plotted in Figs. 10 and 11. It will be noted that the b terms converge much more

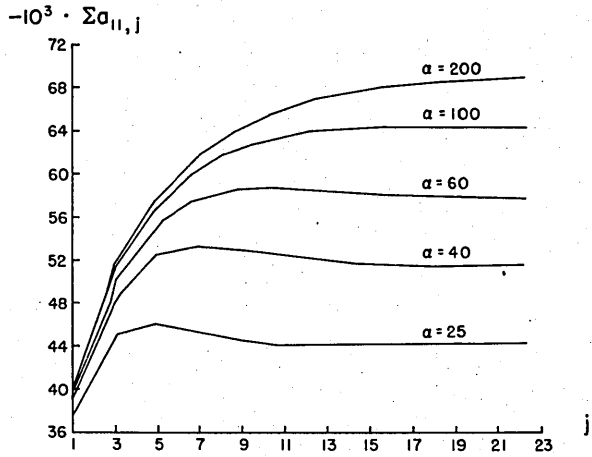


FIG. 10. Values of $-\Sigma a_{11,j}$ as a function of j for various values of α.

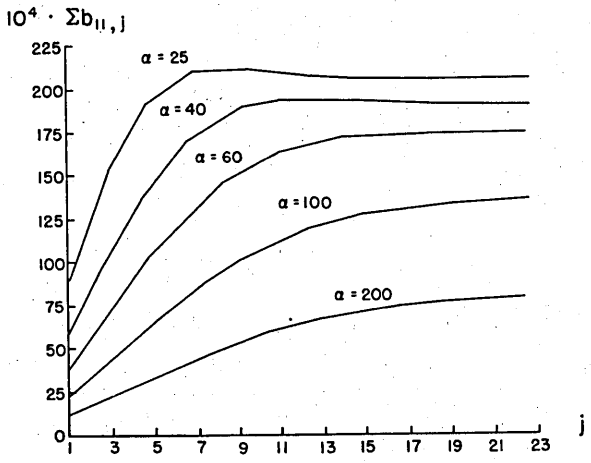


FIG. 11. Values of $\Sigma b_{11,j}$ as a function of j for various values of α.

slowly than do the a terms. It will also be noted that the series for $\alpha = 200$ and $\alpha = 100$ clearly have not converged. It is felt that, for $\alpha \leqslant 60$, the error that has been committed is well under 1 percent, but that, for $\alpha = 80$, 100, or 200, the results listed in Tables 1 and 2 are significantly in error for all but the very small values of p. These "bad" results have been included in the tables in the hope that they may shed some light on the question of convergence. If they had been included in Fig. 7, for the case $p = \infty$, they would lie significantly lower than the curve shown, so that the extrapolated value of f at $\chi = 0$ would be about 0.75 instead of the value 1 as it should be and as, in fact, it appears to be when the values at $\alpha = 80$, 100, and 200 are omitted.

It is unfortunate that the series of b terms converges slowly when the sum of the series is small, thus making the proportional error committed in the truncation of the series that much greater. It should be noted that, as α approaches ∞, all terms in the series $\Sigma b_{11,j}$ approach zero and thus approach each other. If the sum of this series is desired to within a certain "percentage error," the number of terms required becomes infinite as α approaches ∞.

In order to study the propagation of errors, the case $\alpha = 100$ was run with an intentional error introduced in the value of $a_{00,0}$. As the successive terms were computed, the effects of this error became less and less. Therefore it is believed that roundoff errors do not accumulate and need not be considered.

The final control tape in the problem governed the computation of tan 2χ and f^2. Since only three or four decimal places of accuracy were required in the results, in the interests of economy it was decided to obtain χ and f from tables by hand methods. Thus it was possible to complete this stage of the job in 1 day of machine use.

The deck of cards containing the terms $a_{11,j}$ and $b_{11,j}$ for all the values of α was fed into the calculator along with another deck of cards containing R and R^2 [Eq. (1)] for each of several values of p. Then, by means of Eqs. (13) and (14), tan 2χ and f^2 were obtained directly by multiplications and a division. The multiplications in the computation of tan 2χ were checked by the distributive law, those in the computation of f^2 by the associative law, and the division by multiplying the quotient by the divisor. All results were printed in duplicate and were also punched on cards in order that they might be available for future use if desired.

Using the two sets of printed results, the values of χ and f were computed by two people working independently and later comparing the two sets of computations.[37]

It is believed that no error has been made at any stage of the computation except in the truncation of the series of $a_{11,j}$ and $b_{11,j}$ terms. Thus it is felt that the results contained in Tables 1 and 2 are entirely accurate when either α or p is very small, are significantly in error when α is very large (unless p is very small), and are probably accurate to well within 1 percent in all cases except when $\alpha = 80$, 100, or 200.

The total machine time consumed by this problem was slightly over two weeks. It is felt that here is a perfect example of the situation where a large-scale automatic calculator has a tremendous advantage over hand or desk computers, perhaps not so much in the matter of speed as in the problem of organization.

The method of double refraction of flow in systems containing large asymmetric molecules gives experimental data which, when interpreted in light of the theory of Peterlin and Stuart, enable one to calculate molecular lengths; information about the polydispersity of the system and about the optical properties of the solute particles may also be obtained from such data.

Heretofore, this theory had been developed so that the data could be interpreted only under the limiting condition of low velocity gradient where the degree of orientation of the solute particles is very small. With the aid of the Mark I computer, the necessary equations have been solved to give numerical values over a much wider range of velocity gradients, thus greatly increasing the usefulness of flow-birefringence measurements for the study of macromolecular systems.

REFERENCES

1. Edsall, "Streaming birefringence and its relation to particle size and shape," *Advances in Colloid Sci.* **1,** 269 (1942).

2. Cohn and Edsall, *Proteins, amino acids, and peptides* (Reinhold, New York, 1943), p. 506.

3. Peterlin and Stuart, *Handbuch und Jahrbuch der Chemische Physik*, Bd. 8, Abt. IB (1943).

4. Snellman and Bjornstahl, *Kolloid-Beihefte* **52,** 403 (1941).

5. Signer and Gross, *Z. physik. Chem.* [A.] **165,** 161 (1933).

6. Mehl, *Cold Spring Harbor Symposia in Quantitative Biology*, vol. 6, (1938) p. 218.

7. Lauffer and Stanley, *J. Biol. Chem.* **123,** 507 (1938).

8. Foster and Edsall, *J. Am. Chem. Soc.* **67,** 617 (1945).

9. Edsall, Foster, and Scheinberg, *J. Am. Chem. Soc.* **69,** 273 (1947).

10. Edsall and Foster, *J. Am. Chem. Soc.* **70,** 1860 (1948).

11. Edsall, Gordon, Mehl, Scheinberg, and Mann, *Rev. Sci. Instruments* **15,** 243 (1944).

12. von Muralt and Edsall, *J. Biol. Chem.* **89,** 315, 351 (1930).

13. Edsall and Mehl, *J. Biol. Chem.* **133,** 409 (1940).

14. Sadron, *J. phys. radium* [7] **7,** 263 (1936).

15. de Rosset, *J. Chem. Phys.* **9,** 766 (1941).

16. Lawrence, Needham, and Shen, *J. Gen. Physiol.* **27,** 201 (1944).

17. Above a critical speed of rotation, the flow becomes turbulent (see reference 1), but we are here concerned only with conditions in which the flow is laminar.

18. A rigid ellipsoid of revolution is considered a moderately good approximation to the shape of a protein molecule, but not to the shapes of flexible coiling molecules such as those of many synthetic polymers.

19. Jeffery, *Proc. Roy. Soc.* (*London*) [A] **102,** 161 (1922–23).

20. Boeder, *Z. Physik* **75,** 258 (1932).

21. Peterlin, *Z. Physik* **111,** 232 (1938).

22. Peterlin and Stuart, *Z. Physik* **112,** 1, 129 (1939).

23. Perrin, *J. phys. radium* [7] **5,** 497 (1934).

24. For example, for prolate ellipsoids ($a > b$), rotary Brownian movement of the a-axis about the b-axis is characterized[23] by the rotary diffusion constant Θ_b and a relaxation time τ_a, where

$$\frac{\Theta_b}{\Theta_0} = \frac{\tau_0}{\tau_a} = \frac{\dfrac{3q^2(2-q^2)}{\sqrt{1-q^2}} \ln \dfrac{1+\sqrt{1-q^2}}{q} - 3q^2}{2(1-q^4)},$$

where $q = 1/p = b/a$, and the zero subscript refers to quantities for a sphere of the same volume.

If $a > 5b$, the following approximation is valid within 1 percent:

$$\Theta_b = \frac{1}{2\tau_a} = \frac{3kT}{16\pi\eta a^3}(2 \ln 2p - 1),$$

where η is the viscosity of the solvent.

It is thus easily seen that Θ_b is not a very sensitive function of the axial ratio p as compared to the length of the semimajor axis a. Therefore p can be determined with sufficient accuracy for present purposes from viscosity measurements, and may be taken as a known quantity. A determination of Θ_b from flow-birefringence data thus gives the molecular length $2a$.

25. This quantity is denoted by b by Peterlin and Stuart.

26. The problem of calculations based on double refraction of flow measurements in polydisperse systems will be considered in a later publication.

27. The expression for the distribution function for the steady state may also be given in the equivalent form[4]

$$\Theta \Delta F = \operatorname{div} F\omega,$$

where ω is the angular velocity of the particle and is a function of G and R; Θ is the rotary diffusion constant showing the analogy between the rotational problem and the similarly expressed problem of translational diffusion embodied in Fick's laws. The substitution of Jeffery's expressions[19] for ω and performance of the indicated vector operations leads to Eq. (2).

28. Jahnke and Emde, *Tables of functions* (Dover, New York, 1945), p. 107.

29. The series appearing in Eqs. (13) and (14) can be shown to be equal to

$$\sum_{j=1}^{\infty} R^{j-1} a_{11,j} = \frac{-5f}{16\pi R} \cos 2\chi, \text{ and } \sum_{j=1}^{\infty} R^{j-1} b_{11,j} = \frac{5f}{16\pi R} \sin 2\chi.$$

These terms appear explicitly in the treatment of flow birefringence in polydisperse systems.[26] Tables of these functions would be of significant help for such computations.

30. The change of sign in the equation for f, when p changes to $1/p$, is due to the factor R before the parentheses in Eq. (14).

31. $a_{11,1} = -\dfrac{1}{8\pi} \dfrac{1}{1 + 36/\alpha^2}$; $b_{11,1} = \dfrac{3}{4\pi\alpha} \dfrac{1}{1 + 36/\alpha^2}$.

32. Simha, *J. Phys. Chem.* **44,** 25 (1940).

33. Eirich, *Reports on progress in physics*, vol. 7 (1940), p. 329.

34. Personal communications with Drs. Onsager, Peterlin, and Simha.

35. Kuhn and Kuhn, *Helv. Chim. Acta* **28,** 97 (1945); see especially Eqs. (73) and (74).

36. This distribution function appears to be correct although its convergence has not been established. An indication of the possibility of the convergence was obtained during the computation procedure described in the next section.

37. We should like to express our thanks to Dr. Eric Ellenbogen for aid in computing one of these sets.

L-SHELL INTERNAL CONVERSION

MORRIS E. ROSE

Oak Ridge National Laboratory

While the problem of internal conversion is of considerable interest to nuclear physicists, and may be of some interest in connection with the proceedings of this Symposium as an example of an intricate and imposing calculation, part of which has already been carried out on the Automatic Sequence Relay Calculator (Mark I), I wish to use it merely as a jumping-off point to discuss a more general problem. This more general problem, to which I may give the title "Interaction of electrons and electromagnetic radiation," is one that is ripe, so to speak, for the utilization of modern computing machinery. My principal thesis is that as a by-product of the internal-conversion work we obtain a very important contribution to the problem of numerical solution for the description of other processes which are of prime interest to the physicist. These processes are:

(*a*) Bremstrahlung, or the emission of light by an electron in the neighborhood of an atom.

(*b*) Pair formation, or the transition of an electron from a negative energy state to a positive energy state under the influence of an external electromagnetic field. Again, this process takes place when there is an atom nearby. The result of the transition is to create an electron-positron pair.

(*c*) One-quantum pair annihilation, the reverse of process (*b*).

(*d*) Photoelectric effect, the absorption of light by a bound electron.

(*e*) Compton scattering, the scattering of light by bound electrons.

To this list we may add the internal-conversion process that is to be described. One considers a system of nucleus plus atomic electrons. As a result of a nuclear transmutation the nucleus is often left in an excited state. It can get rid of its excitation energy by one of several mechanisms, of which the most important are: (1) emission of high-frequency light (γ-ray) or (2) transfer of the energy to one of the atomic electrons, say one of those in the K-shell, which is usually the most probable event of this type. This electron is ejected from the atom and appears as a sharp line in an energy spectrum measured with an electron spectrometer. Processes (1) and (2) are alternative modes of decay and from a measurement of the ratio of the rate of process (2) to that of process (1)—that is, the internal-conversion coefficient —one obtains the following vital statistics concerning nuclear structures. While there is a great deal we do not know about a nucleus we do know that in each state, in addition to the energy, the angular momentum of the nucleus is a constant of the motion. We also know that the parity is a constant of the motion; this is a two-valued (even-odd) quantity describing the behavior of nuclear wave functions under space inversion. Now the fact that these two quantities and the energy are constants of the motion is the only nuclear information inserted

into the problem. What comes out of an experimental and theoretical study of the internal-conversion coefficient, together with other data of nuclear spectroscopy, is a quantitative knowledge of all the constants of the motion, energy, angular momentum, and parity for the pertinent nuclear states. There is no doubt in the minds of many physicists that at the moment the most promising approach to an understanding of nuclear structure is through the accumulation of information on the quantum numbers of nuclear states. It is also a well-known fact that completely detailed information about the structure of quantum systems is not always necessary in order to make some useful applications. An example is the role of the quantum mechanics of molecular structure in elucidating the empirical rules of chemical valence. It is to be expected that the study of nuclear spectroscopy will be equally useful.

In order to make clear the thesis as originally stated, it is necessary to say a few words about the details of the internal-conversion calculations. The rate of electron ejection appears as a sum over all possible final states of the electron, selected according to conservation laws of energy, angular momentum, and parity, of squares of matrix elements referring to each final state. These matrix elements, as always, involve certain averages over the configuration space of the electron of an equivalent electromagnetic field corresponding to the nuclear transition. Thus one deals with integrals of the form

$$\int_0^\infty F(r;Z,p,J)\chi_l(kr)f(r;Z,j)dr \tag{1}$$

in which there appear certain physical parameters Z, p, k, J, l, j. Of these one is interested in a particular j, and the conservations laws are such that once j is chosen there are only three free parameters—Z, p, l, for example. The χ_l is a spherical Hankel function of the first kind of half-integral order. The functions F and f will be referred to as wave functions, and, in a relativistic treatment of the problem, they appear as solutions of coupled linear homogeneous differential equations to which certain boundary conditions are applied.[1] These boundary conditions constitute an eigenvalue problem which has a discrete part and a continuous part; F belongs to the continuous part, f to the discrete part.

When the electron is in the K-shell, which is closer to the nucleus than any other shell, it is permissible to neglect the effect of all the other electrons in the atom, and then analytic representations of the wave functions are available. The integrals (1), of which there are seven for each Z, p, l, can be represented in terms of functions that have known properties but are untabulated, namely, hypergeometric functions with complex parameters. The complications involved in the computations for the number of points in Z, p, l space which was required was sufficiently imposing to bring up the question of an alternative procedure, namely, numerical integration of the differential equations for at least the F function, which brings in most of the complication, and then evaluation of matrix elements, typified by (1), by numerical quadrature. For the purpose in hand this procedure turned out to involve a considerably greater number of operations and the first-mentioned procedure, computation of hypergeometric series, was adopted. As mentioned, this work was done at the Computation

Laboratory under Professor H. H. Aiken. In the light of later developments and a broadening of our point of view, it can be said in afterthought that the second procedure of computing wave functions would have been more desirable.

For the L-shell internal conversion it is not at all legitimate to forget the presence of other electrons. They will produce a net field which must be added to the nuclear field and this will modify the motion of the ejected electron in an important way. The modification of the potential field in which the electron moves can be determined and this itself is a problem of no mean proportions. Fortunately, this problem has been solved by J. Reitz, assisted considerably by the ENIAC. The potential field, which we call a screened field, must be inserted in the differential equations, which then determine the wave functions, and, since the field is known only numerically, it is necessary to integrate the wave equation numerically. Other possibilities, such as analytic representation of the potential function, or perturbation or variational techniques, for finding the wave functions turn out to be highly impractical.

I will omit any discussion of the many interesting problems that arise in connection with the numerical solution of the wave equation, except to remark that one of the most difficult parts of this problem is to obtain well-behaved solutions in the discrete spectrum. For this purpose it is well to remember that a highly accurate eigenvalue may give a poor wave function and that some interpolation procedure for applying the boundary conditions is required. This and a number of other problems of methodology have been solved and it is hoped to put the L-shell internal-conversion computations on the Mark III shortly.

In the process of calculating the internal conversion it will be necessary to tabulate $3N$ discrete-spectrum wave-function pairs where N is the number of atoms for which the calculations will be made; N will be about 10. In addition, for radiation fields of angular momenta $l = 1$ through $l = 5$ and for six values of the energy (parameter p or k) and $N = 10$ it will be necessary to tabulate 840 wave-function pairs in the continuum. These will comprise final-state electron waves of all angular momenta J up to 13/2.[2] With these wave functions about 10^4 matrix elements (quadratures) will have to be computed. This in brief is the program for the L-shell internal conversion.

The 840 continuous-spectrum wave functions thus obtained, which are solutions for the relativistic Dirac electron in a screened field, would represent a compendium of the most accurate set of wave functions available. In fact, no wave functions, even without screening, have been available in tabulated form. These same wave functions are involved in the quantitative description of all the processes involving interaction of electrons and light which were mentioned in the preceding. Hitherto all these processes have been calculated by approximate methods involving wave functions describing electrons subject to no atomic or nuclear fields at all. These approximate calculations suffice where high energies of the electrons are involved (say 10 Mev or more) although even here 10- to 15-percent discrepancies between theory and experiment have been observed. At low energies very large errors may be incurred by the use of these approximate wave functions. Thus for pair production at 1.5 Mev in lead the error in the calculated cross section is 100 percent.

The wave functions used in the internal conversion work correspond to a representation in which the angular momentum and one of its components are constants of the motion. For the more general class of interaction problems one needs a representation which corresponds to an outgoing current in a definite direction and this implies a linear combination of angular-momentum wave equations. In order that component functions with $J > 13/2$ shall contribute inappreciably we are restricted to low electron energies (up to about 2 Mev). This restriction could be removed, of course, by a more extended program of computation of wave functions, although even with the restricted number of angular-momentum values some interesting work could be done in connection with all the processes that do not involve pairs. In the case of the pair phenomena there is another difficulty. The description of these processes requires wave functions belonging to the negative-energy continuum. Essentially, these are obtained from the positive-energy continuum by changing the sign of the parameter Z. While this is a more or less trivial change in an analytic representation of the wave functions, it is far from trivial when the wave functions must be obtained by computational methods and are to be given in numerical form. The negative-energy wave functions can be obtained in exactly the same manner as were the positive-energy functions. This again calls for an even more extended program for computing these wave functions.

It seems to me that such a program would be very much worth while and would constitute an invaluable contribution to physics.

NOTES

1. The function F is coupled to another function G, f to a function g. Both G and g appear in other matrix elements of the form (1).

2. The unit of angular momentum is $\hbar$ (Planck's constant divided by 2π); $\hbar = 1.05 \times 10^{-27}$ erg sec.

The Use of Fast Computing Machines in the Theory of Primary Cosmic Radiation

MANUEL S. VALLARTA

*University of Mexico**

Computing machines have been used in the theory of primary cosmic radiation in two cases: (*a*) in the theory of the geomagnetic effects, that is, to solve the equations of motion of a charged particle in the field of a magnetic dipole; (*b*) in connection with problems involved in the theory of the emission of cosmic rays from the sun, that is, to solve the equations of motion of a charged particle in the variable magnetic field of a sunspot, and the equations of motion of a charged particle in the field of two crossed magnetic dipoles.

It is well known from Gauss's analysis of the magnetic field of the earth as observed at the earth's surface that the potential component of this field is the superposition of a dipole and a quadripole field, of which the former is by far the more important. Further, the former varies as the inverse cube of the distance from the dipole, while the latter varies as the inverse fourth power of the distance from the quadripole. It follows that at distances from the earth of the order of magnitude of a few earth's radii the dipole field controls the motion of a charged particle, while the quadripole field plays the role of a small perturbation. Since primary cosmic rays are charged particles coming to the earth from distances large compared with the earth's radius, only the dipole component of the geomagnetic field has to be taken into account.

The equations of motion are readily set up from the classical laws of motion. The force acting on the particle is simply the Lorentz force which, since it always acts perpendicular to the path, does no work and hence the kinetic energy of the particle is a constant of the motion. As a consequence the particle's mass is not its rest mass but its relativistic mass which remains constant throughout the motion.

The equations of motion of a charged particle in the field of a magnetic dipole are integrable in terms of known (elliptic) functions only in the case of motion in the plane perpendicular to the dipole, that is, the plane of the geomagnetic equator. Elsewhere they must be integrated by making use of methods of numerical or mechanical integration. Methods of numerical point-by-point integration have been extensively used by Störmer and his assistants at the University of Oslo, Norway. As the geomagnetic field varies rapidly, particularly in the region close to the earth, the interval of integration must be chosen correspondingly short to reach adequate precision, and this means that a very large amount of labor is required. Depending on the kind of trajectory, a numerical integration by standard methods requires from a day to more than a week, and this circumstance rules out the possibility of solving problems where

* Read at the Symposium by J. C. Street, *Harvard University*.

a large number of complicated trajectories are needed. Hence the necessity of using fast modern computing machines.

When the field has axial symmetry, as in the case of a dipole, a second integral of the motion exists in addition to the kinetic-energy integral already mentioned. This second integral is the projection of the moment of momentum of the particle at infinity, relative to the dipole axis, on this axis. As a consequence the motion can be split up into two motions: (*a*) a motion in the meridian plane, that is, in a plane containing the dipole axis; (*b*) a rotation of the meridian plane. Most of the important problems related to the theory of primary cosmic radiation, in particular the geomagnetic effects, do not require the knowledge of the motion of the meridian plane.

A particle of given energy and angular momentum can move only within certain regions of space known as the Störmer regions. Outside of these regions the kinetic energy would become negative, or, what amounts to the same thing, the time would become imaginary. The shape of these regions is determined from the value of the energy and the angular momentum.

The main result of the analysis of the motion of a charged particle in the field of a magnetic dipole is that all particles of a given energy and sign must arrive at any point on the earth from directions within a certain cone, known as the allowed cone. This cone has its vertex at the observer on the earth and its generators are further described below. The allowed cone is a cone of many sheets: the first sheet determines the so-called main cone, which has the property that any direction within is an allowed direction; the last sheet determines the shadow cone, which has the property that any direction outside is a forbidden direction. Between the main cone and the shadow cone the sheets of the allowed cone determine alternately bands of allowed and bonds of forbidden directions of very complex structure. This has been called the region of penumbra.

The generators of the main cone are trajectories, which are asymptotic to members of the family of unstable periodic orbits that are farthest removed from the earth, and which do not form loops, known as trajectories of the first kind; the generators of the shadow cone are trajectories which do not form loops and are tangent to the earth before reaching the observer. These trajectories are known as trajectories of the second kind. The generators of the penumbra bands are trajectories of either the first or second kinds which form one or more loops before reaching the observer. An asymptotic trajectory may at the same time be tangent to the earth and thus mark a direction along which two sheets of the allowed cone, for instance the main cone and the shadow cone, touch each other. Such trajectories are known as trajectories of the third kind.

The equations of motion of the particle in the meridian plane, in appropriate coördinates, are

$$\frac{d^2x}{d\sigma^2} = \left(\frac{1}{2\gamma_1}\right)^4 e^{2x} - e^{-x} + e^{-2x}\cos^2\lambda,$$

$$\frac{d^2\lambda}{d\sigma^2} = e^{-2x}\sin\lambda\cos\lambda - \frac{\sin\lambda}{\cos^3\lambda},$$

and the equation of motion of the meridian plane is

$$\frac{d\varphi}{d\sigma} = \frac{1}{\cos^2 \lambda} - e^{-x}.$$

No use of this last equation will be made in what follows.

The determination of the main cone requires (*a*) the knowledge of the members of the family of periodic orbits that lie farthest from the dipole; (*b*) the knowledge of the asymptotic orbits to each member of the family of periodic orbits. To find the shadow cone, all trajectories that do not form loops and are tangent to the earth at some point apart from the point of observation must be known. To determine the penumbra bands all trajectories, either asymptotic or tangent, making one or more loops must be available. All these problems depend for their solution on the integration of the equations of motion.

The integration of the equations of motion was carried out by means of the first differential analyzer at the Massachusetts Institute of Technology during the years from 1933 to 1939. The first integrals may be written

$$\frac{dx}{d\sigma} = \int \left(\frac{e^{4x}}{16\gamma_1^4} - e^x + \cos^2 \lambda \right) e^{-2x} d\sigma,$$

$$\frac{d\lambda}{d\sigma} = \int \left(e^{-2x} - \frac{1}{\cos^4 \lambda} \right) \sin \lambda \cos \lambda d\sigma.$$

In this form the system of differential equations may be immediately set up and solved in the differential analyzer. Five input tables are needed to introduce the functions $(e^x/2\gamma_1)^4 - e^x$, e^{-2x}, $\cos^2 \lambda$, $\cos^{-4} \lambda$, and $\sin \lambda \cos \lambda$. Four integrators are required to integrate the product of the two functions under the integral signs and two more to integrate $dx/d\sigma$ and $d\lambda/d\sigma$. The output is plotted with x as abscissa and λ as ordinate.

The initial conditions are introduced in the machine from a knowledge of x and the initial slope. The knowledge of x and λ is sufficient to set the starting points on the input tables. To set the starting points on the six integrators the values of the functions e^{-2x}, $(e^x/2\gamma_1)^4 - e^x$, $\cos^2 \lambda$, $-\cos^{-4} \lambda$, $\sin \lambda \cos \lambda$, $dx/d\sigma$, and $d\lambda/d\sigma$ must be calculated. The first four are computed from the known values of x and λ, the last two from x, λ, and the initial slope.

One way to find the family of periodic orbits is to start the trajectory at right angles to the boundary of the Störmer forbidden region and continue the integration until the point where the trajectory has a tangent parallel to the λ-axis is reached. If the point of "vertical" tangent is on the equator, the required periodic orbit has been found; if it is not, then the starting point is moved along the boundary of the forbidden region until the required trajectory has been discovered. In this way it was found that periodic orbits exist only for a limited range of values of the angular momentum, and for any value of angular momentum within this interval they exist in pairs, of which one is stable and the other unstable.

To find the asymptotic trajectories a point is chosen on the equator between the earth and the outer (unstable) periodic orbit and a trajectory is started in the direction toward this orbit; the initial slope is then adjusted by trial until the trajectory neither falls short of

nor intersects the periodic orbit. Five to ten trials are necessary, and the critical initial slope is determined by this method with a precision of a few thousandths of a radian, as shown by independent calculation.[1, 2]

To determine shadow orbits one may choose a given value of the energy and angular momentum and start trajectories at fixed intervals of a few degrees of latitude, tangent to the earth. These trajectories are then continued until they strike the earth at some other point. Two limits of latitude are determined in this way. The lower limit is characterized by the fact that the orbit through this point is a self-reversing orbit, that is, an orbit that reaches the boundary of the forbidden region of Störmer and reverses along itself. The upper limit corresponds to an orbit that has an inflection at the point of tangency and is therefore a transition orbit between simple orbits of the second kind and orbits having maxima at the point of tangency and minima at points within the earth. All orbits of the latter class are clearly in "shadow" and are not generators of the shadow cone.

The determination of penumbra orbits is very much more complicated but follows in general along the same lines.[4, 5, 6]

The precision of the trajectories determined with the help of the differential analyzer is far from constant and depends on two factors—the number and sharpness of turns and the length of a trajectory. For short runs without sharp turns the precision reached may be of the order of a few thousandths of a radian; for long runs with many sharp turns the error may be as high as half a radian.

Another problem requiring the use of high-speed modern computing machines is the emission of cosmic rays from the sun. A few years ago it was found that the appearance of certain solar flares was followed by a sudden large increase in the intensity of the cosmic radiation as observed everywhere on the earth except at equatorial latitudes. It seems certain that charged particles present in the neighborhood of sunspots can be accelerated up to cosmic-ray energies by the action of the variable magnetic fields of sunspots and that they can escape only when the proper conditions are satisfied between the permanent dipole magnetic moment of the sun and the transient dipole moment of the pair of sunspots associated with the flare. To find out the actual trajectory followed by such charged particles from the sun to the earth requires the integration of the equations of motion in the combined field of the permanent and transient dipoles. This problem has no axial symmetry; consequently the angular-momentum integral is lost and only the kinetic-energy integral remains. In phase space the trajectories are therefore subject to the condition that they must remain on the surface $v_x^2 + v_y^2 + v_z^2 = 1$ (in appropriate units). This condition, translated into configuration space, yields the result that, provided the ratio between the permanent and the transient dipoles and their relative orientation is within certain limits, and provided also the ratio between the field values is above a certain constant, a tunnel is drilled through the Störmer forbidden region of the permanent dipole, and through this tunnel the particles accelerated by the variable sunspot field can then escape. This condition is necessary but not

sufficient; in other words, it is not known in advance whether trajectories exist that start from the sun and come out of the tunnel.

The equations of motion in the field of the two dipoles are

$$\frac{1}{K^2}\frac{d^2x}{ds^2} = \frac{dy}{ds}\left\{k\left[\frac{3(z-z_0)P\cdot\rho}{r^5} - \frac{\cos\gamma}{r^3}\right] + \left[\frac{3(z-z_0)Q\cdot\rho}{R^5} - \frac{\cos c}{R^3}\right]\right\}$$
$$- \frac{dz}{ds}\left\{k\left[\frac{3(y-y_0)P\cdot\rho}{r^5} - \frac{\cos\beta}{r^3}\right] + \left[\frac{3(y-y_0)Q\cdot\rho'}{R^5} - \frac{\cos b}{R^3}\right]\right\},$$

and two other equations obtained by cyclic interchange, where $K^2 = 9M_s/mv$, $k = M_{ss}/M_s$, M_s is the magnetic moment of the permanent dipole, M_{ss} is that of the transient dipole; ρ, P, ρ', Q are the vectors whose components are, respectively,

$(\cos\alpha, \cos\beta, \cos\gamma)$, the direction cosines of M_{ss};

$(x-x_0, y-y_0, z-z_0)$, x_0, y_0, z_0 being the position coördinates of M_{ss};

$(\cos a, \cos b, \cos c)$, the direction cosines of M_s;

$(x-X_0, y-Y_0, z-Z_0)$, X_0, Y_0, Z_0 being the position coördinates of M_s;

$r^2 = (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2$, $R^2 = (x-X_0)^2 + (y-Y_0)^2 + (z-Z_0)^2$;

and the origin of coördinates is at the mouth of the tunnel nearest the sun.

In the region close to the sun the field is large and changes rapidly. A preliminary integration has shown that in this region the trajectory starts out as a tight spiral of a few hundred kilometers radius and a few thousand kilometers pitch. As a consequence the velocity vector changes rapidly but the displacement is small. Further, the interval of integration must be taken very small to keep within the required precision.

The U.S. Army has very kindly made available the ENIAC machine at Aberdeen, Maryland for the purpose of carrying out the integration of the equations of motion. For the reason mentioned above, even this fast machine is unable to carry through the integration in cartesian coördinates without prohibitive labor. In order to circumvent this difficulty we have made use of the fact that both the radius and the pitch of the spiral trajectory are slowly varying functions of the distance along the trajectory, and only the angle turned through is a rapidly varying function. We have therefore introduced helical coördinates defined by the transformation $x = M\cos\theta$, $y = M\sin\theta$, $z = N\theta$, where M, N, and θ are functions of the arc length s measured along the trajectory. The transformed equations are

$$\frac{d^2M}{ds^2}\cos\theta - 2\frac{dM}{ds}\frac{d\theta}{ds}\sin\theta - M\cos\theta\left(\frac{d\theta}{ds}\right)^2 - M\sin\theta\frac{d^2\theta}{ds^2}$$
$$= \left(\frac{dM}{ds}\sin\theta + M\cos\theta\frac{d\theta}{ds}\right)\{\cdots\} - \left(\frac{dN}{ds}\theta + N\frac{d\theta}{ds}\right)\{\cdots\},$$

and two more equations obtained from these as in the previous case. It is hoped to start the integrations with the help of the ENIAC this coming fall. The interval of integration will be a few thousand kilometers at the start, and will be doubled every other step. From fifty to a hundred trajectories will be required, and in all some ten million operations will be needed.

Without taking account of the time required for necessary machine repairs, the time required for the actual integration on the machine will be about a month.

We are indebted to Professor John von Neumann of the Institute for Advanced Study, Princeton, New Jersey for his interest and help with the problem of the emission of cosmic rays from the sun.

REFERENCES

1. G. Lemaitre and M. S. Vallarta, *Phys. Rev.* **49,** 719 (1936).

2. G. Lemaitre and M. S. Vallarta, *Phys. Rev.* **50,** 493 (1936).

3. G. Lemaitre and M. S. Vallarta, *Ann. soc. sci. Bruxelles* **56,** 102 (1936).

4. E. J. Schremp, *Phys. Rev.* **54,** 158 (1938).

5. R. A. Hutner, *Phys. Rev.* **55,** 15 (1939).

6. R. A. Hutner, *Phys. Rev.* **55,** 614 (1939).

7. M. S. Vallarta, *An outline of the theory of the allowed cone of cosmic radiation* (University of Toronto Press, Toronto, 1938).

8. S. E. Forbush, P. S. Gill, and M. S. Vallarta, *Rev. Mod. Phys.* **21,** 44 (1949).

Computational Problems in Nuclear Physics

HERMAN FESHBACH

Massachusetts Institute of Technology

At present the computing machine is employed by theoretical nuclear physicists as a tool of research. This is perhaps a more exciting role than such an instrument commonly plays; on the other hand, there is the concomitant danger that many of the results record an experiment, an attempt to explain the properties of the atomic nucleus and its constituents as following from some initially chosen hypothesis. This is particularly true now when our knowledge of the forces that hold nuclear particles together is vague and fragmentary. It is usual that such an attempt will be a failure and the numerical results obtained will be replaced in the future by ones of greater validity. It is thus of the greatest importance to choose problems that are presumed to be most fruitful in exposing the inner workings of the nucleus.

As a consequence of this view of the type of calculations likely to occur, it is highly desirable that the computing machine be flexible. Thus, it should be possible to make rapid changes in the input of the computer—for example, to change the numerical values of some of the parameters; or one might need to change some of the computational details, for in many problems where the final results cannot be envisaged easily a priori it is difficult to have all such details in order for every choice of parameters. Note also the human element here for someone must judge how the parameters are to be changed. Thus, for some of the very fast computing machines under construction, it seems likely that the nuclear problem cannot be run continuously. Rather it must be possible, particularly for economic reasons, to switch to some other problem while the decision is being made.

Before discussing some of the computational problems that occur in nuclear physics, it is useful to tabulate the types of data which nuclear theory must be expected to correlate and explain. Some notation is necessary: A is the total number of particles in the nucleus, Z is the nuclear charge and therefore the number of protons, $A-Z$ is the number of neutrons. (1) The *mass* of the nucleus may be measured. Its deviation from the sum of the masses of the individual neutrons and protons that make up the nucleus is called the *binding energy*, constituting the first body of nuclear data to be explained. (2) Under the influence of external forces such as those produced by γ-rays, electrons, or nuclear projectiles, the nucleus can gain energy and exist for a short while in an excited state. The energies required to excite each of these various excited states form a second set of data. (3) In comparing two nuclei with the same number of nuclear particles, that is, the same value of A, it is found that one of them is more stable. The optimum value of Z for each A, and the differences in binding energy between the stable nucleus and the unstable nuclei, must be explained. (4) Another set of important data is found by applying electric or magnetic fields to the nucleus. From these

measurements we may obtain the magnetic moment and the electric quadripole moment of the nucleus. A related datum is the total angular momentum of the nucleus about its center of mass.

As an example, consider the case $A = 2$. The most stable system with this value of A is the deuteron, consisting of a neutron and proton. The other systems with this value of A, two neutrons or two protons, are less stable; actually, they are unstable. The deuteron has a binding energy of 2.23 Mev. It has an electric quadripole moment of 2.73×10^{-27} cm.^{2}, and a magnetic moment of 0.8565 nuclear magnetons. The angular momentum is $\hbar$ (Planck's constant divided by 2π). There are no excited states for the deuteron in which the neutron and proton remain bound to each other.

The scattering and absorption of particles by nuclei, together with the resultant transmutations of the target nuclei, provide another important source of data. In this type of experiment, the incident particle—neutron, proton or alpha particle—upon striking the nuclear surface may be reflected, giving rise to scattering, or it may be absorbed. Upon absorption, the resultant nucleus will be unstable emitting then a particle (it can be the same type as the incident particle). The residual nucleus is generally not the same as the target nucleus. A simple example is

$$B^{10} + n \rightarrow Li^7 + \alpha.$$

Deuterons, and lately tritium (H^3), are also employed as projectiles. Deuteron reactions are of a rather different nature inasmuch as the deuteron is a rather loosely bound combination of neutron and proton. When the deuteron approaches the nucleus, the electrostatic field of the nucleus repels the proton and thus polarizes the deuteron, so that whenever the neutron lies between the nuclear surface and the proton, the deuteron is stretched. If the neutron should strike the nucleus with the proton outside, the neutron is absorbed or reflected by the nucleus. In either event, the bond that held neutron and proton together in the deuteron combination is not strong enough to keep them together under this impact, the proton going off independently of the neutron. Upon some occasions, of course, the complete deuteron may strike the nucleus and be absorbed but, at least for small energies, the electrostatic field of the nucleus tends to prevent the proton from reaching the nucleus, making this process relatively improbable.

Considerably more fundamental experiments occur when the elementary nuclear particles—neutrons and protons—scatter from each other as in neutron-proton and proton-proton scattering, for then we are dealing directly with the nuclear forces between particles.

In this field of nuclear reactions and scattering a number of functions occur which should be and indeed are in part tabulated. This circumstance arises from the fact that one factor in describing the probability of an event, for example, absorption by a nucleus, is the probability that the particle will strike the nucleus. Since this depends on the motion of the incident particle while it is outside the nucleus where the forces acting are known, it becomes possible to tabulate this probability for various energies and charges of the incident particle. If the

incident particle is a neutron, then the probability may be stated in terms of the solutions of the equation

$$\frac{d^2u_l}{dr^2} + \left[k^2 - \frac{l(l+1)}{r^2}\right] u_l = 0, \qquad (l \text{ an integer}) \tag{1}$$

subject to the boundary condition $u_l \underset{r\to\infty}{\to} e^{ikr}$. Here $k = \sqrt{2ME/\hbar^2}$, where E is the energy of the incident particle and M is its mass. The required solutions are well known to be kr times the spherical Hankel function $krh_l(kr)$, where $h_l(kr) = \sqrt{\pi/2kr}\, H_{l+\frac{1}{2}}^{(1)}(kr)$. The important physical quantities are the phase and amplitude of this function. These have been tabulated.

If the incident particle is charged (for example, a proton, a triton, or an α-particle), then the solutions of the following differential equation are required:

$$\frac{d^2u_l}{dr^2} + \left[k^2 - \frac{2\eta}{r} - \frac{l(l+1)}{r^2}\right] u_l = 0, \qquad (l \text{ an integer}) \tag{2}$$

where $\eta = MZZ'e^2/\hbar^2$, Z is the charge on the target nucleus, and Z' is the charge on the incident particle.

It is more convenient to use a dimensionless independent variable

$$\rho = kr. \tag{3}$$

Then Eq. (2) becomes

$$\frac{d^2u_l}{d\rho^2} + \left[1 - \frac{2ZZ'e^2}{\hbar v}\frac{1}{\rho} - \frac{l(l+1)}{\rho^2}\right] u_l = 0, \tag{4}$$

where v is the velocity of the incident particle. The solutions of Eq. (4) are known as Coulomb wave functions. The solutions of interest must satisfy boundary conditions similar to those given in Eq. (2):

$$u_l \underset{\rho\to\infty}{\to} e^{i(\rho - \alpha \ln 2\rho)}, \qquad \alpha = \frac{ZZ'e^2}{\hbar v} \tag{5}$$

Eq. (4) may be reduced to the equation for the confluent hypergeometric function. Thus, power-series expansions of this function with an infinite radius of convergence exist as well as expansions in terms of the spherical and cylindrical Bessel functions. Extensive tabulations have been made of the imaginary part of the solution u_0 by the Computation Laboratory in New York City. The higher l values may be obtained by recurrence formulas, although successive application of some will result in loss of accuracy. Note that by making appropriate changes in the parameters α and k the solutions may be utilized in the discussion of proton-proton scattering.

The problem of the motion of a deuteron in the electrostatic field of a nucleus has not yet been solved. In this case one does not expect to obtain exact analytic solutions; rather the attempt is made to reduce the partial differential equation to a form that would be suitable for machine calculation. The equation is

$$\left\{\nabla_n^2 + \nabla_p^2 + \frac{2M}{\hbar^2}\left[E - V(|\mathbf{r}_n - \mathbf{r}_p|) - \frac{2Ze^2}{r_p}\right]\right\} \psi(\mathbf{r}_n, \mathbf{r}_p) = 0. \tag{6}$$

Here ∇_n^2 and ∇_p^2 are the Laplacians in the coördinates $\mathbf{r}_n$ and $\mathbf{r}_p$ respectively. The function

V is the neutron-proton potential. It is appreciable only when the neutron and proton are separated by less than about 2×10^{-13} cm. The function $\psi(\mathbf{r}_n, \mathbf{r}_p)$ should behave as given by Eq. (5) as the center of mass of the deuteron goes to infinity. At the surface of the nucleus, according to recent models of nuclear reactions, the logarithmic radial derivative with respect to r_n and also to r_p must satisfy specified boundary conditions. Actually this statement of the boundary conditions already contains considerable simplifications. It would, however, take us too far afield to discuss these here.

Coulomb wave functions for electrons also are needed in many strategic places. Here we are again dealing with known forces and consequently known equations of motion. For low electron velocities and "bare" nuclei (that is, disregarding the extranuclear electrons) the electronic Coulomb wave functions satisfy Eq. (4) with $Z' = -1$. The effect of the extranuclear electrons, particularly for low velocities, complicates the calculation considerably, as is clear from the earlier discussion by M. E. Rose. However, when the electrons are moving rapidly, it is necessary to employ the Dirac equation which satisfies the requirements of relativity. Again the wave functions satisfy Eq. (4), except that d is no longer an integer. There is the additional complication that in the Dirac case there are two such solutions which must be properly combined to give the final solution. Wave functions of this type would be useful in a large number of problems, all of which are more or less concerned with nuclear structure. Besides the problems of internal conversion (see the paper by M. E. Rose) and internal pair production, there are the problems of electron excitation and disintegration of nuclei, electron production of mesons, and, of great importance in electrodynamics, the production of x-rays, the production of electron-positron pairs, and the scattering of electrons by nuclei, where it is necessary to take into account the distribution of charge within the nucleus itself.

We have now exhausted the catalogue of functions whose tabulation would be of permanent value, particularly in nuclear physics. It should be emphasized, however, that these functions are not descriptions in any way of the nuclei but rather are tools by means of which the analyst may extract the salient features of such a description from the experimental data. For example, in the scattering and absorption of particles by nuclei the energy at which resonance occurs and the width of the resonance may, by means of the functions discussed above, be translated into the value of the logarithmic derivative at the surface of the nucleus at or near the resonance. This is a property of the interior of the nucleus. The problem of understanding these facts in terms of a theory of nuclear structure remains.

Let us now turn to the problem of determining nuclear structure itself. Here we attempt the calculation of such properties as binding energy, the energies of the excited states and the associated widths, stability questions, electromagnetic properties of nuclei, relative yields in nuclear reactions, and so on. The general plan of action consists in utilizing the properties of the simpler nuclear systems to test and finally choose the law of force between nucleons. Then this law of force is to be employed to predict the properties of more complicated systems. Is it possible for such a program to succeed? It is not altogether clear that it is; for example, the recent discovery of "shell structure" in nuclei indicates at least the possibility that the

heavier nuclei are in some ways less complex than the very light nuclei. Or it may be that the notion fundamental to this program, that there are laws of force which depend only upon the coördinates of the nucleons, may be incorrect. We begin to see that some calculations may prove to possess only an ephemeral value.

Of course we cannot discuss this entire program in detail here. It will suffice to point out the mathematical questions involved and some suggested methods of solution. It will be seen that much machine computation is involved, barring some revolutionary discovery that would succeed in reducing the present complication to simplicity much as the Copernican theory of the motion of the planets reduced the older Ptolemaic theory.

The mathematical problem to be solved is that of the Schroedinger equation:

$$\left[\frac{-\hbar^2}{2M}\sum_i \nabla_i^2 + \sum_{\substack{i,j \\ i>j}} V_{ij}\right]\psi = E\psi, \tag{7}$$

where ψ is a function of all the coördinates of each particle; the subscript i denotes the particle involved, so that V_{ij} is the potential energy between the ith and jth particles. The word coördinate as employed here includes not only the space coördinates but the charge (1 or 0) and spin (intrinsic angular momentum of each nucleon) coördinates as well. The energy of the system is denoted by E; it has its lowest value for the ground state of the system where its value is the negative of the binding energy. The excited states of the system all have larger energies. The outstanding characteristic of these forces is their short range and their consequent rapid variation as the distance between the particles changes. These features of the nuclear forces make calculations difficult to perform.

Equation (7) can be reduced to a system of linear second-order *ordinary* differential equations only for the two-particle nuclear systems, for which $A = 2$. These equations may be integrated by either numerical or analytic methods if possible. However, even for the two-body case it is often more economical to adopt approximate procedures which lead to the desired results more rapidly and easily. For all other nuclear systems, three-body and more, approximate methods must be employed.

There are three such methods which have been employed in the past and upon which we may expect to rely in the future. These are (1) the perturbation method, (2) the Rayleigh-Ritz method, and (3) the variational-iterational method.

In the first of these, the solution ψ is expanded in terms of the eigenfunctions of some approximate problem. We rewrite Eq. (7) symbolically in the form

$$H\psi = E\psi,$$

where H is an operator. Suppose that

$$H = H^0 + H',$$

and that the approximate problem with eigenfunctions χ_n and energies E_n is

$$H^0\chi_n = E_n\chi_n.$$

The energy E is then the solution of a secular determinant

$$|H'_{nm} - (E - E_n)N_{nm}| = 0, \tag{8}$$

where

$$N_{nm} = (\chi_n, \chi_n)\delta_{nm}, \quad H_{nm}' = (\chi_n, H'\chi_m).$$

We are employing Hilbert space notation; δ_{nm} is the Kronecker δ. Approximate formulas which assume that the difference between H and H^0 is small have been developed[2] and are customarily employed rather than the full secular determinant. Calculations of this type have been made for the lighter nuclei, employing as the unperturbed wave functions the harmonic-oscillator wave functions[3] (Hermite functions) for each particle, assuming particle independence. For the heavy nuclei, plane-wave approximations for each particle have been assumed.[4] Perturbation methods have also been employed in the resonating-group method where the nucleus is presumed to exist for a time in certain subgroups; it is decided a priori which groups are most likely. This essentially provides a scheme wherein it becomes possible to base the properties of a nucleus A on those of $A - 1$. Because of the length of the calculations, the perturbation method has never been pushed far enough to obtain convergence; usually it has stopped early with but a few terms permitting a qualitative understanding of some nuclear properties, and on the other hand leading to some very grave misconceptions. It may not be the most appropriate method but certainly it is one that is readily adapted to machine calculation.

The Rayleigh-Ritz method enjoyed a very great success in the theory of the atom and consequently has been employed in nuclear problems. The method as it is customarily employed consists of two parts. First (note that one may use either the independent particle picture or the resonating-group method) one assumes a form for the function ψ involving nonlinear parameters. These nonlinear parameters are then determined by the variational principle. However, as is well known, this procedure yields only an upper bound to the eigenvalue. The second step is an attempt to determine the eigenvalue itself. The initial choice for ψ is made the first term in an infinite series of functions with linear undetermined coefficients, the functions involved forming a complete set for the problem under discussion. Introducing this series into the variational principle yields a secular determinant similar to Eq. (8). By considering the successive values of E as the number of terms in the series is increased, one may estimate the convergence and thus the final value of the eigenvalue E. Unfortunately, this procedure is not foolproof,[5] for sometimes the convergence obtained may be false. This is caused in part by the faulty choice of the type of unperturbed problem. However, a considerable fraction of the difficulty lies in the rapid variation of nuclear potentials with interparticle distance, implying the need to employ a considerable number of eigenfunctions with fairly large quantum number. Hence the lack of convergence.

In the discussion of both the perturbation and the variational methods we have concentrated on the calculation of binding energies and the energy levels of nuclei. However, it should be noted that both of these methods apply as well to scattering and nuclear-reaction problems. Their application to these problems has been made for only the very light nuclei.

The variation-iteration method adds to the variation method (*a*) a systematic method of improving the initial trial function and (*b*) a method of obtaining a lower bound which,

combined with the upper bound given by the Rayleigh principle, determines the required eigenvalue to a certain accuracy. The method has been applied successfully by several authors. Convergence is rapid and security in the results is available because of the existence of a lower bound.

The problems under discussion may all be written in the form

$$A\psi = \lambda B\psi, \tag{9}$$

where A and B are Hermitian operators and λ is an eigenvalue. For example, in Eq. (7), $A = -(\hbar^2/2M)\sum_i \nabla_i^2 + \sum_{\substack{i,j \\ i>j}} V_{ij}$, $B = 1$, $\lambda = E$. There is another possibility, however. Let $V_{ij} = qf_{ij}$, where f_{ij} is just the form of the dependence of the internucleon potential energy, and q is the measure of the strength of that potential. The number q may also be considered as the eigenvalue λ, $A = E + (\hbar^2/2M)\sum_i \nabla_i^2$, $B = \sum_{\substack{i,j \\ i>j}} f_{ij}$. In this formulation E is assumed to be known, and the necessary strength of potential needed to obtain this value of E is computed. At the present stage in the history of nuclear physics, this is actually a more convenient order, for we are now interested in determining what V_{ij} will yield the known experimental binding energies. It is this formulation which has been employed in the calculations that so far have been made with this method.

The technique goes as follows. By some means or other, either by the Rayleigh principle or by knowing a reasonably good approximation to the correct eigenfunction, an initial wave function φ_0 is chosen. The iteration method is employed to improve the initial trial function. It is important to employ in the iteration an operator that essentially involves an integration rather than a differentiation. In the problem under consideration, therefore, the successive iterates φ_n are generated as follows:

$$\begin{aligned} \varphi_1 &= A^{-1}B\varphi_0, \\ \varphi_2 &= A^{-1}B\varphi_1, \\ &\cdot \quad\quad \cdot \\ &\cdot \quad\quad \cdot \\ &\cdot \quad\quad \cdot \\ \varphi_{n+1} &= A^{-1}B\varphi_n. \end{aligned} \tag{10}$$

It is easy to see how the successive applications of $A^{-1}B$ improve φ_0. Since the solutions of Eq. (9) form a complete orthonormal set $\{\psi_n\}$ we may expand φ_0 in terms of the set:

$$\varphi_0 = \sum_0^\infty a_p\psi_p.$$

Then

$$\varphi_n = \sum_{p=0}^{\infty} \frac{a_p}{\lambda_p{}^n}\psi_p,$$

where λ_p is the eigenvalue associated with ψ_p. Inasmuch as there is an eigenvalue in the set λ_p, say λ_0, which has the lowest absolute value, $\varphi_n \underset{n\to\infty}{\to}$ (constant) ψ_0.

From these successive iterates it is possible to form successive approximations to the eigenvalue by employing the iterates as trial functions in the two variational principles:

$$\lambda_0 = \text{stat. value of } \frac{(\psi, A\psi)}{(\psi, B\psi)} \tag{11}$$

and

$$\lambda_0 = \text{stat. value of } \frac{(\psi, B\psi)}{(\psi, BA^{-1}B\psi)}. \tag{12}$$

Introduce into these expressions $\varphi_1, \varphi_2, \ldots$ for ψ. The resultant values of the ratio are

$$\begin{aligned}
\lambda_0^{(1)} &= \frac{(1,0)}{(1,1)}, \\
\lambda_0^{3/2} &= \frac{(1,1)}{(1,2)}, \\
\lambda_0^{(n-1/2)} &= \frac{(n-1, n-1)}{(n,n-1)}, \\
\lambda_0^{(n)} &= \frac{(n,n-1)}{(n,n)}, \\
\lambda_0^{(n+1/2)} &= \frac{(n,n)}{(n,n+1)},
\end{aligned} \tag{13}$$

where $(n, m) = (\varphi_n, B\varphi_m) = (\varphi_n, A\varphi_{m+1})$. We now tabulate some theorems with regard to the quantities $\lambda_0^{(n)}$ and $\lambda_0^{(n+1/2)}$. Two cases are to be distinguished. The inequalities are special cases of a more general inequality which may be readily found.

Case 1: A and B are positive definite operators.

(*a*) $\left.\begin{matrix}\lambda_0^{(n)} \\ \lambda_0^{(n+1/2)}\end{matrix}\right\} \underset{n\to\infty}{\to} \lambda_0$;

(*b*) $\lambda_0^{(n-1/2)} \geqslant \lambda_0^{(n)} \geqslant \lambda_0^{(n+1/2)} \geqslant \lambda_0$;

(*c*) $\lambda_0 \geqslant \lambda_0^{(n+1)}\left(1 - \dfrac{\lambda_0^{(n+1/2)} - \lambda_0^{(n+1)}}{\lambda_1 - \lambda_0^{(n+1)}}\right)$,

$\lambda_0 \geqslant \lambda_0^{(n+1/2)}\left(1 - \dfrac{\lambda_0^{n} - \lambda_0^{n+1/2}}{\lambda_1 - \lambda_0^{n+1/2}}\right)$;

(*d*) $\min \dfrac{A\psi}{B\psi} \leq \lambda_0 \leq \max \dfrac{A\psi}{B\psi}$;

(*e*) the error decreases in the ratio λ_0/λ_1 in going from $\lambda_0^{(n}$ to $\lambda_0^{(n+1/2)}$.

Case 2: B is positive definite; A is not positive definite.

(*a*) $\lambda_0^{(n-1/2)} \geqslant \lambda_0$,

$\lambda_0^{(n-q-1/2)} \geqslant \lambda_0^{(n+p)}$; $p \geqslant 0$, $q \geqslant 0$

(*b*) $\lambda_0^{(n+1/2)} \geqslant \lambda_0^{(n+3/2)} \cdots \geqslant \lambda_0$ if $\lambda_1^2 \geqslant \lambda_0^{(n+1)}\lambda_0^{(n+3/2)}$;

(c) $\lambda_0^{(n-1/2)}\lambda_0^{(n)} \geqslant \lambda_0^{(n+1/2)}\lambda_0^{(n+1)} \geqslant \cdots \lambda_0^2$,

$$\lambda_0^2 \geqslant \lambda_0^{(n+1)}\left(1 - \frac{\lambda_0^{(n+1/2)} - \lambda_0^{(n+3/2)}}{|\lambda_1| - \lambda_0^{(n+3/2)}}\right),$$

$$\lambda_0^2 \geqslant \lambda_0^{(n+3/2)}\lambda_0^{(n+2)}\left(1 - \frac{\lambda_0^{n+1/2}\lambda_0^{n+1} - \lambda_0^{(n+3/2)}\lambda_0^{n+2}}{\lambda_1^2 - \lambda_0^{n+3/2}\lambda_0^{n+2}}\right),$$

$$\lambda_0 \geqslant \lambda_0^{(n+1)}\left(1 - \frac{\lambda_0^{(n+1/2)} - \lambda_0^{(n+1)}}{|\lambda_2| - \lambda_0^{(n+1)}}\right) \text{ if } \lambda_1 \leq 0,$$

where the problem has been adjusted so that λ_0 is > 0. There is one substantial difference between the two cases that should be mentioned here. When A and B are positive definite, the successive approximations to λ_0 approach λ_0 monotonically. This is not true in the other case discussed.

Here λ_1 is the next eigenvalue above λ_0 in absolute value, λ_2 the one above that. It is generally necessary to have a lower-bound estimate of λ_1. This may be obtained in several ways. In one we employ the relation

$$\text{Spur } (A^{-1}B)^2 = \sum_p (1/\lambda_p^2). \tag{14}$$

Hence

$$\lambda_1^2 \geqslant \text{Spur } [(A^{-1}B)^2 - (1/\lambda_0^{(n+1)})^2]^{-1}. \tag{15}$$

The variation-iteration method may be improved by several simple methods, of which we shall give two here. One involves using the functions φ_n generated by the iteration as base functions for the Ritz method discussed earlier. This leads to a secular determinant whose elements may be expressed in terms of $\lambda_0^{(n)}$ and $\lambda_0^{(n+1/2)}$ and have therefore been already computed. The secular determinant is

$$\begin{vmatrix} \vdots & \vdots & \vdots \\ \cdots 1 - \dfrac{\lambda_0}{\lambda_0^{(n-1)}} & \dfrac{1}{\lambda_0^{(n-1)}}\left(1 - \dfrac{\lambda_0}{\lambda_0^{(n-1/2)}}\right) & \dfrac{1}{\lambda_0^{(n-1)}\lambda_0^{(n-1/2)}}\left(1 - \dfrac{\lambda_0}{\lambda_0^{(n)}}\right)\cdots \\ \cdots 1 - \dfrac{\lambda_0}{\lambda_0^{(n-1/2)}} & \dfrac{1}{\lambda_0^{(n-1/2)}}\left(1 - \dfrac{\lambda_0}{\lambda_0^{(n)}}\right) & \dfrac{1}{\lambda_0^{(n-1/2)}\lambda_0^{(n)}}\left(1 - \dfrac{\lambda_0}{\lambda_0^{(n+1/2)}}\right)\cdots \\ \cdots 1 - \dfrac{\lambda_0}{\lambda_0^{(n)}} & \dfrac{1}{\lambda_0^{(n)}}\left(1 - \dfrac{\lambda_0}{\lambda_0^{(n+1/2)}}\right) & \dfrac{1}{\lambda_0^{(n)}\lambda_0^{(n+1/2)}}\left(1 - \dfrac{\lambda_0}{\lambda_0^{(n+1)}}\right) \\ \vdots & \vdots & \vdots \end{vmatrix}. \tag{16}$$

From the solution of this equation one obtains an upper bound not only to λ_0 but also to λ_1 and λ_2, etc., depending upon the size of the determinant. Employing an upper bound for λ_1, it is possible from the spur in Eq. (14) to obtain a lower bound for λ_2 required in one of

the inequalities above. Finally, it also becomes possible to give another set of lower bounds based on the new approximations to ψ_0 obtained by solving the secular determinant.

The variation-iteration method may also be combined with the relaxation method to yield another procedure for automatically improving the initial trial function. These results are equivalent to those incorporated in Eq. (16), except that the quantities involved occur in a somewhat different order.

Applications of this method to problems in nuclear physics have been made by Thomas[6] and Svartholm.[7] In recent months a series of extensive calculations on the properties of the deuteron have been made in which the method was utilized with great success in a rather difficult problem. This case was, as a matter of fact, rather interesting, for it involved a non-positive definite operator, with the consequence that the eigenvalues extend from $-\infty$ to $+\infty$. The convergence of the method in this case depends upon the ratio $|\lambda_0/\lambda_1|$ and in some cases λ_1 was very close to $-\lambda_0$, corresponding to a degeneracy in the iterated eigenvalue, so that convergence would be slow if the method was applied without modification. One may improve the convergence by employing the secular determinant (16) after the first two iterations, or as it turned out it was easy to derive an expression which extrapolates to the final answer and which is particularly applicable to the nearly degenerate case.

The variation-iteration method has only been employed for the lightest nuclei, $A \leq 4$. The problem of extending this type of calculation, or indeed any of the others, to heavier nuclei lies in the large number of coördinates involved and the consequent large number of multiple integrals of many dimensions that would be required. Probably some approximate technique such as that given by the Monte Carlo method would be necessary. In any event, it would seem foolhardy to extend the calculations much above $A = 4$, in view of the present uncertainty in nuclear forces and the imminent possibility that some simplifying notions in the physics may turn up in the near future.

In conclusion, we would like to compare the ease with which computing machines could be utilized in each of the three methods mentioned. It is rather clear from the outset that the variation-iteration scheme is much more easily adapted to machine computation than either the perturbation or the variation method. This is primarily because of the repetitive nature of the operations involved, which simplifies considerably the number of directions—the number of stored functions—that need to be fed into the input side of the device. When we combine this considerable advantage with those already mentioned, it seems to be not too risky to predict the increasing use of the variation-iteration method in nuclear problems.

REFERENCES

1. The real and imaginary parts of the spherical Hankel functions have been tabulated by the Mathematical Tables Project, *Tables of spherical Bessel functions* (Columbia University Press), vols. 1 and 2. The required phases and amplitudes have been computed by Morse, Lowan, Feshbach, and Lax in a report issued by NDRC, Division 6, entitled "Scattering and Radiation from Circular Cylinders and Spheres."

2. For the most recent version of these perturbation formulas see E. Feenberg, *Phys. Rev.* **74,** 206 (1948); H. Feshbach, *Phys. Rev.* **74,** 1548 (1948).

3. H. Margenau and W. A. Tyrrell, Jr., *Phys. Rev.* **54,** 422 (1938); H. Margenau and D. T. Warren, *Phys. Rev.* **52,** 790 (1937); H. Margenau and H. Carroll, *Phys. Rev.* **54,** 705 (1938); H. Margenau, *Phys. Rev.* **55,** 1173 (1939); D. T. Warren and H. Margenau, *Phys. Rev.* **52,** 1027 (1937).

4. H. Euler, *Z. Physik* **105,** 353 (1937); S. Watanak, *Z. Physik* **113,** 482 (1939).

5. D. T. Warren and H. Margenau, *Phys. Rev.* **52,** 1027 (1937).

6. L. H. Thomas, *Phys. Rev.* **47,** 903 (1935).

7. N. Svartholm, *Thesis* (Hagon Ohlssons Boktryckeri, Lund, 1945).

SIXTH SESSION

Thursday, September 15, 1949
2:00 P.M. to 5:00 P.M.

AERONAUTICS AND APPLIED MECHANICS

Presiding

Harald M. Westergaard
Harvard University

Computing Machines in Aeronautical Research

R. D. O'NEAL

University of Michigan

I shall attempt to outline the possible application of computing machines, both digital and analog, to some of the principal fields of aeronautical research. Although this symposium is mainly concerned with high-speed digital machines, a discussion of the application of computing machines to aeronautical research could not be complete without considering analog machines, because they have already proved themselves extremely useful. The fields of research that I shall consider and the order in which I shall consider them are as follows:

1. Over-all flight-path problems for aircraft,
2. Stability of aircraft in flight,
3. Airflow studies to determine aerodynamic coefficients,
4. Structural analysis of aircraft,
5. Dynamic simulation of aircraft,
6. Traffic handling.

Although the use of computers in the last field will likely be more as a part of a system rather than for aeronautical research, I shall want to discuss it briefly because I believe it is one of the problems which most requires high-speed digital computing machines.

Let us first consider the general flight-path problem for an aircraft. I shall use the term aircraft to mean both airplanes and guided missiles. The equations of motion can easily be derived from considering the forces acting on the craft. If we neglect external forces such as that of wind, which may or may not be small depending upon the ratio of the velocity of the aircraft to that of the wind, the Coriolis force, which is a small effect even for very fast supersonic aircraft, and variations in the acceleration due to gravity, which is also definitely a small effect, then the equations of motion for flight in a plane are (see Fig. 1):

$$M\ddot{x} = T \sin \varphi - C_D S\rho(\dot{x}^2 + \dot{y}^2) \sin \theta - C_L S\rho(\dot{x}^2 + \dot{y}^2) \cos \theta, \tag{1}$$

$$M\ddot{y} = T \cos \varphi - C_D S\rho(\dot{x}^2 + \dot{y}^2) \cos \theta + C_L S\rho(\dot{x}^2 + \dot{y}^2) \sin \theta - Mg, \tag{2}$$

where T is the thrust exerted on the aircraft by whatever propulsive system is used to drive it, which will, in general, vary with altitude y; C_D is the drag coefficient, which is a function of the Mach number and angle of attack, becoming quite high in the transonic range; C_L is the lift coefficient, which is also a function of the Mach number and angle of attack; S is a characteristic cross section of the aircraft; ρ is the air density; θ is the angle that the velocity vector makes with the vertical; φ is the angle that the thrust vector makes with the vertical; M is the mass of the aircraft; g is the acceleration due to gravity.

Since no analytic solution has been found for the above set of nonlinear differential equations, it is necessary to use numerical or analog methods to obtain the flight path. Obviously, hand methods of numerical solution can be very long and laborious for some flight paths. The automatic digital machine can be extremely useful in solving these problems. The Mark I, or a similar machine, can handle the problem quite well. The use of computing machines should make parametric studies of a preliminary aircraft design much easier and more economical. For instance, the effects of various climb programs, of various thrusts, and of other parameters can be studied by varying each parameter within reasonable limits and studying the effects of these variations on the over-all flight path. These parametric studies can be carried out with analog equipment, providing a high degree of accuracy is not required and

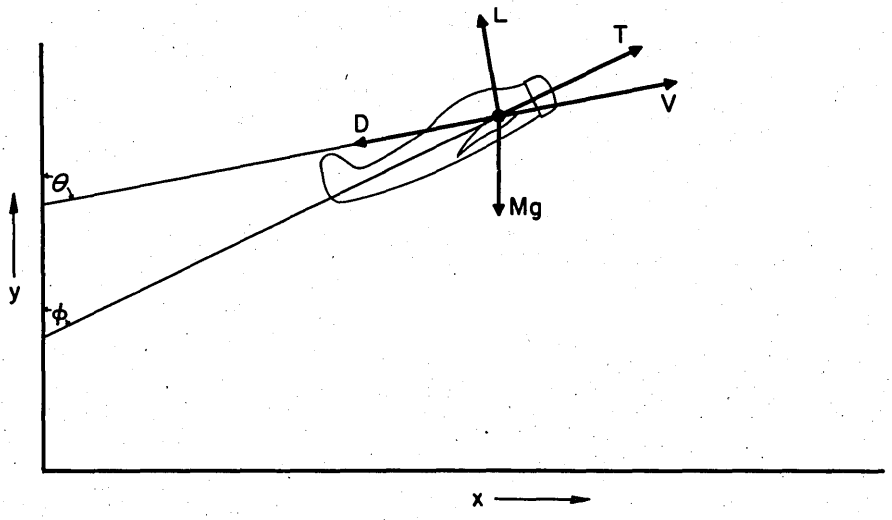


FIG. 1. Forces on airplane during flight in a plane.

providing suitable equipment is used. Special function equipment is required for representing variation of C_D and C_L with Mach number, variation of air density with altitude, and possibly variation of thrust with altitude. However, if high accuracy is required, digital machines must be used.

In the foregoing discussion of flight-path problems, a stable aircraft was assumed. Some of the major problems in aeronautical research are those involving stability. The problems of stability and control are closely related. For instance, in order to keep the control forces small, the static stability should be low. In fact, highly maneuverable planes such as fighters may actually be statically unstable. The problems of stability and control may initially be studied separately, particularly in case the aircraft is statically stable, and later tied together when the over-all aircraft system including pilot (either a human pilot or an autopilot) is studied. I shall not discuss the stability problem in detail because it is discussed fully by E. T. Welmers in the next paper. However, I would like to say that for some stability problems,

such as the linearized pitch-plane stability problem, electronic analog computing equipment can be used. If one makes the assumptions that deviations from given flight conditions are small, that acceleration has negligible effect upon the aerodynamic parameters, and that the coefficients are constant for the duration of the analysis, then the stability equations are ordinary linear differential equations that can be solved easily by electronic analog computing equipment. Parametric studies to indicate the effects of variations over wide limits of each of the parameters can be made fairly easily, and these are very useful in design. The accuracies achieved with analog equipment are usually sufficient. It may be that even problems such as these, which can be adequately handled by analog equipment, will be handled by digital computing centers when such centers are more plentiful and when the programming time can be reduced. Whether the problems will continue to be solved by analog equipment or whether most of them will be solved by the digital centers will probably be determined mainly by economic factors.

In order to make studies of the performance and flight characteristics of an aircraft, we have already seen that it is necessary to know certain aerodynamic coefficients, such as the drag and lift coefficients. The solutions of the airflow problems are aimed at furnishing these coefficients. In the past the computations for airflow problems have been largely performed with the aid of standard desk machines and at a considerable expenditure of time and effort. Digital computers can and will be very useful in solving airflow problems.

Let us consider, as an example of an airflow problem, the partial differential equations describing in cartesian coördinates the flow about a body for an axially symmetric case:

$$\left(1-\frac{u^2}{a^2}\right)\frac{\partial u}{\partial x}+\left(1-\frac{v^2}{a^2}\right)\frac{\partial v}{\partial y}-\frac{uv}{a^2}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)+\frac{v}{y}=0, \tag{3}$$

$$\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}=0. \tag{4}$$

In these equations, x and y are the rectangular coördinates of a point in a fixed meridian plane of the body, the direction of the x-axis being along the axis of symmetry; u and v are x- and y-components of the velocity of the air relative to the body; and a denotes the local speed of sound. Of course, the appropriate boundary conditions must be satisfied for any particular configuration being considered. These boundary conditions are (*a*) that the component of air velocity normal to the aerodynamic body be zero and (*b*) that there shall be conservation of mass, conservation of energy, and conservation of momentum across the shock.

Eq. (3) is hyperbolic for supersonic flow and elliptic for subsonic flow. Eq. (4) is the condition for irrotational flow. Aerodynamicists have usually used one of two methods for solving these equations for the condition of supersonic flow: (*a*) the method of linearization; or (*b*) the method of characteristics, by which the foregoing system of partial differential equations is reduced to an equivalent but simpler system of ordinary differential equations. The method of linearization is not exact enough for many cases but often the results obtained with the

method of characteristics have been no more accurate, because the labor of getting many points in the step-by-step numerical solution, required when the method of characteristics is used, has been too great, and so insufficient points were obtained to assure good accuracy. However, if high-speed digital computing machines are used, this difficulty will be overcome and the greater exactness of the method of characteristics can be realized. Indeed, at the meeting of the Association for Computing Machinery, held at the Ballistic Research Laboratory at Aberdeen in December 1947, the computation of the airflow about a cone-cylinder for the axially symmetric case and for irrotational flow, that is, the flow governed by the equations above, was discussed. At that time it was planned that a solution to this problem would be carried out on the ENIAC, and I believe that this has been done. The general three-dimensional problem for supersonic flow has not yet been worked out, although Ferri has recently published a technical note on "The method of characteristics for the determination of supersonic flow over bodies of revolution at small angles of attack."[1]

For airflow studies, large-scale digital computing machines will probably be most useful in two general types of problems: (*a*) those involving a large set of similar problems such as the work done at the M.I.T. Center of Analysis on tables of supersonic flow about cones,[2, 3] and (*b*) as a research tool in helping to obtain a better understanding of the nature of supersonic flow, particularly in studying the effects of viscosity and interference effects in aircraft. It must be emphasized that the previous discussions of airflow have been for the nonviscous case. Actually, the effects of viscosity may be rather large in some cases. For this reason exact solutions of the above expressions would certainly not replace the wind tunnel for obtaining aerodynamic coefficients of lift and drag. Rather, the high-speed digital computer can be a valuable tool in conjunction with the wind tunnel for basic research. The effects of various parameters can be more easily isolated by a method of numerical experimentation with a digital machine than by physical experimentation with the wind tunnel. The aerodynamicists with whom I have discussed these problems believe that the digital computing machine will be a valuable tool not only in working out the theory of visco-compressible flow, but also in studying the interference effects in supersonic flow.

It is axiomatic that aircraft must be built structurally strong, but still as light as possible. For this reason a large amount of effort has gone into structural analysis and the stress-analysis problem on a modern aircraft is an extremely long and time-consuming task. The method which is principally used is the "unit method" described first by F. R. Shanley and F. P. Cozzone.[4] Their paper presented not only improvements in the methods of analysis for determining axial and shear stresses in box beams, but also a tabular method which permitted a considerable saving of time and which can be adapted to machine methods. This method consists essentially of dividing the beam structure, for instance, an aircraft wing, into a number of parts, taking cross sections normal to the length or longitudinal axis of the plane and further subdividing the structure by longitudinal planes. To facilitate computations, flange material is assumed to be concentrated into effective units that coincide with these spanwise divisions.

The analysis begins at the outer or free end of the structure and works stepwise in to the fixed end, each step using the results of the previous step. Only simple algebraic expressions have to be solved for each step but the total amount of computation is large. IBM equipment is being used to very good advantage by aircraft companies in this analysis. In fact, before the IBM equipment was used, I understand that it was often difficult to do a complete structural analysis for an airplane, especially taking into account all of the various loadings that might occur for different flight conditions. It is quite an advantage to be able to do so, because weak spots can then be found before planes are built and so a considerable saving in cost can be effected.

Let us consider briefly the problem of simulating an aircraft system—that is, the aircraft, and the pilot or autopilot that exerts the control on the aircraft. Such a system is a closed servo-loop and may be treated as such. We may consider as an example the control of an airplane in elevation or pitch by means of an autopilot. This case is considered in a report[5] by Hagelbarger, Howe, and Howe. First the equations of motion of (1) the airplane, (2) the autopilot, and (3) the elevator are determined. Then each of these equations of motion is

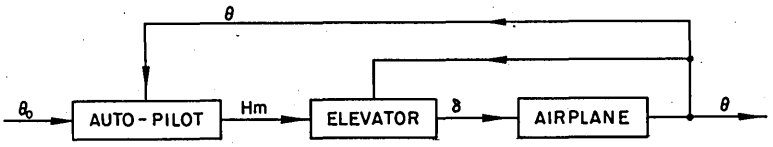


FIG. 2. Block diagram of aircraft system.

simulated by means of analog computers. Finally, the three components are tied together so as to represent the complete system. This system is shown in block-diagram form in Fig. 2.

The equation of motion about the center of gravity for the airplane is

$$\frac{1}{M_\delta}\ddot{\theta} + \frac{C_1}{M_\delta}\theta - \frac{C_2}{M_\delta}\theta = \delta - z_\omega \int \delta dt, \tag{5}$$

where θ is the angle of pitch, or the angle that the airplane makes with the horizontal; δ is the angle of the elevator with respect to the stabilizer, M_δ, C_1, C_2, and z_ω are constants, ω being the forcing frequency applied to the elevator. This single equation does not, of course, completely represent the motion of the airplane, but is illustrative of the type of equation that is used.

The constants used in the work mentioned above were taken from a report covering the steady-state response of a B–25J airplane to sinusoidal oscillation. Eq. (5) is derived on the assumption that the angles δ and θ are small so that the forward velocity of the airplane remains constant.

When the calculated steady-state response of the airplane was compared with points obtained from the computer, there was agreement to within the limits of error of the recorder used in conjunction with the computer.

A circuit was designed that would give about the same gain and phase characteristics

as for a B–24 autopilot amplifier, whose response curves were known, except for a frequency ratio of approximately two, since it was assumed that the frequencies for a B–25 would be, for the same response, about twice those for a B–24.

It was assumed that the equation of motion of the elevator is of the form

$$I\frac{d^2\delta}{dt^2} + C\frac{d\delta}{dt} + K\delta = T_\delta(t), \tag{6}$$

where $T_\delta(t)$ is the net torque applied to the elevator, I is the moment of inertia of the elevator, C is the aerodynamic damping coefficient, and K is the aerodynamic restoring torque for a unit deflection of δ.

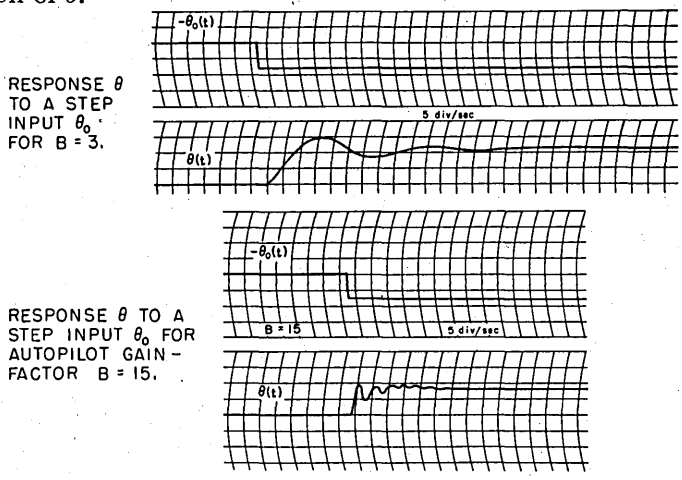


FIG. 3. Effect of change in gain on stability.

It was further assumed that

$$T_\delta = K_1 H_m + K_2\theta,$$

in which H_m is the hinge moment applied to the elevator and is equal to Be_0, where e_0 is the output of the autopilot circuit and B is the autopilot gain factor.

Now these three elements were tied together to form the complete aircraft system. Again checks were made of the steady-state frequency response to compare the calculated curves with the curves determined by the computer. Again the agreement was good, as would be expected. The resonant frequency of the system was measured, and the degree of stability was studied as a function of the autopilot gain factor by using a step input signal. The effect of this change in gain can be seen in Fig. 3. It is to be noted that for the higher gain, θ follows θ_0 much more closely and the static error is almost zero.

With such a simulator it is quite easy to make parametric studies to find how changes in various parameters affect the over-all stability. Various arbitrary disturbances can be put into the system and the response of the system to these disturbances studied. A human pilot might be substituted for the autopilot provided the forces that he would undergo in flight could be put upon him, and so his reactions to various design changes could be studied.

Although the analog computing equipment is limited in accuracy, there seem to be many simulation problems that can be handled adequately by it. Again, as in my comments on stability, I should like to say that digital computing machines may be used for these simulation problems as high-speed digital computing centers become plentiful and provided it is economically advantageous.

I believe that one of the most important uses of high-speed digital computing machines in aeronautics of the future may well be in the handling of air traffic. As the speed of airplanes increases, as the number of flights operating into and out of each major airport increases by a large factor, and as all-weather flying becomes a reality, the present handling systems, which depend upon human reaction times, will be inadequate. It is reasonable to believe that the vastly shorter reaction time of a high-speed digital computer will be required to take the positional information on each plane in the neighborhood of an airport and perform the necessary computations to determine where it should fit into a complicated and rapidly moving landing pattern. Airplanes with very different flying and holding speeds (holding speeds of the airplanes already vary by about a factor of two) and landing characteristics will have to be handled and so a fairly large number of rules and airplane-performance data will have to be in the machine in order that decisions can be made automatically by the machine, such as the altitude at which the airplane will enter the landing pattern, the speed at which it shall fly and rate of descent, how close it will be allowed to come to other planes, and the turning program to be followed. I do not mean to imply that the system of landing aircraft would necessarily be entirely automatic. Actually, there would still be a pilot in the airplane, but he would be receiving his instructions from a high speed digital computer instead of from a human controller. A standby human controller would have to be available for emergencies.

Such an automatic system for handling aircraft probably cannot be realized for several years to come, not only because of lack of high-speed computing machinery, but also because terminal equipment is not available. By terminal equipment I mean the devices for converting the positional information, obtained from radars or other devices, into digital data that can be handled by a computer, and for performing the reverse function of converting the digital commands into intelligence that can be used by a pilot.

I have not discussed the use of computers for such problems as the reduction of flight data, or for the preparation of design tables such as Professor Aiken's work done at the Harvard Computation Laboratory entitled "Tables for the design of missiles,"[6] but I believe that I have discussed enough examples of the use of computing machines in aeronautical research to indicate clearly that the aeronautical industries and the aeronautical research centers have been using available computing equipment as it is developed. I believe that nearly every aeronautical research center and industry now has analog computing equipment which has either been purchased or built. Many of them have IBM installations, and I believe that nearly all the large digital computing machines that are now working have already solved problems in aeronautical research. I am very certain that the field of aeronautics will continue

to use new computing aids as quickly as they are available. Problems such as the traffic-handling problem may well tax the handling capacity of even the fastest and biggest machines now contemplated.

REFERENCES

1. A. Ferri, NACA Technical Note No. 1809 (February 1949).

2. M.I.T. center of Analysis, Technical Report No. 1, *Tables of supersonic flow around cones*; work performed under the direction of Z. Kopal under NOrd contract No. 9169.

3. M.I.T. Center of Analysis Technical Report, *Tables of supersonic flow around yawing cones.*

4. F. R. Shanley and F. P. Cozzone, "Unit method of beam analysis," *J. Aeronaut. Sci.* **8,** 246–255 (1941).

5. D. W. Hagelberger, C. E. Howe, and R. M. Howe, "Investigation of the utility of an electronic analog computer in engineering problems," University of Michigan Report, UMM–28, Project MX–794, USAF Contract W33–038–ac–14222 (April 1, 1949).

6. Staff of the Computation Laboratory, *Tables for the design of missiles* (Harvard University Press, Cambridge, 1948).

Problems of Aircraft Dynamics

EVERETT T. WELMERS

Bell Aircraft Corporation

The upsurge of interest in computational devices and techniques, so well exemplified in the Harvard Computation Laboratory, has been welcomed in all fields of aeronautics, but perhaps nowhere so warmly as in the field of dynamics. The analysis of the flutter or stability characteristics of a modern airplane involves so much computational work that in many aircraft companies an equality sign is understood to exist between the words "flutter calculations."

A detailed discussion here of the derivation of the equations for the problems of dynamics would be too lengthy and actually unnecessary. However, the relation of these problems to modern computing methods may be made more clear if some indications are given of their sources. Three major problems will be considered, namely, flutter, aerodynamic stability, and servomechanisms. The order in which they are discussed is chosen only to permit certain comments about their interdependence. Although generalizations concerning standard methods of solving these problems cannot be made, the flutter problem has usually been solved by digital computation, while the problems of servomechanisms have frequently been studied by analog methods.

The experience obtained from continued analysis and flight testing has established certain broad policies for the guaranteeing of stability. To avoid excessive weight or configuration penalties, to allow for unconventional designs or speeds, and to permit consideration of the aircraft as a whole, all these require a careful analysis. The dangers involved in testing an airplane, especially for flutter, are such that the time spent on analytical work can usually be justified. The result has been that a rather complete theory exists for idealized problems and, in many instances, experimental verification has been sought for a developed theory, rather than theories devised for the explanation of the physical phenomenon.

It must be emphasized that the problem statements and methods of solution discussed here are not unique. In particular, adaptation to modern digital techniques may suggest more convenient methods of attack.

Flutter can be described as a self-induced oscillation involving aerodynamic, inertial, and elastic forces; at least two degrees of freedom are usually required, for example, wing bending and torsion, or wing bending-torsion-aileron rotation. Above a critical velocity (or in a certain velocity range) any slight disturbance of the airfoil results in an oscillation of increasing amplitude, frequently sufficient to cause structural failure; below this critical velocity (or outside the range) such a disturbance causes a damped oscillation. Symmetric and unsymmetric motions are usually considered separately.

The determination of the inertial and elastic forces involved in the problem is relatively straightforward. First attempts at determining the aerodynamic forces, in which they were taken to be proportional to the instantaneous position of the airfoil, were not satisfactory. The gradual development of a theory that included the oscillation frequency and the phase relations culminated in a complete solution for the forces on an oscillating airfoil in incompressible flow by Theodorsen,[1] Kussner,[2] Cicala,[3] and Kassner and Fingado[4] about 1935. The parameter on which these forces depend is $b\omega/v$, usually called the reduced frequency, where v is the forward velocity of the airfoil, b the semichord, and ω the frequency of oscillation. In extending oscillating-airfoil theory to compressible fluids, at each Mach number aerodynamic forces must be determined for various values of the reduced frequency.

The mathematical structure of the problem can be illustrated by the simple example of wing bending-torsion flutter, and more involved cases can be discussed in terms of the notation used there. Applying the differential equations for vibrating beams to the weight, inertia, and stiffness distributions of the actual airfoil, or using an influence coefficient method or a Rayleigh-Ritz variational method, the fundamental bending and torsion frequencies ω_h and ω_α, and normalized mode shapes $h(x)$ and $\alpha(x)$, can be determined. The mode shapes $h(x)$ and $\alpha(x)$ are used as generalized coördinates q_i. Because the airfoil properties are not simple mathematical functions, the solution for q_i is usually a step-by-step digital or a matrix process; however, it is not an obvious problem for very large-scale digital computers, since no family of solutions is required for a given structure nor is the same coding likely to be convenient for different structures.

In matrix form, the differential equations of motion for a strip of unit width can be written

$$(A_{\ddot{q}} + I_{\ddot{q}})\ddot{q} + A_{\dot{q}}\,\dot{q} + (A_q + E_q)q = 0, \tag{1}$$

where the matrices have the following interpretations: q is the column matrix of generalized coördinates; dots denote differentiation with respect to time; the A's are square matrices of aerodynamic terms whose subscripts indicate association with acceleration, velocity, or displacement in the generalized coördinates; the elements depend on the reduced frequency and the geometry of the airfoil, and will usually be complex to include phase differences; $I_{\ddot{q}}$ is a square matrix of inertia terms; the main diagonal involves weight or moment-of-inertia terms; the nondiagonal (or coupling) elements involve unbalances or products of inertia; E is a diagonal matrix of elastic terms, usually expressed as frequencies in the fundamental modes. If a three-dimensional theory is considered, integrations over the span of the airfoil are necessary. Eq. (1) is not changed in form, but the matrices A and $I_{\ddot{q}}$ then involve the assumed deflection mode shapes.

Using a method standardized by the Air Forces,[5] consider the amount of structural damping required to maintain simple harmonic motion. To do this, let $q = q_0 e^{i\omega t}$ and add another term to the coefficient of q, namely iG_q. For convenience, damping coefficients in all modes are assumed equal, and thus

$$q = q_0 e^{i\omega t},\ iG_q = igI;\ \text{last term of Eq. (1) becomes } (A_q + E_q + igI)q. \tag{2}$$

If $g > 0$, damping must be added to give a steady oscillation, implying instability under actual circumstances. Substituting Eqs. (2) into Eq. (1) and combining the aerodynamic matrices gives

$$(A + I_q + E\Omega)q_0 = 0, \tag{3}$$

where Ω is a complex eigenvalue,

$$\Omega = \left(\frac{\omega_\alpha}{\omega}\right)^2 (1 + ig), \tag{4}$$

involving a known reference frequency ω_α, and the required unknowns, frequency ω and damping g. Vanishing of the determinant of the matrix $(A + I_q + E\Omega)$ is required for a solution and determines values of Ω. Substitution of an Ω_i allows solution for $n - 1$ of the q_0 elements in terms of the other.

Change in the reduced frequency $b\omega/v$ will be reflected only in the matrix A. For an assumed value of reduced frequency, the eigenvalues Ω_i and the associated eigenvectors q_{0i}

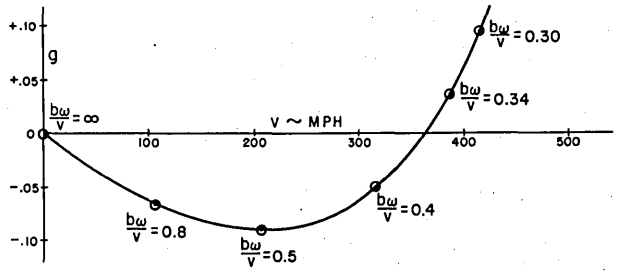


FIG. 1. Damping g and velocity v for various values of $b\omega/v$.

can be determined. The status of stability is indicated by the sign of the imaginary part of Ω_i, a positive sign denoting instability. Substituting the known ω_α into the computed complex eigenvalue, the frequency ω and the damping g can both be determined. For the reduced frequency chosen, substitution of the computed frequency ω and the semichord b, gives

$$\frac{b\omega}{v} = k, \quad v = \left(\frac{1}{k}\right)b\omega. \tag{5}$$

A graph (Fig. 1) of the damping g and associated velocity v for a range of values of $b\omega/v$ is sufficient to determine the critical flutter velocity.

The determination of complex eigenvalues and eigenvectors in a problem of this type is an excellent example of digital calculation; the expansion of the 2×2 determinant, the solution of the resulting quadratic, and the determination of relative amplitudes (or the eigenvector) in this illustration is only a small problem for a desk calculator. I would like to indicate directions in which the problem expands sufficiently to make its consideration at this symposium justified. Complications of three types are obvious—enlargement of the matrices by introduction of more degrees of freedom, modifications of the aerodynamic matrix A, and modification of the structural matrices I and E. These will be considered in inverse order.

Two possible reasons for modification of the matrices I and E can be proposed. First, it

is not always possible to estimate with the desired accuracy the elements of the matrices. Uncertainties as to root fixities, or even manufacturing tolerances in a wing attachment, will cause a frequency variation; it is almost impossible to calculate the frequency of control-surface rotation considering the linkages and supports involved; elastic axes of structures are difficult to determine and, for structures with discontinuities, somewhat meaningless; effective weights and inertias for concentrated masses are frequently difficult to evaluate. As a result, it is often desirable to solve the flutter problem not only for the best estimate of values for elements of E and I, but also for variations of these which may lead to more critical conditions. If such conditions appear, changes can be made in the final design or in fabrication that will reduce the probability of their occurrence.

Rather than stemming from ignorance like the first, the other reason arises from the variety of configurations under which the aircraft may fly, some of which will change its flutter characteristics. Wing-tip external fuel tanks may be full, empty, or may have been jettisoned; rockets or bombs may be loaded under the wings; different control-surface actuation methods may be used. These modifications may also change the aerodynamic matrix, as in the case of external wing-tip tanks extending far ahead of the wing. Thus a family of solutions may be required for a single airplane.

Modification of the aerodynamic matrix A can be likewise justified on the basis of ignorance; this has frequently been done in studying the flutter of tabs attached to movable control surfaces. However, the chief cause for modification is to introduce compressibility or Mach-number effects. The aerodynamic forces determined by Theodorsen[1] were for an incompressible fluid, or $M = 0$; corrections to the $M = 0$ flutter speed based on the Glauert factor $(1 - M^2)^{-\frac{1}{2}}$ were used for high subsonic speeds.[5, 6] More exactly, the actual aerodynamic forces in compressible flow have been determined by Dietze[7] and others, for values of M below the critical value at which shock waves form. Above a Mach number of about 0.8, no theory exists until definitely supersonic speeds are reached. The papers of Garrick and Rubinow,[8] Temple and Jahn,[9] and others, based on the fundamental work of Possio,[10] permit a consideration of the aerodynamic forces from $M > 1$ to hypersonic speeds.

One method of analyzing flutter in a compressible fluid is to construct a stability graph similar to Fig. 1 for each value of M. Only one velocity on each graph will correspond to the Mach number for which the graph was drawn. The locus of these points gives the complete stability graph (Fig. 2). A gap will exist in the locus from approximately $0.8 < M < 1.2$.

The extension to more degrees of freedom perhaps is most quickly suggested by the existence of high-speed digital calculators. Hand calculation passes rapidly from the difficult to the inefficient to the impossible beyond four degrees of freedom.

Recent analytic and experimental work has demonstrated the importance of including the rigid-body degrees of freedom, vertical translation and pitch in symmetric motion, and roll for unsymmetric motion. Since there are no elastic restraints in these modes, the square matrix E is of higher order than rank; thus, for symmetric motion the number of eigenvalues

Ω_i will be two less than the number of degrees of freedom, and one less for unsymmetric motions.

The inclusion of additional portions of the airplane in the analysis is a second source of additional degrees of freedom. To the simple wing bending-torsion problem can be added aileron rotation and tab rotation. Or the bending-torsion of a stabilizer becomes a minor in the problem of fuselage vertical bending, stabilizer bending, stabilizer torsion, elevator rotation, tab rotation, and the rigid-body motions. After noting how a vibrator on the wing can excite the most remote parts of the airplane when at the correct frequency, it becomes rather questionable to consider a wing-flutter analysis as separate from a tail-flutter analysis. A unified analysis of a complete airplane may be unnecessarily involved, even for our best computers, particularly if it is possible to reorder the degrees of freedom in such a way that

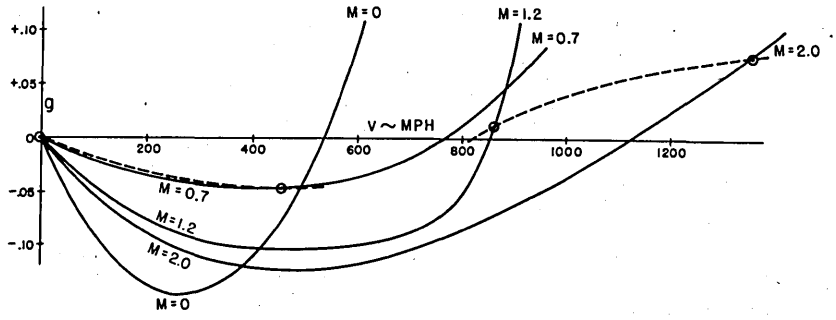


FIG. 2. Stability graph for a compressible fluid.

elements of the matrices far off the main diagonal tend to vanish. There may be some advantage in coding a problem for perhaps ten degrees of freedom and forcing all flutter calculations into that pattern.

In connection with an analysis of flutter in swept wings, Spielberg, Fettis, and Toney[11] have adopted the following method of introducing the boundary conditions at the wing roots. The various bending and torsion modes and frequencies are computed for the actual wing of the airplane cantilevered without sweep at its root. Several of the lower bending and torsion modes are used as generalized coördinates; the boundary conditions inherent in sweep are introduced into the equations, and the eigenvalue problem is solved as before. Unless the wing is complicated by concentrated masses, two or three bending and one or two torsion modes are usually sufficient, making five to seven degrees of freedom for a simple bending-torsion symmetric flutter including rigid-body motions. The same method can be used in the solution of the flutter problem for unswept wings.

Another desirable increase in the number of degrees of freedom is suggested by the preceding. A possible way of determining the vibration modes of a nonuniform beam or wing is to assume modes for a uniform beam that satisfy the boundary conditions and combine them, by variational principles, into the modes of the actual wing. Similarly, modes for a uniform beam can be considered as generalized coördinates in the flutter equation. Accuracy of the same

order as that obtained in the previous method will likely require two to four more degrees of freedom, a very costly penalty to pay without high-speed digital computers. Today, however, it is not unlikely that uniform deflection modes for all problems will sufficiently simplify the preparation of the problem for a computer to justify the added computation.

The problem of stability of an aircraft differs somewhat from that of flutter in several respects. Whereas we have been interested in determining a critical speed for flutter instability, we are here primarily interested in the frequency and damping of oscillations at specific velocities. The aircraft is considered to be a rigid body and only rigid-body motions are studied. Finally, the air forces involved are those of steady-state aerodynamics, as contrasted with forces for oscillating airfoils studied in flutter. These differences, however, are more traditional than essential. In fact, they are rapidly breaking down.

The basic problem may be described as the determination of the dynamic response of an aircraft due to the introduction of forces or moments, either externally or by control-surface deflections. The forces to be considered are inertial and aerodynamic. Since quasi-steady aerodynamics is assumed, there will be no phase lags and the aerodynamic terms will be real. The complete system of equations of motion consists of six simultaneous differential equations if control surfaces are assumed fixed, and nine such equations if the three control surfaces are assumed free, that is, themselves movable under the action of aerodynamic and inertial forces.

In many instances subsections of the whole problem are of interest; for illustration, the equations for motions of an aircraft in its plane of symmetry with free controls will be written.[12] The dependent variables are $\Delta V/V_0$, the velocity change divided by original velocity; $\Delta\gamma$, the change in flight path angle; $\Delta\theta$, the change in angle of pitch. Unprimed terms of the form a_{ij} involve aerodynamic coefficients and perhaps certain initial conditions; primed terms involve both aerodynamic and inertial forces. The differential operator d/dt is represented by D; the $f_i(t)$, frequently step functions, serve to introduce the disturbance instigating the dynamic response. We have

$$\begin{bmatrix} D + a_{11} & a_{12} & a_{13} \\ a_{21} & D + a_{22} & a_{23} \\ a_{31}' & a_{32}' & D^2 + a_{33}'D + a_{33}'' \end{bmatrix} \cdot \begin{bmatrix} \frac{\Delta V}{V_0} \\ \Delta\gamma \\ \Delta\theta \end{bmatrix} = \begin{bmatrix} a_{14} + f_1(t) \\ a_{24} + f_2(t) \\ a_{34} + f_3(t) \end{bmatrix}. \tag{6}$$

Terms a_{i1} which multiply the velocity variable are dependent on the Mach number M, and thus solutions must be found for various values of M. As in the case of oscillating air forces, the transonic region, $0.8 < M < 1.2$, remains questionable.

Since only real elements in the matrices are involved, classical criteria such as Routh's discriminant can be used for the indication of stable or unstable solutions. Laplace-transform methods are ideally suited for determining the actual analytic solution. Analog computers can usually be applied directly, and are especially useful for surveys involving numerous parameter changes. Solution by a digital computer requires establishment of a sufficient

number of cycles to permit determination of the frequency and damping. The degree to which digital computers can furnish a solution to the more involved problems of this nature, for example, those requiring nonlinear aerodynamic terms, is primarily a function of coding difficulties. Standardization to a few basic and inclusive types may be the best method of attack. Any efforts being made to narrow the gap between digital and analog machines will be particularly useful here.

The traditional assumptions on the stability equations are no longer always valid. Aeroelastic effects introduced by the relatively flexible airplanes of today frequently dominate the stability problem. In the case of a missile, this may only require introducing one additional degree of freedom—fuselage bending ; the consideration of aeroelastic stability may require doubling a fuselage skin thickness that was satisfactory in all other respects. For an entire airplane, such as a swept-wing bomber, the complete problem has probably never been stated mathematically. The availability of large calculators bears directly on the interest shown and effort expended on such a problem.

It should not be assumed that the only difficulties in the problem are calculational. The aeronautical engineer has considerable information concerning aerodynamics "in the large," total lifts and moments for rigid members. The effects of local deflections, or aerodynamics "in the small," are on a much less satisfactory basis. Even though the equation coefficients are not sufficiently well determined to permit a reliable solution, variations of the coefficients may lead to certain relatively invariant properties which may be of interest. The attempt to obtain results with inaccurate data frequently points out the important errors, or comparison of the unreliable calculated results with experiment may lead to better estimates of the coefficients.

Another of the assumptions associated with stability—that aerodynamic forces depended only on the instantaneous position of the airfoil—is now being questioned. German research during the war indicated that, under certain circumstances, the use of the oscillating air forces from flutter theory gave different results, even for the long-period phugoid oscillations. Although the oscillations in the Theodorsen theory[1] are assumed to be harmonic, other unsteady air-force theories can be used for highly damped motions.[13] A proper evaluation of the conditions under which unsteady theories are of importance in stability must await additional research activity.

Various electrical and mechanical analogs to inertia, damping, and elastic forces are well known. The basic components of servosystems lend themselves to the same analogs. As a result, the appearance of systems of equations similar to Eqs. (3) and (6) in studying the internal dynamics of servos is to be expected. The use of digital computers in solving problems relating to a servosystem itself is not immediately likely. The similarity between the components of analog computers and servomechanisms is so pronounced that analog methods seem to be simpler. Also, it is possible to combine servo units with analog computers to test systems without knowing all the analytic details of the servo components.

Rather than illustrate the servomechanism aspect by considering its internal stability, it may be more informative to mention the stability or dynamic equations for a servo-controlled airframe. Again the system lends itself to analog solution in many cases; availability, coding simplicity, and unification are factors that will influence the use of digital machines. Mention was made that stick-free stability involved additional equations, one for each control surface, which described the inertia and aerodynamic forces on it. The control surface angle is here determined by autopilot intelligence and is applied by a servomotor, rather than by inertial and aerodynamic forces. Two additional equations are required for each control surface.

Let E indicate the autopilot output voltage and $\Delta\delta$ the change in elevator angle. Only motions in the plane of symmetry of the aircraft are considered. In one instance, Eq. (6) was augmented in the following way:

$$\begin{bmatrix} D+a_{11} & a_{12} & a_{13} & 0 & a_{15} \\ a_{21} & D+a_{22} & a_{23} & 0 & a_{25} \\ a_{31}' & a_{32}' & D^2+a_{33}'D+a_{33}'' & 0 & a_{35}' \\ 0 & 0 & D^2+K_1D+K_2 & D^2+K_3D+K_4 & 0 \\ 0 & 0 & 0 & K_5 & D+K_6 \end{bmatrix} \cdot \begin{bmatrix} \frac{\Delta V}{V_0} \\ \Delta\gamma \\ \Delta\theta \\ E \\ \Delta\delta \end{bmatrix} = \begin{bmatrix} a_{14}+f_1(t) \\ a_{24}+f_2(t) \\ a_{34}+f_3(t) \\ a_{44}+f_4(t) \\ a_{54}+f_5(t) \end{bmatrix}, \quad (7)$$

where the K_i are autopilot and servo constants. In many instances rapid changes of altitude or of Mach number will require variation of the coefficients a_{ij}.

The general comments of the preceding section can be repeated for this case. In one respect the problem here is more serious; the desirable high natural frequencies of a servo-system are closer to the frequencies associated with flutter than to stability oscillation frequencies. The flutter matrix will be augmented in much the same way as Eq. (7). The effect of a high-frequency control-surface oscillation on the autopilot intelligence is difficult to evaluate. The addition of only one degree of freedom involving servo natural frequency may present a sufficiently accurate picture. Servo flutter has attracted some analytic attention recently,[14] but the importance of the problem cannot now be properly evaluated.

If the "general analysis" philosophy of E. H. Moore is applied to the dynamic problems being discussed, the mere association of the word "stability" with these three cases implies the existence of a unified theory. This is being realized in practice, partly because efficiency requires that methods of attack on similar problems not be contradictory, and partly because a satisfactory airplane design depends on their interrelation. Realizing that sheer bulk of the problem prevented any logical unified attack, the dynamicist has been unable to influence design to the proper extent. The jumps in the order of magnitude of possible problems that have been brought about in the last few years now will permit him to attempt the unification and contribute more directly to the design.

Perhaps one way of describing this unified dynamic problem is to consider frequency responses. If the response of the complete airplane is known at all frequencies from zero to

beyond the highest structural natural frequency present, all the problems individually discussed will be solved. The response of the airplane to slow control-surface deflections will be found near zero frequency; abrupt deflections can be analyzed by considering various frequency components. Aerodynamic stability oscillations, with periods from a minute down to a second, will supply the next peaks in the frequency spectrum. Flutter will contribute resonances with frequencies as high as 40 to 60 c/sec.

Although the flutter problem discussed here is shown as a system of simultaneous algebraic equations [Eq. (3)], and the stability and servo aspects as simultaneous differential equations [Eqs. (6) and (7)], the unity still exists. The differences in statements reflect traditional ways of handling problems previously considered independent. In comparing calculation with flight test, it is desirable to know solutions of the basic flutter equation [Eq. (1)] rather than to introduce harmonic motions. The particular trends in solution methods for the unified dynamic problem will depend on the computing machines used; thus the problem and its method of solution are related.

As has been indicated, the problems of flutter have been solved digitally from the beginning, while problems involving servos have been frequently treated by analog devices. Hand calculators have been replaced by punch-card methods as the number of degrees of flutter freedom has increased. In several instances as many as eight degrees of freedom are consistently studied with IBM equipment. In some cases, direct expansion is carried out; in others, iterative methods are used. Within the last few months a flutter problem involving five degrees of freedom and four eigenvalues has been solved on the Mark I; in the solution, the 5×5 determinant with complex elements was expanded directly and the resulting quartic in Ω solved by approximation methods. The equations listed for stability have been used on various digital machines in the calculation of flight paths for missiles and general dynamic-response problems of aircraft.

Two factors influence the use of digital computers in aircraft dynamics. The first is primarily educational. The type of problem considered, the degree of complication, and the interpretation of results should all be influenced by the tremendously increased calculational capacity. "Shotgun" methods are possible, that is, a variety of problems can be solved which surround the somewhat uncertain location of the actual. Research activity should tend to fill in the gaps of a unified theory. Only by a proper realization of the power of methods of solution now available can worthwhile problems be proposed.

The second factor relates to coding difficulties. Many problems which could be done quickly on large-scale digital machines require a large amount of coding and analysis time for trivial machine time. This immediately tends to discourage attempts at machine solution and allows tedious and inaccurate hand calculation to compete in efficiency. I have attempted to indicate in this paper that numerous parameter changes require repeated solutions of similar problems; also that it is possible to consider a large, inclusive problem as the basis, and extract from it pertinent sections. Thus, if a flutter problem involving ten degrees of freedom and eight eigenvalues were already coded, it would be possible to force a large variety of flutter

problems into this form; in some cases many of the matrix elements would be zero, but use of the same code would still be justified by the resulting standardization. Similar standardization of the other dynamic problems is possible. And as efficient high-speed computers become more available, a single coding tape for the complete dynamic problem may be possible.

In conclusion, the problems of aircraft dynamics do have much in common. The increasing complexity of aircraft and missiles has required extension of various aspects of dynamic problems beyond the possibilities of hand calculators. Utilization of large-scale computing machines permits this necessary extension and also allows, for the first time, a study of the coupling between these problems, which have previously been separated for simplicity in analysis.

REFERENCES

1. T. Theodorsen, "General theory of aerodynamic instability and the mechanism of flutter," NACA Report No. 496 (1935).

2. H. G. Küssner, "Zusammenfassender Bericht über den Instationären Auftrieb von Flügeln," *Luftfahrt-Forschung* **13,** 410–424 (1936).

3. P. Cicala, "Le azione aerodinamiche sui profili di ala oscillanti in presenza di corrente uniforme," *Mem. reale accad. sci. Torino* [**2, pt. I**] **68,** 73–98 (1934–35).

4. R. Kassner and H. Fingado, "Das ebene Problem der Flugelschwingung," *Luftfahrt-Forschung* **13,** 374–387 (1936).

5. B. Smilg and L. Wasserman, "Application of three-dimensional flutter theory to aircraft structures," Air Corps Tech. Report No. 4798 (9 July 1942).

6. T. Theodorsen and I. E. Garrick, "Mechanism of flutter," NACA Report No. 685 (1940).

7. F. Dietz, "Die Luftkräfte des harmonisch Schwingenden Flügels im kompressiblen Medium bei Unterschall-geschwendigkeit," *DVL* (Jan. 1943); AAF Translation F–TS–506–RE (Nov. 1946).

8. I. E. Garrick and S. I. Rubinow, "Flutter and oscillating air-force calculations for an airfoil in a two-dimensional supersonic flow," NACA Report No. 846 (1946).

9. G. Temple and H. A. Jahn, "Flutter at supersonic speeds," R.A.E. Report No. S.M.E. 3314 (April 1945).

10. C. Possio, "L'azione aerodinamica sul profilo oscillante alle velocità ultrasonora," *Pontif. Acad. Sci. Acta* **1,** 93–105 (No. 11, 1937).

11. Spielberg, Fettis, and Toney, "Methods for calculating the flutter and vibration characteristics of swept wings," Air Materiel Command Eng. Div. Memorandum Report MCREXA5–4595–8–4 (3 Aug. 1948).

12. W. L. Mitchell, "Dynamics of aircraft flight," Bell Aircraft Report 02–981–006 (1948).

13. H. Wagner, "Über die Entstehung des Dynamischen Auftriebes von Tragflugeln," *Z.A.M.M.* **5,** 17–35 (1925).

14. J. Winson, "The flutter of servo-controlled aircraft," *J. Aeronaut. Sci.* **16,** No. 7 (July 1949).

A Statistical Method for Certain Nonlinear Dynamical Systems

GEORGE R. STIBITZ

Consultant in Applied Mathematics, Burlington, Vermont

I would like to describe a way of treating certain nonlinear dynamical systems that reverses the usual trend of expertness in scientific study. Everybody knows that the trend of expertness is in the direction of learning more and more about less and less until one knows everything about nothing. I do not want to suggest carrying the reverse process to a limit, but I do want to propose, quite seriously, a method that gives us no information at all about the behavior of a dynamical system in any particular cases, but tells us quite a bit about its behavior in a lot of cases. In other words, the results are statistical.

This paper is divided into three parts; it discusses first why statistical results may be useful; second, the theory of a statistical method that has been found useful in some problems; and third, the application of the method, particularly showing how automatic computers will be needed in this application.

The parameters of the physical systems that we treat in applied mathematics are never known exactly. Sometimes this ignorance is the kind we are born with and sometimes it is acquired. In other words, there are problems in which parameters are not known exactly because it is impossible to make perfect measurements, and there are problems in which we do not want to bother about fixing their values. For instance, when we set manufacturing tolerances, we are in effect saying that we could measure certain dimensions quite accurately, but for convenience we prefer to give them some latitude.

Sometimes small errors or variations are unimportant, and sometimes they are very important. For instance, if I calculate my time of departure for the railroad station and adjust my speed of travel aiming to get to the station just as the train is about to leave, a small variation one way or the other determines whether I catch the train or miss it. I find, of course, that the time of my arrival at my final destination depends very critically on minute variations in the values of the parameters I choose. In this dynamical system there is a discontinuity in my response to the initial values, and in the thermal energy dissipated at the station.

On the other hand, if I were to drive all the way to my final destination, then small variations in the speed and time of starting would not be so important.

Now, suppose I were commuting daily. Then the data that are most important to me are not what will happen on any particular trip, but the number of times per year that I miss the train. I could make a statistical study of this factor as a function of my probable error in estimating speed, of the time I use in eating another piece of toast, and so on.

The subject of commuting has been treated rather fully by Streeter and Williams, so I will leave that subject to them, and turn to a somewhat more dignified problem. The problem that started the present study of a statistical method actually has many features in common with that of commuting, but arose during the war when we tried to make a device called the dynamic tester. The dynamic tester included a servo that made 60 attempts every second to catch one of two trains in opposite directions. We did not care whether it caught the right train on any particular trip, provided it did not get too far off schedule on the average.

Figure 1 is a rough schematic diagram of the servo. The tester is intended to put some equipment through its paces at normal speeds, following a precalculated course. A motor M drives the equipment being tested; every sixtieth of a second a control device C reads from a punched tape what the motor position ought to be, and compares this position with the actual position of the motor at that instant. If the motor is lagging, the control device fires a thyratron tube and kicks the motor forward, while if it is leading the required position, the control kicks it backward. Between kicks, the motor coasts at practically constant speed.

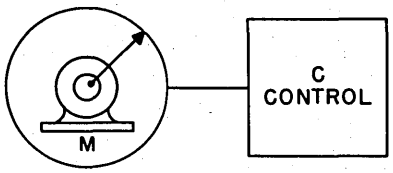


FIG. 1. Schematic diagram of a control servo.

Theoretically, the kicks are all of the same size, and the motor catches one or the other every time. In practice, the system was stabilized by supplying extra kicks whenever the sign of the error changed, but we will ignore this detail for the moment.

Several things are quite clear about this system. By no stretch of the imagination could it be solved as a linear system. The impulses to the motor are not at all proportional to the error. They are either full-sized positive or full-sized negative kicks. They cannot be linearized by averaging over long periods of time, because the sampling interval is not short enough to be negligible. On the other hand, we do not care about the exact history of the servo motor's position, provided its probable error in following the course is small enough, and provided the error never (or hardly ever) gets so large that the motor gets out of step.

Dynamically, this servo is very simple, if all the parameters, such as the initial position and velocity of the motor, the course, and the size of the kicks are known exactly. Knowing the position at any sampling instant, we can compare it with the position required by the course data, and determine whether the motor will be accelerated or decelerated. Adding or subtracting the resulting increment of velocity, we can easily calculate the motor position and velocity at the next sampling instant, 1/60 sec later. This completes one cycle of the operation, and we can repeat it as often as we like.

However, like the time of arrival at the end of a railroad trip, the motor position and velocity are discontinuous functions of the initial conditions, and these are never known exactly. Tiny variations in the initial position or speed would make the servo lead or lag at the next sampling instant, and hence change the sign of the kick; and the entire future path would be thereby altered, not infinitesimally, but by a finite amount. Luckily, this is one of the cases where we are not interested in the details of individual runs, but in the statistics of many runs.

A STATISTICAL METHOD

One of the vital bits of statistical information we need is the probability that the error will exceed one half of a revolution, for when this happens, the motor gets out of step and is lost. With the servo taking 60 chances every second, this probability must be very small indeed, if the number of failures is to be admissible. The probability of failure should not exceed 10^{-5} at most. If we want to estimate such a probability by the usual method, we would need to calculate something of the order of a million steps, for each proposed design. Such a program is staggeringly inefficient.

Evidently a new approach was needed. The servo I have been talking about is one of the simplest cases of this kind of problem, but it is not the only one. In general, a dynamical system that is subject to random influences of some kind, and that has severe nonlinearity or actual discontinuity so that simple linear theory cannot be applied, may be suitable subject matter for the statistical method.

With this brief statement of the reasons for wanting a method of treating the effects of random influences in highly nonlinear dynamical systems, I shall outline the scheme that finally solved the dynamic-tester servo problem, and suggest its extension to similar but more complicated problems. It is the extension to larger problems that calls for automatic computers.

As background for this outline, let us recall some ideas that are probably quite familiar, but that we shall use in a rather different way. The notion of a "phase space" for a dynamical system is very old, as such things go. The phase space for the dynamic-tester servo is very simple. It can be represented on paper by making a graph with the position and velocity (or momentum) of the motor as coördinates.

This is true because the state of the servo at any instant is completely determined when we know its position and its velocity, and we can represent these quantities by a single point on the graph. So we can calculate the future behavior of the servo under any set of forces when the location of its representative point in phase space is given.

As time goes on, the servo motor is kicked back and forth, and its position and velocity change. Its representative point therefore moves about in phase space, and we could trace its trajectory if we like; or, if we prefer, we can imagine that we take a series of motion-picture frames of the phase space, each frame showing the representative point in a slightly different position.

The usual method of dynamics traces the path of the representative point through phase space, starting at an arbitrarily chosen initial position. The tracing may be done by an analytic solution in some few cases, or it may be done step by step, as we said the dynamic-tester servo could be solved. The essential thing is that a representative point is tagged, so to speak, and followed as long as its history is of interest.

We have noted that the values of the variables and parameters (including the initial values) cannot be measured or established with absolute exactness. Even if we tried to start the system from a given point in phase space, we could not do so. All we can do is to start it so that its representative point is somewhere near a given spot in phase space. If we made a

great many tries and plotted the representative point in each case, a microscopic examination of the phase space near this spot would look as if a shotgun had sprayed it, or as if a swarm of bees had settled there. As time goes on, each representative point moves through phase space, so the swarm of bees drifts along.

This swarm (to use the bee analogy) may stay in a compact mass, actually condensing more and more, or it may disperse. In a stable dynamic system, the swarm keeps together more or less, but in an unstable system the swarm expands without limit, either actually or practically. If the system is linear or mildly nonlinear, the members of the swarm are well-behaved bees, seldom crossing each other's paths, and in general going along nicely side by side. If the system is discontinuous, however, every bee makes sudden decisions, changing its mind, and darting through the swarm. Despite this erratic behavior, the swarm may keep within fairly definite limits.

In terms of this picture of a swarm moving through phase space, it is easy to point out the distinction between the usual dynamic methods and the statistical approach. The usual methods follow the flight of one particular bee, whereas in the statistical method we shall study the motion of the swarm as a whole, observing its density and its tendency to disperse, and noting what proportion of the swarm gets lost.

Now that we have decided upon a point of view we can see whether it shows us a means for the practical solution of problems that arise.

Suppose that at any instant of time we have a picture of the phase space of a dynamical system, showing a swarm of representative points. For brevity, I shall call them simply dots. We can arbitrarily partition off the phase space into small cells, and we can count the number of dots in each cell. Then we can record these counts in the cells of our graph.

As time passes, the dots move about from cell to cell, according to the dynamical laws of the system. For the moment we can concentrate on two cells, A and B, that lie not too far apart in phase space. Suppose that at time $t = 0$ there are D_A dots in cell A and none elsewhere. The first frame of our imaginary movie then shows a density D_A in cell A, and zero density everywhere else. The next frame, which we shall say is 1/60 sec later, shows these dots somewhat scattered. A certain fraction of them, say $N(A,B)$, has moved from cell A into cell B, so that at this time there are $N(A,B)D_A$ dots in cell B.

Using the dynamical relations of the system, we can calculate the "transfer ratio" $N(A,B)$ for every pair of cells in phase space. When we have found the function $N(A,B)$, we can put our imaginary camera out of focus, so it no longer shows individual dots, but merely records the densities of dots in each cell.

That is, we are no longer interested in individual runs of the system, but only in the density of dots in a given cell at a given time. Obviously, we interpret this density as a probability. The density in a given cell at a given time is proportional to the probability that the system will be found at that instant to have the velocity and position corresponding to that cell.

I have passed very lightly over the construction of the transfer ratio $N(A,B)$. Evaluating this ratio is analogous to setting up the computing routine for the numerical solution of the

dynamical system, in the usual treatment. It varies from problem to problem, and I will explain later how it is set up for the particular example of the dynamic-tester servo.

I wish first, however, to review the general features of the statistical method. It will be recalled that the whole scheme depends upon plotting the state of a dynamical system in a phase space, so that when a representative point or dot is given, the future behavior of the system can be calculated from a knowledge of the forces acting on it. In the case of the simple dynamic-tester servo, the phase space is two-dimensional, with position and velocity of the motor as coördinates.

The dynamic specifications for the system tell where any dot moves to in an interval of time, but instead of following one dot throughout its path, we propose to deal with the density of these dots. Using the same dynamic specifications as in the usual method, we construct a transfer ratio that tells us how the density in each cell at the end of a short time interval is related to the density distribution throughout the phase space at the start of that interval.

Instead of starting with an assumed initial point in phase space and following its path, then, we start with an assumed initial density of dots representing the initial probability distribution for the system, and apply the transfer operation to see how this distribution changes with time. We frequently find that we can choose the coördinates so that the distribution quickly settles down into a static pattern, unless the system itself is unstable. In the latter case, of course, the density distribution spreads out more and more, approaching zero everywhere in any finite region.

We have seen that the simple servo has a phase space of two dimensions, but if it had more degrees of mechanical freedom, the phase space would have more dimensions. I think it is clear why such a problem would require an automatic computer.

Having mentioned some problems in which a statistical method seems desirable, and having outlined the scheme that was found useful in solving one of those problems, I want to give a summary of the numerical methods and results of one simple example.

We shall reduce the dynamic-tester servo to its simplest form, and shall suppose that the "course" called for by the control mechanism is identically zero. Let p be the angular position of a brush carried on the motor shaft, measured in arbitrary units clockwise from the zero position. For convenience in computation, these units may coincide with the size of the cell; they might, for instance, be 10° of actual motor rotation. In the same way, the velocity q may be measured in units of one position unit per unit of time. It is convenient to use the interval between samplings as the unit of time. Then 1 sec is 60 units of time.

We draw a picture (Fig. 2) of the phase space for this system, with coördinates p and q. A dot at P_0 represents the servo at an instant when the motor passes through a position 3 units clockwise from its zero position, with a velocity of 2 units. Ideally, the control would deliver a negative impulse to the motor whenever the dot is to the right of $p = 0$ in the graph. We make the simple assumption that the kick is of unit magnitude, so that it instantaneously changes the velocity by 1 unit, and we assume that the motor moves with constant speed between kicks.

In the picture, we see two possible paths for the dot. If the motor were not kicked, its dot would move to the right 2 units of position, because its velocity is 2 units. Since the motor does get a kick, the dot moves down one space, because its velocity is reduced from 2 units to 1 unit. The dot therefore moves to the right one space, since the motor travels for 1 unit of time with 1 unit of velocity. Similar paths can be found for dots in each cell.

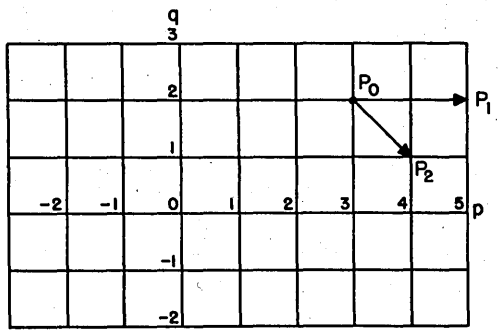


FIG. 2. Phase space for the simplest form of dynamic-tester servo.

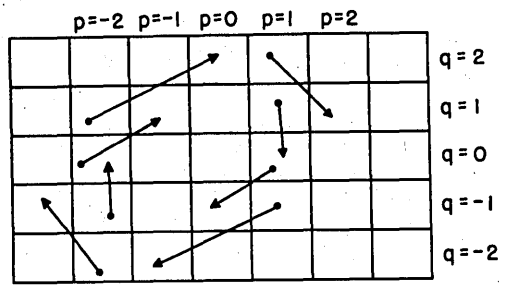


FIG. 3. Motions of contents of some of the cells of the phase space.

It will be convenient to express the change in position of the dot during one unit of time by writing equations for the changes Δp and Δq in the coördinate values p and q. Then the simple assumptions we have made are equivalent to saying that

$$\Delta q = -\operatorname{sgn} p,$$

$$\Delta p = q + \Delta q,$$

where p and q are the coördinates at the start of the interval.

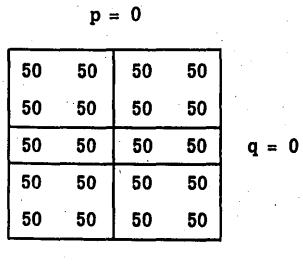


FIG. 4. Uniform distribution of 1000 dots in 20 cells.

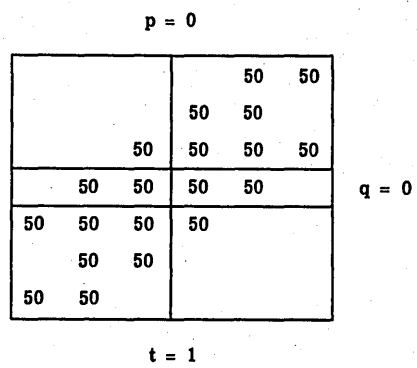


FIG. 5. The distribution of Fig. 4 after one application of the transfer scheme.

In this trivial case, it is easy to see that the entire contents of each cell moves into another cell whose position is defined by the conditions on Δq and Δp just stated. Figure 3 shows how the contents of some of the cells will move. We will carry the trivial case one step further by distributing a thousand dots uniformly over the 20 cells shown in Fig. 4. Each cell has 50

dots, but since we are not interested in their individualities, we simply mark each cell with the number of dots it contains. We apply the transfer scheme already discussed to this distribution, and find that the densities have shifted to the pattern of Fig. 5. One more application of the transfer (Fig. 6) will be enough of this rather uninteresting example. We have already

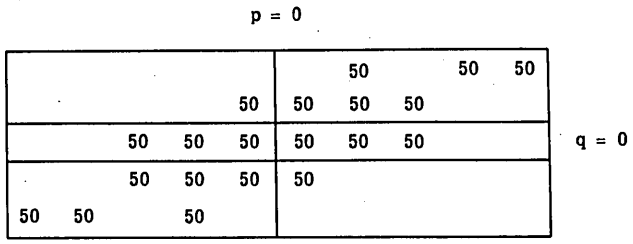


FIG. 6. The distribution of Fig. 4 after two applications of the transfer scheme.

gone far enough to see that the distribution shows signs of dissipating and, as a matter of fact, the system will be found to be unstable.

To make an interesting and useful example, we need, first, to take account of the imperfections of the system, and second, to introduce a stabilizing mechanism.

It is clear that in practice the control cannot decide without error whether p is positive or negative. Because of the finite width of the control brush, as well as backlash, vibration, errors in timing the sample, and so on, there will be occasions when the motor gets a kick of

Table 1. Distribution of probability $\times$ 1000 at $t = 0$.

																		q
0	0	0	0	0	0	0	50	50	50	50	0	0	0	0	0	0	0	2
0	0	0	0	0	0	0	50	50	50	50	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	50	50	50	50	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	50	50	50	50	0	0	0	0	0	0	0	-1
0	0	0	0	0	0	0	50	50	50	50	0	0	0	0	0	0	0	-2
-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	p

the wrong sign. Of course, the farther the motor is from zero position the less chance there is for a mistake of this kind. In an actual application of the method, an estimate of the probability of such mistakes would be made by examining the mechanism. For our example, we shall simply say that 30 percent of the dots in the blocks between + 1 and — 1 are subject to error. Then 30 percent of the contents of the first column of cells to the right of $p = 0$ will move as if they received positive kicks, and 70 percent as if they received negative kicks, with corresponding conditions in the left-hand side. All cells further removed will receive proper kicks.

There are many factors that affect the size of the impulses, as well as their time of application. Again, if this were an actual problem, we would need to examine the mechanism to estimate the probabilities involved, but as this is merely an example, we arbitrarily choose a

Table 2. Distribution of probability × 1000 at $t = 1$.

																			q
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	3	2	5
0	0	0	0	0	0	0	0	0	0	0	0	0	40	35	15	2	0	4
0	0	0	0	0	0	0	0	0	0	0	5	43	35	15	2	0	0	3
0	0	0	0	0	0	0	0	0	0	5	43	35	15	2	0	0	0	2
0	0	0	0	0	0	0	0	2	8	48	35	15	2	0	0	0	0	1
0	0	0	0	0	0	12	30	48	48	30	12	0	0	0	0	0	0	0
0	0	0	0	2	15	35	48	8	2	0	0	0	0	0	0	0	0	-1
0	0	0	2	15	35	43	5	0	0	0	0	0	0	0	0	0	0	-2
0	0	2	15	35	43	5	0	0	0	0	0	0	0	0	0	0	0	-3
0	2	15	33	40	0	0	0	0	0	0	0	0	0	0	0	0	0	-4
2	3	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-5
2	5	22	52	92	93	95	83	58	58	83	95	93	92	52	22	5	2	Total
-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	p

Table 3. Distribution of probability × 1000 at $t = 2$.

																				q
0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	5	5	7	4	2	3
0	0	0	0	0	0	0	0	0	0	0	11	28	32	48	31	14	2	0	0	2
0	0	0	0	0	0	0	0	2	15	35	51	49	40	15	2	0	0	0	0	1
0	0	0	0	0	0	2	15	35	49	49	35	15	2	0	0	0	0	0	0	0
0	0	0	0	2	15	40	49	51	35	15	2	0	0	0	0	0	0	0	0	-1
0	0	2	14	31	48	32	28	11	0	0	0	0	0	0	0	0	0	0	0	-2
2	4	7	5	5	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-3
2	4	9	19	38	67	75	92	99	99	99	99	92	75	67	38	19	9	4	2	Total
-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	p

simple law, and say that 10 percent of the contents of any cell will get kicks that are smaller than normal by 1 unit, and 10 percent kicks that are larger. We shall select as the normal impulse, for this example, one that makes Δq equal to 2 units of velocity.

Next, we consider the question of stability. Actually, the dynamic-tester servo was stabilized by making the impulses much greater whenever the error changed sign, but the analysis will be a little simpler if we calculate the distribution when a friction flywheel is used. This flywheel

Table 4. Distribution of probability $\times$ 1000 at $t = 3$.

																				q
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	4
0	0	0	0	0	0	0	0	0	0	0	0	0	3	12	12	2	0	0	0	3
0	0	0	0	0	0	0	0	0	0	3	16	34	34	14	1	1	0	1	1	2
0	0	0	0	0	0	0	3	15	38	46	49	28	11	9	9	5	2	0	0	1
0	0	0	0	2	12	29	43	31	39	39	31	43	27	12	2	0	0	0	0	0
0	0	2	5	9	9	11	28	49	46	38	15	3	0	0	0	0	0	0	0	-1
1	1	0	1	1	14	34	34	16	3	0	0	0	0	0	0	0	0	0	0	-2
0	0	0	2	12	12	3	0	0	0	0	0	0	0	0	0	0	0	0	0	-3
0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-4
1	1	3	9	25	47	75	108	111	126	126	111	108	75	47	25	9	3	1	1	Total
-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	p

Table 5. Distribution of probability $\times$ 1000 at $t = 7$.

																			q
0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	1	0	4
0	0	0	0	0	0	0	0	0	0	0	1	11	19	13	2	0	0	3
0	0	0	0	0	0	0	1	3	5	16	33	34	16	4	3	1	0	2
0	0	0	0	0	0	2	9	19	36	56	37	23	13	2	1	1	1	1
0	0	0	1	6	7	23	43	54	54	43	23	7	6	1	0	0	0	0
1	1	1	2	13	23	37	56	36	19	9	2	0	0	0	0	0	0	-1
0	1	3	4	16	34	33	16	5	3	1	0	0	0	0	0	0	0	-2
0	0	2	13	19	11	1	0	0	0	0	0	0	0	0	0	0	0	-3
0	1	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-4
1	3	9	22	55	75	96	125	117	117	125	96	75	55	22	9	3	1	Total
-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	p

is loosely coupled to the motor, and slips for a short period after the application of each impulse. Therefore the change in position is greater than it would be if no such slippage occurred. It can be shown that this is a stabilizing influence. Let Δy be equal to $p + 2\Delta q$, the figure 2 representing a stabilizing factor. Again, we ignore effects that would be taken into account

in an actual problem. The friction is not constant, the timing is not exact, the velocity of the motor is not strictly constant, and so on. These things are easily taken into account, but we omit them for simplicity.

Table 6. Distribution of probability × 1000 at $t = 11$.

																q
0	0	0	0	0	0	0	0	0	0	0	0	1	3	3	1	4
0	0	0	0	0	0	0	0	0	0	2	11	16	13	1	0	3
0	0	0	0	0	0	1	3	5	17	33	34	16	5	4	1	2
0	0	0	0	0	2	11	22	38	54	37	23	12	2	0	0	1
0	0	1	3	6	19	40	53	53	40	19	6	3	1	0	0	0
0	0	2	12	23	37	54	38	22	11	2	0	0	0	0	0	-1
1	4	3	16	34	33	17	5	3	1	0	0	0	0	0	0	-2
0	1	13	16	11	2	0	0	0	0	0	0	0	0	0	0	-3
1	3	3	1	0	0	0	0	0	0	0	0	0	0	0	0	-4
2	8	24	48	74	93	123	121	121	123	93	74	48	24	8	2	Total
-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	p

Table 7. Distribution of probability × 1000 at $t = 12$.

																q
0	0	0	0	0	0	0	0	0	0	0	0	1	3	3	1	4
0	0	0	0	0	0	0	0	0	0	2	11	16	13	2	0	3
0	0	0	0	0	0	1	3	6	19	37	36	17	5	4	1	2
0	0	0	0	0	2	11	22	38	54	37	20	13	1	0	0	1
0	0	1	3	5	17	36	50	50	36	17	5	3	1	0	0	0
0	0	1	13	20	37	54	38	22	11	2	0	0	0	0	0	-1
1	4	5	17	36	37	19	6	3	1	0	0	0	0	0	0	-2
0	2	13	16	11	2	0	0	0	0	0	0	0	0	0	0	-3
1	3	3	1	0	0	0	0	0	0	0	0	0	0	0	0	-4
2	9	22	50	72	95	121	119	119	121	95	72	50	22	9	2	Total
-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	p

Often it is simpler to tell where the contents of a cell go than to tell where the cell gets its contribution from. In other words, it is easier to express the distribution of density at the end of an interval as the sum of a number of partial densities, each consisting of the flow of probability due to one of the contributory causes. This formulation seems to make for easier

application of machine computation, also, since it reduces the amount of sorting required. In the present example, for instance, we see that, except for the first column of cells, 10 percent of the contents of each cell on the right move down one cell. If q_0 is the value of q from which it moves, then this density moves to the right $q - 2$ cells. Only 70 percent of the first column to the right of zero is affected by this move, and only 30 percent of the first column on the left is affected. This is a simple transfer, which is easily mechanized. It gives us a partial distribution. Similar partial distributions can be formed of the 80 percent that moves down two

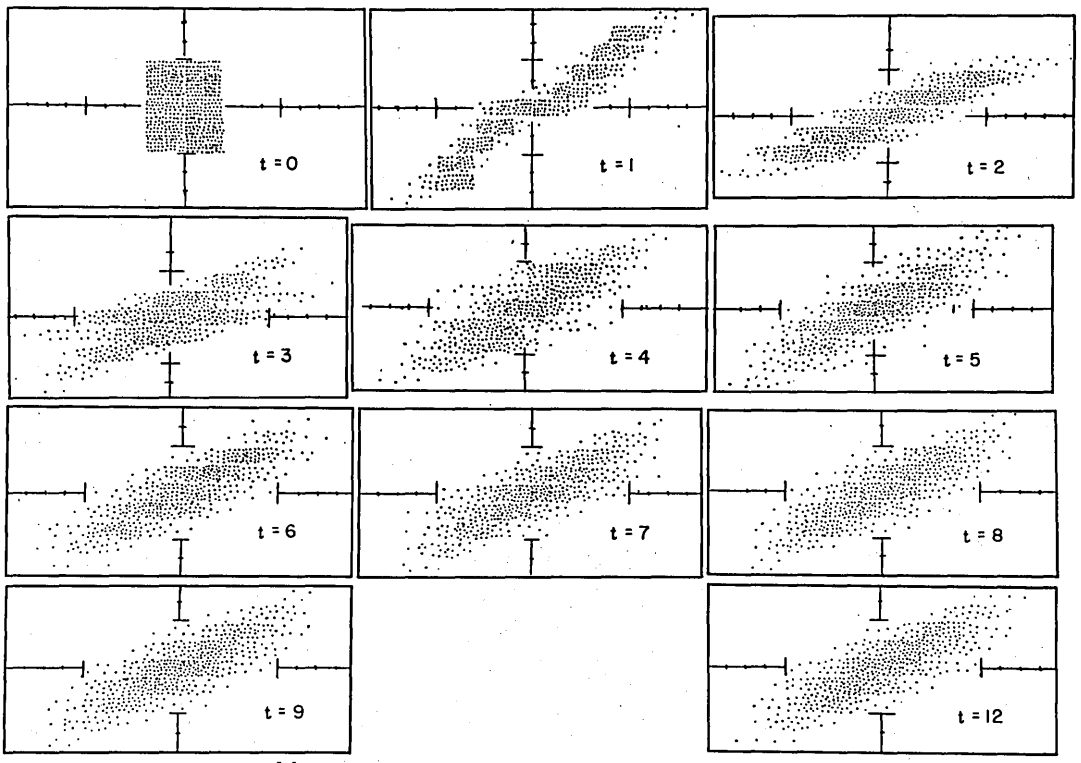


FIG. 7. Results of applying the transfer scheme to an initial rectangular distribution of density (see Tables 1 to 7).

rows and right $q - 4$ columns, and so on. Summing all such partial transfers, we have the total transfer of probability or density.

Tables 1 to 7 give the results of applying the transfer for the simplified dynamic-tester servo to the same rectangular distribution of density that we used as the initial distribution for the trivial example. The same results are shown graphically in Fig. 7.

In this paper we have discussed dynamical problems, notably those about systems having severe nonlinearities or actual discontinuities, for which we want statistical information. One of these is the dynamic-tester servo.

Next, a statistical method has been worked out for following a distribution of runs of

such a system instead of dealing with individual runs. The method consists of plotting densities or probabilities in a phase space, and of calculating transformations for such densities. In practice the transformation can be carried out by automatic computers, and if the number of degrees of freedom for the dynamical system is large, there will be a great many cells, and automatic computers will really be needed.

Finally, we have glanced at a simple example, and have seen how the probability flows around in phase space and gradually settles down to a steady flow pattern.

Combustion Aerodynamics

HOWARD W. EMMONS

Harvard University

Fluid-mechanics problems have been developed to date with restrictive assumptions based in part upon the problems whose solutions were sought and in part on mathematical convenience. By far the largest amount of work has been done on questions involving the flow of incompressible ideal fluids, that is, fluids of constant density and zero viscosity. Since the beginning of this century considerable progress has been made in the extension of our knowledge of fluid mechanics through the addition of studies of the effects of viscosity—thus abandoning the ideal fluid—and, more recently, of problems in the flow of compressible fluids —thus abandoning the assumption of constant density. Most work on compressible fluids has involved a specific type of compressibility, namely, that of the ideal gas.

In the study of compressible fluids the interesting phenomenon of shock waves appears. These disturbances are studied directly in order to determine their fundamental nature, that is, the variation of temperature, pressure, velocity, etc., through the shock wave itself. They are studied indirectly by considering ideal nonsteady flow to determine when and how shock waves first develop, and by studying the flow in regions between shock waves, the shock waves themselves being treated as discontinuities.

In a few cases authors have treated problems that involve, besides the various phenomena already mentioned, heat transfer between various parts of the fluid and between the fluid and the walls. There is one very important phenomenon, however, that is excluded from most of these treatments, namely, combustion. Fluid mechanics has progressed to the point where this phenomenon, or at least the simpler aspects of it, can be added. Take, for example, a phenomenon with which everyone is familiar, the flame on a Bunsen burner. When this flame is small, there is a very steady, sharp, central cone surrounded by a stream of hot, somewhat luminous gas which fans out above it upon the top of the burner. When the flame is large, everyone is familiar with its rather random oscillations. This phenomenon, which involves the stability of a jet and the stability of the combustion process simultaneously, would undoubtedly be difficult to analyze. However, a small quiescent flame appears innocent enough and suggests itself as an object for study. Curiously enough, this apparently simple phenomenon has not to my knowledge been computed.

I need not spend any appreciable time enumerating practical problems in which a knowledge of combustion aerodynamics is important. Besides the simple gas flames such as that already mentioned, which are used not only for laboratory work but also for cutting torches, welding torches, and other devices, there are all the various furnaces and combustion chambers used throughout industry. In every case the aerodynamic phenomena are responsible for

accomplishing the distribution of heat in the desired way. At present our knowledge of this field is almost entirely empirical. Certain general principles are yet unknown. This is brought out most clearly when one observes the difficulty encountered in attempting to change the scale of a piece of equipment involving a flame.

Perhaps the most important approach available to the engineer on problems that are inherently too complex for present-day computation is the possibility of testing on a model scale and then using the resulting information for the design of the prototype. With combustion this is impossible, since the development of a small furnace gives few clues to the performance of another, say twice as large. In the present paper, the relatively simple problem of the Bunsen burner will be considered.

It is clear that the most important addition to present-day aerodynamics requiring consideration in order to include combustion is the interaction of chemical reactions with the motion of the fluids. Consideration here will be limited to the combustion of premixed gases. We might note that this does not include all of the phenomena of importance, since very frequently the fuel and air are not premixed, but, for example, a liquid fuel is sprayed in fine droplets from a nozzle. These droplets must evaporate and the resulting vapor must mix with air before it can burn.

Since we are here interested only in the aerodynamics of combustion, we need consider only the chemistry of the problem to the extent to which this influences the fluid motions. At this point we encounter a setback since, while the aerodynamics of combustion has an almost nonexistent literature, the chemistry of combustion has a most extensive literature. In spite of this, however, the setback is serious since an examination of this literature shows that in spite of valiant attempts the chemistry of combustion, which in large part is the chemistry of reaction rates, is by no means well understood. In fact, for even the simplest of reactions the lack of understanding is still tremendous. From our point of view a review of what is known brings out the following as important phenomena. The reaction which obviously propagates into the unburned material from the burned gases, thus continuing the reaction in the neighborhood of the relatively stationary flame front, is propagated through a combination of effects. The diffusion of chemical species from the burned into the unburned material may act as chain carriers and thus initiate further reaction. Heat from the hot, burned gases may propagate forward by thermal conduction (and perhaps in some cases by radiation) into the unburned gases and thus bring about, through the dependence of reaction rate on temperature, the reaction in the unburned gas. The reactions themselves depend of course upon the precise chemical nature of the combustible mixture. Ions and free radicals and various molecular species are present in concentrations that vary from place to place through the region of combustion. In fact, probably the best definition of the region of combustion would be that region in which the composition of the gases differs significantly from the reactants and combustion products. The chemistry of the reaction region is not only exceedingly complex but is still covered by an extremely dense veil of ignorance.

Thus there appear open to us in the study of the aerodynamics of combustion the same two possible approaches as are available and are used in connection with shock-wave phenomena. On the one hand, we might focus our attention upon the aerodynamics of the flame itself and ask for the variation of temperature, pressure, velocity, and, what is more important, composition, and so forth, through the reaction region. Numerous such studies have been attempted with more or less success in the past in order to clarify the chemistry of the problem. A great deal remains to be done in this direction and it is to be hoped that the study of combustion aerodynamics may indeed lead ultimately to the understanding of flame propagation.

On the other hand, we can focus our attention not on the reaction region but on the flow before and after that region. Our success in this approach, like our success in the corresponding approach to flow with shock waves, depends very largely upon the physical dimensions of the combustion region. If, like shock waves, the combustion region is indeed a small fraction of an inch in thickness, then the assumption that it can be replaced by a mathematical discontinuity will not invalidate our results. However, if the reaction region is wide compared with other pertinent dimensions of the apparatus, our results would be without significance. We will here make the assumption that the reaction region is thin and can be replaced by a discontinuity. Thus we are considering those cases of combustion in which the reaction is essentially completed over a very short distance. At present it appears that for most premixed gases this assumption is correct. Even in the case of luminous flames in which carbon particles are obviously burning over a large region, the primary combustion which liberates these excess carbon particles takes place very rapidly, the luminous region being present only because of the mixing of the "burned" gases with additional air, thus permitting secondary combustion of the carbon in a diffusion flame.

We will thus assume that a mixture of combustible gases in an equilibrium mixture arrives at the flame front. These gases then instantaneously react to an equilibrium mixture of combustion products. Flame-front relations can be derived by the application of the continuity, momentum, and energy laws to an element of the flame front. Such analyses have been made many times under fairly general circumstances.

Since we are here intending to set up for solution the entire flow field, it is desirable to simplify the problem as completely as possible. Thus it is to be observed that for a Bunsen burner and other low-velocity (laminar) combustion processes, the pressure variations throughout the flowing gases are relatively small, as are also the temperature variations except across the flame front itself. We may assume, therefore, that the unburned gases flow as an incompressible fluid, and in addition that the burned gases also flow as an incompressible fluid. We must, however, take into account the discontinuous change of density across the flame front. We will denote the density ratio by n. Then

$$n = \frac{\rho_1}{\rho_2} = \frac{q_{n_2}}{q_{n_1}}, \tag{1}$$

where ρ is the density of the gas, q_n is the component of velocity of the fluid normal to the flame front, the subscript 1 refers to the unburned mixture and the subscript 2 to the products

of combustion. For incompressible fluids, the value of n completely determines all of the flame properties except the flame-propagation rate S_t. The flame-propagation rate, which is the velocity of propagation of the flame front normal to itself into the unburned gases, will be assumed to be constant. Then

$$\begin{aligned} q_{n_1} &= q_1 \sin \theta_{w_1} = S_t, \\ q_{n_2} &= q_2 \sin \theta_{w_2} = nS_t, \end{aligned} \tag{2}$$

where q is the velocity at the flame front and θ_w is the angle between the velocity vector and the flame front.

It is clear that both n and S_t must be determined for the particular combustible being used. For a given combustible, n can be computed with considerable precision from the equilibrium properties of the burned and unburned fluids. The transformation rate S_t, however, cannot yet be computed; moreover, current experimental data, as interpreted, do not show S_t to be an absolute constant. In fact, one of the principal uses to which the present theory would be put is to take accurately into account those aerodynamic aspects of flames which must be understood in order to determine accurately whether or not S_t is constant for a given flame.

With these two constants the flame-front relations can be completely determined. We note from Eqs. (2) that the velocity normal to the flame front changes discontinuously from S_t to n times this value. The momentum equation written for an axis along the flame front shows that the tangential velocity component q_t does not change;

$$q_{t_1} = q_{t_2}. \tag{3}$$

Thus the resultant velocity changes discontinuously from q_1 to a larger value q_2 which deviates in direction from q_1 by an angle δ. All these relations are shown in Fig. 1. The difference in velocity components in a given direction making an angle ν with the flame front is shown by the geometry of Fig. 1. Thus

$$q_{\nu_2} - q_{\nu_1} = (n-1)S_t \sin \nu. \tag{4}$$

At this point we shall restrict further consideration to two-dimensional flow. Thus, in Fig. 1 we show x- and y-axes with a flame front making an angle α with the x-axis. For this case, with $\nu = \alpha$, we get

$$u_2 - u_1 = (n-1)S_t \sin \alpha, \tag{5}$$

and with $\nu = \alpha - 90°$,

$$v_2 - v_1 = -(n-1)S_t \cos \alpha, \tag{6}$$

where u and v are the components of q along the coördinate axes.

As will be seen in the following, the only other flame-front relation we need is the discontinuous change in total pressure p_0. To derive this relation we start with the momentum equation written for an axis normal to the flame front. This gives

$$p_2 - p_1 = \rho_1 S_t^2 (1-n), \tag{7}$$

where p is the static pressure. The total pressure p_0, which is constant on streamlines between discontinuities, is now found from the Bernoulli equation,

$$p + \tfrac{1}{2}\rho q^2 = p_0. \tag{8}$$

Thus we get, for the difference in total pressure,

$$p_{0_2} - p_{0_1} = \frac{\rho_1 S_t^2}{2}(n-1)\left(1 + \frac{\cot^2 \theta_{w_1}}{n}\right). \tag{9}$$

The flow of an incompressible fluid in two dimensions is described by the continuity and irrotationality relations (using the usual hydrodynamic notation),

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \tag{10}$$

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 2\omega. \tag{11}$$

We introduce the volume-flow stream function ψ by

$$u = -\frac{\partial \psi}{\partial y}, \quad v = \frac{\partial \psi}{\partial x}. \tag{12}$$

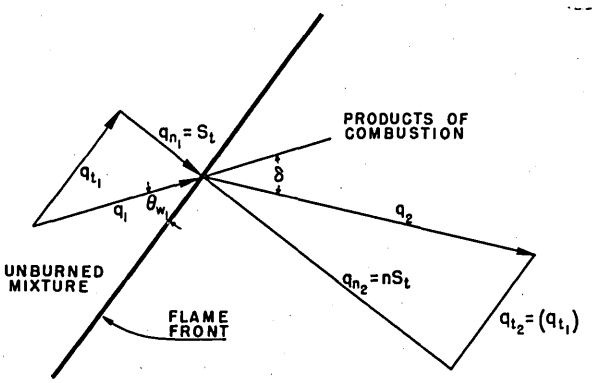


FIG. 1. Aerodynamic flame-front relations.

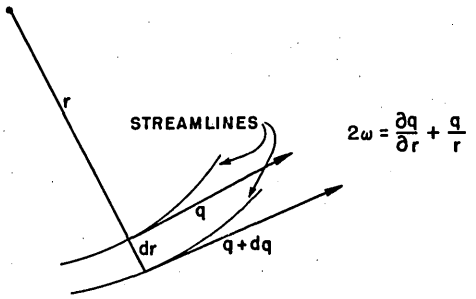


FIG. 2. Relation between rotation and velocity variation.

Thus by Eq. (11) ψ is given by

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 2\omega. \tag{13}$$

For an incompressible fluid the rate of rotation is related to the total pressure of the fluid; the relation is most easily derived from Fig. 2. We have

$$2\omega = \frac{\partial q}{\partial r} + \frac{q}{r}; \tag{14}$$

but, by Eq. (8),

$$\frac{\partial p}{\partial r} + \rho q \frac{\partial q}{\partial r} = \frac{\partial p_0}{\partial r}, \tag{15}$$

and by the radial-momentum equation,

$$\frac{\partial p}{\partial r} = \frac{\rho q^2}{r}. \tag{16}$$

Thus

$$2\omega = \frac{1}{\rho q}\frac{\partial p_0}{\partial r}. \tag{17}$$

But

$$q = \frac{\partial \psi}{\partial r}. \tag{18}$$

Thus finally the desired relation is obtained:

$$2\omega = \frac{1}{\rho}\frac{dp_0}{d\psi}. \tag{19}$$

Since p_0, the Bernoulli constant, is a function of ψ, the rate of rotation is fixed on streamlines, $\omega(\psi)$. We are now in a position to write the equations that have to be solved in order to understand the aerodynamics of a simple two-dimensional combustion problem.

For the unburned mixture (subscript 1),

$$\frac{\partial^2 \psi_1}{\partial x^2} + \frac{\partial^2 \psi_1}{\partial y^2} = 2\omega_1; \tag{20}$$

for the products of combustion (subscript 2)

$$\frac{\partial^2 \psi_2}{\partial x^2} + \frac{\partial^2 \psi_2}{\partial y^2} = 2\omega_2. \tag{21}$$

For boundary conditions we must rely upon our physical knowledge to assure ourselves that we have a sufficient set to obtain a solution. For the unburned mixture we specify the flow passage, by specifying, for example, the channel walls. In addition the velocity distribution must be given at some upstream point (perhaps $x = -\infty$).

Thus setting $\psi_1 = 0$ on one wall, we compute the value of ψ_1 on the upstream section from the given velocity distribution. The value of ψ on the other channel wall is then found and is equal to the total volume flow of combustible gas. These boundary conditions are not sufficient to determine ψ_1, since as yet no conditions closing the domain on the flame side have been given. We note that the vorticity distribution in the inlet stream is given by the given velocity distribution by use of Eqs. (8) and (19). Since the rate of rotation ω is constant on streamlines, the vorticity distribution ω_1 is determined simultaneously with ψ_1.

For the products of combustion, we again can specify a priori boundary conditions on three sides. Channel walls may be given if the combustion takes place within a passage, or free streamlines may be specified as for a Bunsen flame. In either case, the stream function is known by continuity:

$$\psi_2 = n\psi_1. \tag{22}$$

For free streamlines the additional fact of constant pressure, hence constant velocity, is needed. The free-streamline location will be given by the solution. At some downstream section (perhaps $x = \infty$), the pressure is taken as constant. This is a sufficient condition, since Eq. (8) is a relation between p_{0_2} and q_2 and hence between a function of ψ, $p_{0_2}(\psi)$ and its first derivative $q_2 = \partial\psi/\partial n$, where n is normal to the as yet undetermined streamlines.

To complete the specification of the problem we must add sufficient conditions connecting ψ_1 and ψ_2 along the flame front so that ψ_1, ψ_2 and the flame-front location can be found. A sufficient set of conditions is provided by Eqs. (5), (6), and (22) in the form

$$\psi_2 = n\psi_1, \tag{22}$$

$$\frac{\partial \psi_1}{\partial y} - \frac{\partial \psi_2}{\partial y} = (n-1)S_t \sin \alpha, \tag{23}$$

$$\frac{\partial \psi_1}{\partial x} - \frac{\partial \psi_2}{\partial x} = (n-1)S_t \cos \alpha. \tag{24}$$

To make the problem soluble, we yet require a method of finding ω_2. This is supplied by Eqs. (9) and (19). From these we find at the flame front the relation

$$2\omega_2 = 2\omega_1 - \frac{(n-1)}{2} S_t^2 \frac{d \cot^2 \theta_{w_1}}{d\psi_2}. \tag{25}$$

The only analytical solution so far found to the above system of equations is the plane oblique flame separating two uniform parallel streams.

If the combustible is flowing along the x-axis with constant velocity U, the unburned stream function is

$$\psi_1 = -Uy. \tag{26}$$

If a plane flame at angle $\alpha = \theta_w$ passes through the z-axis its equation is

$$y' = x \tan \alpha = mx, \tag{27}$$

and the resulting stream function of the products of combustion is

$$\psi_2 = -\frac{(n-1)mU}{1+m^2} x - \frac{1+nm^2}{1+m^2} Uy. \tag{28}$$

Equations (26), (27), and (28) are the analytic expressions for the flame of Fig. 1 (with $q_1 = U$, parallel to the x-axis).

The channel problem shown in Fig. 3 has not yet been solved but will be set up for solution by two different methods.

The first is an integral-equation method in which the equations can be solved numerically by an iteration procedure. It is based upon the observation that the flame front can be considered a line source of strength density

$$\left(\frac{\partial \psi'}{\partial l}\right)_{ff} = (n-1)S_t, \tag{29}$$

where l is distance along the flame front measured to the right when crossing with the fluid, and the subscript indicates differentiation along the flame front. In addition to the constant source strength, the flame front position is determined by the condition that it propagates at the constant rate S_t;

$$\frac{\partial \psi_1}{\partial l} = S_t. \tag{30}$$

We now have the entire flow specified by Eqs. (12) and (13). The boundary conditions on four sides—inlet, outlet, and channel—are the same as before, no distinction being made between burned and unburned fluid. The separation of burned and unburned fluid is accomplished by setting a line source of strength given by Eq. (29) in such a location as to satisfy Eq. (30).

The channel of Fig. 3 imposes the specific boundary conditions

$$\psi = 0 \text{ at } x = 0, \tag{31}$$

$$\psi = 1 \text{ at } x = 1, \tag{32}$$

$$\psi = x \text{ at } y \to -\infty, \quad \text{(uniform parallel stream of combustible at inlet with velocity } v = 1) \tag{33}$$

$$\frac{\partial \psi}{\partial y} = 0 \text{ at } y \to \infty. \quad \text{(asymptotically constant velocity on each streamline—this is equivalent to the constant-pressure condition)} \tag{34}$$

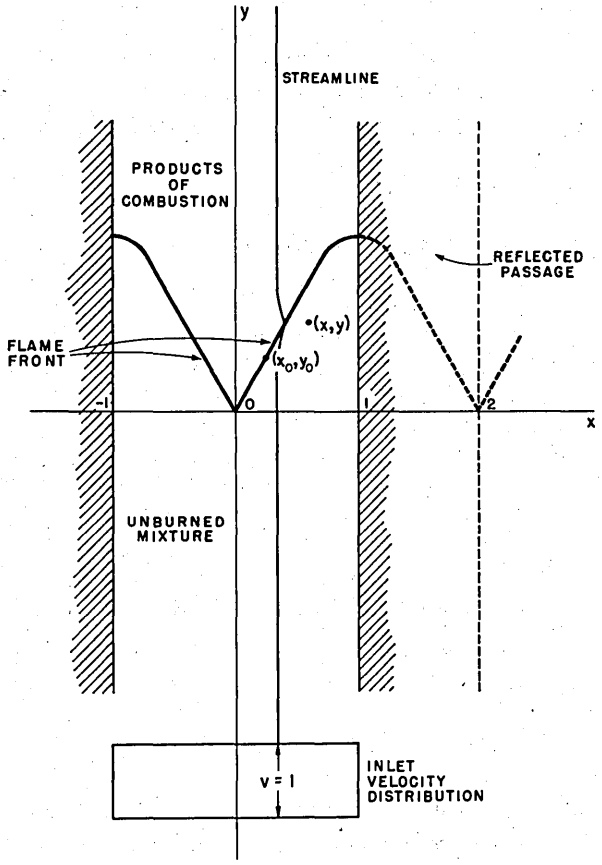


FIG. 3. Combustion in a channel.

The foregoing variables may be considered dimensionless if we use as the unit dimensions the width of the channel, the fluid velocity at the inlet, and the total inlet volume flow.

The velocities induced at any point (x, y) by a source of strength Q at the point (x_0, y_0) (including $(-x_0, y_0)$ and all image points) to satisfy the boundary conditions are

$$u_Q = \frac{Q}{4}\left[\frac{\sin \pi(x - x_0)}{\cosh \pi(y - y_0) - \cos \pi(x - x_0)} + \frac{\sin \pi(x + x_0)}{\cosh \pi(y - y_0) - \cos \pi(x + x_0)}\right]$$

$$= Qq_u(x, y, x_0, y_0), \tag{35}$$

$$v_Q = \frac{Q}{4}\left[\frac{\sinh \pi(y-y_0)}{\cosh \pi(y-y_0) - \cos \pi(x-x_0)} + \frac{\sinh \pi(y+y_0)}{\cosh \pi(y-y_0) - \cos \pi(x+x_0)}\right]$$

$$= Qq_v(x,y,x_0,y_0), \tag{36}$$

The corresponding velocities induced by a vortex of strength Γ at the point (x_0, y_0) meeting the same boundary conditions are

$$U_\Gamma = -\Gamma q_v(x,y,x_0,y_0), \tag{37}$$

$$v_\Gamma = \Gamma q_u(x,y,x_0,y_0). \tag{38}$$

Let the flame be situated along the line $y_f = y_f(x_f)$. If we suppose this line known we can write down the velocity components at any point (x, y) for the solution to Eq. (13) using as the source strength

$$Q = d\psi' = (n-1)S_t dl = (n-1)S_t\left[1 + \left(\frac{dy_f}{dx_f}\right)^2\right]^{\frac{1}{2}} dx_f \tag{39}$$

and as the vortex strength

$$\Gamma = \omega(x_0,y_0)\, dx_0 dy_0. \tag{40}$$

The velocity components are

$$u(x,y) = (n-1)S_t \int_0^1 \left[1 + \left(\frac{dy_f'}{dx_f}\right)^2\right]^{\frac{1}{2}} q_u[x,y,x_f',y_f'(x_f')]dx_f'$$

$$- \int_0^1 dx_0 \int_{-\infty}^{\infty} \omega(x_0,y_0)\, q_v(x,y,x_0,y_0)dy_0, \tag{41}$$

$$v(x,y) = (n-1)S_t \int_0^1 \left\{1 + \left(\frac{dy_f'}{dx_f'}\right)^2\right\}^{\frac{1}{2}} q_v(x,y,x_f',y_f'(x_f'))dx_f'$$

$$+ \int_0^1 dx_0 \int_{-\infty}^{\infty} \omega(x_0y_0) q_u(x,y,x_0,y_0)dy_0 + v_0, \tag{42}$$

where v_0 is a constant to be selected to provide the given inlet velocity. We now use the condition that the flame propagates at a fixed rate s_t—Eq. (30)—in the form

$$S_t = v\frac{dx_f}{dl} - u\frac{dy_f}{dl} = \left(v - u\frac{dy_f}{dx_f}\right)\left[1 + \left(\frac{dy_f}{dx_f}\right)^2\right]^{-\frac{1}{2}} \tag{43}$$

Thus the integral equation to be solved for $y_f(x_f)$ becomes

$$\left[1 + \left(\frac{dy_f}{dx_f}\right)^2\right]^{\frac{1}{2}} = (n-1)\int_0^1 \left[1 + \left(\frac{dy_f'}{dx_f'}\right)^2\right]^{\frac{1}{2}} \left[q_v(x_f,y_f,x_f',y_f') - \frac{dy_f}{dx_f} q_u(x_f,y_f,x_f',y_f')\right]dx_f'$$

$$+ \int_0^1 dx_0 \int_{-\infty}^{\infty} \frac{\omega(x_0,y_0)}{S_t}\left[q_u(x_f,y_f,x_0,y_0) + \frac{dy_f}{dx_f} q_v(x_f,y_f,x_0,y_0)\right]dy_0 + \frac{v_0}{S_t}. \tag{44}$$

The functions q_u, q_v have poles at $(x_f', y_f') = (x_f, y_f)$. The first integral on the right is to be taken around the pole on the side of the burned fluid, that is $y_f' > y_f$ at $x_f' = x_f$. If the

principal value of the first right-hand integral is taken, the value obtained in passing around the pole must be added. This value is

$$-\frac{(n-1)}{2}\left[1+\left(\frac{dy_f}{dx_f}\right)^2\right]^{\frac{1}{2}}. \tag{45}$$

To determine the constant v_0 it is simplest to integrate Eq. (29) to get the total source strength as

$$\psi' = (n-1). \tag{46}$$

Since at $y = -\infty$ (the inlet) this fluid is uniformly confined between $x = 0$ and $x = 1$, the velocity induced at the inlet by the flame source is

$$v_i = \frac{-\psi'}{2\Delta x} = \frac{-(n-1)}{2}. \tag{47}$$

The vorticity which is confined to the burned gases in the present channel problem induces no velocity at $x = -\infty$. Hence, since the resultant velocity was given as unity,

$$1 = v_0 + v_i. \tag{48}$$

Thus finally the integral equation becomes

$$\frac{n+1}{2}\left[1+\left(\frac{dy_f}{dx_f}\right)^2\right]^{\frac{1}{2}} = (n-1)\int_0^1\left[1+\left(\frac{dy_f'}{dx_f'}\right)^2\right]^{\frac{1}{2}}\left[q_v(x_f,y_f,x_f',y_f') - \frac{dy_f}{dx_f}q_u(x_f,y_f,x_f',y_f')\right]dx_f'$$
$$+\frac{(n+1)}{2S_t} + \int_0^1 dx_0\,\frac{w(x_0,y_0)}{S_t}\left[q_u(x_f,y_f,x_0,y_0) + \frac{dy_f}{dx_f}q_v(x_f,y_f,x_0,y_0)\right]dy_0, \tag{49}$$

where principal values are to be taken of all integrals.

The solution proceeds by first assuming a straight flame front,

$$y_f^0 = \left(\frac{1}{S_t^2}-1\right)^{\frac{1}{2}}x_f^0, \tag{50}$$

where the slope is obtained by supposing the inlet velocity to be unaltered up to the flame. Note that (50) satisfies Eq. (49) if the integrals are ignored.

Now substitute (50) in the first integral on the right of Eq. (49), ignore the second integral, and solve for dy_f'/dx_f'. By integration compute $y_{f_1}'(x_f')$.

In the next approximation it is necessary (in a channel) to include the second integral in ω. To do this we note that $\omega_1 = 0$ while $\omega_2(\psi)$ is not zero but is given by Eq. (25). To find ψ we start with the complex potential for a source and integrate over the flame front and for a vortex and integrate over the products of combustion.

$$\psi = \psi_0 + \psi_1 x - \frac{(n-1)}{2\pi}S_t ip\int_0^{1+iy_f'(1)} \ln\sin\frac{\pi(z-z_0)}{2}\sin\frac{\pi(z-\bar{z}_0)}{2}\,dz_0$$
$$+\frac{1}{2\pi}v\cdot p\int_{\substack{\text{region behind}\\ \text{flame front}}}\omega(x,y)\ln\sin\frac{\pi(z-z_0)}{2}\sin\frac{\pi(z+\bar{z}_0)}{2}\,dx_0dy_0. \tag{51}$$

The constants ψ_0 and ψ_1 are found from the conditions at the inlet, while $\omega(x,y) = \omega_2(x,y)$

is found from the previous approximation. Taking the real and imaginary parts of Eq. (51), as indicated,

$$\psi = \psi_0 + \psi_1 x - \frac{(n-1)S_t}{2\pi} \int_{\substack{\text{along} \\ \text{flame front}}}^{1} \left(\tan^{-1} \frac{\tanh \frac{\pi}{2}(y - y_0)}{\tan \frac{\pi}{2}(x - x_0)} + \tan^{-1} \frac{\tanh \frac{\pi}{2}(y - y_0)}{\tan \frac{\pi}{2}(x + x_0)} \right) \left(1 + \frac{dy_0}{dx_0}\right)^{\frac{1}{2}} dx_0$$

$$+ \frac{1}{2\pi} \int_{\substack{\text{region behind} \\ \text{flame front}}} \omega(x_0, y_0) \ln \frac{\cosh \pi(y - y_0) - \cos \pi(x - x_0)}{\cosh \pi(y - y_0) - \cos \pi(x + x_0)} \, dx_0 \, dy_0. \tag{52}$$

At $y = -\infty$, $\psi = x$, but by Eq. (52)

$$\psi = \psi_0 + \psi_1 x + \frac{(n-1)}{2}(1 - x). \tag{53}$$

Thus

$$\psi_0 = -\frac{n-1}{2}, \quad \psi_1 = \frac{n+1}{2}, \tag{54}$$

in agreement with Eqs. (47) and (48). For the first approximation we find ψ_2' from Eq. (52) by inserting the value of y_f' just found and $\omega_2{}^0 = 0$. Now ω_2 can be computed from y_f', ψ_2' and Eq. (25). An iterative solution now proceeds by replacing Eq. (50) by $y_f'(x')$ and repeating the previous steps. This time, however, $\omega_2{}^1$ is used in the integrals.

In view of the involved nature of the integral-equation attack just outlined it seems desirable to consider also another method based upon difference equations.

For this purpose it is most convenient to introduce the stream function Ψ based upon mass flow. This can be conveniently done in terms of the previous stream function

$$\Psi_1 = \psi_1, \quad \Psi_2 = \frac{\psi_2}{n}. \tag{55}$$

The basic Eqs. (20) and (21) applied to the flame in a channel yield

$$\begin{aligned} \Psi_1{}^1 + \Psi_1{}^2 + \Psi_1{}^3 + \Psi_1{}^4 - 4\Psi_1{}^0 &= 0, \\ \Psi_2{}^1 + \Psi_2{}^2 + \Psi_2{}^3 + \Psi_2{}^4 - 4\Psi_2{}^0 &= 2\delta^2\omega_2{}^0, \end{aligned} \tag{56}$$

where the superscripts refer to values at a point 0 and its surrounding points 1, 2, 3, 4 on a square net of points of spacing δ.

The boundary conditions of Eq. (31), (32), and (33) are now used directly. The boundary condition of Eq. (34) is applied by assuming that $\Psi_2 = x$ at first, and then progressively correcting this by the equation

$$\Psi_2 = x - \frac{x}{n} \int_0^1 d\zeta \int_0^\zeta \omega_2(\eta, \infty) d\eta + \frac{1}{n} \int_0^x d\zeta \int_0^\zeta \omega_2(\eta, \infty) d\eta, \tag{57}$$

which can be used as soon as an approximation to $\omega_2(x, y)$ is available, that is, as soon as a first approximate solution to Ψ and the flame front has been obtained.

In deriving Eq. (57) use was made of the flame internal boundary condition

$$\Psi_2 = \Psi_1, \tag{58}$$

which follows from Eq. (22).

Equations (23) and (24) expressed in finite-difference form in terms of Ψ_1 and Ψ_2 complete the boundary conditions required to find the location of the flame front and the stream-function solution.

Any technique for the solution of the finite-difference system is suitable. The following appears to be a good method of handling the flame-front conditions, for either a relaxation calculation or a computing machine. A square net of points is placed to cover the entire channel. A flame front is placed in the channel by guess. This flame of course falls between net points almost everywhere. Values of Ψ_1 and Ψ_2 are placed at the net points in their respective regions by guess in the usual way. At every net point adjacent to the flame front a value of Ψ_1 and Ψ_2 is placed. Thus in a band of points along the flame front there are values of the stream function Ψ appropriate to both burned and unburned material.

Equations (23) and (24) expressed for pairs of points between which the flame passes provide a relation between the boundary values of Ψ, that is, the value of Ψ_1 on the (2) side and vice versa. During the course of solution of the (1) and (2) regions the flame-front boundary equations become increasingly in error. Periodically, therefore, these equations are used to readjust the boundary values by starting at the center line of the channel where the flame is assumed held and recomputing the Ψ boundary values in steps along the flame front. This is the finite-difference manner of flame-front characteristic propagation.

Finally, the flame front itself is located between net points by use of the equality of Ψ_1 and Ψ_2 at the front, using linear interpolation between nearest points on either side.

The theory of flames herein described is a first attempt to solve the whole aerodynamic problem of combustion in a form that will permit a careful check of the streamlines and flame-front position with those obtained experimentally.

Although the theory is stripped to its barest essentials, it is still of such complexity as to be solvable by the relaxation method only with considerable labor. It is to the computing machine that those wishing to solve nonlinear boundary-value problems must increasingly turn if our growing empirical knowledge is to be supported and guided by a real understanding of the phenomena involved.

Application of Computing Machinery to Research of the Oil Industry

MORRIS MUSKAT

Gulf Research & Development Company

The writer is neither an analyst nor an electronics expert. Nor is he qualified to speak at all about computational problems as such. At the most, he can claim to be playing the role of an interested spectator of the rapidly developing science of computation.

This paper will be unique in this Symposium. In contrast to the others on the program it will contribute nothing to the problem of computing-machinery development. It will not discuss the analytic aspects of specific computing problems, nor exhibit any completed solutions of mathematical equations treated numerically or by large-scale computers. If it will serve any purpose at all beyond fulfilling a promise made to Doctor Aiken, it may be that of stimulating the development of computing services from the point of view of the industrial user as well as that of the computing organization itself.

Having been in the oil industry for 20 years, I might well be expected to be able to discuss authoritatively the subject of computational problems in all phases of the oil industry, as the title of this paper suggests. Unfortunately, however, specialization in the oil industry is as severe as in many other engineering fields, and any one individual must content himself with developing complete knowledge in at most very restricted aspects of the whole industry. In the case of the author, his personal technical activities have been largely confined to the physics of oil production. Nevertheless, for completeness, brief reference will be made to the basic types of mathematical problems arising in other phases of the oil industry, although no attempt will be made to do more than exhibit some of the fundamental equations involved. It is only with respect to the equations arising in the physics of oil production that the author has had direct experience involving the use of large-scale computing machinery.

To the author's knowledge no serious attempt has yet been made to apply large-scale computing machinery to solve the basic problems of geophysical prospecting, oil refining, or lubrication. The reason, of course, is that the fundamental equations underlying these subjects were formulated long before the development of large-scale computing equipment (this term is used throughout the paper to mean digital rather than analog computing machinery) and in those cases where solutions were urgently required direct numerical methods were applied or such approximations were introduced as to make the equations analytically tractable. While undoubtedly the availability of the powerful computing facilities currently being developed will stimulate their application to problems arising in future research, no active interest in immediate applications in the fields of geophysical prospecting, refining,

or lubrication seems to have yet materialized. Accordingly, we shall merely list some of the governing equations pertaining to these fields as indicative of the types of problems that may be proposed for computational analysis in the future.

The three major types of geophysical prospecting having widespread application in the oil industry are known as the gravity, magnetic, and seismic methods. In effect, they are all composed of procedures of making measurements at the surface and inferring from these the nature and geometry of the subsurface rocks which presumably give rise to the surface data. Gravity and magnetic prospecting are both based on potential-theory principles, and their analytic aspects are quite similar in many respects. It will suffice for our present purposes to note merely that the problem of "gravity interpretation" is essentially equivalent to that of solving the integral equation:

$$g_z(x,y) = k\int \frac{z\sigma(\tau)d\tau}{r^3}, \tag{1}$$

where $g_z(x,y)$ is the vertical component of the acceleration due to gravity measured at the surface—the x,y-plane; $\sigma(\tau)$ is the density "anomaly" at the volume element $d\tau$ lying at the depth z below the surface and at the distance r from the origin; and k is a constant. The quantity g_z is to be considered as the "reduced" value of the acceleration due to gravity after correction for surface terrain and the uniform contribution due to an ideal subsurface of constant density.

There is a voluminous literature on the practical solution of Eq. (1) and its analogs for magnetic prospecting by indirect and approximate procedures,[1] and the direct solutions of the integral equations corresponding to simplified forms of Eq. (1), including questions of their uniqueness, have been investigated quite thoroughly.[2] From a practical standpoint, therefore, there seems to be but little urgent need for undertaking additional analysis by large-scale computing equipment.

In seismic prospecting the situation is essentially the same, although there have been very few fundamental investigations of the mathematical aspects of the seismic method. Virtually all procedures for interpreting seismic data are limited to evaluations of the times of travel of the various reflected or refracted waves in terms of depths and velocities of assumed surfaces of discontinuity in the underground strata. The wave equation, as such, plays no direct role in the application of seismic data. It seems unlikely that machine computation will be called on in the study of seismic prospecting, except possibly for long term investigations of phenomena which may arise in media of continuously variable elastic properties.

The term "refining" encompasses such a vast scope of technical activities that no single problem can be properly considered as typical. The theories of catalysis, fractionation, solvent extraction, distillation, and chemical kinetics all provide potential subjects for detailed investigation by machine computational methods. Of these only the last will be exhibited as an illustration. This may be expressed by the set of equations

$$\frac{dN_i}{dt} = \sum_j a_{ij}N_j + \sum_{jk} b_{ijk}N_jN_k + \cdots, \tag{2}$$

describing the homogeneous-phase kinetics of the interactions and transformations between m molecular species of instantaneous concentrations $N_1, \ldots, N_m$ with reaction-rate coefficients $a_{ij}, b_{ijk}, \ldots$ which are to be considered as empirically determinable constants. The $N_i(0)$ are also to be assumed as known.

The nonlinearity of these equations evidently makes the development of general analytic solutions impractical except for extremely specialized simplified cases. While it is doubtful whether the oil industry would support large-scale computing programs for solving these equations as a matter of general interest, it is not inconceivable that special circumstances may arise where the immediate potential applicability of the solutions would warrant the computational treatment of specific sets of these equations. The problem of determining the equilibrium distributions—where $dN_i/dt = 0$—among the N_i for heterogeneous reactions has been given special analytic study in such a form as to facilitate computational treatment,[3] but no similar analysis of the transient problems even for gas reactions has yet been reported.

From an analytic standpoint the only phase of lubrication which has been sufficiently well crystallized to lead to strict mathematical formulation is that known as "thick-film" or "hydrodynamic" lubrication, in contrast to "boundary" lubrication where surface phenomena and interactions are superposed to an important if not predominating degree on the strictly hydrodynamic effects. The basic equation was developed as long ago as 1886 by O. Reynolds. When generalized to include thermal effects on the density and viscosity of the lubricant, though neglecting centrifugal forces and heat transfer by thermal conductivity within the film, two interdependent equations are required, namely,[4]

$$\frac{\partial}{\partial x}\left[\gamma h\left(1-\frac{h^2}{6\mu U}\frac{\partial p}{\partial x}\right)\right]-\frac{\partial}{\partial y}\left(\frac{\gamma h^3}{6\mu U}\frac{\partial p}{\partial y}\right)=0, \tag{3a}$$

$$\left(1-\frac{h^2}{6\mu U}\frac{\partial p}{\partial x}\right)\frac{\partial T}{\partial x}-\frac{h^2}{6\mu U}\frac{\partial p}{\partial y}\frac{\partial T}{\partial y}=\frac{2\mu U}{\gamma C h^2}\left\{1+\frac{h^4}{12\mu^2U^2}\left[\left(\frac{\partial p}{\partial x}\right)^2+\left(\frac{\partial p}{\partial y}\right)^2\right]\right\}, \tag{3b}$$

where p, T are the lubricant-film pressure and temperature at (x,y); γ, μ are the density and viscosity; C is the specific heat; U is the velocity, in the x-direction, of the moving surface; and h is the lubricant-film thickness. When the thermal effects are completely neglected and γ, μ are taken as constant or functions only of the pressure, Eq. (3a) becomes independent of Eq. (3b), which can then be solved, in principle, in sequence. However, even then Eq. (3a) still remains virtually intractable analytically except when the film thickness h is of extremely simple form, such as the thrust-bearing wedge, or when the bearing is assumed to have infinite width ($\partial/\partial y = 0$).

The complexity of Eqs. (3) would suggest that only large-scale machine computation could cope with their solution in a practical manner. However, it has been found[5] that a numerical treatment by relaxation methods is quite feasible even when the thermal effects are taken into account. There are many special lubrication systems for which the specific solutions of Eqs. (3) would be of considerable interest. But in view of the power of the relaxation method it is doubtful whether these will call for the application of large-scale digital computing equipment. The only immediate possibility would appear to lie in the investigation of dynamically

loaded journal bearings when the finite bearing width, film discontinuity, and thermal reactions are taken into account, although when these latter effects are neglected the equations describing the gross dynamical features of the journal-bearing motion can be solved[6] by mechanical integration.

These brief remarks about geophysical prospecting, refining and lubrication should not be interpreted as implying that each is in the status of a closed book. Research in these fields is being vigorously prosecuted. While at the moment computational problems do not constitute major bottlenecks to progress, it may well be that, once computational facilities and services become generally available on a practical basis, many old problems which were previously dropped because they did not warrant the laborious and time-consuming hand calculation and new problems arising in current research will be submitted for machine computation. These phases of the oil industry therefore should not be written off as having no interest in large-scale machine computation from a long-term standpoint.

In the field of oil production the history of the mathematical developments has been marked by a sequence of evolutionary steps. The first serious attempt to treat analytically problems of fluid flow in porous media appears to have been that pertaining to a study of flow of artesian water into well bores, reported[7] in 1863. While this was based on Darcy's law expressing the linearity of the relation between the fluid velocity and the hydraulic gradient, which is the fundamental basis of all viscous-flow phenomena in porous materials, no general formulation was developed. The latter first appeared[8] in 1897 in the form of Laplace's equation for the pressure distribution, namely,

$$\nabla^2 p = 0. \tag{4}$$

In addition to illustrative solutions exhibited in this original work, many others have since been reported[9] for systems simulating in some degree those of interest in oil production. These involved little more than the application of conventional potential-theory techniques.

The scope of problems in fluid flow through porous media governed by Laplace's equation is limited to incompressible liquids fully saturating the porous media. An extension of Darcy's law[10] to gas flow in 1931 led to a nonlinear differential equation, which can be expressed in the form

$$\nabla^2 \gamma^{(1+m)/m} = \frac{(1+m) f \mu \gamma_0^{1/m}}{k} \frac{\partial \gamma}{\partial t}, \tag{5}$$

where γ is the gas density, t the time, f the porosity of the medium, k its permeability, μ the gas viscosity, γ_0 the atmospheric density, and m a quantity that defines the thermodynamic character of the gas expansion.

A further extension[11] of the single-phase fluid theory to the flow of compressible liquids, with constant compressibility, showed that such flow systems could be described by the heat-conduction equation with the liquid density γ as the dependent variable, that is,

$$\nabla^2 \gamma = \frac{f \mu \kappa}{k} \frac{\partial \gamma}{\partial t}, \tag{6}$$

κ being the compressibility of the liquid and f, μ, k having the same meaning as in Eq. (5).

Both Eqs. (5) and (6) have been applied to practical problems.[12] The former, being nonlinear for the transient case, has been treated by approximation methods. And the well-known procedures for solving the Fourier equation, supplemented by direct electrical-circuit analogs,[13] have sufficed for solving a great variety of problems in flow of compressible liquids.

Mathematical problems of a much higher order of complexity arise when the physical situation is generalized to the actual practical conditions obtaining in most oil-producing reservoirs, namely, the simultaneous flow of two or more fluid phases—gas, oil, and water—through the same porous medium. For the case of simultaneous flow of oil and gas in a producing well bore, representing the operation of a "solution gas drive" reservoir, the corresponding equations may be written

$$\left.\begin{aligned} a(p)\frac{\partial}{\partial u}\left[F_1(\rho)b(p)\frac{\partial p}{\partial u}\right] + \frac{\partial}{\partial u}\left[F_2(\rho)e(p)\frac{\partial p}{\partial u}\right] + F_1(\rho)c(p)\left(\frac{\partial p}{\partial u}\right)^2 \\ = e^{2u}\left[f(p) - g(p)\rho\right]\frac{\partial p}{\partial t}, \\ \frac{\partial}{\partial u}\left[F_1(\rho)b(p)\frac{\partial p}{\partial u}\right] = e^{2u}\left[i(p)\frac{\partial \rho}{\partial t} - j(p)\rho\frac{\partial p}{\partial t}\right], \end{aligned}\right\} \tag{7}$$

where p is the fluid pressure, ρ the oil saturation, t the time, and u the logarithm of the radial coördinate. The functions $a(p)$, $b(p)$, . . ., $j(p)$ are to be considered as known functions of p, determined by the thermodynamic properties of the gas and oil, and $F_1(\rho)$, $F_2(\rho)$ as known functions of ρ, reflecting the dynamical characteristics of the porous medium. The quantity ρ itself, which expresses the fraction of the pore space of the rock occupied by oil, may be assumed to have initially a uniform value, less than 1, and must always remain positive and never exceed its initial uniform value. The pressure p likewise may be taken to be uniform initially, and must subsequently always be lower than this value, though positive. At a closed sand body, $\partial p/\partial u$ and $\partial \rho/\partial u$ will vanish. And at the producing well one may impose the history of either p or of the flux: $F_1(\rho)b(p)\,\partial p/\partial u$.

It is this last set of equations that has been the basis of the writer's personal interest in the subject of computing machinery. The equations, in their essential aspects, had been formulated[14] in 1936 on the basis of experimental work done[15] at the Gulf Research & Development Company on the fundamental laws of multiphase fluid flow through porous media. In principle, these hydrodynamic equations, when suitably generalized, govern the whole complex of physical processes underlying the recovery of oil from underground reservoirs. They are evidently too complicated to permit analytic solution. So in order to show that the equations were somewhat more than academic curiosities, a numerical solution for an equivalent simplified linear system was carried through. While the results seemed physically reasonable, they were in no sense precise and had been subjected to considerable smoothing, guided only by physical intuition. However, while this situation was by no means satisfactory,

the six months computing labor required even for the simple ideal system completely discouraged undertaking the analysis of more complex and practical systems.

In lieu of practical methods of solving Eqs. (7) directly, approximations have been introduced. By neglecting the pressure gradients in Eqs. (7) or by an equivalent derivation from first principles, one can obtain[16] an ordinary nonlinear equation of the first order relating ρ and p which suffices to give the gross production history of the reservoir. This has been applied extensively by numerical integration in predicting oil recoveries, the pressure versus oil-recovery history, and the effect of returning the produced gas to the oil-bearing formation.[17] Unfortunately, however, there are a number of important questions relating to oil production which cannot be answered by such simplified treatments, since they pertain to effects of the neglected pressure gradients. Perhaps the two major problems of this type are (1) the effect of the spacing between the producing wells on the ultimate recovery, and (2) the effect of the rate of production on the recovery. Attempts to evaluate these effects have all been beclouded by the uncertainty whether the approximations that have been made have not automatically predetermined the quantitative aspects of the conclusions. Yet well spacing and production rates are among the most important parameters that are subject to the choice of the operator in controlling the ultimate oil recoveries.

Except for the original attempt at hand calculation already referred to, the problem of solving Eqs. (7) directly remained dormant until July 1946, when an announcement appeared of the war development of the ENIAC. Negotiations with the government were then entered into for applying the ENIAC to the solution of these equations. Although several plans were developed over a period of more than a year for carrying out this project, it was not found feasible, because of legal difficulties, to arrange for the required coöperative effort between the government and an industrial concern. While this situation was subsequently resolved by one of the governmental agencies becoming interested in the problem and assuming sponsorship for the work, it was ultimately found, much to the embarrassment of the author, that the project had to be abandoned anyway because the memory capacity of the ENIAC would not suffice for handling the large number of operational orders required.

This unhappy history is referred to here to serve as an illustration of what *can* happen when one unfamiliar with the science of computation and its ramifications is left to the mercy of his own naïve optimism. The writer has learned the "hard way" that there is more to the computational solution of complex equations than the desire to have them solved. We shall discuss this matter further below.

To complete the record, following the realization that the ENIAC was not sufficiently powerful to solve Eqs. (7), the equations were submitted to the International Business Machines Corporation for their consideration. After a number of preliminary discussions, the IBM Corporation undertook to place the problem on the Selective Sequence Electronic Calculator. This work is still in progress. Needless to note, this project is being given a thorough preliminary analytic formulation by the IBM staff prior to final machine computation.

With respect to the general field of the physics of oil production, it should be noted that

Eqs. (7) themselves represent highly simplified systems in which it is assumed that the oil-producing rocks are everywhere uniform. Such reservoirs actually never occur in practice. While the development of the implications of Eqs. (7) would in itself be a constructive accomplishment, it will ultimately be of considerable interest to investigate their generalization to nonuniform systems. Moreover, effects of gravity have been ignored in constructing Eqs. (7). Their inclusion would make the physical problem three-dimensional, which would lead to an additional order of complexity. And even aside from studies of the gravity effects, the investigation of the three-dimensional analogs of Eqs. (7) will be of importance in treating stratified producing systems with mutual cross flow. Finally, Eqs. (7) take no account of interfacial capillary phenomena, which may be of importance under special conditions, and especially when gravity is an important factor in the producing operations.

In fact, the detailed study of Eqs. (7) constitutes only a beginning in the establishment of the quantitative aspects of what is now generally known as "reservoir engineering." Even without anticipating new developments in this field as research continues, it is clear that large-scale computing machinery will find wide and important applications in oil production for many years to come. And it is not inconceivable that while only a passive interest in extensive digital computation has thus far developed in other branches of the oil industry, comparable applications in refining and geophysical prospecting may ultimately be found once the practical availability of these powerful tools becomes disseminated throughout these other fields of activity.

A single and obviously unique experience is a dangerous basis for generalization. The following remarks are not to be construed as direct implications of the above outlined personal contact of the author with problems of computation. On the other hand, the program of this Symposium itself is evidence that outside of government organizations and academic institutions the application of computing equipment to the solution of specific problems apparently has thus far been rather fragmentary. It therefore seems appropriate to explore the general subject of computing-machinery service for industrial applications, even though much of the discussion must be of a speculative character.

The computing-machinery service to be considered here is that which would require the use of large-scale equipment localized at computing centers such as the Computation Laboratory at Harvard, the IBM Corporation, the Bureau of Standards, and similar organizations which may provide their facilities, at least in part, for the investigation of industrial problems. The specific question involved is essentially that of defining the term "service."

There are two aspects of the composite problem of application of computing equipment about which there will be little question. The first is that the one who is primarily interested in the solution must provide both the analytic and the physical statements of the problem. Second, the computing-service organization must carry out both the actual machine operation and the coding of the problem. It is in the intermediate coupling of these two contributions that the situation remains uncertain. And it is in this link that the efficiency and value of the

computing project may be ultimately determined. It is here that control may be applied on the accuracy of the solution, and often on its physical reality and convergence. It is the writer's belief that this bridge of analytic programming should be made available, when necessary, by the computational organization.

It may well appear that the very program of this Symposium belies the suggestion that analytic preparation and programming should be a part of computational services. For many of the papers presented here report on investigations in which those with whom the problems originated carried through all aspects of the problem short only of the machine operations themselves. You will note, however, that in almost all cases the authors represent academic or similar research institutions. The same may be expected with respect to many problems arising in the aircraft, automobile, shipbuilding, explosives, telephone, and railroad industries, or in the larger companies in the electrical, steel, radio, and glass industries. By their very nature the major industrial concerns in these fields require virtually self-contained large technical staffs capable of handling all phases of their engineering activities. However, in spite of the great contribution to our total industrial effort made by such organizations, by far the larger part of industry as a whole is comprised of the composite resultant of the hundreds of intermediate- and small-sized concerns engaged in some form of technical activity.

In their own specialized fields the engineering problems of these companies are essentially the same as those encountered by the large corporations. Yet in contrast to the latter they cannot afford to maintain the permanent, complete, and well-rounded research organizations which can attack effectively virtually any problem that may arise. In particular, with respect to mathematical problems, or such where analytic treatment may be required at least to guide experimental research or design, these smaller firms may be fortunate if their engineers have enough mathematical background merely to construct the equations to be solved. Of the members of the American Mathematical Society who gave their employment affiliation on the membership list, fewer than 325, or 9.0 percent, indicated connections with industrial concerns, including those who have a direct interest in the development of computing machinery.

As the writer himself has learned by painful experience, and as any "outsider" attending this Symposium or meetings of the Association for Computing Machinery would quickly observe, the science of computation is a highly specialized technical field. In many respects it is still in its infancy—a war baby—but it is growing with accelerating speed. The practicing engineer or physicist of today literally heard nothing of it during his academic training. The terms coding, programming, the binary system, and many others that are commonplace in the language of the modern computation science are quite foreign to those in the engineering professions.

Among those on the membership list of the Association for Computing Machinery who have given their employment affiliation more than 82 percent are in government agencies, on academic or research institute staffs, in computing organizations, or are employed by industrial concerns that are obviously engaged in some phase of computing-machinery

development. More than half of the remainder are employed by aircraft and insurance concerns, and very probably a number of the 38 "residuals" also are primarily interested in the equipment development itself. It is thus clear that to the extent that membership in the Association is an index of interest and contact with the computing profession, such interest has yet been disseminated but slightly into industry as a whole.

If the engineer or industrial physicist or chemist in the average commercial firm must stop to take a training course in the theory of computational machinery, even if he should be temperamentally suited to absorb such specialized disciplines, before having his problem accepted by the computing organization, the probability is great that he will drop or circumvent the problem. And even if he were willing and could make arrangements to "study up" on the basic elements involved, it is still very unlikely that he would thus develop the required analytic skill to guide the choice of the mesh to be used, decide on the differencing procedures that may be required for convergence, carry through the preliminary numerical solutions, or even prepare functional representations for the empirically variable functions in his equations, if this should be necessary.

At present most of the applications of computing equipment are being made by members of governmental agencies or academic institutions. Many of these are well staffed for analytic work. Moreover, in spite of the importance of these problems, it is doubtful whether the pressure for their speedy solution is comparable to that in industry, where diversions into the purely computational aspects of the problem may not be accepted without prejudice. Undoubtedly, there is a large backlog of demands by such organizations for the use of presently operating computing machinery. So there may appear to be no need to cater to and accept computation proposals from those who are unprepared or unable to submit a completely programmed problem. Such, however, it is believed, would be a shortsighted policy and would lessen the long-term possibilities of growth of the science of computation.

It is not suggested that the argument is one-sided. No doubt the provision of this type of service by computing organizations will involve difficult personnel problems, though these same difficulties would be even more serious in most industrial firms. It is also true that such service would increase the total charges, which might discourage the interest in them by small concerns with very limited engineering development budgets. But at the same time it would make it possible to extend the applications of large-scale computing equipment to many organizations that would otherwise simply have to give up because of lack of qualifications.

Perhaps the strongest reason for centering the intermediate analytic facilities within the computing organizations lies in the importance of experience in this phase of numerical computation. In a science as young as this, virtually each problem gives rise to new questions of detailed treatment. It is not yet ready for standardization and the preparation of tabulated instructions. The analyst who is continually engaged in programming will no doubt accumulate a wealth of experience which will be of inestimable value both to the computing organization and to its clients. To have each problem prepared and analyzed for computation by a beginner will be pitifully inefficient as compared to their handling by personnel for whom

such work is their daily professional business. In fact, the writer ventures to predict that if and when the "buyer's market" overtakes the computing industry the burden of selling computing services will fall on competitive claims of the *experience* of the organization and the *completeness* of the service rather than on the number of milliseconds the machines take for a multiplication or whether the price is x or $x - \Delta x$ dollars per hour.

It is to be understood, of course, that even if the computing organization provides the analytic preparation of the problem the sponsor must still accept the responsibility of interpretation and evaluation of the solutions. Except possibly when solving purely arithmetic problems, as systems of algebraic equations or function-table preparation, it is the sponsor who must supply the guidance in dropping terms if such should be necessary to make the problem tractable, in fixing the order of accuracy required, and in evaluating the physical significance of the solution. This may well call for visits by the sponsor to the computing organization during the planning, programming, and coding, and in most cases his continuous presence there during the time the problem is actually on the machine. Indeed, the experience and background of the sponsor in the technical field giving rise to the problem may be just as indispensable in achieving a satisfactory solution as the experience of the computing staff with respect to its analytic aspects.

There is no easy way to accomplish difficult tasks. Coöperative effort by all parties concerned is required. The ultimate impact of the science of computation on our technology and industrial life will most certainly be tremendous compared to what has already materialized and what can now be envisioned. It has already evoked an absorbing interest from and recruited into its ranks some of the outstanding leaders in the fields of engineering and electronic design and mathematical analysis. Let us therefore plan to guide this important growing effort so that its fruits may be enjoyed by the maximum number for the greatest benefit of our nation as a whole.

REFERENCES

1. See, for example, L. L. Nettleton, *Geophysical prospecting for oil* (McGraw-Hill Book Co., New York, 1940).

2. H. M. Evjen, *Geophysics* **1,** 127 (1936); H. Bateman, *J. App. Phys.* **17,** 91 (1946); E. C. Bullard and R. I. B. Cooper, *Proc. Roy. Soc. (London)* [*A*] **194,** 332 (1948); G. Kreisel, *Proc. Roy. Soc. (London)* [*A*] **197,** 160 (1949); L. J. Peters, *Geophysics* **14,** 290 (1949).

3. S. R. Brinkley, Jr., *J. Chem. Phys.* **14,** 563 (1946); **15,** 107 (1947).

4. See W. F. Cope, *Proc. Roy. Soc. (London)* [*A*] **197,** 201 (1949).

5. D. G. Christopherson, *Proc. Inst. Mech. Engrs. (London)* **146,** 126 (1942).

6. J. T. Burwell, *J. Applied Mechanics* **14,** A–231 (1947).

7. J. Dupuit, *Etudes théoriques et pratiques sur le mouvement des eaux* (1863).

8. C. S. Slichter, U.S.G.S. 19th Annual Report (1897–98).

9. M. Muskat, *Flow of homogeneous fluids through porous media* (McGraw-Hill Book Co., New York, 1937; J. W. Edwards, 1946).

10. M. Muskat and H. G. Botset, *Physics* **1,** 27 (1931).

11. T. V. Moore, R. J. Schilthuis, and W. Hurst, *Oil Weekly* **69,** 19 (May 22, 1933); M. Muskat, *Physics* **5,** 71 (1934).

12. M. Muskat, reference 9.

13. V. Paschkis and H. D. Baker, *AIME Trans.* **64,** 105 (1942); W. A. Bruce, *AIME Trans.* **151,** 112 (1943).

14. M. Muskat and M. W. Meres, *Physics* **7,** 346 (1936).

15. R. D. Wyckoff and H. G. Botset, *Physics* **7,** 325 (1936).

16. M. Muskat, *J. Applied Physics,* **16,** 147 (1945).

17. M. Muskat, *Physical principles of oil production* (McGraw-Hill Book Co., New York, 1949).

The 603–405 Computer

WILLIAM W. WOODBURY

Northrop Aircraft, Inc.

For the past year and a half Northrop Aircraft, Inc. has operated a computing machine built by the International Business Machines Corporation, which differs radically in its treatment of problems from the usual computing installation of IBM accounting machines. This machine consists of three standard machines—a Model 405 printer, a Model 603 electronic multiplier, and a Model 517 summary punch. They are interconnected to function as a single unit.

The printer, commonly known as a tabulator, is a machine of parts. Two sets of brushes which read punched cards are the machine's point of entry. Each card is read consecutively, first by the "upper" brushes, then by the "lower" brushes. A card has room for 80 decimal digits, indicated by punching. A number is represented on a card by the vertical position of punches in the 80 columns. It is represented in the machine by the relative time at which the brushes make contact through these punched holes. All machine elements are synchronized to the movement of a card past both sets of brushes. The card actually has not ten, but 12 vertical positions. Ten are used for digits, while the remaining two are used principally for algebraic signs, or for control of switches (selectors) which will be described below.

The machine has a counter capacity of 80 decimal digits. These are arranged in groups —four each of 2, 4, 6, and 8 digits—but the groups may be combined to produce individual accumulators up to 80 decimal digits. Eighty-seven type bars, through which any information in the machine may be printed in a single cycle of the machine's operation, are also important. A detachable plug board may be wired according to the arrangement of counters and type bars desired for a particular problem. This is a convenient feature of the machine, since several of these may be wired for various problems in advance, thus permitting immediate change-over as soon as a problem is finished. Auxiliary equipment includes six 10-pole and 16 single-pole double-throw switches or selectors; 20 positions for numerical comparison which yield impulses if the numbers entered are unequal; and two distributors for separating impulses in a circuit with respect to time.

The items mentioned above represent regularly available accessories to the machine. Special additions have also been made, such as the multiplier entry, exit, and control connections which will be described in connection with the multiplier. There are 16 8-pole, four 4-pole, and 40 single-pole double-throw switches, plus five 8-pole quadruple-throw switches called chain selectors. A second plug board is provided for wiring these additional elements.

The multiplier, an electronic device, develops a 12-digit product from two 6-digit factors. These factors are entered at the same time that the product from the previous entry is read

out. The multiplication is executed during the time between cards. Since the tabulator does its adding and subtracting in the "nines complement" system, provision has been made for reversing the factor entries in time when a counter is standing in complement. For example, take the number 999998. Counting forward in time to the number yields 999998. Counting from the number to 999999, which is equivalent to zero for this machine, gives — 1. If 2 is added and the leftmost position carry entered in the units position, the correct answer, + 1, is obtained. Since the absolute value of the product is developed, provision is made for a negative-sign impulse when a negative product is read out. This impulse reverses the add or subtract instruction given to the receiving counter. Provision is also made for round-off, by addition of five in the leftmost position dropped. The multiplier control is logically complete; numbers may be entered from any source, either as complements or as absolute values with signs; and the product with its sign may be taken to any place, including reëntry as a factor of a succeeding product.

The remaining machine—the summary punch—provides a means of punching the information from the counters on cards.

It is now possible to compare this machine to the present conception of an adequate all-purpose computer. It has an arithmetic organ—the tabulator counters plus the multiplier. It has an input and two outputs, one of which is the input medium. An internal memory is achieved by apportioning a part of the 80 counters to memory. The external memory, in the form of punched cards, is indefinitely large. Control could be established from the counters. This, however, is not expedient. The control instructions must be punched into cards for entry into the machine, and are just as well left there. Besides, the limited internal memory capacity will hold only a trivial program. Programs are sometimes wired implicitly, however, so that only blank cards need be fed after the initial data have been entered.

The general-purpose computer is presently conceived of as being organized around one or two channels. These may be either serial, in which case words are moved about digit by digit, or parallel, where the entire word is moved at once. This machine has no channel as such, but its array of switches is used to construct the channels best suited to the problem at hand. This frequently permits a kind of multiple-parallel operation, in which several computations are made simultaneously. A table look-up operation from the control cards may be channeled into one counter, while higher derivatives are being integrated (Δt = a power of ten), and while the multiplier and a counter or two are iterating for a square root. This kind of operation is commonly performed in actual problems. The machine's speed is determined by the card-feeding rate. When no output is required, it accomplishes 150 cycles/min, performing one multiplication and one or more additions or transfers in 400 msec. When transactions must be printed, the speed drops to 75 cycles/min. In this kind of printing cycle, called a list cycle, the operations mentioned above can be performed in 800 msec. Another kind of cycle prints the contents of the counters, and may or may not clear the counters. The duration of this cycle is equal to that of the list cycle, but no computation can be performed while it is in progress. This cycle is called a total cycle.

In order to punch cards, the machine must be stopped for approximately 1 sec. The design of the machine makes it necessary to take a total cycle at this time.

One limitation on the speed of the machine is that it can perform only one multiplication in a cycle. Thus, problems involving many multiplications but few other operations take more time than problems containing relatively few multiplications.

The direct-printing feature is worthy of emphasis. It involves no subsidiary machines. It does not await manual operations. It simply prints whatever is in or passing through the machine when the machine is so instructed. In trouble-shooting and in checking, the results of previous cycles are there for comparison with the result of the present cycle. When a problem is complicated, intermediate results may be printed at will, to provide a picture of the relative magnitude of various factors. When final answers only are required, printing can be restricted to these answers. In the inching process—that is, printing every cycle—the exact nature of errors, not only in the program but frequently in the mathematical formulation, are immediately apparent. Indeed, we dispense almost completely with checking at the transcription level, since errors are discovered so easily in this manner.

This brings up an interesting point. On many problems this machine produces results about as fast as they can be apprehended. That is, for a problem of an investigative nature wherein various configurations are to be tried, additional speed is not so desirable as another machine when someone else wishes to work with another problem. It is true that much work is done on this machine that is of the nature of tabulation of functions. For the moment, no one is much concerned with the development of the answers. In this work great speed would be an advantage.

A résumé of the kind and scope of problems with which the writer has had experience follows, for those concerned with the industrial application of this equipment. The computer was built to integrate a system of six nonlinear differential equations in a single independent variable. These equations were of the first order in four of the dependent variables, and of the second order in the other two dependent variables. In addition, the sine and the cosine of one of the dependent variables entered four of the equations as coefficients. Since the continuity of the solution was good, the sine and cosine were integrated stepwise along with the equations themselves. A second system of four nonlinear second-order equations was integrated, with the interesting program variation generated through a relation between two of the dependent variables: $x/(x^2 + y^2)^{\frac{1}{2}}$. This expression was evaluated through a table look-up operation and, because of the wide variations in x and y, was done with a floating decimal point. The above equations all represent work in connection with servomechanisms having several degrees of freedom, and with cross-product terms of considerable magnitude. Stochastic processes have been part of the bread-and-butter work for the machine and are especially adapted to it because of the multiple-channel operation and the further possibility of making several simultaneous discriminations for future choices, even as the consequences of the last choice are being computed. Run-of-the-mill work has involved the reduction of test data and structural analysis. In an aircraft company, test data mean wind-tunnel and strain-gage

information, which may be programmed to completion in one machine passage. Structural analysis has been limited by the need of the machine for other more pressing problems, but the machine is quite capable of handling problems in this field.

An investigation of the behavior of the biharmonic difference equation for cantilever plates has been going on, as time has been available. I would like to be able to give some definitive results from this investigation, but possibilities remain to be explored. It seems probable, however, that the number of computations required for a reasonable convergence of an iterative process is of considerably higher order than the number of computations required to invert the matrix of the points, the ratio being possibly somewhere in the neighborhood of n^2, where n is the number of points in the lattice. This is for an unsophisticated pattern which simply substitutes the new value for a point when it is obtained. Convergence could probably be improved by second-order corrections, but I doubt whether it could be improved enough to equal matrix-inversion speed. The inversion of high-order matrices using an elimination algorithm requires about $(n/20)^3$ working days. This time is slow and this operation one of the weakest for the machine because of the preponderance of multiplication. Each multiplication, incidentally, involves 3 cycles to retain sufficient accuracy.

The work with the biharmonic equation indicates a limit of application. Unless results are of sufficient value to justify the expenditure of several months' time, $\nabla^4\varphi = -q$ with three free boundaries is beyond the machine's power. With respect to the simpler harmonic equation, we hardly feel ready to compete with the relaxation technique described by Southwell.

Checking is accomplished in various ways, according to the problem. In integrating differential equations, continuity of the solution is often a sufficient check. For final structures reports, when balances are not available as a check, duplicate runs are made and compared mechanically. Sample calculations are made on a desk calculator in order to avoid systematic errors. The machine is operated 24 hours a day, 5 days a week, and has averaged about 10 percent down time.

In conclusion, I wish to say that this card-controlled computer was thrown together in about four weeks to meet a need for a powerful computer to do a complicated integration. It seems to fill a useful place in its ability to integrate differential equations in a single independent variable and to do routine calculations involved in engineering design work with an over-all efficiency of better than ten times that of any other generally available equipment. We have here a machine different in nature from most computers. It can perform a multiplication per cycle and several additions and transfers simultaneously. Thus, it is more efficient in use of its rate than other computers for which each operation is exclusive. This machine never has to wait to find out what to do next. Even if what it is to do next is dependent on the solution so far, this is readily incorporated in the wiring through the use of a selector, so that no time is lost. One is led to feel that as the clock rate of an internally programmed machine is increased it should be easy to increase the input rate of the externally programmed machine, so that the time loss inherent in program operations is still large. This is to emphasize the time cost of internal operations upon program instructions. The work with the biharmonic

equation suggests that partial differential equations will require far too much time for the iterative solution which can be accomplished with a relatively small high-speed memory. This indicates that the memory capacity should be based on the storage requirement for inverting large matrices, which is of the order of n^2 words where the matrix is $n \times n$. Less storage than this will require continuing use of the input and output, with the consequent loss of time. From these considerations we at Northrop who are close to this work feel that the internally programmed machine will require perhaps ten times as much high-speed memory as has been considered to date to our knowledge, the possibility of executing program revisions simultaneously with explicit computation, and the elimination of access time through the use of multiple registers.

SEVENTH SESSION

Friday, September 16, 1949
9:00 A.M. to 12:00 P.M.

THE ECONOMIC AND SOCIAL SCIENCES

Presiding

Edwin B. Wilson
Office of Naval Research

Application of Computing Machinery to the Solution of Problems of the Social Sciences

FREDERICK MOSTELLER

Harvard University

This paper discusses some of the applications and limitations of the use of modern computing machinery in the social sciences. Such a discussion could scarcely be expected to be exhaustive, but merely indicative of the kinds of applications occurring now and in the near future. Of course, some remarks must be included about the use of computing machines in social science's principal quantitative tool, statistics. We have not included economics in the social sciences because the title of the program—the Economic and Social Sciences—indicates that these fields are to be considered separately. For our purposes the social sciences might be regarded as including education, social psychology, and sociology. It would be a tour de force at this time to include cultural anthropology, history, political science, and similar largely nonquantitative subjects, in a discussion of modern computing machinery.

Thus far, most direct applications of computing machinery to social-science problems are associated with routine problems of solving simultaneous linear equations, either homogeneous or nonhomogeneous. The commonest and most widely applied technique is multiple regression. Here we have one dependent or criterion variable Y which we desire to predict from a flock of independent variables $X_1, X_2, \ldots, X_k$. The standard approach is by means of least squares, where we are required to find weights a_i to minimize the function

$$\sum_{j=1}^{N} (y_j - \sum_{i=0}^{k} a_i x_{ij})^2, \quad x_{0j} = 1. \tag{1}$$

In the relation (1) the subscript j refers to the observations. This well-known minimization produces a set of k simultaneous linear nonhomogeneous equations which we solve for the weights a_i. From this solution we get a linear prediction equation

$$Y = \sum_{i=0}^{k} a_i X_i, \; X_0 = 1. \tag{2}$$

We are often asked if it would not be better to try to fit some function of the X's to Y. It certainly would, but we do not ordinarily know the function. The reasonable thing, therefore, is to take the plane given in Eq. (2) as a first approximation to this function and to hope that the range of the variables is sufficiently small that this method will be adequate for predictive purposes. If we go to the quadratic approximation, we will have $k(k + 1)/2$ additional terms to fit. Even if k is as small as 6, we will have $7 + 21 = 28$ simultaneous equations to solve for the second approximation. The resulting reduction of the residual sum of squares is seldom

worth the work. A more common device is to adjust the scale on which variables are measured to make the linearity assumption of Eq. (2) more realistic.

The applications of computing machines in the above examples are rather obvious. First, of course, we want the equations solved. But second, and perhaps more important, we want to know how much faith we can put in such an equation. We must remember that the observations ordinarily are good to only one or two significant figures.

A problem closely related to multiple regression is the discriminant function. In its simplest form we are asked to divide a population into two groups. Whereas in multiple regression we might be asked to predict degree of marital success, or degree of adjustment of a paroled man to the outside world, or degree of success as a pilot, in the case of the discriminant function we are asked to produce a function that will separate sheep from goats. Will the postulated marriage end in divorce; if we parole this man will he return to jail; will the candidate get his wings or not? A little further afield, we may even ask whether Alexander Hamilton wrote this particular essay from the *Federalist Papers*—or was it James Madison?

The distinction between multiple regression and the discriminant function, then, lies in the nature of the dependent variable. In the case of the discriminant function it is dichotomous. We want to construct an index number

$$Z = \sum_{i=1}^{k} \lambda_i X_i \tag{3}$$

and establish a criterion number C, so that according as $Z \gtrless C$ we can predict successful marriage or divorce, good citizenship or recidivism, Hamilton or Madison, with a reasonable percentage of success.

Another way of looking at this problem is that we want to find λ's such that we can maximize

$$G = \frac{(\bar{z}_1 - \bar{z}_2)^2}{\sum_{i=1}^{2} \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2}, \tag{4}$$

where $\bar{z}_1$ and $\bar{z}_2$ are the Z means of the success and failure groups, and z_{ij}, $j = 1, \ldots, n_i$, $i = 1,2$ are the Z values for particular individuals. In other words, we want to maximize the ratio of the between-groups square to the within-groups sum of squares. The numerator of G measures the separation of the groups; the denominator measures the variabilities of the groups within themselves. This method of looking at the problem is chosen because of a connection with a later problem. It turns out, after some manipulation due to R. A. Fisher, that the λ's will be obtained by solving the equation

$$\sum_q \lambda_q S_{pq} = cd_p, \quad p = 1, \ldots, k \tag{5}$$

where

$$d_p = \bar{x}_{p1} - \bar{x}_{p2},$$

$$S_{pq} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (x_{pij} - \bar{x}_{pi})(x_{qij} - \bar{x}_{qi}), \tag{6}$$

and c is an arbitrary nonzero constant. Here the d's are the distances between the means

of the two groups on the independent variables and the S's are weighted covariances between pairs of independent variables summed for the two groups. As in multiple regression, we are left with simultaneous linear nonhomogeneous equations to solve (c is arbitrary). Once this is done we can compute the Z values for each group and discuss the effects of choosing various values of the cutoff point C. How we choose this point will depend on the costs of making wrong decisions of either kind. If plenty of pilot candidates are available, and training costs are high, we will make the cutoff on the index quite high to reduce the washout percentage, realizing that we are discarding numerous candidates who would have made good pilots.

Up to now, little has been done about discriminant functions when it is desired to split the population into three or more groups. This lack of progress may be due partly to the very heavy computational work that would undoubtedly be associated with a decent formulation of the problem. As modern computing machinery becomes available to scientists, it is likely that they will no longer be so reluctant to formulate problems that require heavy computation.

An example of the application of homogeneous equations is supplied by Guttman's scaling theory. One such problem is that of scaling attitudes. More detailed expositions of this problem are given in Paul Horst, *The Prediction of Personal Adjustment* (Bulletin 48, Social Science Research Council, 230 Park Avenue, New York, 1941), and *The American Soldier*, vol. IV (Princeton University Press, Princeton, N.J., to be published shortly). We will restrict ourselves to dichotomous questions (answer yes or no) for the explanation. If we have six such questions, they would form a perfect Guttman scale if the responses to all the questions by all respondents could be arranged into one of the six forms shown in Table 1. In this table,

Table 1. Guttman scale for the responses to six questions.

	Question					
	1	2	3	4	5	6
Favorable	X	X	X	X	X	X
	0	X	X	X	X	X
	0	0	X	X	X	X
	0	0	0	X	X	X
	0	0	0	0	X	X
	0	0	0	0	0	X
Unfavorable	0	0	0	0	0	0

X corresponds to Yes and 0 to No. The numbers attached to the questions are dummies. If we could achieve such a perfect state of affairs we would clearly have formed a scale on the favorable–unfavorable axis which could be thought of in Steven's classification as ordinal.

The direction of the scale is determined by the content of the questions. The perfection displayed in Table 1 can scarcely be expected in practice. Therefore we request that scores be assigned to individual patterns of responses to accomplish this ordering of patterns of responses as nearly as possible. The criterion used is that we should maximize a certain correlation ratio. This maximization leads to a set of simultaneous homogeneous equations. Actually there is a perfectly decent and workable approximation scheme (called the scalogram method) that can be used to get the initial rankings of the people and the questions, and we could usually avoid computation in practical applications were it not for some further developments. The

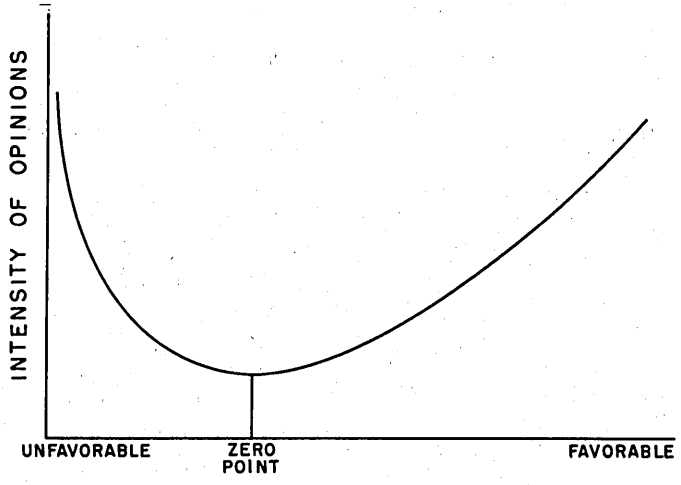


FIG. 1. Diagram showing schematically a curve of intensity as measured by the strength with which opinions expressed are held, plotted against the score as obtained from a Guttman scale. It is conjectured that the lowest point on the intensity curve corresponds to neutrality and should be regarded as the psychological zero point. If the first component is regarded as the score on the attitude scale, and the score on the second component is plotted against the first, similar *U*-shaped curves appear, with minima quite close to those of the intensity curve.

scoring that is achieved represents mathematically the principal component of the system. Since there are more components available, and since these have been found to have meanings in other fields of endeavor, it is not unreasonable for the psychologist to wonder whether these further components might not have further meaning for him. In particular, the second component has been found in some attitude studies to correlate extremely well with the concept of intensity, where intensity has a separate definition. More recently, Guttman has worked on a possible interpretation of the third component. I must admit that I take a rather different view of these components and that I feel it is rather a fortuitous accident that intensity is closely related to the second component. Intensity with which an opinion is held, as it is ordinarily defined, leads to *U*-shaped functions when graphed against the favorable–unfavorable scale (see Fig. 1). In so far as the first component is arranged in a roughly linear fashion against

this scale, the second component, which is orthogonal to the first, must be rather *U*-shaped. Considering the reliability of the observations, all *U*-shaped functions look pretty much alike. We need not try to decide this matter here. The main point is that social scientists are interested in these further components, but we have no very good practical way to get them except by direct computation. When the questions are numerous, as they often are, the work requires heavy computation.

If we agree to identify the second component with intensity, it is possible to get at a zero point on an attitude scale by agreeing to take the lowest point on the intensity scale as determining the score on the first component, which will be regarded as neutral (see Fig. 1).

It might be useful to indicate a type of scale, not unlike Guttman's, in which it is easier to explain the criterion for obtaining the scores. We might take k attitude items on a special topic and ask the subject to endorse the r that come closest to his opinions. Out of such an experiment we would ideally obtain a set of responses like those in Table 2. In this example

Table 2. A second type of attitude scale.

	Item							
	1	2	3	4	5	6	7	8
Favorable	*X*	*X*	*X*	0	0	0	0	0
	0	*X*	*X*	*X*	0	0	0	0
	0	0	*X*	*X*	*X*	0	0	0
	0	0	0	*X*	*X*	*X*	0	0
	0	0	0	0	*X*	*X*	*X*	0
Unfavorable	0	0	0	0	0	*X*	*X*	*X*

$k = 8$, $r = 3$. If individuals chose only the response patterns indicated above, we would have a perfect scale. Actually, there will be response patterns with gaps between the checked items. We formulate the problem this way. We want to assign weights to the items so that when an individual t chooses items i, j, k, we can give him the score $S_t = W_i + W_j + W_k$. As our criterion we take the ratio of the variability of the scores S_t to the variability of the weights making up a score, the latter summed over the individuals. This view of the situation is entirely analogous to the criterion given earlier for Fisher's discriminant function. From the point of view of analysis of variance, we want to maximize the ratio of the sum of squares between individuals to the total sum of squares (because there is an additive relation between "between individuals," "within individuals," and "total sum of squares"). This ratio of "between" to "total" is proportional to the correlation ratio. If we try to maximize this ratio we are led again to solutions of homogeneous linear equations. If the number of items is

large, we have a long computational problem. The method just suggested is in some ways a variation of Thurstone's method of equal-appearing intervals.

Similar problems arise in education. For example, we may have letter grades in four courses for a number of individuals. We would like to pool these letter grades to form a scale of scholastic achievement. However, the distributions of the grades in the several courses are quite different. We want to assign numerical values to the grades in the different courses, and then add these to get a score for the individual. The problem is not unlike the one just treated, except that for each course (item) we need several weights.

The problems discussed above are common to many of the fields of social science: sociology, education, social psychology, and perhaps even economics. We could continue to multiply these examples from scaling theory without difficulty. We have not touched on the problem of factor analysis—the attempt to find the meaningful psychological or sociological dimensions of a space of test scores, while reducing the dimensionality of the space—although these problems are again concerned largely with matrix manipulation. Nor have we discussed the analysis of time series. However, time-series problems are so general in all sciences these days that social scientists can expect generous contributions on this problem from their more mathematically minded friends in the natural sciences.

To dream a little, I think that certain social problems may be capable of being formulated in terms of game theory. Then certainly computing machines will be useful, but this application waits on two developments—first, the ability to describe a social problem in terms of a game, and second, the development of good methods of finding solutions to games. I have no doubt that progress on the second problem will be more rapid than on the first. Similarly, the application of the computing machine as a model for certain problems in clinical psychology seems to me extremely speculative at this time.

We move from direct to indirect applications of computing machinery in the social sciences when we discuss problems in theoretical statistics. I would like to call a few of these to the attention of computing experts. Both theoretical and practical reasons make the normal distribution one of the most important of all distributions. Therefore, estimates of its parameters from samples is a constant problem. For a long time the view was held that efficient statistics (in a technical sense) for estimating parameters were the best ones to use. Efficiency (or relative precision) of two unbiased estimates of the same parameter is measured by the ratio of the variances of the two computing estimates; it is the ratio of the smaller variance to the larger variance. However, it has turned out that efficient statistics are not always the easiest ones to compute.

It has been found that a few carefully selected observations from a large sample can produce extremely good estimates of the mean and the standard deviation with little calculation. Similarly, in very small samples it turns out that little efficiency is lost by estimating the mean from the average of the largest and smallest values, and that the standard deviation can be very adequately estimated from the range instead of from the cumbersome root-mean-square.

The result of these practical findings has been an interest in order statistics. If we draw a sample from a distribution and order the n observations from least to greatest,

$$x_1 \leq x_2 \leq \cdots \leq x_i \leq \cdots \leq x_n, \tag{7}$$

then x_i is called the ith order statistic. Statistics constructed from these order statistics—for example, range, median—are called systematic statistics when they take cognizance of the order (the mean does not). In studying the worth of these systematic statistics, it is of the greatest interest to know certain properties of the order statistics. In particular, we wish to know for the normal distribution the mean, the variance of any order statistic, and the covariances between pairs.

For microstatistics ($n \leq 10$) we have good tables of these quantities. The first attempt to get covariances by numerical integration resulted in two-decimal accuracy in spite of eight-decimal initial values. The latest attempt is much improved (five-decimal accuracy) because of the discovery of a method of exact integration which works up to $n = 10$, but does not seem to want to go further. For macrostatistics (say $n > 100$) we are in fairly decent shape with asymptotic theory helping us. However in the middle range ($100 \geq n > 10$) we are in trouble. This is a fairly standard situation in statistics, the middle-sized samples causing us considerable worry because it is not clear when the asymptotic theory will be accurate enough to take over from the computer.

The probability element of the ith order statistic from a sample of n drawn from a continuous probability-density function $f(x)$ with cumulative distribution $F(x)$ is

$$g(x_i)dx_i = \frac{n!}{(i-1)!\,(n-i)!}\,[F(x_i)]^{i-1}[1 - F(x_i)]^{n-i}f(x_i)dx_i, \tag{8}$$

while the probability element of the joint distribution of x_i and x_j, $i < j$, is given by

$$h(x_i, x_j)dx_i dx_j = \frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!}\,[F(x_i)]^{i-1}[F(x_j) - F(x_i)]^{j-i-1}[1 - F(x_j)]^{n-j}f(x_i)f(x_j)dx_i dx_j. \tag{9}$$

The quantities we are particularly interested in are

$$\begin{aligned} E(x_i) &= \int_{-\infty}^{\infty} x_i g(x_i)\,dx_i, \\ E(x_i^2) &= \int_{-\infty}^{\infty} x_i^2 g(x_i)\,dx_i, \\ E(x_i, x_j) &= \int_{-\infty}^{\infty}\int_{-\infty}^{x_j} x_i x_j h(x_i, x_j)\,dx_i\,dx_j, \end{aligned} \tag{10}$$

for n in the middle range. This would make it possible to construct and discuss the efficiency of any linear systematic statistic. It would also open the door to improving approximations which would be useful in noncomputational theoretical investigations.

A statistic used in social sciences, where data are frequently ordinal rather than metric, is the rank correlation coefficient (I choose this example rather than some others for ease of

exposition). We have objects ranked from greatest to least on two characteristics. The rank correlation depends entirely on the sum of the squares of the differences of the pairs of ranks given to the objects. There are $n!$ arrangements of the second ranking when we hold the first one fixed. These $n!$ rankings produce a distribution of the sum of squares. We refer an obtained sum of squares to this distribution to decide whether there is reason to believe that there is really correlation between the rankings or whether such a sum of squares might have arisen by chance. For example, with $n = 4$ the distribution is given by

d^2	$f(\Sigma d^2)$
0	1
2	3
4	1
6	4
8	2
10	2
12	2
14	4
16	1
18	3
20	1
Total	$24 = 4!$

For such a small n we probably would not feel much confidence in any correlation unless the rankings agreed perfectly. Such distributions have been tabulated by Olds, and independently by Kendall, by hand up to $n = 8$. We know from work of Hotelling and Pabst that for large n the distribution tends to normality. However, even for $n = 8$, the normality is not close enough for us to get very good approximations to the percentage points of the distribution function. Without the help of high-speed computing machinery, we cannot push this simple calculation much further. The real bother is that $n!$ goes up so rapidly that after n is pushed a few steps further even modern computing machines are bound to be defeated.

We have numerous problems like this in statistics. Many of the attempts to create useful nonparametric statistics bog down at exactly this computational point. Some of these difficulties can be solved in time by sufficiently clever combinatorial devices. But those of us who want methods for practical use, rather than the sheer joy of mathematical investigation, are beginning to wonder whether we might not get more work done in the long run by having tables made by computing machine. It is often easier to solve combinatorial problems when the answer is essentially known. And certainly the table is what we often want for practical work. In other words, why hold up the practical problem for the theoretical investigation when machines can solve the problem, often more accurately, directly? The result of such a trend would be to leave more time for thinking about problems and their solutions and reduce

the time required for arithmetic manipulation. At the same time we relieve the statistician of a side condition. He ordinarily thinks in terms of solutions that he can compute. Computing machines should extend his horizon. For example, very extensive work has been done in multivariate analysis of interest to educators and economists among others, but although the theory is in good shape nothing much has been tabled. Computing machinery can get us these tables and put some of these methods to work.

As a final application I might mention the use of sampling experiments. In some statistical work we cannot get a very workable formulation of the problem mathematically, or, even if we do, the computation becomes too nasty for even modern computing machines. In such cases we are leaning more and more to the sampling experiment. A simple example is supplied by the problem of the truncated normal distribution. We sample from a normal distribution, but part of one or both of its tails has been removed. We would like to compare several methods of estimating the original parameters of the untruncated normal from such truncated samples. With a sample of 15 or 20 the integrations required seem quite unreasonable. Instead, we try the various methods a number of times and use the empirical results in place of the theory. It seems to me that such experiments are admirably suited to computing machines because they involve many repetitions of the same procedure.

By giving these examples of applications of computing machines to social-science problems, and to statistical problems and theory which in turn can be applied to social-science problems, I do not care to give the impression that the uses are really very general, or that modern computing machines will make very fundamental contributions to social science in the near future. Most studies are not very large and the calculations can be handled with a desk computer. Some exceptions are psychological investigations as carried out by Thurstone at the University of Chicago, those done by the Educational Testing Service at Princeton, and censuses and sample censuses as carried out by the Bureau of the Census. Social scientists generally think in nonquantitative terms and, except for economics, there is no large body of mathematical theory available to make quantitative studies on a grand scale sensible. Only in the last ten years has any progress been made in applying mathematics to the social sciences, and the authors of these attempts are quite agreed that little has been done. There has been vague mention of the use of computing machines as logic choppers, the notion being that this is what the social scientist needs because he thinks qualitatively. Until some definite use in the social sciences for a logic machine is suggested, I cannot see how it would apply, however interested I might be in the development of such a device. Some of the burden of limitation falls, of course, on the computing-machinery people. The social scientist interested in population and sociology problems would like to be able to play around with the census data. He would like to make tabulations himself, or to his own order (the Bureau of the Census will make sample studies for him). One might think this would be a job for high-speed computers, and no doubt it will be, but just now I do not think we are very good at high-speed scanning and tabulation of large masses of original data.

At present, most of the direct applications of computing machines in the social sciences outside economics fall into the realm of simultaneous linear equations—homogeneous or nonhomogeneous. Examples are multiple regression, discriminant function, scaling theory, factor analysis. Computing machines can help statistics, the tool subject of quantitative social science, in numerous ways. Some of these are by tabulating functions, by computing properties of statistics, both directly and by means of sampling experiments. This process will help in the development of statistical theory, and make possible the practical use of a large body of theory which is little used because of computational difficulties. Outside economics there is not yet a large body of mathematical theory in the social sciences. This fact sets severe limits on the direct applications of high-speed computing machines except in some special cases mentioned earlier.

Dynamic Analysis of Economic Equilibrium

WASSILY W. LEONTIEF

Harvard University

At a similar occasion three years ago I had the opportunity to discuss a computational problem arising in connection with quantitative analysis of mutual interrelations of the different sectors of a national economy. Since the subject of this paper represents the next step along the same path of inquiry, it can best be introduced through a short recapitulation of the original problem.[1]

The mutual interdependence of the many different branches of production, transportation, distribution, etc. (I will refer to all of them from now on as the different "industries") is basically due to the fact that they exist by "taking in each other's wash." The inputs of any one industry are the outputs of the others: Let X_i represent the annual rate of total output (measured in appropriate physical units) of industry i, x_{ik} the amount of the product of industry k absorbed annually by industry i, and x_{nk} the amount of the same product k made available for "outside use," that is, for consumption not by any one of the m industries explicitly included in the economic system under consideration. The over-all input-output balance of a whole national economy comprising m separate industries can be described in terms of m linear equations

$$X_i - \sum_{k=1}^{k=m} x_{ki} = x_{ni}. \qquad i = 1, 2, \ldots, m. \tag{1}$$

Turning to the internal input-output structure of any particular industry, we find that there exists a definite relation, rather narrowly determined by technological—in the widest sense of the word—considerations, between the rate of its output and the quantities of all the various materials and services required to achieve it. As a first empirically justified approximation, the assumption can be made that the quantity of each kind of input absorbed by an industry per unit of its output is fixed. Thus the magnitudes included in the balance equations are subject to the set of structural relations

$$x_{ki} = a_{ki}X_k. \qquad i = 1, 2, \ldots, m; \; k = 1, 2, \ldots, m. \tag{2}$$

Substituting Eq. (2) in Eq. (1) we have

$$X_i - \sum_{k=1}^{k=m} a_{ki}X_k = x_{ni}, \quad i = 1, 2, \ldots, m \tag{3}$$

which, solved for the X_i's, gives

$$X_i = \sum_{k=1}^{k=m} A_{ik}X_{nk}, \qquad i = 1, 2, \ldots, m \tag{4}$$

where the A_{ik}'s are elements of the inverse of the structural matrix $|a_{ki}|$.

Given an "outside bill of goods" $x_{n1}, x_{n2}, \ldots, x_{nm}$, be it the final domestic consumers'

demand, allocations to the foreign countries under the Marshall Plan aid, or itemized material requirements of a military mobilization program, the last system enables us to determine the corresponding level of output in all the individual sectors of the economy.

Only a few years ago the computational difficulties involved in the inversion of a square matrix of order 40, 100, or 150 would have been considered practically insurmountable. Now our main concern is that of collecting sufficiently accurate primary quantitative information on the basis of which such matrices are being set up.

But new computational problems arise as the thinking on the subject advances.

The theoretical scheme of interindustrial relations as presented above is entirely static. All variables occurring in it are time rates of input and output flows. The actual economic process involves, however, not only *flows* but also *stocks* of commodities: stocks of machinery, stocks of buildings, inventories of raw materials or goods-in-process and of finished commodities.

Explicit incorporation of stocks as well as of flows in the model of the national economy leads to formulation of a dynamic theory. The change in the magnitude of any particular kind of stock—if it is at all possible—is achieved through accumulation or decumulation of a flow over time. Let s_{ik} represent the stock of commodity k used in industry i at the time t; $\dot{s}_{ik}$ describes, then, the flow of additions to (or subtractions from) that particular stock. The balance equations (1) can now be rewritten

$$X_i - \sum_{k=1}^{k=m} x_{ki} - \sum_{k=1}^{k=m} \dot{s}_{ki} = x_{ni}. \quad i = 1, 2, \ldots, m. \tag{5}$$

The flow of commodities from industry i to industry k is being split here explicitly into two components, x_{ki} representing that part of it which is being used "on current account" and $\dot{s}_{ki}$ the other part added (if $\dot{s}_{ki} > 0$) or subtracted (if $\dot{s}_{ki} < 0$) from the stock s_{ki}.

A corresponding modification must be introduced also into the description of the internal structure of the separate industries. The equations of set (2) as formulated above refer only to technical input requirements on current account. All additions to stock have, in the original static formulation, been treated as parts of the independent "outside demand," that is, the vector $x_{n1}, x_{n2}, \ldots, x_{nm}$; they were not explained but rather treated as known parameters. The more comprehensive dynamic formulation contains an additional, second set of structural equations in which each stock or capital requirement of a particular industry is related to its rate of output,

$$s_{ki} = b_{ki}X_k, \quad i = 1, 2, \ldots, m \tag{6}$$

or differentiating,

$$\dot{s}_{ki} = b_{ki}\dot{X}_k. \tag{6a}$$

The constants b_{ki} can be referred to as the capital coefficients.

Substitution of Eqs. (2) and (6*a*) in Eq. (5) gives a set of m linear differential equations with constant coefficients,

$$X_i - \sum_{k=1}^{k=m} b_{ki}\dot{X}_k - \sum_{k=1}^{k=m} a_{ki}X_k = x_{ni}. \quad i = 1, 2, \ldots, m. \tag{7}$$

A general solution of this dynamic system can be written

$$X_i(t) = \sum_{k=1}^{k=m} C_k(X_1^0, X_2^0, \ldots, X_m^0)K_{ki}e^{\lambda_k t} + L_i(x_{n1}, x_{n2}, \ldots, x_{nm}), \qquad (8)$$
$$i = 1, 2, \ldots, m$$

where the λ's are the characteristic roots of system (7); the C_k's, linear functions of the initial conditions (expressed in terms of the rates of output of all industries at some point of time t_0); the K_i's, appropriate functions of the constants, the a's, b's, and λ's; while $L_i(x_{n1}, x_{n2}, \ldots, x_{nm})$ are linear functions of the outside demand $x_{n1}, x_{n2}, \ldots, x_{nm}$.

Once obtained in numerical form, this solution makes it possible to answer various types of questions arising in connection with the explanation of the behavior of the economic system over time.

Western Europe, for example, is striving to accomplish an investment program which in a certain number of years would make it independent of the negative "outside demand," that is, outside supplies currently being made available to it under the Marshall Plan. Had the necessary primary information been available one could have determined the m rates of surplus imports of various commodities (the x_{ni}'s) that would be required to raise the domestic output from a given original level $X_1(t_0), X_2(t_0), \ldots$, to some prescribed higher level $X_1(t_1), X_2(t_1), \ldots$ over a stated period of time $t_1 - t_0$.

To obtain the desired answer it would be only necessary to insert in Eq. (8) the original levels of output as the initial condition, set t in the exponentials equal to the prescribed recovery period $t_1 - t_0$, and equate the right-hand terms of the equations to the desired final levels of output $X_1(t_1), X_2(t_1), \ldots$ The resulting m linear equations can then be solved for the m unknown quantities $x_{n1}, x_{n2}, \ldots, x_{nm}$ of surplus imports.

If some of the characteristic roots of the differential equations (7) turn out to be complex, the outputs of the individual industries will display a typical periodic pattern of motion with increasing constants of diminishing amplitudes depending upon the magnitude of the real parts of the roots. The consideration of such periodic solutions constitutes the theoretical basis of many a contemporary business-cycle theory.

Although very attractive because of its obvious simplicity, this explanation of alternative booms and depressions has a serious weakness which, if it is overcome by appropriate theoretical reformulation, leads to a new and interesting computational problem.

The original static system has been transformed into a dynamic one by the introduction of stock-flow relations as described by a set of appropriate capital coefficients. Not all capital stocks can, however, be decumulated, that is, reduced in the same way in which they are being accumulated. Investments in raw materials, goods-in-process, and finished commodities, in short, stocks which are associated with the concept of working capital, can indeed move downward as easily as they can go up; not so with fixed capital, that is, machinery, buildings, permanent investment in roadbeds, soil conservation, etc. Provided sufficient sources of supply exist, these stocks can change in the upward direction as readily as working capital.

In the case of contracting demand, however, fixed capital cannot be as readily reduced as, say, inventories of raw materials. For obvious technological reasons, the rate at which machinery and buildings, not to speak of so-called permanent land improvements, can be used up is strictly limited and at best is very low. The appearance of unused capacity of idle fixed capital (in times of downward production trends and during the initial phases of recovery when output has not yet reached the previous high) constitutes one of the most characteristic aspects of modern business fluctuations.

In the light of these observations, the use of unchanging stock-flow ratios well suited—at least in the first approximation—to the analysis of the dynamics of working capital is inappropriate in application to fixed capital. As soon as the rate of output of an industry begins to diminish, or reaches a certain critical rate of decrease at which idle investment makes its appearance, the technologically necessary stock-flow relations lose their economic significance so far as the stocks of fixed investment used in this industry are concerned. Since and to the extent that the previously accumulated stocks of durables cannot be reduced through consumption on current account these become quasi-free goods. The economic connection between such stocks and the current rate of output ceases to exist and it is reestablished only when and if, in the course of a subsequent increase in the rate of production, all existing idle capacity has again been reabsorbed, and thus additional investment becomes necessary for further expansion.

In terms of a previously described theoretical model, it means that, in the course of the time intervals during which the fixed capital stock of a particular industry exceeds the technologically required magnitude, the corresponding capital coefficients in the system of differential equations (8) become zero. They acquire, however, the original values if and as soon as the output of the industry again increases to the level at which the surplus stocks become reabsorbed.

The movement of the whole economic system can thus be described as a succession of alternative phases. Within each phase its path is defined by an appropriate system of linear differential equations with constant coefficients. Each of these alternative systems is obtained by suppressing a certain subset of capital coefficients of the original, complete system. The initial conditions of every phase are determined by the state in which the system found itself at the end of the previous phase. (Specifically, they are described in terms of the corresponding rates of output of the m individual industries); thus, throughout the process as a whole all variables are continuous functions of time but their derivatives are, in general, discontinuous at the point of junction of the successive phases.

The rules governing the transition from one phase to another are as follows:

1. The current phase terminates

(*a*) whenever the *downward* change in the rate of output of any industry that uses any particular kind of durable capital goods reaches a certain fixed critical magnitude below which the rate of effective capital decumulation (that is, the rate of reduction in the magnitude of the particular kind of capital stock) cannot fall, and

(*b*) whenever the rate of output of any capital-using industry reaches, in the course of its *upward* movement, a level at which the stock of any one kind of fixed capital becomes again fully utilized, that is, becomes equal to the stock that existed at the beginning of the latest previous phase during which it started to be idle, minus the maximum effective rate of attrition, multiplied by the length of time elapsed since then.

2. The set of differential equations of a succeeding phase is obtained from that of a previous phase

in case (*a*) by suppressing the capital coefficients b_{ki} of industry k that refer to stocks of durable commodities, and

in case (*b*) by reintroduction of capital coefficients that were suppressed at the beginning of the terminating phase.

From an economist's point of view, it is particularly interesting to note that the process described might display alternating upward and downward movements in the rate of output of the various industries, even if none of the characteristic roots of the sets of differential equations governing any one of its separate phases are complex, that is, even if neither one of them contains periodic components. The unrealistically smooth symmetry and exact periodicity implied by conventional business-cycle models is thus entirely eliminated.

Given a set of empirically observed technical flow and capital coefficients and the initial state of the system, the course of the ensuing-phase periodic process is uniquely determined. The nature of the underlying dynamic relation is such, however, that the description of the resulting movement cannot be reduced to a simple general formula. The movement of the system has to be found stepwise from one phase to another. Modern high-speed machines are, for obvious reasons, peculiarly suited to efficient solution of the resulting computational problem.

REFERENCE

1. W. W. Leontief, "Computational problems arising in connection with economic analysis of interindustrial relationships," *Proceedings of a Symposium on Large-Scale Digital Calculating Machinery* (Harvard University Press, Cambridge, 1948), p. 169.

Some Computational Problems in Psychology

LEDYARD R. TUCKER

Educational Testing Service, Princeton, New Jersey

During the major portion of this paper we shall consider in detail two problems of special interest to the speaker. One of these problems is concerned with a point in general methodology and can be applied to a number of psychological studies. The other problem is of a more theoretic nature involving quantitative assumptions of particular psychological relations. The first problem concerns the question of *how* to perform the requisite computations; the second problem is more a question of *what* are the results of a set of theoretic constructs. Practicable means of obtaining solutions in either case, making use of normally available methods, have not been found. It is the hope of the author that large-scale digital computing machines will be of assistance in solving these and similar problems in psychology.

Before going into the details of these problems, I wish to emphasize the point that they are not typical of the majority of computations in psychology. In general, workers in the quantified aspects of psychology are plagued with the necessity of working with masses of what might be termed low-grade data. Here we are using the term "low-grade data" to indicate that each datum has limited pertinence owing to the influence of uncontrollable factors in experimental or observational situations. As a result of inconsistencies in the observations and of the many attributes that must be considered simultaneously in a number of studies, it has often been necessary to use linear types of equations in quantitative psychological theories. Even these simpler mathematical formulations have yielded gratifying results. The point for consideration here is that, typically, computing problems in psychology involve simple calculations on many observations. At present, psychologists place general emphasis on computing aids in matrix manipulations such as multiplications and inversions. The first problem we will describe in detail is atypical only in the fact that the matrices are interrelated in a more complex manner. There are some areas in psychology where complex functions are employed, but computational-type solutions are seldom indicated as necessary. In many cases, adequate mathematical constructs have not been developed and reliance is placed on graphical aids and the judgment of the experimenter.

In addition to the mathematical type of psychological theories, considerable dependence is placed on statistical methods, from which methods a number of quantity-type computational problems arise.

During the foregoing discussion, the characteristics of simple calculations on a large quantity of data have been emphasized as more typical of the more critical computational problems in psychology. For these problems use of the large-scale digital computing machines is not

appropriate, from what I understand about these machines. In consequence, I have selected for discussion two more complex computational problems.

Problem 1. *Determination of maximum-likelihood estimates of a factorial analysis structure.*

As previously indicated, this problem deals with procedural implimentation of a theoretic method. The theoretic work was performed by D. N. Lawley at the University of Edinburgh, who first published a paper on the subject in 1940.

Table 1. An illustrative covariance matrix C (based on observations of 50 cases).

	1	2	3	4	5	6	7	8
1	*0.35*	− 0.60	0.88	− 3.92	− 1.29	− 1.15	− 1.00	− 0.30
2	− 0.60	*6.80*	− 5.84	31.36	5.52	11.20	1.60	4.80
3	0.88	− 5.84	*11.20*	− 22.40	− 10.80	− 6.00	− 9.92	− 0.48
4	− 3.92	31.36	− 22.40	*225.40*	18.90	73.50	− 9.52	36.12
5	− 1.29	5.52	− 10.80	18.90	*29.25*	0.75	26.76	− 6.66
6	− 1.15	11.20	− 6.00	73.50	0.75	*61.25*	− 23.40	18.90
7	− 1.00	1.60	− 9.92	− 9.52	26.76	− 23.40	*58.40*	− 8.40
8	− 0.30	4.80	− 0.48	36.12	− 6.66	18.90	− 8.40	*45.00*

Consider the matrix C in Table 1. It is square and symmetric. The off-diagonal cell entries are known as covariances; the diagonal entries are known as variances. In this case this is a fictitious setup, but the values recorded could have been obtained from observation of the performance of 50 people on eight different tests. Each variance, or diagonal cell entry, is a measure of the variability of scores of the 50 subjects on one of the tests. Each covariance, or off-diagonal cell entry, is a measure of similarity of the rank-order position of the 50 subjects on a pair of the tests. There is one row and one column for each test included in the study. Although eight variable tables of covariances are rather common, factorial-analysis methods should not usually be applied to so small a set of tests. The size of the study that would involve factor analysis has, more typically, 20 to 100 variables. The number of cases may be over 1,000. The covariance matrix C is our starting point for the present computing problem.

Let n represent the number of tests and N the number of people tested. It is desired to obtain a matrix A with n rows and r columns to satisfy Eqs. (1) to (7) and Conditions I and II. The numerical work for Eqs. (1) and (2) is given in Table 2 for the case where there are two columns in matrix A, that is, where r is 2. The simplicity in the setup of the illustrative example is apparent in the use of simple values for entries in matrix B and for the u_j's. Our procedure for the illustrative example was the reverse of the computational problem because we set

Table 2. Numerical example for Eqs. (1) and (2).

Variable number	c_{jj}	Matrix A I	Matrix A II	u_j^2	u_j	Matrix B I	Matrix B II
1	0.35	0.3	— 0.1	0.25	0.5	0.6	— 0.2
2	6.80	— 2.4	— 0.2	1.00	1.0	— 2.4	— 0.2
3	11.20	2.4	— 1.2	4.00	2.0	1.2	— 0.6
4	225.40	— 12.6	— 4.2	49.00	7.0	— 1.8	— 0.6
5	29.25	— 2.7	3.6	9.00	3.0	— 0.9	1.2
6	61.25	— 4.5	— 4.0	25.00	5.0	— 0.9	— 0.8
7	58.40	— 1.2	6.4	16.00	4.0	— 0.3	1.6
8	45.00	— 1.8	— 2.4	36.00	6.0	— 0.3	— 0.4

up the solution first and then worked backward to matrix C. Unfortunately, it is much more difficult to obtain the solution working in the normal direction from matrix C. In Table 2, the column for c_{jj} contains the diagonal entries of matrix C. Matrix A is the desired solution. Each entry in the column headed u_j^2 is obtained by subtracting the sum of squares of the entries in matrix A for the row from the c_{jj} in the row. Equation (1) summarizes the relation of the u_j's to the c_{jj}'s and entries in row j of matrix A

$$u_j = (c_{jj} - \sum_m a_{jm}^2)^{\frac{1}{2}}. \tag{1}$$

Each entry in matrix B is obtained by dividing the corresponding entry in A by the u_j for the row. This step can be summarized in matrix notation by Eq. (2), where the matrix U is a diagonal matrix with diagonal entries u_j;

$$B = U^{-1}A. \tag{2}$$

Table 3. The matrix E.

	1	2	3	4	5	6	7	8
1	*0.40*	— 1.20	0.88	— 1.12	— 0.86	— 0.46	— 0.50	— 0.10
2	— 1.20	*5.80*	— 2.92	4.48	1.84	2.24	0.40	0.80
3	0.88	— 2.92	*1.80*	— 1.60	— 1.80	— 0.60	— 1.24	— 0.04
4	— 1.12	4.48	— 1.60	*3.60*	0.90	2.10	— 0.34	0.86
5	— 0.86	1.84	— 1.80	0.90	*2.25*	0.05	2.23	— 0.37
6	— 0.46	2.24	— 0.60	2.10	0.05	*1.45*	— 1.17	0.63
7	— 0.50	0.40	— 1.24	— 0.34	2.23	— 1.17	*2.65*	— 0.35
8	— 0.10	0.80	— 0.04	0.86	— 0.37	0.63	— 0.35	*0.25*

With Eqs. (3) we return to matrix C and define a new matrix E with the same number of rows and columns as C:

$$E = U^{-1}CU^{-1} - I. \tag{3}$$

Entries in each row and column of C are divided by the corresponding u_i's; and, in addition, unity is subtracted from the diagonal values. Table 3 is the result for our example.

Equations (4) and (5) and Condition I are the definitive relations of matrix B to matrix E:

$$B = EBK, \tag{4}$$

$$K^{-2} = B'EB. \tag{5}$$

Condition I. *Matrix B has the r latent vectors of E corresponding to the largest latent roots of E.*

Table 4 gives the numerical example for these equations. Matrix B is repeated from Table 2. The matrix product EB has been computed. Matrix K^{-2} is a diagonal matrix with diagonal

Table 4. Numerical example for Eqs. (4) and (5).

Variable number	Matrix B		Product EB	
	I	II	I	II
1	0.6	− 0.2	7.56	− 1.12
2	− 2.4	− 0.2	− 30.24	− 1.12
3	1.2	− 0.6	15.12	− 3.36
4	− 1.8	− 0.6	− 22.68	− 3.36
5	− 0.9	1.2	− 11.34	6.72
6	− 0.9	− 0.8	− 11.34	− 4.48
7	− 0.3	1.6	− 3.78	8.96
8	− 0.3	− 0.4	− 3.78	− 2.24

	Matrix K^{-2}			Matrix K	
	I	II		I	II
I	158.76	0	I	$\frac{1}{12.6}$	0
II	0	31.36	II	0	$\frac{1}{5.6}$

entries equal to the sum of products between corresponding entries in identical columns of B and EB. The off-diagonal entries in K^{-2} are sums of products between entries in different columns of B and EB and must come out zero. Matrix K is also a diagonal matrix with diagonal cell entries equal to the reciprocal of the square roots of the corresponding entries

in K^{-2}. When the entries in each column of EB are multiplied by the corresponding diagonal entry in K, matrix B is reproduced.

It is to be noted that Eqs. (1) to (5) and Condition I apply for any given number, r, of columns of A. The matrices are so interdependent in these five equations that direct solutions starting only with matrix C are quite impracticable and it is necessary to resort to successive trial solutions. Discussion of the trial procedures is postponed until the theory of the method for determining the number of columns of A has been considered.

Table 5. The matrix G. (For $j > k$: $\sum_j \sum_k g_{jk}^2 = 0.3200$, $W_2 = 16.00$.)

	1	2	3	4	5	6	7	8
1	*1.00*	0.20	0.04	− 0.16	− 0.08	− 0.08	0.00	0.00
2	0.20	*1.00*	− 0.16	0.04	− 0.08	− 0.08	0.00	0.00
3	0.04	− 0.16	*1.00*	0.20	0.00	0.00	0.08	0.08
4	− 0.16	0.04	0.20	*1.00*	0.00	0.00	0.08	0.08
5	− 0.08	− 0.08	0.00	0.00	*1.00*	0.20	0.04	− 0.16
6	− 0.08	− 0.08	0.00	0.00	0.20	*1.00*	− 0.16	0.04
7	0.00	0.00	0.08	0.08	0.04	− 0.16	*1.00*	0.20
8	0.00	0.00	0.08	0.08	− 0.16	0.04	0.20	*1.00*

The matrix G of Table 5 is defined by Eq. (6):

$$G = E - BB' + I. \tag{6}$$

Off-diagonal entries in G are interpreted as the residual portion of E not accounted for by the matrix B. The diagonal entries must be unity as a consequence of Eqs. (1) to (5). The general size of the off-diagonal entries is of special interest. First, as the number of factors —that is, columns of matrix A—is increased, the general size of these off-diagonal entries decreases. A second important relation is with the number of cases on which observations were made. When observations are made on an extremely large group of cases, the off-diagonal entries in G are vanishingly small for a given set of variables and a given number of columns in matrix A. As smaller groups of cases are considered, these off-diagonal entries will usually increase in size. Consider the index W_r defined in Eq. (7):

$$W_r = N \sum_j \sum_k g_{jk}^2, \qquad j > k \tag{7}$$

where N is the number of cases in the sample. The double summation is of squares of the off-diagonal entries of G below the diagonal. When the off-diagonal entries differ from zero only because of chance sampling effects, W_r for a number of samples of size N will have a frequency distribution in accordance with the chi-square distribution with appropriate degrees of freedom. Values of chi-square have been tabulated. It is possible, therefore, to set up a

value above which we would believe the possibilities of observing W_r owing solely to chance sampling effects to be negligible. Thus, whenever a W_r is observed above this critical value, it will be concluded that more columns of A are required. This reasoning leads to

Condition II. *The number of columns r of matrix A is the smallest number where W_r is less than chi-square, with $\frac{1}{2}[(n-r)^2-n-r]$ degrees of freedom, for a given level of significance.*

When a 10-percent level of confidence is used, a chi-square can be computed that will be exceeded 10 percent of the time by chance sampling effects. This 10-percent level of confidence seems, to the author, to be appropriate for the present problem.

By making use of an excellent approximation devised by Edwin B. Wilson and Margaret M. Hilferty, a direct computational procedure can be designated. Here r is to be the smallest integer for which

$$W_r < \tfrac{1}{2}[(n-r)^2-n-r]\left\{1-\frac{4}{9[(n-r)^2-n-r]}+1.2816\sqrt{\frac{4}{9(n-r)^2-n-r}}\right\}^3. \quad (8)$$

Solutions have been obtained for the illustrative problem for 0, 1, and 2 columns of matrix A. Values of W_r, degrees of freedom, and actual chi-squares and approximate chi-squares for a 10-percent level of confidence are listed in Table 6. For no factors and for one factor, that

Table 6. Values of W_r, number of degrees of freedom, and actual and approximate chi-square for 10-percent level of confidence for illustrative example.

r	W_r	Degrees of freedom	Chi-square	
			Actual	Approximate
0	228	28	37.92	37.91
1	106	20	28.41	28.40
2	16	13	19.81	19.80

is, r equal to 0 or 1, W_r is greater than the corresponding chi-square. For two factors, W_r is less than the chi-square. We therefore conclude that two factors are appropriate for the illustrative example. The last two columns of the table indicate the excellence of the approximation.

In review, the computational problem is to begin with a matrix C and to obtain matrices A and values of W_r for successively increasing numbers of factors until the inequality (8) is satisfied. The matrix A with the smallest number of columns for which the inequality is satisfied is the solution.

The major difficulty lies in obtaining a matrix A, with any specified number of columns, from matrix C in order to satisfy Eqs. (1) to (5) and Condition I. Direct solutions seem impracticable. One possible successive-approximation method is based on a procedure developed by Harold Hotelling for obtaining principal components. An initial trial matrix

A_1 is obtained by any of several methods. Random numbers may be used, provided the u_{j1}^2 of Eq. (9) is greater than zero:

$$u_{j1}^2 = c_{jj} - \sum_m a_{jm1}^2. \tag{9}$$

It is possible to have zero-entries for all but two cells in any one column of A_1. Computations for one trial for the illustrative example are presented in Table 7. The case illustrated is that

Table 7. One trial of a successive-approximation method.

Variable number	A_1 I	A_1 II	u^2_{j1}	L_1 I	L_1 II	M_1 I	M_1 II	A_2 I	A_2 II
1	− 0.28	0	0.27	− 1.04	0	− 3.4	− 0.1	− 0.28	0.10
2	2.44	0	0.85	2.87	0	30.1	2.6	2.44	− 0.19
3	− 2.22	0	6.27	− 0.35	0	− 26.9	0.6	− 2.18	1.41
4	12.81	0	61.30	0.21	0	159.6	21.2	12.93	2.13
5	2.18	0	24.50	0.09	0	25.8	− 5.7	2.09	− 3.53
6	4.68	4.66	17.63	0.27	0.26	63.3	16.4	5.13	4.23
7	0.53	− 5.66	26.08	0.02	− 0.22	3.4	− 13.3	0.28	− 5.79
8	2.04	0	40.84	0.05	0	26.4	6.8	2.14	1.75

	K_1^{-2} I	K_1^{-2} II
I	153.66	15.71
II	15.80	7.19

	K_1^{-1} I′	K_1^{-1} II′
I	12.40	0
II	1.27	2.36

	K_1' I′	K_1' II′
I	0.081	− 0.043
II	0	0.424

in which two factors are under consideration. A similar series of trials was performed for the case of one factor, and the first column of A_1 in Table 7 is the approximate solution obtained. The values in cells for rows 6 and 7 were obtained by solving Eqs. (10), (11), and (12):

$$f_{jk} = c_{jk} - a_{j1}a_{k1}, \tag{10}$$

$$a_{62}a_{72} = f_{67}, \tag{11}$$

$$\frac{a_{62}^2}{a_{72}^2} = \frac{f_{66}}{f_{77}}. \tag{12}$$

The entire matrix F with cell entries f_{jk} was computed and the largest off-diagonal entry selected. This was for tests 6 and 7. Equations (11) and (12) yield reasonable values for a_{62} and a_{72}. Obtaining a trial matrix A_1, the u_{j1}^2's are computed by Eq. 9.

Equations (13), (14), and (15) indicate the computations for subsequent steps:

$$L_1 = U_1^{-2}A_1, \tag{13}$$

$$M_1 = CL_1 - A_1, \tag{14}$$

$$K_1^{-2} = L_1' M_1. \tag{15}$$

The matrix L_1 is obtained by dividing each entry in a row of A_1 by the u_{j1}^2 for that row. Matrices M_1 and K_1^{-2} are obtained as matrix products. Matrix K_1^{-2} should be symmetric except as a result of rounding-off errors in M_1. Usually minor errors in symmetry will not materially affect succeeding steps. The matrix K_1^{-1} is computed to satisfy Eq. (16) and to have zero entries above the diagonal:

$$(K_1^{-1})(K_1^{-1})' = K_1^{-2}. \tag{16}$$

The inverse of K_1^{-1} and the matrix product of Eq. (17) are obtained:

$$M_1 K_1' = A_2. \tag{17}$$

The matrix A_2 is the second approximation.

Usually this system should converge on the desired solution. The rate of convergence is likely to be low, and many trials may be required for each number of factors. As a result, the kind of computational assistance offered by large-scale digital computing machines is essential. Other simpler factorial methods are in use by psychologists mainly to avoid the computational labor involved in the procedure we have outlined. These other procedures have not answered the question of the number of factors to be considered, and they suffer from the usual effects of approximations. It would be quite desirable to be able to apply Lawley's method.

Another computational problem with which many psychologists are concerned can be mentioned here, but we will not spend much time on it. The matrix A can be considered to give the coördinates of n vectors, one for each test, in r-dimensional space, one dimension for each factor. It is possible to restate Eqs. (4) and (5) and Condition I to allow the rotation of axes within the space defined by matrix A. L. L. Thurstone has stated his principle of simple structure to solve the problem of where to rotate the axes. It is sufficient to state here that much computational labor is involved and assistance will be welcomed by psychologists. Unfortunately, Thurstone's principle is qualitative in nature. Until his principle has been successfully restated in precise mathematical relations, the solution will depend on the judgment of the analyst, and the applicability of large-scale calculating machines will be doubtful.

We shall now turn to our second problem.

Problem 2. Determination of Distribution of Items in Difficulty to Yield Maximum Test-Ability Correlation.

The computational problem can be stated quite simply. Consider Eq. (18):

$$V = \sum_h r_s \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t_h^2} \left\{ \sum_h \int_{t_h}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_h^2} dx_h - \left(\sum_h \int_{t_h}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_h^2} dx_h \right)^2 \right.$$
$$\left. + \sum_h \sum_{i \neq h} \int_{t_h}^{\infty} \int_{t_i}^{\infty} \frac{1}{2\pi\sqrt{1 - r_s^4}} \exp \left[- \frac{1}{2(1 - r_s^4)} (x_h^2 + x_i^2 - 2x_h x_i r_s^2) \right] dx_i dx_h \right\}^{-\frac{1}{2}}. \tag{18}$$

This equation arises in a mathematical development related to the assembly of standardized aptitude examinations. Let us take as an example a test in addition, where it is desired to differentiate among people most validly with respect to rapidity and accuracy of adding columns of numbers. When, however, a large number of problems are tried out on a group of people, it is found that some problems are easy and others are more difficult in terms of the percentage of the group that gives the correct answers in the time allowed. These differences in difficulty could arise from differing lengths of the columns to be added, or from differind numbers of digits in each number, or from use of numbers differing in ease of addition. It is easier to add 1's and 2's than 6's, 7's, or 8's. Thus the test technician has at his disposal a large number of problems of different difficulty. How should the problems be selected as to difficulty? If only those problems are used to which half of the group gives the correct answer in the time allotted, then the people who are very poor in addition will be grouped at low scores and the people who are very good in addition will be grouped at high scores, there being, therefore, little differentiation among members in either of the extreme groups. It thus seems that the problems should be spread out in difficulty. Exactly what is the best distribution of problems in difficulty is not known. Our Eq. (18) is an attempt to obtain the answer.

We have assumed a scale of ability. The coefficient V is the correlation between the test scores and scores on this ability. The quantity r_s is a measure of the relation of each test problem and the underlying ability. The t_h's are indices of the problems' difficulties. The subscript i is used alternatively with h to designate the problems. We would like, for any given number of problems and value of r_s, to select that set of t_h's which would yield a maximum value for V. It will be noted that the independent variables—the t_h's—appear as exponents and as limits of definite integrals for which only numeric solutions exist. Tabulated sets of values of the t_h's that maximize V for several conditions of numbers of problems between 10 and 500 and of values of r_s between zero and unity would be of considerable assistance as guides in the construction of tests. I have not tried a computational type of solution, but the reports of capabilities of the large-scale digital computers indicate that they should yield the desired results.

In this paper I have outlined two problems on which the large-scale computing machines should be of assistance to psychologists. These problems were chosen because (*a*) they were so stated that computational-type solutions were feasible, (*b*) the volume and complexity of the calculations indicated a need for assistance from a large-scale machine, and (*c*) the problems were of special interest to the author.

REFERENCES

H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educational Psychol.* **24,** 417–441, 498–520 (1933).

D. N. Lawley, "The estimation of factor loadings by the method of maximum likelihood," *Proc. Roy. Soc. (Edinburgh)* **60,** 64–82 (1940).

D. N. Lawley, "Further investigations of factor estimation," *Proc. Roy. Soc.* (*Edinburgh*) **61,** 176–185 (1941).

D. N. Lawley, "The application of the maximum likelihood method to factor analysis," *Brit. J. Psychol.* **33,** 172–175 (1943).

L. L. Thurstone, *Multiple factor analysis* (University of Chicago Press, Chicago, 1947).

Edwin B. Wilson, and Margaret M. Hilferty, "The distribution of Chi-square," *Proc. Nat. Acad. Sci. U.S.* **17,** 684–688 (1931).

Computational Aspects of Certain Econometric Problems

HERMAN CHERNOFF

University of Chicago

Economists are frequently interested in constructing models of economic behavior and estimating the parameters involved. In many cases the traditional least-squares treatment fails and it becomes necessary to maximize a likelihood function of many variables. Such problems frequently involve many matrix operations where it is of great importance for computers to have high computing speeds, either large internal memory or rapid transfer from internal to external memories, or both, and the ability to take advantage of matrices which are very simple in that most elements are zero.

As an illustrative example, suppose that y_{t1} represents the quantity sold of a certain good in year t; y_{t2}, the price of the good in year t; z_{t1}, the national income in year t; and z_{t2}, the wage rate in the producing industry in year t. Then the economist hypothesizes the following model, consisting of the demand and supply equations

$$\begin{aligned} y_{t1} &= \alpha_1 y_{t2} + \alpha_2 z_{t1} \qquad\qquad + \alpha_3 + u_{t1}, \\ y_{t1} &= \beta_1 y_{t2} \qquad\qquad + \beta_2 z_{t2} + \beta_3 + u_{t2}, \end{aligned}$$

where u_{t1} and u_{t2} are random unobserved disturbances. It is desired to estimate the parameters α_1, α_2, α_3, β_1, β_2, β_3 because they can be used to forecast the effect of an excise tax, say, on the price and quantity of a good (once the national income and the wage rate in the producing industry are known). Furthermore, these parameters in themselves have important theoretical significance.

One should refrain from using the traditional least-squares regression method to estimate the parameters, for the conditions justifying the use of least squares are not satisfied and that method may give meaningless results even in large samples. (To justify least squares one must assume that u_{t1}, u_{t2} are distributed independently of y_{t2}, z_{t1}, z_{t2}; this is not the case since y_{t1} and y_{t2} are determined by z_{t1}, z_{t2}, u_{t1}, u_{t2}.) In this case the maximum-likelihood method should be applied.

In general, our system of equations may be written

$$\begin{aligned} &\beta_{11} y_{t1} + \beta_{12} y_{t2} + \cdots + \gamma_{11} z_{t1} + \gamma_{21} z_{t2} + \cdots = u_{t1}, \\ &\beta_{21} y_{t1} + \beta_{22} y_{t2} + \cdots + \gamma_{21} z_{t1} + \gamma_{22} z_{t2} + \cdots = u_{t2}, \\ &\quad\vdots \\ &\beta_{g1} y_{t1} + \beta_{g2} y_{t2} + \cdots + \gamma_{g1} z_{t1} + \gamma_{g2} z_{t2} + \cdots = u_{tg}, \end{aligned}$$

where the y's and z's are observed variables and the parameters β_{ij}, γ_{ik} are subject to certain restrictions derived from economic theory. A case which occurs frequently is that where linear functions of these parameters are known to vanish. In this case a considerable number of simplifications may enter. The u_{ti} are unobserved random disturbances which are assumed to be jointly normally distributed, independently of the z_{ti}.[1]

The above equations may be written

$$\mathrm{B}y_t' + \Gamma z_t' = u_t',$$

or

$$\mathrm{A}x_t' = u_t',$$

where

$$\mathrm{A} = (\mathrm{B}\ \Gamma), \quad x_t' = \begin{pmatrix} y_t' \\ z_t' \end{pmatrix}.$$

Then it can be shown that the maximum-likelihood method involves maximizing

$$\log L(\mathrm{A}) = \tfrac{1}{2} \log \det (\mathrm{A}\ W\ \mathrm{A}') - \tfrac{1}{2} \log \det (\mathrm{A}\ M\ \mathrm{A}')$$

as a function of A, where A is subject to the a priori restrictions due to economic theory. Here

$$M = \frac{1}{T} \sum_{t=1}^{T} x_t' x_t = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},$$

$$M_{11} = \frac{1}{T} \sum_{t=1}^{T} y_t' y_t, \quad M_{12} = \frac{1}{T} \sum_{t=1}^{T} y_t' z_t, \text{ etc.},$$

$$W = \begin{pmatrix} W_{11} & 0 \\ 0 & 0 \end{pmatrix},$$

$$W_{11} = M_{11} - M_{12} M_{22}^{-1} M_{21}.$$

Iterative methods have been applied to maximize L. These are gradient methods where one starts with a certain approximation P and takes a step in the direction of steepest ascent.[2] In particular, the Newton method is a special case of a gradient method where convergence per iteration is very rapid but the amount of calculation per iteration is quite large. This method converges faster as the approximation gets closer to the maximizing value. To apply gradient methods one must consider first derivatives of L. For the Newton method, second derivatives are also required. It is found convenient to work with the Taylor expansion

$$\log L(\mathrm{P} + h\mathrm{D}) = \log L(\mathrm{P})\ h \operatorname{tr} \{[(\mathrm{P}W\mathrm{P}')^{-1}\ \mathrm{P}W - (\mathrm{P}M\mathrm{P}')^{-1}\ \mathrm{P}M]\mathrm{D}'\} + O(h^2),$$

where tr represents the trace or the sum of the elements along the main diagonal. From this expansion it is seen that to compute the first-order derivatives one must essentially compute $(\mathrm{P}W\mathrm{P}')^{-1}\ \mathrm{P}W - (\mathrm{P}M\mathrm{P}')^{-1}\ \mathrm{P}M$. In a typical case of an eleven-equation system, the number of rows and columns corresponding to P, M, W_{11}, respectively, are (11×40), (40×40), (11×11).

In the case where each of the restrictions is linear in the coefficients of one equation, one can reduce the number of computations per iteration considerably at the cost of introducing a few initial operations. For example, consider the following formula for the direction of

steepest ascent with respect to the linearly independent parameters a in terms of which P may be computed:

$$d = a\Phi[(PWP')^{-1} \otimes W - (PMP')^{-1} \otimes M]\Phi' B^{-1},$$

where $B = \Phi[(PMP')^{-1} \otimes (M - W)]\Phi'$, Φ is a constant matrix determined by the restrictions, a is in a special sense the vector of the independent or unrestricted variables involved in P, and $\otimes$ indicates the Kronecker product of two matrices.

Among the circumstances that can be used to reduce the number of computations per iteration is the fact that Φ is usually an extremely simple matrix most of whose elements are zero or one. In the above-mentioned eleven-equation system, the sizes of the matrices d, a, Φ, B, respectively, are (1×60), (1×60), $[60 \times (40 \times 11)]$, (60×60).

Thus to treat the above system or larger ones it is necessary to have a machine which can rapidly refer to large matrices which are kept constant throughout the iterations to perform over a dozen matrix operations per iteration and preferably to make good use of the simplicity of certain matrices (Φ).

REFERENCES

1. The condition of normality can be relaxed without seriously affecting the properties of the estimates obtained in the manner to be described below. The possible generalizations of conditions are considered in more detail in a forthcoming monograph of the Cowles Commission.

2. See H. Chernoff and J. Bronfenbrenner, "Gradient methods of maximization," Cowles Commission Discussion Paper 332, and T. C. Koopmans, H. Rubin and R. Leipnik, "Measuring the equation systems of dynamic economics," Art. II (esp. Sec. 4) in "Statistical inference in dynamic economic models," Cowles Commission Monograph 10.

Physiology and Computation Devices

WILLIAM J. CROZIER

Harvard University

In discussing briefly the significance of high-speed calculation instruments for inquiry in the realm of physiology I am putting to one side, for present purposes, those aspects of organic activity which involve interactions between and among individuals and between groups of individuals and environmental forces or conditions. Undoubtedly, as data accumulate and analytical insight grows (although the two may not march in close company), the service of such devices can be great in reducing the labor of calculations required for the testing of hypotheses and the making of predictions. Even here, however, one can doubt whether it will be possible or even desirable to resort to the push-button selection of hypotheses concerning ethology, ecology, population genetics, and the like, let alone for the setting up of differential equations basic to the erection of quantitative conceptions. The problem must be stated before a solution can be tested.

It is rather in connection with questions of the nature of the individual organism and its constituent processes that some less obvious considerations are required. Two foci of interest are clear. The sheer complexity of these processes requires that as understanding of them improves it will be more and more necessary to recognize, for theory building, that the organic property under scrutiny is essentially a multivariate one, in which many parameters are significant. In the end, now only dimly approached, it is likely that all available mathematical devices and aids to calculation will be called upon for real progress toward formulation and perhaps comprehension.

This first point of interest finds parallels in, for instance, weather prediction or even in cosmology, where the basic difficulty is rather a mechanical one, arising from the degree of appropriateness of the mathematical system applied and the onerousness of using it, granted data that are sufficient. Beyond this is the second focus of interest, essentially more attractive and more stimulating. It centers upon the possibility of creating models of individual biological processes. Such models, both mathematical and material, have of course played a considerable role in general physiology—as aids to the clarification and the concreteness of thinking, and as springboards for new experiments. Somewhat crudely put, but not unfairly, the question has arisen: Do complex, fast, computation devices provide an effective model of mental processes, or even of one general class of human cerebral operations—even to the extent that such machines, with developments of kinds now foreseeable, may be used to serve as surrogates for human decisions or actions?

Note that there are here two distinct questions. The adequate imitation of a given kind of end result as achieved by an organism does not at all imply that the mechanism whereby

the organism acts or decides has been duplicated. For engineering purposes, as in the "no hands" operation of a production line, this may be quite immaterial (so long as men keep the surrogate in good working order). But the physiologist's job is different. What he seeks is not merely an over-all model. He really looks for an understanding of the actual mechanisms whereby the organic, biological machine operates. An even partially successful model may be enormously helpful in furthering this quest, in a particular case; but its character as an analogical crutch must not be lost sight of.

The biologist is not necessarily too grumbly about this—or at least I should not like to have it thought that he is. He is perfectly able to accept, for the time being, a system of kinetic equations embracing the data of photosynthesis as picturing the known essence of the mechanism of this process. He also wants to learn the molecular inwardness of the matter, the kinetic mechanism with all its defects serving as a ladder toward specific experimental inquiry. He is not so crude as to look for a nexus of springs pulling dashpots through baths of hydraulic oil when he peers at muscle fibers in electron-microscope pictures, though this kind of model has had brilliant uses for some purposes. On the other hand, he knows in a practical way, as the logician knows, that reasoning by analogy from dynamical properties of the model is liable to stumble over imperfections in the analogy. Gross instances are available in which properties of the mechanical model alone, not recognized at the time as such, have misled inquiry into blind alleys.

An illustrative case is in point. It is drawn from the field of visual excitation, which has many advantages for my present purpose. It is significant as illustrating several principles important for the method of evaluating analogs for organic events and processes. It happens that a variety of visual data have been submitted to description in terms of the notion that a visual (seen) threshold effect is brought about when the incidence of light produces a fixed amount of freshly formed decomposition product of a photosensitive material in the retina (investigations by S. Hecht). One type of such data was that presented by the contour of critical flicker frequency as a function of flash intensity. The equation, derived from an experimental foundation, used for description (by visual, aesthetic test of fitness) implied certain consequences as necessary if the temperature of the organism were to be altered or if the light-dark sequence in the flash cycle were to be changed from that commonly characteristic in such experiments, namely, 1:1. The test of the cogency of this formulation cannot be made by any process of curve fitting. It has to be made by the use of what I may be allowed to call *parametric analysis*—the deliberate modification of experimental conditions in such ways as to reveal the number and the (or some) quantitative properties of the parameters of the necessary formulation. In this specific case, it was easily shown (Crozier) that the photostationary-state equations were incompetent because they completely failed to predict even the direction of the modifications of the flicker contours when tests such as those already referred to were made.

It is interesting to note in outline just what had happened here. It provides a useful commentary on the dangers of analogy. The visual systems of animals exhibit reversible

changes of excitability according to the prevailing illumination to which they are exposed. As a first approximation, it was inevitable that help should be sought from the roughly parallel kinetic properties of known reversible-reaction systems in physical chemistry. This involved, however, the crass assumption that seeing, as a decisive behavioral act, occurs in direct proportion to retinal events. It also involved the delusion consequent upon the implicit assumption that the retina is, in the sense of chemical kinetics, a homogeneous medium. It likewise avoided recognition of the patent fact that the equations to which the initial assumption led were not in fact unique, even accepting their apparent descriptive adequacy for the data. In form the equation is identically a logistic in which the light intensity enters as the logarithm. This of course cannot be distinguished from a Gaussian integral except by very precise measurements extending close to the asymptotes. Data of this kind, as well as the quantitative properties of the parameters, help to decide in favor of the logarithmic-Gaussian equation.

The important, central fact is of course that the visual system—even at the retina, and even in a single receptor unit—is obviously a microheterogeneous system. It comprises assemblages of semi-isolated reaction chambers. Thereby is generated an essentially statistical situation. Without enlarging here upon details, it can pretty certainly be shown that here one really has a situation in which the assumptions used by Gauss are implicit; hence sensory effect is expressed as a Gaussian integral, which is logarithmic in the abscissa in common cases because the basic parameter fluctuates spontaneously.

There is involved here the conception, for which there is direct observational support, that at sundry levels in the hierarchy of assemblages of neural units and subunits implicated in a behavioral decision or act one has to do with populations of units which individually fluctuate in their contributions to the end result. In the light of this we may examine briefly several concrete cases. From a study of them there emerge two considerations, forming the core of what is here suggested. The first is that any model of a mental process, which to be analytically significant must include in its dynamical structure the essential nature of the process, must in this context operate in such a way that decisions are automatically achieved by thorough statistical comparisons. The second is that it is precisely in this way that elementary "mental" decisions are arrived at.

One is thinking here of elementary decisions neurally mediated. If one is to have a model of mental processes rather than merely a mimicking of their results, it is necessary to know sufficiently the quantitative properties of that which it is desired to imitate. It is only with respect to such elementary decisions as those, for example, involved in the conscious discrimination between the neural effects due to two intensities of illumination, that one has anything like data adequate for the beginnings of quantitative treatment. The concrete cases exhibited are therefore chosen from this category. They are given as illustrative only, and no real description of their analysis is attempted here.

Take first a case having intimate relation to the problem of willful indeterminacy. How many quanta of radiation in the visible range are required to evoke the elementary visual act (threshold effect)? It has been asserted that on ordinary probability considerations

involving the particulate aspect of the structure of light, and because the number of available quanta may be quite small, the frequency of seeing of intensities differing near the median threshold by small steps forms a Poisson integral, reflecting randomization of the quantal content of repeated flashes intended to be identical. Two curious assumptions are involved here. One is that only the stimulus is variable, and that the organism is constant (during the experiment) in its capacity for excitation. The other is that one sees with the retina. Prejudice aside, each of these assumptions has to be definitively rejected. For formal purposes it is sufficient to show that the seeing-frequency function is not a Poisson integral. It is a simple matter to demonstrate that this function is symmetrical in log ΔI_0 (or in log $t_{exp.}$), and that its mean and its variance are totally independent quantities. This latter is proved by elementary application of parametric analysis: mean and standard deviation are shown to be independently variable, by the empirical findings when such relevant conditions as wavelength of light, image area, retinal location, oxygen partial pressure in the air respired, and the like, are systematically changed.

To the question, therefore, how many quanta are required for minimal visual excitation, there is required a complex answer. It includes the limitation of "how often." It is further complicated by the fact of photosensitization, which has been shown to lower visual thresholds by a very considerable amount. For such reasons, as well as others, I purposely do not deal here with the asserted correspondence, in fact illusory, between the results of seeing-frequency tests and the outcome of calculations based upon absorptive losses of light between the cornea and the receptors.

The notion that a small number of quanta may suffice for a neural decision as to the conscious presence of light introduces the possibility of disordered capriciousness—indeterminacy—in this mode of conduct. In no respect of real analytical interest can this be supported. In fact, in every instance for which we have data it is found that the standard deviation of the critical intensity is rationally related to the mean value of this intensity. The measurements have been given in a number of publications. Where, as in the observations based upon marginal recognition of visual flicker, it is possible to consider also the converse experiment, in which at fixed flash intensities the variation in critical flash frequency can be determined, it is found that the properties of this variation can be forecast from the tests made in the other way. Comparable results have been obtained in measurements of the discrimination of intensities. The fact that the index of internal correlation computed from the *scatter* of $\sigma_{\Delta I}$ is systematically related to the magnitudes of experimental conditions inescapably reinforces the conviction that although quantitative fluctuation in the basis of simple organic decisions is real, it is not lawless. Rules can be written for it, and successful predictions can be made as to its properties under novel circumstances. Therefore it is not capricious. The root basis of such elementary neural acts is not one in which indeterminacy rules.

The reason of course is that, even if a single photon should on occasion elicit a valid visual discriminatory response, this must be by way of a multiplicative amplifier central to the retina. The evidence shows that discriminatory responses are in effect the result of competitive action

between groups or populations of neural effects. The internal rules of this competition seem to be at least closely akin to those dictated by really elementary statistical considerations. For example, the standard deviation of the seeing-frequency function, as dependent on wave-number, is readily understood on the basis, substantiated by collateral evidence, that the standard deviation is larger the greater the number of potentially excitable elements of neural effects.

Measurements of intensity discrimination have played a huge part in psychophysics. Their treatment has been curiously raddled by the view that there is *a* value of ΔI, or of $\Delta I/I$, which if the universe were only kind one could get accurately. Actually, it fluctuates, under given conditions. This fluctuation, instead of being due to a cloud of irrelevant "errors," is actually the key to the matter. Illustrations are at hand showing how, from a knowledge of the scatter of critical intensities in certain kinds of experiments (flicker, visual acuity) it is possible to compute precisely the form of the curve relating ΔI to I_1. Instances of this sort decidedly encourage the conviction, which can be supported through other lines of evidence, that the measured fluctuation in organic performance is really more important analytically than the usually considered average value, and that in a valid sense these average values are determined by the capacity for variation, rather than the reverse.

These calculations of ΔI are made by "asking" the organism to discriminate between two intensities on the basis of the probability band formed by flash-frequency F *versus* log $(\bar{I} \pm k\sigma_I)$, where $\bar{I}$ is the mean critical intensity for response. With I_1 fixed, there is an associated range of effects (F). A mean intensity $\bar{I}_2$ to be just distinguished from I_1 must give a range of effects for which the mean differs from that associated with I_1 by the factor $k'\sigma_{\text{Diff}}$. On this basis the directly determined $\Delta I/I_1$ curve can be exactly duplicated. Various other properties of $\Delta\bar{I}$ as measured can also be computed—for example, those based on the fact that $+\ \Delta I$ is greater than $-\ \Delta I$.

Quite detailed examination of the simple statistical properties of other aspects of visual excitation have given additional evidence in support. Much less complete evidence, consistent with this, is available for other sensory phenomena. If it is correct, as it seems to be, that elementary neural decisions appear to be carried out on an orthodox statistical basis, then an effective surrogate in the form of a swift computer device would have to involve a very large number of elements with the capacity for fluctuation in performance. Despite the implication of a wide extrapolation, it is probably unsafe to assume that more complex neural (mental) operations can depend on other than an elaboration of the kind of complexity glimpsed in the case of simple neurosensory discriminations. A nonliving dynamical model would at least occupy a very large space, and doubtless would require considerable maintenance attention.

Until the nature of mental activities is better understood it cannot be said that the construction of a "thinking" machine is impossible. The question can be put, however, whether it would be desirable to construct such devices. We already have such mechanisms in some abundance. The task is to have it seen to that through biological engineering good brains

are produced, recognized, fostered, and assisted. Even Swift's monkeys at their type machines would require someone to tell them (directly, or by a device he had built) when they had really achieved the accepted text of Shakespeare. It will still be necessary to know the problem for which an answer is sought before "push-button" help can be invoked for the calculations. The invention of fruitful problems by nonliving means is probably quite remote. The possibility of it is, of course, rather revolting, to even a mechanist; on selfish aesthetic grounds the revulsion is akin to that aroused by the horridly inaccurate statement of Bertrand Russell's concerning artificial parthenogenesis, that "Loeb had shown that a sea urchin could have a pin for a father."

In any case the physiologist will welcome the mechanical aid and the stimulus provided by fast computers, and certainly the advice of those who have gained wisdom in dealing with them. He will be on his guard, however, about using them as giving models of neural processes, even when some end results are mimicked. He will have the conviction that, however complex their form, these models will not serve for automatic invention, but will function solely in terms of effects actually (even if sometimes unwittingly) built into them at the start.

The Science of Prosperity

FREDERICK V. WAUGH

*Council of Economic Advisers**

The science of economics attempts to determine the "best" use of our economic resources; to tell us what adjustments are needed in the use of our land, labor, and capital in order to increase our wealth and prosperity.

This is obviously an important task, and a difficult one. We must be concerned with the economic decisions made by millions of primary producers, investors, distributors, and consumers. We must gather and analyze great masses of statistics. The large-scale computer may prove to be of great value in such research.

But before all economists and statisticians jump on the electronic bandwagon, it might be well to note that there is still room for economic research of kinds that can be done with the simplest of statistical techniques—often without even an ordinary calculating machine, or even a slide rule.

This is true of two important groups of studies: first, those dealing with small, isolated segments of the economy; and, second, those dealing with a few aggregates for the economy as a whole.

As examples of research dealing with small segments of the economy, take such questions as:

1. How much grain should be fed to a dairy cow?
2. Will it pay a particular manufacturer to buy a new machine?
3. What kinds and amounts of advertising will be most profitable in a given situation?
4. What changes are needed in the Boston fruit and vegetable market to make it more efficient?

These questions—and thousands of similar questions—can be studied by gathering very simple kinds of statistics and by using elementary methods of analysis. The first three questions deal with problems confronting individuals, or single firms. In such cases it can often be assumed that the decision made by the individual or firm will not affect the rest of the economy significantly. For example, if the dairy farmer feeds his cow a little more grain we can neglect the effect of this action on the prices of grain and milk. The fourth question, concerning the Boston fruit and vegetable market, is somewhat more complicated. To answer it we do need some detailed statistics on the costs, demands, and habits of several hundred buyers and sellers. But, basically, the analysis is simple, requiring little more than the addition of figures to determine peak volume, degree of overlapping in transportation, possible savings by changing location, and so on.

* Read at the Symposium by Leon Moses, *Harvard University*.

As examples of research dealing with a few aggregates for the whole economy, take such problems as:

1. The relation of potato consumption in the United States to the average retail price of potatoes, and to the disposable income of consumers.
2. The relation of consumer expenditure (and saving) to the level of income.
3. Cycles in production, employment, and prices.

In these cases—and many others like them—we deal with only a few variables. We should, of course, remember that the "system is not complete" because we neglect many minor variables, and because we lump together in one aggregate some things that may not be strictly the same for our purpose. Such studies require only the simplest kind of analysis by multiple, or joint, correlation. Often graphic methods are preferable to mathematical computations.

Now these types of problems probably are uninteresting to most mathematicians and computers, but they are extremely important, nonetheless. The vast majority of economic studies probably always will be, and should be, devoted to work of this kind. I think it is worth while to recognize this at the start. We have fashions in economic research, as elsewhere. Twenty-five years ago, when I was beginning to do research in agricultural economics, the current fad was multiple correlation. No self-respecting graduate student would write a thesis that did not include at least one multiple regression equation, even if neither he nor his professors knew what the results meant. Let us avoid a similar fad of large-scale computation, applied indiscriminately to all economic problems. The economist should apply economy to the number of variables, and to the complexity of the analysis.

After these words of caution, it is necessary to go on, and to say that there is a crying need for a few basic studies dealing in some detail with the whole economy.

This need was first recognized during the industrial mobilization for war production. A thorough analysis of our economic potentials by large-scale input-output techniques would have been extremely valuable. It is likely to become a vital part of industrial and military planning in the future.

We need these comprehensive studies in times of peace, as well as in times of war. This need was brought into focus by the Employment Act of 1946. That Act declared that the policy of the Federal Government is to utilize all its resources "to promote maximum employment, production, and purchasing power." It requires the President, with the assistance of the Council of Economic Advisers, to determine existing, and needed, levels of employment, production, and purchasing power. Also, it requires him to transmit to the Congress a program to bring about the needed levels.

This responsibility cannot be fully met by studies of aggregates alone. True, we have made real progress in developing a Nation's Economic Budget which provides estimates of total incomes and expenditures of four main economic groups: consumers, business, government, and international. Doctor Colm, of our staff, is working with economists in the government service and in the universities to develop the Nation's Economic Budget into a powerful analytical tool that will tell us the main adjustments that are needed. But I know that Doctor

Colm, and the other experts in the analysis of national budgets, clearly see the need for going beyond these major aggregates, for making rather elaborate breakdowns of many items, and for detailed study of the interrelations among the thousands of variables that make up the totals. The national budget analysis supplements this more detailed work, coördinates it, makes it more useful—but is not at all a substitute for it.

What do we mean by "maximum employment, production, and purchasing power?" Do we mean simply that x million persons should be employed, producing y billion dollars' worth of goods and services, and spending z billion dollars?

The answer obviously is no. We need to know how much of each *kind* of employment and production is needed, and how expenditures need to be allocated among various groups of consumer-goods and investment categories. We can have overemployment in agriculture and underemployment in manufacturing. We can have overproduction of nondurable goods and scarcity of durables. We can have too much spending by some groups and too little by others.

The "needed levels" must be broken down in considerable detail as rapidly as we can. Since the beginning of the war, the Department of Agriculture has determined production goals for the major crops and livestock products. These goals indicate state by state how much corn, cotton, or milk is needed. We need similar specific goals for clothing, housing, coal, steel, automobiles, roads, schools, health facilities, and so on.

Here, as I see it, is where the large-scale computer comes in. Our goals are worthless unless it is feasible to reach them. We must find out what combinations of goods and services it is feasible to produce. The structural equations of Leontief presumably can provide a basis for making usable estimates of the combinations that are technically possible with a given labor force, plant and equipment, and land.

This will, of course, involve the inversion of very large matrices—perhaps matrices with over 100 rows and columns. But, as I understand it, this computational job can be completely licked by the large-scale computer. True, some methods of inversion may give us large errors due to the compounding of rounding errors. But in the Leontief matrix

$$S = [I - A],$$

the norm of A (or the sum of the absolute values of the elements in any column) is necessarily less than 1; and therefore S^{-1} can be computed to any desired degree of accuracy by iteration based upon the equations

$$S^{-1} = I + A + A^2 + A^3 + \cdots,$$

or

$$S^{-1} = [I + A][I + A^2][I + A^4] \cdots$$

This is essentially the method used by the Bureau of Labor Statistics.

Of course, the results are no more accurate than the original data. Doctor Morgenstern[1] has performed a service by emphasizing the need for more reliable economic data. But I am not sure that this problem of faulty data is any worse in connection with large Leontief matrices than with analyses involving a few variables.

For example, let

$$S = [I \pm D][A],$$

where S is the true (unknown) structural matrix, A is the estimated matrix, and DA is the error matrix. Then

$$S^{-1} = [A^{-1}][I \pm D]^{-1}.$$

In many practical cases I believe we might assume $[I \pm D]$ to be a diagonal matrix, in which the element d_{kk} is our estimate of the highest likely proportional error in the coefficients representing purchases by the kth industry. Then $[I \pm D]^{-1}$ is a diagonal matrix, the kth element being $I/(1 \pm d_{kk})$. So multiplying the kth row of A^{-1} by $1/(1 \pm d_{kk})$, we have simple limits for the elements of the true inverse. These limits do not appear to be affected by the size of the matrix. The problem appears to be no worse for a problem of 100 variables than for one of two variables. Of course, it may be bad in both cases.

We need better data, and we need more detailed breakdowns. But I am optimistic enough to think that we can in the near future make usable approximations of the combinations of goods and services that could be produced. If you can help the economist do this you will have made great progress.

This, in itself, of course is not enough to determine a good set of goals for the economy. It is an essential first step. But in a democracy of free people the goals are finally determined by the citizens. They take into account many special factors, such as the amount and kinds of work they want to do, and the minimum standards of health and diet that are guaranteed to low-income families. But to make intelligent choices, we citizens have to know what the alternatives are, and presumably the large-scale computer can give us useful information about these alternatives.

Our goals cannot be static. The choices we make this year affect our economic output next year. We may choose high savings and investment now in order to build up our productive capacity for the future. We may choose a temporary deficit in the government budget in order to forestall a depression. We may store up surplus foods to avoid excessive fluctuations in market supplies and prices.

One of the main objectives of the Employment Act is to find practicable ways of avoiding alternate periods of boom and depression, and to encourage a steadily rising level of living for all groups. Timing is an essential feature of economic policy in this field. Policies that would be excellent in an inflationary boom would be bad in a depression.

Can the mathematician, with the help of large-scale computers, give us the information we need to make intelligent choices about timing? This is an intriguing subject in mathematics and in economic theory. For the economist, at least, it is a very difficult subject. Dynamic economic theory is in an elementary stage. I doubt whether many of us economists know what questions to ask the computer.

Fortunately, the need for a dynamic economic theory is widely recognized, and some progress is being made in this field, especially by the mathematical economists and the econometricians. They will doubtless have work for the large-scale computer to do. If this

work is to be fruitful, it must avoid formalistic mechanical computations. It must be designed to help us understand real economic problems, and to show the probable results of various economic policies that are being considered.

The dynamics of the whole economy is necessarily complicated, and we need many studies of particular segments of the economy. For example, one of the really effective stabilization devices we have developed in the United States has been the price-support program for farm products. But this program can be greatly improved—not only as a means of supporting farm income, but as a part of a general economic program to maintain high employment, production, and purchasing power throughout the economy. What price should be supported for wheat or cotton? If the supports are set too low they will not prevent a serious drop in farm income. If they are set too high they may reduce domestic consumption, lose foreign markets, result in excessive storage, and perpetuate either overproduction or strict controls over acreage and marketings.

What levels of agricultural price support are feasible, and would be of lasting benefit to farmers and to non-farmers? The question can be answered only in dynamic terms, because we are concerned with the effects of today's support levels upon future production, marketings, and stocks. Research on this more limited subject would be considerably less ambitious than an attempt to analyze the dynamics of the whole economic system. But even this limited research would involve a large number of variables, and a fairly complicated analysis.

Farm-price supports represent, of course, only one example of the many dynamic studies that are needed to provide a good basis for sound economic policies. We need to know about the dynamics of taxes, of credit and monetary policies, of public-works programs, and of social security. The e all involve rather elaborate and difficult problems of economic theory. The economist must get his theory stated in terms that can be tested and quantified by statistical analysis.

Doubtless the large-scale computers will have plenty of work to do in this general field.

Can the scientific economist, mathematician, or statistician really determine *maximum* employment, production, and purchasing power? The word "maximum" is a mathematical word. I have great respect for the mathematician, and I assume that if we could define what it is we want to maximize he could tell us how to do it.

The trouble is, I think, that we really want an *optimum* allocation of resources in some broad sense that is difficult to pin down in precise mathematical terms. Certainly it is conceivable that we could determine the allocation that would maximize the Gross National Product, or that would minimize the hours of work necessary to reach a given level of living. With large-scale computers, we should be bright enough to work out such maximum solutions, or minimum solutions.

Some work along these lines will doubtless be useful. I will confess, however, to some personal skepticism as to the finality of such studies. With the help of methods developed by Dantzig,[2] I recently worked through the interesting problem proposed by Stigler:[3] to choose from the list of foods quoted by the BLS a diet that provides at the lowest possible cost at least

the minimum needed amounts of nine nutrients. With only nine variables it can be handled with an ordinary computing machine. My answer is that in 1939 the minimum-cost adequate diet could have been purchased for $39.66 a year, or 10.87 cents a day, for each person. Over one-half the expenditure would have been for dried navy beans. The only other foods on the list would have been wheat flour, beef liver, spinach, and cabbage.

I don't want the minimum-cost adequate diet. I am not sure that we want to maximize the GNP, nor to minimize the labor in producing a given bill of goods.

The problem of goals is necessarily complicated. The results of mathematical maximization probably should be taken *cum grano salis*. Still they doubtless can, at least, tell us some of the main adjustments that would be needed to raise national income, or to increase the total value of goods and services available to the consumer.

These studies, in themselves, do not provide goals to be imposed upon the public, and to be enforced by government policies and programs. Their value is simply in showing the public what is possible, and feasible, with the resources we have at our disposal.

With large-scale computers we should not limit our work to the search for the maximum maximorum. We should analyze all of the main alternatives. This is true not only of war-mobilization studies, or the analysis of foreign-aid programs; it is even more true of research on our peacetime economic potentials. We need to know what the alternatives are. We need then to know what policies or programs would be required to reach any goals we might set—for example, tax and fiscal policies, farm price supports, wage policies, and so on.

If our research can do this, it will give the Congress and the general public a scientific basis for deciding what is really an optimum solution. In a democracy the majority of the citizens determine our economic goals, and the policies and programs to be used to reach the goals. If their decisions are based upon full and correct information, the goals they set are really optimum solutions in a more basic sense than any solution that could be computed by a mathematician or economist. So I will close with this awful thought: that the elaborate economic studies made possible by the large-scale computer will be worth very little unless, or until, the results are explained in simple terms to the general public.

REFERENCES

1. O. Morgenstern, *Am. Econ. Rev.*, 238–240 (May 1949).

2. G. B. Dantzig, "Procedure for maximizing a linear function subject to linear inequalities," USAF Comptroller (January 1948).

3. G. J. Stigler, "The cost of subsistence," *J. Farm Econ.*, 303–314 (May 1945).

EIGHTH SESSION

Friday, September 16, 1949

2:00 P.M. to 4:00 P.M.

DISCUSSION AND CONCLUSIONS

Presiding

Willard E. Bleick

U.S. Naval Academy Post Graduate School

The Selectron

JAN RAJCHMAN

Radio Corporation of America

The initial work on a special type of electrostatic memory tube, called the Selectron, was reported on January 8, 1947, at the first Symposium on Large-Scale Computing Machinery.[1] The present paper is a report of the tube developed as a result of work in progress since that date. Important changes in the initial tube were found necessary to obtain a practical and reliable device for use in electronic computers.

The memory of an electronic computer can be idealized as a large set of cells, each identified by a coded address and each capable of retaining a single on-off signal. A combination of such signals occurring simultaneously on several channels or sequentially on a single channel constitutes a number. The memory will be particularly useful if the occurrence of the set of pulses specifying the address will give access to the signal stored, or to be stored, in the shortest possible time without consideration of any previous selection. A device with such a digitalized address system and such direct access to any stored signal can be used singly or in groups in a most flexible manner, since no amplitude-sensitive qualities have to be dealt with and no specific sequences are intrinsic to the memory.

The Selectron (Fig. 1) is a vacuum tube designed in an attempt to realize such an ideal memory device. The principle of the tube depends on quantizing both the address of the stored information and the information itself. The selection of the address is obtained by means of two orthogonal sets of parallel spaced metallic bars forming a checkerboard of windows. A shower of electrons impinges on this checkerboard. Electrons are stopped in all windows except in a selected one by applying address-selecting voltages to certain groups of bars connected into appropriate combinations. The storage is in terms of the two stable potentials that tiny floating metallic elements, located in register with the windows, assume under continuous electron bombardment. The reading signals are sizable electron currents passing through a hole in the storing elements. The signals produce also a visual monitoring display.

The basic principle of the Selectron has not been changed. The main improvement is the use of discrete metallic eyelets as the storing elements. In addition to very reliable storage, these eyelets have a "grid-action" effect yielding strong electronic reading signals.

The Selectron tube, called SE256, has 256 storing elements, is 3 in. in diameter and 7 in. long, and utilizes a 40-lead stem. The diametral and axial cross sections of the tube are shown in Figs. 2 and 3. Eight elongated cathodes of rectangular cross section are located in a diametral plane of the tube. Between and parallel to the cathodes are a set of nine selecting bars of square cross section. These vertical selecting bars are connected into six groups: V_1, V_2, V_3, V_4,

and V_1', V_2', as shown in Fig. 4. On either side of the plane of the cathodes and V bars there is a set of 18 parallel bars of square cross section at right angles to the V set. These two sets of horizontal selecting bars sandwich the cathodes and V bars as do all subsequent electrodes

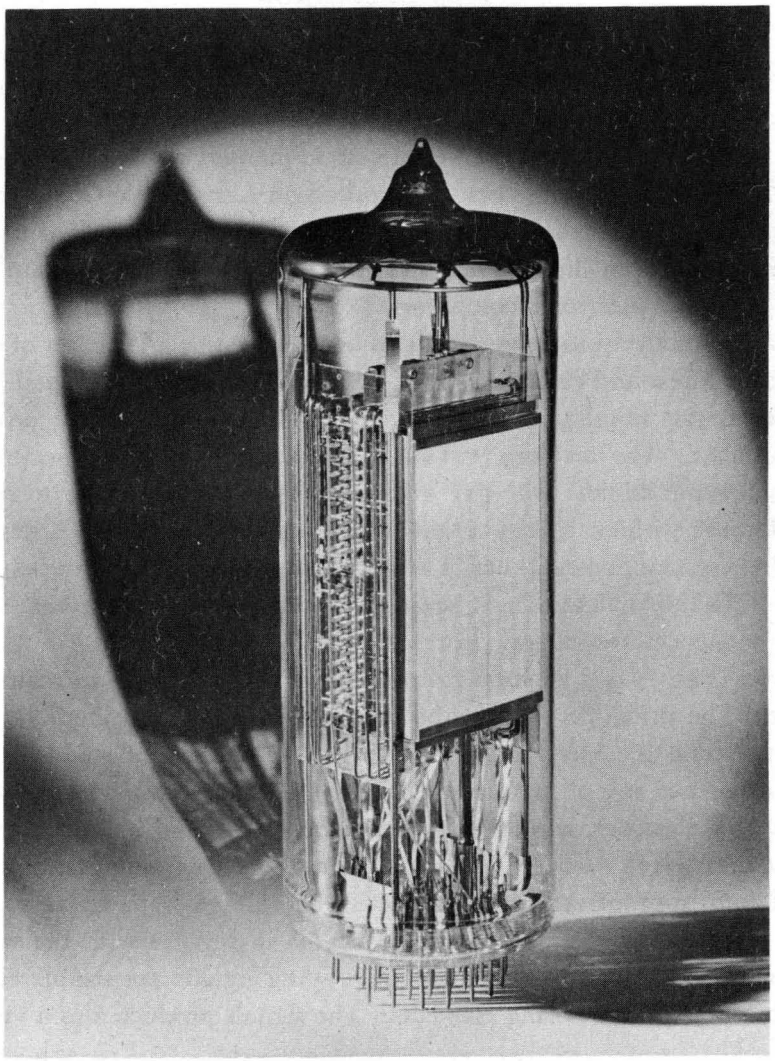


FIG. 1. The Selectron.

of the tube, the tube being symmetrical with respect to the cathode plane. The 36 horizontal selecting bars are connected in 12 groups: H_1 to H_4 and H_1' to H_8', as shown in Fig. 4. There are nine vertical bars for eight gates and 36 horizontal bars for 32 gates, the excess bars taking care of the end effects.

On either side beyond the horizontal bars there is a collector made of two flat plates perforated with round holes whose centers match the centers of the windows formed by the

V and *H*-bars. Adjacent to the collector plates there are two perforated mica sheets holding between them 128 metallic eyelets. These eyelets, made on automatic screw machines, have a conical head, a center hole, a holding collar and a shielding tail. They are nickel-plated

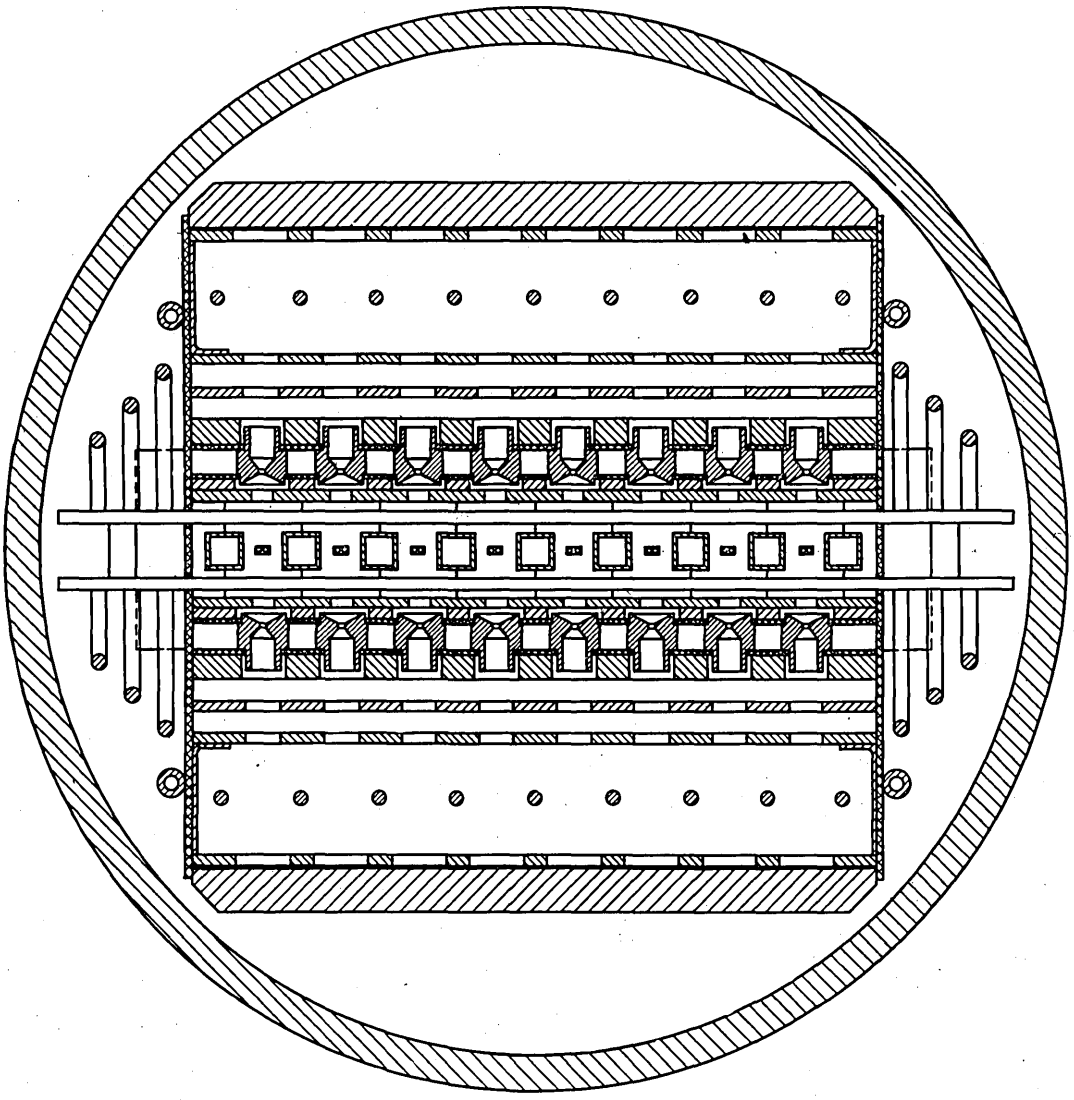


FIG. 2. Diametral view of Selectron.

steel. On the other side of the two mica plates is another perforated metal plate—the writing plate. The two collector plates, the two eyelet mica plates, and the writing plate form a tight assembly riveted together at the ends and in the center.

Beyond the writing plate is another metal plate—the reading plate—perforated with holes in register with the holes of the other plates. Beyond it is a Faraday cage formed by two perforated plates spaced some distance apart and closed on all four sides by a metallic wall.

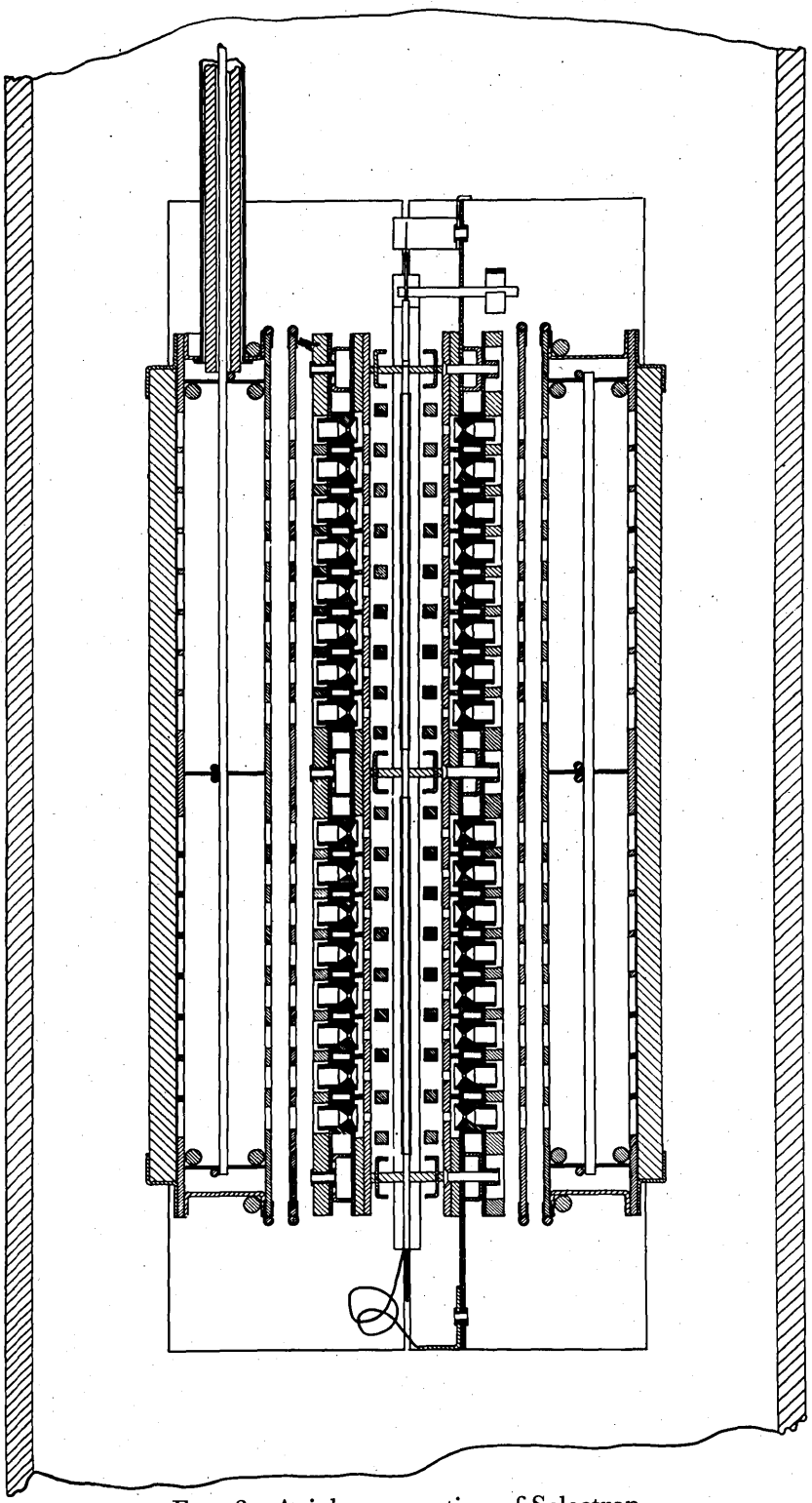


FIG. 3. Axial cross section of Selectron.

A glass plate coated with a fluorescent material is placed against the outer plate of the cage. In the central plane of the cage there are nine wires which are spaced so as to be between the holes of the perforated plates. These reading wires are connected together and the corresponding lead to the stem is shielded.

In the quiescent state of the tube storing information previously written-in, all the selecting bars are at the potential of the cathodes (0 v) and all other electrodes at potentials indicated in Fig. 5. In this condition electrons emitted from the cathodes are focused into 256 beams by the combined action of the *V* and *H* bars at zero potential and the collector plate at some positive potential, such as 180 v. These beams are focused through the centers of the collector holes and are directed on the eyelets. Since the eyelets are not connected anywhere—are electrically floating—their potentials will adjust themselves so that the net electron current to them is exactly zero. It turns out that there are two naturally stable potentials for which this is the case. This can be understood by examining the current to the eyelet as a function of its potential as shown in Fig. 6. When the eyelet is more negative than the cathode, no current reaches it because it repels any incurring electrons. As the eyelet is made more positive, some electrons strike it, producing a negative current. At a still more positive potential, secondary emission from the surface of the eyelet starts as a result of the primary bombardment and tends to cancel the negative current, being a loss of negative charge. Eventually, the two are equal at the so-called first crossover. For still more positive potentials, the secondary emission is greater than the primary emission and a positive current is obtained. Finally, when the eyelet reaches the collector potential and becomes more positive, the secondary electrons are suppressed owing to a retarding field at the surface of the eyelet. The current therefore passes through zero again to become negative. It will be recognized that the cathode and the collector potentials are stable, because a deviation from the zero-current potentials tends to produce a current in a direction tending to restore the equilibrium potential. The first crossover point, on the other hand, is unstable. The restoring current at the two stable potentials makes up for any possible detrimental ohmic or ionic currents. Therefore, any

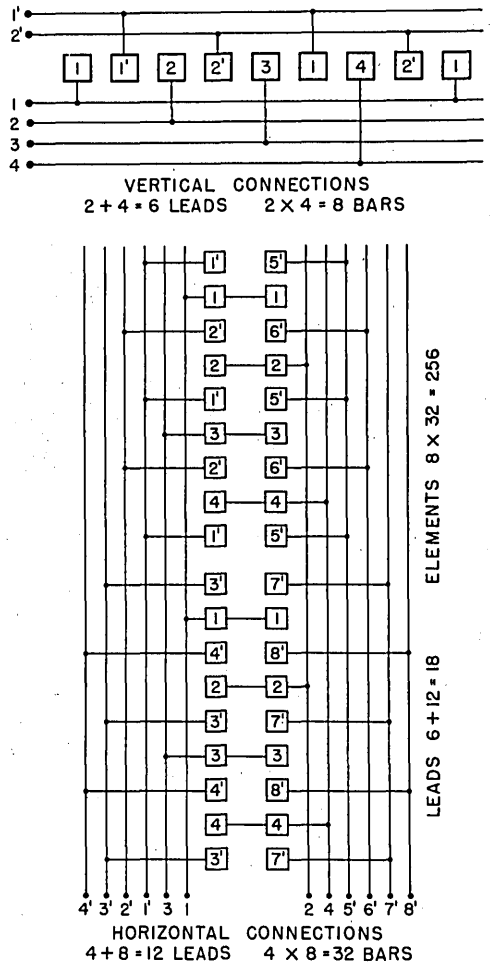


FIG. 4. Connections of selecting bars.

eyelet left in one or the other of the two potentials will keep it indefinitely (as long as power is on the tube) without any deterioration of information whatsoever.

To write or read into or from the memory, the quiescent state of the selecting V and H bars is momentarily disturbed so that the current reaches only the one selected eyelet into which writing or from which reading is desired. This is accomplished by applying a negative pulse to all the selecting V and H bars except one in each of the four groups V, V', H, and H'. The bars are connected in such a way that one and only one gate in each of the V and H

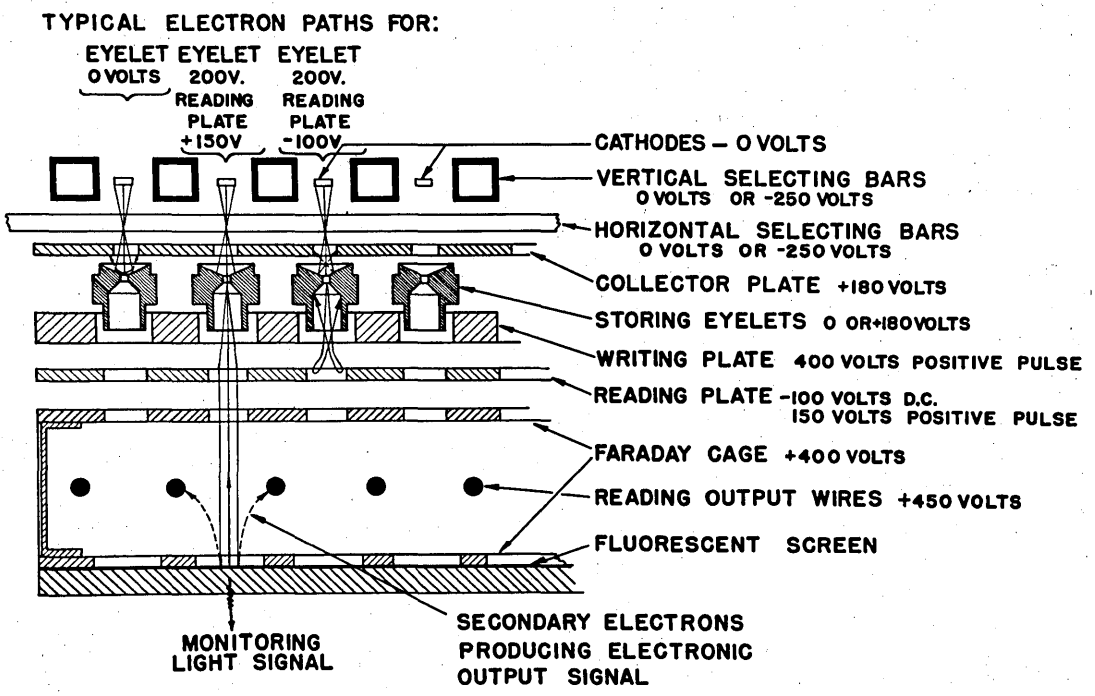


FIG. 5. Operating potentials of Selectron elements.

directions will have its two limiting bars at cathode potential, while all others will have one or both limiting bars at the pulsed negative potential, as can be seen by examining Fig. 4. When a V or H bar is sufficiently negative it cuts off almost entirely the current from the adjacent cathode or cathode location and the small remaining part is deflected and does not reach the hole of the collector. When both sides of a gate are negative, a potential barrier is formed through which no electrons can pass. It follows, therefore, that only the particular selected window with its four bars at zero potential will still have its original current, while all others will be completely cut off.

This principle of selection operates on the basic idea that both sides of a gate have control of the passage of electrons through it and that therefore combinatorial systems of connections are possible by connecting each side of the gate to appropriate sides of other gates. In fact, since this is done in both directions, a fourth-power relation exists, in general, between the

number of necessary connection groups and the number of controlled windows. Since each connection group is connected through the vacuum envelope of the tube and is controlled by an external circuit, the economy in the number of connections is of particular interest when tubes with larger capacity are contemplated. The fourth-power relation has of course a spectacular effect in this case; for example, 128 leads can be made to control 1,049,576 windows.

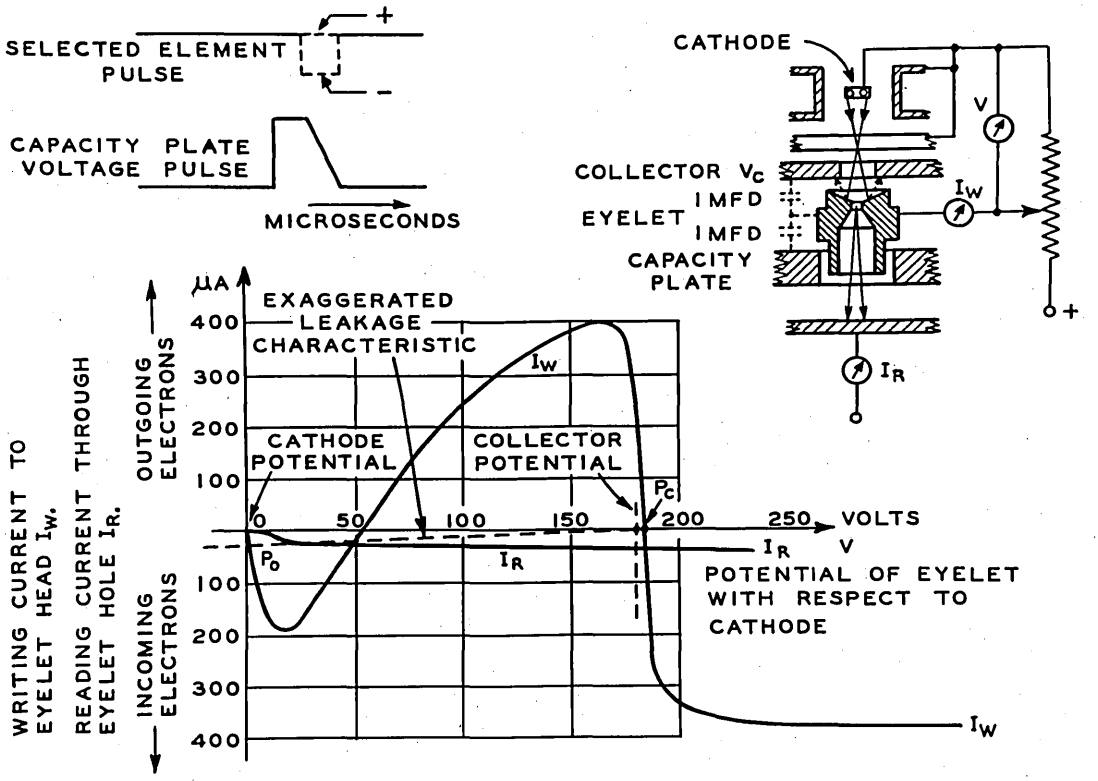


FIG. 6. Current to eyelet as a function of its potential.

Writing and reading are done one element at a time (or two if the tube is used as a two-channel device) and require selection.

To write into a particular element, current is interrupted everywhere except to that element. Then a voltage pulse of the shape shown in Fig. 6 is applied to the writing plate. Because of the capacitive coupling between the eyelet and the writing plate, the rapid rise of this pulse will cause the eyelet to jump up in potential by an amount adjusted to be a substantial proportion of the collector potential or more. If the eyelet was initially at cathode potential, it will now have been brought near collector potential and will settle at that potential during the plateau of the pulse. If it had initially the collector potential, it will acquire momentarily twice the collector potential and will receive substantial negative current (see Fig. 6) which will also bring it to the collector potential during the plateau time. Whatever

the initial condition, at the end of the plateau time the eyelet will be at collector potential. At this instant the choice is made between positive and negative writing. For positive writing, no additional pulses are applied to the selecting bars, and the current remains on the eyelet during the relatively slow decay of the writing pulse. The decay is slow enough to allow the electronic locking current to keep the eyelet at the collector potential in spite of the displacement capacitive current tending to drag it to cathode potential. This "slow" decay is in fact only one to several microseconds. For negative writing, an additional pulse is applied to one or more of the four selecting bars in the groups V, V', H, and H', which cuts off the current to the selected eyelet during the decay time of the writing pulse. The capacitive down drag is therefore not counteracted and the eyelet is brought to cathode potential.

Immediately after the end of the writing pulse the selection pulses end, and current is reëstablished to all eyelets. Only residual ohmic (on other second-order electron or ionic currents) affect the unselected eyelets during the short selection time, and therefore at the end of the writing pulse they have almost their original potential. This potential is reached almost immediately thereafter by virtue of the stabilizing currents.

The reading signal is derived from the current passing through the central hole in the eyelets. Part of the current directed at the eyelet is directed at that tiny hole. When the eyelet is positive, at collector potential, the electrons directed at the hole go through it by virtue of their inertia. When the eyelet is negative, at cathode potential, it exercises "grid action" and electrons are repelled and do not go through the hole. The electrons' paths are shown in Fig. 5 for the three cases, while the current characteristics are shown in Fig. 6. The presence or absence of the current through the eyelet is therefore an indication of the state of the eyelet.

In the quiescent state of the tube the reading plate is biased off negatively and the reading current going through all the positive eyelets (any number from 0 to 256) does not reach the reading circuits. To read, an element is selected by applying negative pulses to all but four bars, as explained above. Immediately thereafter a positive pulse is applied to the reading plate which allows the current through the selected element, if current there is, to proceed to the output electrodes. The electrons penetrate into the Faraday cage, strike the fluorescent screen, producing a light signal, and also cause the emission of secondary electrons. These secondary electrons are collected by the reading wires which are connected in parallel and constitute the reading output signal. The reading wires have a low electrostatic capacity and are well shielded from capacity pick-up by the Faraday cage.

For monitoring purposes it is convenient to bias positively the reading plate. A display of the stored pattern appears then on the fluorescent screen.

The main characteristics of the Selectron SE256 may be summarized as follows. The tube has a capacity of 256 on-off signals. The storage time is indefinite. The access time to any element is approximately 10 μsec and is independent of all previous accesses to other elements. The address selection is by means of combinations of non-amplitude-critical pulses of about 200 v applied to circuits with pure capacitive loading of 10 to 20 $\mu\mu$f. The writing and reading require also pulses whose amplitude and duration have considerable tolerances and are applied

to pure capacitive loading, 200 $\mu\mu$f for writing and 50 $\mu\mu$f for reading. The output is a direct electronic current of 20 to 40 μamp per element. The tube is its own monitor. The supply voltages have wide tolerances. The total power dissipation is 40 w.

About a score of tubes have been made to date. These tubes were tested first by d.c. or simple pulse tests. Uniform characteristics of selection and control have been observed in all tubes, as these depend on geometric factors that are easily reproducible. The cathode emissions and secondary emissions of the eyelets were also found essentially uniform. The period of quiescent-state storage has, of course, been found to be as long as desired or as there was patience to observe it.

A program has been initiated to test the tubes in conditions as similar as possible to those of an actual computer straining its memory severely. The system consists of taking two Selectrons, setting an arbitrary pattern of stored information in one of them, interrogating the elements of that tube one by one in succession, and registering the answers in the corresponding windows of the other tube. The stored pattern will thus be transferred from tube No. 1 to tube No. 2. The pattern is then transferred in a similar manner from tube No. 2 back into tube No. 1, but this time the polarity is reversed so that positive elements in one tube correspond to negative ones in the other. The life test consists of letting this back-and-forth transfer proceed automatically at a reasonably high repetition rate and observing whether the initially set pattern remains unspoiled in the system.

To date, runs of 20 hr without any failures have been observed. The over-all characteristics of the pair of tubes in the life-test circuit did not change measurably in 700 hr. We are engaged at present in improving the testing circuits to be certain that they are not the cause of the occasional failures that still occur in long runs. We are also attempting to gain greater safety factors in the tubes themselves.

The research has reached the stage at which a Selectron of a capacity of 256 elements has been designed. It is practical and reliable in its operation and reasonably easy to build. While the life tests are still in progress and data from them are incomplete, there is every reason to believe that tubes with fairly long life can be made. The fast access time, the digitalized operation for address reading and information registering, the relatively intense output signals and self-monitoring by luminous display make the tube particularly useful for electronic computing machines and other information-handling machines.

REFERENCE

1. J. Rajchman, "The selectron—a tube for selective electrostatic storage," *Proceedings of a Symposium on Large-Scale Digital Calculating Machinery* (Harvard University Press, Cambridge, 1948), p. 133.

Traits Caractéristiques de la Calculatrice de la Machine à Calculer Universelle de L'Institut Blaise Pascal

L. COUFFIGNAL

*Institut Blaise Pascal**

I. CONSIDÉRATIONS GÉNÉRALES

La destination même du laboratoire de calcul mécanique de l'Institut Blaise Pascal est de poursuivre des recherches relatives à des matériels de calcul numérique, et spécialement à des machines arithmétiques, et aussi des recherches relatives au mode d'utilisation de ces matériels, c'est-à-dire aux méthodes de calcul.

C'est l'une des raisons pour lesquelles les caractères constructifs de la machine à calculer universelle de l'Institut Blaise Pascal n'ont pas été arrêtés a priori et de façon définitive. Même après sa mise en service, cette machine pourra subir des modifications, soit par remplacement de certains organes par des organes nouveaux, soit par adjonction d'autres organes; elle sera, par elle-même, une sorte de laboratoire.

Cette souplesse, cette aisance de transformation, est peut-être le plus caractéristique de ses traits; c'en est du moins un trait fondamental.

Son rôle d'instrument de recherche lui impose d'être véritablement universelle, c'est-à-dire de pouvoir être équipée de manière à exécuter toute sorte de calculs. Une telle exigence serait excessive pour la machine à calculer d'un laboratoire de recherche consacré à des travaux déterminés, et dont les calculs sont d'un nombre limité de types bien définis; il suffit dans ce cas d'une machine permettant d'effectuer ces calculs dans les meilleures conditions de rapidité, d'économie, et aussi dans les meilleures conditions de simplicité de manipulations; puisque les opérateurs d'une telle machine ne sont pas en général des spécialistes du calcul mécanique, et qu'une machine à calculer est pour eux l'un des nombreux appareils de leur laboratoire dont ils ont à apprendre la manipulation. Il y a aussi grand avantage à ce qu'une telle machine ne soit que d'un faible encombrement. Nous pensons que la machine-laboratoire de l'IBP servira à déterminer les caractéristiques de machines plus réduites, destinées à des laboratoires particuliers, adaptées le mieux possible aux besoins de ces laboratoires, et de manipulation simple. Cette considération nous a conduit à étudier avec un soin particulier la réalisation matérielle des éléments de la machine, en vue d'une fabrication de type industriel et de l'échange standard des unités sujettes à usure ou accident, notamment celles qui comportent des tubes à vide; c'est là, pensons-nous, un second trait caractéristique de nos recherches et des parties de la machine déjà construites; on verra dans quelques instants les résultats obtenus dans cette voie.

* Read at the Symposium by Léon Brillouin, *Harvard University*

Considérant que l'élément essentiel d'une machine à calculer universelle est le mécanisme calculateur, nous avons d'abord fait porter nos efforts sur cette partie de la machine.

L'expérience acquise dans l'utilisation de machines mécaniques nous y incitait déjà, et nous a guidé utilement. Nos recherches en ce domaine en sont au point où nous pensons avoir obtenu des résultats à peu près définitifs, du moins si l'on se borne à utiliser comme matériel élémentaire celui que peuvent actuellement fournir les fabricants de matériel de radio. C'est donc cette partie de la machine sur laquelle je me propose de donner quelques détails.

Nous l'appelons la *calculatrice*. J'espère que les renseignements relatifs à notre calculatrice donneront une idée nette de l'orientation de nos recherches.

Il est clair, enfin, que la plupart des travaux mathématiques qu'une machine à calculer peut être appelée à faire ont pour origine des recherches concernant la technique ou les sciences de la nature. Les travaux de mathématiques pures nécessitent rarement des calculs numériques importants; l'utilité de ces calculs ne semble pas aussi impérieuse. Cette remarque nous a conduit à étudier de façon approfondie le calcul mécanique de la racine carrée, opération qui intervient fréquemment dans les calculs techniques.

L'étude, poursuivie sur ces bases, nous a confirmé dans la préférence d'une *calculatrice parallèle* à l'exclusion d'une calculatrice à *séquence*; une analyse rapide de l'exécution des opérations fondamentales, chiffrage, addition, soustraction, multiplication, racine carrée, dans une calculatrice parallèle, donnera, avec l'explication logique de la structure de la calculatrice de la machine de l'IBP, la justification de notre choix.

II. LES OPÉRATIONS FONDAMENTALES

Chiffrage. Le *chiffrage*, opération consistant à représenter matériellement un nombre, exige, dans le système de numération binaire, un organe par ordre binaire capable de prendre deux états distincts, et un second organe capable de maintenir le premier dans l'état qu'on lui a fait prendre; nous appelons le premier organe un *inscripteur élémentaire*, le second un *verrou* et l'ensemble des deux, un *chiffreur élémentaire*. Les chiffreurs élementaires des divers ordres binaires constituent un *chiffreur binaire*; leur nombre est la *capacité* du chiffreur, un chiffreur de capacité k peut représenter tous les entiers de 0 à $2k - 1$.

Addition. L'addition nécessite, pour être automatique, un *reporteur*, dispositif effectuant le report des retenues de telle sorte qu'après inscription successive de deux nombres sur le chiffreur, ce dernier représente la somme des deux nombres. Nous appelons *totalisateur* l'ensemble d'un chiffreur et d'un reporteur.

Exemple (Fig. 1): $x = a + b$, $a =$ II0II, $b =$ I00I, $k = 6$.

Dans cet exemple, on suppose conformément à la plupart des réalisations mécaniques, électromécaniques, ou électroniques, que le reporteur est constitué par un chiffreur auxiliaire qui enregistre les reports à faire pendant l'inscription du second terme de la somme et le transmet ensuite au chiffreur. Le reporteur de la machine IBP qui va être décrit n'est pas de ce type.

Soustraction. La *soustraction* peut se ramener à l'addition par la méthode bien connue des

compléments. Le complément de b pour la capacité k est $2k - b$, et l'on sait que, si l'on inscrit sur un totalisateur de capacité k les nombres a et $2k - b$, le totalisateur marque $a - b$, le chiffre I dans l'ordre k ne pouvant pas être représenté par la machine.

La méthode que nous utilisons dérive de la méthode des compléments (Fig. 2). Le complément du *retrait* est remplacé par le *permuté*, qui s'obtient en permutant les chiffres 0 et I

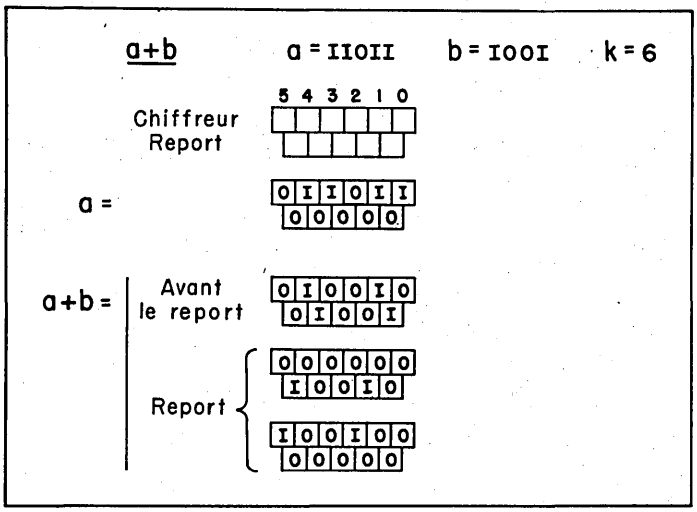


FIG. 1. Exemple d'addition.

dans la figuration de ce nombre, et le reporteur est complété par un élément enregistrant les reports provenant du chiffreur élémentaire de l'ordre le plus élevé pour les transmettre au chiffreur élémentaire de l'ordre le plus faible; ce report, appelé *report sans fin* est l'application au système binaire d'un procédé déjà en usage dans certaines machines décimales mécaniques.

L'avantage de cette méthode est que le calcul mécanique du permuté est beaucoup plus aisé que celui du complément; ce dernier s'obtient en permutant les chiffres 0 et I, sauf le dernier I à droite et les 0 qui le suivent; il exige donc une commande conditionnelle que n'exige pas le calcul du permuté.

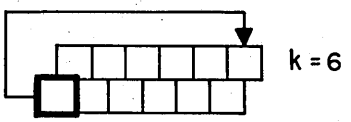


FIG. 2. Schéma d'une soustraction.

Nous ne donnerons pas la démonstration de l'équivalence des deux méthodes, qui est très facile, mais il est utile de noter deux particularités.

D'abord, l'opération $a - b$, où a et b sont positifs, n'est exacte que si $a > b$, car rien ne distingue, sur le totalisateur, la différence $0 - b$, et la somme $0 + b'$, en designant par b' le permuté de b qui se présente comme nombre arithmétique; il faut donc que le permuté du retrait, qui joue le rôle d'un *ajouté négatif*, soit accompagné du signe $-$; on voit aisément qu'il suffit pour cela d'ajouter un chiffreur élémentaire à la gauche du chiffreur et de lui attribuer un reporteur, en convenant que, dans ce *chiffreur de signe* le signe $+$ soit représenté comme le chiffre 0 et le signe $-$ comme le chiffre I. Un tel totalisateur peut être appelé totalisateur algébrique.

Exemple (Fig. 3): $x = a - b$, $a = 100100$, $b = 1001$, $k = 6$.

Notons au passage que le système binaire est le seul où soit possible l'assimilation des signes distinctifs des nombres positifs et des nombres négatifs à des chiffres de la numération; c'est là un avantage du système binaire qui a déjà été utilisé, mais ne paraît pas avoir été souligné de façon nette.

La seconde particularité de cette méthode tient à la nature de la réalisation mécanique de la soustraction. Si l'on applique la méthode précédente au calcul de $a - a$, on trouve une figuration formée de I dans tous les chiffreurs élémentaires, y compris le chiffreur de signe.

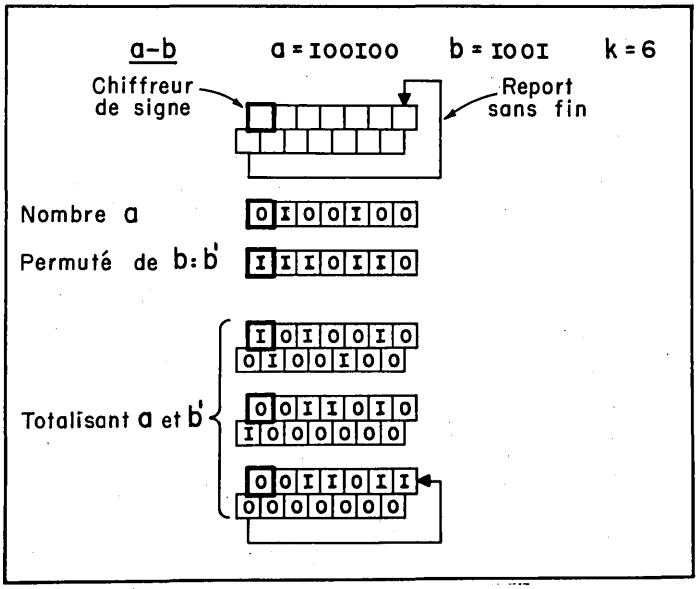


FIG. 3. Exemple de soustraction.

Exemple: $a = 1001$, $k = 6$.

inscription de a $= \overline{0}00100\overline{1}$
figuration de $-a$ $= \overline{1}10110$
somme $= \overline{1}111111$

Il faut considérer cette figuration comme représentant 0. Comme le chiffreur de signe porte le signe —, nous l'appellerons le *zéro négatif*, et par opposition, nous appellerons *zéro positif*, la figuration $\overline{0}000000$ ($k = 6$).

Multiplication. Pour la *multiplication*, dans notre premier modèle de calculatrice, nous nous sommes arrêtés à la méthode classique d'additions répétées. Le multiplicande m est inscrit dans un chiffreur M, le multiplicateur x dans un chiffreur X. Le multiplicande est transféré à un totalisateur P ou non selon que le premier chiffre de x est I ou 0; puis le multiplicande m subit un déplacement d'un pas vers la droite dans le chiffreur M, tandis que le multiplicateur x subit un déplacement d'un pas vers la gauche dans le chiffreur X; la même suite d'opérations se reproduit jusqu'à épuisement des chiffres de x.

Notons en particulier, d'une part que, si la capacité de M est k et celle de X, k', la capacité de P doit être $k + k'$; et d'autre part que la multiplication comporte une commande conditionnelle, dépendant de la nature 0 ou I du chiffre figuré dans un certain chiffreur élémentaire (le premier chiffreur élémentaire de X).

Division et racine carrée. Les méthodes opératoires précédentes sont déjà connues dans leur ensemble. Au contraire, la méthode de la division et celle de l'extraction d'une racine carrée, que nous allons exposer, nous paraissent nouvelles. L'exposé précédent éclaire dans une certaine mesure la théorie de la division et de l'extraction d'une racine carrée; en outre, il contribuera à mettre en relief la condensation des mécanismes que permettent les méthodes que nous décrivons.

On peut développer pour la division et l'extraction d'une racine carrée des théories analogues, qui même s'étendraient aisément à des racines d'ordre supérieur à 2.

Soit à diviser a par b. Désignons par $\bar{q}_n$ le nombre formé par les premiers chiffres du quotient jusqu'au chiffre d'ordre n et par q_{n-1} le chiffre suivant.

Par définition:

$$b\bar{q}_n 2^n \leq a < b(\bar{q}_n + 1)2^n,$$
$$b(2\bar{q}_n + q_{n-1})2^{n-1} \leq a < b(2\bar{q}_n + q_{n-1} + 1)2^{n-1}. \tag{1}$$

Posons:

$$r_{n,+} = a - b\bar{q}_n 2^n,$$
$$r_{n,-} = a - b(\bar{q}_n + 1)2^n. \tag{2}$$

Des relations (1) et (2) on tire:

$$(q_{n-1} - 1)b2^{n-1} \leq r_{n,+} - b2^{n-1} < q_{n-1}b2^{n-1},$$
$$(q_{n-1} - 1)b2^{n-1} \leq r_{n,-} + b2^{n-1} < q_{n-1}b2^{n-1}. \tag{3}$$

Pour chacune des inégalités doubles (3), si le terme médian est positif, q_{n-1} est égal à 1 d'après la seconde inégalité, et si le terme médian est négatif, q_{n-1} est égal à 0 d'après la première inégalité. Les réciproques se démontrent de même. En outre, le terme médian est égal à $r_{n-1,+}$ ou à $r_{n-1,-}$ selon qu'il est positif ou négatif. D'où:

Règle de division: Selon qu'un reste partiel est positif ou négatif, on inscrit le chiffre I *ou le chiffre* 0 *au quotient à la droite des chiffres précédents, et on retranche de ce reste, ou on lui ajoute, le diviseur déplacé d'un rang vers la droite pour obtenir le reste partiel suivant.*

Soit maintenant à extraire la racine carrée de a. Désignons encore par $\bar{q}_n$ le nombre formé par les chiffres de la racine jusqu'à l'ordre n et par q_{n-1} le chiffre suivant. Par définition:

$$\bar{q}_n{}^2 2^{2n} \leq a < (\bar{q}_n + 1)^2 2^{2n},$$
$$(2\bar{q}_n + q_{n-1})^2 2^{2(n-1)} \leq a < (2\bar{q}_n + q_{n-1} + 1)^2 2^{2(n-1)}. \tag{4}$$

Posons:

$$r_{n,+} = a - \bar{q}_n{}^2 2^{2n},$$
$$r_{n,-} = a - (\bar{q}_n + 1)^2 2^{2n}. \tag{5}$$

Des relations (4) et (5) on tire:

$$(q_{n-1}-1)[\bar{q}_n 2^{2n}+(q_{n-1}+1)2^{2(n-1)}] \leq r_{n,+} - (\bar{q}_n 2^{2n}+2^{2(n-1)}) < q_{n-1}[\bar{q}_n 2^{2n}+(q_{n-1}+2)2^{2(n-1)}],$$

$$(q_{n-1}-1)[\bar{q}_n 2^{2n}+(q_{n-1}+1)2^{(2n-1)}] \leq r_{n,-} + (\bar{q}_n 2^{2n}+3\cdot 2^{2(n-1)}) < q_{n-1}[\bar{q}_n 2^{2n}+(q_{n-1}+2)2^{2(n-1)}]. \quad (6)$$

Pour chacune des inégalités doubles (6), si le terme médian est positif, q_{n-1} est égal à 1, d'après la seconde inégalité; si le terme médian est négatif, q_{n-1} est égal à 0, d'après la première inégalité.

Les réciproques se démontrent de même. En outre, le terme médian est égal à $r_{(n-1),+}$ s'il est positif et à $r_{(n-1),-}$ s'il est négatif. D'où:

Règle d'extraction de racine carrée: Selon qu'un reste partiel est positif ou négatif, on inscrit le chiffre I *ou le chiffre* 0 *à la droite des chiffres de la racine déjà obtenus, et on retranche de ce reste, ou on lui ajoute, la racine ainsi obtenue déplacée d'un rang vers la droite et suivie des chiffres* 0I *ou* II, *pour obtenir le reste partiel suivant.*

Les règles d'opérations qui viennent d'être formulées ramènent la division et l'extraction d'une racine carrée à des suites de transferts et d'additions ou de soustractions, dont le nombre est sensiblement le même que pour une multiplication; cette remarque met en évidence l'énorme gain de temps (90% au moins) qu'elles procurent par rapport aux méthodes d'itération en usage jusqu'à présent, par exemple la formule $x_{i+1}=\frac{1}{2}[x_i+(a/x_i)]$ pour l'extraction de la racine carrée du nombre a.

Ces opérations de transfert, addition et soustraction, peuvent être effectuées au moyen des chiffreurs M et X et du totalisateur P qui ont servi à la multiplication pourvu qu'on leur adjoigne des moyens de réalisation de la permutation et du déplacement. Nous allons voir avec quelle simplicité de moyens matériels ces fonctions peuvent être réalisées.

III. L'ÉTAGE BINAIRE ET LA CALCULATRICE IBP

Le schéma (Fig. 4) représente un élément de totalisateur binaire, réalisant ces fonctions, que nous appelons couramment un *étage binaire.* Il se prête également à *l'effaçage*, opération évidemment nécessaire à tout dispositif de calcul.

Chiffrage. La triode F_c de la paire F marque I quand elle débite et 0 quand elle ne débite pas; la triode F_v montée en flip-flop avec elle en constitue le verrou.

Inscription. Elle s'effectue en attaquant simultanément les deux cathodes du flip-flop F en 15; cette attaque s'effectue à travers une triode de régularisation L_c; il faut comprendre que la borne de sortie 6 est reliée à la borne d'entrée 15 de l'élément suivant.

L'attaque est donc commandée de l'extérieur par la borne 7 d'entrée de la triode L_c.

Le totalisateur étant parallèle, tous les chiffres sont inscrits à la fois.

Un tube à néon branché sur la triode F_v est éclairé lorsque le chiffre marqué par F_c est I.

Report. Lorsque la triode de chiffrage F_c passe de I à 0, par l'inscription successive de deux chiffres I, elle est parcourue par une impulsion positive que l'on transmet à la borne d'entrée 15 de la triode de chiffrage de l'étage suivant à travers la triode de régularisation L_c, après l'avoir retardée, dans la ligne de retard *ES*, le temps nécessaire pour l'achèvement de l'inscription directe dans cet étage. Ce dispositif supprime l'inscription du report sur un chiffreur auxiliaire.

La durée d'une addition dans un totalisateur de capacité k est ainsi de $(k + 1)\theta$, θ désignant la durée de basculement d'un flip-flop.

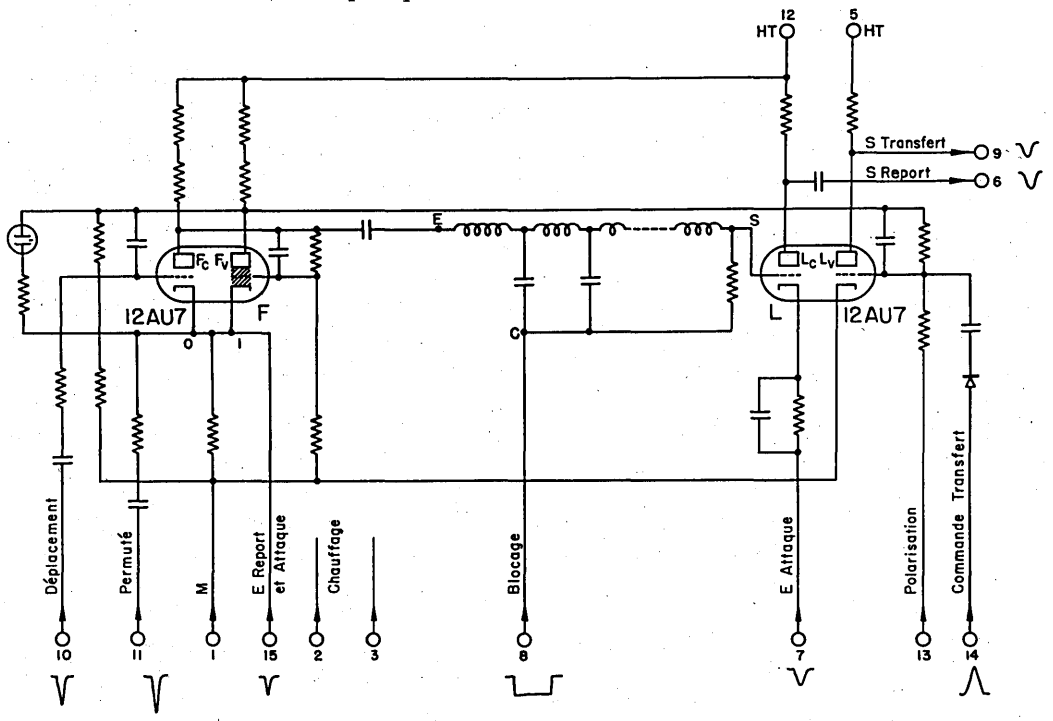


FIG. 4. Elément de totalisateur binaire.

Permutation. La permutation est obtenue par l'attaque de tous les étages du totalisateur par la borne 11, qui est reliée à un générateur d'impulsion unique, et par le blocage simultané des reporteurs par la borne 8, qui est reliée à un autre générateur d'impulsions.

Transfert. Le transfert s'effectue d'un étage à tous les étages du même ordre *binaire* des *totalisateurs* auxquels peut être transféré le nombre marqué par le totalisateur auquel appartient l'étage considéré; la borne de sortie 9 est reliée à cet effet à toutes les bornes d'entrée 7 des étages du même ordre binaire de ces totalisateurs, mais ceux qui ne doivent pas recevoir de nombre sont bloqués en 8, comme pour la permutation. Le transfert est réalisé par une impulsion positive envoyée en 14, sur tous les étages simultanément; cette impulsion n'est pas suffisante pour que la triode L_v atteigne le cut-off, mais la grille de cette triode peut recevoir une polarisation positive statique de la plaque de la triode F_v, qui est en tension haute lorsque

la triode de chiffrage F_c débite, et ainsi marque le chiffre I; l'impulsion 14 peut alors mettre en débit la triode L_v, qui envoie une impulsion négative d'inscription dans le circuit 9.

Déplacement. Une impulsion négative, envoyée en 10 dans tous les étages simultanément, ramène à 0 les triodes F_c qui marquent I et n'agit pas sur celles qui marquent 0; cette commande produit donc le même effet que la commande de l'addition du nombre à lui-même. Par l'action du reporteur, l'addition devient effective, or, l'addition du nombre à lui-même est identique à la multiplication de ce nombre par 2, c'est-à-dire à son déplacement d'un pas vers les positions hautes. Pour le déplacer vers les positions basses il suffit de monter le reporteur en sens contraire.

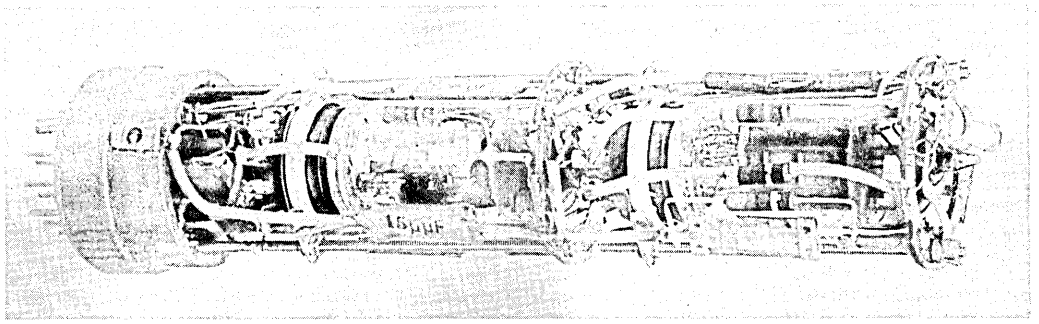


FIG. 5. Ligne de retard (organes intérieurs).

Effaçage. L'impulsion de déplacement en 10 ramenant tous les chiffreurs à 0, il suffit de bloquer en même temps les reporteurs par une impulsion en 8 pour obtenir l'effaçage.

On voit que toutes les opérations élémentaires ont une durée de moins de 2θ, sauf le report dont la durée peut atteindre $k\theta$.

Réalisation matérielle d'un étage binaire IBP. Le schéma montre qu'il nous suffit pour constituer un étage binaire de deux doubles triodes et d'une ligne de retard. Matériellement la ligne de retard est constituée par quelques bobines plates enroulées sur un tube de carton fort, et les autres pièces sont montées en un ensemble compact porté par un socle à 14 broches (Fig. 5). Cet ensemble est coiffé par le tube support de la ligne de retard qui lui sert de carter (Fig. 6). L'étage binaire ainsi constitué a 5 pouces de haut et $1\frac{1}{2}$ pouces de diamètre. Dans un souci de standardisation, on a pris pour L_c et L_v les triodes d'une double triode identique à celle qui sert au chiffrage, bien que L_c et L_v aient des fonctions indépendantes et ne soient pas montées en flip-flop. En outre, pour faciliter le remplacement des tubes usés, le montage s'ouvre transversalement vers le milieu de sa hauteur.

Réalisation matérielle d'une calculatrice. Puisqu'un totalisateur porte en lui-même des moyens

de permutation et de déplacement il suffit de remplacer par des totalisateurs les chiffreurs M et X considérés dans la théorie des opérations algébriques, pour constituer les organes calculateurs d'une calculatrice parallèle.

Les étages binaires qui constituent ces totalisateurs sont engagés dans des douilles placées côte à côte sur une plaque de fondation commune. Les connexions entre étages sont réalisées de façon fixe sous cette plaque, qui constitue elle-même le couvercle d'une boîte dans laquelle souffle un vent suffisant pour refroidir les tubes à vide, en circulant à l'intérieur de chacun des tubes carter de chacun des étages binaires.

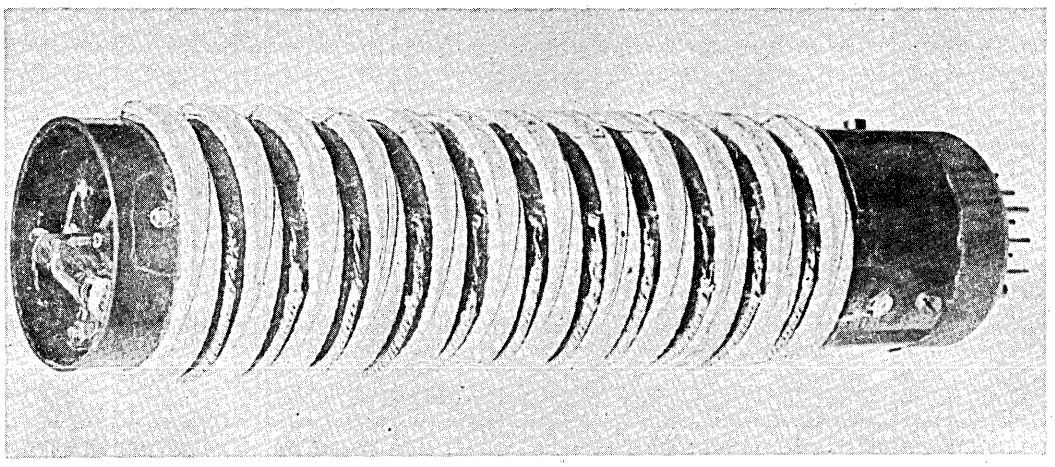


FIG. 6. Ligne de retard (vue d'ensemble).

La hauteur totale de cet ensemble est de 8 pouces environ; sa surface, celle de $6k$ carrés de $1\frac{1}{2}$ pouce de côté, k désignant la capacité du multiplicande et du multiplicateur; par exemple, pour la machine IBP, qui travaille sur 15 chiffres décimaux, $k = 50$ et la surface des totalisateurs de la calculatrice est de moins de 700 pouces carrés.

Les dispositifs de commande et les générateurs d'impulsions demandent une cinquantaine de tubes, quelle que soit la capacité des totalisateurs. Ces tubes sont du même type que ceux des totalisateurs à l'exception de quelques pentodes et thyratrons.

Nous croyons pouvoir insister sur la réduction d'encombrement, le caractère industriel et le haut degré de standardisation atteint dans la réalisation de cette partie de notre machine.

La suite des opérations élémentaires est indiquée par les schémas suivants pour la division, elle est analogue pour l'extraction de racine carrée.

Exemple (Fig. 7): $a : b$, $a = \text{I00000I}$, $b = \text{I0I}$, $k = 6$, $k' = 5$.

Nous noterons, d'une part, que le zéro négatif doit commander le transfert de $\overline{\text{I}}$ à X tout

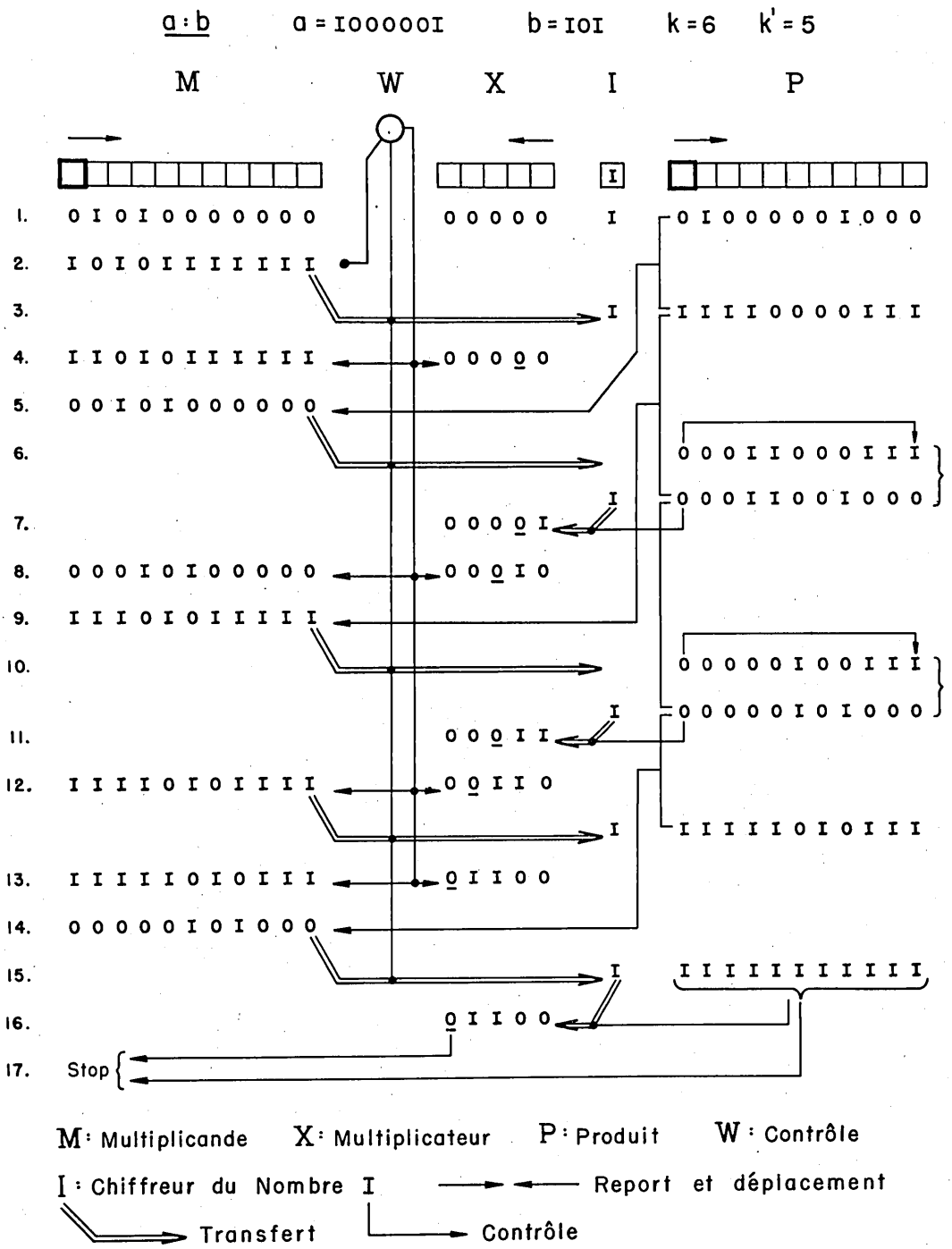


FIG. 7. Exemple de division.

comme le zéro du chiffreur de signe de P; et d'autre part, que la permutation de M dépend de la comparaison des signes successifs de P; on voit apparaître deux nouvelles commandes conditionnelles spéciales, tenant à la fois à la structure de la calculatrice et à la méthode de calcul utilisée.

IV. UN PRINCIPE DE RECHERCHE. LA QUESTION DE LA MEMOIRE

Nous voudrions, à cette occasion, rappeler un principe que nous formulions dès 1933, et dont les confirmations se sont multipliées. L'observation de l'évolution des machines existant à cette date nous conduisait à avancer que le progrès, en calcul mécanique, résultait d'une adaptation mutuelle des machines à calculer et des méthodes de calcul. Un exemple particulièrement typique d'adaptation des méthodes aux machines est, dans les analyseurs différentiels, la détermination des fonctions élémentaires, sin x, L x, etc., par des analyseurs différentiels auxiliaires, c'est-à-dire, mathématiquement, la substitution à une fonction d'une équation différentielle dont elle est solution. L'exposé qui précède offre de nombreux exemples de détail, de réaction mutuelle des recherches mathématiques et des recherches techniques; en particulier, la simplicité des méthodes de division et d'extraction de racine carrée est fort accrue par la simplicité de la technique du déplacement et de la permutation.

Les confirmations renouvelées de ce principe nous conduisent à considérer comme inefficace, dans l'état actuel de la technique, une discussion logique a priori de la réalisation matérielle d'une machine à calculer universelle.

Par exemple, le débit très élevé d'une calculatrice telle que celle dont nous venons de donner une description schématique, met en question la méthode de calcul des fonctions élémentaires, et nous conduira vraisemblablement à abandonner les tables mécaniques, que nous conseillions, en 1938, pour une machine électromécanique, sous une forme voisine de celle que l'on peut admirer dans la machine Mark I du professeur Aiken.

On comprendra aussi, pensons-nous, pourquoi nous avons déclaré, en plusieurs circonstances, que nous ne savons pas encore quelle sera la nature de la mémoire de notre machine.

Fonctionellement, nous considérons comme nécessaires une mémoire interne et une mémoire externe, et comme avantageuse la séparation de la mémoire des nombres et de la mémoire des commandes.

La structure de ces diverses mémoires doit dépendre, à notre avis, des calculs à faire et de la méthode adoptée. Par exemple, la mémoire n'intervient pas dans les mêmes conditions si l'on calcule des trajectoires ou si l'on résout un système de 50 équations linéaires à 50 inconnues; dans le second cas, les phases sont de une ou deux opérations, dans le premier cas, elles peuvent atteindre la centaine d'opérations.

C'est-à-dire que notre machine-laboratoire comportera plusieurs types de mémoire dont le mode d'emploi aura à être étudié systématiquement, en liaison avec les problèmes traités.

Nous donnerons pour terminer le schéma d'une mémoire que les essais poursuivis jusqu'à présent nous conduisent à considérer comme avantageuse dans la fonction de mémoire interne d'une calculatrice parallèle (Fig. 8).

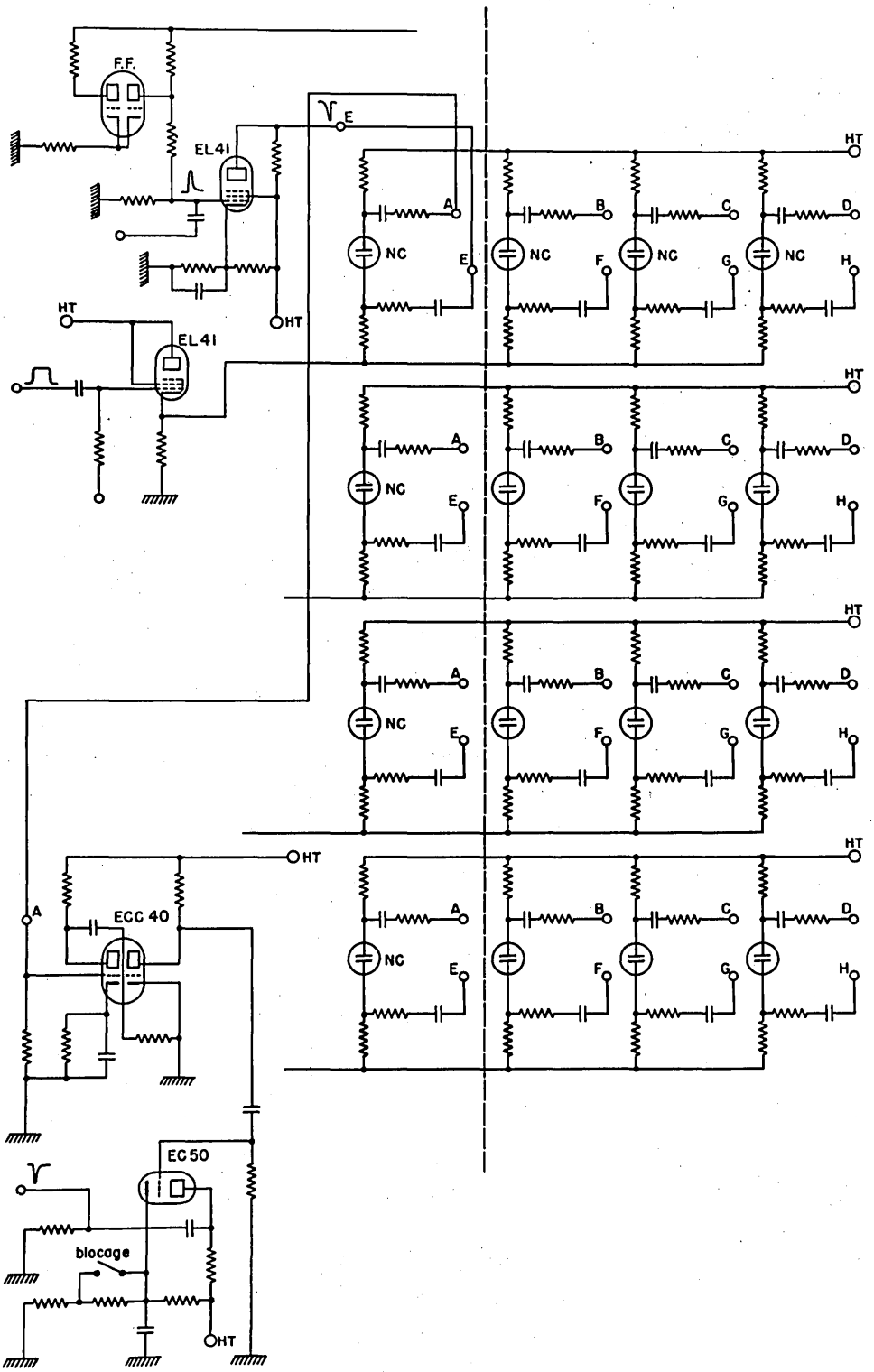


FIG. 8. Schéma d'une mémoire.

Chaque chiffre est enregistré au moyen d'une diode à gaz *NC*. Les diodes constituant les chiffreurs élémentaires d'une même chiffreur binaire sont figurés sur une même ligne horizontale; les diodes du même ordre binaire dans les divers chiffreurs sont figurés sur une même ligne verticale; le schéma montre donc une mémoire de 4 nombres de 4 chiffres; il faut comprendre en outre que les plots représentés par la même lettre sont réunis entre eux.

Le fonctionnement du dispositif se fonde sur la remarque que le seuil de tension d'allumage d'une diode est nettement plus élevé que son seuil de tension d'extinction.

L'inscription s'effectue en envoyant une impulsion positive par les bornes *A*, *B*, *C*, . . . dans tous les ordres binaires où doit être représenté le chiffre I, et en bloquant les tubes des chiffreurs où l'inscription ne doit pas être faite par une impulsion opposée.

La lecture s'effectue en envoyant, par les bornes *E*, *F*, *G*. . . . une impulsion négative trop faible et trop brève pour provoquer l'extinction des diodes. L'effaçage s'obtient en prolongeant l'impulsion de lecture.

L'impulsion de lecture peut avoir pour durée θ, durée de basculement d'un flip-flop des totalisateurs; c'est, croyons-nous, la plus faible durée atteinte pour l'extraction d'un nombre d'une mémoire et son transfert à un chiffreur. C'est cette caractéristique de fonctionnement de la mémoire à diodes qui en fait l'intérêt; cette mémoire ne retarde en rien la calculatrice, car l'inscription dans la mémoire et l'effaçage peuvent se poursuivre pendant que la calculatrice travaille isolément.

Malgré le nombre des diodes, qui peut paraître élevé, ce dispositif reste simple et sûr, parce que les diodes à gaz sont des tubes robustes, et que l'on peut les utiliser dans des conditions où leur fonctionnement ne produit guère d'usure. En outre, ces tubes sont peu coûteux.

The Future of Computing Machinery

LOUIS N. RIDENOUR

University of Illinois

The title of these remarks is somewhat misleading, in that one of the things Professor Aiken has requested of me is to give a very brief critical summary of the proceedings of the present Symposium; following this, I venture a few speculations regarding the principal directions in which the research and development on computing machinery seem to be tending.

The central interests and concerns of the more than 700 people in attendance at the present symposium are extremely diverse; the fields in which papers have been presented are various and wide. There have been papers on computing machinery, on methods of numerical analysis, on the solution of problems involving numerical analysis in the fields of physics, engineering, economics, and social science. No doubt, the fact that interest in this Symposium has been so splendidly sustained in spite of this diversity of subject matter can be explained by observing that, once a problem has been reduced to a mathematical form, then what proceeds from that point onward is of common interest to those concerned with numerical analysis, almost without regard to the way in which the original equations to be solved arose.

Thus a prominent effect of the development of computing machines is likely to be that of producing important unifications and sharings of viewpoint among various scientific disciplines which present problems amenable to attack by numerical analysis. The reports presented at this Symposium encourage the belief that the art of computing machines may be entering a new phase—a phase of increased maturity. We are assembled here to celebrate the completion of the Harvard Mark III machine, and many of the papers presented here in the sessions on computing machines have described completed and operating machines, rather than the plans for constructing machines not yet built. It is clear that powerful methods of numerical analysis are being developed, and that the new numerical problems posed by the extreme speeds of modern machines are becoming evident and are beginning to be attacked. Many of the numerical problems that have been described here—in physics, engineering, and social science—are not merely proposed for solution, but actually have been attacked and solved in whole or in part. The keynote of the present meeting thus seems to be achievement, even if limited achievement, rather than promise.

Let us now consider some of the papers presented here, in the order in which they appear in the program. It will not be possible to mention each of the some forty papers presented, but an effort will be made to deal with typical ones in each category.

No comment on the Harvard Mark III machine is offered beyond saying that we have all had an opportunity to inspect this machine and to learn something of its design and its properties. I should like to remark upon the very considerable debt that the entire high-speed

computer art owes to the early, continued, and effective work of Professor Aiken and his group.

The Bell Telephone Laboratories computer that was described seemed remarkable principally for its complete avoidance of the use of conventional vacuum tubes. There were used as computing elements mainly electromechanical relays, together with fewer than one hundred vacuum tubes. Possibly because of this unconventional design, this machine and its relatives in the series of Bell machines have achieved a very remarkable record of continued reliability.

Very interesting progress reports on machines under construction were offered by the Massachusetts Institute of Technology, the Raytheon Manufacturing Company, the General Electric Company, the National Bureau of Standards, Mr. Elliott for the British, and the Institut Blaise Pascal. Many of the machines described are scheduled for completion in the year 1950; that year should be a very interesting one for those concerned with computing machines.

One aspect of the British developments seems worthy of special remark. This is the quite evident difference in the approach to the problem of constructing a large computing machine adopted respectively by British workers and by American workers. Before launching upon the construction of large machines, the British prefer to make preliminary experiments, and to gain experience, with small machines of admittedly limited scope which, however, possess sufficient generality to be educational. American practice has been, on the other hand, to embark from the beginning on the construction of quite ambitious machines, usually without preliminary experience on small-scale models. To some degree, this may express the greater availability of research funds from the American government, but I think that it goes deeper than that; I think that it expresses a difference in the national character.

During the recent war, I was frequently distressed by what seemed evidences of stupidity and ineptitude in our Air Force operations as they concerned my area of interest—airborne radar. On one occasion I was complaining to a general officer about this, and pointing out to him how much better the Royal Air Force managed its affairs. He said: "Well, you have to expect that. There are two ways to fight a war: you can fight a smart war, or you can fight an overwhelming war, but you can't do both. The British are fighting a smart war, but we aren't. We made our choice a long time ago; we decided to fight an overwhelming war, and that's what we're doing. Don't expect us to be smart." It seems that this approach has been carried over to the computer field; we Americans have a tendency to overwhelm our difficulties.

Several papers were presented on the subject of components for computing machines. It was clear from these that the outstanding component problem still is—as it has been for some time—that of an adequate high-speed storage device, or inner memory, for a computing machine. While special methods for reducing demands on an inner memory can usually be devised for any particular problem, nevertheless the scope of a machine increases and its operation becomes simpler as the capacity of the inner memory rises. Quite a lot of work is being done on this problem. The work of F. C. Williams, and that of the Eckert-Mauchly

group, which appears to be derived from it, seems to be very promising; so is the success that has recently been obtained in the use of mercury delay lines as high-speed storage elements.

Further, two papers given here reported on novel and interesting devices whose further development seems very promising. These are the magnetic delay lines and memory elements, on the one hand, and the highly suggestive work on electrochemical storage elements and relays, on the other.

Mr. Engstrom called to our attention the importance of special-purpose machines. Naturally enough, attention has mainly been focused on what are called "general-purpose" machines; but it is desirable to remember that for many purposes, notably those of industry and government, special-purpose machines are quite adequate and can often be realized for fewer dollars per function performed than could a general-purpose machine. An interesting example of a special-purpose machine is the Northrop assemblage of IBM equipment to make a simple, rapidly assembled, useful, and quite reliable machine.

In the session on numerical methods, Mr. Brown proposed a scheme for solving certain types of problems by playing a game. He has consulted with workers here at the Computation Laboratory of Harvard, and finds that their conservative judgment is that a 40×40 matrix can be dealt with completely in a thousand steps, and with an error of one part in a thousand, in a total time somewhat less than one hour. The complete program has not been prepared, and this is only an estimate, but it seems a promising one.

All those concerned with machine design should be grateful for Mr. Lehmer's elegant scheme for the generation of pseudorandom numbers by machines. Such numbers, and their production by a simple scheme, will take on increasing importance as lengthy analysis is replaced by statistical experiments conducted on machines, in the fashion of the Monte Carlo method described to us by Mr. Ulam.

Other papers in the session on numerical methods dealt with important problems in numerical analysis. Mr. Milne catalogued the outstanding needs in iterative schemes for the solution of the Laplace equation and other elliptical partial differential equations. Further work is needed, first, on the development of ways of programming for machine use such rapid systems of error removal in iterative solutions as the relaxation methods of Southwell; second, on ways for dealing with curved boundaries; third, on schemes for selecting good initial values of the functions being dealt with; and fourth, on better methods for handling mixed boundary conditions.

In the session on applications to physics, Mr. Furry made the general observation that the high-speed computing machine permits experimentation in theoretical physics with less labor and better results than have ever been accessible before. Such "theoretical experimentation" (if this is a good term) includes the testing of theories, the decision among competing hypotheses, the determination of ranges of validity of various approximations, and so on.

Examples of the actual use of machines in the solution of problems were offered. Problems presented included the birefringence produced by viscous flow, the trajectories of cosmic-ray particles, and the interaction of atomic electrons with electromagnetic radiation. In connection

with the last, Mr. Rose remarked that he has assembled the first compendium of wave functions for Dirac electrons in the screened nuclear field. It is indubitable that this catalogue of wave functions will be important and useful in many other investigations. It is to be hoped that Mr. Rose or others will carry on in the direction he has marked out, computing such wave functions for higher values of angular momentum, and for electron states in the negative-energy continuum.

In the sessions on aeronautics and applied mechanics, the papers presented made it clear that the use of computers in these fields will be very extensive. I mention particularly Mr. Welmer's prediction that the complete solution of the now separate problems of flutter, aerodynamic stability, and servomechanism performance of an airframe may soon be found in terms of the complete frequency-response spectrum for a particular airplane. Mr. Emmons and Mr. Muskat outlined two other practical applications in which computers will be extremely useful.

Mr. Mosteller set forth, in the session on economics and social science, the types of problems likely to be dealt with. He asserted that these are mainly solutions of simultaneous linear equations, both homogeneous and inhomogeneous, giving as specific examples problems in multiple regression, the finding of discriminant functions, scaling theory, and factor analysis. He further remarked that the present lack of adequate mathematical theory outside the field of economics now gives machines little to do in social science, while at the same time it points up very clearly the major present job of those interested in a quantitative social science. Examples of specific problems suitable for attack by machines were given by Mr. Tucker and Mr. Chernoff, and Mr. Waugh pointed out that many important economic problems do *not* require the use of machines. He urged that those present at the Symposium would assist the economist in preventing the establishment of a fad for using high-speed computing machines for all purposes, whether justified or not. This having been said, Mr. Waugh remarked that machines have a very important place in the solution of very important problems, notably those of Government in these days of increased central control. He reminded us that formulation of such problems is difficult and that economists often do not know what questions to ask, what answers to seek, or how to secure the public acceptance of policies necessary to implement the answers found.

On the ground, no doubt, that physiology can be regarded as an elementary sort of social science, Mr. Crozier found himself on the social-science program. He pointed out first of all that the multivariant character of organic processes almost certainly means that, when a proper mathematical description of such processes is formulated, it will be so complicated that machine computation will be demanded. He then addressed himself to the question of the validity of using the physiology of high-speed computing machines as an analogy for the physiology of the nervous systems of living organisms. He reminded us of the dangers of misleading analogies and concluded, from the example concerning vision which he quoted and from other evidence, that elementary neural decisions in a living organism are reached statistically. Thus, according to Crozier, a true thinking machine would have to have a very

large redundancy in individual elements whose individual performances fluctuate, in order to imitate in any meaningful way the performance of the neural system of a living organism.

Let us turn now to the questions more directly suggested by the title: The Future of Computing Machinery. The first such question, in this time of vigorous development, design, and construction, is perhaps: "Who is likely to possess large high-speed computing machines in the future?" Some workers in the computing-machine field, and some people interested in the field, are quite pessimistic about the ultimate wide availability of large high-speed computing machines, on the grounds that such machines are complicated and expensive to build, expensive to maintain and operate, and therefore cannot ever be afforded by institutions such as the normal middle-sized university. This is a point of view with which I disagree completely. I strongly believe that a competent high-speed computing machine will very soon be recognized as an important and inevitable part of the research equipment of any university having even modest research pretentions.

Thus I regard the computing machine as being not in the category of the large astronomical telescope, which is a pleasant but optional luxury for a university, but rather in the category of the electronuclear particle accelerator, which is a necessity for any university that desires to cultivate modern nuclear physics. In the early and middle nineteen-thirties, when Lawrence was having his first successes with the cyclotron, I remember many discussions of whether this or that institution should build a cyclotron. There were always those who argued that the cyclotron was expensive to build and run, that it had a limited field of usefulness, that there were already plans to build all of them that the country needed or could support, and therefore that the institution concerned in the discussion need not and should not build a cyclotron. Now it is not quite true that the only universities that have made substantial contributions to nuclear physics are the ones who ignored such skeptical notions and built cyclotrons, but it is nearly enough true to be significant. And the successful institutions that do not have cyclotrons do have, in all cases, some competing form of particle accelerator.

By analogy, I suggest that high-speed computing machines will be part of the routine and necessary research equipment of universities, industrial laboratories, government research establishments, and indeed any institution where any substantial volume of scientific research, in any field, is carried on. Possibly this trend will be readily discernible in a year or two, and surely its full implementation is less than a decade off.

Of course, the wide availability of high-speed computing machines will be greatly forwarded by improvements in reliability and reductions in cost, in consequence of continued developments of improved components and better logical design.

Let us now ask: "How large, how fast, and how complicated should a large, high-speed, general-purpose computing machine be?" The ENIAC still holds the record for the total number of vacuum tubes. More recent designs are considerably more ambitious in terms of the speed of individual operations, the size of the inner memory, and the general competence of the device; yet in spite of this they have fewer tubes, which they use harder, so to speak.

I should like to propose that the answer to the question of where to draw the line in designing a general-purpose machine is set entirely by considerations of reliability. That is, a large general-purpose machine ought to be as big and fancy and competent as it can be made, subject to the limitation that it must not commit errors oftener than once in, say, four hours. There is no other significant limitation on the total complexity of the device; for the machines that we have now, even those that have not yet been realized, but are in design, are still inadequate to deal with many problems we should like to put to them.

Professor Aiken has quoted to me a remark of Hartree's. Hartree said that the fastest computing machine that has yet been designed is still some 10^{10} times too slow to solve completely the problem of the wave equation for the copper atom. Mr. O'Neal, in his remarks before this Symposium, said that the solution of the traffic-handling problem for aircraft on the airlines of this country would tax the capacity of the biggest and fastest computing machines now in existence.

Warren McCulloch, a professor of psychiatry at my university, has interested himself in the sort of analogy between computing machines and neurophysiology that Mr. Crozier regards as being so dangerous. He has remarked that the over-all complexity of the largest and most complicated computing machine now in existence or proposed is just about equivalent to the complexity of the nervous system of the flatworm. You may or may not regard this as being a fair comparison. It is based upon drawing a parallel between a single flip-flop in the machine and a single neuron in the nervous system of the flatworm, and I think that it is safe and suitable for our present purpose.

There is little question that, so far as the carrying out of numerical computations is concerned, the computing machine is more useful than the flatworm. There are two obvious major reasons for this. First, the machine is specialized in its function, while the nervous system of the worm is not. The machine can deal only with special classes of situations, while the delimitation of the flatworm's competence is far less narrow. Second, the machine works about a thousand times faster than any organic nervous system.

Without claiming in any way that a computing machine "thinks" in the sense of origination, we must admit that it relieves human computers of a tremendous burden of routine mental effort which is ordinarily classified as thinking. This thinking is special, in the sense that it is governed by formal logical rules of manipulation, but in the past it has had to be managed by human nervous systems. With the help of machines, it can be *directed* by human nervous systems, but carried out without human intervention or assistance.

Thus, we are not talking about machines possessed of the ability to "think" in the sense to which Mr. Crozier was objecting, but rather machines which can perform logical processes in a rapid, uniform, and unerring way. The faster and more competent we can make such a machine, the bigger will be the burden of routine thought that it can take away from men. If we can make a machine large enough and competent enough, and if in the meanwhile we have learned more than we know now about the logical organization of the nervous systems of living organisms, we may at last be able to make a machine capable of origination and

problem-solving behavior. But this lies in the rather distant future; our present problem is to make large and reliable the machines of unitary function which are designed simply for the straightforward application of the logical rules built into their design.

We want, therefore, to make computing machines as large and as complicated as we can; for the fanciest machine that we can realize today is powerless in the face of problems that we can readily pose, but not yet solve. The limitation on size and complexity is set by reliability; for a machine will be useless to us if it is not sufficiently reliable to be depended upon for hours at a time.

This leads us to my final question: "How can a computing machine be made more reliable, so that its complexity can be increased without increasing the chance of failure?" Of course, there is no simple or evident answer to this question, or such an answer already would have been exploited in machine design. There are some promising indications on the horizon. I suggest that the first thing that should be done is to look toward as complete as possible an elimination from computing machines of vacuum tubes and electromechanical relays. These two components are presently the major sources of failure in existing machines, partly because they are so numerous, and partly because they wear out with continued use. What we need is computing elements that can perform the same nonlinear functions as those we now achieve with tubes or relays, but elements that are far less prone to depreciation in use.

Another drawback of the vacuum tube, of course, is the ridiculously large amount of energy that must be expended to boil off free electrons from its cathode. At the time McCulloch made his remark about the flatworm, he also observed that if a computer built on present principles should be made to have the same number of individual elements—let us call them "neurons" —as there are in the human central nervous system, then all the power of Niagara Falls would be required to light the tubes, and the complete water flow over the Falls would be required to keep the device cool. The human nervous system, though slow in electronic terms, is incomparably efficient in terms of energy expenditure per individual computing element.

What is needed is to replace the present basic nonlinear elements used for computers with another type of element that does not require enormous quantities of stand-by power, and is not depreciated by continued operation—an element that, once installed, can be relied on indefinitely unless it is abused. There are some hints as to the possible nature of such a device; some of these have been reported upon at this Symposium. The most promising ones visible today are, first, semiconductor devices of the sort of the recently announced transistor; second, magnetic devices like those reported here; and third, the electrochemical devices that may be developed from the pioneer work of which Mr. Bowman has told us. A great deal of intensive work on promising unconventional elements for computer use will be repaid if the over-all reliability of computers can thereby be increased. Reliability, as we have seen, brings in its train larger, more complex, and more competent computing machines. Presumably it also brings in its train a greater availability and a lower cost for the computers of present size and scope; inevitably, it will bring wider general use and acceptance of computing machines of all sorts.