I

W. A.

UNCLASSIFIED

PROCEEDINGS OF
A SYMPOSIUM ON LARGE SCALE
DIGITAL COMPUTING MACHINERY

CODE IV

SPONSORED BY
THE ARGONNE NATIONAL LABORATORY

August 3-5, 1953

SECRET

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

II

EDITOR'S NOTE

This volume of the Proceedings is a collection of the transcripts of papers presented at the Computer Symposium sponsored by Argonne National Laboratory on August 3, 4, and 5, 1953. These papers deal mainly with the engineering aspects of the design of digital computers. The problems encountered in the employment of cathode ray tubes as storage devices and the advances in circuit techniques are discussed in six papers on Williams Tube Memory. Theory of operation and the practical application of large scale ferromagnetic memory are presented in the four papers on this subject. A survey of the work done at Bell Laboratories in connection with the application of transistors to high-speed computers is covered in the papers presented by the Staff of the Bell Laboratories. Other topics of current interest, such as magnetic recording, ferroelectrics, programming experience, experimental mathematics, as well as weather computations, are discussed.

The success of this symposium was due in a large measure to the splendid work of the Program Committee and the Session Chairmen and the stimulating talks delivered by the various speakers.

On behalf of the Argonne National Laboratory, I wish to express my appreciation to the members of the symposium for their attendance, and personally I must express my gratitude for the assistance I obtained from Lester Furney, R. F. Kramer, Margaret Isom, and many other staff members of both ANL and Oak Ridge National Laboratory (on loan to ANL).

During the symposium many speculations arose as to why this meeting was held on August 3, 4, and 5. For the completeness of this Proceeding a few of the interesting speculations are recorded as follows:

1. It is intended to be the first and last Computer Symposium ever held in August, especially in the Chicago area.
2. It is the tenth year anniversary of the completion of the first all-electronic computer, the ENIAC.
3. It is the anniversary date of the start of the voyage of Columbus which resulted in the discovery of America.

However, whatever the reason may be, I must pay tribute to all of our guests that they lived through it and seemingly enjoyed it.

J. C. Chu
Physics Division

Lemont, Illinois
November 1953

III

CONTENTS

	Page
Officers of the Symposium	vii
Members of the Symposium	viii
Program of the Symposium	xxi
A Transistor Pulse Amplifier Using External Regeneration	W. A. Cornell 1
Tradic, a Transistor Digital Computer	J. R. Harris 21
Institute for Advanced Study Williams Memory	J. Pomerene 37
The ORACLE Memory System	R. J. Klein, Jr. 47
Relative Merits of Williams Memory Display	D. Jacobsohn 59
The Illiac Memory	J. M. Wier 72
Design and Manufacturing Considerations of the Oscillograph Type Electrostatic Storage Tube	E. M. Smith 83
A Myriabit Magnetic-Core Matrix Memory	J. Rajchman 84
Ferromagnetic Cores with Microsecond Access	I. L. Auerbach 118
Coincident-Current Magnetic Computer Memory Developments at M.I.T.	J. Forrester 150
Computer Components Research at Mellon Institute	F. A. Schwertz 159
Characteristics of the ORACLE	E. W. Burdette 194
Physical Aspects of Magnetic Computer Materials	S. E. Harrison 202
Magnetic Reading-Recording Head Design for UNIVAC	V. J. Porter 213

CONTENTS

		Page
40 Channel Magnetic Tape Auxiliary Memory	L. Merrill	219
Tape Sprocketing Device	A. Ostroff	236
A Preliminary Report on the Prototype Diode-Capacitor Memory	A. W. Holt	247
Some Topics in Experimental Mathematics	N. Metropolis	265
Survey of Ferroelectrics	B. T. Mathias	271
Some Experience in Coding and Programing with the Institute Computer	H. Goldstine	273
"Computing the Weather"	J. W. Mauchly	279

OFFICERS OF THE SYMPOSIUM

Chairman of the Symposium

J. C. Chu - Argonne National Laboratory, Lemont, Ill.

Program Committee:

J. P. Eckert

Remington-Rand Corporation, Philadelphia, Pa.

J. H. Felker

Bell Telephone Laboratories, Whippany, N. J.

D. A. Flanders

Argonne National Laboratory, Lemont, Ill.

Herman Goldstine

Institute for Advanced Study, Princeton, N. J.

A. S. Householder

Oak Ridge National Laboratory, Oak Ridge, Tenn.

N. Metropolis

Los Alamos Scientific Laboratory, Los Alamos, N. M.

J. A. Rajchman

RCA Research Laboratories, Princeton, N. J.

V

MEMBERS OF THE SYMPOSIUM

Irving Abrams	New York University New York 3, New York
P. A. Adamson	Hughes Aircraft Co. Culver City, California
Francis J. Alterman	New York University New York 3, New York
Gene M. Amdahl	IBM Corporation Poughkeepsie, New York
Frederick J. Anderson	Sylvania Electric Boston, Massachusetts
Isaac L. Auerbach	Burroughs Corporation Philadelphia 23, Pennsylvania
W. Scott Bartky	Illinois Institute of Technology Chicago, Illinois
Carl Bergstrom	Argonne National Laboratory Lemont, Illinois
J. C. Boyce	Argonne National Laboratory Lemont, Illinois
A. E. Brandt	NYOO, U. S. AEC New York 23, New York
Thomas Brill	Argonne National Laboratory Lemont, Illinois
E. Bromberg	Institute of Mathematical Sciences New York, New York
George W. Brown	International Telemeter Corp. Los Angeles, California
Donald Burbeck	Hughes Aircraft Company Culver City, California
Earl Burdette	Oak Ridge National Laboratory (on loan to ANL)

VII

W. H. Burkhardt	Monroe Calculating Machine Co. Morris Plains, New Jersey
James R. Burnett	Purdue University Lafayette, Indiana
J. W. Butler	Argonne National Laboratory Lemont, Illinois
Margaret Butler	Argonne National Laboratory Lemont, Illinois
A. E. Brennemann	IBM Corporation Poughkeepsie, New York
Richard M. Bloch	Raytheon Manufacturing Co. Waltham, Massachusetts
E. H. Cabaniss	General Electric Schenectady, New York
Jeffrey C. Chu	Argonne National Laboratory Lemont, Illinois
George I. Cohn	Illinois Institute of Technology Chicago, Illinois
Warren A. Cornell	Bell Telephone Laboratories Murray Hill, New Jersey
H. D. Crane	Institute for Advanced Study Princeton, New Jersey
Charles H. Davidson	University of Wisconsin Madison 6, Wisconsin
Ernest J. Dietrich	Raytheon Manufacturing Co. Waltham, Massachusetts
Joseph J. Eachus	National Security Agency Washington 25, D. C.
Douglas E. Eastwood	Armour Research Foundation Chicago, Illinois
J. Presper Eckert, Jr.	Remington-Rand, Inc. Eckert-Mauchly Div. Philadelphia, Pennsylvania

VIII

Howard O. Edfeldt	Argonne National Laboratory Lemont, Illinois
Charles Eggler	Argonne National Laboratory Lemont, Illinois
Robert D. Elbourn	National Bureau of Standards Washington, D. C.
Cornelius Eldert	University of Pennsylvania Philadelphia 4, Pennsylvania
Roger H. Farrell	Museum of Science and Industry Chicago, Illinois
Jean H. Felker	Bell Telephone Laboratories Whippany, New Jersey
P. R. Fields	Argonne National Laboratory Lemont, Illinois
Donald A. Flanders	Argonne National Laboratory Lemont, Illinois
Gail T. Flesher	Illinois Institute of Technology Chicago, Illinois
Jay W. Forrester	Massachusetts Institute of Technology Cambridge, Massachusetts
David C. Friedman	National Bureau of Standards Washington, D. C.
Irving B. Friedman	Armour Research Foundation Chicago, Illinois
L. Furney	Argonne National Laboratory Lemont, Illinois
Robertson Gannaway	Raytheon Manufacturing Co. Chicago 41, Illinois
William Gerhard	Oak Ridge National Laboratory (on loan to ANL)
Herman H. Goldstine	Institute for Advanced Study Princeton, New Jersey

IX

Harold Greenspan	Argonne National Laboratory Lemont, Illinois
Sidney Greenwald	National Bureau of Standards Washington, D. C.
Louis Eugene Grosh, Jr.	Purdue University Lafayette, Indiana
Walter H. Gray	Raytheon Manufacturing Co. Waltham, Massachusetts
Jean Hall	Argonne National Laboratory Lemont, Illinois
Preston C. Hammer	University of Wisconsin Madison 6, Wisconsin
J. M. Hansen	Hughes Aircraft Company Culver City, California
James R. Harris	Bell Telephone Laboratories Whippany, New Jersey
Sol E. Harrison	Remington-Rand, Incorporated Philadelphia, Pennsylvania
(Miss) Ruth C. Haueter	National Bureau of Standards Washington, D. C.
Donald L. Herr	American Electronic Manufacturing Culver City, California
N. Hilberry	Argonne National Laboratory Lemont, Illinois
John Lindsay Hill	Engineering Research Assoc. St. Paul, Minnesota
Thomas C. Hill	Computing Lab., B.R.L. Aberdeen Proving Grounds, Maryland
G. W. Hobbs	General Electric Schenectady, New York
Arthur W. Holt	National Bureau of Standards Washington, D. C.

X

Arthur W. Horton, Jr.	Bell Telephone Laboratories Murray Hill, New Jersey
Alston S. Householder	Oak Ridge National Laboratory Oak Ridge, Tennessee
Charles Hubbard	Oak Ridge National Laboratory (on loan to ANL)
E. S. Hughes	IBM Corp. Endicott, New York
Harry D. Huskey	Wayne University Detroit, Michigan
L. Roy Harper	IBM Corporation Poughkeepsie, New York
R. A. Henle	IBM Corporation Poughkeepsie, New York
Merle Hayward	Phillips Control Corp. Joliet, Illinois
J. J. Iffland	Capehart-Farnsworth Corp. Fort Wayne, Indiana
W. B. Jacob ₁	Hughes Aircraft Company Culver City, California
David Jacobsohn	Argonne National Laboratory Lemont, Illinois
Kenneth W. Jarvis	Jarvis Electronics Corp. Chicago, Illinois
Eleanor Jutzi	Argonne National Laboratory Lemont, Illinois
S. Kaufmann	Argonne National Laboratory Lemont, Illinois
Theodore J. Kavis	Argonne National Laboratory Lemont, Illinois
R. J. Klein, Jr.	Oak Ridge National Laboratory (on loan to ANL)

John F. Koch, Jr.	Technitrol Engineering Co. Philadelphia, Pennsylvania
Raymond F. Kramer	Argonne National Laboratory Lemont, Illinois
H. E. Kubitschek	Argonne National Laboratory Lemont, Illinois
Robert Kudlich	University of Illinois Urbana, Illinois
Edward Lach	Argonne National Laboratory Lemont, Illinois
W. J. Ladniak	Oak Ridge National Laboratory Oak Ridge, Tennessee
Robert M. Leger	Illinois Institute of Technology Chicago, Illinois
J. Lerner	Argonne National Laboratory Lemont, Illinois
Samuel Lubkin	Underwood Corporation Long Island City, New York
Frank J. Lynch	Argonne National Laboratory Lemont, Illinois
Harold F. Lanier	Goodyear Aircraft Corporation Akron, Ohio
Frank J. Maginniss	General Electric Schenectady, New York
Theodore E. Martin	University of Wisconsin Madison, Wisconsin
B. T. Matthias	Bell Telephone Laboratories Murray Hill, New York
John W. Mauchly	Remington-Rand, Incorporated Eckert-Mauchly Division Philadelphia, Pennsylvania
Harold A. May	Argonne National Laboratory Lemont, Illinois

E. G. Mayer	RCA - Victor Division Lancaster, Pennsylvania
Robert H. McDowell	Argonne National Laboratory Lemont, Illinois
Torben Meisling	University of California Berkeley 4, California
Leslie Merrill	Argonne National Laboratory Lemont, Illinois
Nicholas Metropolis	Los Alamos Scientific Laboratory Los Alamos, New Mexico
Charles A. Miller	Argonne National Laboratory Lemont, Illinois
William F. Miller	Argonne National Laboratory Lemont, Illinois
John B. Miller	University of Wisconsin Madison, Wisconsin
Dorothea Minden	Argonne National Laboratory Lemont, Illinois
Morse Minkow	Underwood Corporation Long Island City, New York
Harry Moses	Argonne National Laboratory Lemont, Illinois
Roger D. Moulton	National Security Agency Washington 25, D. C.
John Muerle	University of Illinois Urbana, Illinois
David E. Muller	University of Illinois Urbana, Illinois
Walter E. Mutter	IBM Corporation Poughkeepsie, New York
Joseph B. Naines	Technological Institute Northwestern University Evanston, Illinois

J. P. Nash	University of Illinois Urbana, Illinois
Marilyn M. Nelson	Argonne National Laboratory Lemont, Illinois
B. J. Norris	Capehart-Farnsworth Corporation Fort Wayne, Indiana
John A. O'Brien	Massachusetts Institute of Technology Cambridge, Massachusetts
Kenneth H. Olson	Massachusetts Institute of Technology Cambridge, Massachusetts
Arthur Ostroff	Underwood Corporation Long Island City, New York
James H. Paquin	Consolidated Engineering Corp. Pasadena, California
Robert J. Parent	University of Wisconsin Madison, Wisconsin
Alan J. Perlis	Purdue University Lafayette, Indiana
Clay Perry	Oak Ridge National Laboratory Oak Ridge, Tennessee
George W. Platzman	University of Chicago Chicago 37, Illinois
James Pomerene	Institute for Advanced Study Princeton, New Jersey
Vincent J. Porter	Remington-Rand, Inc. Philadelphia, Pennsylvania
William H. Potter	Naval Ordnance Plant Indianapolis, Indiana
Charles F. Pulvari	The Catholic University of America Washington, D. C.
Richard M. Purinton	American Phenolic Corporation Chicago 50, Illinois

XIV

William C. Purple, Jr.	Melpar, Incorporated Alexandria, Virginia
Jan Rajchman	RCA Laboratories Princeton, New Jersey
E. Rauh	Argonne National Laboratory Lemont, Illinois
T. J. Reilly	General Electric Schenectady, New York
James Richardson	Los Alamos Scientific Laboratory Los Alamos, New Mexico
Vincent C. Rideout	University of Wisconsin Madison, Wisconsin
James E. Robertson	University of Illinois Urbana, Illinois
Juergen Roedel	Pullman Std. Car Mfg. Company Hammond, Indiana
Herbert V. Ross	Argonne National Laboratory Lemont, Illinois
W. H. Rous	American Phenolic Corp. Chicago 50, Illinois
Ward Sangren	Oak Ridge National Laboratory Oak Ridge, Tennessee
Arthur J. Schmitt	American Phenolic Corp. Chicago 50, Illinois
Herbert W. Schultz	Technological Institute Northwestern University Evanston, Illinois
M. G. Schlapkohl	Argonne National Laboratory Lemont, Illinois
Otto A. Schulze	Argonne National Laboratory Lemont, Illinois
Robert W. Schumann	Engineering Research Assoc. St. Paul, Minnesota

F. A. Schwertz	Mellon Institute Pittsburgh, Pennsylvania
Allen K. Scidmore	University of Wisconsin Madison, Wisconsin
David H. Shaftman	Argonne National Laboratory Lemont, Illinois
Theodore Shapin, Jr.	University of Illinois Urbana, Illinois
Morton Shepetin	Institute of Math. Sciences New York University New York 3, New York
Dr. Ralph J. Slutz	National Bureau of Standards Washington, D. C.
Blanchard D. Smith, Jr.	Melpar, Incorporated Alexandria, Virginia
Charles V. L. Smith	Office of Naval Research Washington 25, D. C.
E. M. Smith	RCA Victor Division Lancaster, Pennsylvania
William E. Smith	North American Aviation, Inc. Downey, California
Albert E. Smith	Office of Naval Research Washington 25, D. C.
Rodolfo M. Soria	American Phenolic Corp. Chicago 50, Illinois
William S. Speer	Computer Research Corp. Hawthorne, California
Robert K. Swank	Argonne National Laboratory Lemont, Illinois
L. E. Swedlund	RCA Victor Division Lancaster, Pennsylvania
Homer W. Spence	Computing Laboratory, B. R. L. Aberdeen Proving Grounds, Maryland

Norman H. Taylor	Massachusetts Institute of Technology Cambridge, Massachusetts
John Terandy	Argonne National Laboratory Lemont, Illinois
C. Todd	Purdue University Lafayette, Indiana
Dolan Harold Toth	Engineering Research Assoc. St. Paul, Minnesota
L. A. Turner	Argonne National Laboratory Lemont, Illinois
James D. Tillman, Jr.	University of Tennessee Knoxville, Tennessee
Charles F. Valach	American Phenolic Corp. Chicago 50, Illinois
Donald L. Waidelich	Argonne National Laboratory Lemont, Illinois
Harry K. Walker	Oak Ridge National Laboratory Oak Ridge, Tennessee
Willis H. Ware	Rand Corporation Santa Monica, California
Richard Wehrle	International Harvester Co. Chicago, Illinois
Leon H. Weiss	American Electronic Mfg., Inc. Culver City, California
Charles F. West	Air Force Missile Test Center Patrick Air Force Base, Florida
Joseph M. Wier	University of Illinois Urbana, Illinois
G. H. Winslow	Argonne National Laboratory Lemont, Illinois
Way Dong Woo	Raytheon Manufacturing Co. Waltham, Massachusetts

James Woody	Oak Ridge National Laboratory (loaned to ANL)
J. Howard Wright	National Bureau of Standards Washington 25, D. C.
William W. Youden	National Bureau of Standards Washington, D. C.
Edward D. Zimmer	Engineering Research Assoc. St. Paul, Minnesota
W. H. Zinn	Argonne National Laboratory Lemont, Illinois

XVIII

PROGRAM OF THE SYMPOSIUM

FOREWORD

Babbage's vision of long ago of an automatic digital computer has been realized in this decade, and in abundant measure. It sometimes seems that the machine age is about to change into the computing machine age. This has come about as the result of technological discoveries that Babbage could scarcely have dreamed of. The many machines differ in a variety of ways but are alike in being bulky and consuming much power. Recent discoveries seem to give hope that these undesirable characteristics can be overcome. It seems worthwhile to pause briefly and see where we have been, where we are, and whither we are bound. The Argonne National Laboratory is sponsoring a small, informal conference of workers in the computer field in the hope that such an assessment of progress will lead to a wider and more profound appreciation of the new developments and contribute to the early realization of smaller, cooler, and still better automatic digital computers.

Louis A. Turner
Director, Physics Division

XIX

Monday, August 3, 1953

9:30 A.M.

Museum of Science and Industry,
West Wing

REGISTRATION

10:00 A.M. - 12:00 M.

The Little Theatre - Museum of Science
and Industry

OPENING ADDRESS

L. A. Turner, Director, Physics Division
Argonne National Laboratory

APPLICATION OF TRANSISTORS TO HIGH SPEED COMPUTERS

Chairman: Jan Rajchman - RCA Research Laboratories
Princeton, N. J.

Transistor Circuitry for a 3-Megacycle Serial Computer

W. A. Cornell - Bell Telephone Laboratories, Whippany, N. J.

Tradic, a Transistor Digital Computer

J. R. Harris - Bell Telephone Laboratories, Whippany, N. J.

Future Developments in Semiconductor Computers

J. H. Felker - Bell Telephone Laboratories, Whippany, N. J.

1:30 P.M. - 4:30 P.M.

The Little Theatre - Museum of Science
and Industry

TECHNICAL SESSION
WILLIAMS TUBE MEMORY

Chairman: R. F. Kramer - Argonne National Laboratory
Lemont, Ill.

The IAS Williams Tube Memory

James Pomerene - Institute for Advanced Study, Princeton, N. J.

XX

ORACLE Memory System

R. J. Klein, Jr. - Oak Ridge National Laboratory, Oak Ridge, Tenn.

Relative Merits of Williams Memory Display

David Jacobsohn - Argonne National Laboratory, Lemont, Ill.

Illiac Memory System

J. M. Wier - University of Illinois, Urbana, Ill.

Maniac Memory System

J. Richardson - Los Alamos Scientific Laboratory, Los Alamos, N. M.

Design and Manufacturing Considerations of Oscillograph Type Electrostatic Storage Tubes

E. M. Smith - Cathode-ray Tube Development Group
RCA Victor Division, Lancaster, Pa.

Tuesday, August 4, 1953

9:30 A.M. - 12:30 P.M.

Chemistry Building Auditorium,
Argonne National Laboratory

WELCOME TO LABORATORY

N. Hilberry, Deputy Director
Argonne National Laboratory

FERROMAGNETIC MEMORY

Chairman: J. H. Falker - Bell Telephone Laboratories
Whippany, N. J.

Myriabit Magnetic Core Memory

Jan Rajchman - RCA Research Laboratories, Princeton, N. J.

Ferromagnetic Memory with Microsecond Access

Isaac L. Auerbach - Burroughs Corporation, Philadelphia, Pa.

Random Access 3-Dimension Magnetic Memory

Jay Forrester - Massachusetts Institute of Technology
Cambridge, Mass.

Computer Component Research at Mellon Institute

F. A. Schwertz - Mellon Institute of Industrial Research
University of Pittsburgh, Pittsburgh, Pa.

Characteristics of the ORACLE

E. W. Burdette - Oak Ridge National Laboratory, Oak Ridge, Tenn.

1:45 P.M. - 5:00 P.M.

Inspection Tour

History, Organization and Major Achievements
of Argonne National Laboratory

L. C. Furney, Assistant to Laboratory Director
Argonne National Laboratory

Inspection of AVIDAC, ORACLE, Cyclotron,
Nuclear Reactor and Van de Graaff Generator

The group will be divided into small parties. Guides will be provided and explanations will be made by persons associated with the projects on display; as follows:

AVIDAC:	Jean Hall
	C. Bergstrom
ORACLE:	William Gerhard
	J. W. Woody
	Clay Perry
Cyclotron:	J. P. Fitzpatrick
	Warren Ramler
Van de Graaff	
Generator:	Alexander Langsdorf
Nuclear Reactor:	J. T. Weills

7:00 P.M.

Hotel Windermere

BANQUET

"Computing the Weather"

Speaker: John W. Mauchly - Remington-Rand, Inc.
Eckert-Mauchly Division, Philadelphia, Pa.

Wednesday, August 5, 1953

9:30 A.M. - 12:30 P.M.

The Little Theatre - Museum of Science
and Industry

SPECIAL DEVICES

Chairman: J. P. Eckert - Remington-Rand, Inc.
Philadelphia, Pa.

Physical Aspects of Magnetic Computer Materials

S. E. Harrison - Remington-Rand, Inc., Philadelphia, Pa.

Magnetic Reading Recording Head Design

Vincent J. Porter - Remington-Rand, Inc., Philadelphia, Pa.

40-Channel Magnetic Tape Auxiliary Memory

Leslie Merrill - Argonne National Laboratory, Lemont, Ill.

Magnetic Tape Sprocket

Arthur Ostroff - Underwood Corporation, Long Island City 6, N. Y.

The Diode Capacitor Memory

Arthur W. Holt - National Bureau of Standards, Washington, D. C.

2:00 P.M. - 4:00 P.M.

The Little Theatre - Museum of Science
and Industry

GENERAL SESSION

Chairman: D. A. Flanders - Argonne National Laboratory
Lemont, Illinois

Some Topics on Experimental Mathematics

N. Metropolis - Los Alamos Scientific Laboratory, Los Alamos, N. M.

General Survey of Ferroelectrics

B. T. Mathias - Bell Telephone Laboratories, Murray Hill, N. J.

Some Experience in Coding and Programing with the Institute Computer

Herman Goldstine - Institute for Advanced Study, Princeton, N. J.

-1-

ANAL-512

A TRANSISTOR PULSE AMPLIFIER USING EXTERNAL REGENERATION

by

J. H. Vogelsong
Presented by W. A. Cornell

Bell Telephone Laboratories, Inc.,
Murray Hill, N. J.

SUMMARY

A pulse-regenerative amplifier using a point-contact transistor has been operated at a basic frequency of 3 megacycles. To produce regenerated pulses with waveshapes which are practically independent of the waveshapes of the input pulses, a germanium diode circuit has been used in conjunction with an external feedback path. This arrangement also provides for the synchronization of the output pulses with a master clock. Transformer coupling has been incorporated into the circuit to provide DC restoration.

INTRODUCTION

In synchronous serial computers of even a moderate degree of complexity, regeneration of pulses in amplitude, shape and time must be performed throughout the system. When logical operations are executed in circuits containing active elements such as vacuum tubes or transistors, this regeneration need not be separated from these operations since both can be performed simultaneously by the same circuit. On the other hand, logical operations can be carried out in circuits containing only elements which are essentially passive, such as diodes, to reduce the number of active elements required. In such instances regeneration can be provided by other circuits containing active elements.

Regenerative circuits of this kind, using point-contact transistors as the active elements, have been designed. In the one to be described, the transistor operates as an over-driven amplifier with external pulse feedback used in conjunction with germanium diodes to provide the regeneration. This mode of operation is in contrast to that of bistable circuits which utilize internal feedback supplied by the base resistance of the transistor.

The complete unit, to which a sinusoidal clock voltage is supplied, regenerates 0.17 microsecond pulses at a basic repetition rate of 3 megacycles. One of the desirable features of its operation at this frequency is the fact that its performance is relatively unimpaired by a slow recovery of reverse impedance which is found in many diodes.

2 GENERAL

Block diagrams of two possible logical configurations for the regenerative amplifier are shown in Figure 1. Both arrangements are well known in the computer field, the principle of external feedback being used in SEAC,¹ for instance. In either circuit a coincidence between an input signal and the proper portion of the clock wave is necessary to initiate an output pulse. Therefore, if the leading edge of the input signal is arranged to arrive slightly prior to this portion of the clock wave, the clock determines the timing of the start of the output pulse. To assure that the clock also determines the duration of the output pulse, a feedback path returns the output pulse in such a manner as to make it equivalent to the original input signal. In this fashion, the output pulse can be sustained until terminated by the clock, even if the original input signal is removed.

DETAILED CIRCUIT OPERATION

Figures 2 and 3 are circuit schematics of two slightly different implementations of the block diagram of Figure 1b. The ensuing description is concerned more directly with the circuit of Figure 2, but since the circuit of Figure 3 differs only in that a more conventional configuration of the AND Circuit is used, it will not be treated separately. To aid the description, the performance with no input signal applied will be considered first, and this will be followed by the modification resulting when an input signal is present.

As a starting point, consider the applied clock potential to be at the negative peak of the sine wave. At this time diodes D_1 and D_2 and the transistor emitter are all in a reverse impedance state. If these reverse impedances were infinite, the emitter potential would be equal to the applied clock potential; actually, as a result of finite back impedances the emitter potential will be somewhat less negative at this time. While diode D_2 is essentially nonconducting, the potential (V_2) at the junction of diodes D_2 , D_4 and D_6 and resistor R_3 is E_{clamp} minus the series forward voltage drops across diodes D_4 and D_5 . (This is based on the assumption that no current flows through D_6 or R_4 toward this junction.)

As the applied clock potential becomes less negative the emitter potential also rises. Diodes D_1 and D_2 and the transistor emitter remain nonconducting until the emitter reaches the potential, V_2 , at which time diode D_2 changes to the low forward impedance state. From this point, as the clock potential becomes still less negative and then positive, current flows through resistors R_1 and R_2 and diode D_2 from the clock source to the diode junction (D_2 , D_4 , D_6) at which V_2 is measured. This current does not appreciably change the potential, V_2 , but rather, replaces a portion of the current which had been flowing through diodes D_4 and D_5 into resistor R_3 . Thus the change in V_2 is only the small change in forward voltage drops across these diodes resulting from the decrease in current through them.

---

¹R. D. Elbourn and R. P. Witt, "Dynamic Circuit Techniques Used in SEAC and DYSEAC," Transactions of the I.R.E., Vol. EC-2, No. 1, pp.2-9, March 1953.

When the applied clock voltage becomes sufficiently positive to raise the potential (V_1) at the junction of resistors R_1 and R_2 and diode D_1 to the value E_{clip} , diode D_1 begins to conduct. Through this means the clock voltage is clipped, V_1 remaining at E_{clip} (plus the small forward voltage drop across diode D_1) until the positive peak of the clock voltage sine wave has passed and the clock voltage is no longer sufficiently positive to maintain V_1 at E_{clip} . During this clipping interval both V_2 and the emitter potential are essentially constant and the emitter remains slightly below cutoff. As V_1 becomes less positive than E_{clip} , diode D_1 again becomes nonconducting.

The cycle of input circuit operation, with no input signal present, is completed as the clock voltage swings toward the negative peak of the sine wave. Diode D_2 becomes nonconducting as the emitter potential drops below the potential, V_2 , which is again clamped by diodes D_4 and D_5 as at the beginning of the cycle. In Figure 4, waveshapes of some of the input circuit voltages and currents are plotted. The values of the constants which were used for the calculations are those indicated in Figure 2, with the exception that the emitter is assumed to run at ground potential while the clock potential is positive and at the clock potential when this is negative. These assumptions are reasonably valid whether or not an input signal is present. When no input signal is present, as has been assumed to this point, the curve in Figure 4 labeled, " V_1 or emitter current," should be interpreted as, " V_1 or current through diode D_2 ."

Application of positive pulses to the input modifies the performance of the input circuit during the intervals when the clock voltage is positive. For proper operation, which implies that the clock completely determines the timing of the output pulses, the input pulse must be applied prior to the time when the clock voltage swings positive through zero. When this is the case, diode D_2 does not begin to conduct as it does when no input pulse is present since V_2 is more positive as a result of the presence of the input pulse. (The input pulse causes diodes D_4 and D_5 to become nonconducting.) For diode D_2 to conduct, the emitter potential would have to be driven as positive as potential V_2 by the clock voltage. Before this occurs, however, the emitter potential reaches the value² at which the transistor emerges from the cutoff region.

After this point has been reached, the transistor emitter characteristic is such that for most external collector impedances a large increase in emitter current can occur with only a slight change in emitter potential. Moreover, this slight change may be either positive or negative, depending on the particular transistor. Thus, essentially the current which flows through diode D_2 when no input signal is present is diverted into the transistor emitter when such a signal is applied.

---

²This potential is referred to as the "break point," a term coined by B. G. Farley.

4

Voltage-current characteristics for both emitter and collector for a typical unit have been included in Figure 5. To illustrate the action of the transistor a linear load line has been plotted on the collector characteristic and has then been transferred to the emitter characteristic where it is no longer linear. The line, as plotted on the collector characteristic, has a 2000 ohm slope and passes through the point ($I_e = 0$, $V_c = -18$). With the particular transistor that was used to obtain the characteristics, a slight negative resistance is evident at the emitter in the region to the right of the break point; another transistor might have displayed a zero, or even positive, resistance in this region.

If the voltages and circuit constants indicated in Figure 2 are assumed, the presence of an input pulse results in a change in emitter current from a very slightly negative value to a positive value of approximately 3.5 milliamperes. The excursions over the emitter and collector characteristics are indicated by the load lines between these extremities. In Figure 5 the pulse voltage and current developed at the collector have been indicated as ΔV and ΔI . Since the collector is not driven too deeply into saturation, the unpredictable pulse broadening which results from "hole-storage" is avoided.

In general, the load lines corresponding to circuits which the amplifier will be called upon to drive will not be simple straight lines. In fact, if reactances are disregarded, they are apt to be similar to the characteristic plotted in Figure 6b, which is a plot of the input resistance of this amplifier when diode D_2 is cut off. To facilitate the calculations for this plot, the 2-straight line approximation to a diode characteristic, shown in Figure 6a, has been assumed for all diodes. The sharp corner in this assumed characteristic is responsible for the sharp corners in the input resistance characteristic.

As shown in Figures 2 and 3, the transistor collector is coupled to the load through a transformer and diode D_3 . The purpose of the transformer is two-fold. First, it matches impedances (to the extent that non-linear impedances can be matched). Second, it makes possible more satisfactory DC restoration than can be obtained if a condenser is used for coupling. For transistor circuits of this type, capacitive coupling results in a considerable reduction in useful pulse amplitude for a train of pulses as compared to the amplitude of a single isolated pulse. If a properly designed transformer is used, however, the useful amplitude of the pulses can be maintained essentially constant regardless of the number of consecutive pulses.

To obtain complete³ DC restoration, the total energy stored in the transformer mutual inductance and the shunt capacitance of both the transformer and the transistor output during the pulse should be completely removed prior to the start of the next pulse interval. This cannot be done, however, so it must be approximated as closely as possible. The optimum

---

³ Complete is used here to indicate a maintenance of the over-all waveform, not merely amplitude.

condition involves a ringing by the transformer (following a pulse) at a frequency of slightly less than 3 megacycles with damping slightly less than critical. To this end, diode D_3 is arranged to effectively disconnect the load from the transformer as the transformer polarity reverses at the termination of a pulse. The transformer ringing frequency and the damping are then determined by parallel combination of transformer mutual inductance, shunt capacity and effective shunt resistance, and the transistor output resistance and capacity.

It should be noted that DC restoration is not obtained without cost. The price paid is the energy stored in the reactive elements during the signal pulse - a greater amount of energy than would have to be stored if no DC restoration were desired.

As indicated previously, an external feedback path is provided for temporary lock-up to assure proper retiming of the pulse. From Figure 2 it can be seen that diodes D_4 and D_6 constitute an OR-Circuit. Therefore, in returning to the input via diode D_4 and current limiting resistor R_4 , the feedback pulse becomes equivalent to the input pulse. Diode D_5 is a low impedance source of current for clamping the potential at the junction of diodes D_2 , D_4 and D_6 whenever the clock voltage swings negative. With this current source available, the current need not be supplied through the transformer, and a source of 3 megacycle "noise" at the amplifier output is thereby avoided.

The clamping action of diode D_5 would undoubtedly be more effective if its cathode were brought to the junction of diodes D_2 , D_4 and D_6 . To the present time it has been connected as shown, however, to decrease the possibility of its providing a low impedance shunt across resistor R_3 , as a result of poor diode reverse recovery, when an input pulse is present.

Photographs of waveshapes at various points in the amplifier of Figure 2 are included as Figures 7 and 8. To obtain Figure 7, a 500 ohm resistive load and a 3:1 transformer turns ratio were employed, whereas, Figure 8 was obtained with a load consisting of four amplifiers in parallel while the transformer turns ratio was 4:1.

POWER DISSIPATION

For the voltages and circuit constants indicated in Figure 2, the average 3 megacycle clock power dissipation is approximately 12 milliwatts per amplifier. This figure is substantially the same whether or not the amplifier is being pulsed. On the other hand, the DC dissipation within the transistor is very nearly equal to the product of E_2 and $I_C(0, E_2)$.⁴ The exact value, therefore, is directly dependent upon the particular transistor, but is usually less than 50 milliwatts. An additional 25 milliwatts of DC power are dissipated external to the transistor.

---

⁴ $I_C(0, E_2)$ is the collector current corresponding to an emitter current of zero and a collector voltage of E_2 .

COMPARISON OF INPUT CIRCUITS

Two different implementations of the AND Circuit in the input have been indicated in Figures 2 and 3, respectively. The arrangement of Figure 2 seems to be superior to that of Figure 3 in that a slow recovery of reverse impedance in diode D_7 in Figure 3 is detrimental to the performance of the circuit. If the sine wave clock were applied directly to this diode, the transistor emitter potential could be carried over the break point by each of the positive voltage swings of the clock regardless of the presence or absence of an input signal. A clipping circuit consisting of R_1 and D_1 can be used to alleviate this condition, however, by clipping the positive peaks of the clock voltage at a value slightly more positive than the transistor break point. The normal running potential at the emitter could, of course, be made more negative, but this would necessitate larger input pulses to operate the amplifier.

As might be expected, the analytical expressions defining output pulse rise times are different for the two circuit versions. Specifically, these differences reflect the fact that, for the circuit of Figure 2, the emitter current is supplied from a sinusoidal clock generator, whereas, for that of Figure 3, it is supplied from a source of constant potential. A representation, which disregards the transformer and is suitable for either circuit during the rise time interval, is shown in Figure 9. To differentiate the cases, the applied voltage is a suddenly applied sine wave in one instance and a step function in the other. The only frequency-dependent element introduced into the tee-equivalent for the transistor is the mutual impedance, Z_m , which is postulated to have the form

$$\frac{R_m}{1 + j \frac{\omega}{\omega_c}}$$

In this expression, ω is the angular frequency variable, $2\pi f$, and ω_c is the particular angular frequency for which the magnitude of Z_m is down to 0.707 times the DC value. The unit step function is represented by the notation, $u(t)$, in the following material.

For the circuit of Figure 9 the following loop equations can be written in terms of the complex frequency variable, p :

$$[R_1 + r_e + r_b] I_e(p) + r_b I_c(p) = E(p) \quad (1)$$

and

$$[r_b + Z_m(p)] I_e(p) + [r_b + r_c + R_2] I_c(p) = 0 \quad (2)$$

From a simultaneous solution of these equations for $I_c(p)$, the expression for $V(p)$ can be readily obtained, since

7

$$V(p) = - I_c(p) R_2 = \frac{R_2 [r_b + Z_m(p)] E(p)}{[R_1 + r_e + r_b] [R_2 + r_b + r_c] - r_b [r_b + Z_m(p)]} \quad (3)$$

Laplace transforms can now be used to obtain the time response, $V(t)$, when the appropriate representations for $E(p)$ are inserted.

Based on these conditions, the equations for the output voltage response are:

(for a suddenly applied sine wave voltage described here as $E(t) = E_0 \sin \omega_0 t u(t)$)

$$V(t) = \gamma E_0 \left[\frac{(\lambda - \beta) \frac{\omega_c}{\omega_0}}{\left(\beta \frac{\omega_c}{\omega_0}\right)^2 + 1} e^{-\beta \omega_c t} + \sqrt{\frac{\left(\lambda \frac{\omega_c}{\omega_0}\right)^2 + 1}{\left(\beta \frac{\omega_c}{\omega_0}\right)^2 + 1}} \sin(\omega_0 t + \psi) \right] \quad (4)$$

(applies to the circuit as shown in Figure 2)

and,

(for the step function where $E(t) = E_1 u(t)$)

$$V(t) = \gamma E_1 \frac{\lambda}{\beta} - \left(\frac{\lambda}{\beta} - 1\right) e^{-\beta \omega_c t} \quad (5)$$

(applies to the circuit of Figure 3)

where

$$\beta = \frac{[(R_1 + r_e + r_b)(R_2 + r_c + r_b) - r_b R_m]}{(R_1 + r_e + r_b)(R_2 + r_c + r_b) - r_b^2}$$

$$\gamma = \frac{R_2 r_b}{(R_1 + r_e + r_b)(R_2 + r_c + r_b) - r_b^2}$$

$$\lambda = 1 + \frac{R_m}{r_b}$$

$$\psi = \phi_1 - \phi_2$$

$$\phi_1 = \tan^{-1} \frac{\omega_b}{\lambda \omega_0}$$

$$\phi_2 = \tan^{-1} \frac{\omega_0}{\beta \omega_c}$$

$\mathcal{E}$

Generally λ is much larger and β slightly smaller than unity. In addition, $\frac{\omega_c}{\omega_0}$ will be considered only for values of unity or greater, so the following approximations are usually satisfactory:

$$\lambda - \beta \sim \lambda$$

$$\lambda/\beta - 1 \sim \lambda/\beta$$

$$\sqrt{(\lambda \frac{\omega_c}{\omega_0})^2 + 1} \sim \lambda \frac{\omega_c}{\omega_0}$$

$$\tan^{-1} \frac{\omega_0}{\lambda \omega_c} \sim 0$$

With these approximations, equation (4) can be reduced to

$$V(t) \sim \gamma \lambda E_0 \frac{\omega_c}{\omega_0} \left[\frac{1}{(\beta \frac{\omega_c}{\omega_0})^2 + 1} e^{-\beta \omega_c t} + \frac{1}{\sqrt{(\beta \frac{\omega_c}{\omega_0})^2 + 1}} \sin(\omega_0 t - \phi_2) \right] \quad (6)$$

and equation (5) can be reduced to

$$V(t) \sim \frac{\gamma \lambda E_1}{\beta} (1 - e^{-\beta \omega_c t}) \quad (7)$$

Equations (4) and (6) are, of course, valid only prior to the time when the clipping of the clock commences.

It is evident that the circuit of Figure 3 is superior to that of Figure 2 with regard to rise time. This follows from the fact that the limiting best response obtainable by equation (6) is sinusoidal, whereas, equation (7) indicates a limiting best response which is a step function. Use of a square-wave clock with the circuit of Figure 2 can nullify this difference, however, if a satisfactory clock source of this type is available. As a substitute, a sinusoidal clock voltage with excessive amplitude could be used with the clipping level unchanged, but such a course is wasteful of clock power.

Other rise time information is conveyed by the equations. For instance, increasing the load resistance, R_2 , can shorten rise times if the transistor is driven to saturation. Increasing R_2 has the effect of increasing the multiplying factor, γ , and thereby causes $V(t)$ to rise toward a greater ultimate value. When the transistor collector reaches the saturation region, where further increases in emitter current produce little change in collector current, $V(t)$ is arrested at a maximum value which is less than the ultimate value indicated by the equations. Since either a sinusoidal or an exponential rise is being halted short of a goal, a reduction in rise time is effected.

9

Also useful for providing shorter rise times are transistors for which λ is large, in that they can more readily be driven to saturation. To a certain extent, λ is a measure of transistor gain, and it can be seen that it is a multiplying factor in both response expressions. It should be pointed out again, though, that pulse broadening is apt to result from "hole-storage" if the unit is driven too deeply into saturation.

Figures 10 and 11 are included to indicate the extent to which transistor cutoff frequency influences rise time. They have been plotted from equations (6) and (7), respectively, for the single case of an assumed β of 0.9. In either figure the higher cutoff frequencies produce the shorter rise times, as expected, but a point of diminishing returns is reached. It is interesting to note that when transistors are driven into the saturation region, gain can be traded for cutoff frequency and vice versa. Best over-all performance is nevertheless obtained with transistors having both a high cutoff frequency and large current gain.

ACKNOWLEDGEMENTS

This work on the regenerative amplifier has been done under contract with the Bureau of Ordnance of the U. S. Navy.

The author wishes to acknowledge the many useful discussions with W. A. Cornell and Q. W. Simkins and the test work performed by W. Kaminski and M. F. Wazowicz.

10

(a)

(b)

Figure 1

Logical configurations for a pulse-regenerative amplifier.

Figure 2

Regenerative Amplifier I.

12

Figure 3

Regenerative Amplifier II.

Figure 4

Voltage and current waveshapes for Regenerative Amplifier I.

14

Figure 5

Transistor characteristics to illustrate development of a pulse.

Figure 6

Straight line approximations to: (a) diode characteristic, (b) regenerative amplifier input characteristic.

16

Figure 7

Photographs of voltage waveshapes at various points of Regenerative Amplifier I. A 500 ohm resistive load and a 3:1 transformer turns ratio were used.

Figure 8

Photographs of voltage waveshapes at various points of Regenerative Amplifier I. A load consisting of four similar regenerative amplifiers in parallel and a transformer turns ratio of 4:1 were used.

Figure 9

Equivalent circuit for rise-time analysis.

Figure 10

Graphical illustration of the influence of transistor cutoff frequency in determining pulse rise time of Regenerative Amplifier I.

20

Figure 11

Graphical illustration of the influence of transistor cutoff frequency in determining pulse rise time of Regenerative Amplifier II.

21 TRADIC, A TRANSISTOR DIGITAL COMPUTER

by

James R. Harris

Bell Telephone Laboratories
Whippany, New Jersey

INTRODUCTION

In this talk I want to share with you some of the experience of building a high-speed digital computer based on transistors. In particular, I want to give you the feel of what it's been like to work with transistors on this project, project TRADIC, and what kind of computer is now emerging.

The TRADIC computer is intended for real-time military problems. You can think of it as a general purpose computer for problems that have very small memory requirements. It is well adapted to problems that have a few subroutines that are used repeatedly. A typical subroutine is rotation of coordinates, that is, finding the coordinates of a point in a new system that is rotated with respect to the first. The ultimate aim of TRADIC is a highly reliable airborne computer that will occupy 2 or 3 cubic feet and dissipate less than 100 watts.

First, let me say how TRADIC fits into the chronology of the transistor. A little over three years ago, the first experimental high-speed switching transistor, the M-1734, was developed in our Murray Hill Laboratories. The TRADIC computer now being built is based entirely on the M-1734 transistor.

HOW THE TRANSISTOR IS USED IN THE COMPUTER, AND WHY IT IS SO USED

When the high-speed transistor first became a reality, it seemed natural to ask how it could best be used. Should it be a flip-flop, or a pulse amplifier, or a logic element, or what? I want to review how this question was answered, and at the same time bring out some transistor properties and show some circuitry. To do this, let's examine the M-1734 transistor. The first slide shows the transistor in a simple circuit. Resistance in the base circuit is a feedback element. There are two important ideas I want to convey. One is the inherent feedback of the transistor. The other is the charge storage effect.

FEEDBACK

First, let's look at feedback and see the strange things it does. An increment of current driven rightward at the input produces a rightward increment of current at the output. This polarity of transmission is typical of all simple point contact and junction transistors in the active region with the common base connection. Having noted the polarity of transmission, note that the 1734 type transistor produces current gain.

For example, an increment of 1 milliamperes of emitter current produces an increment of about 2 milliamperes in the collector circuit. (This is current gain typical of point contact transistors.) The difference in current is drawn upward through the base electrode. This base current makes the base voltage more negative. What happens to the emitter voltage? The emitter follows the base. This can be understood by examining the emitter-to-base circuit inside the transistor. Note that the "emitter diode" within the transistor is conducting and has a low voltage drop.

Now, what have we found? When a positive current increment is driven into the emitter terminal, feedback produces a negative voltage increment at the emitter. That is, the emitter exhibits negative resistance.

The negative resistance is shown on the right-hand side of the slide, where emitter current is plotted to the right and negative emitter voltage is plotted down. In the region where the transistor has gain, called the active region, an increase in emitter current makes the emitter voltage go negative. Over at the left, we have a region where the emitter diode is reverse biased and shows a high positive resistance. At the right, we have a region of low positive resistance. In this region, the transistor acts almost like a short circuit. The transition from the central to the right region occurs largely because the voltage across the collector diode falls too low to produce gain.

We started out to examine the effects of inherent feedback, so let's ask now if this feedback is truly inherent in the M-1734 transistor. The answer is yes. To understand this, look at the two possible ways of getting rid of feedback. One of these is to raise the external impedance in the collector circuit and so reduce the current gain in the circuit, thus reversing the base current. The high cut-off frequency of the M-1734 transistor makes this impossible. Why? Because stray capacity makes a short circuit from collector to ground. The second way to get rid of feedback is to cut out the base resistance. Unfortunately, many M-1734 units have base resistance of 500 ohms or more within the transistor itself. So we conclude that feedback and the resulting negative resistance are inherent.

44

HOW A NEGATIVE-RESISTANCE CIRCUIT CAN BE USED

Let's see what we can learn from this curve. The first thing to do is to observe that we can add a so-called "load line" to the curve. This is just the same as adding a load line on the plate family of a vacuum tube.

If we put a low resistance across terminals A and B, we have a rather horizontal load line. If we put a proper voltage in series, we shift the load line vertically, and we can make it intersect in three places, and we have a flip-flop. The flip-flop is stable at the left and right intersections, but not at the central intersection.

Next let's ask if we can make an amplifier of the ordinary almost-linear type. The answer is no. The reason is that the circuit won't sit still in the active region: it's unstable. Why and how is it unstable? First why. It is an experimental fact that a circuit with a curve like this is unstable when terminated with a low impedance. At high frequency, say 30 mc., where many 1734 units have gain, stray capacity across terminals A and B is a low enough impedance to cause instability.

Now how does this instability show up? Let me show by an example. Suppose we have a high resistance load line, nearly vertical, with only one intersection. Suppose we place the load line to intersect in the left region and gradually move it rightward until the emitter current is about 100 microamperes. The circuit becomes unstable and the emitter current jumps by a factor of perhaps 50 in a few millimicroseconds.

The very feature that prevents ordinary amplifier operation makes it easy to build a pulse regenerator. Let me describe one way to make a pulse regenerator. We start with a flip-flop sitting in its low current state. We use the input pulse to move the operating point into the active region. Positive feedback snaps the circuit to its high current state and produces a strong step of collector current. Next a steady state clock signal drives the circuit back to its low current state and ends the pulse of collector current.

With this background, you can see that there are two basic ways of building a computer with this transistor. One uses the transistor in an ordinary flip-flop. The second uses the transistor as a pulse regenerator. The TRADIC computer uses the transistor purely as a pulse regenerator, and I want to explain how this choice was dictated by transistor properties.

WHY THE TRANSISTOR IS USED AS A PULSE REGENERATOR

First examine one aspect of a single-transistor flip-flop. Sitting in the high-current saturated state, the transistor has low impedance and requires a rather heavy current signal to switch it off. This is bad if the

24

switch-off signal has to come from an information source. Another disadvantage of the flip-flop comes from the charge storage effect. Think of it this way: collector current flows because of the presence of carriers that come from the emitter. When a transistor is saturated with these carriers, it takes a long time to stop the collector current. In our flip-flop, the time required to switch out of the high-current state can be as much as 4 or 5 microseconds. Of course, 5 microseconds is too long for a high-speed computer. If we could operate just at the edge of the active region, there would be fewer excess carriers, and switching would be quicker. However, we must keep the operating point out of the active region to avoid oscillation. With these limitations, it appeared that a circuit which depends on biasing just to the boundary would be too critical. So the flip-flop used in the ordinary way was not attractive.

Fortunately, both objections are overcome when the flip-flop is used as a pulse regenerator. In this use you will recall that the flip-flop is switched from low to high current by an information pulse and then reset to low current by a steady-state clock signal. The circuit for this is shown on slide No. 2. Note that the emitter must be relatively positive to switch on the transistor. When there is a positive pulse on the emitter (a "one" signal) and the clock wave at the base goes negative, the transistor switches on. Then, when the clock goes positive, the circuit switches off. This action is shown by the waveforms on slide 3. Note that the phase of the clock is delayed with respect to the input signal to allow the emitter capacity to charge positiveward. TRADIC uses a four-phase clock. With this system, the fairly heavy current needed for reset can now be obtained without too much trouble from a sine-wave generator.

It is especially interesting to see how the long switch-off time is overcome. In the pulse regenerator, the transistor never stays ON more than one pulse time. Because of the short time, not many carriers can accumulate in the transistor. With half microsecond pulses the switch-off time is about a tenth of a microsecond in this circuit.

HOW THE PULSE AMPLIFIER IS USED IN A SYSTEM

At this point I hope I've given you some feeling for the transistor properties that led us to use it as a pulse regenerator. Now I want to touch on the way the basic pulse amplifier is used in a system. I can save effort by saying that the TRADIC system is close kin to SEAC, the Bureau of Standards Eastern computer. The elements used for high-speed computing are semiconductor diode gates for the AND and OR functions, inverting transformers for the INHIBIT function, distributed delay lines, and the pulse regenerating amplifiers.

25

In the TRADIC computer being built, all of these elements are in plug-in packages with the functions shown on slide 4. The INHIBIT package requires pulses on the top two leads and no pulse on the bottom lead to produce an output. The bottom package has in it a delay line. In addition to the five packages shown, there is a sixth package containing only passive delay.

OPERATION OF THE PULSE AMPLIFIER CIRCUIT

Now, before going on - to a larger view of the system - let's look again at the basic circuit, slide 2. Let's ask how much gain it has and how this gain is produced. The first thing to notice is that the over-all current gain of the circuit is about 4, greater than the transistor "alpha." This is in spite of loss in the input gate and output capacitor. How can this be? By positive feedback and a purchase of circuit gain with excess transistor bandwidth. Remember that the circuit can be triggered from OFF to ON by just 100 microamperes of emitter current, providing the emitter-to-ground impedance is low. With the poorest M-1734 (low cut-off frequency) 20 micromicrofarads from emitter to ground is enough to provide the low impedance and to cause instability.

What must the signal do? It must charge the emitter capacity to a voltage positive enough for the emitter diode to conduct, and it must provide at least 100 microamperes of emitter current. In the circuit, this current must come from the 22 K resistor, which provides about 300 microamperes. Only about half of that is needed to charge the capacity. Why the extra 150 microamperes? The answer is very interesting. Because the rest is lost as transient reverse current in the emitter clamp diode, even though special diodes are selected by means of a switching time test on the diode.

How is it that transistor bandwidth buys circuit gain? If the transistor bandwidth were less, the emitter capacity would have to be greater and more current would be needed to charge it in the one-quarter microsecond that is available. The switch-on time of the M-1734 transistor is roughly 0.03 microsecond, which is really shorter than necessary for good pulse shape.

Diode logic circuits connect quite naturally to the input of the basic circuit. The output impedance is about right to drive four logic circuits. This is indicated by tests made with transformers in which a 1:1 ratio was found best.

26

MARGINAL CHECKING AND TRANSISTOR TESTS

We have relied on marginal checking of functional packages to be sure we will have good margins in the computer. We try to assure ample margins in all the things we call "elements of performance" - proper triggering by a "one" signal - proper switch-off by the clock - sufficient output, etc. To this end we have worked out a set of acceptance tests which define the M-1734 transistor. The set consists of 6 DC tests and two dynamic tests. It is 100% successful in assuring that a transistor will operate in the circuit and about 97% successful in assuring full safety margins. Even though the tests were developed for a specific circuit they assure that the M-1734 is workable in quite a wide range of circuits.

DESCRIPTION OF THE TRADIC COMPUTER

I promised to say what kind of computer is emerging from this work. Construction is about half finished on what we call the laboratory feasibility model computer. It will have about 850 transistors and will be substantially full-sized except for the number of steps in the main programs. It is a serial machine, purely binary, with a word length of 16 digits. To classify it by number of addresses, I can say the following: Every step of the computation selects a single address in the storage unit. Some steps of computation select an address and also give an operational command.

ARITHMETIC UNIT, STORAGE UNIT, AND INPUT-OUTPUT EQUIPMENT

For further details, let's look at slide No. 5. The arithmetic unit here is serial type. Why? Because serial operation is fast enough for the application of the computer. Addition takes 16 microseconds and multiplication takes 272 microseconds. There are about 60 transistors in the arithmetic unit proper.

The storage unit has space for 18 words for storing intermediate results of computation and 13 words of hand-set constants. These hand-set constants are part of the treated information and have nothing to do with programming. How did we choose a type of storage system? We looked them all over to find something compatible with transistors, something with good inherent reliability, something with no tough problems to solve. The system chosen for the feasibility model computer uses electric delay line storage. Delay sticks made by Technitrol in Philadelphia are used with a regenerating amplifier inserted every four microseconds. About 150 transistors are used in the upper part of the block and 50 in the lower part of the block. This is more transistors than we like to use in the storage system. To cut down on physical size and the number of transistors, we are developing a transistor version of the AVRAMUDAC system that was

27

developed at the Bureau of Standards by Arthur Holt. Our version, called MUDCAT, uses Bell Telephone Laboratories' silicon junction diodes as switch elements. We have been able to get a supply of experimental diodes with reverse current less than 2×10^{-10} ampere. Leakage of charge from storage condensers (500 micromicrofarads) is so low that no programmed regeneration is necessary.

We have a project at Whippany to explore and develop coders and decoders. I think the development of most interest to you is an optical shaft position encoder. This has an accuracy of 13 binary digits (one part in 8192). It uses a code wheel 5 inches in diameter with germanium photo-cells for reading.

Information flows between units in serial form. It flows either on the right-hand bus alone or on the left-hand bus alone. I mentioned that a step of computation carries one address and may carry one command. Let me describe a typical step: A typical step is address: "address 12" and command: "take a number from the storage unit and add it to the number already in the arithmetic unit." The mechanism of selecting a word in storage is as follows: Each of the word slots in the storage unit has one control wire. Selection is made by putting a pulse on the one proper wire.

In order to shift a word from one of the coders into storage, the address in the storage unit is given in the usual way, and the particular coder is designated as part of the command.

CONTROL SYSTEM

The remaining part of the feasibility computer is the control system, and this is shown on slide No. 6. What do we ask the control unit to do? For an example, suppose there are 64 steps in the computation. Then the control unit must have 64 internal states. It must advance from one of these states to the next at the proper time. At each step (state) the control unit must produce an output in this form: A signal on the proper lead in the address lead group and a signal on the proper lead in the operational command lead group. How fast must the control unit act? The shortest time between steps is 16 microseconds, the time required to add or to move a word between units. This is 62.5 kc. How does the control unit know when to advance? It receives a pulse from the minor counter. This pulse says that the previous address has been gated out, and the control unit should start advancing to the next step. The minor counter, or order generator, uses about 140 high-speed packages.

Now the question is what have we elected to put inside the box marked control unit, so let's look at slide No. 7. For simplicity, the subroutine parts are omitted from this slide. There is a counter followed by an AND matrix

followed by an OR matrix. The counter remembers which step of computation is under way. The AND matrix translates the state of the counter to one-hot code to make an electronic stepper. The OR matrix translates the output of the AND matrix to addresses and operational commands. The desired program is set in by joining horizontal and vertical wires at a cross-point. There is one connection on each horizontal lead.

There are some noteworthy features I want to mention. First is the use of silicon junction diodes in the OR matrix. Ordinarily, there is severe attenuation in a large OR matrix because of reverse current in the diodes. We have used nonpremium silicon junction diodes having a reverse current of 10^{-8} ampere at 6 volts. The result is that attenuation due to reverse current is quite small.

In order to reduce attenuation in the AND matrix, a scheme developed at the Bureau of Standards was used. The AND matrix is made up of transformer AND gates, each having two inputs, with one gate at each crosspoint. The input to the matrix must be one-hot in the horizontal group and one-hot in the vertical group. This input is provided in TRADIC by a pair of ring counters made from standard packages. It's noteworthy that high-speed packages "proved in" even though the counter never changes more often than 16 microseconds.

The AND matrix works with half-microsecond pulses. The capacity in the OR matrix is quite high, so that its output tends to be the envelope of blocks of pulses. The output current of the OR matrix is only about 100 microamperes, which is too small to trigger a standard package. A modified package called "Integrating AND" is used here. The trickle of output current is accumulated on a shunt condenser and then dumped into the pulse amplifier when a signal is given by the minor counter.

EVALUATION

I mentioned that the purpose of the machine now being built is to show feasibility. The question, then, is what has been shown by the work so far. Well, the essential parts of all units of the computer have been tested separately, and the results are encouraging. The greatest bulk of experience we have has come from a life test on a multiplier and word generator containing 78 transistors and 500 diodes. This equipment has operated day and night since September of 1952. There have been 3 transistor failures and one diode failure. Transistor failure rate is half of 1% per 1000 hours, which compares very favorably with the vacuum tube experience. Marginal checks show the remaining units to be in good shape. We think a good part of the transistor failures can be eliminated by a new cartridge mounting that would provide better humidity protection.

29

We have studied random errors by circulating a word in a storage register. Over the months, we have found that we lose the circulating word after periods ranging from about a week to a month, but so far there has been a failure of AC power to account for each loss of information.

We have made no attempt to get good performance over the military temperature range. Why? Because we assumed that this would allow us to design for better reliability at room temperature and that power dissipation would be low enough to make it easy to provide an environment with temperature control. This seems to be borne out by experience to date.

Now, what size computer are we building? You can think of it as a block 6 feet long, 1 foot thick, and 5 feet high. Work is going on now at Bell Laboratories which promises a many-fold reduction in size.

In summary, let me say this. The transistor technology developed on TRADIC shows the way to very low power, highly reliable, high-speed arithmetic units and control units. When the memory requirements are small, the technology is available for a complete computer.

ACKNOWLEDGEMENT

Work on this project is supported by the Air Force under contract AF33(600)-21536. Project reports can be obtained from Wright Air Development Center, WCLGB-3, by demonstrating a proper need for the information.

REFERENCE

The basic transistor circuit for pulse regeneration was described and analyzed in Proceedings of the IRE, November, 1952, Vol. 40, pp. 1584-1596, by J. H. Felker.

TRANSISTOR NOTE

The M-1734 transistor is not to be produced by Western Electric. The immediate successor to the M-1734 has a different mechanical form and has the laboratory type number 1760.

30

Figure 1

EQUIVALENT CIRCUIT OF TRANSISTOR TEST CIRCUIT EMITTER VI CURVE

3/

Figure 2
THE BASIC RC COUPLED PULSE AMPLIFIER

32

Figure 3

WAVEFORM FOR THE BASIC RC COUPLED PULSE AMPLIFIER

NOTE:

THE VOLTAGE OF THE BASE LINE OF THE COLLECTOR WAVE DEPENDS ON THE NUMBER OF PULSES THAT HAVE GONE THROUGH THE OUTPUT COUPLING CAPACITOR IN THE RECENT PAST.

Figure 4

THESE FIVE FUNCTIONS ARE USED IN THE HIGH-SPEED SECTIONS OF TRADIC. THERE IS A PLUG-IN PACKAGE TYPE FOR EACH FUNCTION. EACH PACKAGE HAS ONE TRANSISTOR, USED PURELY TO AMPLIFY AND RESHAPE PULSES.

34

Figure 5

INFORMATION PATHS IN TRADIC

35

Figure 6

TRADIC CONTROL SYSTEM

36

Figure 7

TRADIC MASTER AND SUBROUTINE CONTROL UNIT

37

INSTITUTE FOR ADVANCED STUDY WILLIAMS MEMORY*

by

James H. Pomerene

Institute for Advanced Study
Princeton, New Jersey

INTRODUCTION

The computer group at the Institute for Advanced Study began work on the Williams memory system in June 1948 on the basis of a report prepared by Williams and Kilburn. The first effort was naturally concentrated on verifying the essential storage phenomena. We drew freely on the Kilburn report; keeping, for example, the fixed interlace of action and regenerate cycles. However, we adopted from the start the balanced, current adder type of deflection system, now widely used.

Storage on a 1024 raster was achieved in September 1948 using a one tube experimental setup. The next step was to see if this memory effect was stable and reliable enough to work with a large scale digital computer. This was a very important question. The regeneration process itself might gradually degenerate or else the system might be hopelessly sensitive to unavoidable disturbances. In the following eight months we learned a great deal about ground currents, amplifier stability, and the need for very careful shielding of the analogue deflection circuits. However, error free endurance runs of 7 to 34 hours were achieved in May through August of 1949. The next year was spent in design, construction, and tube selection for the 40 tube memory now in use, which was completed and working in July 1950.

THE PRESENT MEMORY

The main features of the IAS memory are given in Figure 1. The 5CP1A cathode ray tube was used because it has a long, separate third anode coating which effectively shields the pickup plate from the gun elements. The second anode potential is restricted to 1000 volts by the present direct coupled deflection system; it will soon be increased to the order of 2000 volts. The 24 microsecond period is very conservative and will probably be considerably reduced.

---

*The construction work of this project was done under contracts with the U. S. Army Contracts No. W-36-034-ORD-7481, DA-36-034-ORD-19, and DA-36-034-ORD-1023.

38

The physical appearance of the computer is shown in Figure 2. There are 20 memory tubes on each side. Each is contained in a double mu metal and copper shield which is effective against external fields up to 15 Gauss. A three stage video amplifier, built in circular form, is mounted within the end of the shield, keeping the low level signal path entirely inside the shield.

Above each tube is an eight tube discriminator which accepts the amplifier output and, on writing cycles, the digit to be stored, and determines the appropriate beam turn on routine.

Four different turn on routines are provided in Figure 3. These correspond to the operations of regenerating dot on dot and dash on dash and of storing dot on dash and dash on dot. These permit separate adjustments for optimum reading and writing.

The state of a point is determined by sampling the first $1/2 \mu\text{sec}$ of amplifier output after initial beam turn on, and a toggle is set to store this information through the remainder of the cycle.

After the dot turn on, the beam is turned off and displaced slightly, or "twitched." The direction of twitch is reversed on alternate columns to reduce inter-spot coupling. The beam is not turned on again until twitching is complete; thus a double dot is used rather than a dash.

After the dash turn on the decision is made as to whether the next cycle will be a regenerate cycle or an action cycle. The appropriate address is then given to the deflection system in readiness for the next cycle.

The deflection system is direct coupled and settles within a micro-second. Each address digit switches a current determined by one resistor. The currents are combined and passed through a cascode arrangement of four triodes (per deflection plate) ending in a resistor across which the deflection signal is generated at the second anode voltage level. Cathode followers are used to drive the 40 deflection plates in parallel.

The IAS memory is now three years old. The life record of the memory tubes is shown in Figure 4. Ten of them are original with almost 8000 hours of active service. Only five tubes actually expired. All other removals were made to improve the initial performance of the memory. No replacements were needed for created flaws or raster burn.

The pressure of mathematical work has prevented any major engineering changes in the memory for the past two years. We have therefore tried to improve the read-around, or spill, performance of the memory by CRT replacement alone. Both RCA and DuMont have cooperated with us in making new tubes available for test. Results on tests of the last 250 tubes are given in Figure 5. Almost 30% of the production tubes are accepted on our flaw test. Of these 70% pass a five point spill test.

37

FUTURE DEVELOPMENT

The experience of the Institute group with the Williams memory has been very encouraging. External noise has been no problem. Internal noise has been controlled by eliminating certain noisy classes of components and by minimizing dust deposits on high voltage points. Flaws are controlled, if tediously, by tube selection.

Inter-spot coupling, or read-around, is still a challenging problem. At present it is possible to achieve quite satisfactory read-around performance with 1024 bits per tube. However, a major increase in bits per tube, say to 2048 or more, seems to require new methods of attack.

For a given amount of coupling, improved sensing techniques such as those developed by Schumann at Argonne National Laboratory and Thorensen at Institute for Numerical Analysis can substantially raise the permissible read-around.

The coupling can probably be reduced, for a given effective spot size, by special pulsing techniques; and the spot size itself has already been reduced by new gun designs. A decrease in effective spot size can also be obtained by operations external to the tube.

Figure 6 shows the shape of the visual spots for a 5CP1A tube focused at the center and with balanced deflection. The size of the spots are proportionate but enlarged for detail.

Hutter¹ has shown that an electrostatic deflection system produces both an additional radial and cylindrical focusing action. A further radial component is due to the mismatch between beam radius and faceplate curvature.

If the radial component is removed by correcting focus voltage at every position, it is possible to make every spot a circle. The correction required is a function only of the radius. Using this correction, we have improved spill two or three to one experimentally.

The cylindrical component due to one pair of plates can be compensated by the other pair. If the average potential of both plates is adjusted as well as focus, every spot can be made circular and as small as the minimum center spot.

---

¹R. G. E. Hutter, Advances in Electronics, (Academic Press, New York, N. Y., 1948), pp. 167-218.

40

Figure 1

IAS WILLIAMS MEMORY

40 standard 5CPI-A tubes

1024 bits per tube

Period: 24 μ sec.

$E_{b_2} = 1000$ v. $E_{b_3} = 2000$ v.

Amp. gain: 30 K

Dash signal ≈ 1 mv.

41

Figure 2

Figure 3

WILLIAMS PULSE CHAIN

*Total Chain (20μs)**Clock Pulse**S.D. and Strobe**Normal and Super Dot**Twitch Settling Delay (with termination pulse)**Normal and Super Dash**Action/Regen. Decision**Clear and Gate to get New Address*

43

Figure 4

MEMORY 5CPI-A CRT STATISTICS

July 1950 - July 1953

Tubes in Memory

10
11
15
4

40

Hours in Service

7500
5000
2500
< 2500

Total Removals

16
29
5
6

56

Reason

Flaws
Spill
Emission
Misc.

44

Figure 5

**FLAW and SPILL TESTS
on
250 PRODUCTION 5CPI-A's**

Worst flaw 0.8 normal or better..... 28%

OK flaw tubes with min. spill 100..... 70%

Overall yield..... 19%

Figure 6

5CPI-A VISUAL SPOT PLOT

No deflection correction

(Normal raster size shown dotted)

$E_{b_2} = \text{Hor. Av.} = \text{Vert. Av.}$

Focus at center

Figure 7

5CPI-A VISUAL SPOT PLOT

$E_{b_2} = \text{Hor. Av.} = \text{Vert. Av.}$

Focus = $f(\text{radius})$

(a)

Vert., Hor. Av. = $f(\text{deflection})$

Focus = $f(\text{radius})$

(b)

47 THE ORACLE MEMORY SYSTEM

by

R. J. Klein, Jr.

Oak Ridge National Laboratory
On loan to Argonne National Laboratory

INTRODUCTION

The preliminary design of the "ORACLE" called for a fast access memory of at least 1024 words of 40 binary digits each. The Williams electrostatic memory seemed to fulfill the general requirements if some method could be found to overcome its weaknesses. Considerable effort was expended investigating these weaknesses and possible remedies for them. As a result of these investigations a one-stage model was designed and tested. When it performed satisfactorily, a frame was designed which would distribute the pulses and voltages necessary for operation of 40 such stages in parallel. This frame and 50 memory stages were then built and tested.

INVESTIGATION OF PROBLEMS

From our own experience with the "Argonne Computer Memory" and the experience of others who had worked with electrostatic storage, it appeared that the principle difficulties with the Williams system were these:

1. Failure of certain storage locations because of imperfections on the storage tube screen.
2. Loss of information in the area around a location which was consulted a large number of times.
3. Frequent adjustment required to maintain optimum operation.
4. Extreme sensitivity of the system to external disturbances.

Each of these problems was investigated with the intention of eliminating it or at least minimizing its effects.

IMPURITIES

With the Williams system, the storage medium is the phosphor screen of a cathode-ray tube. The storage process depends upon the secondary emission properties of this screen. The presence of foreign material, even a very small particle, can prevent satisfactory information storage at the location of this particle.

It was supposed that if these particles were very small, satisfactory storage might be obtained if a modified Williams system were used - that is, some system which utilized a larger area of the storage surface per digit than the usual two dots. With this in mind, the circle dot, and de-focus-focus systems were tried.

The circle-dot system was more tolerant toward impurities but required pulsed, high frequency sine waves on the deflection lines with an exact phase relation between them. After preliminary experiments, it seemed this would be very difficult to produce on a 40-digit parallel memory.

The de-focus-focus system was also more tolerant toward impurities than the original Williams system, but it produces a great reduction in the number of times one location could be consulted without destroying its neighbors.

The most satisfactory approach seemed to be a double storage tube system with a best of two selection circuit. This is quite simple to accomplish as these impurities always produce a negative signal. Therefore, the output of both tubes can be fed into a single amplifier and the inspecting beams of each tube turned on separately. If the output from either tube is positive, a toggle circuit records that information and both tubes are restored to a positive state. Thus, the impurities are ignored unless they lie at the same location on both tubes. Interchanging one tube with some other stage would correct this difficulty.

This dual tube system of operation has the added advantage that it is possible to double the memory capacity if it is decided to code around the locations that have impurities. It is also possible that impurity-free tubes will be available in the near future so this double memory capacity would be available with no special coding considerations for impurities.

The decision was made, therefore, to use this dual tube system in the ORACLE memory, with the dot-dash type of display.

49

READ-AROUND RATIO

One of the most disagreeable features of the Williams storage system is its susceptibility to interaction between adjacent storage locations. As a result of this interaction, a dot, or binary "zero," may be erroneously changed to a dash, or "one," by the spill from a neighboring spot when the neighbor is read a large number of times before the dot is regenerated. The number of times one location can be consulted per regenerating of its neighbors without causing error is commonly called the read-around ratio.

This ratio should be as high as possible to prevent slowing down the computation to wait for the memory to be regenerated. It was estimated that a ratio of about 250 would be satisfactory for the ORACLE memory. No commercial tube was known to possess this read-around ratio so a testing program was initiated to test the many types of commercial cathode-ray tubes to ascertain whether or not one could be found to fill this specification.

The 3JP1 seemed to be the most satisfactory from these tests, so the memory system was designed around this tube. As a result of the very low deflection sensitivity of this tube, special circuits were required to drive the deflection busses, which serve 80 parallel tubes with the 200 volt deflection signals.

The driver on this first slide is capable of providing a pulse current of up to 300 ma either polarity, but it draws a steady state current of only 50 ma. It is entirely DC coupled. The 807 discharge tube provides the heavy current for a negative input step, while the left half of the 3C33 provides the current for a positive step. The right half of the 3C33 is used as a feedback sensing amplifier. The deflection bus, whose capacity to ground is 1500 μmf , moves through its maximum swing of 200 volts in about 1 μsec .

Four of these drivers are built on a single chassis, which is shown as the center chassis on the second slide. The chassis' above and below are the deflection adders which receive the digital address from the deflection register. These adders then convert this digital address to the corresponding analog address on the storage screens. These adder circuits utilize a resistor adding matrix which has proven extremely stable and noise free. No memory error has ever been traced to this deflection system.

The acceleration voltage on the storage tubes is quite high to improve the read-around ratio. At the present time it is set at 3kv total. The cathode of the storage tube is at -2800 volts while the second and third anodes are at +200 volts.

FREQUENCY OF ADJUSTMENT

In order to realize the highest possible read-around ratio a storage tube is capable of delivering, it is necessary to maintain its beam current constant. If this is not done the amplitude of the output signals drift with beam current so that it is necessary to set the recognition level for a "one" very low, otherwise some information would drop out when the beam current is low. Conversely, if the beam current is high, the partially refilled "zeros" will now be recognized as "ones." Beam current drift is quite severe on most cathode-ray tubes due to changes in gun temperature and filament voltage. For example, several three inch tubes had a change of 75% output pulse amplitude with a 10 % change in filament voltage. Amplifier gain change and high voltage power supply drift also tend to alter the amplitude of the output signals.

To correct this difficulty a circuit was arranged to sample the output from one storage location and hold it at a constant amplitude by changing the beam intensity. This intensity is not allowed to change between sampling cycles so that the output from all storage locations is constant.

The circuit on this third slide is called the automatic beam stabilization on ABS circuit. A differential amplifier compares the output of the sample spot with a reference voltage and produces a signal which tends to correct for any discrepancy. This correction pulse is then double stretched and applied to the cathode of the storage tube. If the pulse out of the sampled location is too large, the bias on the cathode-ray tube will be increased thus reducing the output signal, and vice versa.

The second stretching circuit has a 100 second-time constant so that no beam current change takes place between the sampling cycles, which are 20 milliseconds apart. This period is controlled by a free running multivibrator to keep a constant duty cycle on the stretching circuits.

The pulse which admits output of the test spot to the comparison amplifier must be wide enough to admit both the usual positive "one" signal and its beam turn-off pulse. This is necessary because extremely high beam currents, such as those encountered immediately after turning on the power, do not produce the usual storage phenomenon. The "one" signal under these conditions is actually negative but its "beam turn-off pulse" is very positive so this provides the correction signal until the intensity is low enough for the regular "one" signal to take control.

One difficulty appeared when this circuit was first tested. The sampling spot is a "one" because "ones" are more sensitive to beam current changes than "zeros." Usually the amplitude of this "one" is proportional to beam current but when a location nearby is consulted a number of times, the spill caused the test "one" signal to be higher than normal. This difficulty was overcome by arranging the control to regenerate the test spot immediately before sampling it.

51

A potentiometer is provided on the high voltage bleeder to allow setting this ABS circuit to its optimum operating range.

The operation of this ABS circuit has been quite satisfactory.

It appears from our experience that intensity adjustment should seldom if ever be required after the initial setting.

SENSITIVITY TO EXTERNAL DISTURBANCES

The signals from the pickup plate of the Williams storage tube are in the order of $1/2$ millivolt. Therefore, the amplifier must have a gain of about 40,000 to bring this signal up to the gating level. Any external disturbance such as noise on the power supply or capacitatively coupled pulses from the logical circuitry can cause signals out of the amplifier that might be confused with the information pulses from the storage tube. Mechanical shock is also likely to produce spurious signals by vibrating the elements of the low level amplifier stages.

On the ORACLE memory, regulators were installed on the voltage lines to reduce the noise from the power supplies. Rather elaborate shielding and grounding precautions were employed to prevent coupling of unwanted pulses into the amplifier system. The low frequency cut off of the amplifier was placed at 150 kc which somewhat reduces the microphonic problem. This also produces the rapid recovery time required for the time multiplex inspection scheme mentioned previously. To be certain that microphonics will not cause trouble, each memory stage is individually shock mounted.

If a high read-around ratio is expected of the memory, it is necessary that the electron beam of the storage tube be deflected to precisely the same position each time a spot is to be interrogated. All the voltages associated with the address generating circuits must therefore be accurately regulated. On the ORACLE memory the four most critical voltages have their regulators adjacent to the deflection adder chassis.

Quite often with electrostatic storage systems there are one or more stages which make random errors for no easily discernible reason. Various components may cause this condition, but it occasionally is a time consuming process to locate the defective part and replace it. For this reason the ORACLE memory was built with plug-in package units for each memory stage so that any stage can be replaced quickly with a spare unit.

The next slide shows the construction of the memory frame. The control bay is on the left. The pulse generating and distributing equipment and the deflection system are located in this bay. The center bay contains

52

filament transformers for the memory units and air ducts to channel the cooling air. The right hand bay contains three spare memory positions and a seven-inch monitor scope which can be switched to observe the information stored in any stage. The spare positions are used to keep replacement units warmed up and ready for service. These positions are also useful for running quietness tests on units suspected of picking up spurious information.

The next slide shows the monitor tube and neon light board. This neon light board is used for periodic quietness tests on the entire memory.

The next slide shows the rear of the memory frame with its ducts to supply voltages and pulses to the individual units. The pulse generating and gating circuits can be seen at the right of the control bay.

OPERATING EXPERIENCE

The ORACLE has been in operation approximately two months. Much of this time has been spent running tests to make the initial adjustment of parameters on the memory. From the performance to date, the memory seems to be functioning very close to design specifications. Random error failures have been few, and are expected to stop almost completely when the noisy tubes and poor solder joints are weeded out. The read-around ratio is approximately 250 with very few errors at ratios as high as 600. These figures are obtained with what has shown to be the worst possible conditions - namely, freshly written zeros around a "one" which is read "N" times. If the zeros are not freshly written over old "ones" the read-around ratio is considerably higher.

The automatic beam stabilization circuit holds the signal level constant so that no special memory warm-up time is required for satisfactory operation. Often codes are inserted within a minute after the B^+ voltages are applied.

Trouble shooting of individual stages is made quite easy by having the entire memory stage plug-in. The amplifier and logical circuitry are easily removable from this main plug-in chassis.

No storage tube selection was done before the tubes were inserted in the memory. However, about half of the tubes have been rejected because of poor read-around ratio or poor DC centering.

Figure 2

Figure 3

Figure 4

57

Figure 5

Figure 6

59

RELATIVE MERITS OF WILLIAMS MEMORY DISPLAYS

by

David Jacobsohn

Argonne National Laboratory
Lemont, Illinois

The purpose of the work to be described was to compare four methods of inspection of Williams Tube Memory systems with the possible aim of later modifying the memory of the AVIDAC to take whichever of the systems might prove most feasible. This work was done in order to test the relative read-around ratios of the four systems. It is hoped that if and when modifications are to be made that it can be done with the least disruption of our present equipment.

In order to understand the basic operation of electrostatic memory systems, it is necessary to first discuss the particular aspects of the secondary emission phenomena that are applicable.

Assuming an insulated target material whose secondary emission ratio $\delta_s = \frac{\text{secondary current}}{\text{primary current}} = \frac{i_s}{i_p}$, as shown in Figure 1, has the properties graphically illustrated, the following phenomenon will occur. If a target element is bombarded by a beam where the cathode-target potential, V_{kt} , is between V_{cr1} and V_{cr2} then $\delta_s > 1$ and the target will become more positive until the potential V_{cr2} at which $\delta_s = 1$, is reached. If V_{kt} is greater than V_{cr2} at time of bombardment, then $\delta_s < 1$ and the target will become more negative until the potential V_{cr2} , at which $\delta_s = 1$ is reached. Therefore, V_{cr2} is a stable equilibrium potential if $V_{kt} > V_{cr1}$. It is also apparent that if there is an electron cloud in the region of the target where the collector-target potential, V_{ct} , is positive, the target will have a greater affinity for the electrons until it is discharged to the potential $V_{cr} = 0$.

It is now possible, using the information of the previous discussion, to show how a cathode ray tube with a phosphor having the characteristics shown in the graph of Figure 1 can be used for binary storage.

First, consider what happens when a positive pulse is placed on the control grid of a cathode ray tube. If we assume that the phosphor on the inner surface of the tube has a secondary electron emission ratio of 1, the pulse in Figure 2a will appear on the face of the tube. The negative slope is due to the space charge of the beam. The level trough is obtained when

the beam current is in equilibrium with the secondary electrons being emitted by the phosphor. The positive slope is due to the space charge cutoff. If this charge configuration is capacity coupled into an amplifier with a resistance from the pickup plate to ground, the pulse form in Figure 2b appears. This is the picture of the square wave being differentiated by the above mentioned R. C. circuit.

If, however, we have a phosphor whose secondary ratio is described by the graph in Figure 1, when an electron beam strikes a $V_{ct} = 0$ surface, more secondaries are emitted than there are primaries striking the surface. Therefore, a positive potential builds up. As this potential builds up to V_{cr2} the secondary electron ratio decreases, when it reaches unity, the equilibrium condition occurs. The pulse form in Figure 2c appears from the amplifier. The first negative dip is due to the space charge, the first positive slope and peak is due to secondary emission, the drop to the base line is due to the RC network, and the last pulse is due to beam turnoff. If we again bombard this spot before the positive charge has chance to dissipate, it will appear as Figure 2b.

The last case of interest (Figure 2d) occurs when we bombard a position adjacent to a position previously bombarded. In this case, the initial portion of the output is similar to that of Figure 3c; however, this portion is followed by a negative dip due to the secondaries emitted from the presently bombarded position falling into the adjacent positive well. This will occur so long as the positive well is of greater potential than the collector. It should be noted that positive charging is a fast phenomenon, whereas the negative filling is relatively slow. If this second position is bombarded sufficiently long, the first position will be completely discharged.

It should now be noted that at various times there exists potential differences between pulses of the types in Figure 2, representing the bombardment of initial or adjacent spots.

By inspecting the potentials at one or more of these times, it is possible to determine the previous state of the screen at the present position and time of inspection. It is also possible to generate potential configuration of the types in Figures 2b and 2d, as well as to change either type to the other.

Because of the ability to write, read, and alter these two possible configurations, this system lends itself to a binary memory storage system.

Initially, confining the discussion to the classical Williams inspection system which inspects at the time of the secondary electron peak, we find that the inspection of a position previously bombarded reinforces the information that was there and the inspection of an adjacent position destroys the contents of the position and requires a regeneration by bombarding the

61

adjacent position. Information not requiring regeneration after inspection shall be called "zero," information requiring regeneration after inspection shall be called "one." The process of moving the beam to the adjacent position is called twitching and the distance is the twitch distance.

It should now be mentioned that since the phosphor of the cathode ray tube is not a perfect insulator, the charges have a tendency to leak off. Loss of ability to read information from the face of the tube due to leakage occurs in a matter of seconds. It is, therefore, necessary to regenerate the information during this period; this is done about every 0.02 second.

Information is stored on the tube by pairs of points in the form of a square lattice on the face of the cathode ray tube. In the particular test system to be discussed, a 32 x 32 raster was used to store 1,024 point pairs of information.

Having so far described the nature of the electrostatic storage phenomenon, it is now appropriate to describe the equipment used to utilize this phenomenon for actual machine storage. Figure 3 shows the equipment necessary to regenerate and write new information on the Williams tube in the classical system, as previously described.

The equipment used for this work was the Argonne Memory Test Rack. This equipment consists of a standard Williams circuit consisting of a deflection generator producing a 32 x 32 raster and an amplifier and discriminator circuit that will permit the writing and regeneration of the information to be stored on the memory tube. The deflection generator, however, is of such nature as to permit upon completion of the regeneration of a complete raster an interruption in the regeneration cycle followed by repeated interrogation of a position on a superimposed sub-raster for a predetermined number of times, then to resume the regenerating process for another complete raster, and then to go to the next position, and so on.

Since the purpose of this work was to find the best system relative to use on the AVIDAC, the operating conditions were such as to simulate best those now in use on the AVIDAC. It is realized that these conditions may not be optimum for any of the systems investigated; however, it is believed that reasonable figures of relative merit can be obtained.

In testing the four systems the conditions were held constant unless otherwise stated:

1. The second and third anode voltages were at 1400 with respect to the cathode.
2. An amplifier with a gain of 30 thousand and a band pass from several kilocycles to 3.5 megacycles was used. The low frequency cut-off for the dot-slide inspection system 1/2 megacycle.

62

3. Back-to-back twitch was used except in the refill system. Back-to-back twitch means alternate columns are twitched to the right and left. In each case, both regular and back-to-back twitch were tried and the better system used.
4. The over-all pattern size was approximately 3 inches square.
5. All tubes used were Dumont 5CP1A's.

The equipment for each test was set up in the following manner. One of the selected tubes was installed in the test rack, then the pulse durations and timing, and the twitch distances were adjusted for optimum operation of the tube. After these adjustments were made, they were held constant throughout the testing of the particular system. The only adjustments made to compensate for individual tubes were the intensity and focus of the beam and the discrimination level.

Each tube was then tested for marginal operation of read-around with both 8 x 8 and 16 x 16 sub-rasters superimposed upon the 32 x 32 raster. In all cases the tubes had to be capable of writing and remembering ones and zeros in all positions both with and without read-around. Impurities were circumvented or ignored.

Having so far treated the general aspects of electrostatic storage systems, it is now possible to discuss the specific systems tested and the results obtained.

The classical Williams Inspection System operates by the inspection of the initial pulses and discriminating between the trough of the upper illustrated pulse (Figure 4) and the peak of the lower one. If the peak occurs, the adjacent spot, position II, must be bombarded in order to restore position I to its original discharged state. The discharging is expedited by not twitching the beam until after it is on, as this breaks the charge barrier that interferes with the discharging of position I.

The reading pulse in this system is short as it only investigates the information held at the point of investigation. The regeneration pulse is long because it must discharge the adjacent position and this is a long time phenomenon.

Read-around errors in this system occur due to the refilling of position I with stray electrons from nearby storage positions. Zeros go to ones.

The reading pulse is of the order of one microsecond, the regeneration pulse is of the order of six.

63

The second system treated was that of Refill Inspection (Figure 5). In this system, instead of investigating the first pulse due to secondary electron emission, the latter phenomenon of the negative refill trough is investigated. (In these drawings the flat bottom is due to viewing at the input of a ground grid amplifier.) Regeneration is accomplished by redigging the well in position II. The regeneration process is short since only the well need be dug and it is not necessary to refill position I. The reading pulse must be long, however, both to allow for the later inspection time and to completely refill the well in position II if a zero is to be written over a one. The reading pulse is of the order of six microseconds, the regeneration two. At optimum adjustment read-around errors occurred in both directions.

Another system operating in a similar manner to the original Williams system is the Dot-Slide Inspection system (Figure 6). In this system the beam is turned on to position I where information is stored and then twitched to position II where the opposite information exists. By using an amplifier with a short time constant, it is possible to sum the information of the two spots by the amplifier overshoot. The output is inspected immediately after twitching. Both the reading and the regeneration pulses should be of moderately long duration in order to guarantee the refilling of the opposite position. The reading pulse was of approximately two microseconds and the regeneration of five. Read-around errors occurred in both directions.

In the final system to be discussed (Figure 7), that is the double-inspection system, both positions are individually inspected in order to obtain the maximum discrimination between ones and zeros. The first reading pulse being of short duration, as it only inspects the information in position I, the second reading pulse being of long duration as it not only has to inspect position II but must also restore position I to the discharged state. The final and regeneration pulse must be of long duration as it must restore both positions to their original state when it is called for. In this system the first inspection is used to set the D. C. level to the discriminator, the second inspection is when the actual discrimination takes place. This system should to some degree accommodate for local D. C. drift of the cathode ray tube screen potential. The initial reading pulse is about one microsecond long, the other two pulses are about four microseconds long. Read-around errors with this system occur in both directions.

Before stating the read-around results of the various systems, it seems proper to give some incompletely explored observations. The last three systems discussed appear in varying degrees to be less impurity sensitive. Impurities with these three systems caused failure to hold information in both directions, depending on location. Conditions that would cause one type of error with one system would not be present with another system. It is necessary to re-evaluate standards of operation when changing from system.

Table I indicates the relative read-around results.

64
TABLE I

Relative Read Around Ratios of Five 5CPIA CRT's
with Various Inspection Systems

Inspection System	Tube No.	1	2	3	4	5	Average
Williams		1	1	1	1	1	1
Refill		2	2.3	4	2	2.7	2.7
Dot-Slide		2.5	1.7	4	4	3.0	3.0
Double		3	3.4	5.3	4	5.3	4.2

Data normalized with Williams Inspection System equal to 1.

Reference:

An Improved Cathode Ray Tube Storage System
R. Thorensen
NBS Report #2275

Dot- Slide Inspection System
Work done by members of the ILLIAC staff - University of Illinois

Double-Inspection System
Robert Schumann - Toronto Meeting of ACM

65

$V_{KT} = \text{CATHODE - TARGET POTENTIAL}$
 $V_C = \text{CATHODE - COLLECTOR POTENTIAL}$
 $V_{CR1}, V_{CR2}, \text{POTENTIAL WHERE } \delta_s = 1$
 $V_{CT} = \text{COLLECTOR - TARGET POTENTIAL}$

V_C
 $\delta_s \text{ vs } V_{KT}, V_{CT}$

FIGURE 1

66

FIGURE 2

TIMING CHART

FIGURE 3

Figure 4

Williams Inspection System

57

Figure 5

Refill Inspection System

Dot-Slide Inspection System

71

Figure 7

Double Inspection System

CS
CS

THE ILLIAC MEMORY

by

J. M. Wier

University of Illinois
Urbana, Illinois

Williams memory systems have been extensively described since the first paper by Williams and Kilburn appeared in 1949.^{1,2,3,4} For that reason only a short summary of the quantities usually of interest will be presented in this paper. The differences of operation unique to the ILLIAC memory will be more extensively treated.

Considerable effort has been expended at the University of Illinois in experiments on various Williams type memory systems in an attempt to improve the over-all system. One of the experiments involved a change in the time at which the binary information is sensed. This was tried at the indirect suggestion of Dr. Thorenson of SWAC. The system devised differs from that described by Thorenson⁵ in several respects. It is that system which will be described here. As considerable operating experience has been obtained using this system since its installation in March, enough data is available for a fairly accurate judgment of its worth.

---

¹Williams, F. C. and Kilburn, T.: "A Storage System for Use with Binary Digital Computing Machines," Proceedings I.E.E., 1949, Vol. 96, Part III, p. 81.

²Eckert, J. P., Jr., Lukoff, H., and Smoliar, G.: "A Dynamically Regenerated Electrostatic Memory System," Proceedings of the Institute of Radio Engineers, 1951, Vol. 39, p. 900.

³Holt, A. W., and Wright, J. H.: "Progress Report on Electrostatic Storage" (National Bureau of Standards Report No. 1082, 1951).

⁴Logue, J. C., Brennemann, A. E., Koelsch, A. C.: "Operating Experience in the Design and Operation of a Large Scale Electrostatic Memory," Convention Record of the I.R.E., Part 7, Electronic Computers, 1953.

⁵Thorenson, R.: "An Improved Cathode Ray Tube Storage System" (National Bureau of Standards Report No. 2275, 1953).

73

The ILLIAC memory contains 1024 words of 40 binary digits. These digits are written in parallel on 40 standard 3KP1 cathode ray tubes, the raster size being about 1.75 inches square. These tubes are operated with 2000 volts accelerating potential. Each cathode ray tube is operated with separate controls for adjusting focus, astigmatism and intensity. It is perhaps noteworthy that after adjusting tubes to best conditions of operation, the astigmatism voltage is separated from 5 to 20 volts from the mean deflection plate voltage.

The regeneration chassis is of the plug-in type. It contains a four tube amplifier with a band width from 60 to 500 k.c. It is a straightforward condenser coupled circuit with no feedback other than unbypassed cathodes. The input resistor is 15K. The gain to a positive pulse of 2μ sec. duration is about 70,000. Since the coupling circuits are only of 5μ seconds time constant, considerable overshoot exists and recovery time is short. It will later be seen that the overshoot actually plays some useful part in the operation of the memory regeneration system.

The basic regeneration period is 18μ sec. and the system operates on a two dot basis. Although the beam is not turned off when it is moved the move is made as fast as the circuits used allow.

Slide No. 1 shows the manner in which the raster is arranged. The half address stagger of every other row gives a positioning of spots such that every address is surrounded by six other addresses on the corners of an equilateral hexagon. This was done in order to obtain a better average separation. The twitch is vertical and is about 12 mils in length.

Slide No. 2 shows the chassis in position on a test rack. The arrangement is the same as in the ILLIAC but is in the open for ease of repair of faulty chassis. In the machine these chassis are arranged in two rows of 20, one above the other. Each one is separated by about 0.75 inch from its neighbor.

Slide 3 shows the way the tube is mounted within the shield. The tube is plugged into its socket which is permanently mounted on the memory frame. No CRT voltage connections are brought through any plugs but are wired permanently to circuits mounted on the frame.

Slide 4 is a view of the rear of the chassis. All high voltage connections to the tube are heavily bypassed. All cathodes are fastened together and bypassed with one 4 microfarad condenser per five memory positions. 0.5 microfarad condensers bypass both the focus electrode and intensity control point of each position to the cathode. The second anode is bypassed to ground with an RC filter of 0.25 second time constant.

DESCRIPTION OF ILLIAC SYSTEM

The ILLIAC operates on a different pulse scheme than the normal Williams system. In the process of regeneration the beam is turned on at the first of two storage spots at a given address. It is kept on this spot for 1μ second and then abruptly moved to the twitch spot approximately 12 mils distant, with the beam still remaining on. It remains at the second spot for 2.5μ seconds. About one microsecond after moving the beam to the second spot the output signalled is sampled with a strobe pulse of 0.8μ second duration. This time corresponds to the output of a positive pulse of about 40 volts peak for the storage of a "0" and a signal considerably below zero volts for the storage of a "1." The decision level is at 10 volts above the zero volt level and is made on a purely amplitude basis. At this time the beam is again moved back to the original spot. It is turned on there for about 3μ seconds if a "0" is to be written and is turned off before going there if "1" is to be written. Suitable circuitry is also included in a very standard manner for writing in and reading out.

Slide 5 shows the turnon pulses applied to the cathode ray tube grid. They are 50 volts in amplitude with a rise time of about 0.4μ sec. Both extremes are pinned. Essentially no difficulty has been experienced in excessive grid to output coupling. The two pulses shown are those associated with the sensing operation and the writing operation. Each is slightly over three microseconds in duration.

Slide 6 shows the output signals existing for dots and dashes. The sensing takes place during the second positive peak. It is about 3.5 microseconds from the start of the pulses until the second positive pulse goes back down. The first rising positive pulse corresponds to a "1" and is about 50 volts high. It results from the normal Williams phenomena of turning the beam on at a relatively uncharged point. The spot is charged positively by the secondary emission from the spot until equilibrium is reached. When the beam is moved to the second spot there will be a change in the electron cloud about the screen since the charging is no longer taking place. Further, there will be a certain amount of digging done at the second spot to repair the discharging done by the sensing beam. This is a relatively minor amount. Finally there will be the flow of electrons from the second spot back to the first spot to refill that spot. It is to be observed that this leaves a net negative flow of electrons to the screen for a time indicated by the negative portion of the "1" signal. It is also to be noted that a small contribution will be made due to the overshoot of the amplifier after the initial positive signal.

The "0" output is the one which goes negative when the beam is first turned on at the first spot. This is the usual result caused by the arrival of a space charge at the screen after being smoothed out by the amplifier and pickup screen circuits. When the beam is moved from the first spot to the second, the second one will be originally uncharged and so a large charging action occurs to recharge that spot just as the first spot was charged in the

75-

case of the "1" signal. This spot is driven to this relatively uncharged state by the second pulse in the write-in part of the cycle just as is normally done to discharge the first spot for a "1" signal. A further contribution to this positive peak will be made by the change in the space charge which occurs as a result of moving from an equilibrium charged spot to one which requires further charging, an action somewhat inverse to the one which acts toward causing the large negative second pulse in the case of a "1."

EFFECT ON READ-AROUND RATIO AND FLAWS

This type of storage has a few interesting properties which make it behave differently from the normal Williams system with respect to read-around ratio and flaws. Since the information available at the time the strobing of the output is done is a function of the information stored at both spots, one might expect a more balanced behavior than is found in most Williams systems. This is in fact the case. The usual low secondary emission spots cause failures in both directions, depending upon whether they are beneath the first storage spot or the second one. Further, read-around ratio failures which with the normal Williams system always cause "0" to "1" failures, may fairly readily be equalized in this system so that failures in both directions occur with equal probability. This latter characteristic has been the one which has been the most exploited with the ILLIAC system. Although the failures occur in both directions, the minimum read-around ratio at which the first failures begin to occur is raised by a factor which is from two to four. Using this system, the point at which the first failures occur under the worst conditions found on the ILLIAC has been moved upwards by a factor from two to three.

The read-around ratio test used on the ILLIAC is conducted in the following manner. The routine for performing the test is placed in half of the memory. The other half is cleared to "0"'s which corresponds in this system to the longest beam on time. Then an address is selected. First, eight addresses surrounding this one are filled with "1"'s; then the central address is bombarded n times with a "0" as rapidly as possible. The surrounding addresses are then regenerated and the center spot again bombarded n times with a "0." This process prevents random regenerations in the neighborhood from affecting the process adversely. Then the surrounding eight spots are checked for "1" to "0" failures on all positions. If failures occur, suitable indicating information is printed out; if no failures occur, the test is repeated at the point with the central spot surrounded by "0"'s, the central spot bombarded with "0"'s and a check is made for "0" to "1" failures, failures again being printed out in a characteristic manner if failures occur. The remaining spots in half of the memory are tested in this way and then the routine is transferred to the opposite half and the remaining part of the memory tested for read-around ratio in this way. The information printed out gives the address, tube number, direction of failure, and testing read-around ratio number n at which the failure occurred.

76

Prior to the change to this system the read-around ratio was such on the ILLIAC as to begin to indicate failures at a value for n of about 25 on one of the 40,960 storage addresses. After the change this number rose abruptly to about 70 and since this has been variable between 60 and 80 with normal operating conditions depending on the state of the tubes and degree of adjustment of the memory.

It should perhaps be mentioned that since flaws at either spot can now cause failures, the probability of landing on a spot which will cause a flaw failure may be somewhat higher if flaws are present on the storage surface but operation with flaws is hectic at best even with the normal Williams system.

In any case it is felt that the change has been worthwhile since no changes in the regeneration chassis was required in the transition and the time needed to make it was only about 16 hours.

The research and development involved in the evolution of the ILLIAC system was supported in part by the Office of Naval Research.

Figure 1

Figure 2

Figure 3

Figure 4

81

Figure 5

82

Figure 6

83

DESIGN AND MANUFACTURING CONSIDERATIONS OF THE OSCILLOGRAPH TYPE ELECTROSTATIC STORAGE TUBE

by

E. M. Smith

Cathode-ray Tube Development Group
RCA Victor Division
Lancaster, Pennsylvania

SUMMARY*

Fundamental requirements of oscillograph-type electrostatic storage tubes are considered, especially as applied to tube design and production problems. Tubes made for computer operation are analyzed to show the compromises in design for optimum performance in storage systems.

First, a brief, general discussion of the basic operation of charge-controlled storage tubes is presented, and the fundamental principles of oscillograph-type electrostatic storage tubes are reviewed.

The inter-relation of tube design parameters and performance characteristics, such as beam current and masking, magnification, deflection defocusing, life, spot size and current distribution, crossover diameter, grid drive characteristics, and operating voltage is presented. Combinations of those factors which satisfy storage tube requirements are considered.

Manufacturing problems are separated into two general groups: (1) those concerning screen uniformity, and (2) those related to the electron gun. The principal causes of secondary emission blemishes and electron gun defects are mentioned.

---

*See page 288 for complete text.

84

A MYRIABIT MAGNETIC-CORE MATRIX MEMORY*

by

Jan A. Rajchman

RCA Laboratories,
Princeton, New JerseySUMMARY

A coincident-current magnetic-core matrix memory was built giving random access to any of 10,000 bits in a few μ sec. The memory cores, of a diameter less than $1/16$ ", are made on an automatic press from especially synthesized magnetic ceramic material with rectangular hysteresis loop. Automatic test techniques permit rapid sorting to insure uniformity of the magnetic properties of the cores. The memory array is driven by two ten-by-ten dc biased magnetic switches which reduce the number of address-selecting access vacuum tube drivers to 40 and the required driving current to the order of a few hundred milliamperes. The read-out signal originating from a winding linking all 10,000 cores is freed from the disturbing contributions of partially excited cores on the selected lines by integration over a period of time in which these cores are subject to equal and opposite magnetizing forces. The principles and techniques evolved for this experimental model demonstrate the practicability of an extremely reliable fast random access memory with a capacity of hundreds of thousands of bits, ushering in a new era in storage devices for information handling systems.

INTRODUCTION

Magnetic cores with rectangular hysteresis loops have the inherent storage and switching capabilities required from an element of a random access matrix memory based on current coincidence. This was recognized independently by J. W. Forrester¹ and the author, who initiated the early

---

*Reprinted from Proceedings of the I. R. E. October, 1953, permission of the publisher. First presented at the Symposium.

¹J. W. Forrester, "Digital information storage in three dimensions using magnetic cores," Jour. Appl. Phys., vol. 22, pp. 44-48; January, 1951.

85

efforts to develop such a memory at the Massachusetts Institute of Technology and RCA Laboratories, respectively. Last year an operating unit with the relatively modest capacity of 256 bits was made to operate successfully,² and it was predicted² that extension to capacities with thousands of bits was possible in the near future.

This paper describes an operating experimental memory with a capacity of 10,000 bits, named Myriabit. (Myria means 10,000 or many, according to Webster.) This relatively large capacity was attained by basic changes in the cores themselves and in the systems used for driving and interrogating the memory.

Most of the initial work in this field was done with metallic cores which were the first materials to exhibit rectangular hysteresis loops. Cores obtained by winding 5 wraps of $\frac{1}{8}$ mil Mo Permalloy 4-79 on ceramic bobbins were used in the smaller memory built last year.² These cores have relatively long switch-over times of about five microseconds, when used in the coincident memory, and they are intrinsically expensive to fabricate. Owning mostly to their low electrical conductivity, the ferromagnetic spinel materials, often called "ferrites," have much faster switch-over times. Core fabrication, consisting of molding from a powder-like material, is inherently inexpensive. The potentialities of these materials prompted research at General Ceramics Co. and RCA Laboratories towards obtaining sufficient loop rectangularity in materials which otherwise were ideal. Early in 1952 some materials³ were promising, and in the summer of that year both organizations synthesized a material for practical use. The properties of the G.C. material, used by the Digital Computer Laboratory of M.I.T., were described recently.⁴ The nature and actual preparation of the RCA materials, which are "manganese magnesium ferrites," are to be published.⁵

The remarkable success of these ferromagnetic spinel materials offered a challenge to build a memory of a very large capacity, since it became possible to obtain the cores cheaply and in great quantities. Of course, this required an automatic molding machine. We initiated a development to that effect at the Stokes Machine Co. which modified one of its machines for this use.⁶ The necessary machine and tooling work will not be described as it is fairly conventional.

---

²Jan Rajchman, "Static magnetic matrix memory and switching circuits," RCA Review, vol. XIII, pp. 183-201; June, 1952. (Also presented at I.R.E. National Convention, N. Y., March 5, 1952)

³General Ceramic material body 666, RCA Laboratories Body 160B.

⁴D. R. Brown and E. Albers-Shoenberg, "Ferrites speed digital computers," Electronics, pp. 146-149; April, 1953.

⁵I. J. Hegyi, "Ferromagnetic spinel with rectangular hysteresis loops," to be published shortly in Jour. of Appl. Phys.

⁶"A3 Eureka Tablet Machine," F. J. Stokes Co.

56

The operation of a large memory array poses two main problems. driving and obtaining a reading signal free of masking disturbances. The solutions to these problems found last year² were adequate for an array of 256 cores and would be useful for larger ones. We chose deliberately an array size large enough to exceed the limitations of these solutions and to force research into methods applicable to a size one order of magnitude greater. This determined the choice of 10,000 bits. This storage capacity had the incidental advantages of being greater than any obtained in a single electrostatic storage tube and of demonstrating the flexibility of magnetic switching by being an arbitrarily chosen integral power of ten.

The driving magnetic switches for the memory array were worked out to be themselves matrix switches arrays of ten-by-ten cores biased by a dc current. These switches lend themselves to a simple construction in which the limiting effects of the earlier switches are essentially eliminated.

A solution to the read-out problem was found which is applicable to practically any size, even when the array comprises cores with substantial departures from loop rectangularity and considerable variations in the magnetic characteristics. The read-out signal, originating from a winding linking all 10,000 array cores, is essentially freed from the disturbing contributions of partially excited cores on the selected lines by integration over a period of time in which these cores are subject to equal and opposite magnetizing forces.

We concentrated on the building of a memory with access to a single core in a large array because all essentials of switching into and reading from a memory having access to many cores in parallel are encountered in this simpler case. The principles evolved are indeed extendable to the parallel-access type for which the core memory is eminently suitable.

After a review of the principle of the memory itself this paper describes (1) the construction, core testing and characteristics of the memory array, (2) the principles, construction and operation of the driving switches, (3) the reading method, (4) the programming and the functions of the associated circuits. It concludes by predicting a great future for the magnetic matrix memory, which approaches the ideal rapid random access memory closer than any yet constructed.

THE PRINCIPLE OF THE MEMORY

Consider an array of magnetic cores arranged in rows and columns. Each core is linked by three windings. Two of the windings are connected in series by rows and columns, and a third is a common series winding. Single turn windings consisting of straight conductors linking the cores are shown on Fig. 1, as this is the most convenient way of constructing the array.

87

All cores are magnetized to one or the other of the two remanent inductions represented by points P and N on the idealized hysteresis loop of Fig. 1. In this state, the array is holding information previously stored. In order to set any given core to the desired remanence, P or N, pulses of current in one or the other direction are simultaneously applied to the row and column of the core. The amplitude of these currents is so adjusted that their additive, i.e., doubling effects, produce sufficient magnetomotive force to exceed the knee of the loop for the selected core at the intersection of the lines but insufficient to exceed the knee for the other cores in the selected lines where they act singly. Consequently, the selected core will be magnetized in the desired direction while all other cores in the matrix will remain unaffected.

For the read-out, consider the voltage induced in the common winding when read-in currents of a standard direction, e.g., toward P, are sent through the lines of the core to be interrogated. Ideally, the flux can change only in that core, so that the existence of a voltage indicates that the core switched from N to P, while absence of voltage indicates that it was originally at P. This type of read-out signal was used successfully in previous memories.^{1, 2, 4} In order to eliminate the contributions of other cores on the selected lines which actually occur because of imperfections of loop squareness, the interrogating currents in the present memory are pulses of first the one and then the other polarity, as will be explained. In any case the state of the interrogated core is ascertained by changing it, so that a restoring cycle controlled by the read-out signal is necessary, otherwise read-out would be destructive.

THE CONSTRUCTION OF THE MEMORY ARRAY

The 10,000 cores of the myriabit memory were molded from "manganese magnesium ferrite"⁵ in the shape of tiny toroids, with 0.054" O.D., 0.034" I.D. and 0.016" height. These cores, after individual tests, were wired into a plane array in a manner schematized on Fig. 1 and shown on Fig. 2, following page. The size of these cores, which is just about that of the printed letter o in the word core, can be appreciated also by comparison to a match head shown in the detailed Fig. 3 on page 104. The row and column wires, #32 double formex total O.D. of 0.009", pass straight through the cores positioned at 45° and serve as their supports. These line wires 0.080" apart, with addition marker spaces between decades, are stretched between tiny terminals supported on a bakelite frame of about 10" x 10". The reading winding is a series connection of 100 smaller windings each linking the 100 cores of squares 10 by 10. These consist of bare 0.006" wire running back and forth diagonally through the cores.

The miniature size of the cores, the smallest conveniently handled without special microtooling, was chosen to minimize the current and power driving requirements rather than to maximize the bit-per-unit volume packing density. Even so, the current for switching the core, at optimum

87

discrimination in a two-to-one regime, is 1.1 ampere-turns, and the energy is about 5×10^{-8} joule. The brute force wiring of the myriabit array taxed the patience of a person for six weeks. Simple semi-automatic wiring techniques would reduce this time to hours.

AUTOMATIC CORE TESTING

The memory array must not comprise a single core failing to give some specified minimum performance. To insure this, it is absolutely necessary to test every core before it is wired in, because it is practically impossible to control the fabrication of the cores to the absolute degree required even though the average yield may be very high. Manual testing of the cores far exceeds the effort required to make them on automatic molding machines, hence the need for automatic testing.

The principle of an automatic core tester which we devised is shown on Fig. 4, page 105. The cores are poured into the dish of a "syntron elevator." The vertical and angular vibrations of the dish force the cores up a helical track on its periphery. The top core is held by a slotted stop (the only part of the vibrator visible on the left of the figure) and is picked up by one of the radial metal pins of a rotating wheel. Test trains of current pulses obtained through a transformer are sent through the pin as it passes on slide contacts. The resulting induced voltages, characterizing the core, are taken by a second set of contacts and are analyzed according to specified criteria by electronic circuitry which does or does not produce an acceptance signal. This signal activates a relay which positions a slotted arm so as to "wipe off" the accepted core. The unacceptable cores pass the location of the unactivated wiping arm and fall in a reject container. The experimental model of the tester operates at the rate of 40 to 60 cores a minute which is comparable to the rate of the core molding machine. Unfortunately, the tester was not available in time so the cores for the Myriabit were tested by tedious manual methods.

In speculating about the relative merits of memories with discrete storing elements and those with a continuous storing medium, one often considers the latter preferable because no tedious fabrication of a large number of elements is required. But because in most applications of digital storage no single inoperative element can be tolerated, an imperfection in the continuous medium renders the entire device useless. As the imperfections can be detected only by the operation of the completed memory, the possibility of pretesting every element to gain certainty of its operation before assembly shifts the preference to memories with discrete elements. This is particularly true when both manufacture and testing of the elements are automatic.

THE CHARACTERISTICS OF THE MEMORY CORE

The actual properties of the single core, which necessarily depart from ideal rectangularity, ultimately determine the performance of the memory array in any system.⁷ A typical "manganese magnesium ferromagnetic spinel" core of 0.054" O.D., 0.034" I.D. and 0.016" height used in the myriabit has been measured by magnetizing it by various current pulse trains and observing the resulting induced voltages on an oscilloscope synchronized so as to superimpose the traces, as shown on page 106, Fig. 5. The amplitude of the current pulses, adjusted to obtain optimum ratio of full-excited to half-excited responses, turned out to be 1.1 ampere-turns and is referred to P or N depending on the polarity. The voltages resulting from the schedules $P \frac{N}{2} \frac{N}{2} \frac{P}{2} N \frac{P}{2}$ and $P \frac{N}{2} \frac{N}{2} P N$ show the high degree of rectangularity of the hysteresis loop by the smallness of the traces due to half excitations which had to be expanded ten times to be measurable. The hysteresis loops of Fig. 6, page 107, were constructed by integrating the voltage wave shapes of Fig. 5 (including the part due to current decay) and other similar ones obtained for the different pulse schedules, such as $N \frac{P}{2} \frac{N}{2} P$, $P \frac{N}{2} \frac{N}{2} \frac{P}{2} N$ and $P P N$. These minor loops, drawn to an expanded scale, show the essential core behavior, although good accuracy was hard to obtain in spite of averaging over several curves and the consideration of obvious symmetries.

We found it convenient to describe the state of the core by the series of pulses to which it was submitted. Following this convention, the selected core in the memory array is at P after the simultaneous $\frac{P}{2}$ excitations of its selecting lines, since it is subjected to a full P excitation. When these are followed by two $\frac{N}{2}$ excitations of its selecting lines occurring in succession - as in a write-in system to be described - the selected core is brought from P to $P \frac{N}{2}$ and then from $P \frac{N}{2}$ to $P \frac{N}{2} \frac{N}{2}$. When the $\frac{N}{2}$ excitations are simultaneous, the core is brought to N, since the state PN is equivalent to N. The state of the core at the peak of a current pulse is denoted by putting in parentheses the last letter referring to the current pulse yet untruncated as shown on Fig. 6 for the magnetizations of ± 0.55 and ± 1.1 amp. turns. The rate of current rise and the losses⁸ in the cores determine the length of the

---

⁷William N. Papian, "A coincident-current magnetic memory cell for the storage of digital information," PROC. I.R.E., vol. 40, pp. 475-478; April 1952. (Also "A coincident-current magnetic memory unit," Master's thesis, E.E. Dept., M.I.T.; August, 1951).

⁸The relaxation losses are dominant in ferrospinel materials.

voltage pulses and their peak values, the latter not corresponding necessarily to the flux changes of the hysteresis loops such as those of Fig. 6. For very fast rising pulses ($0.05 \mu\text{sec}$), the core switches over in about $0.3 \mu\text{sec}$ and the voltage pulse exhibits the familiar double maxima observed with metal cores.⁹ The slower current pulses of Fig. 5, rising in 0.2 to $0.3 \mu\text{sec}$, are more significant for memory applications, since they simulate more closely the pulses furnished by cores of the magnetic switch described below. The turn-over time is now about $0.6 \mu\text{sec}$, almost half of it being due to the current pulse rise.

When access is made to some core(s) of a line many times in succession, the remanent state of the other cores, suffering many half excitations, must remain near P or N, so as to retain the stored information. With a magnetic switch drive, every access necessarily produces a $\frac{P}{2}$ excitation followed by an $\frac{N}{2}$ one, and cumulative demagnetizations tend to be prevented, as was pointed out previously.² This may be verified on Fig. 6. The loop $P \frac{N}{2} \frac{N}{2}$ to $P \frac{N}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2}$ is closed within measurable accuracy. The loop N to $N \frac{P}{2} \frac{N}{2}$ is not closed, but it is found that the loop from $N \frac{P}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2}$ to $N \frac{P}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2}$ is closed, so that on the third access the core has essentially reached its asymptotic minor loop. We found, incidentally, that the schedule $N \frac{P}{2} \frac{P}{2} \frac{P}{2} \frac{P}{2} \dots$ leads also to an asymptotic value $N \left(\frac{P}{2} \right)_{\infty}$ near the value $N \frac{P}{2} \frac{P}{2}$ for this core. On the other hand, for cores made of a less "rectangular" material we found that $N \left(\frac{P}{2} \right)_{\infty}$ was essentially at the origin ($B = H = 0$), while $N \left(\frac{PN}{2} \right)_{\infty}$ was only somewhat lower than P. This shows the virtue of the demagnetization - remagnetization cycle inherent with transformer drive (but not with unscheduled direct tube drive).

SWITCHING THE MEMORY ARRAY

The problem of switching into the n cores of the memory array is reduced to that of switching into $2\sqrt{n}$ channels because the cores themselves provide double coincidence switching. Because both polarities of current are required in each channel, the number of drive tubes or of any other unidirectional electronic device is double. Although $4\sqrt{n}$ can be relatively much smaller than n it is still very large: for example, 400 tubes would be required to drive a 10,000 core array.

---

⁹J. Rajchman, "Static magnetic matrix memory and switching circuits," RCA Review, vol. XIII, pp. 183-201, Fig. 6; June, 1952.

91

"Commutator" magnetic switches were designed last year^{2, 10} in which one out of 2^k channels is selected by k input push-pull channels. When two such switches drive the memory array, the number of input channels is reduced to the minimum possible, i.e. $\log_2 n$, and at the same time the transformer action of the switch cores can provide a better matching of the driving tube impedance. Two commutator switches with 4 inputs and 16 outputs each were used to drive the 256 memory core array,² requiring in all only 16 relatively small driving tubes (8 push-pull channels).

These commutator switches are a most elegant solution for the binary to one-out-of-many conversion or "decoder." However, as the number of binary positions increases, there is a greater number of combinatorially required turns on each core to establish the k -way coincidence, and this sets some practical limitations to the size. The construction becomes a great deal more complicated, and it is more difficult to put sufficient number of turns on the driving primaries to properly match the tube impedances. This is true even when the window of the cores is full of copper because an increasingly large proportion of that copper carries no useful current when the core is switched over. Furthermore, the electrostatic capacities between the many multi-turn windings produce delays and ringing, effects particularly serious with fast ferrite cores. The delays, which differ for the various windings linking a given core, are detrimental to the operation of the type of switch in which the selecting current pulses themselves turn over the selected core, since they must be in exact time coincidence.¹¹ In the type of commutator switch in which the selecting currents are inhibiting - leaving only the selected core uninhibited - and the drive winding links all cores, the effects of the delay and ringing can be reduced when a longer pulse scheduling involving "priming" and "driving" periods is used. As considerable analysis has shown, a practical solution to the many position decoder switch is in a cascaded use of the "commutator" switches in a sort of a tree where the outputs of a first commutator are the drivers for the next. Such an arrangement would no doubt be practical for the 100-way switches required to drive the myriabit memory. In these, full decoding could be obtained from two binary coded decimal digits, the octuple coincidence being obtained, for example, in two quadruple coincidence steps.

DOUBLE COINCIDENT BIASED SWITCHES

Actually, we developed much simpler 100-way switches based on double coincidence. The switch cores are in a matrix array, 10 by 10 in our case, and are linked by row and column windings, as shown on Fig. 7.

---

¹⁰ Similar switches were designed independently by K. Olsen of M.I.T.

¹¹ In spite of this difficulty, R. Stuart-Williams and M. Rosenberg, in June, 1952, drove a 256 array by two 16-way commutator switches, all cores being "ferrites," in a manner identical to the all-metal memory described by the author, previously.

In addition, there is a winding linking in the same sense every core and carrying a dc biasing current. Each of the 100 switch cores has a single turn output winding which energizes a line on one side of the myriabit memory array. Another similar 100-way switch energizes the other side as shown on Figs. 8 and 9.

The combinatorial principle of the switch is as follows: The biasing magnetizing force of one unit (-1) keeps all cores well saturated (point S on Fig. 7). The selecting row and column current pulses produce magnetizing forces of one unit ($+1$), so that only the doubly excited selected core is subject to a positive magnetizing force ($1 + 1 - 1 = 1$), while all singly excited line cores are subject to no magnetizing force ($1 - 1 = 0$) and the remaining cores simply remain biased (-1). The discrimination between selected and unselected cores depends only on the saturation characteristics (or slope of the B - H loop for large H's) of the cores and not on distinguishing between excitations in the ratio two-to-one as is the case for the memory cores, so that the bias and driving currents may be arbitrarily large. When the selecting line pulses rise, the selected core is switched from negative saturation to or towards positive saturation (e.g. points S* on Fig. 7), and when these pulses decay the dc biasing current automatically brings the selected core back to negative saturation without the necessity of a programmed restoring pulse. Correspondingly, a positive and then a negative voltage pulse are induced in the output winding of the selected core. The switch requires, for m outputs, $2\sqrt{m}$ driving tubes and not $4\sqrt{m}$ as it would without common restoration. It is not a decoder down to the binary code, but in our case, the loss of driving tube economy is insignificant because only twenty tubes are required per switch or a total of 40 tubes for both switches, as compared to two sets of 8 pairs of push-pull tubes required in a binary-coded decimal system or a total of 32 driving tubes.

Complete magnetic decoding could be obtained by driving each of these "double coincidence biased" switches by two quadruple coincidence commutator switches. In fact, this would be the preferred way of using magnetic switches cumulatively: Use double coincidence only in the last switch with most cores, and multiple coincidence in the first or early switches whose total number of cores is relatively small. We chose direct drive of the double coincidence switches in this project and conventional electronic binary to one-out-of-ten decoding, in order to devote most development efforts to the memory array itself.

The two ten-by-ten switches operate the memory array of 10,000 cores as follows. Current pulses are applied to the address selecting row and column windings of both switches. All four currents rise simultaneously and remain constant for a time at least equal to the switching time of the switch and memory cores. The two corresponding switch cores turn over and induce currents in the memory array lines to which they are linked. The sense of linkage is so chosen that their effects are additive, so that the selected core in the memory array is set to a given remanence, e.g. point P of Fig. 6. If all four current pulses decay simultaneously,

93

reverse effects are produced and the selected memory core is set from state P to state N. If, however, the two pulses on one of the driving switches terminate at a different time than the two pulses on the other switch, the negative currents on the lines of the memory array will occur in succession and will bring the selected memory core to $P \frac{N}{2} \frac{N}{2}$ after inflicting two partial demagnetizations $\frac{N}{2}$ upon it. This proves a simple method for registry or writing information in the memory: control of the length of the pulses on one switch. Such control is obtained conveniently by feeding all the driving tubes from a common master pulse (see Fig. 15). To write P, the action of the master pulse is inhibited before its termination on one of the switches, while such inhibition is omitted to write N. The simultaneity of the relevant pulse edges is insured, since they originate from a common master pulse.

The biased switches can be used in a variety of ways to drive a memory in which access is in parallel to a group of digits. For example, consider the cores of the 100 x 100 myriabit plane subdivided for parallel access to ten digits into subplanes of 10 x 100. The lines in the 100-way x direction are driven by a 10 x 10 switch, as described above, and are excited by the longer standard current pulses. The ten groups of lines in the y direction are driven by ten, 10-way switches. Every position of these decade switches is controlled (directly or by double coincidence in a 2 x 5 arrangement) by the address selecting pulses which are also the standard longer pulses. For example, core #7 in all ten decade switches is excited if the 7th line in every digit subplane is to be addressed. In addition, each decade switch may be inhibited by the write signal corresponding to its digit position, so as to shorten its standard exciting pulses. The inhibited switches will cause a P write-in (or rather $P \frac{N}{2} \frac{N}{2}$), while the non-inhibited ones will cause an N write-in, in the corresponding subplanes. Of course, a three-dimensional arrangement physically paralleling the digit planes is possible.

CONSTRUCTION OF THE SWITCHES

The ten rows and ten columns primary windings have 16 turns, while the outputs are single turns so that the cores act not only as switches but current step-up-transformers. Each turn of these windings links every core, i.e., any line of cores is wound by a single common winding as shown schematically in Fig. 7. In this way the effects of interturn capacity, which would be very severe if separate 16-turn winding on each core of the line were in series, are greatly reduced. The pulse shape deterioration is equalized and differences in time of pulse occurrence are negligibly small. Also, this winding method lends itself to a very simple switch construction. The row and column windings are set at an angle, 45° in our case, allowing straight passage of both windings through the cylindrical hole of the core which was appreciable height-to-diameter ratio. The two sets of windings are spaced approximately a core

94

diameter apart, which permits them to be the mechanical supports of the cores and keeps the interwinding capacity to a minimum. Fig. 10 shows the diamond-shaped switches.

The cores of the switches are wound of 40 wraps 1/8 mil Mo-Permalloy 1/8" wide tape on ceramic bobbins about 5/16" O.D. and 3/16" I.D. This does not limit the speed of operation, as it does when similar cores are used in a memory array, because the driving current is not limited by the requirements of good discrimination under a two-to-one regime, but can be made as large as required by the losses inherent in the fast turn-over.¹² The turn-over time seems to be, on the basis of a few tests, inversely proportional to the drive. Switching times of about one microsecond are obtained for current drives comparable to those required for turning over presently available ferrosphenel magnetic cores. The large magnetizing force is required in metals because of high losses and in ferrites because of high coercivity. The advantage of the metal core lies in its better saturation characteristic, resulting in smaller disturbing outputs on partially excited cores. To obtain similar performance in an experimental ferrite switch it was necessary to use several cores on each switch position so as to minimize by neutralization the otherwise larger disturbing signals.

THE OPERATION OF THE SWITCHES DRIVING THE MEMORY ARRAY

The detailed analysis of the operation of the switch, whose combinatorial principle was explained above, is beyond the scope of this paper because the cores act as very lossy nonlinear transformers with nonlinear and variable loads. The empirical results will be explained by considering first the behavior in the loads.

Consider the output circuit of a switch core loaded by a memory line as illustrated on Fig. 11. If ϕ_s is the flux in the switch core, ϕ_m the total of the flux of the 100 memory cores on a line, R the total resistance of the circuit and i the instantaneous current, then:

$$\frac{d\phi_s}{dt} = \frac{d\phi_m}{dt} + iR \quad (1)$$

$$\Delta\phi_s = \Delta\phi_m - \int_0^t Ri dt, \quad (2)$$

where $\Delta\phi_s$ and $\Delta\phi_m$ are the changes of flux from 0 to t . The capacities are so small that displacement currents can be neglected. The verification of these equations for the particular values required by the typical memory-core characteristics will be taken as a demonstration of the switch operation.

---

¹² e.g. The voltage responses of Fig. 6 ("Static magnetic matrix memory and switching circuits").

95

The voltage wave shapes V_1 , V_2 and V_3 as defined on Fig. 11, are shown on the oscillographs of Fig. 12. The current i , given as $V_2 - V_3$, rises in $0.2 \mu\text{sec}$ to a peak of 600 ma while remaining above 500 ma for about $0.4 \mu\text{sec}$, so that it is fairly similar to the test current pulses of Fig. 5. Consequently, the memory core operates near optimum drive. The voltage $V_1 - V_2$ is produced by the change of flux in the half-excited 100 memory cores on the line and the resistive drop of the driving wire itself. At the peak, this difference is 2 volts, the resistive drop 0.1 volt and the contribution of the selected core 0.16 volt, so that the 99 unselected cores contribute each about 17 mv. This value is to be compared to the 4.5 mv produced by the core when going from $P \frac{N}{2} \frac{N}{2}$ to $P \frac{N}{2} \frac{N}{2} \left(\frac{P}{2} \right)$ and 8 mv when going from N to $N \left(\frac{P}{2} \right)$, as seen on Fig. 5. The large discrepancy of these results can be partially accounted for by the higher current peak and somewhat faster current rise in the switch as compared to the test pulses. But it is mostly due to the fact that the voltages are very sensitive to the loss characteristics of the cores whose average, in the myriabit, may be slightly different from the typical core chosen for tests.

The agreement of flux changes, independent of core losses in this instance, is much closer. The flux change $\Delta\phi_s = V_1 dt$ is 101 Mx (1 Maxwell = 10^{-8} volt-sec), and $\int V_2 dt$ is 76 Mx, as determined by measuring the areas under the curves on Fig. 12. The integrated voltage drop across the 0.2 ohm of the $10''$ line wires is about 10 Mx. This leaves 15 Mx for $\Delta\phi_s$. This value should be between $99 \times 1.2 + 5.2 = 17.1$ Mx and $99 \times .08 + 5.2 = 12.1$ Mx, depending on the storing state of the 99 cores, on the basis of the measured changes of flux of the typical core as shown on Fig. 6, since the change from N to $N \frac{P}{2}$ is 0.12 Mx, and from $P \frac{N}{2} \frac{N}{2}$ to $P \frac{N}{2} \frac{N}{2} \frac{P}{2}$ is 0.08 Mx and for the selected core from N to P is 5.2 Mx.

The series resistance of 1 ohm was inserted in the myriabit array lines in order to shorten the time constant L/R of the circuit and obtain the desired effective differentiation of the primary pulses. This resistance acts also as a current regulator for the variable countervoltage produced by the half excited line cores which depends on the state of these cores (nominally varying from 0.45 to 0.8 V).

The driving primary current pulses and the dc bias were adjusted so as to obtain the desired outputs just described. The dc bias in 3.8 A.T. (or 1.9A. in 2 turns) and the driving pulses in each switch line 4.4 A.T. (275 ma in 16 turns). The effective positive magnetizing force ($2 \times 4.4 - 3.8 = 5$) is somewhat higher than the negative magnetizing force due to the dc (3.8 A.T.) acting alone after the termination of the exciting pulses (points S and S* on Fig. 7), while the positive output voltages are somewhat smaller than the negative ones. This asymmetry is due to the fact that the voltages are dependent on the variations of current, which, near maximum, rises more slowly than

76

it decays near the end. The switch can be deliberately driven unsymmetrically, the positive drive terminating in points such as Q_1 or Q_2 on Fig. 7 if it is desired to obtain a sharp positive and a broad negative pulse or vice-versa, corresponding naturally to equal positive and negative flux changes.

The voltage peak induced in the primary 16 turns is 50 volts or 3.1 volts per turn, while only 2.8 volts are obtained in the single turn secondary showing some leakage of flux. From the total effective current drive of 8.8 A.T., about half overcomes the dc bias; most of the remaining current overcomes the high losses of the fast-switched metal core while only about $1/2$ A.T. appear in the useful load. Even with this poor performance from the point of view of a power transformer, the over-all performance of the switch is efficient, as it allows the use of 40 tubes capable of carrying 275 ma instead of 400 tubes capable of carrying 550 ma which would be required in a myriabit array driven directly by tubes.

READING FROM THE MEMORY

Reading from the memory or interrogating it at a given address is accomplished by sending definite current pulses through the appropriate lines and observing the induced voltages in a winding linking all 10,000 cores of the array. Consider first the simultaneous $P/2$ drives on both selecting lines used in earlier memories² bringing the core from N to (P) or from $P \frac{N}{2} \frac{N}{2}$ to $P \frac{N}{2} \frac{N}{2} (P)$. If the selected core were the only one on the reading winding, the discrimination between the two peak voltages would be 160 mv to 9.5 mv or 17 to 1 and between the two flux changes 5.2 Mx to 0.2 Mx or 26 to 1, as can be seen on Figs. 5 and 6. Incidentally, the ratio of final flux change N to P to that from $P \frac{N}{2} \frac{N}{2}$ to $P \frac{N}{2} \frac{N}{2} \frac{P}{2}$ is 35 to 1. Unavoidably, each of the 198 cores on the two selecting lines contributes its small share of peak voltage, which is 4.5 mv when going from $P \frac{N}{2} \frac{N}{2}$ to $P \frac{N}{2} \frac{N}{2} \left(\frac{P}{2}\right)$, and 8 mv when going from N to $N \left(\frac{P}{2}\right)$, or almost 8 mv from $N \left(\frac{P}{2} \frac{N}{2}\right)_{\infty}$ to $N \left(\frac{P}{2} \frac{N}{2}\right)_{\infty} \left(\frac{P}{2}\right)$. If the reading winding linked all cores in the same direction, these voltages could add up to $198 \times 8 \text{ mv} \approx 1.6$ volts or ten times the voltage from the selected core. With the checkerboard pattern of linkage polarities, the direction of linkage changing for every core on every line, used already in all previous memories,^{2,4} these voltages tend to cancel each other. At best, when all cores are uniform and storing in the same polarity, the contributions of 98 cores on one line exactly cancel the contributions of the 98 cores on the other and the two remaining cores increase the already high discrimination of the selected core. When the cores on the selected lines are in different remanent states, the cancellation is no longer perfect even if the cores were uniform, since their contributions are 4.5 or 8 mv depending on the polarity of their remanent state. At worst there will be 100 cores contributing 8 mv and 98 contributing 4.5 mv in the opposite direction, so that the total disturbance will be 357 mv or roughly 2.2 times as much as the voltage from the selected core.

97

The peak voltage is seen to be a completely inadequate signal even when the cores are uniform. The strobing of the voltage past the peak provides a much better signal. This system was used in previous memories^{2,4} and is particularly adaptable to the slower metal core arrays. With very fast magnetic ceramic cores and large arrays driven by switches, as in our instance, the strobing technique is more difficult.

By exciting the selected core first in one polarity and then the other, it is possible to neutralize the unwanted contributions of the half excited unselected cores obtained in the first step by those of the second. Such a cancellation method does not depend, in principle, on the remanent state of any core nor whether it has like properties to other cores. One such method consists in integrating the signal over the period of time in which the selected core is excited to P and then to N. Consider first the contribution of the selected core to the integrated signal. If the core is initially at N, it will be driven to P and then returned to N, so that the net change of flux is precisely zero. If initially at P, the core will be driven to PP (very close to P, not shown on Fig. 6) and then to N so that the net change of flux is large, i.e., 5.2 Mx in our case. Consider next the contribution of an unselected core on the selected lines which was initially at P $\frac{N}{2} \frac{N}{2}$ going to

P $\frac{N}{2} \frac{N}{2} \frac{P}{2}$ and then P $\frac{N}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2}$. As far as we could measure, the states

P $\frac{N}{2} \frac{N}{2}$ and P $\frac{N}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2}$ are identical as mentioned before, and the core de-

scribes a closed minor loop with zero contribution to the integrated signal (see Fig. 6). Finally, consider an unselected core initially at N, changing to N $\frac{P}{2}$ and then to N $\frac{P}{2} \frac{N}{2}$. Here the minor loop is not closed and there is a slight net change of flux of about 0.04 Mx.

This non-closure of the minor loop, from N to N $\frac{P}{2} \frac{N}{2}$, is not detrimental because the filling of the memory array by the regular writing-in process necessarily leaves most negatively storing cores in the state

N $\frac{P}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2} \dots$ where the cycle $\frac{P}{2} \frac{N}{2}$ is repeated not less than three

times. As mentioned before, the minor loops after the first excursion are almost closed, and after the second, are completely closed. Therefore it is only the unselected cores in state N which contribute to the integrated signal. There can be at most two such cores, since writing along any line

necessarily subjects all cores to cyclic $\frac{P}{2} \frac{N}{2}$ excursions. (In the five-step

schedule actually there are two cycles, i.e., $\frac{P}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2}$, for each write-in.)

The greatest integrated unwanted reading signal occurs when the selected core changes from N $\left(\frac{P}{2} \frac{N}{2}\right)_{\infty}$ to N and the two unselected cores in their

98

non-asymptotic states N are unfavorably linked by the checkerboard reading winding and change from N to $N \frac{P}{2} \frac{N}{2}$. For our cores this total flux change is about $0.22 Mx$ or about $1/25$ of the wanted signal of $5.2 Mx$ (points $N \frac{P}{2} \frac{N}{2} \frac{P}{2} \frac{N}{2}$ and $N \left(\frac{P}{2} \frac{N}{2} \right)_{\infty}$ not shown on Fig. 6).

The ratio of wanted-to-unwanted reading signal in the integration method is seen to be large and is independent of the size of the array and of the uniformity of core properties. In practice, this ratio is determined primarily by the characteristics of the integrating circuit.

The experimental myriabit was wired with cores which were tested manually, every core only once, while standards of tests were being worked out. This resulted in poor uniformity as the bulk of the cores have outputs varying in the ratio 2 to 1, while about 50 cores are outside that range, including 15 positions giving practically no output (one core missing). The lack of core uniformity is apparent from the typical direct reading signals illustrated for two lines on Fig. 13. Even so, the integrated signal is useful for all but the 15 positions and sets reliably an output flip-flop.

The reading winding can, but does not need to be, in a checkerboard pattern with the integration read-out system, since the disturbances from a given core are cancelled by opposite disturbances from the core itself rather than by disturbances from other cores. Nevertheless, the myriabit array was checkerboard wound in order to reduce the excursion of voltage in the integrating circuit. Also, in order to neutralize the voltages induced in the relatively large read winding loop through straight air linkage the 100 subdivided, checkerboarded read windings were connected themselves in an appropriate checkerboard.

PROGRAMMING

The program of pulses for the memory operation, shown on Fig. 14, can be considered to be made of steps of equal duration. To read from the memory, identical pulses of current are applied to both x and y switches in step 1. During the plateau of these pulses, the selected core is set in P . The minimum length of this plateau, determined by the time constant L/R of the load circuit so as to allow the positive current to decay to zero, was found to be $2.5 \mu\text{sec}$ as has been explained. With a larger resistance, this duration could approach the limit of $0.3 \mu\text{sec}$ set by the switch-over time of the memory core itself, but excessive drive power would be required if it was desired to shorten the length below $1 \mu\text{sec}$. After the end of the pulse the memory core is set to N , but again $2.5 \mu\text{sec}$ must be allowed in step 2 for the decay of the negative current (see Fig. 12). The voltage appearing on the reading winding is integrated over the $5 \mu\text{sec}$ of steps 1 and 2, and its integrated value at the end of step 2 is the signal indicating the

99

initial state of the interrogated core. In step 3 current pulses are again applied to the switches setting the selected core back to P and simultaneously the signal sets the circuits which will control step 4. In step 4, both pairs of currents driving switches x and y remain at their plateau if the core was initially at N while the pair of pulses driving switch y are terminated if the core was initially at P. In step 5, the negative currents in the output die off. The core is seen to be restored to its initial condition in the steps 3, 4 and 5 which constitute the rewrite. These three steps could be omitted if the core were found to be in N, but by their use the time of step 3 is available for the reading integrated signal to set the controlling circuits.

To write or register in the memory, only steps 3, 4, and 5 are required. For simplicity, all 5 steps are used in our system. The write signal is used instead of the read signal to control whether the currents driving the switch y in step 4 are present or not.

The total time for a write-in or read-out, leaving the memory ready for the next write or read operation, is $5 \times 2.5 = 12.5 \mu\text{sec}$. This time, which may be considered as the random access time, could be reduced to $5 \mu\text{sec}$ as mentioned above. Further saving of time could be accomplished by omitting steps 1 and 2 in writing.

ASSOCIATED CIRCUITS

The electronic circuits built to test the myriabit memory are shown in Fig. 16 and in block form on Fig. 15. The address is set on the four address registers in the binary coded decimal form. It is decoded to the straight decimal system by vacuum tubes and it primes one switch driver in each group of ten. The driver tubes are two 6AU5's in parallel. A master pulse derived from the logic control fires the primed drivers at the appropriate time. A four-digit binary coded decimal counter allows scanning through all the memory addresses. The scan can be through all 10,000 cores, or any manually selected 1000, 100 or 10, or it can be stopped on any particular core. The logic control furnishes pulses required for the five step program (or other test programs). The reading integrating circuit is of a "bootstrap" variety and is followed by amplifiers, clamps and discriminating circuits. The output flip-flop controls the restoring circuits in the logic control. The addresses are converted to analogue voltages for deflecting the beam of a cathode-ray tube. The partial or total contents of the myriabit can be displayed on the tube by controlling the beam intensity by the reading signal.

These circuits, for experimenting with the memory, were overdesigned on purpose. Properly engineered circuits for operating the memory in a computer could be simpler, require less power, and would be combined with the computing circuit themselves. While little is said here of almost conventional circuitry, it did constitute a major part of our effort as seems unavoidable in high-speed memory research.

THE EXPERIMENTAL RESULTS

At the time of writing this paper, testing of the memory is still in progress, but the results obtained show that the memory performs as expected. The stored information in any core was always found to be retained, even after a very large number of accesses to other cores on its selecting lines. The integrated signal from any of the 9985 operating cores reliably sets the output flip-flop even when the pattern of stored information is deliberately set to give maximum disturbance. The 15 non-operative core positions include one with a missing core, and others which are cracked or accidentally passed the relatively crude testing to which the cores were subjected. Most of the tests were made with a 5 step program lasting 25 μ sec, step #3 being 15 μ sec, but the memory was operated also with the designed 5 step cycle time of 12.5 μ sec.

CONCLUSIONS

Ferrospinel magnetic materials with rectangular hysteresis loops were developed in the middle of last year. By the use of automatic molding and testing machines it became practical to envisage memories with capacities in tens of thousands of bits. The basic problems in the realization of such memories are the switching and read-out methods by which fast random access can be obtained efficiently. Successful solutions to these problems are demonstrated in the experimental 10,000 bit memory reported in this paper.

Last year, on the basis of a memory with 256 bits, we predicted² that in the near future storage capacities of thousands or tens of thousands would be achieved and capacities of millions, in a more distant future. As the first part of this prediction is now realized we venture to reiterate the second. Of course, this will require great innovations in construction techniques and still further improvements in magnetic switching. But we have no doubt that memories in which random access to millions of information bits in a few microseconds will be available at a relatively low cost. This will usher in a new area in information handling systems which may have profound consequences, not only in scientific computations, but wherever complicated or repetitive computations burden our lives.

101

ACKNOWLEDGEMENT

This project was brought to success by many members of the RCA Laboratories, principally R. Stuart-Williams¹³ and M. Rosenberg¹³ who made important original contributions. The essential ferromagnetic materials research work was done by I. J. Hegyi, R. S. Weisz, and R. L. Harvey. Many of the attendant tests were made by T. Arbaeus, American-Scandinavian trainee at the Laboratories. J. Lehmann and S. Sternberg designed many of the circuits and C. H. Vose the circuit lay-outs. A. Lo and W. Modlinski participated in initial measurements and M. Kaplan from RCA Victor Division did most of the final testing of the system. R. L. Hunyady wired the myriabit array, and J. Valentine contributed in the design and built the automatic core tester.

The project had the constant supervising interests of V. K. Zworykin.

---

¹³With International Telemeter Co. since May, 1953.

Figure 1

COINCIDENT CURRENT MAGNETIC CORE MEMORY ARRAY AND IDEALIZED HYSTERESIS LOOP

103

Figure 2

Figure 3

105

Figure 4

106

Figure 5

MANGANESE; MAGNESIUM; FERROMAGNETIC; SPINEL CORE

0.54" O.D.

0.34" I.D.

.016" H.

ACTUAL TEST CURRENT
PULSE SHAPES ACTUAL VOLTAGE PULSE
SHAPES ON SINGLE TURN

Figure 6

CORE .054" O.D. .034" I.D. .016"H

Figure 7

DOUBLE COINCIDENCE BIASED MAGNETIC SWITCH

109

Figure 8
MYRIABIT MEMORY

Figure 9

0552

Figure 10

0550

Figure 11

OUTPUT CIRCUIT OF SWITCH CORE

Figure 12

VOLTAGE IN SWITCH OUTPUT CIRCUIT

Figure 13

VOLTAGES ON READING WINDING FOR TYPICAL CORE ON LINES WITH (a) WELL BALANCED, (b) POORLY BALANCED CORES

115

Figure 14

FIVE STEP SCHEDULE OF OPERATION

Figure 15

BLOCK DIAGRAM OF MYRIABIT MEMORY CIRCUITS

Figure 16

FERROMAGNETIC CORES WITH MICROSECOND ACCESS

An Application in ENIAC's Static Magnetic Memory

by

Isaac L. Auerbach

Burroughs Corporation
Philadelphia, Pennsylvania

INTRODUCTION

The use of recently developed rectangular hysteresis loop ferromagnetic material as memory elements for digital data processing machines is now achieving widespread popularity. The advantages of magnetic memories are manyfold: power dissipation is low, no standby power is required for storage, and extremely long life with reliable operation can be assured. The important requisite of a ferromagnetic material for a storage element is two stable states of residual magnetization. These properties are closest to ideal when the hysteresis loop approaches a rectangle. The nickel-iron alloys - molybdenum permalloys and ferrites - provide this extraordinary loop rectangularity and form the basis for a whole field of circuit development and memory systems.

A parallel memory system that uses one magnetic core and one diode per bit of storage was developed in 1951. The memory system is characterized by its positive action - the memory is not amplitude sensitive - and large variations of driver tube, diode and core properties can be tolerated. Access times as short as 5 microseconds are possible.

The first large-scale application of this storage technique is in a 100-word high speed internal memory designed for the ENIAC. The memory is at present installed at the Ballistic Research Laboratory, operating on problems with the ENIAC. This one application does not test the true capabilities of the memory system since the ENIAC is a relatively slow computer by today's standards. However, this first embodiment has served to demonstrate the basic advantages of magnetic memory storage elements.

This paper will be divided into two parts: The first, a discussion of the magnetic properties of ferromagnetic materials under pulsed conditions; the second, a discussion of the more outstanding features of the Static Magnetic Memory System for the ENIAC.

119 MAGNETIC CORE PROPERTIES

Ferromagnetic cores dissipate power only when they are sensed or when they are switched from one to the other of their two stable states. A magnetic core stores information for an indefinite period of time as the residual state of magnetization without power applied to the circuit, just as a bar magnet retains its state of magnetization as the retentivity of its North or South polarization. Life of magnetic units may be compared with pulse transformers and with other wire-wound components. The magnetic material that makes up the core need only be protected from humidity and excessively high temperatures, and the copper wire protected from normal corrosive elements. The life of magnetic cores should therefore be greater than 10 years, and life between 25 and 30 years would not be out of reach.

A. Core Characteristics

An idealized DC hysteresis loop for a ferromagnetic core is shown in Figure 1. The characteristics of particular interest in this discussion and that to follow are B_r , the residual magnetization; B_m , the maximum or peak magnetization; and H_c , the coercivity or coercive force. These characteristics of the magnetic core make it possible to store binary information as the sense of the residual magnetization, either $+B_r$ or $-B_r$. We will define $+B_r$ as the state for the storage of binary ONE, and $-B_r$ as the state for storage of binary ZERO.

Two properties of the magnetic core that are of general interest and importance are the squareness ratio and the switching time. The squareness ratio is defined as the ratio of residual magnetization, B_r , to the maximum magnetization, B_m . Typical values for these properties vary from B_r/B_m of .85 to .95. The switching time is defined as the time required for the magnetization state to change from $-B_r$ to $+B_m$ when an unloaded core is subjected to a step of magnetizing force of an amplitude, H_m . Switching times for typical cores vary from less than one microsecond to about thirteen microseconds. The switching time has been found to be inversely proportional to the applied mmf less the coercivity of the material,

$$t_s \sim \frac{1}{H_m - H_c}$$

Another property of the magnetic core that will enable a better engineering understanding of magnetic circuits is that of impedance. Bistable magnetic cores exhibit two impedance values, depending upon their state of magnetization and the direction of the sensing magnetic force. If the magnetic core is in the ZERO state, and the sensing force also is toward ZERO, then the core will appear as a low impedance, and passage of current through the sensing winding will be "easy." Little back magnetomotive force is generated since few lines of flux are changed within the core. If, on the other

hand, the magnetic core had been at state ONE, and the same sensing force applied, a large change in flux would result as the core switches from positive residual magnetization to negative saturation. During this change the impedance of the core would be high, and passage of current through the sensing winding would be "hard." Numerous magnetic circuits can be designed around this "high-low" impedance characteristic of the magnetic core.

B. Core Measurement Methods

To facilitate the testing of large numbers of toroidal core without recourse to winding each and every core, special test equipment was constructed. Figure 2 is a sketch of the core measurement jig that is used. The silver-plated copper tube serves both as a support for the cores and as a part of the single turn output winding. Two insulated conductors threaded through the tube serve as single-turn windings for applying positive and negative magnetizing forces of sufficient amplitude to switch the core through its major hysteresis loop. The test probe consists of a two-tined fork shown in the figure. Each tine makes electrical contact with the silvered tube on each side of the test core; the tube surface and the two tines constitute a single-turn output loop. This loop encloses an area that is much greater than the cross-sectional area of the test core, and the loop has induced in it, by the air-coupled flux, a relatively large voltage. The effect of this undesirable voltage is eliminated by including within the probe several turns oriented in such a way that the voltage induced in them exactly cancels the undesirable voltage. Since these reversed turns do not include the test core they have no effect on the observed core output voltage. The pattern of applied magnetizing force used in the basic core test is shown in Figure 3. The applied pulse pattern switches the core through the following excursions:

Consider the core to be initially in the ZERO state, or at negative residual magnetization. The first pulse, P_1 , switches the core from $-B_r$ to the $+B_m$ state. Upon the removal of the current, or magnetizing force, the core will assume the state $+B_r$, or that of positive residual magnetization. The second pulse, P_2 , switches the core from $+B_r$ to $+B_m$, and upon removal of this force the core will again resume the state of $+B_r$. The third applied force, pulse P_3 , switches the core from $+B_r$ to $-B_m$, and upon the removal of this applied force the core will resume its initial state of $-B_r$.

The voltages generated in the output loop for each of the applied magnetizing forces are shown in Figure 3b. The first applied pulse, P_1 , switched the core and generated the large change in flux, which resulted in the first voltage shown. However, the second applied pulse resulted in only a small change in flux from $+B_r$ to $+B_m$, and the voltage generated is proportional to this small change of flux or that shown as the noise voltage. The third signal restores the core to its initial condition and is not used to measure the properties of the core; however, the voltage generated should be identical to that for the first applied pulse but of opposite polarity.

121

Note that the amplitude of the signal S is of greater magnitude and duration than the noise designated N. The area under these curves is related to the squareness ratio of the core by the expression

$$\frac{S}{N} = \frac{1 + B_r/B_m}{1 - B_r/B_m}$$

By integrating the output voltages, the pattern shown in Figure 3c is derived. The flux pattern affords a simple and direct method for measuring the squareness ratio B_r/B_m to an accuracy that is adequate for most applications. For practical reasons, a multi-turn output winding is generally used to drive the integrator circuit when this technique is employed.

C. Comparison of Core Characteristics

By superimposing the signals designated S and N on the face of an oscilloscope, the pattern shown in Figure 4 can be obtained. This is the switching and non-switching voltage for a magnetic core of 48% nickel-iron, 1 mil thick, with an applied H_m of 1 oersted. The observed switching time was approximately 13 microseconds, and a voltage of 0.08 volts was developed across the one-turn output winding.

Figure 5 shows a comparison between the switching times for three different cores all being switched by 1 oersted magnetizing pulses. Note that for a given value of applied field, faster switching will be realized with a core of thinnest gauge, lowest maximum flux density, and smallest coercivity.

Figure 6 compares the switching voltages of 1/4 mil, 4-79 Permalloy for three different values of applied magnetizing force. Figure 7 shows the switching voltages of the ferrite MF-1118 for three different values of applied magnetizing force. Although transversing the saturation loop of a material is spoken of loosely, none of the applied fields used in this test is sufficient to truly saturate the cores. Each different value of applied H_m causes the core to transverse a slightly different loop, and each loop has different values of H_c , B_r and B_m . The decreased area under Waveform III in Figure 3 definitely indicates a transversal of a minor loop. A comparison of curves I in Figures 6 and 7 shows them to be nearly identical in switching characteristics so that to obtain a switching time less than 1 microsecond, with a 1/4 mil, 4-79 Permalloy and the ferrite MF-1118, a magnetizing force of 1 oersted is required for the former, and 4 oersteds required for the latter.

These switching times give ample indication that memory systems can be designed with microsecond access times. These core measurement methods have enabled the establishment of specifications for core materials with sufficient accuracy so that memory and logical circuits can be designed from the results of the measurements with regularity.

STATIC MAGNETIC MEMORY SYSTEMS

A. Magnetic Matrix1. General

Consider a matrix of magnetic cores and diodes arranged in rows and columns as shown in Figure 8. Each core is linked by two windings; one set of windings is connected in series by columns. One end of the other set of windings is connected in series with a diode; all of the diode terminations are connected by rows in common; all of the winding terminations of the second set of windings are connected by columns in common. Words to be stored correspond to columns in the matrix. The digits within a word correspond to the rows in the matrix. All cores are magnetized to one or the other of two residual states of magnetization designated ONE and ZERO.

To insert information into a core, a difference of potential must be established between the digit bus and the word bus such that the diode connected in series with the read-in winding will conduct. The current through this loop must be of sufficient magnitude to switch the core to state ONE. To read data out of a core, a pulse of current is transmitted through the clear winding of sufficient amplitude and of such direction as to cause all cores in the word column to switch to $-B_m$, or to state ZERO. Thus, cores that contained ONES will be switched back to the ZERO state, and in so switching will induce a large voltage into the winding in series with the diode and this voltage will appear on the digit bus.

A sequence of operations executed in order to introduce data into the memory array is as follows. Initially a current pulse is applied to the clear terminal corresponding to the selected word; this will switch all of the cores to the ZERO state. Next the digit lines are primed by the application of a voltage to designate the ZERO or ONE to be stored. As soon as the digit lines are primed, a voltage pulse may be applied to the read-in terminal corresponding to the word selected. All those cores in which ONES are to be stored will switch by the passage of current through the diode and coil winding. All other cores in which ZEROS are to be stored will not pass current since the diode will be biased in a negative sense.

In order to extract data from the array the same sequence of operations is followed, except that during the application of the clear signal the voltage induced onto the digit busses is detected by electronic circuitry. This voltage is then used to polarize the digit busses from which a ONE was obtained, and the application of a signal to the read-in terminal of the same word to reinsert the data into the cores, thus restoring it. The data originally stored in the magnetic cores has been introduced into the electronic circuitry where it may be used.

123

The matrix described above is the basic circuit of the core diode memory. Figure 9 shows a 4 x 4 magnetic matrix with associated magnetic switches to supply the read-in and clear signals, and cathode followers to control the digit busses. This shows, in a simplified way, the manner in which magnetic switches may be used to operate the array, thereby reducing the number of driver tubes that would otherwise be required. The magnetic switches are primed by the flip-flops shown at the right side of the figure, which prime the switches in pairs. The read-in and clear drivers alternately pulse these switches to generate the signals required in the array itself. The efficiency of the use of magnetic switches increases rapidly as the number of words in the memory increases.

2. Magnetic Matrix Used in the ENIAC Memory

Due to the state of the art and the engineering schedule for the design and construction of the Static Magnetic Memory for the ENIAC, the advantages of the magnetic switch could not be applied in this system. These switches were replaced by vacuum tube drivers. The basic operation of the Memory is identical to that explained above, and each of the operations will be explained in greater detail below.

Figure 10 shows a 2-word, 2-digit-per-word magnetic matrix in which the electronic circuitry and voltages have been specified. Memory elements M-1 and M-2 constitute one word, and elements M-3 and M-4 the second word. Cathode-follower V-1 is a read-in driver for word ONE, and V-2 is the clear driver. Flip-flop FF-1 and FF-2 control the potentials of points A and B, respectively, to determine whether a ONE or ZERO is to be written into a memory element. The flip-flop voltage swing is sufficient to insure that points A and B are clamped at either -15 volts or ground by the diodes shown.

Read-in Operation - Figure 11 illustrates the voltage waveforms that are significant during a read-in operation. A ZERO is being stored in core M-1, and a ONE is being stored in core M-2 (both cores are initially in the ZERO state). At time t_0 , FF-1 is set and point A is raised to ground potential by cathode-follower V₅. Diodes X₃ and X₅ prevent current from flowing in memory element M-1 and M-3. FF-2 remains in the reset state, and point B is clamped at -15 volts by diode X₂.

At time t_1 , a read-in pulse is applied to the control grid of tube V₁, causing it to conduct. Point C is raised to the -2-volt level. The potential difference of 13 volts across diode X₁ and winding N₂ of memory elements M-2 causes current i_w to flow in such a direction as to introduce a ONE into M-2. No current can flow in element M-1 because of the negative bias on diode X₃, and M-1 remains in the ZERO state.

124

At time t_2 , core M-2 is saturated and the impedance offered by winding M-2 becomes negligible. Current i_w increases until point B is clamped at the -2-volt level. The value of R_1 determines the maximum current that must be delivered by tube V_1 for each element that stores a ONE in the selected word.

Clear Operation - Figures 12 and 13 illustrate the clear operation as applied to the second word. Assume that memory element M-3 contains a ONE and element M-4 contains a ZERO. All digit-controlling flip-flops are initially reset, and write driver V_3 is cut off. Hence, points A and B are at -15 volts and point D is at -17 volts. A clear pulse is applied to point F at time t_1 and current I flows through the N_1 windings of all the elements in the second word. Core M-3 is switched from the ONE state to the ZERO state, and point A is raised to ground potential, priming gate G_1 for as long as the flux change lasts. Core M-4, which contains a ZERO, is switched to negative saturation, and, as indicated in Figure 13, a voltage of short duration appears at point B.

A sampling pulse is applied at gates G_1 and G_2 at time t_2 to provide time discrimination against the ZERO voltage. The start of the sampling pulse is delayed beyond the start of the clear pulse by an amount that is greater than the maximum duration of a ZERO voltage. Time t_3 , Figure 13, indicates the end of the switching of core M-3. As previously explained, only gate G_1 is primed at time t_2 ; the sampling pulse passes through gate G_1 and sets FF-1. The sampling pulse to FF-2 is blocked by the unprimed gate, G_2 . (If it is desired only to clear a word without using its stored information, the sampling pulse is not applied.) At time t_4 , the clear pulse is terminated. But FF-1 is now in the set state, and FF-2 is still in the reset state. A complement pulse at time t_5 will change the flip-flops to the conditions they had when the data was initially inserted into the word position. The information transfer from the cores to the flip-flops is complete, and it is immediately available for delivery to a data-processing machine.

Restore - To restore to the memory elements the information that they held prior to the clear operation, it is necessary only to apply a read-in pulse to the grid of write driver V_3 . The complement pulse at time t_5 (Figure 13) has established the proper potentials at the digit control points.

Access Time - Figure 14 is used to define and illustrate the access time of the magnetic memory system. Before time t_1 it is assumed that the address has been selected, that all digit-control flip-flops are reset, and that all clear and read-in drivers are in the non-conducting state. At time t_1 , the clear pulse is applied; it need last only as long as the switching time of the magnetic core. At any time t_3 after the core is switched the digit-control flip-flops can be complemented, and immediately following that the information can be restored to the memory elements with a read-in

125

pulse at t_4 . At the end of the read-in pulse the digit-control flip-flops can be reset with a pulse at t_5 , and the magnetic memory matrix is ready for the next operation.

The access time of the magnetic memory matrix is the time from t_1 to t_5 . Using standard electronic techniques and presently available magnetic cores, the access time has been designed to be less than 5 microseconds. The read and write operations can be performed in about 2 microseconds each, and flip-flops can be complemented or reset in a few tenths of a microsecond.

B. Memory System for the ENIAC

The requirements imposed by the ENIAC for the Static Magnetic Memory System were:

1. that the repetitive access time be less than one word time (200 microseconds),
2. that all data pulses be synchronized with data pulses in the ENIAC,
3. that the control signals used in the memory be derived from the standard control pulses of the ENIAC, and
4. that address selection be derived from two decimal ring counters.

The maximum clock-pulse frequency in the ENIAC is 125 kc. It is also required that the Memory operate at push-button speeds for checking.

Figure 15 is a block diagram of the Memory System; the numbers in the circles indicate the number of lines in the connecting cables.

1. Converter

The ENIAC employs a pulse count code such that the absence of a pulse indicates the number ZERO, and nine sequential pulses on a single wire indicate the decimal number nine. Ten decimal numbers and a sign constitute a single word. A 100-word memory system that did not re-encode the ENIAC code would have to have a capacity of 9100 bits; the Static Magnetic Memory uses a binary coded decimal system and needs only 4100 bits. The converter also provides the circuitry required to modify the signals received from the ENIAC to those used with the Memory System, and re-modifies the output signals to the waveshape required by the ENIAC. The cathode-followers in the converter provide the control for the digit bias in the core matrix. In addition to the above, the converter is also used to manipulate the data to achieve optimum utilization of the clear driver for establishing a unique code for storage of data within the magnetic matrix.

126

2. Address Selectors and Drivers

Address selection for the magnetic memory was determined by the state of two 10-position ring counters in the ENIAC itself. These 20 lines are used to select one of 100 words in the Memory by gating the screen and control grids of the driver tubes. The drivers are their own gates. The driver circuits are designed so that an address change can be effected in about 10 microseconds. There are 100 read-in drivers and 100 clear drivers, each assembled in a 10 x 10 matrix.

3. Control

The function of the control circuit is to govern the operation of the Memory System as dictated by the various synchronization pulses supplied by the ENIAC. The two basic cycles of the machine are store and extract, both of which are made up of the read-in and clear operation.

A store cycle, initiated by a store instruction pulse from the ENIAC after an address has been selected, proceeds as follows in the Memory: the cores in the selected word are cleared by the clear driver without sampling; a digit convertor flip-flop then counts the incoming data pulses from the ENIAC; the state of the flip-flops in each digit converter are then modified in accordance with a preset schedule; and finally, a read-in pulse is applied to the appropriate driver to store the data in the Memory.

An extract cycle, initiated by an instruction pulse from the ENIAC after an address has been selected, proceeds in the Memory as follows: the digit converter flip-flops are reset to ZERO; the clear pulse is applied to the selected word clear driver, and the data is transferred from the cores to the flip-flops by the code sampling pulses; all of the flip-flops are complemented and a read-in pulse is then applied to the appropriate driver, restoring the data to the magnetic core memory; the machine code is now converted back to the operational code and synchronized data pulses are counted back to the ENIAC.

4. Core Array

The magnetic memory array for the input stores 100 words in 4100 memory elements. Each 41-bit word consists of 10 binary coded decimal digits and a sign bit. The array is made up of 4100 packaged magnetic cores and 4100 germanium diodes.

C. Construction Features of the Memory

The entire Static Magnetic Memory System, including power supply, is contained in the cabinet shown in Figure 16. The over-all dimensions of the cabinet are seven feet by eight feet by two feet.

127

Great stress was placed on easy accessibility for testing and maintenance of all components contained within the Memory. Little effort was expended to minimize the size of the equipment. The converter and control sections of the Memory System are mounted on hinged doors on the front of the cabinet, with an access door to the wiring side for accessibility.

Figure 17 shows the wiring side of one of these doors. Printed circuit boards provide for distribution of signals, both horizontally and vertically. No coaxial cabling was used for the transmission of signals.

All interconnecting wires are unshielded; capacitance is minimized by adequate space between lines and by due consideration for impedance level. Filament transformers are mounted below each baffle; fuses, decoupling networks, and voltage-limiting rectifiers are placed above the baffles.

Figure 18 shows the tube side of this same door. Packaged tube and diode units plug in to chassis baffles, which are mounted on the door. The diode cans are isolated from the tube assemblies by metal barriers to prevent excessive heating of the diodes. The resistors used with the diode circuits are mounted on the side of each diode chassis and are separated from the diodes by a metal partition mounted on the cover plate. Blowers, mounted below the chassis baffles, provide adequate cooling even when the door is opened for maintenance work. Deflecting fins on the cover plates direct the air to hot spots. There are 180 tubes and 2500 germanium diodes in the code converter and control circuits.

Figure 19 shows the tube and diode plug-in units. The code converter and control circuits were both constructed with three basic types of tube circuits and nine different assemblies of diodes. The three basic tube circuits are a flip-flop, a cathode-follower, and a line driver. Germanium diodes are used for all logical functions and for signal clamping.

Each magnetic core is packaged in a nylon case. An exploded view of the assembly is shown in Figure 20. A $3/8$ " ceramic bobbin is wound with two wraps of $1/2$ mil, 48% nickel-iron and spot welded. After the metal is annealed, the core is covered with a piece of protective tape to eliminate touching of the metallic ribbon, which is strain sensitive. The core is then wound with #38 copper wire, 750 turns on the read-in coil and 100 turns on the clear coil. The wound core is then mounted in the nylon case and is ready for assembling into the memory array.

Figure 21 shows one of the four panels in the memory array. The cores are mounted in sticks of 41 and clamped so that, if ever necessary, a single core may be removed from the matrix without major difficulty. These sticks are then assembled on the panel and the diodes are connected as shown.

The DC power supply for the Memory is mounted in the end bay of the cabinet (Figure 22). The primary source is a 3-phase supply, and the power supply consists of only selenium rectifiers, transformers, and condensers. Regulation of better than 1% has been achieved on all critical supplies. A total power of 4.5 KVA is required for the Memory System.

CONCLUSIONS

The Static Magnetic Memory System described has been installed with the ENIAC at the Ballistic Research Laboratory, Aberdeen Proving Grounds, Maryland. Only four days were required from the day the equipment left our laboratory in Philadelphia until rigorous acceptance tests had been completed in the combined installation of the Memory with the ENIAC. Since that time long, error-free runs have been effected, which gives some indication that the Memory will continue as a highly reliable adjunct to the ENIAC. During our experience with the machine, we had negligible difficulty with germanium diodes or vacuum tubes. This was attributed to the conservative electronic design followed, and to the care taken to insure adequate cooling for the diodes and tubes.

The Static Magnetic Memory System described has the following outstanding characteristics:

1. One bit of data is stored in one ferromagnetic core.
2. Read-in and clear operations are positive such that, whenever a core is to be switched, a magnetomotive force to give three times greater than the coercive force is applied. Therefore, careful control of current and voltage values are unnecessary.
3. Core properties are not critical. More than 95% of all cores received were usable in the memory matrix.
4. The read-in and clear operations are performed under essentially constant-voltage conditions.
5. Although the signal to noise amplitude ratio is greater than 10 to 1, time discrimination is used to detect the state of the core during the clear operation and provides even greater reliability.
6. Information is stored even during a power failure.
7. Access times of less than 5 microseconds are practical and have been achieved within the laboratory.

129

ACKNOWLEDGEMENTS

The author wishes to acknowledge the great contributions to the design and construction of the Static Magnetic Memory System for the ENIAC by Mr. Lyle G. Thompson, and for the particularly significant contributions on core measurements by Mr. Joseph Wylen. The author also gratefully acknowledges the assistance of Messrs. William O. Boss, Isadore Gross, Bertin Posthill and George L. Miksit for their efforts toward the successful completion of the Memory System.

130

Figure 1. Idealized d-c hysteresis loop.

Figure 2. Test Equipment for toroidal magnetic cores.

Figure 3. Basic magnetic core test

- (a) Applied magnetizing force for basic magnetic core test.
- (b) Typical core output voltage. The switching voltage is e_s ; the non-switching voltage is e_n .
- (c) Magnetic flux or flux density.

Figure 4. Switching and non-switching voltages for 48% nickel iron, $H_m = 1.0$ oersteds.

- (a) Test core Number 1 (1 mil).
 $B_r/B_m = 0.91$
- (b) Test core Number 2 (1 mil).
 $B_r/B_m = 0.93$
- (c) Test core Number 3 (1/2 mil).
 $B_r/B_m = 0.90$

Figure 5. Switching voltages compared for $H_m = 1.0$ oersteds

- (I) 1 mil 48% nickel-iron
(test core Number 1)
- (II) 1/2 mil 48% nickel-iron
(test core Number 3)
- (III) 1/4 mil 4-79 Permalloy
(test core Number 4)

133

Figure 6. The switching voltages of 1/4 mil 4-79 Permalloy (test core Number 4) for three different values of applied magnetizing force.

(I) $H_m = 1.0$ oersteds

(II) $H_m = 0.5$ oersteds

(III) $H_m = 0.25$ oersteds

Figure 7. The switching voltages of the ferrite MF-1118 (test core Number 6) for three different values of applied magnetizing force. The decreased area under waveform III indicates traversal of a minor loop.

(I) $H_m = 4.0$ oersteds

(II) $H_m = 3.0$ oersteds

(III) $H_m = 2.0$ oersteds

134

Figure 8

Core-Diode Memory Array

Figure 9

136

Figure 10

Two-by-two Magnetic Memory, Read-in

Figure 11

Read-in Voltage Wave Forms

138

Figure 12

Two-by-two Magnetic Memory, Read-out

Figure 13

Read-out Voltage Wave Forms

140

Figure 14

Access Time

Figure 15

Static Magnetic Memory for ENIAC

Figure 16

142

143

Figure 17

144

Figure 18

FLIP FLOP 1

LINE DRIVER 3A

CATHODE FOLLOWER 20

Figure 19a

Figure 19b

CRYSTAL UNITS

147

Figure 20

K9C MAGNETIC CORE UNIT
BURROUGHS CORPORATION

148

Figure 21

Ch.

149

Figure 22

COINCIDENT-CURRENT MAGNETIC COMPUTER MEMORY DEVELOPMENTS AT M.I.T.

by

J. W. Forrester

Director, Digital Computer Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

Rather than giving a technical paper, it has been suggested that this symposium might find interesting our development steps in bringing coincident-current magnetic storage to its present successful operation. A second general-purpose digital computer has recently been completed at M.I.T. This machine, which has performance characteristics similar to the Whirlwind I computer, uses a high speed coincident-current magnetic storage and has been operating since May 1953. That machine and the details of its memory system will be reported later in the year by those who were directly responsible for the design and construction. Let's now review steps which led us to the storage system. The steps in themselves may be interesting, and they illustrate again how much time elapses between the idea and the physical realization of a reasonably simple undertaking.

The coordinate selection principle which led to the present magnetic storage originated in 1947¹ around a gas glow discharge cell. In principle, the glow discharge with its high initial breakdown voltage, its low holding voltage, and low forward impedance after breakdown can be used in a coincident-axis selection system. We studied glow discharge storage of this type at the time that the M.I.T. electrostatic storage tube was being started. Glow discharge storage was soon dropped because of the variability of the individual cells with time and from sample to sample. Work on storage with coincident-axis selection was suspended until a suitable basic storage cell became available in 1948.

Rectangular hysteresis loop magnetic materials appeared in magnetic amplifiers in Germany during the war and in this country in magnetic digital stepping registers.² During 1949 the basic principles were developed

---

¹Jay W. Forrester, "Data Storage in Three Dimensions," 13 pages, Project Whirlwind Report No. M-70, the M.I.T. Servomechanisms Laboratory (April 29, 1947).

²An Wang and Way Dong Woo, "Static Magnetic Storage in Delay Lines," Journal of Applied Physics, Vol. 21, No. 1, pages 49-54, January 1950.

151

for using three-dimensional arrays of magnetic storage cells with coincident-current selection.³

In 1950 Mr. William N. Papian, as a thesis research,⁴ studied numerous magnetic materials with respect to the characteristics which would make them suitable for coincident-current memory. He continued by selecting the best available wound-tape metallic core and constructed a 16 by 16 storage array which operated in successful laboratory tests in December 1951 (Figure 1).

With the selection principle demonstrated, the major hurdle became the procurement of suitable core materials to make the magnetic storage usefully fast and sufficiently insensitive to driving current and read-out detector adjustments.

Demands on the magnetic material for this type of storage are more severe than for previous applications of rectangular loop materials. The restriction to two levels of driving current (not fundamental,⁵ but desirable for simplification of associated circuits) means that the material should not respond to one unit of driving current but must reverse its direction of magnetization rapidly when driven by two units of current.

At first the only suitable materials were the metallic cores wound from a few turns of thin steel tape and properly heat treated. Initial experiments used Deltamax.⁶ Deltamax in tape thicknesses of 0.002" required a switching time of 10,000 microseconds in the coincident-current mode of operation. Since 1949 the use of thinner material and materials of higher coercive force has reduced this switching time to about 10 microseconds. The wound tape cores are sensitive to mechanical assembly processes and were difficult to control accurately. Since 1949 the ceramic ferrites have shown greatest ultimate promise as a fast, inexpensive, and reproducible storage cell.

---

³Jay W. Forrester, "Digital Information Storage in Three Dimensions Using Magnetic Cores," Journal of Applied Physics, Vol. 22, No. 1, pages 44-48, January 1951.

⁴William N. Papian, "A Coincident Current Magnetic Memory Unit," Master of Science thesis, M.I.T. Electrical Engineering Department, August 31, 1950.

⁵Robert R. Everett, "Selection Systems for Magnetic Core Storage," 13 pages, Massachusetts Institute of Technology, Servomechanisms Laboratory, Memorandum E-413, August 7, 1951.

⁶Made by Allegheny-Ludlum and distributed by Arnold Engineering Company.

We first began work with ferrites after seeing an article⁷ illustrating the characteristics of materials (Figure 2) made by General Ceramics and Steatite Corporation. This article discussed magnetic ferrites for transformers cores but showed a promising degree of rectangularity in one of the types.

Our first successful operation of a 16 by 16 ferrite core array (Figure 3) was in May 1952 using General Ceramics MF 1118 cores. Improvements have been made steadily in materials since that time and in the production control of magnetic characteristics. Figure 4 shows about the present degree of rectangularity which can be achieved with a useful ferrite material which will switch in one microsecond or less. Figure 5 shows a more rectangular material which, however, requires about 4 microseconds to switch in the coincident-current mode of operation.

It is interesting to note that an inverse relationship exists between switching time and coercive force. The faster materials require the higher driving currents. In this regard, eddy currents in the thinner metallic cores have a minor influence on speed compared to relaxation processes in the material. Metals and ferrites of similar static magnetic characteristics tend to show similar response speeds.⁸ The plausibility of such a relationship can be seen by noting that the magnitude of the available driving force increases with coercive force as shown in Figure 6.

By summer of 1952, coincident -current magnetic memory appeared ready for a full scale test using ferrite cores. Ten months later a general purpose digital computer was operating with a magnetic storage bank. The storage consists of 17 planes, each with a 32 by 32 array of cores. The cores are 0.060" inside diameter and 0.090" outside. They are driven by one-half ampere turn in each axis and switch in one microsecond. Total memory access time, including rewrite, is about 10 microseconds. The system has shown high reliability and there have been no core failures and very little adjustment required.

---

⁷C. L. Snyder, E. Albers-Schoenberg, and H. A. Goldsmith, "Magnetic Ferrites, Core Materials for High Frequencies," Electrical Manufacturing, Vol. 44, No. 6, pages 86-91, December 1949.

⁸J. B. Goodenough and N. Menyuk, "Nucleation of Domains of Reverse Magnetization and Switching Characteristics of Magnetic Materials," 31 pages and 10 figures, Massachusetts Institute of Technology, Lincoln Laboratory, Division 6, Memorandum E-532, March 9, 1953.

Figure 1

Metallic tape cores in a 16 x 16 array

F1443

Figure 2

General Ceramics Materials from Electrical Manufacturing, Dec. 1949

Figure 3

The first 16 x 16 ferrite array to be tested.

156

Figure 4

A typical hysteresis loop shape for a high speed ferrite core

Figure 5

A more rectangular hysteresis loop for a ferrite core with a slower switching time

Figure 6

RELATIONSHIP OF COERCIVE FORCE
TO AVAILABLE DRIVING FORCE

For the hysteresis loop with the smaller coercive force, driving force is limited to D (because of the double coincident selection mode); while for the loop with the larger coercive force, the driving force D' is available to produce rapid domain wall motion.

159

COMPUTER COMPONENTS RESEARCH AT MELLON INSTITUTE*

by

F. A. Schwertz

Mellon Institute
Pittsburgh 13, Pennsylvania

About three years ago we began research on computer components at Mellon Institute on a very small scale. The work was prompted by recognition on the part of our sponsors at the Air Force Cambridge Research Center that the components employed in contemporary computers were not specifically designed for such application and that certain advantages such as increased ruggedness and decreased size, weight, and cost might accrue if they were. Furthermore, since a digital computer comprises a large scale assembly of a small number of basic circuits, it was recognized that these advantages could best be realized by seeking components which could readily be mass-manufactured in multiple arrays.

In trying to gain the advantages enumerated here, there was from the start no thought of engaging in a program of miniaturizing standard types of components. It was thought rather that we should prospect for physical phenomena which could be adapted to the design of components of value in digital computers. For example, Dr. John R. Bowman of Mellon Institute conceived the idea of employing the phenomenon of metallic coherence for digit storage and the phenomenon of electrolytic rectification for switching. The phenomenon of metallic coherence may be readily explained with reference to the idealized coherer depicted in Figure 1. This figure shows a single line of oxide-coated metallic particles, the cross-hatched regions representing the oxide layers. A voltage applied across the line produces the potential gradients shown in the lower part of the diagram. Since the metal particles themselves are highly conducting there is no electrical field inside of them, and the voltage gradients are confined to the oxide layers. At a certain critical voltage the oxide layers break down and, in the process, small metallic bridges are welded between the adjacent particles. The resistance of the metal chain is thus decreased from about 10^9 ohms to about 100 ohms, more or less, depending on the energy dissipated in the coherer. The original resistance may be restored by breaking the chain mechanically.

---

*The research reported in this document has been made possible through the support and sponsorship extended by the Electronics Directorate of the AF Cambridge Research Center, RDC under Contract AF 19/122-376. It is published for technical information only and does not represent the recommendations or conclusions of the sponsoring agency.

The phenomenon of coherence described here could, in principle, be used to design a binary digit store, the high resistance state being used to represent a binary one and the low resistance state a binary zero. It turned out, however, that too much energy and too much time was required to effect a return to the decohered or high resistance state.

The possibility of employing an electrolytic diode rectifier as a switching element was also examined. This device consists of an oxide-coated tantalum or columbium wire serving as one electrode, and a non-corrosive metal electrode serving as the other. The two electrodes are immersed in a suitable electrolyte such as aqueous phosphoric acid. When the oxide-coated electrode is positive, the diode does not conduct appreciably; when the reverse is true, it does. As shown in Figure 2 the voltage-current characteristic of the electrolytic diode rectifier is very similar to that of a germanium crystal rectifier.

Figure 3 indicates the manner in which electrolytic diodes could be fabricated in multiple arrays to produce a complete function switch. This function switch blank consists of two sets of coplanar, parallel wire ribbons cast in an inert plastic material. The first set consists of tantalum (or columbium) and the second set below and at right angles to the first, of a metal inert to the electrolyte such as copper (or gold). The switching elements are formed by drilling small holes through the metal ribbons but not through the entire blank. The electrolyte is filled into the holes, and the exposed tantalum wire is coated with a thin layer of oxide by an anodizing process. The blank shown in the figure is drilled and coded to form a three binary digit line-selection matrix. The interelectrode capacitance of the switching elements is unfortunately quite large and switching rates are accordingly limited to about 10,000 cycles per second.

The coherer and electrolytic diode were described here at some length to illustrate the way in which we liked to go about our work. The phenomenon of metallic coherence interested us because it was potentially adaptable to the problem of digit storage in virtue of its stable hysteresis. The phenomenon of electrolytic rectification interested us because it was potentially adaptable to the problem of switching in virtue of its nonlinear voltage-current characteristic. These observations are important because the whole logical activity of a computer may be regarded as an ordered sequence of switching and storing operations, and if nonlinear phenomena are potentially adaptable to the switching of information pulses and hysteretic phenomena to their storage, then these are the types of phenomena with which we should be concerned. This is not the whole story, however, since the amplitude of the information pulses must be kept above the noise level of gain elements. These gain elements show interesting negative resistance characteristics which can be adapted to the storage of information pulses, and it is indeed instructive to classify the common storage devices into those which exploit hysteretic phenomena and those which employ negative resistance characteristics. This classification is made in Table I.

161

Table I

CLASSIFICATION OF STORAGE DEVICES

<u>Hysteresis</u>	<u>Negative Resistance</u>
Magnetic Drum	Vacuum Tube Flip-Flop
Static Magnetic Register	Single Transistor Flip-Flop
Coincident Current Magnetic Memory	Twin Transistor Flip-Flop
Ferroelectric Sheets	
Williams Storage Tube	Ferroresonant Flip-Flop
Mercury Delay Line	Neon Diode Flip-Flop

By classifying the mercury delay line and the Williams' storage tube under Hysteresis, we take the broad view that any physical phenomenon is hysteretic if the physical effect in question persists after the stimulus has been removed. The storage devices which exploit hysteretic phenomena employ no "holding power" or only intermittent "holding power"; by contrast all of the storage devices employing the phenomenon of negative resistance require continuous "holding power."

The notion that the logical operation of a digital computer involves a complex sequence of switching and storage operations applies not only to the internal workings of the machine but to the input and output units as well. For example, the reduction of output signals to the printed page may certainly be viewed as a form of permanent storage.

We now pass from these general considerations to the work which the Computer Components Fellowship at Mellon Institute has done in the areas of switching and storage. These are summarized in Table II. Some of the phenomena and devices listed in this table have been examined and found unsatisfactory for one reason or another, as for example, the electrolytic diode rectifier for switching and the coherer for storage. Some of them have been examined and found very promising such as the nonlinear semiconductor resistor and the saturable transformer. Others still are in the preliminary experimental stages such as the photoswitch and the optical bistable elements. I should like, however, to say a few brief words about all of them, first to indicate the scope of our work, and secondly, in the hope that some of the phenomena may find application in special problems and devices.

Table II

162

PHENOMENA FOR DIGITAL COMPUTER APPLICATIONSwitching

electrolytic rectification . . . . . electrolytic diode
 gaseous discharge . . . . . cold-cathode gas diode
 barrier resistance . . . . . nonlinear semiconductor resistors
 magnetic saturation . . . . . saturable transformers
 electroluminescence-
   photoconductivity . . . . . photoswitches

Storage

metallic coherence . . . . . coherer
 propagation of magnetic waves . . . . . magnetic delay line
 magnetic hysteresis . . . . . magnetic scrolls
 color hysteresis . . . . . chromotropes
 photoelectricity-
   luminescence . . . . . bistable optical element-vacuum
 electroluminescence-
   photoconductivity . . . . . bistable optical element-solid

SWITCHING

Electrolytic Rectification

The electrolytic diode rectifier already mentioned has two main advantages: it lends itself quite readily to mass production in multiple arrays and it is extremely inexpensive. The unit cost would probably be of the order of a few cents. It is however limited to switching speeds of the order of 10 kilocycles.

Gaseous Discharge

Cold-cathode gas (neon) diodes show both a nonlinear voltage-current characteristic and a negative resistance characteristic as shown on Figure 4. They are therefore adaptable to both switching and storage. We were interested in them for these reasons and also because they are very cheap and probably could be used in multiple arrays by inserting a number of electrode-pairs in the same gas-filled bottle.

Most of our experiments were done with NE-2 neon diodes and, like other experimenters, we found them nonuniform and incapable of switching reliably at rates above 20 kilocycles. In Figure 4 we show a line-selection

163

matrix which employed NE-2 diodes as the switching elements. The output of the line selection matrix was decoded in such a way as to produce a three binary-digit adder. At times the output sum and carry digits could be gated-out at switching rates up to 100 kilocycles; however, over long periods, the operation was unreliable and unsatisfactory.

Nonlinear Semiconductor Resistors

Our work with cold-cathode gas diodes convinced us that the asymmetric resistance of a diode rectifier is not essential to its role as a switching element and that any passive two-terminal network possessing a suitable nonlinear voltage-current characteristic curve may be employed as a switch. For this reason we turned our interest to nonlinear resistors.

I would like to talk at somewhat greater length about the phenomenon of nonlinear resistance since it appears to have some special advantages for switching applications. The phenomenon owes its origin to the electronic states which exist on the surfaces of granular n-type semiconductors. The physical situation is illustrated on Figure 6 where an idealized semiconductor is depicted. We have studied such resistors at considerable length, especially those formed from granular aggregates of silicon carbide, and have derived the following equation which appears to describe their DC characteristics fairly accurately:

$$i = \frac{k A P^{n/m} d^{n-2} V^n}{t^n}$$

In this equation i is the current flowing through the aggregate under the applied potential, V . The cross-sectional area of the resistor is designated by A , its thickness by t , and the average particle size by d . The constants k , m , and n are structure-sensitive and their values depend on the physical-chemical nature of the granular semiconductor.

The AC characteristics of the nonlinear resistors are not so well understood. They do, however, have the peculiarity that both their apparent resistance and capacitance decrease with increasing frequency. Data bearing on this point are shown in Figure 7 which applies to a resistor formed from rubber-bonded silicon carbide. The resistor was $5/8$ " in diameter, about $1/16$ " thick and contained particles averaging 0.02 " in diameter. The data show that the apparent resistance drops by almost an order of magnitude as the frequency is increased from one to 125 kilocycles; the capacitance however decreases only by about 25%.

Nonlinear resistors may be used to replace whole arrays of crystal diodes in certain logical function switches. Where such replacement is possible, considerable savings in fabrication and component costs may be effected because both the nonlinear resistors and the associated connecting busses may be fabricated by applying printed circuit techniques to standard plastic- or ceramic-bonded sheets of semiconductors such as granular silicon carbide.

A "logical and" or "logical or" circuit constitutes perhaps the simplest function switch. Such switches are represented diagrammatically in Figure 8 where crystal diodes are employed as the switching elements. The "logical and" circuit has the property that the output voltage is high only when all of the inputs are high; the "logical or" circuit output is high when any of the input voltages are high. Physically these switches are the same; they differ only in diode polarity and in applied bias voltage. A device which will serve either as a "logical and" or "logical or" circuit is shown in Figure 9. It was fabricated from a rubber-bonded sheet of granular silicon carbide. It contained No. 36GGE grit (average particle diameter approximately 0.02") and was 1/16" thick. Electrodes were placed on the sheet by spraying silver conducting paint through a suitable stencil. The switching resistors were 5/16" in diameter and the load resistor. The switching resistors carried one milliamperere at about 57 volts, and the load resistor the same current at 43 volts.

The output voltage levels for such a device may be readily calculated. Some typical values are shown on Figure 10 for a three-input switch for the case where the index of nonlinearity $n=5$ and the load resistor has twice the diameter of the switching resistor. For comparison the voltage levels obtained from a crystal diode switch are also represented. For a switching voltage of amplitude E the output levels for the "logical and" and circuit are at approximately $E/2$ except when all inputs are high; the output voltage then assumes the value E . For the "logical or" circuit the output levels are approximately $E/2$ except when all inputs are low; the output voltage is then zero. An oscillographic recording of an operating switch is shown in Figure 11. The data apply to the rubber-bonded silicon carbide switch shown in Figure 9 for which $n = 4.7$, a fairly typical value. The "logical and" circuit was driven with the aid of three series-connected binary scalars. The output voltages from the plates of the scalars were used to drive cathode-followers, and these in turn served as low impedance input switches for the "logical and" circuit.

The upper three photographs in Figure 11 show the ungated output of the "logical and" circuit at counting rates of 10,000, 100,000, and 500,000 counts per second. The lower three photographs were obtained from the same pulse trains with the aid of a biased cathode-follower which was employed to separate the output signals from the background noise. It will be seen from the top of the figure that the pulse output waveforms have been considerably deteriorated at a rate of 500,000 counts per second.

The phenomenon of nonlinear resistance has only recently been explored theoretically. With the aid of the theory and new materials now being developed we hope ultimately to design nonlinear resistor switches which will operate at a megacycle rate.

165

Circuits of the type shown in Figure 9 may be conveniently connected in parallel to form, for example, a line-selection matrix. Such a function switch is shown in the photograph in Figure 12. In Figure 13 a decoding matrix has been wired to the line-selection matrix in such a way as to yield a three binary-digit serial adder. In Figure 14 a photographic recording of the output pulses are shown. I must apologize for this figure because it contains results from an adder which was incorrectly designed: the diameter of the load resistors was only half that of the switching resistors rather than the other way around. Also the index of nonlinearity, n , was approximately three rather than the more common value for silicon carbide of 4.5 - 5.0. The switching rate for the output pulses shown in Figure 14 was 100,000 cycles per second.

Magnetic Saturation

The use of magnetic materials displaying square hysteresis loop characteristics for both storing and switching applications is well known and highly developed. For low-level switching applications the hysteretic phenomenon is not essential and any magnetic material showing a high permeability at remanence and low permeability near saturation is suitable for the purpose. The phenomenon in question is illustrated in Figure 15. The information on the figure indicates that the change in magnetic induction, B , near saturation is negligibly small when an incremental field H is added to the magnetizing field. On the other hand when the magnetic material is in the remanent condition a small incremental field will produce a large change in the magnetic induction. This behavior is exploited in the device shown in Figure 16. The input signal passes when the core is not saturated and it is blocked when the core is saturated. In the configuration shown at the top of the figure the control coil does not induce a voltage in either the input or output circuit because the flux divides equally between the halves of the magnetic circuit. Equivalent two-core configurations are shown in the middle and lower portions of the figure.

Because of their high speed of response only ferromagnetic ferrites have been examined for high speed switching applications. Some data obtained using a two-core configuration of Ferramic I are shown on Figure 17. The cores used to obtain these data were 1.78 cms. in mean diameter and had a cross-sectional area of 0.37 cm^2 . There were two signal windings on each core, each comprising 45 turns of No. 32 wire. The control winding contained 89 turns of No. 28 wire and 50 milliamperes were required to yield a biasing or control field of one oersted. The signal pulses were about 2.5 microseconds wide and were roughly triangular in shape. The pulse amplitude was 28 milliamperes and the repetition rate 25 kilocycles per second. The control voltage was turned on and off alternately every four pulse times. A signal to noise ratio of approximately 25:1 was obtained at the maximum value of the biasing field.

A possible application of the phenomenon described here is illustrated in Figure 18. It is equivalent in function to a high-speed relay tree and, like a relay tree, it is bilateral in nature - that is, signals may either be passed from top to bottom of the tree, or in the reverse direction.

Electroluminescence-photoconductivity

We have had the opinion for some time that important advantages are to be gained if photons as well as electrons are used as information carriers in digital data processors. An optical information pulse may readily be transformed into an electrical information pulse and vice versa, and certainly the idea of transferring information pulses through space than through wires is exceedingly attractive.

It is anticipated that optical devices will be useful in both switching and storing applications and a fundamental prerequisite for both applications is a high-speed light source. One phenomenon which appears attractive in this respect is the phenomenon of electroluminescence. This phenomenon can be explained with reference to Figure 19. The electroluminescent light source schematically represented in the figure comprises an intimate mixture of a phosphor and a plastic sandwiched between an evaporated metal electrode and a piece of glass bearing a conductive, transparent coating. The distance between the two electrodes is of the order of a thousandth of an inch. Impressing a voltage gradient of the order of 10^4 to 10^5 volts per cm. across the electrodes causes the phosphor to emit light. Such cells are now available commercially from the Sylvania Electric Products Company, Inc. The pulse characteristics of one of these, measuring 3" x 3", was studied. The results are depicted on Figure 20. The upper curve in each photograph shows the output pulse of a photomultiplier used to detect the output light of the electroluminescent lamp; the lower curve indicates the voltage across the lamp. The pulse width in the lower photograph is about 40 microseconds wide. The variation in light output is seen to follow the rapid variations in voltage. From this limited evidence one can only say that a high speed, broad area light source is available.

A photoconductor, or more generally any two-terminal photosensitive device, in intimate contact with a high-speed light source should produce a high-speed, relay-like switch (or photoswitch). Photoconductors with response and relaxation times of less than a microsecond are known. We are planning to investigate some of these with a view to using them in conjunction with high-speed electroluminescent light sources. This work is just now beginning. We envisage, of course, the simultaneous fabrication, perhaps with the aid of printed circuit techniques, of multiple arrays of photoswitches.

167

STORAGE

We have encountered a number of phenomena which may potentially be developed into high speed storage systems. We have evaluated others and found them inadequate for the purpose. I think a few words about each of them would be of interest.

Metallic Coherence

I have already mentioned briefly the phenomenon of metallic coherence. As pointed out earlier, this phenomenon displays definite and stable hysteresis; however, the time required to effect a transfer from the cohered to the decohered state is prohibitively large.

Propagation of Magnetic Waves

The possibility of storing digital information in a ferromagnetic delay line was experimentally examined. The storage device investigated is shown in Figure 21. It consisted essentially of a long ferromagnetic rod on which were placed two small coils. A current pulse placed on one of the coils magnetized the rod and ultimately induced a voltage in the other coil. The time delay between the input and output pulse was observed on a two-beam oscilloscope.

Apparently the phenomenon in the magnetic delay involves a propagation of the magnetic induction vector. A semi-empirical theory of the phenomenon indicates that the rate of propagation of the magnetic induction as well as the attenuation factor should be independent of the excitation frequency in materials of low conductivity. This was approximately true for the ferromagnetic ceramic material known as Ferramic E. Unfortunately the propagation velocity was of the order of 10^6 cms. per second. This is about an order of magnitude larger than the velocity of sound in solid materials. In addition the attenuation was quite large, that is, about 60 db per meter. For these reasons research on the ferromagnetic delay line was discontinued.

Magnetic Hysteresis

Shortly after we began our work in 1950 we had the idea of exploiting the rectangular hysteresis characteristics of Deltamax in a coincident-current type of magnetic memory. The notions that we had are illustrated in Figure 22. As shown in this figure, a single turn of Deltamax (1.75" long by 1 mm diameter) is wound tightly around, and insulated from a current-carrying (copper) wire. If one assumes that the Deltamax is initially magnetized in a circumferential direction, a properly directed current pulse, either through the wire or the Deltamax itself, will reverse the state of magnetization. Alternately, both the wire and the scroll could carry currents, each of which produces a magnetic field slightly larger than half the field required to reverse the magnetization. This reversal could be sensed by an auxiliary circuit thus providing a basis for memory read-out.

We had done just about enough work on this idea to be convinced that it would work when Dr. Jay W. Forrester's paper on a coincident-current memory appeared in the January 1951 issue of the Journal of Applied Physics. We reluctantly had to conclude that we had been anticipated and further work on this storage device was discontinued.

Color Hysteresis

I would like to say just a word about the phenomenon of color hysteresis. Certain chemicals assume different colors at different temperatures. For example the complex $\text{Cu}_2\text{HgI}_4 - \text{Ag}_2\text{HgI}_4$ (33.9 mol % Cu_2HgI_4) is red at 70°C . If cooled to room temperature it remains red but further cooling to 10°C will cause it to turn yellow. On heating to room temperature the color remains unchanged. The speed of the transition apparently is limited only by the heat capacity of the chemical itself.

There is no thought that this phenomenon can be engineered into a high-speed storage system for computers. It does, however, constitute an erasable ink and as such it may be useful in office-type data-handling systems. The information could be read both by the human eye and by photocells.

Photoelectricity-luminescence

The possibility of storing binary digits by means of a novel diode comprising a photo-emissive cathode and a fluorescent anode is being examined experimentally. Under the application of a steady DC bias of a few hundred volts two stable states exist: an "on" state where electrons flow in one direction and photons in the other. The situation is represented schematically in Figure 23. Transition from the "off" to the "on" stage can be effected by a small pulse of light focussed on the photocathode. Transition from the "on" to the "off" state can be effected in two possible ways. (1) If a high value resistive load is included in the power supply, a strong light pulse momentarily saturates the diode; the electrons then leave insufficient energy to excite the phosphor and flow ceases. (2) Alternatively the photodiodes could be shunted by a photoconductor. Shining light on the photoconductor would lower the potential drop across the diode below the critical level for emission from the phosphor and in this way cause cut-off.

Electroluminescence-photoconductivity

It is possible that a solid state analogue of the photodiode described above could be made with the aid of electroluminescent light sources. Such a cell is represented schematically in Figure 24. In this case the photoconductor plays the role of the photocathode and the electroluminescent light source the role of the luminescent phosphor. Use is again made of internal feedback, the light from the electroluminescent phosphor maintaining the photoconductor in a low resistance state. Read-in can be effected with a luminescent phosphor and read-out by a photoconductor. Transition from the "on" to the "off" state can be effected in a manner analogous to that described above.

CIRCUIT FABRICATION

In addition to the work I have outlined, we have an interest in the fabrication of electronic circuits. For example, we have just begun to examine the Xerographic process with a view to adapting it to the fabrication of printed circuits.* We hope also to make resistive and perhaps more complicated elements with the aid of this technique.

The possibility of assembling "wireless" electronic circuits is being investigated. In principle the method envisages the use of flat sheets of material possessing either resistive, dielectric or magnetic properties and of assembling these, by lamination, into functional electronic circuits. It is conceivable that amplifying devices like ferromagnetic and dielectric amplifiers may be fitted into this scheme.

CONCLUSION

We have been much interested in following developments in the field of high speed digital data processing from a rather special point of view. This view, perhaps not unique, holds that high speed data handling or computation is possible only because electrons are low inertia entities which may be conveniently transferred from place to place. The convective flow of electrons and their associated electromagnetic fields were put to work in the first truly high speed computers. This convective flow was sometimes associated with undesirable side-effects such as the generation of eddy currents in metallic ferromagnetic materials. The invention of ferromagnetic ferrites removed this difficulty and ultimately led to the development of the high-speed coincident current and static magnetic memories.

If it can be said that the discovery of ferrites brought electrons to a complete stop, the discovery of holes in p-type semiconductors put them in reverse. In any case the mutual interaction of holes and electrons has produced a new type of amplifying element and other electronic components. Their mutual annihilation produces photons and, it is our belief, will ultimately yield a practical high-speed light source. The introduction of "holes" into electronic components produced a major revolution in the field of digital data processing. It is our conviction that the availability of a high speed source of photons will eventually effect a similar revolution especially in input-output systems where progress has not kept pace with improvements in switching and storage systems.

---

*This technique was suggested by Mr. C. E. Ryan of the Air Force Cambridge Research Center, Cambridge, Mass.

170

FIGURE 1
A SIMPLIFIED COHERER

171

Figure 2
ASYMMETRIC SWITCHING ELEMENTS

172

Figure 3
A THREE - BINARY INPUT
LINE - SELECTION MATRIX

Figure 4
SYMMETRIC SWITCHING ELEMENTS

FIGURE 5
THREE DIGIT NEON DIODE ADDER

FIGURE 6

Figure 7
SILICON CARBIDE RESISTANCE AND CAPACITANCE AS A FUNCTION OF FREQUENCY AND AMBIENT TEMPERATURE

177

LOGICAL "AND" CIRCUIT

LOGICAL "OR" CIRCUIT

LOGICAL CRYSTAL DIODE CIRCUITS

FIGURE 8

FIGURE 9
LOGICAL NONLINEAR
RESISTOR SWITCH

179

FIGURE 10. QUANTIZATION OF VOLTAGE LEVELS

UNGATED OUTPUTS

GATED OUTPUTS

10,000 counts/sec.

100,000 counts/sec.

500,000 counts/sec.

FIGURE II
OUTPUT OF "LOGICAL AND" CIRCUIT

FRONT FACE

BACK FACE

FIGURE 12 BINARY - TO - OCTAL CONVERTER

182

FIGURE 13 THREE - BINARY - DIGIT ADDER

Sum Digits

0 1 1 0 1 0 0 1

Carry Digits

0 0 0 1 0 1 1 1

15V

Rubber - bonded
Silicon Carbide
Adder

0 1 1 0 1 0 0 1

0 0 0 1 0 1 1 1

8V

Thyrite Resistor
Adder

FIGURE 14. ADDER OUTPUT VOLTAGE PATTERNS (100,000 COUNTS/SEC.)

Figure 15

OPERATING PATHS FOR A HYPOTHETICAL, BUT
TYPICAL, MATERIAL SUITABLE FOR A MAGNETIC GATE

FIGURE 16
DEVELOPMENT OF TWO-CORE CONFIGURATION FROM THREE-LEGGED CORE

186

FIGURE 17 GATED OUTPUTS OF TWO-CORE MAGNETIC GATE FOR DIFFERENT VALUES OF APPLIED GATING FIELD.

Pulse duration 2.5μ sec. Pulse rep. freq. 25 kc.

187

Figure 18
PYRAMID CHANNEL SELECTOR

Signal will appear only at Output Channel 1, assuming the convention that saturating current is flowing from the shaded half of the flip-flop controls. The shaded cores are saturated.

188

FIGURE 19
ELECTROLUMINESCENT SAMPLE

FIGURE 20
PULSE RESPONSE OF ELECTROLUMINESCENT MATERIALS

170
Figure 21 MAGNETIC DELAY LINE

FIGURE 22
MAGNETIC STORAGE

192

FIGURE 23
PHOTOELECTRIC BISTABLE ELEMENT

FIGURE 24 OPTICAL STORAGE CELL

194

CHARACTERISTICS OF THE ORACLE

by

E. W. Burdette

Oak Ridge National Laboratory
Oak Ridge, Tennessee

In December 1950 an agreement was entered into by the Oak Ridge National Laboratory and the Argonne National Laboratory which provided that an electronic digital computer be designed and constructed at ANL for ORNL's use. It was further agreed that Oak Ridge scientific personnel be assigned to ANL to aid in the design and construction of the Oak Ridge computer.

The work under this agreement was to be carried out in two phases: Phase I provided for the development of design and components for the Oak Ridge machine and the completion of the Argonne machine; Phase II provided for the construction and testing of the Oak Ridge machine.

During the development period, which extended from September 1950 to September 1951, investigation and design was conducted on the components of the computer. General logical design followed as a guide the logic discussed by Burks, Goldstine and Von Neumann in "Preliminary Discussion of the Design of an Electronic Computing Instrument." However, the detailed logic together with several new logical approaches was developed at Argonne by the joint Oak Ridge and Argonne staff.

This year of planning allowed time for the Oak Ridge engineers to obtain training in the computer field and at the same time design the computer incorporating the new logic and circuitry.

The design work on the arithmetic unit and electrostatic memory unit ran concurrently; however, more investigation was required on the memory system and construction of the arithmetic unit was begun before the construction of the memory unit. The arithmetic unit was built and tested in eight months and the memory unit was built and tested in nine months. On June 2, 1953 the two units were tied together to operate as a computer.

In September 1952, at the association for computing machinery meeting in Toronto, we announced that the Oak Ridge computer had been named the ORACLE. Previous to this announcement several names had been proposed, all of which were contractions of Oak Ridge and automatic computer. One name was proposed due to a small mishap that occurred during the construction of the computer. A radium spill at the laboratory contaminated certain parts of the machine and the name ORXAC was suggested to mean Oak Ridge Automatic Computer with radiation. The X referred to X-ray. Also ORAC was proposed and this name was tentatively adopted. At the same time a

195

special purpose computer was being installed at Oak Ridge and a sister name was proposed which added two letters onto ORAC to give ORACLE which meant Oak Ridge Automatic Computer for linear equations. Then it was decided to name the general purpose computer ORACLE for two reasons. The Air Force had a computer named OARAC which might be confused with ORAC and the name ORACLE was dropped from the special purpose machine since it will not be used now that the general purpose machine is available. Therefore, ORACLE was changed to mean Oak Ridge Automatic Computer Logical Engine.

The mathematical characteristics of the ORACLE are as follows:

1. It is a binary, parallel type machine with words of 40 bits each.
2. A one address code is used with each 40 bits word containing two 20 bits instructions.
3. Each instruction contains a 10 bits order and a 10 bits address.
4. A fixed binary point is used and a complemented representation of negative numbers is used with -1 permissible and +1 not representable.
5. Multiplication is accomplished in 41 steps by successive addition.
6. Division is accomplished in 41 steps by successive subtraction.

The order code has five basic arithmetic orders with eight types of add orders, two multiply orders, four divide orders, eight right shifts and eight left shifts. There are six basic non-arithmetic orders with ten transfer of control orders, four teletype orders, sixteen magnetic orders and thirty-three transfer of information orders.

The physical arrangement of the ORACLE was designed to have compactness of circuitry as much as possible but good maintenance access dictated the final form. The arithmetic unit has both compactness of circuitry and good access for maintenance. The electrostatic memory has unitized construction allowing accessibility and interchangeability of memory stages. In contrast with computers of IAS design the arithmetic unit and the memory unit have been separated to provide more working area for maintenance.

The engineering characteristics of the ORACLE will be displayed by dividing the computer into seven units: the arithmetic unit, the electrostatic memory unit, the control, input-output, power supply, magnetic tape memory and the console.

Figure 1 shows a block diagram of the arithmetic unit and part of the control. This unit consists of three storage registers, a parallel logical adder, a M to Q selector and four driver chassis. Two of the registers A and Q,

196

are shifting registers and each register contains two banks of 40 flip-flops with transfer gates between the flip-flop banks. The shifting registers use the asymmetrical flip-flop together with a system of clear and transfer such that a shifting operation can occur in 2.5 to 3 microseconds. In all operations the flip-flops are cleared to zero with an overlap time-wise of the transfer pulse and clear pulse. The shifting registers also have an overflow feature that allows easier programming of scaling operations and floating point routines. The third storage register holds the instructions and one of the operands which can be transferred as a number or the complement of the number.

The parallel logical adder is a voltage type adder consisting of gates that follow the logical adder scheme. Each adder stage contains seven double triodes and two double thermionic diodes. The carry time per adder stage is 0.1 microsecond permitting a total carry time for forty stages of 4 microseconds.

The M to Q selector serves as a fast switching and dispatching medium for transfers between the memory and the arithmetic unit. It permits partial substitution from the storage registers and is a direct transfer not requiring a pulse routine of the arithmetic unit.

Drivers for the registers operate the clear and transfer lines and provide a regulated pulse amplitude for all gating signals to insure greater reliability of operation.

Automatic sequencing of machine operations is accomplished by the control. The arithmetic unit has an asynchronous control and the memory unit has a synchronous control. The dynamic programmer, static programmer and shift counter make up the arithmetic control. The dispatch counter, pulsers and mixer make up the memory control. The MOSC (memory order sequence control) is used to synchronize the operation of the arithmetic and memory controls.

The dynamic programmer generates a pulse routine to transfer information in the arithmetic unit after the static programmer decodes the order and sets up the transfer gates. The shift counter is a six-stage scaler type binary counter that keeps track of each arithmetic step and generates a recognition signal at the end of an operation sequence.

The dispatch counter locates the deflection system of the memory at the position specified by the address of the order. It automatically sequences the order pairs and controls the regeneration of the memory.

Ten pulsers provide the pulses necessary to actuate the memory storage and control the operation of the MOSC. The mixer combines various pulses and dispatches them to the memory stages for operation of the memory.

197

Figure 2 summarizes the operational characteristics of the arithmetic unit.

The electrostatic memory consists of forty-three plug-in memory stages, forty in service and three in warm-up, deflection adders and the slave scope. The Williams type storage system is used with a dot dash display. Eighty 3JP1 cathode ray tubes are used with two cathode ray tubes per stage. A time multiplex pulsing system permits identical information to be stored on both tubes of a stage. This method uses a raster of 1024 digits and overcomes screen blemishes which would prevent the storage of a "1." In the event that impurities do not exist on the cathode ray tubes or are not objectionable then the system can store a raster of 1024 digits on each of the eighty tubes to give a 2048 memory.

The stability of the system is greatly increased by using Automatic beam-current stabilization which compensates for drifts that change the amplitude of the ones signal.

The memory cycle is 20 microseconds and at least one regeneration occurs after each action cycle writes or reads information in storage.

Input to the computer at present is a teletype reader which loads the machine at a rate of 24 bits per second. A Ferranti photoelectric reader will be installed next month.

Output is a teletype punch operating at a rate of 80 bits per second. A fast teletype punch will be installed in November.

DC power is taken from five Power Equipment Company rectifiers. Primary power is approximately 50 kw with the computer using 17 kw of AC power and 23 kw of DC power.

The magnetic tape memory is in the design and construction stage. It is due to be installed at Oak Ridge next spring. This auxiliary memory will have four tape drive units with 1000 feet of two inch tape per unit. Each tape will have 42 channels with a packing density of 100 words per inch. Total storage capacity will be 3,600,000 words or 151,200,000 bits. This memory will have a minimum access of six milliseconds.

The console is ready for installation and will be installed next week. All register lights are brought to the console and grouped in a centralized display for inspection by the operator. Adjacent to the register lights is a cathode ray tube monitor for checking any stage of the memory. Switches are provided to operate the teletype input-output equipment. Special orders and corrections can be inserted into the machine via the console. There are also engineering test switches located on the console for trouble shooting purposes.

198

After the arithmetic unit and the memory units were tied together on June 2, the month of June was devoted entirely to trouble shooting, developing test routines, adjusting memory parameters and correcting a few logical and circuit errors. During July work continued on the computer to gain further improvement in the memory and at the same time the mathematicians used the computer for debugging codes and solving problems. Four major problems have been solved, eighteen test routines have been used successfully and several subroutines have been checked and used. The work schedule has been arranged to allow engineering improvement and checking to take place during the regular work week and then the weekends are used exclusively by the mathematicians for problem solution. Ninety-five hours of computer time were used during July for problems and only three hours were lost due to computer failure. Eight hour runs have been made on four different occasions without any adjustments being made on the computer. Three, four and five hour runs were required to solve most of the problems. The exact nature of the problems is classified; however, in general the computer has solved the following:

1. Heat exchange problem for a coolant circulating in a closed system.

839 words in storage   -   339 code   -   500 constants
 7 hours print out
 4 hours computing

2. Gaseous discharge problem.

400 words in storage   -
 1/2 hour print out
 27-1/2 hours computation

3. Determination of characteristics values of matrices.

900 words in storage   -   150 code   -   750 constants
 1 min. print out
 1 hour computation

4. Mathematical problem to test floating point subroutine.

150 words in storage

5. Reactor problem.

20 hours computation

Future plans for the computer require that the ORACLE be installed at Oak Ridge this fall. Figure 3 is an illustration of the computer installation. After the computer is installed and operating satisfactorily work will be undertaken to improve the input-output equipment and develop a translation device for magnetic tape to paper tape.

199

ARITHMETIC UNIT AND CONTROL

Figure 1

200

Figure 2

OAK RIDGE COMPUTER CHARACTERISTICS:

MACHINE TYPE	PARALLEL	
REGISTER CAPACITY	40	
ALLOWED ADDER CARRY TIME	5	μ s.
ALLOWED TIME PER SHIFT	6	μ s.
MULTIPLICATION TIME (eir = 0)	240	μ s.
MULTIPLICATION TIME (eir = $1-2^{-39}$)	440	μ s.
DIVISION TIME	440	μ s.
NUMBER OF DIGITS ASSIGNED TO NUMBER	40	
NUMBER OF DIGITS ASSIGNED TO ORDER	9	
NUMBER OF DIGITS ASSIGNED TO ADDRESS	11	
MEMORY MINOR CYCLE OR MAXIMUM ACCESS TIME	20	μ s.
MEMORY MAJOR CYCLE		
MODE 1024	20.48	m.s.
MODE 2048	40.96	m.s.
NUMBER OF VACUUM TUBES	3000	

201

Figure 3

202

PHYSICAL ASPECTS OF MAGNETIC COMPUTER MATERIALS

by

S. E. Harrison

Remington-Rand, Inc.
Philadelphia, Pennsylvania

Several previous papers in this conference have pointed out the fascinating engineering potentialities of magnetic core memories. This paper deals with the considerations involved in the synthesis of materials for use in these magnetic cores.

INTRODUCTION

The search for new materials and the fundamental understanding of their basic physical properties are very important in the development of magnetic materials for computer use. This search in magnetism may be divided into two themes.

The first broad problem has to do with the intrinsic properties of ferromagnetism, such as the variation of the saturation magnetization with temperature, and the criteria of its occurrence in particular materials. In this latter connection it may be mentioned in passing that the manganites and cobaltites are recently discovered classes of ferromagnetic oxides which promise to supplement the more familiar ferrites as useful magnetic materials. Further developments may be envisaged, in which cubic or pseudo-cubic structures are formed from transition element cations and from anions of the sixth periodic column. The myriad of such possibilities presents a vast field of new compounds to be exploited for computer applications.

The second theme is the variation of the magnetization with magnetic field. This is closely related to final stages of development of magnetic materials for computer use. The variation with field of the magnetization of a ferromagnetic is markedly dependent on structural irregularities. From a physical standpoint it is important to realize that these structural irregularities are of a kind which may equally occur in non-magnetic materials, so that the investigation of the low field behavior is not only interesting and important in itself, but is also contributory toward the explanation of a much broader problem.

203

It is the behavior of magnetic materials in a magnetic field to which this talk will be devoted. If one considers the necessary properties for the use of magnetic components in a memory unit, some of the traits to be desired for the material are:

1. A small output signal ratio for a half to full input pulse,
2. A short switching time,
3. A low coercive force,
4. A high flux density,

and finally, with a full physical understanding of how to achieve these first four aims we will obtain what is a very important practical consideration:

5. A highly uniform end product and simplicity of manufacture.

Under the first characteristic mentioned above, that of small output signal ratio for a full to half input pulse, we consider the problem of obtaining material with a square hysteresis loop; that is, a hysteresis loop with sides as nearly vertical as possible and with the maximum ratio of remanent to saturation induction, (B_r/B_s). In Figure 1, the relevant quantities are defined.

We now examine from a physical viewpoint the several sections of the hysteresis loop in order to determine how we can obtain sufficient insight into the magnetization processes to predict physical properties of materials, and to compare the theoretical predictions with the experimental results. We are concerned with the several mechanisms underlying the complete hysteresis loop and with the role of these mechanisms in the determination of saturation, remanence, coercive force and reversible and irreversible permeabilities. We first review some basic steps in the magnetization process of a ferromagnetic material.

MAGNETIZATION PROCESSES

It is now a well understood fact that a ferromagnetic material consists of an assembly of domains, each of which is magnetized to saturation. In the normal unmagnetized state the vectorial distribution among domains is such that the overall magnetization of the macroscopic volume of material is zero. Because each domain is magnetized to saturation, a change in magnetization can occur only in one of two ways. Either the direction of magnetization in the domain can be changed, so that the vector sum of the magnetization arising from the various domains does not vanish, or domains magnetized in one particular direction can grow in volume at the expense of neighboring domains, which are magnetized in other directions.

In real crystals there are preferred directions of magnetization, closely related to the crystal symmetry, the quantitative measure of this preference is expressed by the crystal anisotropy. In the absence of an

204

external magnetic field, the magnetization in a domain lies along one of the several crystallographically equivalent preferred directions, and varies among these from domain to domain. The application of a field destroys the equivalence of the preferred directions, and influences the magnetization to be along that particular preferred direction which is most nearly parallel to the field. At sufficiently large field strengths the field may dominate the crystal anisotropy, and the magnetization will simply be parallel to the field direction. The orientation of the magnetization is also influenced to a certain extent by prevailing stresses, either homogeneous (such as an applied tension) or statistically distributed (which are always present in any material). The forces resulting from stresses are called magnetoelastic forces.

When a ferromagnetic material is magnetized by an external field, and the magnetization vectors in the domain are drawn toward the field direction, the length of the specimen usually changes. The magnitude of this magnetostrictive effect is expressed by the magnetostriction constant λ , which is the relative change in length in the field direction on magnetizing the specimen from the demagnetized state to saturation. The magnetostriction constant λ can be either positive or negative, depending on whether the sample increases or decreases in length, and is usually of the order of 10^{-6} . Conversely, if a stress is applied the magnetization varies. For example, if a tensile stress is applied to a substance with positive λ the specimen tends to elongate and the magnetization vectors of each domain are drawn toward the direction of the stress. The stress direction is the preferred direction for B_s , no other forces being active. On the other hand, if λ is negative the domain vectors tend to line up perpendicular to the stress direction.

Domains magnetized in different directions are divided by a wall of relatively small thickness compared with the mean domain dimension. The finite thickness of this domain wall, or Bloch wall ($\approx 10^{-6}$ cm in Fe), is due to the lower energy associated with a continuous change in the direction of the spins relative to a discontinuous change, i.e., the elementary atomic magnetic moments in a ferromagnetic prefer to be aligned parallel and, therefore, a gradual rotation of the spins from a direction of easy magnetization in one domain to an easy direction in the neighboring domain will demand less energy. This tendency by itself, however, would cause the domain wall to be infinitely wide. There is a second contribution to the total energy which prevents the infinite wall thickness. When the spins do not lie parallel to an easy axis of magnetization, the atomic moments then have additional energy due to the crystal anisotropy. This increase in energy tends to diminish the number of spins in the wall and competes with the first mechanism, so that the final wall thickness is finite. The crystals of most ferromagnetic materials have cubic symmetry and it is found that the preferred direction of magnetization is either parallel to the cube edges or to the cube diagonals. In general, we have domains whose directions of easy magnetization are either nearly perpendicular or antiparallel to each other and consequently we distinguish between 90° and 180° walls.

205

MAGNETIZATION CURVES

In Figure 2 we have an initial magnetization curve. The lower section of this magnetization curve is a reversible region which corresponds to the reversible movement of 90° and 180° domain walls. The middle region is a result of the irreversible movement of 180° domain walls while the region beyond the knee of the curve is due to the rotations of the magnetization vector into the direction of the field.

In connection with the above mentioned desirability of square hysteresis loops, we see that we would like the value of B_r , the remanent flux density, to approach B_s , the technical saturation. At saturation all of the domains are lined up parallel to the field. As the field decreases from saturation, the domains undergo reversible rotations so that the magnetization vectors will lie along the easy axes of magnetization. The easy axes so chosen will be only those with positive components along the field direction, thus producing only 90° walls. Assuming only 90° walls in this region B_r is roughly a measure of the amount of irreversible magnetization which has taken place during magnetization to technical saturation. Domains of the 180° variety can only be present here if the material contains domains magnetized in the opposite direction. Coming from saturation the magnetization in all domains is parallel and there exist no 180° boundaries until nuclei of antiparallel domains have been produced. The possible sites for such nuclei of reverse magnetization are dislocations, vacancies, grain boundaries, inclusions of non-magnetic material or any portion of the crystal where the homogeneous structure is disturbed.

The initial formation of these nuclei is accompanied by the production of a domain boundary, and the creation of a strong self-demagnetizing field. Both of these make nucleus formation unfavorable because substantial energy is required for either process. It is, however, possible that an irregular demagnetizing field in the neighborhood of the boundaries of the domains, together with free magnetic charge arising from non-uniform domain magnetization, facilitates the formation of nuclei of antiparallel magnetization in individual domains or single crystals. Therefore, bearing this fact in mind, since we desire to avoid the formation of 180° domains as we descend from saturation to remanence on the hysteresis curve, this will lead us to attempt to make materials as homogeneous as possible.

On the basis of no interaction between domains, theoretical calculations for polycrystalline material show that, taking into account only reversible spin rotations, a remanence of .83 (B_r/B_s) is obtained for iron, .86 for nickel and .5 for substances such as cobalt which have only one axis of easy magnetization. In practice B_r varies tremendously depending upon treatment and values lower than .5 are found. For aggregates with preferential orientation we might have any value from zero to one. For materials under applied stress, experimental results are in good agreement with the simple theory which predicts that the remanence B_r approaches one when the magnetostriction constant is greater than zero and B_r approaches zero when the magnetostriction constant is negative.

The remanence value can sometimes be increased by annealing in a magnetic field through the Curie point of the material down to approximately 400°C , provided the Curie point is high enough. The conditions that favor a successful magnetic annealing process are a non-vanishing value of the magnetostriction constant and a relatively small value of the anisotropy constant. In this process, as the spontaneous magnetization increases on cooling below the Curie point, magnetostrictive stresses are set up, which may be relieved if cooling over the plastic range is slow enough. The deformation is such as to make the energy a minimum for the actual direction of magnetization over any uniformly magnetized region, whether this is an entire crystal grain or a smaller magnetized region. When the material has hardened at lower temperatures, these actual directions become relatively easy directions of magnetization of the crystals, compared with others which, in an unstrained crystal, are equivalent. Ordinarily these easy directions of magnetization will be distributed at random among the equivalent easy directions of magnetization of the crystals constituting the polycrystalline aggregate. However, if a field is applied during cooling, the easiest directions will be those nearly parallel to the field.

Another important consideration in the remanence region is the value of the permeability. The calculation of the remanence permeability is essentially similar to that of the calculation of the initial permeability. However, in the demagnetized state we have domains with 180° boundaries which are not present in the remanence state. In iron and nickel, calculations give the result that the remanence permeability is approximately one-third the initial permeability. In the case of large stresses, boundary movements are not important but the rotation of domains are. Therefore, the remanence permeability equals the initial permeability since, except for differences in 180° , domain distribution in the remanent state is equal to that in non-magnetic states.

The concept of a high degree of uniformity can also be shown to be useful if we consider the slope of the sides of the hysteresis loops. If we have a 180° wall in a material that is perfectly uniform, it does not require a large expenditure of energy to move it; therefore, greater wall velocity will then result from more homogeneous materials since the velocity is approximately proportional to the applied field. This will in fact give rise to a shorter switching time. A high flux density will increase the driving force and should also aid in giving us a shorter switching time. However, considerations such as anisotropy and internal stress of materials cause us to make a compromise between the materials with a high flux density and those with other desirable properties.

The coercive field is determined by the ability of the material to nucleate domains of reverse magnetization and the ability of the 180° walls to propagate. So we see that we are led to the position of desiring that reverse domains do not nucleate by their own demagnetizing fields, but by very small applied fields in order that coercive forces may be as low as possible. Once

207

the reverse domains are formed in a starting field, H_s , the field necessary to maintain them can be considerably less than H_s . Here we are faced with the problem of nucleation of these reverse ferromagnetic domains. This is to determine under what conditions and by what processes do these stable nuclei of reverse magnetization arise in a medium which is homogeneous with regard to stress, magnitude and direction of magnetization. The formation and dissolution of domain boundaries are a complex and speculative question, and, generally speaking, not much more than tentative suggestions have been made.

SIMPLE MODEL FOR HYSTERESIS LOOPS

Most of the factors which have been suggested as determining the extent and character of boundary movements can conveniently be indicated with reference to Figure 3. For simplicity, a 180° boundary is considered. The curve in part (a) of the figure represents a possible form for the dependence of the internal free energy of the material on the position of this one boundary. Using the symbol E for the energy per unit area of the boundary and assuming that in zero field the boundary is in equilibrium at $x = 0$, the condition for equilibrium in a field H is

$$\begin{aligned} d/dx (E - 2HI_0s) &= 0 \\ 2HI_0 &= dE/dx \end{aligned}$$

As the field increases the boundary moves reversibly to the right until dE/dx reaches a maximum at b when the equilibrium becomes unstable. Once the value of field $H_c = 1/2I_0 (dE/dx)_{\max}$ is passed, the boundary moves spontaneously without further increase of the field to a new position of equilibrium c , the change from b to c corresponding to a Barkhausen discontinuity,** with further increase of the field reversible movement is resumed. The course of the changes when the field is decreased will be clear from the figure. The diagram (c) shows how the boundary movement contributes to the reversible and irreversible susceptibilities (ranges ab and bc , respectively) in an increasing field, and to the reversible susceptibility again in a decreasing field. The manner in which remanence, coercivity and hysteresis can arise from boundary movements are all illustrated.

**Barkhausen noise does not, as had previously been envisaged, correspond to the discontinuous rotations of whole domains. It results from the irregular fluctuation in the motion of the domain boundary and the corresponding change in the magnetization of parts of a domain. The volumes measures as sources of Barkhausen effect are much less than and have no connection with actual domain volumes, which may be very much larger.

The initial permeability is a measure of the intrinsic restoring force on the wall for small displacements, while the coercive force is a measure of the maximum restoring force. The coercive force tells us the field strength needed to carry a wall from one potential energy valley to another potential energy valley past the highest energy hump.

The assumption of a not unduly specialized dependence of energy on boundary position thus leads to a formal coordination of a wide variety of effects associated with a single elementary process. It is not then unreasonable to suppose that the bulk characteristics of ferromagnetics might arise from the superposition of a large number of similar processes.

INHOMOGENEITIES AND COERCIVE FORCE

A variation of internal energy may arise from deviations from uniformity of the material corresponding to a variation in internal stress or composition, and the theoretical problem is that of linking the two together. Several theoretical attempts to incorporate the concept of inhomogeneities of the material in an explanation of the mechanisms involved in hysteresis loop phenomena have been made. The most important aspect of these treatments has been the calculations of the coercive force.

There have been three distinct lines of approach on the problem of the coercive force which have all contributed somewhat to our understanding. The first treatment is the strain theory.¹ The basic idea of the strain theory is that the variation of energy with boundary position is mainly due to the effect of spatial variation of local stresses on the energy of the boundary itself. The critical field H_C as calculated by this theory is proportional to the maximum stress and the domain boundary width. In this development arguments are advanced that the initial susceptibility is mainly dependent on 90° boundary movements while the coercivity corresponds to an approximate average of the critical fields for 180° boundary movements.

The second theory is the inclusion theory as developed by Kersten.¹ In a ferromagnetic substance containing non-ferromagnetic inclusions a boundary between domains of the matrix may pass through or include within itself impurity particles which do not contribute to the boundary energy. For a given gross area of boundary the energy will depend on the area occupied by the particle, being a maximum when no particles are included. The calculations for this theory show that the coercive force will vary directly as the anisotropy and volume of inclusions and inversely as the saturation magnetization.

---

¹M. Kersten "Probleme der Technischen Magnetisierungskurve," pp 42-72 (R. Becker, ed., Berlin, Springer, 1938).

209

The criticism of the inclusion theory is that the magnetostatic energy associated with inclusions is completely neglected. For some substances, the Kersten inclusion effects are very much less important than the effect of magnetostatic energy. Another fault is the assumption of regular distribution of magnetic voids. The random distribution, which obviously is the one found in nature, gives much lower values of coercive force (approximately of the order .1 to 1 oersted).

The third theory is called the disperse field theory which has been proposed by Neel.² Neel suggests that randomly distributed localized irregularities result in fluctuations in the direction and the magnitude of the intrinsic magnetization. B_s is not strictly uniform even within a single domain and the fluctuation gives rise to non-zero values of the divergence of I , free magnetic poles are developed and corresponding inhomogeneous magnetic field distributions are also obtained. The energy associated with these disperse fields varies with the positions of the domain boundaries. The variable energy may be regarded as arising from the mutual potential energy of free pole islands in regions of positive or negative charge.

Although according to the disperse field theory internal stresses and inclusions may have a negligible direct effect on boundary energy variation, there is no doubt that they are responsible for the diversity of low field properties of different materials, or more specifically, they give rise indirectly to a variation of energy with boundary position. These energy variations caused by the disperse fields may be of the right order of magnitude to determine the experimental results for both hard and soft magnetic materials.

From Neel's calculations in Fe the effect of inclusions are predominant in determining the value of the coercive force while in Ni the effect of stress is most important in this respect.

From what we have said thus far, and we have not included all the possible physical aspects (for example, we have not considered grain orientation) we are immediately aware of the complexity of the problem of predicting and controlling the properties of a magnetic material. Any general conclusions as to the actual success of the prediction of a rectangular hysteresis loop would be premature at this time, but the state of domain theory today makes the solution of this problem a real possibility in the near future.

---

²L. Neel, J. Phys. Radium, 5, 24, (1944).

Fig. 1 Definition of important quantities of the hysteresis loop.

211

Figure 1

Representative magnetization curve, showing the dominant magnetization processes in the different regions of the curve.

2/2

Fig. 3 Simple model for the illustration of hysteresis loop phenomena for a single domain wall (After Stoner and Rhodes, Phil. Mag. 40, 1949).

215
MAGNETIC READING-RECORDING HEAD DESIGN FOR UNIVAC

by

V. J. Porter

Remington-Rand Inc.
Philadelphia, Pennsylvania

ABSTRACT

High-speed multiple-channel tape recording has become a familiar means of digital computer input and output. The problems of designing a head capable of meeting the performance standards of such a tape recording system include:

1. Saturation of the recording medium. Saturation recording results in maximum signal, and hence in maximum signal-to-noise ratio. Furthermore, the emphasis in digital pulse recording is on uniformity of output rather than linearity or fidelity. The head materials must possess a sufficiently high saturation density that they will not be driven to saturation before the tape is. Also the initial core permeability must be high enough to permit the head to deliver a strong signal on playback.
2. Frequency response. The pulse recurrence frequency alone is usually high up in the audio range; the harmonic components of the pulse waveshapes are considerably higher.
3. Interchangeability. Not only must reading and recording operations be possible in the same head, but what is recorded by one head must also be readable on another. Maintaining constant spatial interrelationships between head gaps and tape signals requires establishing rigid standards of manufacture and assembly in making the tape, the head and the transport apparatus.
4. Wear resistance. Since wear alters the dimensions of the head, and so changes the spatial relationship of the gap areas to each other and to other tape records, the heads must be sturdy and wear-resistant. Tape wear must also be held to a minimum.
5. Economy. With all other factors considered, the head must be economical to manufacture. A production process, rather than a custom-machining process, is indicated, even though closer tolerances are possible in the custom process.

Early experiments showed that flat-contact heads were impractical to use with metallic tapes. The superiority of metal tapes resulted, in the Eckert-Mauchly organization, in a decision to concentrate on them, and this

214

was a major factor in the subsequent decision to design the cylindrical contact eight-channel head illustrated in Figure 1. The cylindrical contact surface permits a lowering of applied tensions in the tape path, reducing pressures over the tape-pass surface to a point where, with proper care, tape and head wear are negligible factors. The simplicity of guiding the tape across such a surface aids in maintaining a constant interrelationship between tape and head.

The smaller the radius of the cylindrical configuration, the more efficiently the system functions. However, the fact that metal tapes fatigue when flexed too sharply places a lower limit on this radius. Five-sixteenths of an inch has been chosen as an optimum radius on the head used in the UNIVAC System.

The carcass, or head-block, is made from three pieces of beryllium copper, two sidebars .625" thick and a .125"-thick centerbar. Beryllium copper was chosen because it is nonmagnetic, and because it can be hardened by heat treatment to resist wear. The carcass pieces are gang-slotted to contain the eight unit-heads. Four of the eight heads bridge each of the two junctions. The stacks of four unit-heads are staggered with relation to each other. The eight channels are each .047" wide, with a blank .013"-space between channels ($=.06"$). Thus staggering the gaps permits a total separation of .073" in one direction, and .125", the width of the centerbar, in the other.

(To achieve a .073" separation without staggering the heads would have required a tape at least an inch wide ($.047" + .073" = .12" \times 8 = .96"$), where only a half-inch tape is now used.)

The unit-heads (see Figure 2) are made from laminations of Hymu 80, a nickel-iron alloy whose high initial permeability results in small values of core reluctance, and whose high saturation density permits a fringing flux dense enough to saturate the tape. The laminations are bound together with epoxy resins, whose electrical resistivity inhibits the establishment of eddy currents.

The signal coil is .2" long, with .1" outer diameter. A shorter, thicker coil would have presented greater crosstalk difficulties. The coil's 600 turns are wound onto the C-shaped core piece, and the straight I-piece closes the magnetic circuit. This core configuration permits closer spacing of the heads, and is considerably simpler to wind than a one-piece toroidal coil. The gap in the head is established by a 1-mil thickness of calendared shim paper placed between the core pieces.

The reluctance of the core material, the contact reluctance between head and tape, and the reluctance of the tape material are all kept small by choice of materials, although the tape material is necessarily less permeable than the core material. The reluctance across the back gap, also created by the paper shim, is reduced by doubling the area of apposition at that gap. These four factors are the major considerations of the actual magnetic circuit.

2/5

The reluctance of the recording gap is a compromise; its value must be as high as feasible, but the two ways of increasing this value also seriously affect the efficiency of the system. Making the gap wider reduces the pulse-packing ability of the tape-head system, and reducing the area of apposition, or pole-face area, results in a thinner, less sturdy core piece.

After being wound, the cores are mounted in milled slots in the carcass pieces. The C-shaped core pieces are mounted in the sidebars, complemented by I-pieces mounted in both sides of the centerbar. The C-pieces are potted in place with a minimal-shrinkage varnish. The signal wiring is brought out to a small terminal block mounted inside the head, and thence to a connector socket at the top. Over each side-bar is placed a sheet of shim paper to establish the gaps, and the carcass is then bolted together. Between the unit-heads, thin sheets of Hymu 80 are inserted; by short-circuiting the leakage path, these act as interchannel shields.

After being assembled, the head is finish-ground on a $5/16$ " radius, bringing carcass and exposed core-surfaces into cylindrical conformity. Tape is fed to the head under tension of about twelve ounces, on an approach angle of $27-1/2^\circ$ with the plane tangent to the cylinder at its axis of symmetry.

Any misalignment of the tape relative to the head is accounted for in the reading system. A buffer tape - a $.001$ " plastic spacer which is reeled between the head and the recording tape whenever the tape device is running - reduces both head and tape wear to a negligible amount.

The recording head uses DC erase, and is aided by a simple DC broad-erase head used only during recording. The erase polarity of voltage is always applied to the recording head, and the tape is always saturated to "zero." During recording, specific channels selected by the driving circuits reverse polarity to "one" for a single pulse time (the pulse-envelope system is not used).

Improvement of production techniques and the elimination of some of the more costly components of the head have been more recent subjects of study. The aim is to maintain the electrical standards while reducing the cost and simplifying the manufacturing process. A comparative photograph showing late and early models of the head is given in Figure 3.

FIG. 1
Eight-Channel Cylindrical-Contact Head
used in the UNIVAC System

217

FIG. 2
The Unit-Head

218

FIG. 3
Late (left) and Early (right)
models of UNIVAC head

219

40 CHANNEL MAGNETIC TAPE AUXILIARY MEMORY

by

Leslie Merrill

Argonne National Laboratory
Lemont, Illinois

I find that the previous speaker* saved me a lot of work. When I describe our recording head, you will see that the similarities are very great. I agree with 99% of the requirements, engineering tricks, notes, observations, experience, etc., that Eckert-Mauchly have had, although I must confess that their experience is much more extensive. So I will give a fairly brief description of the auxiliary memory systems as we propose to have them, one for the AVIDAC and one for the ORACLE. After describing the systems, I will discuss for a short while the component parts that we have developed and then discuss auxiliary memory in general.

Figure I is a rough schematic of the auxiliary memory system looking at the front. Items A denote wind-up reels, each capable of holding 1,200 feet of tape. Items B signify the 40-channel recording heads, and Items C denote clutches. Items D are the walls and bottom of the baskets. Item E is a true pair of walls and bottom, while Items D are a series of false bottoms and removable false walls.

The action of the wind-up reels are to be servo controlled by the amount of tape in the basket immediately beneath them. The false bottom of a basket is arranged to swivel up and down slightly with the weight of tape in it. This swiveling action will operate sets of switches and these switches, in turn, will tell a very crude servo system which way to go. The attempt is made at all times to keep the basket roughly half-full of tape so that, as a servo system, it is a fairly easy one. The baskets will hold about 200 feet of tape half full. Thus, the reels have 100 feet of slack either way to catch up with what the fast clutches are attempting to do. The theoretical power input to the reels need not be greater than .01 HP even with a fully loaded reel.

The purpose of the false sides and the true bottom is to enable the programmer to use the tapes in a way most suitable for the problem under computation. Typical examples of various tape lash-ups are shown in Figure II. A denotes the basic form of operation involving use of spliced loops or the reels. B shows how more than one reading station may be used on

---

*Vincent J. Porter, Remington-Rand.

either a loop or reels. C illustrates the use of all four stations on a single closed loop. Since one master plane surface governs the positioning of the elements shown in Figure I, it is possible to construct complex combinations of the type shown in Figure II without distortion of the tape. When I discuss the purpose of the auxiliary memory in general, I think the reasons for these features will come clear. It does give the programmer a good deal of flexibility.

The fast clutches, and the reason for their speed will also be explained, are made by arranging to have continuously driven rollers squeeze the tape between them, thus pulling the tape along. The direction of operation is away from the recording head. Tension in the tape to wrap it around the head is supplied by some simple friction drags. (At the demonstration, they did not look so simple and they were not working very well.)

Figure III is the first basket drive that we constructed. It did everything we expected of it, except for one item. The small bearings on the inner rollers didn't stand up. The solenoids, which activate the clutches, can develop a thrust well over 100 lbs. without overheating and the little bearings, as I remember it, only had a rating of 10 lbs. at the speed we were running, although their static rating was 50 lbs. each. The usual happened. The pair of round caps you see are end caps on the bearings of big rollers which are continuously driven by the motor. Immediately beneath these big rollers are small rollers approximately $1/4$ " in diameter, of which you can see the small end bearings. The tape is next, and underneath the tape is another set of rollers of exactly the same $1/4$ " dimension. Push rods drop from the bearings of the bottom little rollers. The push rods drop down to equalizer bars and thence down to major thrust rods to the plungers of the solenoids which are at the very bottom.

The purpose of the three-roll system is not obvious at first sight. The first attempt at pulling a piece of tape with pinching rollers and using frequent start and stops led to a great deal of damage to the recording surface. The purpose of the little rollers is to remove the friction buffing from the tape. The coefficients of friction are such that the small rollers stay with the tape under all conditions including even an accidental jam. All slippage and buffing takes place between the top little roller and the big driven roller.

Figure IV is the second model we built, just completed, designed primarily to overcome the bearing trouble which we had in the first one. This is the one you saw trying to go in the demonstration. I am sorry the photograph is not larger. It is difficult to see, but the bearings now are much, much larger located near the top of the clutches. The use of large bearings made impractical the use of equally large friction rollers because the high inertia would slow down the starting and stopping action. Hence, we could not use friction buffers in the form of rollers. In this model,

both rollers are quite large, continuously driven and surrounded by friction buffing sleeves. These sleeves are in the form of fairly close fitting cylinders and are about as thin as practical. The coefficient of frictions are such that, at the initial instant of contact, the buffing is between the inside of either sleeve and its roller and not between the sleeve and the tape. This unit is not debugged or tested, but it is expected to yield substantially the same starting and stopping features of our first unit. The first unit had a pick-up to full speed or a complete stop in approximately 3 milliseconds from signal for either operation. The corresponding distances on the tape were $1/8''$ to start, $1/16''$ to stop. By the way, the stop feature is a function of the friction drags. They are twofold in purpose, that is, both to provide tension on the tape and to stop the tape when the clutches let go.

The solenoids are made from standard filament transformers. These happen to be Stancor 6-ampere transformers, but it turns out, from a magnetic design standpoint, that any filament transformer makes a pretty efficient solenoid for developing large thrusts over short throws. The transformers are disassembled and reassembled so that all the E's point one way, and the bars assembled together to form a plunger. One finds that most, almost 90%, of the area in the break is iron and only about 10% is coil. It takes over 250 pounds to pull apart the fully closed halves of this particular solenoid with 10 watts dissipation in the coil. This is a little above manufacturers' recommendations. We do not, of course, use anything like the 250 lbs. for one reason, i.e., we cannot allow the plunger to come fully in contact with the solenoid. Since we put finished ground surfaces on the halves, the resultant closed gap is very close to absolute zero. For such a case, eddy currents can hold a flux in the core when you turn the juice off (even in a laminated system) for a considerable time afterwards, depending on what flux you are talking about. We made some measurements and found that it was very important not only to laminate but also not to allow the plunger to come in contact with the solenoid. In our system, there is a .006" plastic shim placed between the two halves at the bottom. The normally open position of the solenoid is .012" further open than this. The roller system, however, moves only .006" because there is a shock absorbing spring built into the main push rod. The purpose of the shock spring is to allow a relatively thick tape splice to pass between the rollers without having to move the heavy plunger.

Figure V is a direct view of the business end of the prototype recording head. Down inside it bears some remarkable similarities to the one made for the UNIVAC by Remington Rand. The procedures in construction, of course, are different. Our approach to the wear problem is different and it has never been considered that we would seriously use a metal tape. We are sticking to plastic tape. There are two features which I would like to point out that have governed the over-all design of the recording head. One of them is, as you have heard before, the complete interchangeability of the heads, both from the standpoint that more than one head may be operating on a

common section of tape, and that, over a period of months and years, we wish to be able to recover the contents of the auxiliary memory. This means that if a head burns out or wears out in the meantime, we wish to be able to produce more just like it. Therefore, one sets up a standard gap geometry. The tolerances we are trying to hold are $\pm .001$ " in the positioning of the gaps along the length of the tape. This is approximately 0.1 of a pulse length at the packing density we propose to use. The tolerance along the width of the tape (channel to channel) is more like .002". The main guide for the tape is right on the head; the next view will show this quite clearly. This piece functions with a dual purpose of positioning the parts inside and guiding the tape, although I do not mean to imply that we use no other guides.

As for wear, it has been estimated this head will cost us in production (our production only amounting to eight or ten heads) about two thousand, five hundred dollars. Therefore, they should have a reasonable half-life. We approach the standpoint of wear by the use of lavite inserts. Lavite is quite similar to steatite: a hard, very abrasive resistant ceramic. The white things you see in the surface of the head are the lavite inserts. This head was used for fifty hours and checked under the microscope and with indicators. We are very confident that it will be good for several hundred hours.

We have attempted to make the head in a three-piece construction (Figure VI) such that the pieces can be reworked so as to renew the head in case that it should ever wear out. This includes maintaining the track geometry to agree with our set up standard. The whole idea depends on having integral units that can be set up on surface grinders. The assembly of the many component parts is, therefore, done by potting. We use Hysol's standard clear amber potting compound and it works very well. This little line of squares you see is the bottom of the "C," using the terminology of the previous speaker. It is a very good term. We call it the pole piece but I think "C" is probably better. We call this the center piece, by the way. At the top you see the internal facing of the top edge of the "C." The coils are not too visible simply because they are under the plastic now, and, although the plastic is clear, it does not photograph well, but the coils run up and down the length of "C." We use 450 turns of number 42 wire. The little lines are the edges of some semi-circular shield plates which duck inside and come back out. The shield plates do not come up to the top of the head and do not touch the recording tape. This is the center piece; I can't recollect the term used by the previous speaker. It contains the stacks of laminations which abut one-to-one against the pole pieces over here. The distance between the two lines of gaps is governed by the finished ground width of the center piece. Once this is finished ground, we never attempt to grind it again, or obviously it would not fit our standards. But what we can do is take material out of the top of the surface against the tape without changing the distance between gaps. In the head bodies containing the pole pieces or "C's," the positioning of these laterally is governed by slots in the bakelite which forms the main holding medium. Here again, we try to

223

keep everything square so that grinding straight into inside and top surfaces will not affect the geometry of the exposed parts. There is depth enough in the plastic over the coils for four such operations at approximately .008" or .009" each. This, combined with the fact that we think the heads will be good for several hundred hours, means that we have hopes that they will be good for several thousand, at least 1,000.

The gaps are formed by chrome plate. This plate is on the center piece and this I definitely can see doesn't show. The top quarter inch of the inner surface of the lamination stacks are individually plated with .0005" of chrome. We are attempting currently to do this for the twenty stacks simultaneously. This chrome plate makes the gap as hard as anything else in the system and substantially harder than most of the other parts. We can get clean gaps no matter how the finishing stone is used. It is very difficult for the tape to ruin them. They simply stay in their shape no matter what the wear is. The main idea behind it originally was to prevent smudging of mu-metal into mu-metal across the gap, and the chrome has served remarkably well to divide the two halves and form a nice recording gap.

Very roughly, the electrical characteristics of the heads are: the coils measure 25 ohms, 15 millihenrys; saturation of the tape is 95% complete on the tape at 10 mils; using saturation recording (non-return-to-zero) at a tape speed of 75" per second, the signal output is 25 mv peak-to-peak. The signal pulses are well resolved at 150 pulses per inch but we do not intend to operate that closely. The output from coil-to-coil is constant to within 15% and, as I will show later, this is important because we cannot use an individual gain control on each head. From the same standpoint, the coils should have somewhat the same saturation current and they are within 5% in this respect. The cross-talk level found by measuring a pulse all by itself on a channel and then erasing that channel and then putting pulses on all the remaining channels is about 5%. In DB, this is something on the order of 25 DB signal to cross-talk ratio. We have made no attempt to measure the 60-cycle pick up, tape noise, and other stuff on DC recording, but I would guess that it is somewhere between 35 and 40 DB down from peak signal.

Figure VII is another view of the recording head, giving perhaps a better feel for the way the parts fit together. The top pieces, which I said served to guide the tape, also, have undercut surfaces here which position the three parts of the recording head in place before any surface stoning is done. We do not radius grind the top of the head. We merely take the corner off and let the tape do the rest. Set screws, a whole series of them, on either end press the system together. The fine wires of the coils go to eye-lets at the bottom of the head bodies and thence to heavy twisted pairs, which terminate in an Amphenol 100-contact plug. This is merely a bracket to hold the plug. Fastening to the master plate of the tape drive is by two screw holes on the face away from us in the picture.

224

Figure VIII is the test equipment. In a very complex undertaking, it is probably unwise to do too much at once, i.e., to try to trouble-shoot a computer and an auxiliary memory system involving perhaps six or seven hundred tubes at the same time. Therefore, the decision was made very early that a test rig would be needed. It has been built twice. This is the final version. It is capable of checking any forty-channel tape drive up to a pulse rate of about 50 kc, in addition to checking the control channels. By the way, our head is a 42-channel head with two sprockets at the center. The test rig is capable of putting the tape drive through its paces and checking the information as it comes back, bit by bit, digit by digit continuously. In addition to this, we can use it to check the tapes themselves after its life as a tape drive tester is over. We believe this is going to be necessary because, in spite of the fact that the quality of the tapes we get from our manufacturers has been excellent, there are phony spots on the tape every so often, perhaps one in a hundred feet. By using the test rig to find this out, we can avoid laying down any sprocket pulse, thus letting the computer skip over the bad spot. We think this is going to be necessary.

That concludes what I think I should say on the equipment. I would be very happy to talk to any of you later on the details and experiences.

Our auxiliary memory system is designed with the sole purpose in being of serving as an auxiliary memory. Absolutely no attention was paid to input-output features. This one concept has led to several general broad engineering features. I would like to define the auxiliary memory. Quite simply, it is merely the extension of an internal memory that is available to the computer, for the computer, and by the computer. Since the information is written by the computer, access is not intended by any other means. Later on, perhaps, we will try to make it dual as input-output by providing equipment capable of reading the computer's records and transcribing to teletype tape or other type storage that, at the current state of the art, are assumed to be more permanent.

Most of the so-called auxiliary memory systems that I have heard about to date really dual for both purposes, input-output and auxiliary memory. The reason we divorce the two concepts arises from the nature of the problems which the computers will handle. In most cases, the input-output will be relatively negligible. If I might, I can give a fairly typical example of that. You are probably all familiar with the fact people play with quantum-mechanics and like to compute wave functions. A good deal of time has been spent computing wave functions on various machines and probably a good deal of time will be spent in the future. Consider an extremely large problem, for example: trying to compute functions for all the electrons in a heavy atom. This is, of course, quite far-fetched at the present day. The problem will be extremely large, requiring immense storage, but not much input-output, simply because the wave functions, while interesting, are not the basic answer. Operations on the wave functions by other operators result

225

in single valued, or at the most, few valued, answers which need to be printed and which are of primary interest. For example, suppose we do have a million values ψ_j of a wave function in a region. This is not what we basically want. We want operations for example $\sum \sqrt{j} \psi_j^2 / \sum \psi_j^2$ where $\sqrt{j}$ is the radius vector from a point in the region. This is a single answer. It is typical of what you run into, I think, in physics. There is, of course, a whole category of statistical problems which might be thought to involve large input-output, but here the statistics themselves or the individual elements of a statistic would be generated by means of equations or logic, and hence, small input, while the desired answer would be means and averages, etc., and, hence, negligible output. Dr. Mauchly's problem of taking over the weather data is, of course, completely different. There the data has not been produced by a computer and has to be gotten into it. This is a function of input-output.

Now, if an auxiliary system is to be an extension of the internal memory, its only purpose is to extend the range of the problem, the scope of the problem which can be put on a machine. If I might just write down a relation -

$$N_c t_a < T$$

where N_c is the number of auxiliary memory consultation, t_a is the average access time for these consultations, and T is time spent in computing a problem. This relation states that the time spent consulting the memory is less than the time it takes to compute the problem, certainly, since some computations must take place. Now, from what I have seen of what is going on in the coding and programming groups, (things like making the machine do its own coding) I do not think that there is too much doubt that it would be desirable, at least from the standpoint of the mathematicians, that the number of consultations increase without limit - and it's getting high already. From the relation above, if $N_c \rightarrow \infty$, and T is bounded, then t_a must be approaching zero. So the first thing that comes about is a short access time, saying nothing at all about storage capacity, just from the standpoint of the scope of the problem you wish to handle. Now, turning the argument around, if engineers provide the mathematicians with a short access time, it does not necessarily follow that the scope of the problem is going to go up indefinitely because just from the relation itself, you cannot say that because $t_a \rightarrow 0$, $N_c \rightarrow \infty$. That does not have to hold. From the practical standpoint, time must be spent in computation over and above the consulting time, so there is a point of diminishing returns. Where should we stop? Where does it become impractical or uneconomical to push the access time down? The answer is where the order of magnitude of the average access time per consultation is equal to the average time spent in computation to use the results of a consultation. For the AVIDAC and ORACLE, this could be 10 microseconds in a simple addition problem involving add times, or it could be 200, 300, 500, or a fraction of a millisecond. This is the basic reason for all the hard work put into the forty channels. The number of channels is about the only parameter left to play with to get more speed out of the tape drive. It happens

226

that forty channels is convenient from a gating standpoint since the machine is a forty-bit machine. Eighty channels, with double-gating to the fast memory, is not inconceivable. The internal memory speed is even up to the possibility of 120-channel tapes.

The other big thing about the average access time in a tape system is that you have got to use the tape serially in order to be able to use its best feature, namely, that when it is moving, it is pouring words out in quite a hurry. In our case, we are intending to run at 200 microseconds between words on a straight out basis. Aside from basic speed, all the other features are there for flexibility in programming. It is possible for the programmers to program so that locally on each recording head the tape is used in a serial fashion. If not, it is something so close to it on an average basis that, in a large scale problem, it is off by not more than a factor of two. These features are: we are providing for the possibility of loops - we are providing for more than one head on a tape - and we are demanding that a single transfer between memories be not less in extent than 32 whole words. This is done in order that the fast clutches exhibit some sort of decent efficiency during start-stops. Fast as they are, 32 words is a mighty short time. Of course, many problems would be better set 64, 128, etc., but 32, we feel, is an operable minimum that gives the coder plenty of leeway. The fact that, because he only wants a single number at the moment, he has to drag 31 other numbers along with it is not going to cause a serious loss of the high-speed memory. Now, this figure of 32 is no magic number. Actually, it arises from what we thought could do in the way of a clutch and it turns out that a block of 32 is about as long as it is going to take us to start and stop the tape, so that the efficiency will be about 50%, operating at blocks of 32, and rises up to 100% utilization of the tape as the blocks become longer and longer, up to perhaps 512 or 1024.

I should perhaps deal a little with the capacity of the system, the actual storage. In the above discussion, N_c represented numbers of consultations. Of course, this is probably a much larger number than the actual number of storage elements needed. Nevertheless, even though the number of consultations in the problem may be thousands of times as large as the number of storage elements required, the basic capacity needed is already more than we can comfortably put on one piece of tape. Thus, the governing factor in how much should a tape be able to store is how often do we feel we want to be changing tapes. We have chosen 1,200 feet. We find that the 1,200 feet matches well with the speed of the tape drives and that, in the worst possible situation, where a straight-through consultation of a piece of tape is made, the time spent could be as short as three minutes. While a change of tape every three minutes would keep people hopping, in nearly every other case, it probably will turn out to be more like a change every fifteen minutes, half an hour, maybe once every several hours.

227

Another feature that I have not discussed, which we are going to build into the system, is to hunt while computing. In this connection, we are also pulling one more trick and this is a trick against the access time but in favor of the engineering. The four recording heads give a total of 160 coils and it runs roughly 5-6 tubes per coil to handle the system. This is quite a high tube count, so we are going to tie our recording coils in series, channel by channel, and provide clutch interlocks so that the tape moves under one station at a time. Whichever station has tape moving under it will print the record on the tape in the case of driving the heads and will furnish signal to the amplifiers in case of reading back. This means we suffer a good deal of loss in hunting efficiency, because we cannot hunt on more than one station at a time. Also, it means that if you are going to hunt while you compute and you run out of stuff to do, you have to interrupt the hunt before you can consult any other reading station. However, we save a factor of four in the reading and writing equipment, tube counts, etc., necessary to set the system going. The concept of coils in series is what I meant when I said that we cannot put gain controls on each of the recording heads (nor match a driver to them) simply because we have to deal within groups of four coils at a time. Within these four, they simply have to match to within some sort of tolerances or other.

This is about all I have to say, but before closing, I would like to acknowledge the contributions made by Messrs. Al Crowley, Howard Edfeldt, Edward Lach and Herb Ross of the Argonne shops and, in addition, the ideas and help contributed by other people in our Computer Group and Engineering Department. Lastly, I would like to extend my thanks to the Argonne National Laboratory for the availability of some of the finest facilities in the country.

f. 130

Figure 1

(A)

(B)

(C)

Figure 2

Figure 3

231

Figure 4

232

Figure 5

233

Figure 6

234

Figure 7

235

Figure 8

236

TAPE SPROCKETING DEVICE

by

Arthur Ostroff

Underwood Corporation
Long Island City 6, New York

About two years ago, the logic for the Elecom 100, the first general purpose digital computer marketed by the Electronic Computer Division of the Underwood Corporation, was for all intents and purposes frozen. The philosophy of logic set forth by Dr. Samuel Lubkin did not depend on extensive checking circuits but rather upon a small number of reliable computer components. The lack of any type of check circuits suggests a multitude of difficulties associated with the input-output equipment, namely that of the magnetic tape. The straight use of commercial magnetic tape without any sort of an inspection or pre-examination was immediately tossed off as unsatisfactory.

The following discussion relates to a method of examining commercial magnetic tape prior to its use in a digital computer. This examination was deemed necessary because of past performance records of digital machines now in use, which suggests the need for greater reliability in the input-output devices of such machines. The equipment to be described is termed a tape sprocketer. The first model was designed and constructed specifically for the Elecom 100. However, in principle it may be adapted for any computer requiring digital storage on magnetic tape.

By way of review, the use of the magnetic recording tape for the reproduction of speech and music presents no serious problems save that of faithful reproduction and the maintenance of a relatively uniform output level. Bad areas of tape, which may result in lower than normal playback amplitudes are of little consequence in such applications, unless the affected length of tape is relatively great. Present production methods for the manufacture of sound recording tape almost preclude this possibility. In attempting to apply sound recording tape to the handling of digital data, however, it has been found that the rigid reliability standards for this application were not, in general, met. Defects which would be insignificant for sound recording can easily cause loss of pulses when handling digital data. Reliability has been improved considerably by rigid production control and testing, but the resulting "instrumentation" tape is still found to have occasional defects which cannot be tolerated in a digital system.

In brief, the equipment under discussion acts as a "go-no-go-gauge" for the actual areas of tape to be used by the computer instead of seeking generally poor areas which would be sensed by the computer by some electrical or mechanical marking on the magnetic tape.

237

Specifically, the equipment records on a reel of magnetic tape a continuous train of pulses closely related in width and intensity to those which would be recorded by the computer itself in its use of the tape. The device then proceeds to examine this recording for playback amplitude. In this way, a measure of reliability of the tape taken in exactly the manner of use is achieved. The net result of the process is to record a series of pulses or sprockets along one channel of multi-channel tape only where acceptable corresponding regions have been observed on all channels across the tape. This is shown in Figure 1.

In actual usage the logic of the associated computer is designed so that all digit transfers to the tape are initiated by the reading of the sprocket into the machine. The sprocket channel in effect being read by the central computer as an indication of acceptable tape locations in addition to its function as a synchronizing signal. The details of the system referred to above will now be investigated.

We feel that this type of acceptance testing for tape has two primary advantages. First, it does not require stoppage of the tape to physically remove or otherwise mark bad areas. Secondly, the tape is not examined in a macroscopic manner, but rather by specific areas and the usable tape confined to these areas. Hence, there is no possibility of recording in an unexamined region. The first model of this tape sprocketer required 15 passes of the tape in alternate directions. The tape is examined at an approximate rate of six feet per second, so that a 1200 foot reel is completely processed in about one hour.

In the design of any type of equipment to accomplish the previously mentioned results, three basic questions must be answered. First, we must ask ourselves what type of defects will we encounter. Second, we must decide upon the allowable deviation in playback amplitude which will be accepted or tolerated by the computer. Third, we must come to some sort of economical compromise so far as time and equipment are concerned to decide upon a type of testing. Obviously, no type of processing, no matter how accurate, can account for damage subsequent to the examination procedure. Therefore, the usual type of precautions in the use of magnetic tape are to be followed.

It is believed that the most common type of difficulty encountered in the manufacture of magnetic tape is the inclusion of small amounts of foreign matter during the coating process. This results in an elevation on the surface of the tape, smaller than the size of a pinhead, surrounded by a circular depression, much resembling a volcanic crater. During the playback procedure the effect of this elevation is to temporarily lift the tape away from the reading head, resulting in weakened playback areas on either side of this elevation. The examination process to be described detects such shortcomings in the coating of magnetic tape. Another type of difficulty to be encountered with magnetically coated surfaces is one of a discontinuous emulsion. Tiny breaks in the magnetic coating would have two effects. On one hand, they would cause

noise to be present in the case of a saturated surface or DC erase system. On the other hand, they would probably result in the reduction playback amplitude of a normal pulse recorded in the general area of the discontinuity.

Experimental evidence has shown that the noise resulting from minute discontinuities such as pin holes result in noise amplitudes of approximately 18% to 25% of average playback amplitude. This level is low enough so as not to cause any serious difficulties inasmuch as a clipping level is usually set well above 20% of average playback amplitude.

The question of acceptance level has many facets and each one must be examined carefully in order to attain a reasonably economical design. On the one hand, we are interested in economy of tape. From this point of view alone it seems desirable to set a rather low acceptance level. However, if this is carried to an extreme, the conditions imposed upon the amplifiers within the central computer would be such as to make the pre-examination almost completely unnecessary. Conversely, if the question of acceptance level is approached from the economy of tape-reading-amplifiers in the central computer alone, a large amount of tape rejection may result. Let us review this problem in a little greater detail. Presumably the signal as read off the tape after having been amplified in high-gain linear amplifier will be used to set some sort of a trigger device such as a flip-flop. There is a certain level below which this digit-input flip-flop will never set. Also, there is an input level above which this digit-input flip-flop will always set. The difference between these two levels represents a region of uncertainty. To assure any degree of reliability in transferring from the tape to the digit-input flip-flops the setting signals to these flip-flops should be considerably greater in amplitude than this region of uncertainty for the smallest signal to be encountered from the tape. In the Elecom 100, this region of uncertainty is approximately 3 volts, that is, from -3 to 0 volts. The discrimination level was therefore chosen to be -30 volts. This means that a signal of at least 30 volts is required to set the digit-input flip-flop.

The amplifier gains are such that a normal playback signal will have an amplitude of approximately 90 volts from base line to peak. This means that the tape reading amplifier equipment in the central computer has a dynamic range or margin of 3 to 1 and a noise rejection ratio of almost 2 to 1 if we consider that the noise level will be approximately 20% of normal playback signal amplitude. Reduction in clipping or discrimination level would have been toying with danger inasmuch as noise would soon become a problem. Let me hasten to add that the dynamic range of the amplifiers in the central computer should not be made any greater than is necessary to achieve over-all system economy. We therefore established that the acceptance level within the computer itself would be 30% . It is now necessary to set the acceptance level for the tape well above this figure of 30% , probably 2 or 3 times as high. It was chosen to be 80 %.

239

The only question to be answered now was one of tape economy. If most of the tape resulted in playback signals of greater than 80%, the overall system would result in one of economical design. It is believed that the number of playback signals reaching a specific amplitude follow a distribution of the form shown in Figure 2. Actually, the curve shown is of exaggerated form since it indicates a rather wide distribution of signal amplitudes. In reality, almost the entire area under the curve is located about the A_2 amplitude. With nearly none of the signals having an amplitude greater than this, the noise curve (not shown) is assumed to drop to zero amplitude at about 25% of A_2 .

The third question to be considered is the type of examination. It is at this point that the discussion becomes highly specific since reference is made to the Elecom 100. The conversion of synchronous information within the machine to a recording on tape involves a $\frac{1}{2}$ to $1\frac{1}{2}$ minor cycle delay. The width of the recording was approximately 500 microseconds. A minor cycle corresponds to 250 microseconds. Therefore, there is an area of tape corresponding to 750 microseconds at the specific tape speed that was used on which the recording could take place. The head gap dispersion and tape skewing raise this figure to approximately 1,000 microseconds. A cell space on the tape corresponds to approximately 2 milliseconds. Therefore, it is necessary to examine approximately one-half of the linear dimension of the tape. The actual processor was run at 10 times of the machine's tape speed in order to avoid the rather unsavory prospect of a 10-hour examination. Therefore, the final system resolved itself into recording a square wave of 5 kilocycles on a tape running at 75 inches per second, which was just 10 times the machine's tape speed, and examining every pulse. It must be remembered that information is only contained on the leading and trailing edges of the current pulse, inasmuch as the playback is the derivative of the flux pattern on the tape. Since this is true, only a fraction of the pulse area can be sampled during a single reading operation provided we are not near the limit of resolution. To place the process in the direction of safety, we chose to examine the central region 2 and the two end regions 1 and 3, which actually extend beyond the area in question. This is shown more clearly in Figure 3.

One might justifiably question the entire method of examination herein described inasmuch as it appears that a much simpler way of accomplishing this task would be to merely record at a high enough density so that the entire area of the tape is examined instead of concerning ourselves with the examination of discrete areas in which the recording for a particular machine would take place. The reason for not doing this in the first processor was that it would require a sampling or examination of the playback signal in the presence of recording on another channel. This was impossible utilizing only one 6-channel head. At present we are planning a multi-head processor which does not require reading and writing simultaneously in the same head assembly, and plans call for this high density examination.

5-10

If the amplitude of a playback signal reaches the acceptance level a pulse is simultaneously recorded in other channels, in such a position so as to allow the examination of the next region in question in those channels on which the recording was made. The magnetic coupling between channels at the head requires a 2-head separation between reading and writing channels to prevent blocking of the reading amplifiers. Consequently, the process is divided into two sequences, each sequence of which processes three channels. In sequence 1, channels 1, 4, and 5 are processed. In sequence 2, channels 6, 2, and 3 are processed. This provides a 2-head separation between reading and recording heads at all times. Obviously, the process can be simplified by the design of a reading amplifier which is incapable of blocking, such as a DC coupled amplifier. The previous slide indicated that in order to examine all 3 regions a delay of $\frac{1}{4}$ pulse between reading and recording may be necessary. Furthermore, in order to simplify the sequence a polarity reversal is suggested on playback. Since each sequence involves the examination of 3 regions of each of 2 groups, 6 passes of the tape are required for each sequence. In addition to the 2 sequences of 6 passes each there is an initial recording, a transfer and a rewind test, making a total of 15 passes. The entire sequence of operation is shown in Figure 4. The arrows denote the direction of tape movement. Polarity reversal switches are so chosen such that the shaded lobe of the playback signal corresponds to the region being examined on that pass which are the regions indicated on the top of the figure.

A blocked diagram of the entire logic for the processor is shown in Figure 5. The triangles are amplifiers with the following polarity convention:

1. Positive input causes positive output at apex of triangle and negative output off side of triangle.
2. The triangle with a line through the center is a DC amplifier.

Semicircles containing straight through lines are buffers or logical "or" circuits. Empty semicircles are gates or logical "and" circuits. Relay 9 is assumed to be up for the examination of the A group (Ch. 1 or 6) and down for the examination of the B group (Ch. 3 and 4 or Ch. 2 and 5).

An additional advantage to be derived from this type of processing is that one may resort to the use of magnetic tapes which would otherwise be unacceptable. For example, it is known that acetate base tapes are highly hygroscopic and do not display high mechanical stability. Mylar base tapes, a du Pont product, are superior to acetate base tapes from the standpoint of physical stability, but very much inferior from the standpoint of electrical quality. A digital examination as described here today permits us to resort to tapes which are not in themselves acceptable for use in digital computers without some sort of processing. We frankly admit at this point that we have run no tests which give conclusive evidence that the system of sprocketed tape is substantially better than a system employing instrument tape together with some sort of checking circuits.

341

In defense of this system, I might add that the two machines in the field at the present time operate with extreme reliability, especially in regard to the magnetic tape units. New computer development at Elcom calls for the use of sprocketed tape, exclusively, since it seems to offer a rather simple solution to what can be a very nasty problem under certain circumstances.

242

Fig.1*

X denotes examined regions

Figure 1

Figure 2

Figure 3

245

Figure 4

245

ELECOM 100
TAPE SPROCKETING DEVICE
LOGICAL

ELECOM 100
TAPE SPROCKETING DEVICE
MECHANICAL

NOTE:
Tape speed = 72 %sec
Pulse density = 5000/72 ≈ 69 r.p.m.
Pulse Space = 200 μsec

Figure 5

247

A PRELIMINARY REPORT ON THE PROTOTYPE DIODE-CAPACITOR MEMORY

by

Arthur W. Holt

National Bureau of Standards
Washington, D. C.

INTRODUCTION

This paper describes the general construction and early operation of a large prototype model of a very rapid access memory using diodes and capacitors.¹ This type of memory has the following characteristic features: simplicity, high access rate, and the absence of any points of low power level. The order of simplicity is such that it is possible to design equipment on paper, without extensive laboratory experimentation; such simplicity also leads to reliability and easy maintenance.

The access rate of the present model is 83,000 randomly selected words per second (8 bits per word at the present time) and may be modified to give 125,000 words per second; output level of the memory proper is one volt at 200 ohms impedance. Although the initial cost of such a system is higher at present than for other high-speed systems (mainly due to the requirement of two diodes per bit), it is possible that the expense may be much reduced when silicon junction diodes come into volume production.

BASIC CIRCUIT

The basic circuit consists of two diodes (either semi-conductor or thermionic) connected in series, with one end of a capacitor connected to the midpoint. The capacitor may be charged either positive or negative by driving the appropriate diode; this charge may then be retained by biasing the diodes in their back direction. Since the back resistance of the diodes is not infinite, the storage state must then be read and rewritten at regular intervals. The storage cells are arranged in rows and columns such that a single access driver can activate an entire row, which can then be used as a group to store a word.

The access driver used is a new type of two-input coincidence gate called the "transformer and-gate."¹ This circuit is a single diode connected in series with the primary winding of an ordinary pulse transformer; one input is connected to the free end of the diode, and the other input is connected to the free end of the primary winding. These inputs are of opposite

248

pulse polarity and are so biased that current flows through the diode and transformer primary only when both inputs are activated. Output is obtained from any desired number of secondary windings. When numbers of these circuits are combined, there results a matrix switch in which all power supplied to the matrix is delivered to the desired outputs. There is no standby power, and no power is wasted on unselected circuits.

THE CONSTRUCTION OF THE MODEL B PROTOTYPE

A list of the basic characteristics of the prototype equipment is given in Table 1. Although it may seem that there are rather a large number of ways to operate the model, this considerable flexibility has proved to be useful. (Table 1). By way of note, the access rate when used with SEAC* is limited by the speed of the arithmetic unit; when in use with SEAC, Model B takes over eight positions normally used for Williams memory. In further note, the "repetition-rate" tolerance check can be extended to 24,000 microseconds by throwing the switch which gives alternate regeneration and reference addresses.

A front view photograph of the present prototype is shown in Figure 1. At the top is the one-megacycle sine-wave "clock" which synchronizes the circuitry, together with a group of switches used in the various modes of operation. The next six rows of base connectors represent the entire control equipment and "first level" selection matrix; etched circuit packages, following the SEAC² pulse technique and the DYSEAC³ construction method, are plugged into the back of these connectors. The first two rows of this circuitry are devoted to a fairly sophisticated control system which allows a number of different modes of operation, either in conjunction with the computer or in independent laboratory operation. Of particular interest is the technique used for checking the memory cells, which is based on the fact that the length of time a cell can store information correctly between regenerations is an absolute figure of merit for the cell. The switches at the top of the cabinet, then, merely lengthen the period between regenerations according to the desired test. This is called the "repetition-rate" check.

The third row of the DYSEAC-type circuitry contains an eight-stage counter which provides sequential address for regeneration. The output of this counter is fed into an eight-stage address register, occupying the fourth row. This register can accept (in parallel) either the regeneration address, supplied by the counter, or a reference address, which is the address at which the computer desires to read or write information. When the computer is calling for references at a maximum available rate, the regenerations and references are interlaced in a one-to-one ratio: the first six-microsecond period is occupied by a reference cycle, the next by a regeneration, that following by a reference cycle, and so forth. Any reference period not required by the computer is used for regeneration.

---

*Standards Eastern Automatic Computer

249

Outputs of the address register drive the last two rows of DYSEAC-type circuitry which are arranged to form what is called the "First-Level Matrix Selector." These are two sets of 16 packages (standard except for higher-inductance transformers) wired so that each has a four-input diode and-gate driving the 6AN5 amplifier tube. The first four address digits thus activate one out of 16 of these packages, and the last four address digits activate one of the other set of 16 packages.

With the aid of the transformer and-gate the final step is accomplished in the conversion of an eight-digit address to the selection of one out of 256 words. This circuit, which consists of a diode and transformer, is packaged in an octal base, 128 of which may be seen plugged into the bottom chassis shown in Figure 1. Total elapsed time from the activation of the address register until the desired word is selected is 0.6 microsecond, only 0.1 microsecond being due to the transformer and-gate. It is worthwhile to note two points: first, that the transformer and-gate is suitable for use in very much larger matrices, such as 8192; and, second, that the entire conversion from binary to final selection can be accomplished by using several levels of transformer and-gates, interlevel amplifiers not being needed since the gating efficiency is so high.

The "Memory Proper," which is the assembly of diodes and storage capacitors plugged into the bottom half of the cabinet, is packaged in groups of eight bits. These eight bits represent a word in the Model B prototype. Figure 2 shows one side of this package (each side contains four bits). Although the matrix selector has 256 output possibilities, positions for only 128 packages have been constructed, and only 43 are filled in Figure 1.

The back view of the Model B cabinet shown in Figure 3, indicates first that the equipment could have been housed in a good deal smaller cabinet. Starting from the top, as before, the back side of the clock chassis may be seen, followed by the six rows of DYSEAC-type packages. These packages, of which there are only two types (tube package and delay line package) are shown in closeup view in Figure 4. The flexibility of these units is demonstrated not only in this equipment, but in the fact that the entire central core of the DYSEAC computer, with the exception of the acoustic memory is constructed of only these two types.

In the lower half of the cabinet may be seen the wiring side of the Memory Proper chassis and also the eight gating amplifiers, four on each side. All of the equipment in the lower half of the cabinet can be conveniently built using etched circuit techniques, as will be discussed in a later section. (Indeed, even the clock circuit can be duplicated using somewhat modified etched packages.) Details of the gating-amplifiers may be seen in Figure 5. Three 6AN5 tubes and about 15 germanium diodes are used in the circuit.

ASSEMBLY AND OPERATION

The assembly and operational record of this equipment has supported the premise that a high-speed memory of this type would be easy to build for reliable operation. Model B was demonstrated on June 5, 12 days after the final assembly was begun, and for the next month was operated in a non-scheduled manner while final mechanical work was performed. In the middle of July it was disassembled and moved downstairs to SEAC, being out of commission seven hours. No faults were introduced by the move. A July 19 trial with SEAC was unsuccessful because too much crosstalk was experienced in the 25-foot unshielded connecting wires. On July 26, Model B was moved closer to SEAC and connected with shorter wires; a continuous run of 12 hours was then accomplished without error. The program was one in which continuous writing, reading, and checking were employed. Three hundred fifty-two (352) storage cells (704 diodes) were involved in the test. The equipment was not electrostatically shielded in any way; three of the DC voltages and all of the AC power sources were unregulated.

From August 5 to September 23, 1953*, Model B was operated on an eight-hour per day "shakedown" schedule. During this period a total of 111 hours of error-free operation was obtained, and preventive maintenance techniques were evaluated. The main difficulties were with defective wiring in the one-megacycle sine-wave clock and a tube in the clock which was an intermittent short circuit. Difficulties caused by the standard etched-circuit packages have not been serious. In one case, an intermittent short in a delay line package was traced to the fact that the package had been improperly assembled; in several other cases, diodes in the packages have been replaced when the preventive maintenance check showed that they were out of specification, although they were not causing operational errors.

No difficulties at all have been encountered with the transformer-and-gates, the gating amplifiers, and the Memory Proper. Table 2 shows the specifications of the various types of storage diodes which are being seriously considered for this application. (Table 2). Diodes are removed when the cell storage time decreases to less than 3,000 microseconds as determined by the "rep-rate" check; the necessary storage time in "standard" laboratory operation is 1,500 microseconds. (Table 2). The equation given in Reference 1 for the storage time as a function of the back resistances of the diodes was developed from the assumption that the back characteristics of the diodes could be represented by a resistance; experimentally, it is found that the diodes are more nearly constant current in the range of 1 to 5 volts than constant resistance. Based on this assumption the lower bound of the storage time for a cell would be the time required to discharge

---

*The following operation information, although not available at the time of the talk, is presented here for the sake of completeness.

251

the capacitor by a constant current equal to the back current specification of a single diode with 5 volts applied. Excellent calculations of the storage time bounds can be made without difficulty.

On September 24 the equipment was put on a 24-hour per day schedule, one or two hours each morning being used for "rep-rate" check and general preventive maintenance. Although the length of time of operation is still too short to be statistically important, it is encouraging to note that there were no errors during 160 hours of storage of 488 bits for the week of September 24 to September 30.

ETCHED CIRCUIT CONSTRUCTION

While the mechanical design of the present Memory Proper was suitable for small experimental equipment, it was felt that there was much to be gained by using specialized techniques. To this end, an etched circuit design has been developed by this laboratory in conjunction with the L. J. Kamm Co. This circuit includes both the transformer and-gates and the Memory Proper. Figure 6 shows a front view of one of the drawers in position in the rack. Each drawer contains all the wiring for eight complete words of 48 bits each including gates; components are dip-soldered into place. The rack is 17" high, 24" wide, and 13" deep, and contains positions for eight drawers. Thus, the rack shown has a capacity of 64 words of 48 bits each. The conductors on the back panel are etched on phenolic: 48 of these conductors are the connections to the 48 gating amplifiers; six of these carry selection pulses to the gates and two carry DC voltages to the gates.

Figure 7 shows a top view of the drawer. The center configuration is the wiring for the gates, while the broad stripes are the word buses. The 54 spring connectors are husky and wiping, and have plenty of over-travel and tolerance. The bottom side of the drawer as pictured in Figure 8 shows the gating amplifier buses (24 on each side), the six gating pulse buses, and the two DC buses. The broadened points along the gating-amplifier buses allow the use of either axial or tangential leads on the capacitors.

The sample panel shown in Figure 9 illustrates how the components are placed. The capacitors stand on end, flanked on either side by the two diodes; the junction of the diodes and capacitor is unsupported, thus eliminating any possible external leakage path. The diodes are connected to the broad word buses, shown on this side of the panel. The capacitors, on the other hand, go through the panel and are connected to their appropriate buses on the other side.

In the transformer gate assembly on the Figure 9 photograph are two type 1N56 diodes (only one is visible), one bypass capacitor, and one damping resistor. After insertion of components, the entire panel is dip-soldered from the bottom side. All holes are through-plated.

Several other systems involving various levels of plug-in parts were considered in detail; the present design was chosen for the model because of its basic simplicity. The dimensions chosen for this model were predicated on accepting the largest components that might conceivably be used. Assuming the use of type 1821 silicon junction diodes and smaller capacitors, it is likely that the volume can be cut to one-half of that of the present model.

The mechanical support of the upper end of the capacitor has not been given specific solution in this model because a number of different methods appear plausible. The most probable method of support at this time would appear to be the use of strips of sponge rubber between the columns of capacitors.

Turning briefly to the question of suitable packaging for future gating amplifiers, it presently appears that the most likely method is to use the same type of mechanical construction as is used in the standard DYSEAC type computing package. This will make for fewer mechanical techniques. In any case, a highly specialized type of construction does not gain very much, as only 30 to 50 gating amplifiers (a number equal to the number of bits per word) are needed per complete memory system.

CONCLUSION

Although the scheduled operation time of the Model B prototype has been far too short to provide meaningful data on the reliability of this type of memory, the construction and assembly of this equipment have shown:

1. that calculated design agrees with final result;
2. that no adjustments or tinkering are necessary during final assembly or operation;
3. that simple, nearly automatic, methods of maintenance for the memory proper are effective;
4. that the components can take a lot of physical abuse.

The work with the etched circuit model has shown that the equipment can be packaged economically and that the entire structure can be made rugged, compact, and light weight. If future application warrants it, the technique can be used to provide even higher access rates.

ACKNOWLEDGEMENTS

The engineering design of this equipment was accomplished by E. R. Toense. The construction was done for the most part by J. Lane, P. Lebowitz, R. Senator, and F. Woolson. The author gratefully acknowledges many valuable discussions with Dr. R. J. Slutz. The work was sponsored in part by the Wright Air Development Center.

253

Table 1

CHARACTERISTICS OF THE MODEL B PROTOTYPE MEMORY

Mode: Parallel

No. of possible word positions: 128, expandable to 256

No. of bits per word: 8, expandable to 48

No. of words actually operating: about 60 (September, 1953)

Maximum no. of random references: 83,000 words/sec. in Lab mode
20,000 words/sec. in SEAC mode

Basic cycle: 0.25 microsec. read; 1.50 microsec. write (or rewrite);
4.25 microsec. passive

Minimum time between basic cycles: 6 microsec.

Cycle during regeneration only: one regeneration every 6 microsec.

Cycle during maximum references: in Lab mode - regen., ref., regen.,
ref., etc.
in SEAC mode - regen., regen., regen.,
regen., regen., regen.,
regen., reference cy-
cle takes 48 microsec.

"Standard operation" in lab mode: regen. only; 6 microsec. between basic cycles; time allowed for 256 words to be regenerated; period between regens. of any one word is 6 times 256 or about 1,500 microsec.

"Repetition Rate" tolerance check: By means of switches, the time between basic cycles may be extended from 6 microsec. to 12, 24, or 48 microsecs., thus giving possible periods between regeneration of 1,500, 3,000, 6,000, and 12,000 microsec.

"Neurotic Write" Test: In this test, if the gating-amplifier senses a "one," it writes a "zero," and vice versa. This is the hardest writing test.

Table 2

DIODE SPECIFICATIONS AND REMOVALS FROM
AUGUST 5 TO SEPTEMBER 23, 1953

Manufacturer	A	B	C
Type	1N67 germanium whisker	1N67 germanium whisker	Silicon junction
Initial acceptance specification	Standard spec; 1 meg. between -5 and -50 V.; not checked before installation 4 ma at 1 V.	Selected to: 2 meg. at -5 V.; 4 ma. at 1 V.	Selected to: 100,000 meg. at -10 V.; 5 ma. at 1 V.
Total hours of operation	235	235	235
No. of diodes	414	224	32
No. removed for failure to hold 3,000 microsec. during "rep-rate" check	14	0	0

255

REFERENCES

1. A. W. Holt: "An Experimental Rapid Access Memory Using Diodes and Capacitors," Proc. ACM, Toronto Meeting, September, 1952.
2. S. Greenwald, R. Haueter, and S. N. Alexander: "SEAC," Proc. IRE, October, 1953.
3. R. D. Elbourn and R. P. Witt: "Circuitry Techniques Used in NBS Computers," Proc. IRE, October, 1953.

256

Figure 1

257

Figure 2

258

Figure 3

259

FRONT

BACK

FRONT

BACK

DELAY LINE PACKAGE
OPENED

TUBE PACKAGE
OPENED

Figure 4

260

Figure 5

276-A

261

Figure 6

Figure 7

263

Figure 8

207-

Figure 9

278A

265

SOME TOPICS IN EXPERIMENTAL MATHEMATICS

by

N. Metropolis

Los Alamos Scientific Laboratory
Los Alamos, New Mexico

The advent of general purpose, electronic computers, such as the two we have seen at this Symposium, has stimulated new approaches to many problems in the theoretical branches of the natural sciences as well as in pure mathematics.

This development might be referred to as experimental mathematics. Such a description may at first seem almost paradoxical, since mathematics is considered, after all, the most exact and theoretical of the sciences. However, the description may appear appropriate when one reflects that, although our knowledge in pure and applied mathematics is vast and the experience is great, there is still an infinity of unexplored terrain. Some of this terrain, while it may not be concerned with the profound problems of mathematics, is usually of a complicated nature; e.g., there are familiar examples in combinatorics and probability theory, astronomy, meteorology, and biology, not to mention physics and chemistry. Therefore heuristic studies are of great value.

The motivation of experimental mathematics is quite different from that of programmatic studies. In the latter, the underlying principles are presumed to be known. The detailed interactions of various kinds may, however, be quite involved. One is concerned here with a systematic study of a variety of numerical cases, where some educated guesses have been made about the initial conditions or boundary values. The objective is usually to optimize a particular variable or parameter, or a set of such quantities.

On the other hand, in experimental mathematics, the emphasis is more on trying to understand and to formulate the underlying principles at play. One possible approach is to calculate for the given problem a few preliminary examples based on some relatively simplified theory. This initial set of computations may already lead to a critical examination of the heuristic assumptions; i. e., whether they are adequate pro tem, or if not, the results may indicate the direction of change for the hypotheses. With success at the modest level, one may then attempt a more ambitious program. Although some of the methods are common to both types of research, the one differs radically in spirit from the other.

266

I should like to describe a few simple examples that we have studied at times when we were not supposed to be working on deadline commitments. The various problems are concerned with pure mathematics, meson physics, statistical mechanics and astrophysics. They are all in relatively early stages of development. Perforce, the discussion will be very sketchy.

Middle squares - The general class of Monte Carlo problems makes use of large sets of random numbers corresponding to various distribution functions. It is quite easy to get such sets from a collection of uniformly distributed random numbers. The latter may be obtained in a variety of ways. One method, suggested by J. von Neumann, that has been used extensively by us is the so-called "middle square" scheme. It is an iterative procedure.

Consider x to be an n (even) digit number (either decimal, binary, or other base). Form x^2 and extract the middle n digits; this is the first iterate. Successive iterates are obtained by repetition of this process.

It is quite clear that once the numbers, 0 or $2^{n/2}$, appear in the chain of iterates, they will merely reproduce themselves. It may happen that successive application of the iterative scheme produces a number that has already occurred at some earlier point in the chain. We may describe this as a "loop."

Some interesting questions suggest themselves:

1. Starting with some arbitrary number, what chain length on the average is obtained before a loop is reached.
2. What is the average size of the loops.
3. What is the expected number of distinct loops.

We have studied in detail a few binary cases for small n . Some of the preliminary results are interesting, and perhaps surprising. Most numbers eventually terminate at 0 or $2^{n/2}$, about equally divided. There are only a few chains distinct from these two, and such chains are relatively short. All numbers less than $2^{n/2}$ eventually go to 0. There are also a few "isolated" numbers, that generate themselves by this iteration process but are never generated by any other number; i.e., they form a chain of length one.

A more sophisticated question of randomness was posed by S. Ulam.¹ We merely cite the reference here.

---

¹cf. N. Metropolis and S. Ulam, "A Property of Randomness of an Arithmetical Function," Am. Math. Monthly, Vol. LX, No. 4, 252 (1953).

247

At the practical level, it has been possible to predict long (binary) chains for the case $n = 38$. Indeed, the first choice led to a chain some 750,000 in length before terminating at 0. Some preliminary statistical analyses of the "randomness" of this aggregate have been made and are quite satisfactory. These will be reported in detail elsewhere.

Meson Scattering by Hydrogen - Fermi and his collaborators at The University of Chicago have produced mesons with their new synchrocyclotron. These nuclear particles are, of course, of great interest to physicists and it is quite natural to study their interaction with the simplest of nuclei, namely hydrogen. There are at least two kinds of mesons produced, namely positive and negative mesons. It is possible to collimate a beam of such mesons and study the intensity of scattering by a hydrogen target at various angles. This has been done for three angles. Unlike the positive mesons which suffer pure scattering only, the negative mesons can, in addition, exchange charge with the protons, with the result that a neutron and a neutral meson are produced, i.e.,

$$\pi^- + P \rightarrow \pi^0 + N.$$

In other words, for each of the three angles, there are three experimental data, corresponding to pure scattering of positive and negative mesons, and charge exchange by negative mesons. Present day ideas of nuclear physics can interpret these nine data in terms of six parameters called phase shift angles.

In principle, one could use six experimental data and determine the phase angles by solving a system of six non-linear equations. However, because of sizeable experimental errors, Fermi has been interested in a least squares fit of all nine experimental values. Mathematically, the problem is to minimize a function in a six-dimensional space, and the question arises how this can be done with a minimum of computing. It is not at all obvious that one should apply "steepest descent" techniques because, computing-wise, it may require frequent computation of the steepest direction, and such computations may be relatively lengthy. Another possible procedure is to select the steepest of the coordinate directions and minimize first with respect to that coordinate, again re-examine for the steepest coordinate direction, etc. Still a third possibility is to minimize successively along the various coordinate directions. This latter method involves no preliminary investigation. It may be mentioned that all these methods lead to relative minima only. There is also the question of what is the optimum step to take irrespective of the particular method used.

We have tried the first and the third methods described above and for the particular function at hand they were about equally good. A relatively coarse step was used at first, and when a relative minimum was

reached the process was continued with a somewhat smaller step. A more detailed discussion is given elsewhere.²

The general problem of finding the minima for non-linear functions of n variables is of considerable interest and occurs in many applications. It would be desirable to make further experiments.

Equation of State of Individually Interacting Molecules - The methods of statistical mechanics tell us that if we wish to evaluate some property F of a system of particles, we must average that value over all possible states. This averaging process is essentially an n -fold integration, where n is very large--of the order of the number of particles in the system. The particular problem investigated was an equation of state of individually interacting particles. Mathematically, this is a relation between the pressure P , the volume V , and the temperature T . It may be written as

$$PV + D(\rho_0) = kT$$

where k is the Boltzmann constant and in the present problem $F = D(\rho)$ is essentially the density distribution function of particles surrounding a given particle; ρ is the distance between two particles. Our interest is in the value of D at the surface ($\rho = \rho_0$) of a particle.

We have attempted to find out how well a statistical approach using a simplified model agrees with more elaborate theories. With electrostatic storage (we did not have available the magnetic drum when the present investigation was started), it was necessary to limit our system to 224 hard spheres, initially arranged in a trigonal two-dimensional lattice (almost close-packed). The computational problem consisted of two steps. In the first, each particle was moved a distance determined by a uniformly distributed random number, i. e.,

$$\begin{aligned} x^{(1)} &= x^{(0)} + \alpha_1 \xi \\ y^{(1)} &= y^{(0)} + \alpha_2 \xi \end{aligned}$$

where α is a random number between 0 and 1, ξ is the maximum allowed displacement, the superscript (0) refers to the original position and (1) to the new value. If a displacement leads to an overlap of two particles, that move is forbidden and the displaced particle is reset to its old value. The maximum displacement ξ per move must be chosen judiciously, for if it is too large, there will be many forbidden moves. On the other hand, if ξ is too small, the system will change so little that a prohibitive number of cycles will be necessary.

---

²E. Fermi and N. Metropolis, Los Alamos Unclassified Report, LA-1492, (1952).

269

When all particles have been processed, the second step then determines a provisional distribution function for the distance between all possible pairs. The cycle is repeated until the distribution has stabilized. If one has a potential of interaction, the overlapping criterion is replaced by a statistical treatment to determine whether a given move is to be allowed.

This statistical approach to the case of an idealized two-dimensional liquid gives results in quite good agreement with an analytical treatment.³ This gives stimulus to try the three-dimensional case including a potential, where the more conventional approaches are not so tractable.

An Astrophysical Problem - Astronomers have long been interested in calculating the equilibrium configuration of a star. The problem amounts to a numerical integration of a differential equation coupled with several ordinary equations. Usually several integrations are made before one has a solution which is consistent with the boundary conditions. Because of the tremendous amount of desk computing involved, astronomers have considerably simplified the models. With large scale computers, it is quite straightforward to consider more realistically the relevant opacity relations, the energy production within a star, and the various equations of state.

In collaboration with R. Williamson and R. Richtmyer, the author is now attempting to set up a quite general formulation of this problem. It is hoped that one will learn what are the critical or sensitive factors, so that improvements may be focussed on them. If this is successful, it is planned to study typical stars in the main sequence.

Concluding Remarks - There are certain developments in modern electronic computers which when accomplished will make the computer a tool of direct interest to the theoreticians, so that they would be willing to press the buttons and throw the switches. It seems to me that among the more obvious factors, the more important may be:

1. to have simple control panels, keeping those required for maintenance and testing separate;
2. to have some convenient display of the progress of a calculation, exhibiting perhaps one or more of the relevant parameters in some sort of counters;
3. to have easy access to the control, so that the course of calculation may be changed from time to time if desired. This is part of the philosophy that computer and operator should supplement each other, each doing the part that is relatively easier;

---

³ N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. 21, 1087 (1953).

270

4. to simplify as much as possible the task of problem preparation. An extensive subroutine library, relative coding and other techniques will assist greatly.

Finally, one may say that with the technical progress that has been made in recent years with electronic computers, the outlook for the understanding of such topics as turbulence and other cooperative phenomena is less pessimistic.

SURVEY OF FERROELECTRICS

by

B. T. Mathias

Bell Telephone Laboratories
Murray Hill, New Jersey

Ferroelectricity is the electric analog to Ferromagnetism.

From a phenomenological viewpoint this name is justified as the macroscopic dielectric behavior of Ferroelectrics shows many of those features that have been found for ferromagnetic materials. The three outstanding common traits are:

- a. A Curie temperature below which a spontaneous polarization exists.
- b. Dielectric constants (as compared to permeabilities) that vary between 10^2 and 10^5 .
- c. Hysteresis between polarization and applied field.

From a microscopic viewpoint, however, there is no connection between Ferromagnetism and Ferroelectricity. Whereas in the first case the spin-spin exchange is responsible, the Coulomb forces of electric charges bring about a spontaneous electric polarization in the latter. It is therefore very unlikely that any element will ever be found to be ferroelectric and that this phenomenon will be restricted to compounds of quite different atoms.

The known Ferroelectrics to-day can be classified in three groups:

- I. The tartrates: Rochelle salt (ferroelectric between -20°C and $+24^\circ\text{C}$)

- II. The tetragonal acid phosphates and tartrates

The Curie points vary between 90°K and 220°K .

III. Double and single oxides of the 4th, 5th and 6th column of the periodic system.

BaTiO_3 , PbTiO_3 , $(\text{Ba,Pb})\text{ZrO}_3$, LiNbO_3 , NaNbO_3 , *

KNbO_3 , $\text{Cd}_2\text{Nb}_2\text{O}_7$, LiTaO_3 , NaTaO_3 , * KTaO_3 , WO_3 .

*possibly antiferroelectric

The Curie points vary between 11°K and 870°K.

This third group, though last to be discovered, received considerably more theoretical attention than the two other groups.¹ It has thus become quite certain that one has to locate the carrier of the high polarizability in the oxygen octahedra, which are present in all of these ferroelectric oxides.²

The mechanism which causes a spontaneous polarization in group II is according to Slater³ an ordering of the hydrogen bonds. This results, according to Frazer,⁴ in an induced polarization of the PO_4 tetrahedra.

No satisfactory explanation has as yet been given for either of the two Curie points of Rochelle salt.

---

¹ J. C. Slater, Phys. Rev. 78, 748, (1950).

² B. T. Mathias, Science, 113, 591, (1951).

³ J. C. Slater, J. Chem. Physics 9, 16, (1941).

⁴ B. C. Frazer, Acta Cryst. 6, 273, (1953).

273

SOME EXPERIENCE IN CODING AND PROGRAMING WITH THE INSTITUTE COMPUTER

by

Herman Goldstine

Institute for Advanced Study
Princeton, New Jersey

There is some merit in dividing the work of a programmer or coder into two main categories: there is that phase of his activities which occurs before he places a given problem on the machine and that one which occurs during the actual machine operation. This latter activity also consists of two parts: the "debugging" stage and the real running of the problem. I should like now to discuss mainly the complex of difficulties that arise in connection with machine operation.

Ideally one would hope to establish a sharp dichotomy between the mathematical and engineering phases of the question. In point of fact this is not feasible. In practice the coder must not only devote considerable attention to the mathematical aspects of a problem but also to their relationship to the engineering ones. For example, a highly inductive code may have at least in some portion a too high "read around" ratio, or again the reliability of the machine may influence the amount of checking that one is required to perform. This interplay between mathematics and engineering places a severe burden on the coding group of two sorts: on the one hand this group must prepare problems in such a fashion that they observe all safety precautions dictated by the particular requirements of the machine itself, and secondly, they must introduce some safeguards into their codes which will insure that the solutions obtained by the machine are indeed valid results and are not contaminated by machine errors.

It is, therefore, not possible with present day instruments to divorce the engineering and mathematical difficulties which arise in program preparation and program solution. It is some aspects of this interplay which I should like to discuss with you now. In this same connection it is well to point out that in addition to safeguarding codes against both possible mathematical and engineering difficulties the programmer must constantly assist the engineering staff in the localization of breakdowns in the machine. That is to say, our experience has been that the more specific the analysis of a calculational error which can be given the engineers the more quickly they can diagnose the fault in the machine. Present day instruments are so complicated that one must attempt in so far as possible to work cooperatively with the maintenance group not only in the tests but also in the localization of machine breakdowns.

274

One of the most useful means for treating this complex of difficulties is to make a reasonable use of mathematical redundancies in our codes. The use of such techniques is oftentimes of considerable aid in finding errors in codes and in detecting and even in localizing machine failures. It is, of course, important that the redundancies check as significant as possible portions of the calculation. If, for example, one is integrating the equations of motion for a conservative dynamical system, it is prudent to arrange the code in such a fashion that from time to time the total energy of the system is calculated and compared to the initial energy. If these two representations of the energy do not agree to within a reasonable tolerance, it is desirable to have the code stop the machine. In the initial stages of "debugging" such a problem this form of check is an aid in diagnosing code errors and during the actual running of the problem it is a substantial insurance of the reliability of the machine. Clearly the more use that is made of redundancies, the more one can ensure the accuracy of his results and the reliability of the machine; the more such tests exist in a code, the more nearly one can localize for the maintenance group any machine breakdowns. Innumerable examples present themselves in various problems of possible redundancies. I shall mention only two which we found of help in some problems. In some "rotation of axes problems" we had occasion to deal extensively with matrices which were transforms of each other by unitary processes. Throughout the calculations we checked that the norms and traces of the transformed matrices were invariant to within certain tolerances. Again, in a hydrodynamical problem we were able to use an extra condition at the shock front to verify the accuracy of our results as the problem progressed. One can, of course, adduce many other illustrative examples. Generally one finds in virtually every problem some superfluous condition which will serve in the manner indicated.

I should like to return for the moment to discuss briefly the "debugging" of codes. This is one of the most critical phases from the mathematical point of view. What I mean by "debugging" a code is ascertaining the initial errors in programming and coding in a problem which has just been placed on the machine. Furthermore, this check procedure must also determine whether any portions of the code which will become operative at a later time are also correct. I think it is fair to say that the most difficult problem in coding is this whole complex of considerations centered around the determination of the correctness of a code. There are, of course, many possible approaches to the problem but it has been my observation that entirely too little attention is given to it. Almost invariably codes are prepared with little or no reference to the initial debugging. Too often a coder brings a new problem to the machine with few, if any, ideas of how to find coding errors. The crudest technique, and unfortunately the one too many people tend to follow, is to attempt by manual techniques to go through large blocks of orders one at a time and to verify whether or not the code performs the algorithm in question. Such a technique can, of course, be extensively time consuming on the one hand and on the other it frequently fails

to reveal difficulties that arise in highly inductive systems simply because not all possible alternative paths are explored thoroughly. Indeed, such an analysis is often completely prohibitive in time spent by this primitive technique.

The problem of debugging is, of course, one of great difficulty and contains an apparent paradox. That is to say, if one knew the a priori results of a calculation, he would not need to perform the calculation at all. Therefore, in testing the correctness of a program one must devise techniques for gaining confidence that a given set of tests will reliably insure the accuracy of the entire code. We have tried a number of different attacks on this problem and although we are far from feeling satisfied with our tentative conclusions, we do at least feel considerably more assured than we formerly did. In order to analyze the correctness of a code it is perhaps desirable to consider some of the aspects of the problem which must be confronted and solved. There are two main points to consider: both the arithmetics and logics of a code. In handling the former aspect of the problem it is frequently desirable to arrange the code in a series of subroutines and to consider the main code itself as that portion which is concerned solely or at least primarily with the logics of the problem. In this fashion the problem at hand is broken down into the debugging of the individual subroutines, i.e. into a consideration of the arithmetics, and into an analysis of the main code itself, i.e. into the logics. If a code is arranged in some such fashion as this, one may introduce suitable numerical material into the introduction of each subroutine and determine the validity of the algorithm supposedly characterized by this subroutine. In this connection I should like to make a few more remarks. Ideally a subroutine of this sort, i.e., an arithmetic one, is intended to represent a valid approximation of some formula over a range of variability of a certain parameter or parameters. It is, therefore important in a debugging program to check the validity of a subroutine at various points in the range or ranges of variability. By a systematic program of this sort one can satisfy himself that a given code is arithmetically accurate within the limits prescribed by the programmer. Once this fact has been ascertained the balance of the debugging consists in a critical evaluation of "connective tissue," i.e., of the main code itself. That is to say the problem now resolves itself into an analysis of the path followed by the control in utilizing the subroutines. In this connection it is sometimes valuable to arrange the main code in such a fashion that a record is made of the number of times and the sequential order in which each subroutine is used in making a complete transit through the entire code. This generally can be carried out by a very few orders and a bare handful of extra storage positions. Such a procedure has the advantage that in a very short time the path that has been followed by the control while under the influence of the code can be determined.

Another technique which is often of considerable use and which is particularly useful after the initial testing indicated above has been performed is this. Often one can by specializing some parameters in a problem reduce

it to a known situation. If this can be done, a test comparison can be made between the analytically known results and those produced by the machine. Such a test is often valuable in that it can be used to give information about rounding errors, truncation errors, etc. Frequently a problem without many degrees of freedom can be put into such a form by regarding, e.g., fixed numerical coefficients as parameters which are available to the coder for checking purposes. A word of caution is, however, desirable in connection with this type of testing. Often particularization of parameters may cause some important portion of a code to be eliminated or used in a trivial way, or again some subroutine may be used over a very restricted range of its characteristic parameters. Thus it may happen that the results of a special calculation are quite reasonable in spite of serious errors still present in the code. It would therefore seem reasonable to use procedures of this sort as a secondary method and not as a primary one.

Let us for the moment leave these aspects of the problem and consider the time spent on the machine during debugging. In various laboratories one finds considerable variation in the amount of time allowed for checking out a new code. This varies in some places from many hours per person to others in which some fixed time such as 5 - 10 minutes is allowed. It is, of course, quite difficult to lay down a priori any reasonable rules on this point. However, it seems to us quite desirable to require the coders to go to the machine with a carefully thought out plan of action which will not require an undue amount of time to carry out. In this connection we found a significant saving in time if debugging occurred when we changed our input-output system from teletype to IBM equipment. This improvement came about because of the superior speed of the latter machines which, for an expenditure of only one or two minutes, produce a complete printing out of the contents of a memory at a given time. If, then, the coder has arranged things in such a fashion that he places the relevant arithmetic portions in subroutines, he can at the end of an extremely short calculation obtain a virtually complete record of what transpired. He is then able, in his office, to analyze in great detail what occurred during the machine calculation. Such a procedure almost invariably in our experience allows one to determine quite narrowly that area of a code in which a difficulty has arisen. Then an inspection of the input data itself is usually sufficient to find the source of error and almost all of this without recourse to the machine.

The final checking out of the code which is important is the running of it for some reasonable length of time to determine whether such considerations as "read around" are important. Frequently this can be best ascertained by running the same problem several times and comparing in their entirety the contents of the memory on each of the runs. This is very simply done with the IBM equipment.

We may then summarize our *modus operandi* for checking out a given code as follows: we attempt, in so far as possible, to arrange our codes in such a fashion that virtually all the arithmetics are handled by special

277

subroutines and the logics by the main code. We then analyze each subroutine by one or more runs on the machine with complete punch outs at the end of each run. We analyze these printings of the memory to localize in the subroutines those portions which are in error. Having determined the correctness of each subroutine we turn to the main code itself and by techniques such as enumerated determine the correctness of this portion of the code. Next, we attempt to run a special case of the problem about which we may have information and compare results to the last calculation. Finally, by a few runs of the same problem we attempt to gauge the seriousness of the read-around situation. In this connection it should be remarked that we require a detailed analysis beforehand of the read-around situation. This is helpful because it serves to indicate where the most dangerous - from a read-around standpoint - portions of the code lie.

So far I have talked only of the debugging of a code; however, the techniques I have just mentioned will also serve to help the maintenance group in the localizing of non-transient errors. In the same connection it should be noted that the duplications of runs mentioned earlier form the basis for another kind of error checking, that is to say, once it has been determined by several machine runs that the results of a calculation are correct the code can then be arranged in such a fashion that from time to time this test calculation is run and the results at a given time compared to the ones which have previously been obtained.

So far I have concerned myself solely with checking out of a new code and the ascertainment of errors in one which is being run on a machine solely by mathematical and logical considerations. There is, of course, some place in this problem for the use of engineering apparatus itself. One place in particular we have found it desirable to add apparatus to our machine is in the checking of "spills," i.e., the determination of whether a number has grown out of the size permitted by the machine. In our case this can occur with the additions, left shifts or divisions. One can, of course, by coding techniques guard against "spills" but since they are related to the most frequently used arithmetic operations it seems to us desirable to guard against them by circuits built into the machine. Errors which arise from such spilling of numbers are particularly insidious in that a calculation may proceed for hours or even days before such a spill occurs; accordingly it is difficult to say how much initial checking out of a code can safeguard against this type of error in some calculations. As I said before it can, of course, be guarded against by coding checks but in many cases this might lead to an undesirable elaboration of a given code. We have accordingly made the decision during the coming year to install circuitry to guard against such types of difficulties. Again these can occur either in an insignificant mathematical situation or a breakdown of the machine. In either case, it is clearly important to know at once the nature of the error. Such errors are particularly insidious in that they may not occur in a portion of the calculation which is relevant at the moment when the error occurs but which only becomes important at a later moment thus masking the true origin of the trouble.

In closing I should like to turn attention now from the mechanics of localizing errors in codes to a closely related task, namely that of "keeping books on a given problem from its initiation." We feel strongly that successful operation of a laboratory requires stringent rules for keeping exact records of a number of facts related to a given problem. The actual algorithm adopted for a given problem should be carefully recorded together with a record of the sizes of all quantities which can be so bounded. Furthermore, there should be a statement of which parameters cannot be handled in this fashion, i.e., those for which some "floating points" technique will be needed. Again all changes in codes should be recorded, not only changes themselves but the reasons why the changes were made. Furthermore, during the debugging of a code careful records must be kept of all the conditions which characterize any given printing out of the memory. Without these facts it is virtually impossible to make an analysis of a given portion of a problem. It is my opinion that it is impossible to overemphasize at this stage of machine experience the rigor with which one treats these bookkeeping problems.

I hope that I have not said too much about many of these seemingly trivial considerations, but I should like to close by reiterating our experience which has been that the strictest attention paid to these trivial operations means the difference between a reasonable efficiency of operation and a poor one. Undoubtedly many of these conditions become less important as one gains more experience with the operating economy of a computing establishment; however, in these early stages it is of utmost importance to maintain a most careful control of all tests.

279

COMPUTING THE WEATHER

by

John W. Mauchly

Remington-Rand, Inc.
Eckert-Mauchly Division
Philadelphia, Pa.

This symposium is direct testimony to the wide-spread interest which now exists in the field of automatic electronic computation. Yet we are, even now, but upon the threshold of the coming era of information processing systems. It is less than eight years since the very first truly electronic instrument was put into service. Just ten years ago, at the Moore School of the University of Pennsylvania, a group of engineers began the design of the ENIAC for Army Ordnance. Mr. Chu, who worked with Mr. Eckert and myself on that pioneer project, can well take pride in the new computers which we celebrate at this symposium.

However, the aim of this symposium is not to look back on the accomplishments behind us, but to look forward and to try, so far as we are able, to envision what lies ahead. While others on the program are considering new components and developments in computer design, my topic tonight has to do with computer applications to a specific field - meteorology and weather forecasting. In keeping with the aim just noted, I shall have more to say about the future possibilities than about past achievements.

First let me state that I am not a meteorologist, nor do I claim any expertness in the present-day methods of weather forecasting. However, I have been deeply interested in this subject for sixteen years, and it was a direct consequence of that interest which projected me into the field of electronic digital computers. I became convinced that certain vexing and important problems in meteorology could be answered by proper statistical analysis of existing data. When I attempted to carry out even a small part of what I had in mind, I found myself bogged down with tedious figure-work.

If, at that time, I could have turned all of this statistical arithmetic over to a battery of punched-card devices, I might never have thought of proposing an automatic electronic computer. However, my computational resources consisted of one desk calculator, one adding machine, and a small number of part-time college students. Having set these helpers at various clerical tasks, I began dreaming about a new research tool which might be built from electronic counting or 'scaling' circuits. Some of these things I tried out in the laboratory, but my plans went far beyond anything that I could hope to realize with the small funds at my command.

With the approach of World War II, I laid aside my statistical researches into meteorology and joined the staff of the Moore School, where I met Mr. Eckert. He and I had many a discussion regarding the practical possibilities of building a large automatic digital computer, but it remained for Dr. Hermann Goldstine, then representing Army Ordnance, to convert our plans into hardware through a government contract with the University of Pennsylvania. The ENIAC which thus came into being was, of course, destined primarily for ballistics work, and not for weather research.

Nevertheless, it is fair to say that weather research was responsible for guiding my thoughts toward electronic computers. Now let us turn to the question: 'What can electronic computers do for the weather?'

No doubt many of you have already read newspaper and magazine articles about 'numerical weather prediction.' I shall say more about this development in a moment. We may be certain that the day will come when almost all weather forecasts will be made by electronic computers. When that time arrives, the computer will be serving as a production tool. The processes being carried out by the forecasting computers will be very complex, but for the most part they will be monotonous repetition of routine steps.

We cannot, however, expect to arrive at this well-organized use of computers as production tools until we know what, precisely, we want the computers to do. Should we ask the computer to follow the same procedures now used in present-day forecasting? Or can new methods, entirely impractical without a computer, be used with better results? If so, how do we find and develop those new methods? It should be clear that, in our search for new methods, and in testing new methods, we must use computers. The computer is then the essential research tool for the development of better forecasting methods.

There are actually so many ways in which our new and powerful computing instruments can accelerate research in meteorology that I shall have to be content to suggest but a few to be considered as examples. It would be a mistake to limit ourselves to just those phenomena which seem, with our present knowledge, to be related to the weather elements which we hope to predict. The history of science is filled with instances in which new understanding of one problem has come from some quite remote investigation. Meteorology, in its broadest sense, has to do with all of the Earth's atmosphere, not just the portions near the ground. As we look with new techniques, such as radio and rockets, into the higher reaches, we find many things which are not yet well understood, and processes unknown to us yesterday may, in the not too distant future, be found of direct concern in our efforts to compute the weather.

At this point I wish to read a few sentences from a report on forecasting, in the hope that you will find this brief capsule directly relevant to our theme.

281

"The forecasts now issued by meteorological services are based upon skill acquired by study, observation, and experience in calculating the development, movements, increase, and decrease in intensity, and magnitude of areas of high and low barometric pressure that appear within areas covered by the . . . . reports of the several services. . . ."

"The next advance in meteorological science must result from extensive research in solar physics and terrestrial magnetism and comparison of solar and terrestrial phenomena."

"Thus, meteorology is really a very closely allied but difficult branch of solar physics, and it ought to be studied with the aid of a fully equipped laboratory devoted especially to such researches."

"In this field are found the best examples to physical and mathematical problems, because it is nature's great laboratory."

Although I have carefully selected four different passages from a text of 68 pages, I do not think I have done great violence to the opinions of the authors in lifting these portions from their context. They came from an official Weather Bureau publication, Long-Range Weather Forecasts, issued in 1904.

Although written nearly fifty years ago, these passages may still be quite timely. The description of current forecasting methods could apply with slight change today, since only the kind of data and the interpretation of that data has changed, and not the element of skill in dealing with it. My main emphasis, however, is on the statement that in meteorology are to be found some of the most difficult problems in physics and mathematics. To achieve better forecasts, we should be attacking those problems, and among the tools at our disposal for the research which needs to be done, the electronic digital computer is of primary importance.

Let us admit that any forecast, no matter how we arrive at it, is an extrapolation. If we know very little about the fundamental laws which govern a system, we are not able to make good use of our prior observations, and our extrapolations become very unreliable as we attempt to predict for longer periods into the future.

For a simple system, a great deal may be learned empirically. For instance, a very accurate mathematical description of the motion of a pendulum could be derived from a number of observations distributed over many periods. We could then predict with considerable confidence the motion which would be observed during many future oscillations - without any knowledge of Newton's Laws of Motion or of the Law of Gravity.

We are indeed very far from this simple situation in dealing with the Earth's atmosphere, which consists of more than 10^{44} molecules, each with both external and internal degrees of freedom. How can we ever expect to cope with a system of this complexity? The answer is, in part, that we must not try to follow every individual particle and know exactly its motion and energy state, but we must be content to deal with average values. Hence we use the concepts of pressure and temperature, which are time-averages of individual molecular quantities which we need not measure at all.

Both the time-scale and the space-scale of the fluctuations in the atmosphere have a wide range. We enter the fields of macroclimatology, microclimatology, or various other branches of meteorology, according to what extent of space or time we choose to average over. We may think of climatology as concerned with the long period trends, including annual and seasonal variations, and weather as concerned with some of the deviations from such long period trends. However, exceedingly short period deviations, such as the sound waves set up by a train whistle, are certainly not a part of weather. Thus, what we choose to call weather lies within a certain "frequency band." Weather is also characterized by scale factors in space as well as time.

These seemingly trite observations point to the root of certain difficulties which must be overcome when attempting a numerical weather prediction. The equations of hydrodynamics which apply to any fluid could conceivably be used to calculate any future state of our atmosphere from certain initial conditions. The fact that we do not know all of the initial conditions accurately is only a part of the difficulty. A second and very fundamental difficulty lies in the very generality of these equations, which are capable of representing sound waves as well as weather. As a consequence, the general equations have properties which make them extremely difficult to handle by finite difference methods. Fortunately, a derived equation which involves a quantity known as vorticity is, by reason of its specialization, much better suited to the needs of weather prediction.

At the Institute for Advanced Study in Princeton notable progress is being made toward the integration of the vorticity equation for limited regions of our atmosphere. The Princeton group first set up a single-layer model of the atmosphere, and then proceeded to study multi-layer models of increasing complexity. The initial data used for tests in each case is derived from actual observational data for historical situations just preceding large storms. The numerical predictions from the computer are then compared with the actual developments which are a matter of record. I believe I am correct in saying that the agreement between prediction and observation has improved as more complicated models have been tried. This work at Princeton is undoubtedly one of the major contributions to short-range meteorological forecasting. I shall not go into the details of the Princeton work, but refer you to the paper by Dr. Jule Charney and Dr. Norman Phillips which was published in the April 1953 issue of the Journal of Meteorology.

The numerical integration techniques just discussed have to do with weather, rather than climatology. Let me now propose to you a problem which appears to lie in the realm of climatology, but which may very well be related to weather prediction. For many stations in the United States there are detailed weather records for every day over a period of fifty or sixty years. From such records it is customary to calculate "normals" which correspond roughly to the statistician's concept of "expected value." Since we recognize the existence of annual and seasonal variations in meteorological quantities, such normals are derived separately for each month.

We would like to have normal values for individual days as well as for months. Here we find difficulties. With only a sixty year record, there are only sixty observations of the precipitation on July 31st, say. The differences between the average values found for one day and the immediately following day, by such a scheme, are often quite large. Precipitation for one day may average twice that for the next! Are such results to be explained entirely as sampling errors due to the high variability inherent in precipitation observation and the small number of days in each sample?

The honest fact of the matter is that no one yet has a good answer for this question. Those who believe that the answer is "yes" will use some kind of smoothing technique to iron out the variations from day to day in the normals. Everyone has a different opinion as to how this ought to be done. However, there is a contrary school of thought which maintains that these day-to-day differences are, at least in part, real differences in the expectation and consequently should not be removed by smoothing.

It is evident that we have here a statistical problem of extreme complexity. One such peculiar variation in temperature (and in other weather elements) occurs near the end of January, and has come to be known as the "January thaw." Dr. Wahl, of the Geophysical Research Directorate in Cambridge, has studied this January thaw intensively, and has come to the conclusion that it is not merely a sampling variation but that there is a real tendency for warmer weather to occur in New England near certain calendar dates.

Such a calendar-fixed peculiarity in weather is known to meteorologists as a "singularity." Perhaps singularities should be considered a part of climatology rather than weather, but this seems an academic question. However, a very real and important problem demanding a great deal of statistical effort is the identification of "genuine" singularities. For the most part, meteorologists have avoided the investigation of singularities, preferring to regard them as statistical accidents in the absence of a good explanation as to why singularities should occur.

284

I believe that a full-scale investigation of the problems raised by singularities will be highly rewarding. There is a vast government storehouse of meteorological data in Asheville, maintained by the Air Weather Service. To analyze even a fraction of this data for singularities would be difficult indeed without high speed automatic computing equipment. Here then is one way in which the electronic computer of today might be used in meteorological research. The results would, I am sure, be applicable to weather forecasting.

Let me now cite another field of research in which the electronic digital computer could be the major tool. It has been known for a long time that the sun and the moon set up tidal motions in our atmosphere, as well as in our oceans. Tidal forces vary inversely as the cube of the distance between the earth and the tide-producing object. Hence the moon, in spite of its much smaller mass, has a greater tide-producing power than the sun.

In the case of ocean tides, the lunar tide is larger than the solar tide, by a factor which agrees with theory. The analysis of barometric records has shown that the atmospheric tidal component attributable to the sun is not smaller, but eighty times larger, than the tidal component due to the moon. Of course, some part of this twelve-hour solar wave is caused by the heating of the earth's surface by solar radiation. That is, there is a thermal tide as well as a gravitational tide due to the sun. However, the thermal excitation is by no means large enough to explain the gross factor by which the solar atmospheric tide exceeds in amplitude the lunar atmospheric tide.

A possible explanation, which for many years remained difficult to test, is that the atmosphere has a natural resonance period sufficiently close to twelve hours so that amplification by resonance may be invoked. However, prior to our recent knowledge concerning temperatures in the upper atmosphere, there seemed no way by which a resonance period close to twelve hours could come about. Then radio soundings and other new methods of observation showed that the temperature of the atmosphere does not continue to decrease with altitude, but has one or more relative maxima corresponding to certain special layers in the upper atmosphere.

Pekeris then showed that under such temperature conditions in the upper atmosphere a resonance period close to twelve hours might be possible. However, for a quantitative check rather than a simple qualitative answer it was necessary to resort to a large computer. Weekes and Wilkes studied this problem in 1947 with the aid of the differential analyzer at Cambridge University. More recently, Jacchia and Kopal made further calculations using the differential analyzer at M.I.T., as reported in the Journal of Meteorology for February 1952. Available to them were two different temperature curves for the upper atmosphere. One was obtained from V-2 rocket soundings, and the second was a provisional standard agreed upon

285

by the National Advisory Committee for Aeronautics. The M.I.T. work showed that the required resonance amplification could not be obtained from either of these curves, but that an intermediate curve, which lay within the bounds of possible error for the other two, was able to explain satisfactorily the period and resonance amplification as required.

The story does not end there, however. The M.I.T. work seems to indicate that, although this twelve-hour solar tide is not too sensitive to surface atmospheric conditions, it is extremely sensitive to the temperature of the atmosphere at the height of approximately fifty kilometres. I find such a result extremely interesting because, in effect, this provides an indirect measure of atmospheric temperature at a height of fifty kilometres, simply by the analysis of surface barometric records.

The firing of a V-2 rocket is a rather expensive affair. Presumably we can look forward to the development of much cheaper methods for getting instruments into the upper atmosphere, but any direct instrumental method of making observations of the upper atmosphere temperature will have to be extremely cheap before it can compete with the method just suggested. A direct measurement will, of course, give a record localized in time and place, while the barometric tidal analysis should yield a world-average for a period measured in days rather than minutes. It would certainly be impractical to obtain this kind of an average by rocket soundings. Moreover, we have just begun to explore the upper atmosphere directly with instruments, but there are more than a half century of surface pressure records available for analysis now.

As with all atmospheric phenomena, the solar atmospheric tide is variable from season to season, and from month to month. Chapman and others have computed seasonal variations for many stations. This is a tremendous job, and much more remains to be done. It is certainly a field in which the best of our computing machines could well be employed.

What have these tides to do with weather and weather prediction? Of course, there is the answer which applies quite generally to all research in meteorology - the more we know about our atmosphere and the better we understand the processes taking place in it, the more favorable are our chances for understanding the science of weather prediction.

It is my belief, however, that there is a much closer connection. From radio measurements, we know that the variability of the sun's ultra-violet radiation causes extreme changes in the higher portions of our atmosphere. The atmosphere strongly absorbs such short wave-length radiations. Hence, any direct measurement of solar radiation at ground level shows only a slight trace of the high variability which must exist at the top of our atmosphere. At the intermediate level of fifty kilometres, we find the region of maximum ozone concentration. The role of ozone in atmospheric processes is not yet well understood, but more information on the behavior of this layer ought to help in understanding, and perhaps even in forecasting, surface weather.

286

We have seen that the solar tidal oscillations may help us to estimate the average temperature of the upper atmosphere during any month or season. Why stop there?

Instead of combining the data for every January at a particular station, thus deriving the "normal" January tide at that station, let us combine the data from many stations for a particular January, thus estimating quite accurately the amplitude of the solar semi-diurnal tide for that one January. I strongly suspect that when this is done for many months we will find that the amplitude of the twelve-hour oscillation will differ significantly from one year to another. As early as 1940, I carried out some preliminary tests along these lines. The results suggest that a comprehensive analysis will ultimately give us important information about the time-variation of conditions in the upper atmosphere. Perhaps these variations will disclose a correlation with solar activity. These same variations may also be found related to the atmospheric circulation patterns which have a great deal to do with our surface weather.

I have discussed only a few of the many problems in meteorology which could well use the capabilities of large automatic digital computers. There must be a great deal of information resident in the vast bulk of weather records accumulated through the years. Much of this information could be abstracted by appropriate statistical methods. More and more data is constantly being accumulated, and the rate of acquiring data increases each year. Once such data has served its purpose for short-range forecasting, very little further use is made of it. To be sure, certain climatological means and also various contingency tables and frequency distribution charts are prepared. We should, however, analyse the historical data by more sophisticated methods designed to improve our understanding of atmospheric processes. The world spends hundreds of millions of dollars each year on meteorology. Too small a fraction of this is devoted to the analysis of existing data.

As Dr. Conant emphasized in his book, "On Understanding Science," there is no such thing as the scientific method. There are many ways in which new information may be acquired, and our scientific understanding of nature can be advanced. The atmosphere of the earth is an extremely complex system, and we need every resource which we can command to help us in discovering how it operates.

On several occasions I have given papers in which the hypothesis that solar variations have an important influence on surface weather has been tested statistically. The tests have, in each case, strongly favored the hypothesis. An electronic computing machine could, in numerous ways, further the statistical research needed to explore such questions.

287

I have emphasized the statistical uses of computers because I feel that those aspects have been most neglected. However, we must remember that there are many nonstatistical applications, too. Mathematical equations may often be proposed which cannot be solved except by numerical methods. Ultimately, I think, we shall face a large number of extremely difficult mathematical problems which, although appearing presently as of purely astronomical interest, will be found pertinent to the computation of weather. We shall have to learn how to predict the behavior of the sun. Also, many phenomena of interest to the solar physicist depend upon magneto-hydrodynamic equations. Astronomers have recently taken a great deal of interest in these equations for their own reasons. Eventually, the difficult problems lying in this field will need for their investigation the best electronic computers which can be built.

There are today only a few meteorologists who are willing to grant any close connection between solar physics and weather. The pros and cons of this long-pending controversy need not concern us here. The point to be stressed is that we have here a question of fundamental importance to the science of meteorology which might be resolved by the proper application of digital computers. As Prof. Willett of M.I.T. has said, a definite answer either way - pro or con - would be of immediate and vital importance to the efficient conduct of meteorological research.

I have discussed only a few of the many ways in which present-day electronic computers might serve as essential tools for the prosecution of research. With one exception, I have had to suggest what might be done, rather than describe what has been done. There are many potential applications, but only one actual accomplishment.

At present, the only meteorological research project centered about a large high-speed digital computer is at the Institute for Advanced Study. It seems obvious that the excellent progress so far reported on numerical prediction techniques for short-range forecasts would not have been possible without an adequate computer. It is my conviction that, in many other fields of weather research, progress could be greatly accelerated by adopting the electronic computer as a basic tool for research.

288

DESIGN AND MANUFACTURING CONSIDERATIONS OF THE OSCILLOGRAPH-TYPE ELECTROSTATIC STORAGE TUBE

by

E. M. Smith

Tube Department
Radio Corporation of America
Lancaster, Pa.

Although oscillograph-type electrostatic storage tubes have been described in terms of computing-machine parameters, these descriptions generally have not defined the basic factors of tube design.¹ Some of the factors which determine tube performance, and their effect on manufacturing processes, may be indicated by a brief consideration of the more important characteristics of such tubes, the inter-relation of tube design and performance characteristics, and the design of the RCA developmental oscillograph-type electrostatic storage tube.

Samples of this RCA developmental tube supplied to the U.S. Navy under a Bureau of Ships development contract* were reported to be superior to conventional oscillograph tubes in computer storage systems. Tests by computer groups using the RCA tube have been valuable in guiding the design for future developmental types having improved performance.

REQUIRED CHARACTERISTICS

Characteristics of primary importance in the oscillograph-type electrostatic storage tube are uniformity of the secondary emission of the storage surface and storage resolution (the amount of information that can be stored by the tube). In the effort to obtain improved secondary-emission uniformity, an extensive investigation has been made of screening methods and materials which will produce a satisfactory storage surface. Although considerable progress has been made, the preparation of the storage surface remains a difficult manufacturing problem.

Improved storage resolution depends to a large extent upon reduction of the size and upon good control of the current distribution of the focused spot. Conventional oscillograph tubes, such as the 3 KP1, have a relatively large fringe of low current density surrounding the visible spot. For

---

*U. S. Navy Department, Bureau of Ships Development Contract NObsr-52144.

oscilloscope applications and at low "read-around ratios"* in computer storage systems, this fringe is not important. At higher "read-around-ratios," however, this fringe limits storage resolution. The size of the spot focused at the center of the tube face increases as the beam is deflected due to deflection defocusing. For a pattern of equally spaced storage elements, deflection defocusing may be a major factor in limiting storage resolution.

The ideal storage tube, therefore, should have a target surface having uniform secondary emission and a focused spot having a sharp boundary and a small size which does not increase with deflection.

THE ELECTRON GUN

Characteristics of the focused spot are determined largely by the design and quality of workmanship of the electron gun. The electron gun used in the RCA developmental oscillograph-type electrostatic storage tube is a modification of the type used in the electrostatic-deflection and electrostatic-focus oscillograph tubes.² This gun, shown in Figure 1, consists of a cathode which provides a source of electrons, a grid No. 1 (control grid) which regulates the amount of cathode current, and a grid No. 2 which accelerates the electrons. Grid No. 3 (the focusing grid), the top grid-No. 2 aperture, and grid No. 4 form a symmetrical focusing system. The deflecting electrodes are used to position the beam on the target. Grid No. 2, grid No. 4, and the collector are connected within the tube and have a common base-pin connection and are known as the Ultor. When grid No. 1 is biased at less than beam-current cutoff, electrons leave the cathode and are accelerated by the grid-No. 2 potential. The cathode, grid No. 1, and grid No. 2 form an immersion-type lens which causes electrons to converge rapidly after leaving the cathode, producing a region having a minimum cross-section a short distance from grid No. 1. The image of this cross section of the electron beam, called the "cross-over," is focused on the target to produce the minimum spot size.

Grid-No. 2 Voltage and Magnification

Two factors affecting spot size are immediately apparent: the diameter of the "cross-over" cross-section and the magnification of the focusing system. Because the "cross-over" diameter is inversely proportional to the square root of the grid No. 2 voltage, it may be reduced by operation of the tube at higher grid No. 2 voltage. Magnification is reduced by increasing the length of the gun, i.e., the distance between the "cross-over" and grid No. 3 with respect to the grid No. 3-to-target distance.

---

*The "read-around ratio" specifies the number of references which can be made to a storage element without changing the signal from an adjacent element by more than a specified amount.

290

The grid No. 2 voltage is limited by several considerations other than voltage breakdown between electrodes. For example, deflection voltage is directly proportional to the grid No. 2 voltage. In addition, signal amplitude from the oscillograph-type electrostatic storage tube is appreciably higher at low voltages, although it is difficult to determine whether the reduction in signal amplitude with increasing voltage is due to changes in secondary emission of the target surface or to smaller spot size. The grid No. 2 voltage is at least 1500 volts in most cases and may be as high as 2500 volts in many cases.

Deflection Defocusing and Beam Masking

Deflection defocusing is an important limitation in oscillograph-type electrostatic storage tubes. For a given array of equally spaced target elements, the minimum separation between elements is determined by the largest element. When grid No. 3 is adjusted for best focus at the center of the tube, spot size increases with the deflection angle. The spot can be partially refocused at the edges by increasing the grid No. 3 voltage so that grid No. 3 is adjusted for best focus in an intermediate zone between the center and the edge of the array. For a given deflecting-electrode arrangement, the amount of deflection defocusing is determined principally by the diameter of the beam as it passes through the deflecting electrodes and by the deflection sensitivity.

Masking the beam in the grid No. 2 cylinder reduces the effects of aberrations in the focusing lens, thereby improving the characteristics of the focused spot. The improvement effected by the masking aperture depends on the size of the aperture, its position in the grid No. 2 cylinder, and the quality of the part.

Grid Characteristics and Life

Grid-cutoff voltage and grid-drive characteristics affect "cross-over" diameter, the divergence of the beam after "cross-over," and cathode loading (current density from the cathode surface). An increase in grid-cutoff voltage produces a smaller "cross-over" diameter, but the beam divergence angle and cathode loading are increased. Because of aberrations in the immersion lens and space-charge effects, the "cross-over" diameter increases with grid drive. Beam divergence and cathode loading also increase with grid drive.

Life tests on oscillograph-type electrostatic storage tubes have shown that the life of the tube is primarily dependent on the life of the cathode. The life of the cathode is more directly related to cathode loading than to total cathode current.

Design Considerations

291

All the factors mentioned above must be considered in the determination of a suitable design for the oscillograph-type electrostatic storage tube. For example, because the required screen current is small compared with that of conventional oscillograph types, more beam masking may be employed. Masking, however, reduces the amount of screen current with respect to total cathode current so that greater cathode loading and more grid drive are required to produce the optimum screen current. The amount of masking is limited, therefore, by the cathode loading which will provide reasonable life and grid-drive characteristics which maintain a small "cross-over" diameter. One of the factors determining the magnification of the electron gun is the length of grid No. 2. For a given masking-aperture size, a longer grid No. 2 increases masking because of the smaller solid angle formed by the "cross-over" and the edge of the masking aperture. Grid No. 2 may be lengthened without change in the amount of masking provided the diameter of the masking aperture is increased in the same ratio as the distance from the "cross-over" to the masking aperture. This arrangement, however, produces a greater beam cross section in the region of the deflecting electrodes and a corresponding increase in deflection defocusing. Reducing magnification by lengthening grid No. 2, therefore, is limited by beam masking and deflection defocusing.

RCA DEVELOPMENTAL TUBE

The RCA developmental oscillograph-type electrostatic storage tube has a grid No. 1-cathode assembly similar to that used in conventional oscillograph types. The diameter of the grid No. 1 aperture is 0.025 inch and the spacing between the grid No. 1 and the cathode is approximately 0.005 inch. The grid No. 2 assembly is made long to reduce magnification. The masking aperture has a diameter of 0.031 inch and is spaced slightly behind the top grid No. 2 aperture. Location of the masking aperture in a relatively field-free region reduces the distortions introduced by masking at the top grid No. 2 aperture, which is in the region affected by the focusing field. This arrangement produces a more nearly round spot having a sharp beam boundary. The masking aperture, which plays a very important part in determining the shape and size of the focused spot, is manufactured to close tolerances and is carefully cleaned and gold-plated to prevent the accumulation of non-conducting oxides in or around the hole. An insulating coating such as the oxide can acquire a charge under electron bombardment and thus produce an electric field which can distort the beam. Masking in the RCA developmental tube is almost two times greater than that used in the 3KP1 oscillograph tube. Although there is a considerable variation from tube to tube, approximately 94 per cent of the current that leaves the cathode is masked in the grid No. 2 cylinder. This measurement is made with a screen current of 2 microamperes.

The focusing lens of the RCA developmental tube, which is similar to that used in the 3KP1, consists of the top grid No. 2 aperture, the grid No. 3 and the grid No. 4. Concentricity with the tube axis, roundness of the apertures, and parallel disc surfaces are important. As a result of the closer spacing of the deflecting electrodes, the smaller beam diameter, and the narrower deflection angle, this tube has effectively less deflection defocusing than the corresponding oscillograph type. In general, alignment of the apertures, spacing, and quality of the parts are very important to the performance of the electron gun of an oscillograph-type electrostatic storage tube.

Screen Defects

A difficult problem in the manufacture of storage tubes is that of defects in the screen which prevent storage in certain areas of the screen. This problem is sometimes due to non-uniform secondary emission from the storage surface as a result of such factors as leakage, surface roughness, grain size, or photo-electric effects. Usually the screen defects are small particles of foreign material which have settled on, or with, the phosphor and produce variations in the amplitude of the positive output signal. Attempts to locate these defects before the tube is finished have been only partially successful. In many cases, the defects are similar in appearance to the phosphor and can be observed only under electron bombardment. To date, the most effective method of reducing such defects has been the use of extreme care in obtaining clean phosphor, settling the screens, and handling the tubes during processing.

Fortunately, a means of salvaging some tubes which have screen defects has been developed. Sparking the face of the tube with a high-voltage spark coil, as described by W. E. Mutter,³ is often effective in removing these screen defects. It is estimated that the yield can be increased by 50 to 75 per cent by this salvaging process. Striking the face of the tube sharply with a plastic rod to knock the defects from the screen is also helpful.

Electrical Tests

Testing of the RCA oscillograph-type electrostatic storage tube includes both the regular oscillograph-type tests related to the electron gun and tests peculiar to storage tubes. The first group of tests determine cathode emission, deflection sensitivity, gun alignment (which is essentially a check of screen current compared with cathode current), cutoff and grid-drive characteristics, leakage currents, voltage breakdown, and centering. Storage-characteristics tests are designed to detect and evaluate screen defects, and to indicate the amount of information which can be stored on the tube.

Several different types of screen-blemish tests have been investigated. The most sensitive test consists of scanning the storage tube with an unmodulated beam and displaying the output signals on a monitor oscilloscope; this test, however, does not provide a reliable evaluation of a screen defect. Another test is the so-called "line-of-dots,"⁴ in which a series of dots is superimposed on an unmodulated line sweep. The output signals from the dots appear as an envelope of positive signals on the monitor oscilloscope. The blemish is observed on the monitor as a notch in this envelope; the ratio of the depth of the notch to the average amplitude of the envelope evaluates the defect.

The "spill" test³ to determine storage resolution measures the separation between two dot elements at which the redistribution of secondary electrons from one element reduces the dot signal of the other by a specified amount. The test equipment provides for writing one dot a number of times before the neighboring dot signal is observed. This type of test is more reliable than one in which the two dots are written and read alternately because the low-current-density fringe surrounding the focused spots will not affect the dot signal unless repeated reference is made to nearby storage elements.

In addition to determining whether or not a given tube will perform satisfactorily in the equipment, testing has a second important function in manufacturing. The results of tests frequently indicate variations in production which can be corrected before a serious deterioration of quality occurs.

CONCLUSION

This paper has presented some of the fundamental considerations in the design and manufacture of oscillograph-type electrostatic storage tubes, and has indicated the compromises required to achieve satisfactory performance. It should be emphasized, however, that the optimum design of a storage tube is determined from the results of performance tests in computers. A number of factors, such as life of the storage tube, may suggest re-adjustment of tube design when computer performance is evaluated.

ACKNOWLEDGEMENT

The author wishes to acknowledge the assistance of Dr. L. B. Headrick and L. E. Swedlund in advising and directing the development program for the RCA developmental oscillograph-type electrostatic storage tube and their help and advice in the preparation of this paper.

294
REFERENCES

1. Williams, F. C., and Kilburn, T., "Storage System For Use With Binary-Digital Computing Machines," PROC. IEE (London), Vol. 96 Part II, 1949 pages 183 to 202.
2. L. E. Swedlund, "Improved Electron Gun for Cathode-ray Tubes," ELECTRONICS, March 1945, pages 122 to 124.
3. W. E. Mutter, "Improved Cathode-ray Tubes for Application in Williams Memory System," ELECTRICAL ENGINEERING, Vol. 71, No. 4, April 1952, page 352.
4. Bureau of Standards Report No. 12.3-9R, April 1951.

295

Figure 1

Diagram of electron gun used in RCA developmental oscillograph-type electrostatic storage tube.