HIGH-SPEED COMPUTING DEVICES

HIGH-SPEED
COMPUTING DEVICES

By the staff of
ENGINEERING RESEARCH ASSOCIATES, INC.

Supervised by
C. B. Tompkins
and
J. H. Wakelin

Edited by
W. W. Stifler, Jr.

FIRST EDITION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK   TORONTO   LONDON
1950

HIGH-SPEED COMPUTING DEVICES

FOREWORD

This volume represents the cooperative efforts of a number of members of the staff of Engineering Research Associates, Inc. C. B. Tompkins planned the report from which this volume has evolved and prepared much of the initial draft. The report was completed by the staff under the direction of J. H. Wakelin and was edited for publication in its present form by W. W. Stifler, Jr. Other contributing authors from the staff of this company were W. R. Boenning, W. W. Butler, A. A. Cohen, E. C. Olofson, L. R. Steinhardt, and Erwin Tomash.

The contents of this volume were first assembled in the form of a report to the Office of Naval Research, prepared under a provision of contract N6-ONR-240, Task 1, which called for "an investigation and report on the status of development of computing machine components." Publication in the present form came about as a result of a suggestion of Dr. Mina Rees, then head of the Mathematical Sciences Division of the Office of Naval Research, who felt that publication over a wider range would be worth while. The cooperation and assistance of the Mathematical Sciences Division of ONR in bringing about the publication of this volume is gratefully acknowledged.

As those readers who are familiar with this field know, many of the large-scale computing machines already built or under construction have been under Government sponsorship. The two principal all-electronic machines now in operation, the Harvard Mark III and the ENIAC, were sponsored by the Navy Bureau of Ordnance and the Army Ordnance Department respectively. The names of many of the men who have developed computing components and machines are mentioned in this text. Many of their developments have been made possible by those representatives of the sponsoring agencies, both civilian and military, whose foresight and whose confidence in the future of computing have made them willing to support the necessary research and development work.

H. T. Engstrom
Vice-President
Engineering Research Associates, Inc.

Arlington, Va.
May, 1950

v

PREFACE

This volume is primarily a discussion of the mechanical devices and electrical circuits which can be incorporated into computing machines. It is not a detailed comparison of various machines. However, we have included descriptions of a few computers, to provide examples of the integration of techniques and components into complete systems. Because the computers built in this country are so much more familiar to the authors than those which have been built or planned in England or in other countries, we have used American machines as examples.

A list of all those leaders in the computing machine field without whose assistance this book could not have been written, would assume the proportions of a separate bibliography. We are therefore omitting such a list, at the same time acknowledging with sincere thanks and appreciation the personal assistance of many of the men whose names do appear in the various chapter references and bibliography lists.

In the preparation of this manuscript and the correction of the proofs the assistance of Bettie Frankl, Nancy Hall, Ann Kelley, and Irene Painter has been vital. We are indebted to Francis X. Kennelly for drawing most of the figures in this text, and to Harvey L. Waterman for the circuit diagrams in some of the earlier chapters.

THE AUTHORS

ARLINGTON, VA.
May, 1950

vii

CONTENTS

PART II. COMPUTING SYSTEMS

Part I

THE BASIC ELEMENTS OF
MACHINE COMPUTATION

CHAPTER 1

INTRODUCTION

Many of the practical problems of science, business, and warfare are essentially computational. For example, scientific and business conclusions may be drawn by subjecting large quantities of observational data to appropriate statistical treatment. Military plans may be based on calculations from quantitative hypotheses involving numbers of men, numbers of items, numbers of distance units, and the like. The need for extensive calculations in the development of science has existed and continually expanded.

The existence and importance of these and an infinite number of other general computational problems have fostered the development of machine aids to computation. This volume touches on numerical methods and introduces the reader to existing computing techniques and machines. The presentation is intended to show how to formulate computational problems in ways related to the physical structure of machines. The final chapters list both the physical components of which the present-day computing instruments are comprised and other components most likely to find uses in the near future. A discussion of factors governing the choice of components is included also.

The term *component* is used throughout this survey to define any physical mechanism or mathematical method which is used as a tool in automatic computation. The term is applied either to an abstract concept or to an item of physical equipment. The scope of this survey includes not only a treatment of the design and operation of physical mechanisms but also the arithmetical and analytical procedures which form the basis of solution of problems reduced to numerical form.

Computational machines are classified as either digital or analog computers. A digital device, as the name implies, is one which performs mathematical operations with numbers expressed in the form of digits which can assume only discrete

values.   The results yielded by such a device are expressed in digits.   The precision of the computed results from a digital machine depends upon the number of digits it can handle, assuming that the actual operations are all performed accurately. (This is the equivalent of assuming that the machine is not out of order and that the operator makes no mistakes.   A more detailed discussion of the concepts of precision and accuracy is included in Chap. 11.)

An analog computer is one in which numbers are converted for purposes of computation into physically measurable quantities such as lengths, voltages, or angles of displacement.   Computed results are obtained by the interaction of moving parts or electrical signals related in such a manner as to solve an equation or perform a given set of arithmetical or mathematical operations.

The precision of the results which can be expected from any given analog device, as contrasted with a digital device, depends upon the precision with which the device is fabricated, the skill and uniformity with which it is operated (including its dependence upon outside factors such as line-voltage variations), and the precision with which the answer can be read if the final conversion to digital form is made by reading a calibrated scale. In short, it is subject to the systematic and human errors which are inherent in the use of any measuring apparatus.

A slide rule is an example of an analog computer.   In this device, lengths correspond to the logarithms of numbers.   Logarithms are added and subtracted, and numerical answers are read from the scale.   An example of a digital computer is the abacus. The Chinese who uses this device assigns discrete numerical values to the beads and obtains precise digital answers to his problems by proper manipulation of the beads.

It is hoped that those readers who are familiar with the mathematical background on which this survey has been based will find the descriptions of physical elements and their interrelationships of interest.   For those well grounded in electrical engineering and particularly in electronics, it is hoped that the chapters dealing with mathematical methods and with computing systems will be sufficiently straightforward to require no unreasonable special study for thorough understanding.

For those readers who are particularly interested in any isolated phase (mathematical components, physical components,

or computing systems), the following classification of the several chapters will be found useful. The chapters in each class are essentially independent of those in other classes.

A. *General*
   Chapter 1. Introduction
   Chapter 2. Preliminary Considerations
B. *Mathematical Components*
   Chapter 6. Arithmetic Systems
   Chapter 7. Numerical Analysis
C. *Physical Components and Methods*
   Chapter 3. Counters as Elementary Components
   Chapter 4. Switches and Gates
   Chapter 13. Arithmetic Elements
   Chapter 14. Transfer Mediums
   Chapter 15. Data-conversion Equipment
   Chapter 16. Special Techniques and Equipment for Possible Use in Computing Systems
   Chapter 17. Factors Affecting Choice of Equipment
D. *Computing Systems*
   Chapter 5. A Functional Approach to Machine Design
   Chapter 8. Desk Calculators
   Chapter 9. Punched-card Computing Systems
   Chapter 10. Large-scale Digital Computing Systems
   Chapter 11. Analog Computing Systems
   Chapter 12. The Form of a Digital Computer

Additional introductory material on the subject of computation both by analog and by digital methods will be found in the following bibliography, containing three volumes all of which have been published since most of the material in the following pages was prepared. Hartree[2] is concerned primarily with the mechanisms of computing, while Berkeley[1] and Wiener[3] devote their efforts respectively to more popular discussion and to the philosophy of the subject.

REFERENCES

1. Berkeley, E. C., *Giant Brains* (John Wiley and Sons, Inc., New York, 1949).
2. Hartree, D. R., *Calculating Instruments and Machines* (University of Illinois Press, Urbana, Ill., 1949).
3. Wiener, Norbert, *Cybernetics* (John Wiley and Sons, Inc., New York, 1948).

CHAPTER 2

PRELIMINARY CONSIDERATIONS

2-1. Basic Components of Automatic Computation

The basic components required for a general-purpose computing machine consist of (1) the input system; (2) the arithmetic techniques; (3) the system of operations, which includes arithmetic elements, storage, and control; and (4) the output system. Each of these components will be treated here in order to present an introductory view of the fundamental requirements of automatic computation.

2-1-1. Input Systems. This is a general term for the components and processes by means of which problems are introduced to the machine. For example, the input system associated with a standard desk calculator is comprised of a manually operated keyboard of decimal digits and commands (*e.g.*, *add*, *multiply*). The input more nearly defines the use to which a machine can be put than does any other component; there is no universally applicable input medium.

Machines can be divided roughly into two classes according to their expected use, and each of these classes can be divided further into two subdivisions according to versatility:

Class IA—Data-reduction equipment, general-purpose.
Class IB—Data-reduction equipment, special-purpose.
Class IIA—Mathematical equipment, general-purpose.
Class IIB—Mathematical equipment, special-purpose.

The input requirements of these four types differ radically.

Data-reduction machines typically receive voluminous data from physical measurements. Usually manual transcription time (the time spent in reading and transcribing the data from the measuring instrument) exceeds manual reduction time (the time spent on the ensuing arithmetic calculation). The use of automatic computing equipment for large-scale reduction of data will be strikingly successful only if means are provided for the

automatic transcription of these data to a form suitable for automatic entry into the machine. For some applications, of which the most prominent are those in which the reduced data are used to control the process being measured, the input must be developed for *on-line* operation. In on-line operation the input is communicated directly and without delay to the data-reduction device. For other applications, *off-line* operation, involving automatic transcription of data in a form suitable for later introduction to the machine, may be tolerated. These requirements may be compared with teletype operating requirements. For example, some teletype machines operate on line. Their operators are in instantaneous communication. Other teletype machines are operated off line, through the intervention of punched paper tape. The message is preserved by means of holes punched in the tape and is transmitted later by feeding the tape to another machine. This method permits fast regular transmission independent of operator variations. Line time is thus conserved at the expense of elapsed communications time.

The typical mathematical machine, as contrasted with the data-reduction machines, receives input data generated in a human brain. These data are not likely to be numerous, and automatic entry from the human brain is not presently feasible.

2-1-2. Arithmetic Techniques. When a project involving extensive computation has reached the point in its development where the use of large-scale computing equipment is being planned, it is usually true that a fairly well defined field of problems has been formulated. These are the problems the computing equipment must solve. Between the general formulation of this field and the production of solutions of problems from the field come the specific arithmetic formulations of problems and the development and construction of equipment which will yield solutions economically. Roughly, then, the project can be thought of as including a mathematical aspect and an engineering aspect. A precise definition of the boundary between these aspects is not worth attempting; one merges into the other. Neither part of the problem can be solved independently of the other. The solution must be evolved from a utilization of mathematical processes (including arithmetic, logic, etc.) and engineering processes which together yield the desired solution with acceptable economy.

The central meeting ground of these two aspects of the computing program is an area which includes both the set of elementary operations which the device is capable of performing and the general procedure for communicating to the device the set of orders it must perform in solving a problem. The term *arithmetic techniques* is used to describe this area. The term denotes, therefore, those arithmetic and logical operations and combinations of operations which are used by the machine operator in the application to any problem of abstract arithmetic processes. Thus, when an operator uses a simple adding machine to obtain a product of two numbers, he is performing the process known as multiplication via the arithmetic technique of iterative addition.

The speed with which a given problem can be solved with a particular machine depends primarily upon the number of operations (*e.g.*, multiplications) which the machine must perform in solving the problem and upon the time spent by the operator in preparing a program of commands ordering the machine to perform the required operations. For maximum versatility and speed, the machine must be capable of a maximum number of different operations. Generally speaking, the more extensive the variety of possible operations, the less complex will be the program of operations for a given problem. On the other hand, the variety of operations must be restricted to some extent for practical and economic reasons. Ingenuity in preparing programs of operations must be substituted to some extent for variety of available operations.

Therefore, in setting performance specifications for a machine, the choice of an optimum set of available operations depends upon the field of problems which the machine must solve, the speed with which solutions are actually required, and the practicability of constructing physical components which will perform the various mathematical processes which might be desirable. These factors are related to the choice of an optimum set of operations through the arithmetic techniques which will be employed.

2-1-3. System of Operations. The portion of a computer which actually performs the mathematical functions in the solution of a problem may be termed the *operations system*. It consists of arithmetic units, which perform the operations of

addition, subtraction, multiplication, and division; storage units, which provide means for holding information for reuse; and control elements, which interpret commands and initiate arithmetic processes.

The arithmetic units of the machine are only partially determined by the determination of the elementary arithmetic operations they are to carry out. The required accuracy, speed, and versatility of the machine all affect the design of the arithmetic units. If the operation has to be synchronized with outside phenomena, which is the case frequently in on-line operation, this too must be taken into account in the design of the arithmetic elements. Finally, it is true that the arithmetic elements must be designed in such a way as to permit convenient information transfer between the arithmetic units and the storage units chosen for the machine.

It has already been stated that storage units are required to hold information for later reuse. This implies a concept of time which may be used to distinguish between types of storage: quick-access, or high-speed, storage; slow-access, or low-speed, storage; and intermediate storage. These terms, however, merely denote the form of the storage and omit any reference to the actual purpose of storage.

A more descriptive definition of storage facilities may be obtained by considering the use of storage in a typical example. When a piece of paper, a pencil, the physical actions of writing, and the mental processes of calculations are involved in the solution of an arithmetic problem, it is readily seen that there are four kinds of storage required: (1) storage of the original data of the problem upon the paper; (2) mental storage of carry-overs in processes of addition, subtraction, multiplication, and division; (3) storage of intermediate results upon the paper; and (4) storage of the final solution upon the paper. Of these four kinds of storage, all but the second require transfer to a medium external to that in which the fundamental arithmetic processes are performed. This suggests that storage may be further defined as either internal or external; *i.e.*, either for immediate automatic reuse in the arithmetic units, or for reuse in the computation process at a time determined by the programming of operational commands. For the former, quick access time is convenient; for the latter, a lower speed is per-

missible. In fact, a lower speed may be preferable. The factors influencing the choice of storage speeds, the amount of storage, and storage mediums are more fully discussed in Chap. 14.

The third part of the system of operations is the device which controls the computation process, ordering the arithmetic units to perform their various functions upon various numbers in proper sequence. These commands or orders to the arithmetic units may be of two types: (1) orders included in a program of such orders, drawn up by a human operator; or (2) orders derived from previous machine operations. The complexity of the program of orders required for the control of the machine depends upon the type of application for which the machine was designed. Consider, for example, a machine designed for solving a single complex problem, requiring many basic operations to be performed in a particular sequence. This is a special-purpose machine. In theory, a machine may be so highly specialized that after the values of parameters are specified, the only external command required is *start*. The rest of the commands are permanently built into the machine in some manner or are derived automatically from operations performed earlier in the course of the machine's solution. At the other extreme, a machine may be so general that an infinite variety of sequences of operations is possible, either by rearrangement of some temporary wiring system, by revision of the program encoded on the input medium, or by both.

For general-purpose machines, the present tendency is to encode the commands in a numerical code. The machine reads the coded commands inscribed in proper sequence on the input medium. The code is usually composed of groups of digits; each group is composed of a part which has to do with the operation involved and a part which has to do with the location in storage of the number or numbers to be operated upon. There may also be a part which has to do with the choice of the next command to be carried out.

2-1-4. Output System. The output problem is similar to the input problem. The output, as well as the input, must be carefully matched to the utilization expected of the machine. If the instrument is to control a process, and if this process continues during the calculation, then the output may be some sort of on-line device possibly including facilities for translating a

digital output to continuously variable control signals. If the output is a printing device, it is important to avoid letting this printer retard the operation of the whole machine, as it will if the output plans are overextensive. Another consequence of overextensive output plans is the generation of more printed material than anyone can read.

CHAPTER 3

COUNTERS AS ELEMENTARY COMPONENTS

3-1. Introduction

In effect, every high-speed computing machine is composed of an arrangement of fundamental components. It seems reasonable here to study these elementary components in some detail; later it will be seen how they may be assembled into a complete equipment. It is fortunate that the number of different types of such fundamental components is limited. This makes it possible to deal with much of the subject by treating only one general type of *elementary potential digital computing component* (hereafter referred to as EPDCC).

We define an EPDCC as any system, mechanical, electrical, electronic, electromechanical, electromagnetic, or other, which has the following properties:

1. It may assume any one of a discrete set of stable states, the number of stable states in the discrete set being fixed in advance.
2. It may influence other components, including other EPDCC's, or be influenced by them in different ways, depending upon the state in which the components are set.

The application of these components to computing has been indicated in the introductory material; it will be more explicitly indicated in later chapters. In this chapter, discussion will be limited to a few elementary mechanical devices, to electrical circuits involving vacuum tubes, and in particular, to those circuits based on the flip-flop principle.

3-2. Mechanical and Electromechanical Devices

A common EPDCC is the 10-position adding-machine wheel which is used in mechanical desk computers, and which satisfies the definition of an EPDCC in all respects. This mechanical

counter can be transformed into an electromechanical counter by the addition of 10 position contacts and pickup wipers. With this modification, the stable state of the wheel is determined electromechanically, and the influence of the wheel on other components is completely electrical.   An example of such a device is the counter developed for card-tabulating purposes. This device comprises a 10-position wheel, the exact position of which is controlled by an electromechanically operated clutch. The wheel position is read by means of a set of 10 contacts. This type of electromechanical unit also fits the definition of an EPDCC.   Among the few existing large-scale computing machines there is at least one, the IBM Automatic Sequence Controlled Calculator presented to Harvard University in 1944 and generally referred to as the Mark I, the entire action of which is based on the use of such electromechanical devices.[1,13,14]

There is a fundamental limitation, however, to the maximum speed at which electromechanical devices can operate, owing to the inertia of the mechanically moving parts.   Present practice indicates that the maximum speed for reliable operation of any electromechanical EPDCC is approximately 100 counts per second.[21,36]   Since the great majority of future large-scale computing machines will demand basic speeds considerably in excess of this figure, it is necessary to go to some other type of EPDCC to achieve the desired results.   The use of electron tubes offers such a possibility, and the remainder of this chapter will deal only with those EPDCC's which utilize electron tubes.

3-3. The Flip-flop Principle

Historically, the flip-flop principle has been known since 1919,[11] but its application to a full-scale digital computing machine does not appear to have been realized until 20 years later.[9]   Prior to this conception certain special forms of flip-flop had been developed, but mainly for the purposes of counting atomic particles.   Since that time the basic principle of the flip-flop and the ring counters has been altered variously for particular purposes.

The flip-flop configuration of vacuum tubes is characterized by the fact that there are two plate-to-grid couplings and a common bias arrangement between the two halves of the circuit; this coupling and bias arrangement causes the circuit as a whole

to remain in either one of its two stable states until the application of the signal which changes it to the other.

Because the flip-flop circuit is so fundamental to practically every type of EPDCC, we shall proceed first to a description of this arrangement of two triodes (Fig. 3-1) having two stable states.[11,31,32,34,38,43]   In this circuit the plate of the first triode V_1 is connected to the grid of the second V_2, and the plate of the
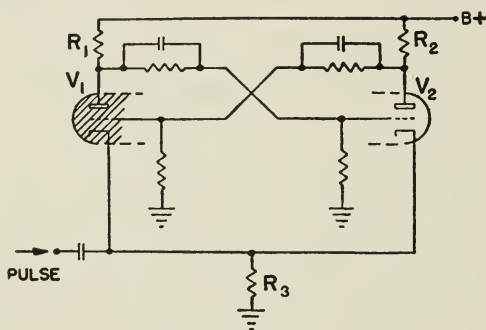
Fig. 3-1.   Flip-flop arrangement of two triodes.   Each successive negative pulse on the single input lead will reverse the operating state of the circuit. In the state shown (V_1 conducting) a negative pulse lowers the cathode potential of V_2 below its grid potential, causing V_2 to start conducting, which in turn lowers the grid potential of V_1.

second is connected to the grid of the first.   An increase in current through either triode tends to lower the grid potential of the other triode.   For example, an increase of the plate current of V_1 causes a greater potential drop across the resistor R_1; this lowers the grid potential on V_2.   Accordingly, the plate current of tube V_2 is decreased, which, in turn, lowers the potential drop across the resistor R_2, causing an increase in the grid potential of tube V_1.   With a higher grid potential on tube V_1, the plate current through this tube is increased.   This general condition of instability prevails and continues to drive the operation just described until tube V_1 is conducting at saturation and tube V_2 is cut off.   The plate current of the conducting tube produces a voltage drop across both its own plate resistor and the common cathode resistor R_3.   The values of the resistors are selected so that the potential at both cathodes, with either tube fully conducting, is above the grid potential of the nonconducting tube. This elementary configuration can be caused to shift from one

stable condition to the other by introducing circuitry designed to respond to various stimuli. There are two main types of shifting stimuli which can be introduced to the flip-flop circuit, and both of these have application in computing devices.

The first type of stimulus utilizes pulses of a fixed polarity which are introduced at a single input terminal. The action of the flip-flop is such that it reverses between its two stable states each time a new input signal is introduced: it is similar in action

FIG. 3-2. Binary counting circuit with a single input lead.

to a pull-chain light switch which reverses its state (on or off) each time the chain is pulled. Thus it indicates, by its state, whether the total number of input pulses which have been applied is odd or even.

Such a device is classified as a *modulo 2 counter*. The term modulo means, literally, "with respect to the modulus, or standard of measurement." Two numbers are congruent with respect to an integral modulus if their difference is divisible by that modulus. That is, $a \equiv b \pmod{c}$ if $a - b$ is divisible by c (or if a and b have the same positive remainder when divided by c).

A modulo 2 (or binary) counter is essentially the main elementary component from which counters and accumulators are assembled. A practical example of such a binary ring circuit having a single input lead, as found in the plus-minus indicating device of the ENIAC, is illustrated in Fig. 3-2. In this illustration the tubes V_1 and V_2 act as trigger tubes. Whenever a positive signal arrives on the one input lead, both of these tubes

begin to conduct.   This causes the grids and plates of both V_3
and V_4 to take a negative swing.   Since V_4 is already noncon-
ducting, it has no effect there, but V_3 is cut off.   As V_3 cuts off,
its plate becomes more positive, and, through condenser C_3,
the grid of V_4 also takes a positive swing.   As V_4 begins to con-
duct, its plate takes a negative swing, and condenser C_2 further
causes the tube V_3 to cut off.   Note that the tubes utilized are
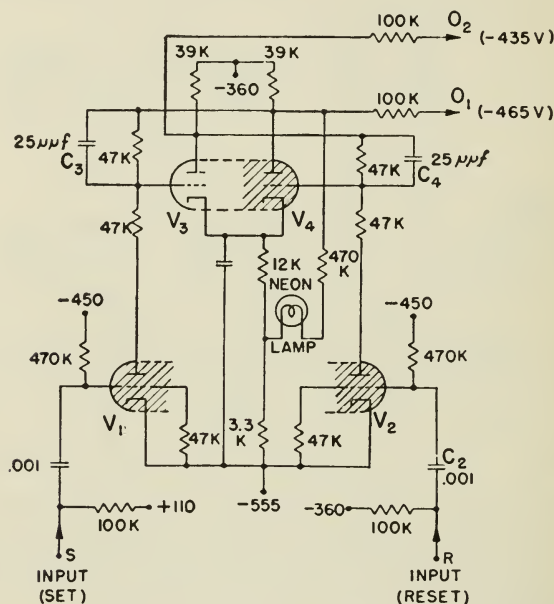
Fig. 3-3.   Binary counting circuit with two input leads.

actually dual triodes so that V_1 and V_2, and likewise V_3 and V_4,
are in a single envelope.

The second type of actuating signal, or stimulus, for altering
the stable state of the flip-flop is introduced on one of two input
terminals.   If the input signal arrives on the first terminal, it
causes the flip-flop to assume the first of its two stable states and
remain in this state thereafter until such time as a signal arrives
at its other input terminal, whereupon it will reverse.   This
action is analogous to that of the ordinary wall-type toggle switch,
which can be thrown in either of two directions and which will
remain thrown in this direction until such time as it is thrown the

other way. A practical example of such a side-stable electronic circuit is shown in Fig. 3-3. Except for a few circuit values, the only difference between this and the circuit of Fig. 3-2 is in the number of input leads for introducing the actuating signal.

3-4. Pulse-sensitive Flip-flop Circuits

Of the two types of flip-flop circuits described above, the one having a single input terminal is of interest mainly because a whole series of pulse-sensitive counters can be built from it. When considered as a single unit operating by itself, it is essentially an odd-even, or modulo 2, counter, as mentioned previously. If a number of such modulo 2 counters are connected so that the output of one feeds the input of the next, a radix 2 counter is formed, each successive flip-flop registering the next binary digit. (The reader who is not familiar with the concepts of binary arithmetic is referred to Chap. 6 for an exposition which will explain more fully the meaning of this paragraph.) If the number of tubes in the flip-flop itself is increased beyond two, the circuit can count to a modulus higher than 2. Stating the proposition generally, a number of modulo r flip-flops used together constitute a counter system operating to the radix r. It is also possible to connect a number of flip-flop circuits in such a manner that each complete flip-flop represents one element in a ring of such twin elements, thus forming another type of modulo r counter, which can be compounded into a radix r counter.

All these configurations are characterized by the presence of a single input terminal and further by the fact that they are all fundamentally compounded from a number of modulo 2 flip-flop elements. Such compounding is necessary because, although it is easy to think of all radices as having substantially equal mathematical merit, the electron tube is unable to assume more than two discrete stable states with any assurance of reliability.

3-5. Radix 2 Counters

One of the simplest schemes for counting the pulses in a series is the radix 2 counter, which may be assembled by coupling together several modulo 2 counters.[15,24,34,43] Figure 3-4 illustrates how this may be done. The circuit shown contains four
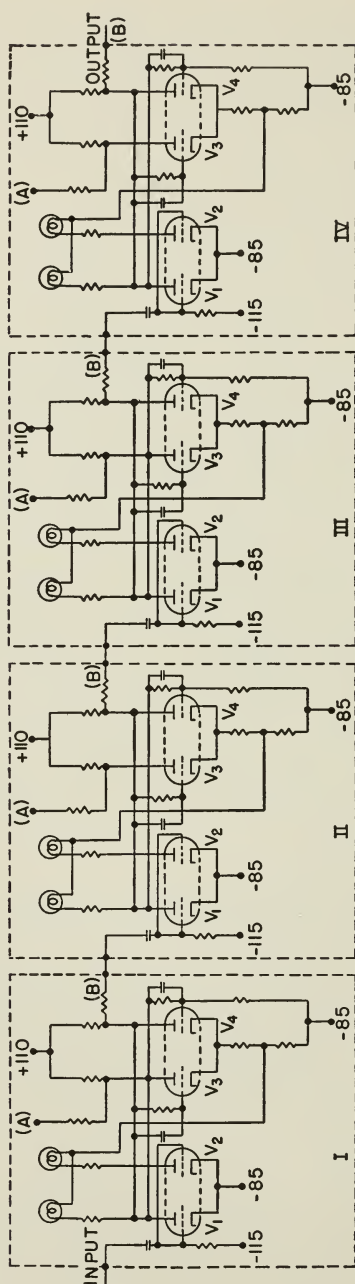
FIG. 3-4.   Circuit assembled from four modulo 2 (or binary) counters.   This circuit has 2^4 or 16 stable states.

modulo 2 counters denoted by the blocks with the Roman numerals. Counter I receives the pulses to be counted; these pulses are introduced via the input lead of a modulo 2 counter of the type illustrated in Fig. 3-2. Each pulse causes counter I to reverse its state. The output from counter I drives counter II. The output is taken from point B in Fig. 3-2 and is coupled, possibly through a pulse-sharpening circuit, to the input lead of counter II. In the same way the output from point B of counter II is taken to the input of counter III, and the output at point B of counter III is taken to the input of counter IV. Now it is clear that each time tube V_4 of counter I, II, or III changes from a conducting to a nonconducting condition, a positive pulse will be introduced to the input of the next counter in the series, causing it to change its state. Assume that a counter's condition is denoted by the symbol 0 if tube V_4 is blocked and tube V_3 is conducting. Then a pulse changing the counter's status from 0 to 1 introduces a negative pulse which has no effect on the next counter in the series. However, if the status is changed from 1 to 0, a positive pulse is introduced to the next counter, and this pulse changes the status of the next counter. Thus, in the example of Fig. 3-4, if the configuration of the four counters were 0000 at the beginning and if 16 pulses were introduced, the successive configurations would be 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111, and 0000.

The representations 0000 through 1111 are the radix 2, or binary, representations of the decimal numbers from 0 through 15. The 16th pulse returns the system to zero. This representation is described more fully in Chap. 6; it should be noted here, however, that from the configuration of the system the number of pulses which have been introduced can be deduced, provided this number is less than 16. By adding more counters modulo 2 with similar coupling, the capacity of the circuit can be increased. The number of pulses which can be counted with m modulo 2 counters is $2^m - 1$.

The reading of these counters is brought about without substantial decrease in their functioning frequencies by establishing reading leads at points A of the individual modulo 2 counters. Simple radix 2 counters can be made to operate at a counting rate of 10 pulses per microsecond.

3-6. The r-triode Counter, Modulo r

For small values of r, such as 3, 4, and 5, it is feasible to extend the circuit of Fig. 3-1 beyond two, so that, of r triodes, one is conducting and $r - 1$ are blocked.[2,15,17,37]   Thus, a pulse introduced to the input lead causes the originally conducting tube to
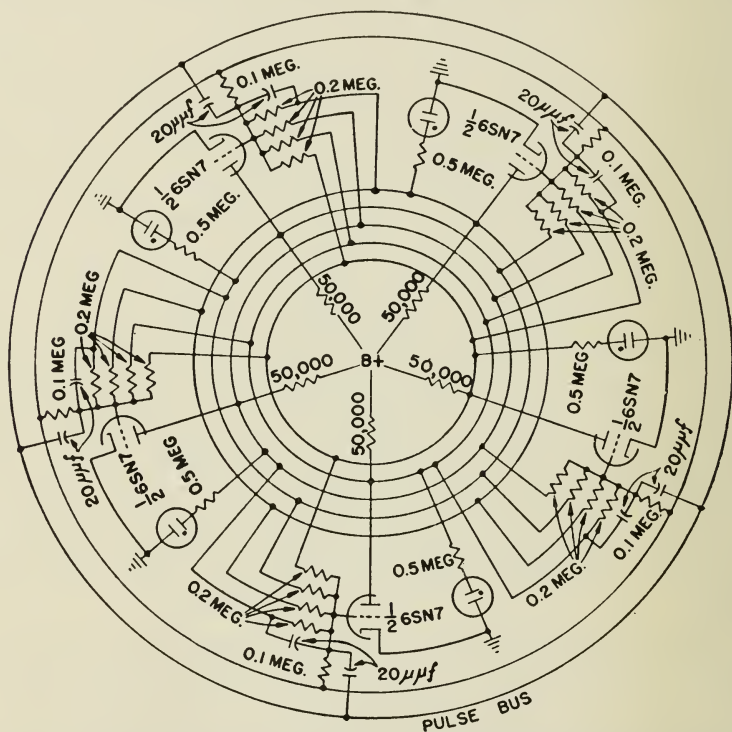
FIG. 3-5.   Five-triode ring.   This circuit has five stable states, in each of which one, and only one, tube is conducting.   It is a modulo 5 counter.

become blocked and causes a tube cyclically advanced beyond the originally conducting tube to begin to conduct.   Figure 3-5 illustrates such a circuit[34] with five triodes so wired that the direction of progress is counterclockwise.   In this circuit the grid resistors are chosen carefully so that the condition in which none, or more than one, of the tubes is conducting leads to a grid voltage too high or too low for stable operation.   The grid of each tube obtains its operating potential from a voltage divider comprising five resistors, four from the plates of the four other

tubes, and one from a permanent bias bus. The values are such that, for any one tube, only two conditions can obtain: (1) if the other four are all cut off, the grid voltage will be sufficiently high that full conduction of this tube is assured, or (2) if one of the other four tubes is conducting, then the grid voltage will be low enough to assure that this tube is cut off, thus preserving the one-on–four-off ratio.

A negative pulse on the input-pulse bus will cause the conducting triode to cut off, thus raising its plate potential. This rise is then transmitted to the other four tubes. However, only the associated voltage-divider resistor of the next tube cyclically advanced is shunted by the necessary crossover condenser; hence it is the only one to achieve conduction as a result of the input pulse. [It is also possible[41] to construct a circuit such that only one tube is cut off and the rest are fully conducting. Another possibility[23] is a ring composed of an odd number of tubes, r, of which in each of the ring's r stable positions $(r + 1)/2$ tubes are nonconducting.]

Read-out from the modulo r counter illustrated in Fig. 3-5 is accomplished at the plates of the tubes, as illustrated by the small neon indicator lamps. Its operation is limited to values of r which are sufficiently small to permit adequate discrimination between conditions governing operation of the individual tubes. In the configuration shown, the plate potential of the conducting tube may be well below that of the nonconducting tubes. The difference in potential between the plates and the grids of the nonconducting tubes is $1/r$ times the difference between the plate of the conducting tube and the plates of the nonconducting tubes. If r becomes sufficiently large, $1/r$ becomes sufficiently small to make operation of the circuit uncertain. In practice, with careful design, values of r up to 7 or 8 appear to be feasible; values much larger than this have led to difficulty. It does not appear possible to make $r = 10$ (for decimal arithmetic) with this scheme, except at considerable decrease in reliability. However, a binary counter and a quinary counter can be combined to form another type of decimal counter.[15,23,34]

There is another form of r-triode modulo r counter which is practical and which has achieved wide usage in certain specialized applications where the ultimate in speed is not required. This is the ring consisting of gas-filled triodes,[15,40,42,43] popularly

known as thyratrons. The thyratron, as contrasted with the conventional tube, has only two possible stable states for all grid potentials: fully conducting and completely nonconducting. The grid of such a tube, by itself, serves to trigger the tube from the nonconducting to the fully conducting state, after which it loses control of the action of the thyratron until the tube is extinguished by other means. Figure 3-6 represents a decade of thyratrons, forming a modulo 10 counter. In each of its 10 stable states, one, and only one, tube is conducting. Because of the
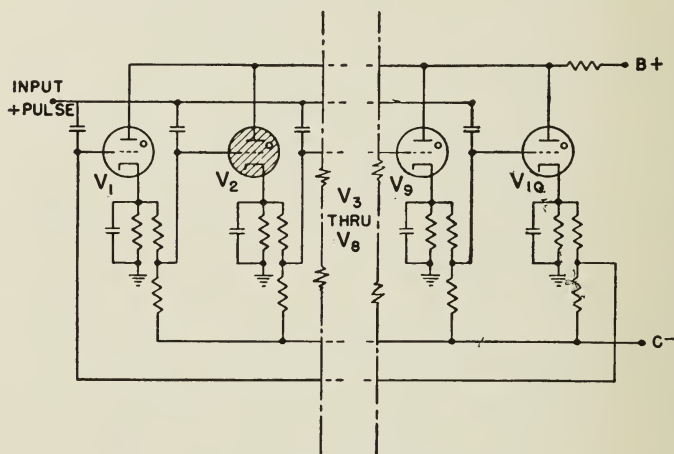
FIG. 3-6. Thyratron ring counter.

voltage divider in the cathode circuit of the conducting tube, the grid voltage of the tube following it is higher than the grid voltages of the other nonconducting tubes. The tube is therefore said to be *primed*. The next positive pulse raises its grid sufficiently to cause it to start conducting. Almost simultaneously the preceding tube is cut off. This is the consequence of a drop in its plate voltage while its cathode voltage remains nearly constant. The drop in plate voltage is occasioned by the additional current (to the tube which has begun to conduct) through the common plate resistor; at the same time, the cathode potential is held nearly constant by the action of the capacitor in the cathode circuit.

This gas-tube ring counter can be used in applications requiring a rate up to about 10,000 counts per second, this limit being established probably by the time constant of the grid circuit.

The ultimate limitation on counting speed is set by the deionization time of the gas molecules. It is probable that 50,000 counts per second is about the best that will be reliably accomplished with production gas-tube counter circuits. Higher speeds, of the order of 1 million counts per second, demand the use of vacuum tubes exclusively, although experimental gas-tube rings have been made to work as fast as 110,000 to 200,000 counts per second.[39]

3-7. The $2r$-triode Counter, Modulo r

For high-speed counting to a radix greater than 2, the $2r$-triode counter modulo r unit is probably the only answer. The most advanced present form of this type of counter will be found in the decade rings of the ENIAC, where it forms the basic counting unit. An example of such an arrangement taken from the ENIAC[6,34,39] is shown in Fig. 3-7.

It will be noticed that there is a strong resemblance between the circuit of Fig. 3-7 and the more elementary circuit of Fig. 3-1. The former is merely a tenfold repetition of the latter. It is interesting to note, however, that this circuit can be operated more reliably if the coupling between flip-flop stages is effected through the regular control grids while the single input terminal is connected to the cathode circuits on only one side of each flip-flop. Thus the condition of the ring as a whole assures that there is one flip-flop in the "set" or abnormal state and that the other nine are in the "reset" or normal state. Successive input signals cause the flip-flop stage previously in the abnormal state to resume its normal condition, meanwhile simultaneously driving the next succeeding flip-flop into the abnormal state.

As compared with the counter described in the previous section (*i.e.*, the one with a single triode per digit of the modulus), the particular counter described above does not appear to have any practical limit as to the size of the modulus, although of course considerably more driving power may be required in the case of the larger rings. For computing purposes, 10 seems to be the highest modulus usually desired.

It is possible to combine a number of modulo 10 counters in such a manner that a carry-over is provided between successive rings; this in effect gives us a radix 10, or decimal, counter capable of handling as many digits as there are rings provided.
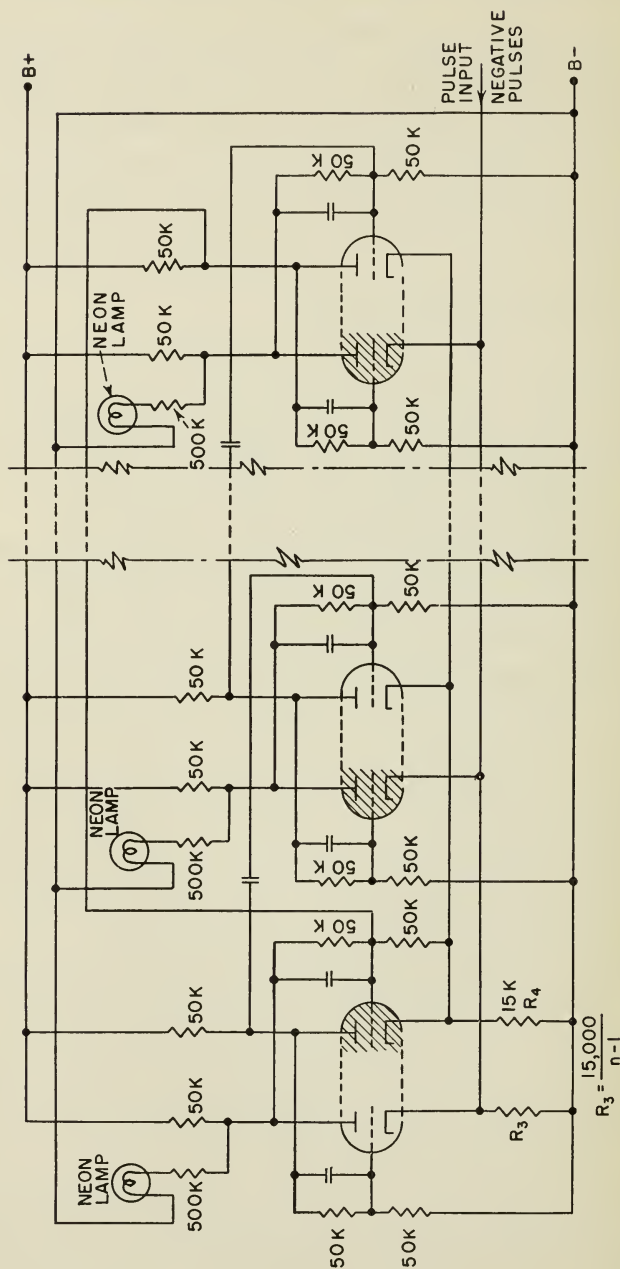
FIG. 3-7.   A counting circuit used in the ENIAC.   The number of stable states which can be assumed by such a circuit is equal to the number of pairs of triodes it contains.   To count to the modulus r it must contain $2r$ triodes.

Again there seems to be no limit to the number of rings that can be combined in this carry-over fashion.

As a final example of this type of counter, we refer again to the ENIAC, which is capable of handling up to 20-digit decimal numbers. One accumulator in the ENIAC contains 10 rings of 10 flip-flops each, or a total of 200 tubes, which are enclosed in 100 envelopes. Actually each such 10-digit accumulator requires almost three times this number of tubes in order to handle certain auxiliary functions such as signal shaping, carry-over, add outputs, subtract outputs, and resets.[34,38]

The choice between the r-triode counter and the $2r$-triode counter is completely dependent upon the value of r itself. In the ENIAC, for example, there are parts of the machine where $r = 10$ (all the decimal arithmetic), others where $r = 6$, 9, 11, 13, 14, or 20 (various special-purpose rings), and several parts where $r = 2$ (plus-minus indicators). In all the cases where $r > 2$, experiments showed that the $2r$-triode counter was required; for $r = 2$, however, the r-triode ring, or elementary flip-flop, was preferable and, in fact, was capable of operating about twice as fast as the $2r$ decade rings.

3-8. Other Types of Flip-flop and EPDCC

In addition to the foregoing general types of EPDCC's, there are a number of special ones which differ in engineering aspects and which have application in certain special circumstances. These will be summarized briefly below.

There is one type of flip-flop which is particularly interesting in view of its simplicity and the small number of component parts required. This is the pentode-coupled flip-flop (Fig. 3-8), in
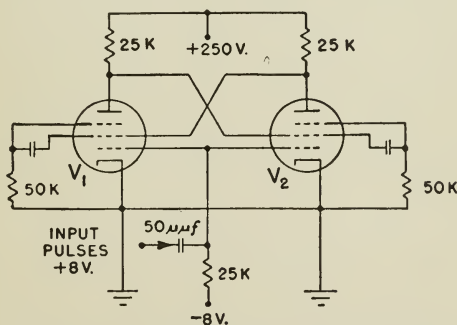
FIG. 3-8. Pentode-coupled flip-flop circuit.

which the screens and plates of the two pentodes are cross-connected.[3,28] The action of this circuit can be readily understood by considering the case in which V_1 is fully conducting; its plate must be at a low potential, and therefore the screen of V_2 is at an equally low potential. This being the case, the right pentode must be essentially nonconducting, which means that its plate is at a high potential. Hence the screen of the left pentode, which must be at the same high potential, is serving to keep the left pentode in its fully conducting state. Positive input impulses serve to switch this flip-flop alternately into its two stable states, just as in the case of the triode flip-flop in Fig. 3-2. This circuit does not appear to have had the exploitation and development which its simplicity warrants. It operates satisfactorily at a speed of at least 200,000 counts per second; however, it appears likely that with an appropriate amount of effort directed toward its fullest exploitation this circuit is capable of the same speeds as are now achieved in the more fully developed triode flip-flop. It is quite possible that the advent of printed-circuit techniques,[5] mentioned in Sec. 16-6 of this book, and the recently released wartime developments of miniaturized vacuum tubes may serve to direct further interest toward the pentode–flip-flop circuit.

The pentode–flip-flop circuit has been incorporated into a decade counter.[29] This particular decade differs from others previously described in that there are only five pairs of pentodes constituting the entire decade ring; thus, we might classify this as an r-pentode counter, modulo r (with the proviso that r be an even number). This particular circuit is also interesting in that its normal mode of operation is one in which there are $r/2$ consecutive pentodes in the conducting state and $r/2$ consecutive pentodes in the cutoff state. (The two pentodes forming any one pair are $r/2$ apart in the ring.) As opposed to the more conventional counter ring, which has one element conducting and $r - 1$ elements nonconducting, this may appear to complicate the read-out problem; however, a rather ingenious and unambiguous visual read-out system has been developed for this pentode ring, and doubtless the same principle could be applied to the read-in and read-out problems which would arise if such a ring were integrated into a complete computing machine. This five-pentode-pair ring has been built and tested up to

100,000 counts per second or more. It compares favorably with the performance of the standard production models of triode-pair decade counters (180,000 pulses per second), which were brought to an advanced state of development for the ENIAC. More recently this same circuit has been further developed,[33] and it now appears capable of achieving a speed of 500,000 counts per second.

It will be noted that all the foregoing decade rings operate with a fixed direction for the progression of the count. This is true for all the rings (r-triode, 2r-triode, r-pentode) except, of course, for the trivial case of the binary counter. In general, this is quite satisfactory for any type of computing machine, since the forward direction can be made to serve for addition, while a system of complements to the forward direction can be made to serve for subtractive purposes. It is a relatively simple problem to provide a few extra tubes for such a system of numbers and complements; this is actually the method employed in the ENIAC.[38] However, it is possible to fabricate a ring which can be driven either forward or backward. At least one such device has actually been constructed; the choice of drive direction in this ring was determined by the choice of which of two pairs of input terminals received the input pulses.[30] Since these flip-flops and counter rings are now in a state of rapid development, it is not possible to include here all the numerous finer engineering points.

One more type of electronic counter should be included here, although its use to date has been directed more toward a simple counting device than as a component in a full-scale computer. This is the *decade scaler*, or decade counter, which counts in the decimal system, but not in a manner similar to previously described counters (all of which embodied a closed ring or chain of 10 electronic elements). This counter actually employs a chain of four binary-system counters similar to the ones described earlier in this chapter. These would, of course, normally count on a modulus of $2^4 = 16$, but certain trick feedback circuits are employed so that effectively six of the 16 stable states are circumvented in one way or another, thus reducing each chain of four flip-flops to what is actually a chain of decimal counters. This type of circuit has been known for several years;[4,18,26,41] several manufacturers have brought out commercial versions of these circuits. The principal merit of this

modified-binary type of decimal counter as compared to the more conventional ring of 10 lies in the lowering of the input electrostatic capacity, which permits the circuit to achieve a higher operating speed. There is also considerable tube economy as compared to the $2r$-triode ring when r becomes 10 or more.

Two other types of flip-flop might be included here for the sake of completeness and because of the theoretical interest they hold. Neither of them has to date achieved any acceptance in practical computer application. The first of these is a standard cathode-ray tube modified in such a way that the beam, instead of falling on a phosphorescent screen, will fall on one of two collector plates, where it will be picked up and used to control the deflection of the beam so that it continues to remain on this first collector plate.[39] By suitably arranging the geometry of the various internal members and the external circuitry of the tube, a satisfactory flip-flop element has been obtained. If there were enough demand for such a tube, it would not be too difficult to effect the necessary design and development work; such a tube, however, does not appear to offer any great advantage over presently available components, and there appear to be a number of engineering disadvantages to such a scheme.

Another type of flip-flop may be built around the use of any single tube which shows an inversion in its characteristic curve. When such an inversion exists, it may be used to represent, effectively, an area of negative resistance, and, with the proper choice of operating conditions, this may be used to obtain the necessary two stable states.[15] However, these regions of negative resistance are caused by phenomena such as secondary emission, which is notoriously unstable and unreliable. This type of tube does not appear to promise any major gain over present conventional components.

3-9. Economic Considerations

Experience with the cost of existing and contemplated machines (described in Chap. 10) does not serve as a very accurate guide toward establishing a cost of the elementary components, because the cost of the few existing equipments is intimately associated with developmental charges and certain military considerations. For this reason it appears best to look to other sources of data. For the purposes of this estimate, we shall exclude EPDCC's of the elementary mechanical type. The cost figures for desk

computers given in Chap. 8 may be taken as indicative of what mechanical devices are likely to cost.

Manufacturers are presently engaged in making standard models of electronic counting equipment on a limited scale. The Radio Corporation of America is now producing a six-decade decimal counter (operating on the modified-binary system described in Sec. 3-9) which is capable of operating on pulses spaced 1 million per second.[27] This device is supplied with an input for normal pulse counting or with a start-stop input control which operates a self-contained, 1-megacycle, crystal-controlled oscillator. The Potter Instrument Company is also producing a Decade Scaler having only three decades (also modified-binary type) but capable of a speed of only 200,000 pulses per second.[25]

In general, it may be stated that the cost of an elementary flip-flop or a decade ring is not closely governed by the operating speeds involved. This statement is approximately true over quite a wide range of speeds: from about 100 cycles per second, the limit of mechanical counters, to several million cycles per second, the present limit of electronic counters. Speed requirements may dictate the choice of tubes involved and their associated circuitry. So far as flip-flops and rings are concerned, the choice of tubes and circuitry does not greatly influence the cost of the over-all assembly, since labor appears to be by far the greatest factor. For example, a top-speed modified-binary decade employing four flip-flops (eight tubes plus a few extra tubes) will have about as many tubes as a low-speed thyratron ring of the conventional type. The two may, therefore, be regarded as about equally expensive even though their operating speeds may differ by a factor of over 100 to 1.

Labor is the principal item of expense per EPDCC. The use of recently announced laborsaving techniques such as the printed circuit may be expected eventually to reduce this; for example, printed circuits have been accepted in fabrication of radios and hearing aids.[5] Their potentialities have not yet been exploited in connection with high-speed computers. It is not unreasonable to suppose that eventually the quantity cost per EPDCC may be brought down to the neighborhood of $1.

REFERENCES

1. Aiken, H. H., and G. M. Hopper, The Automatic Sequence Controlled Calculator—I, *Electrical Engineering*, Vol. 65, Nos. 8 and 9, pp. 384–391

(August–September, 1946); The Automatic Sequence Controlled Calculator—II, *ibid.*, No. 10, pp. 449–454 (October, 1946); The Automatic Sequence Controlled Calculator—III, *ibid.*, No. 11, pp. 522–528 (November, 1946).

2. Alfven, H., A Simple Scale-of-Two Counter, *Proceedings of the Physical Society (London)*, Vol. 50, pp. 358–359 (May, 1938).

3. Blair, J. M., An Improved Current Integrator, *Review of Scientific Instruments*, Vol. 14, pp. 64–67 (March, 1943).

4. Blume, Richard J., Predetermined Counter for Process Control, *Electronics*, Vol. 21, No. 2, pp. 88–93 (February, 1948).

5. Brunetti, Cledo, and R. W. Curtis, Printed Circuit Techniques, *Proceedings of the Institute of Radio Engineers*, Vol. 36, pp. 121–161 (January, 1948).

6. Burks, A. W., Electronic Computing Circuits of the ENIAC, *Proceedings of the Institute of Radio Engineers*, Vol. 35, pp. 756–767 (August, 1947).

7. Compton, J. S., U.S. Patent 2,402,372; filed Nov. 5, 1942; issued Sept. 10, 1946.

8. Couffignal, P. L., U.S. Patent 2,318,591; filed March 26, 1937; issued May 11, 1943.

9. Crawford, Perry, Jr., unpublished thesis, Massachusetts Institute of Technology (1939).

10. Desch, J. R., U.S. Patent 2,404,697; filed March 21, 1942; issued July 23, 1946.

11. Eccles, W. H., and F. W. Jordan, A Trigger Relay, *Radio Review*, Vol. 1, pp. 143–146 (October, 1919).

12. Grosdorf, I. E., Electronic Counters, *RCA Review*, Vol. VIII, pp. 438–447 (September, 1946).

13. Harvard University, Staff of the Computation Laboratory, *A Manual of Operation for the Automatic Sequence Controlled Calculator, Annals of the Computation Laboratory of Harvard University*, Vol. I (Harvard University Press, Cambridge, Mass., 1946).

14. International Business Machines Corporation, *IBM Automatic Sequence Controlled Calculator*. Pamphlet published by the International Business Machines Corporation (1945).

15. Lewis, W. B., *Electrical Counting* (Cambridge University Press, New York, 1942), Chap. VIII.

16. Lifschutz, H., and J. L. Lawson, Triode Valve Scale-of-Two Circuit, *Review of Scientific Instruments*, Vol. 9, pp. 83–89 (March, 1938).

17. Lifschutz, H., A Complete Geiger-Müller Counting System, *Review of Scientific Instruments*, Vol. 10, pp. 21–26 (January, 1939).

18. Lifschutz, H., New Vacuum Tube Scaling Circuits of Arbitrary Integral or Fractional Scaling Ratio, *Physical Review*, Vol. 57, pp. 243–244 (1940).

19. Michael, F. R., Tube Failures in ENIAC, *Electronics*, Vol. 20, No. 10, pp. 116–119 (October, 1947).

20. Miller, O. R., U.S. Patent 2,407,320; filed Nov. 5, 1942; issued Sept. 10, 1946.

21. Neher, H. V., High Speed Mechanical Recorder, *Review of Scientific Instruments*, Vol. 10, pp. 29–31 (1939).

22. Overbeck, W. P., U.S. Patent 2,404,918; filed May 1, 1940; issued July 30, 1946.

23. Overbeck, W. P., U.S. Patent 2,470,716; filed June 11, 1943; issued May 17, 1949.

24. Phelps, B. E., Dual-Triode Trigger Circuits, *Electronics*, Vol. 18, No. 7, pp. 110–113 (July, 1945).

25. Potter Instrument Company, Flushing, N.Y., *Decade Scaler, Model 2092.*

26. Potter, J. T., A Four-Tube Counter Decade, *Electronics*, Vol. 17, No. 6, pp. 110–113, 358, 360 (June, 1944).

27. Radio Corporation of America, Camden, N.J., *Time Interval Counter, Type WF-99B.*

28. Regener, V. H., Design and Use of Directly Coupled Pentode Trigger Pairs, *Review of Scientific Instruments*, Vol. 17, pp. 180–184 (May, 1946).

29. Regener, V. H., Decade Counting Circuits, *Review of Scientific Instruments*, Vol. 17, pp. 185–189 (May, 1946).

30. Regener, V. H., Reversible Decade Counting Circuit, *Review of Scientific Instruments*, Vol. 17, pp. 375–376 (October, 1946).

31. Reich, H. J., Trigger Circuits, *Electronics*, Vol. 12, No. 8, pp. 14–17 (August, 1939).

32. Reich, H. J., Electronic Switch for Simultaneous Observation of 2 Waves with the Cathode-ray Oscilloscope, *Review of Scientific Instruments*, Vol. 12, pp. 191–192 (April, 1941).

33. Seren, Leo, Decade-Ring Scaling Circuit, *Review of Scientific Instruments*, Vol. 18, pp. 654–659 (September, 1947).

34. Sharpless, T. K., High Speed N-Scale Counters, *Electronics*, Vol. 21, No. 3, pp. 122–125 (March, 1948).

35. Shea, H. G., Electronic True Decade Counters, *Electronic Industries*, pp. 82–84, 136 (Sept. 1946).

36. Shepherd, W. G., and R. O. Haxby, Scale of Eight Impulse Counters, *Review of Scientific Instruments*, Vol. 7, pp. 425–426 (1936).

37. Stevenson, E. C., and I. A. Getting, A Vacuum Tube Circuit for Scaling Down Counting Rates, *Review of Scientific Instruments*, Vol. 8, pp. 414–416 (November, 1937).

38. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *The ENIAC* (University of Pennsylvania, Philadelphia, Pa., June, 1946). 5 vols.

39. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *Theory and Techniques for Design of Electronic Digital Computers* (University of Pennsylvania, Philadelphia, Pa., 1946). 2 vols.

40. Wellman, B., and K. Roeder, Electronic Counter for Rapid Impulses, *Electronics*, Vol. 15, No. 10, pp. 74, 140–142 (October, 1942).

41. Wild, John J., Predetermined Counters, *Electronics*, Vol. 20, No. 3, pp. 121–123 (March, 1947).

42. Wynn-Williams, C. E., The Use of Thyratrons for High Speed-Automatic Counting of Physical Phenomena, *Proceedings of the Royal Society (London)*, Vol. 132, pp. 295–310 (1931).

43. Wynn-Williams, C. E., Thyratron "Scale-of-Two" Automatic Counter, *Proceedings of the Royal Society (London)*, Vol. 136, pp. 312–324 (May, 1932).

CHAPTER 4

SWITCHES AND GATES

4-1. Introduction

The basic operations of any digital computer are the ordinary elementary arithmetic operations, addition, subtraction, multiplication, and division, all of which can be performed on most of the inexpensive desk computers as explained in Chap. 8. A principal feature which distinguishes the organized digital computer (described in Chap. 5) from a mere aggregation of desk computers is the ability to switch data rapidly from one component of the computer to another without the necessity of transcribing the intermediate results to paper either by a hand process or by a printing procedure. Other features are the speed with which the individual operations are performed and the extent of the internal storage. The most distinguishing characteristic, however, is the ability to switch numbers, representing either intermediate solutions or coded instructions, from one arithmetic or control unit to another, between the arithmetic or control units and storage, and between the storage and input or output equipment. This chapter describes the various uses of switching equipment, the various types of switching equipment which are presently available or which may be developed and exploited in the near future, and examples of the practical application of these switches to portions of a complete digital computer.

In the subsequent discussion of such equipment, a much-used type of switch is denoted as a *gate*. This is an on-off switch in which the passage of one electrical signal is controlled by the presence of one or more other signals which hold the switch on or off.

4-2. Switch Functions in Computers

Switches and gating equipment are used in many places throughout a digital computer, and, in fact, the same type of equipment may be used in a number of different places in a computer for widely different purposes. Each application of

such a switch or gate, however, falls into one of three general categories outlined below.

4-2-1. Data Switching. The Automatic Sequence Controlled Calculator, or the Harvard Mark I machine, offers what is probably the most easily visualized example of the switching of data from one part of a machine to another.[1,5] Briefly stated, this device consists of a number of components including storage registers, constant registers, and multiply-divide units, each of which has both an input and an output circuit and a single *number-transfer bus*. Each elementary partial step in a computation is carried out by connecting the output of one unit and the input of another unit to the number-transfer bus. The result attained is the transfer of the contents of one unit to the other unit by means of the number-transfer bus. At any instant, there is at most one output circuit and at most one input circuit connected to the number-transfer bus. After this arithmetic operation has been completed, another pair of registers may be connected to the common number-transfer bus. The choice of which units shall be so connected is made through other elements in the machine (such as the main sequence tapes, which are more fully described in Chap. 10). The actual means of carrying out the connections, disconnections, and reconnections is to be found in a large bank of electromechanical relays.

Much the same sort of data switching may be found in other types of computing machinery. In the ENIAC, for example, instead of just one such bus there are several number-transfer busses, so that elementary operations previously described may be carried out simultaneously in various parts of the machine.[2,11] However, the net effect is that each individual number-transfer bus may have connected to it at any instant at most one output circuit and at most one input circuit. In the EDVAC there is again a single number-transfer bus.[12]

4-2-2. Data Conversion. It frequently happens in computers that data must be converted from one representation to another. An example is the ENIAC, in which data are sometimes stored statically in vacuum-tube circuits and transmitted serially in trains of pulses. The conversion is one in which a given number represented on the ring counter (the digit 7 represented by the seventh flip-flop in the ring, for example) is transformed into a train of (seven) pulses for transmission to some other part of the equipment, such as another ring.

4-2-3. Signal Preservation. A large number of high-speed switches or gates are employed in computing machines for purposes which have nothing to do with the actual arithmetic operations; they are required solely for signal preservation by pulse shaping. This application of gating circuits is to be found only in high-speed computers; such devices as the Mark I and Mark II computers and the Bell Telephone Laboratories and IBM relay computers are limited by physical inertia rather than electrical signal distortion.

Since discrete values are represented by the presence or absence of pulses in various parts of digital computers, it is obvious that the identity of the individual pulses must be properly preserved. In a conventionally designed machine, the pulses are of proper size, shape, and timing at the time and place originated. However, after traveling through a long cable for transfer to some other part of the computer, or after being operated upon in a device such as an adder, or after being temporarily stored in some memory equipment, these pulses may undergo considerable deterioration unless proper steps are taken to reestablish their original form. The pulses may not only be diminished in amplitude (a situation corrected by the insertion of an amplifier), but also may be delayed in time of arrival or spread out on the time scale to such an extent that they may give a false representation of the digital value involved. Electronic gates offer the means for overcoming this undesirable situation. A gate may be used to allow a standardized master pulse, which is continuously generated and available throughout the computer, to be combined with the distorted signal in such a way that a properly timed and shaped portion of the distorted signal is selected by the master signal, and this selected portion is then passed on to the rest of the equipment. This type of signal preservation or pulse standardization is required in all machines which use pulses spaced 10 microseconds apart or less. In order to achieve reliable operation with a high-speed digital computer, this pulse standardization is inserted at frequent locations throughout the machine.

4-3. Types of Switches

Four types of switches, including vacuum-tube gates, are presently used in computers. These types vary widely in cost

and speed of operation; each is suitable for one or more particular purposes.

4-3-1. Nonautomatic Switching. Nonautomatic switching is characterized by cable connections and electrical switch settings which are manually adjustable; they are set up initially for a problem and are not altered during its run. Only the simplicity of this type of switching allows its inclusion here since it is incapable of being operated by any other part of a computer. A number of such plug connectors and hand-set switches are used for the insertion of constants in the Mark I and ENIAC. In the ENIAC there are a number of plug connectors for setting up the operations of a problem. Such preset switches and cables constitute a form of storage of the lowest order. In the ENIAC the various units are permanently wired to do the specific operations of addition, subtraction, multiplication, division, taking the square root, and looking up function values.* The particular operations which these units perform in a given problem and the order in which they do them depend to a great extent upon how units are interconnected by the various cables and how the manual switches are set. With a machine of this type, the setup time is so long that it can be operated profitably only if it is called upon to handle a large number of similar problems which can all be run on a single manual setup. This is true of the type of problem which the ENIAC was designed to handle.

4-3-2. Electromechanical Relays. The electromechanical relay exists in a wide variety of sizes and shapes. It is essentially a metallic switch, or several switches, which can be operated by an electrical signal. A relay includes an electromagnet which receives an incoming signal to operate or not operate. When the signal to operate is received, the coil is energized, and a nearby piece of magnetizable material called the *armature* is attracted and moves toward the coil. This motion, by means of a mechanical linkage, causes an assembly of contacts to open or close.

There is almost no limit to the arrangements possible with these contacts. For example, the contacts of a single relay may be made so that in one operation they will close certain circuits, open others, or transfer a circuit from one place to another. The

* This statement is true of the machine as originally used. The philosophy of programming the machine has been altered since then, as explained in Chap. 10, to make the machine more versatile.

removal of the input signal from a simple relay will deenergize the coil, and a spring will cause the contacts to return to the original position. Some relays are provided with mechanical or electrical latches which effectively maintain the contact in the operating position until a reset stimulus arrives through a different circuit. There are other relays, popularly called stepping switches, which must receive a series of alternate operate and nonoperate stimuli before returning to the original state. These, too, have limited application in computing equipment. For most computing purposes, however, the most useful type of relay is one which is capable of closing a number of electrical circuits when operated and reopening them when the operate signal is removed. Small relays have been designed for closing or for transferring 6 to 12 independent electrical circuits, and these are used effectively in the Mark I and Mark II machines.

Since the operation of this type of relay is mechanical, a considerable length of time must elapse from the instant the relay coil is energized until contact is made, even though the distance traveled may be only a few hundredths of an inch. Six- and 12-contact relays of the type mentioned are capable of operating reliably in about 15 milliseconds. Recently a number of relays capable of operating in 1 to 2 milliseconds have been announced by a relay manufacturer,[10] but these high-speed mechanical relays tend to be limited to one or two circuits. The *wire-brush relay* used by the IBM Company has an operating time of 3 milliseconds and can operate eight contacts. The speed limitation of an electromechanical relay is a fundamental one, and it does not appear that an inexpensive mechanically operating relay can be produced which will overcome this limitation.

Despite the limited speed of the mechanical relay, there are a number of significant advantages to this device. First, the relay is nearly a perfect switch; *i.e.*, when the contacts are closed, they offer practically zero resistance to the passage of current, and when they are open, they offer an extremely high resistance. Second, as compared to electronic equipment of equal cost, the relays are capable of handling far more current. Also, one relay can handle 10 or more contacts. Therefore, a relay is attractive to a circuit designer because of its economy, relatively small size, and the large number of combinations of circuits which can be switched with a single relay.

Some relays are used in electronic computers in locations where speed is of lesser importance, *e.g.*, in the reading of input information and its transfer to internal storage, or in the removal of data from internal storage to an output printer. There are numerous noncomputational requirements for relays in computers, such as in the control of power supplies, ventilators, etc.

4-3-3. Vacuum-tube Gating Circuits. An electronic gate is a circuit with a single output and two (or more) inputs so designed that an output signal is produced when, and only when, input signals are received on both (or on a particular set of) input leads. Such circuits are variously known as gates, coincidence circuits, Rossi circuits, or *logical and* circuits.

Figures 4-1a, b, and c illustrate three types of gating circuits. All perform the same operations and differ only in structural details and the polarity of signal required to operate them. All three of these circuits operate with nearly equal speed and are adequate for machines now in existence or presently contemplated.
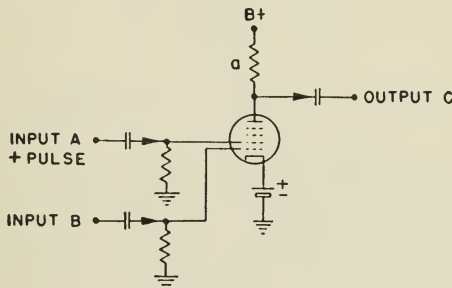
Fig. 4-1a.   Dual grid gate.

The simplest electronic gate employs a single tube which has two or more input grids. Tubes such as the 6SA7 and 6L7 are widely used in radio communications for frequency mixing; they may be used in a circuit such as shown in Fig. 4-1a to form an effective and reliable gate. Except when input signals are applied, the two input grids are negative with respect to the cathode and the tube is therefore cut off. If either grid is driven positive, the condition remains substantially unaltered. However, if both grids are simultaneously driven positive, the full electron current begins to flow, and the plate takes a large nega-

tive swing which appears as an output signal.   (The similarity between this circuit and the pentode-coupled flip-flop, Fig. 3-8, in which plate current could flow only when screen and control grids were simultaneously positive, should be noted.)   One advantage of this circuit is that only a single tube is required.   This single-tube circuit has the disadvantage that the two grids of conventional tubes do not have equal cutoff potentials; hence the two input grids are not strictly interchangeable.

Figure 4-1*b* illustrates a Rossi circuit modified so that it is suitable for computing purposes.[8]   In this circuit there are two tubes, possibly in the same envelope, the plates of which are tied
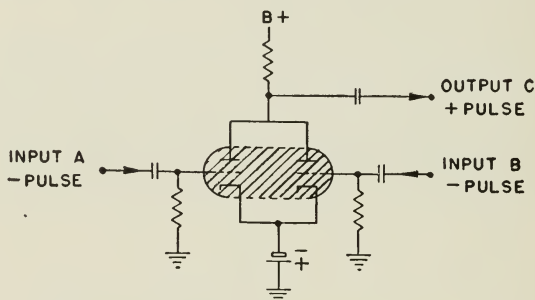
Fig. 4-1*b*.   Rossi circuit, or parallel gate.

together and fed through a single plate resistor.   Both grids are normally positive with respect to the cathodes; the tubes are therefore conducting (as shown).   Circuit parameters are so chosen that either of the tubes is capable of drawing all the plate current supplied by the batteries and plate-resistor combination. Thus, if one tube receives a negative input signal, the output voltage at the common plate connection rises very slightly. However, if both input circuits simultaneously receive negative pulses, then neither will draw current, and the output voltage will rise abruptly to the full plate battery potential.   This is the action desired.   The input stimuli are negative pulses, and the output (upon coincidence) is a positive voltage rise, in contradistinction to the multigrid-tube coincidence circuit previously described.   The necessity for proper polarization of input signals is characteristic of all electronic switching equipment.   It is obvious that a large number of switches may not be cascaded unless proper attention is paid to the polarity of the input pulses

and the resulting polarity of output pulses which are to be used elsewhere as inputs.

In computing machines the two-stage circuit illustrated is generally adequate for most purposes. However, it is possible to extend this multitube coincidence circuit further; coincidence circuits with as many as 100 input stages have been successfully built and operated. Where only two input circuits are required, it is possible to use two simple triodes, which are readily available in a single envelope. For coincidence circuits involving a large number of input stages or for those which demand the ultimate in operating speeds, the more elaborate pentode is preferred.

A circuit which apparently has not been used to date in computing machines, but one which merits consideration in the
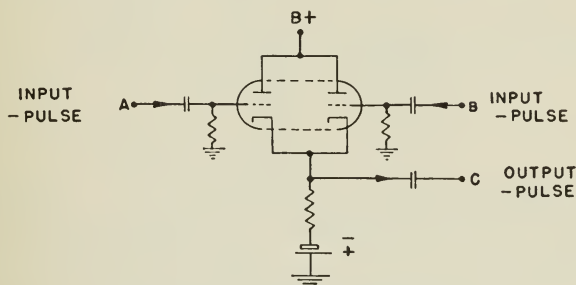
Fig. 4-1c.   Common cathode gate.

design of future equipments, is shown in Fig. 4-1c. This is a combination of the gating circuit with the *cathode follower*.[4] This circuit is similar to Fig. 4-1b except that the common load resistance is changed from the plate to the cathode circuit. In the normal condition, both tubes are conducting. The values of the common cathode resistor and the bias-battery voltage are such that, with the tubes conducting, the cathodes are positive with respect to ground. If the input stimulus in the form of a negative voltage arrives at one grid, it will serve to cut off this one tube. As in the previous case, this will have only a trivial effect on the potential of the common cathodes or output circuit. However, if negative signals arrive on both grids, then both tubes are cut off, cathode current ceases to flow, and the output-circuit potential drops. An important feature of this circuit is that the polarity of the output signal is the same as the polarity of the input signals, thus providing the possibility of cascading a

number of such circuits.   This gate has the important advantage
of any cathode-follower circuit in that it presents a considerably
lower output impedance to the next circuit.   This is an important
consideration if the next unit in the computer is at a considerable
distance and must be connected by a length of cable with a
resulting high electrostatic capacity.   This circuit and also the
one of Fig. 4-1*b* can have identical tubes at each input and hence
can be made with input leads which are interchangeable.
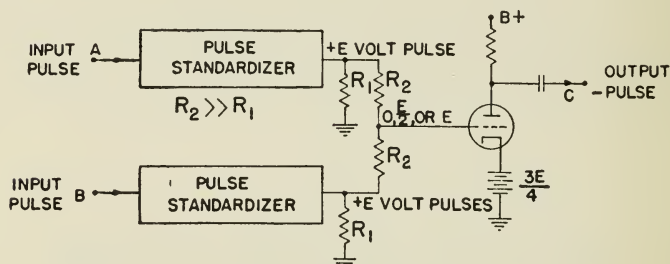
Fig. 4-2*a*.   Resistance grid gate.

4-3-4. Resistor Matrices and Diode Matrices.   Attempts have
frequently been made to simplify the electronic gating circuit by
using pulses of equal size in a matrix of resistors or nonlinear
elements such as crystal rectifiers.   An elementary example of a
resistor gate is shown in Fig. 4-2*a*.   In this circuit the pulses
must first be rigorously standardized as to size.   In the example
shown, the presence of a pulse is indicated by a potential of 20
volts.   The common junction point which feeds the grid circuit
of the output tube may thus assume a potential of 0, +10, or +20
volts, depending on whether zero, one, or two pulses arrive on the
two input circuits.   If the cathode is maintained at a potential of
+15 volts by the batteries shown in the elementary configuration,
the relative grid-cathode potential will be negative for the cases
of zero or one input signal and positive only when two input
signals are simultaneously present.   In the latter case, plate
current is permitted to flow and produces an output signal in the
form of a negative pulse.   By properly rearranging the circuit
parameters, it is possible to cause this circuit to operate in the
opposite fashion so that the coincidence of negative input pulses
will produce an output pulse of positive polarity.   This circuit
works dependably with two input points and reasonably well

for slightly larger numbers. With a number of input circuits n, the threshold between operation and nonoperation is only the difference between $(n - 1)/n$ and 1. This precludes its use for values of n much larger than 5. Precisely the same limitation was found to exist in the r-triode counter, modulo r, described in the previous chapter. Furthermore, the circuit described fails to afford any significant economy over the ones illustrated in Figs. 4-1a, b, and c, because of the number of accessory tubes required to standardize the amplitude of the input signals.

A considerable improvement in matrices can be obtained by the use of miniature crystal rectifiers,[3,7,9] which were brought to a
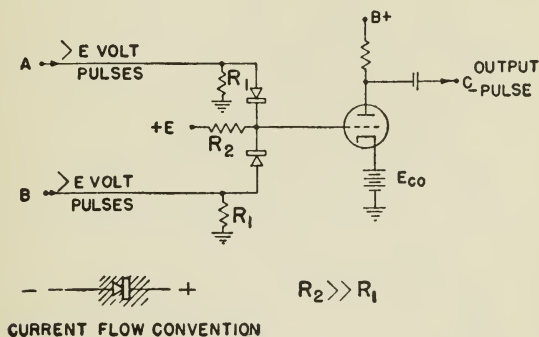
FIG. 4-2b. Diode-coupled grid gate.

state of perfection for radar application during the recent war. Each of these diodes or crystal rectifiers contains a small piece of some semiconductor, such as germanium or silicon, with which contact is made by a small sharpened point, or *cat whisker*. Such a device presents a unilateral impedance. It is capable of carrying a small current in one direction while resisting up to a limit the passage of current in the reverse direction. A circuit similar to Fig. 4-2a is shown in Fig. 4-2b in which the pulse standardizers and resistor pair are replaced by two of these crystal rectifiers. A negative potential on either of the two crystals will hold the grid of the output tube below cutoff. However, if both crystals are simultaneously driven positive, they effectively disconnect their input circuits from the common connection point to the grid, thereby causing the potential at this point to rise, allowing the output circuit to conduct. Here, again, the circuit parameters may be rearranged so that negative

input signals are used and a positive output signal is obtained. The main improvement afforded by the crystals is the snap action of the potential change at the common tie point. This common point assumes one potential for both zero and one input signal and changes to its two-input potential only when two input signals are received.

It is possible to expand this circuit to a large number of inputs since the snap action is obtained only when all inputs are driven
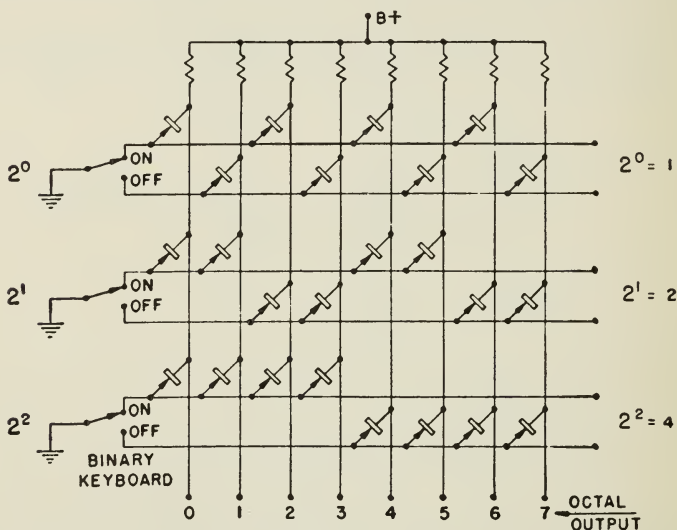
FIG. 4-3a. Binary to octal converter. For any one of the eight possible combinations of three on-off switch positions one, and only one, of the eight output terminals will have a positive voltage; each of the other seven will be grounded through one or more of the diodes.

in the proper direction. The main disadvantage of such crystal circuits is the relatively small amount of power which they are capable of handling. In practical application, an amplifier tube will usually be required to restore the power level after each diode array. However, for some applications with a number of diode switches at the same location, it is possible to defer this amplification until after several stages of crystal diodes.

An example of such a situation is shown in Fig. 4-3a. This diagram represents a hypothetical converter which could be used for translating a number from binary to octal notation. The number 7 is shown on the binary keyboard on the left-hand side

of Fig. 4-3a as $2^2 + 2^1 + 2^0$. The row of terminals labeled 1 to 7 at the bottom of the diagram represents those digits in the octal system. It can be seen that every one of these terminals, except 7, is grounded. The seventh terminal is the only one on which the voltage from the positive battery supply will appear.

Figure 4-3b illustrates a translator which operates in the opposite direction. The input switch is set on number 7.
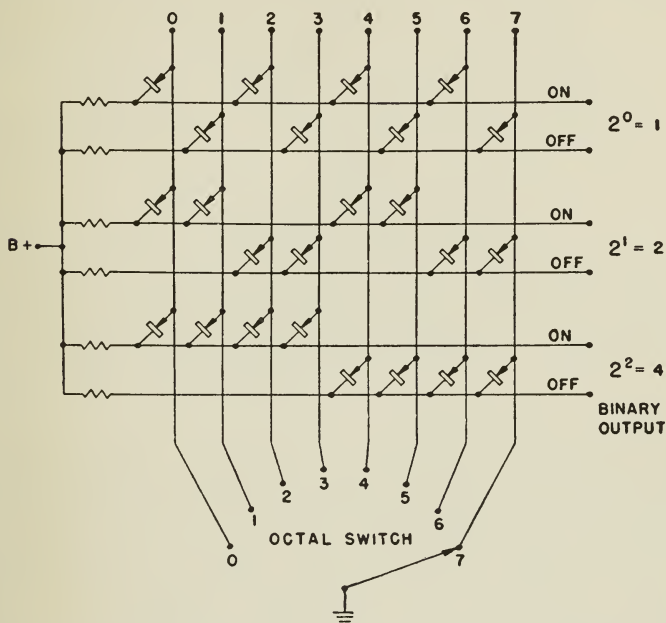
FIG. 4-3b. Octal to binary converter. For each one of the eight possible switch positions a different combination of output leads will be energized.

It can be seen that all the *on* leads for the binary output are energized. It can also be seen that if the switch were on any of the other octal numbers, 0 through 6, one or more of the *on* leads corresponding to some power of 2 would be grounded through the octal input switch.

There is some inherent limitation to the rate at which pulses can be passed through a diode matrix; this is due primarily to the time constants of the diodes themselves. For computing applications, however, this limitation need not be considered a material design factor.

4-4. Inverse Gates, or Buffers

The gating circuits discussed in a previous paragraph are frequently described as being the electronic expression of a *logical and; i.e.*, the output terminal C produces a signal only when the two input terminals A *and* B receive signals simultaneously. In computer design there is frequently a need for a circuit which functions as a *logical or*, to supply an output if either A *or* B is signaled. These *logical or* circuits are referred to as anti-Rossi circuits, isolating circuits, or simply buffers. The term *logical or* is used in the sense that A *or* B includes A *and* B.

The difference between such a buffer and the gates previously described is simply the interchange of the electronic states representing the normal, or stand-by, condition and the operate condition. For example, the single-tube gate of Fig. 4-1a can be changed to a buffer merely by biasing both grids positively and furnishing input signals as negative pulses (the reverse of the gate conditions). Then, if either input is signaled, the plate current will be terminated and a positive output signal will result. This same transformation from gate to buffer can be effected on any of the other circuits, Figs. 4-1b and c and 4-2b. The only engineering difference between a gate and a buffer is in the amount of energy dissipated during the stand-by periods.

At first glance it may appear that the same results can be achieved by connecting both A and B to C so that either A or B can send a signal through to the output C. Unfortunately such an arrangement also connects A and B together in a "sneak circuit" having definitely undesirable results. For example, pulses originating at A and destined for C via a simple Y connection also flow into B, where they are either dissipated entirely or injuriously affect the apparatus connected to input B. The buffer connection eliminates this possibility.

Buffers are commonly used in the input circuits of an electronic register, *e.g.*, in the decade rings of a device like the ENIAC. Here it is necessary to be able to step the ring from any of three inputs, a normal input, a carry-over from the decade of the next lower order, or a read-out input consisting of a series of 10 cycling pulses used to cycle the ring once for read-out. A three-input *logical or* circuit is used to accomplish this, although it is difficult to point to precisely the three tubes which constitute the buffer since this function and certain gating operations are

combined in some of the tubes.   The combination of functions is common when pulses traverse gates and buffers alternately, because a buffer may be inherent in the circuit so that additional tubes are not required for circuit isolation.   An example of such
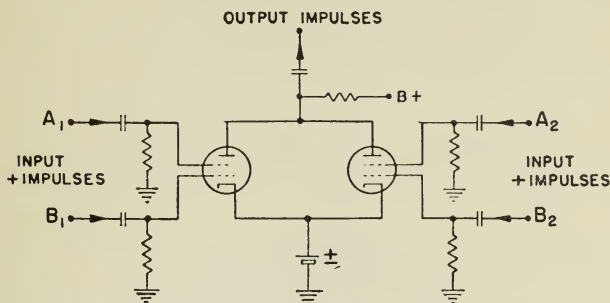
Fig. 4-4.   A combination of *and* and *or* circuits.

a circuit is Fig. 4-4, in which the output circuit will respond if inputs A_1 and B_1 are energized or if A_2 and B_2 are energized (or, also, if both combinations are energized, "or" being used to include "and" as stated above).

4-5. Assemblies of Gates, Buffers, and EPDCC's

Combinations of gates and buffers with the EPDCC's described in Chap. 3 make possible a complete device capable of carrying out an elementary arithmetic operation (addition or subtraction). In the descriptions which follow, three examples are given in order of increasing complexity.   Each involves high-speed electronic techniques which permit a complete addition in less than a millisecond, and each performs decimal arithmetic.   These simple examples illustrate principles which are used in large-scale computing machines.

4-5-1. A Single-digit Adder.   In the upper part of Figure 4-5a is a gate–flip-flop combination.   The flip-flop opens the gate when a stimulus is applied on the *set* lead into the flip-flop.   A stimulus on the *reset* lead closes the gate.   The master oscillator generates a continuous stream of pulses at some convenient frequency such as 100 kilocycles.   The decimal pulsing counter is a modulo 10 ring (of the 20-triode type, for example).

A *start* signal causes the flip-flop to open the gate.   The stream of pulses from the master oscillator is permitted by the open

gate to reach the pulsing counter and also to follow the lines shown in the diagram to other parts of the circuit. The decimal pulsing counter counts exactly 10 pulses, going through one complete cycle in doing so. When it reaches zero again, the tenth stage, it generates an impulse on the *reset* input to the flip-flop. This closes the gate, suppressing the passage of the 11th, 12th, and succeeding pulses from the master oscillator. In
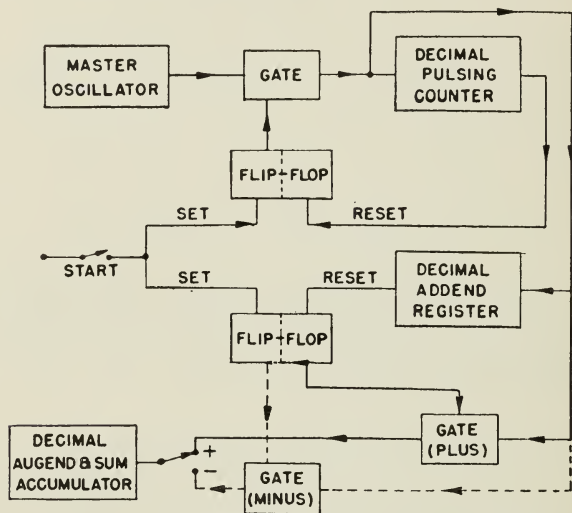
FIG. 4-5a.   An idealized single-digit decimal adder.

short, the upper part of Fig. 4-5a illustrates a device for introducing exactly 10 pulses on a pulse bus each time a *start* signal is received.

Let us consider the use of this circuit in solving a simple single-digit decimal addition problem involving no carry-over. Suppose it is required to add the single-digit decimal numbers 3 and 5. First, the addend, 3, is stored in the decimal addend register, which is a ring counter of the type described in Chap. 3 which can be set to represent a particular digit. The number 3 is stored in this counter simply by setting the counter on its third position.

The accumulator shown in the lower left-hand corner of Fig. 4-5a is a similar ring counter, in which the augend 5 is stored.

The *start* signal, as described above, permits 10 pulses from the master oscillator to be introduced via the bus to the addend

register and to the gate labeled *gate (plus)*. This gate is closed, however, at the beginning of the operation. The pulses therefore do not enter.

The addend register, as indicated above, was set on the third step of its cycle prior to the introduction of the train of 10 pulses. The first seven pulses in the train advance it through the rest of its cycle to zero. When the register transfers from the ninth step to zero, it opens the gate labeled *gate (plus)*. With this gate open, the last three pulses of the train of 10 reach the decimal accumulator, where the number 5 has already been stored. These three pulses advance the accumulator from 5 to 8, which is the required sum.

At the close of the operation, therefore, the sum is read from the accumulator where the augend had originally been set. 3, the number set into the addend register at the start of the operation, appears there again at the close of the operation, because 10 pulses have produced a complete cycle of this register, returning it to 3.

For subtraction, another gate–flip-flop combination is provided, as shown by the dotted lines of Fig. 4-5a. Suppose it is required to subtract 3 from 5. 3, which is the subtrahend in this example, is set into the same register in which it appeared in the preceding example as the addend. 5, the minuend, is set into the register used for the augend in the addition example. The *start* pulse opens the subtract gate labeled *gate (minus)*. It also causes the gate leading from the master oscillator to the counter to be opened. The first seven pulses reach the accumulator via this gate, cycling it from 5 to 2. This gate is closed after the seventh pulse has passed, just as the *gate (plus)* was opened in the preceding example. The required difference, 2, is thus obtained in the accumulator by a process which amounts to the addition to 5 of 7, the 10's complement of 3. (The carry-over is disregarded.)

In practice it is very difficult to operate a counter in the fashion described. Close timing is required when the register counter switches from 9 to 0 because the same signal which effects this advance in the register must also be routed through the proper add or subtract gate. For this reason the elementary configuration shown is not actually used. Instead, a separate source of pulses is used to drive the add or subtract gate, and this stream is delayed slightly in relation to the series of pulses which

drives the register.   These two streams of pulses are accurately synchronized, and the delay is held constant within close limits.

If this system is expanded to handle more than one decade, it is necessary to use complements with respect to 9 instead of 10 in all but the last decade in order to carry out subtraction.   However, the register counter requires 10 pulses in order to effect a complete cycle.   For this reason the register in a practical
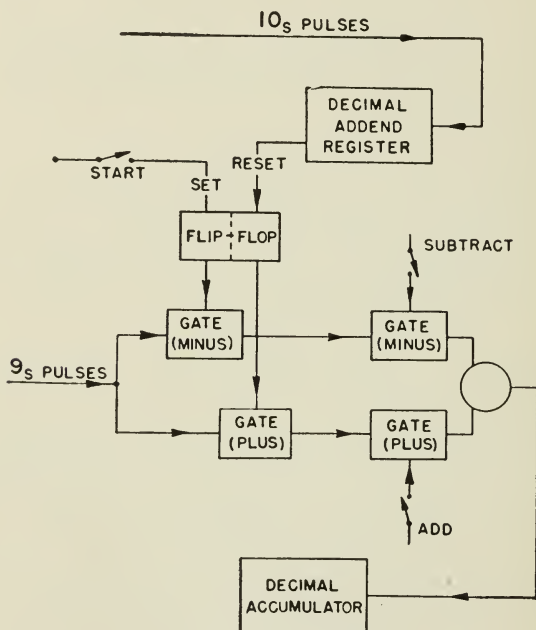
Fig. 4-5*b*.   A practical decimal adder element.

machine is driven by a series of 10 pulses while the gates leading into the accumulator are driven by series of 9 pulses which are synchronized with the spaces between the 10's pulses.   No matter how complex the computer, however, only one standard pulse generator is needed to drive everything else.   With a source of accurate 10's pulses and synchronized interspersed 9's pulses in the computer, an elementary single-digit adder resembles the device shown in Fig. 4-5*b*.   In this figure the 10's pulses are routed to the register as before and cause this device to cycle once around.   The 9's pulses meanwhile are reaching the two gates associated with the flip-flop.   One of these gates is closed

from the time the *start* signal is received until the register transfers from 9 to 0, whereupon it opens. The other gate, which is used for subtraction, does the opposite. Two additional gates are shown, one each for addition and subtraction; one of these two is kept open throughout the operation and causes the operation to be either addition or subtraction. The outputs of these two gates are buffered together and brought to the input of the accumulator. The accumulator thus receives either the contents of the register (add) or the 9's complement of the register contents (subtract). This arrangement of pairs completely avoids the difficulty of having a *set* and a *reset* pulse reaching the flip-flop at almost the same time when the register flips from 9 to 0. This is essentially the scheme of operation of the register-accumulator connections for a single digit in the ENIAC, which will be described more completely in the third example. The example just given shows the use of gates for data switching (*i.e.*, switching streams to indicate add or subtract) and for data conversion (the pulsing of the register by the 10's pulses to convert the register contents to pulses).

4-5-2. An Elementary Parallel Adder. The second example illustrates a method for handling multiple-digit numbers, which involves carry-overs. This method is not the most efficient but will serve for the purpose of illustration. The combination of units shown in Fig. 4-6 can add two three-digit decimal numbers and perform the required carries. It can also take the difference by means of 9's complement addition. In this example let us assume that 379 is stored in the register and 468 is stored in the accumulator. They can be stored by the use of the settable counter or by first setting the counters to 000 and pulsing each digit ring the appropriate number of times. For addition, three distinct operations are involved. First, the units ring of the register is cycled with the 10's pulses as described in the previous example, and the output is used to control the add gate routing the selected pulses from the stream of 9's pulses into the units ring of the accumulator, thereby setting this ring to its new value (7). Meanwhile, another flip-flop connected to the units ring of an accumulator takes cognizance of the fact that this ring has passed from 9 to 0 in the process of bringing up the 7, and it serves to open a special carry gate. After the stream of 10 pulses has been completed, a special 11th pulse is sent to the

carry gate. Since this gate is open because of the carry-over, the 11th pulse can enter the 10's ring of the accumulator. The 11th pulse is supplied by the same master pulse-generating equipment which supplies the 10's pulses and the 9's pulses, and it is properly phased with them. So far the digit 9 has been added

Fig. 4-6. A multiple-digit decimal adder.

to the numbers 468 in the accumulator and has produced the new number 477, which is correct, including the handling of the carry-over. The entire process is repeated except that the next stream of 10's pulses is routed into the 10's ring of the register, where they serve to rotate the 10's ring of the accumulator by an appropriate amount, and the 11th pulse again causes a carry-over, this time into the 100's ring of the accumulator. One more repetition of the process on the 100's ring of the register completes the entire operation, and the number 847 stands in the

accumulator. Switching of the input streams of 10's and 9's pulses would be accomplished by additional gates.

On a desk machine, subtraction is carried out by running the machine wheels and carry backward. For a machine involving electronic counter rings this is not convenient, and some scheme which has the same effect must be sought. Except for carry, it would suffice to turn each ring forward a number of steps equal to 10 minus the digit appearing in the subtrahend. When carry is taken into account, it turns out (as every desk-calculator operator knows) that the lowest order digit which differs from 0 is complemented with respect to 10, the higher order digits are complemented with respect to 9, and the zeros at the low end of the number, if any, are left unchanged; the resulting number is added to the minuend, and the final carry to the left is ignored. (The term *complemented with respect to* 10 means that the digit is replaced by the digit to which it must be added to give the number 10.) Thus to subtract 3,790 from 4,681, the complement of 3,790 is written 6,210, and this is added to 4,681, ignoring the final carry, to give 0,891. Since the position of the digit to be complemented with respect to 10 depends on the number being complemented, this method is not easily applicable to automatic machinery. An equivalent scheme results from noticing that the complement obtained in this fashion exceeds by 1 the number which would be obtained by taking the 9's complements of all digits. Subtraction on many machines including the ENIAC is carried out in this fashion: the 9's complement of the subtrahend is added to the minuend, ignoring the final carry, and 1 is added to the result. Thus, for the example above, 6,209 would be added to 4,681, ignoring the final carry, to obtain 0,890, and 1 is added to the result to obtain 0,891. The correcting pulse which is used to add 1 may initiate some carry-overs. Therefore, after it occurs, it is necessary to operate each of the special carry gates once, beginning with the one of lowest order. Alternatively, the digit of the lowest order might be corrected immediately after adding the complement of the lowest order digit in the register. No net gain would result from this scheme since the same amount of time would have to be consumed when pulsing the digits of high order.

Some improvement in the scheme of addition and subtraction is realized when carry-over is accomplished while cycling the

counters, thus eliminating the need for an 11th pulse to follow later. This is analogous to the scheme used on ordinary mechanical counters, in which the input directly turns over the number of wheels required. The elementary parallel adder is a comparatively slow device since it is capable of handling only one digit at a time. Thus, in the example cited, an addition takes 30 pulse times. In general, if the machine is handling numbers of m digits on the radix r, the total time for addition in this type of device is mr. This becomes an appreciable factor if m is 10 or more. The more advanced type of parallel adder described in the following example is faster.

4-5-3. An Improved Parallel Decimal Adder. The method of addition employed in the ENIAC requires the simultaneous use of two complete accumulators and furnishes an example of an elaborate usage of gates, buffers, and EPDCC's. Altogether, each accumulator decade requires the use of 45 tubes (actually contained in 32 envelopes) to carry out the functions of addition, subtraction, carry-over, accumulative carry, and clearing, plus a number of necessary electronic operations of no direct arithmetic importance. Thus, to carry out a simple addition (or subtraction) of two 10-digit decimal numbers, 900 tubes are required for digit accumulators and another 14 for plus-minus sign indication. This does not include the common pulse-generating equipment and the program controls which bring these accumulators into play.

Figure 4-7, a block diagram of a single ENIAC accumulator decade, illustrates the functions of the various tubes. One decade ring is used to store each digit of the addend and one to store each digit of the augend and later the sum. The former unit is in the transmit condition and the latter in the receive condition. In addition to the 20 tubes of the decade ring itself, there are the following components listed by identification number:

1. Buffer for cycling the ring when the accumulator is used to transmit a number (add or subtract).
2. Gate for receiving pulses when accumulator is receiving.
3, 4, 5, 6. Pulse standardizer and driver tubes to drive the ring on any operation.
7. Gate for control of carry-initiating signal.

8. Gate for synchronizing the timing of the carry and for resetting the carry flip-flops.
9, 10, 11, 12. Flip-flop and associated gates for remembering when a carry is required and for implementing the carry
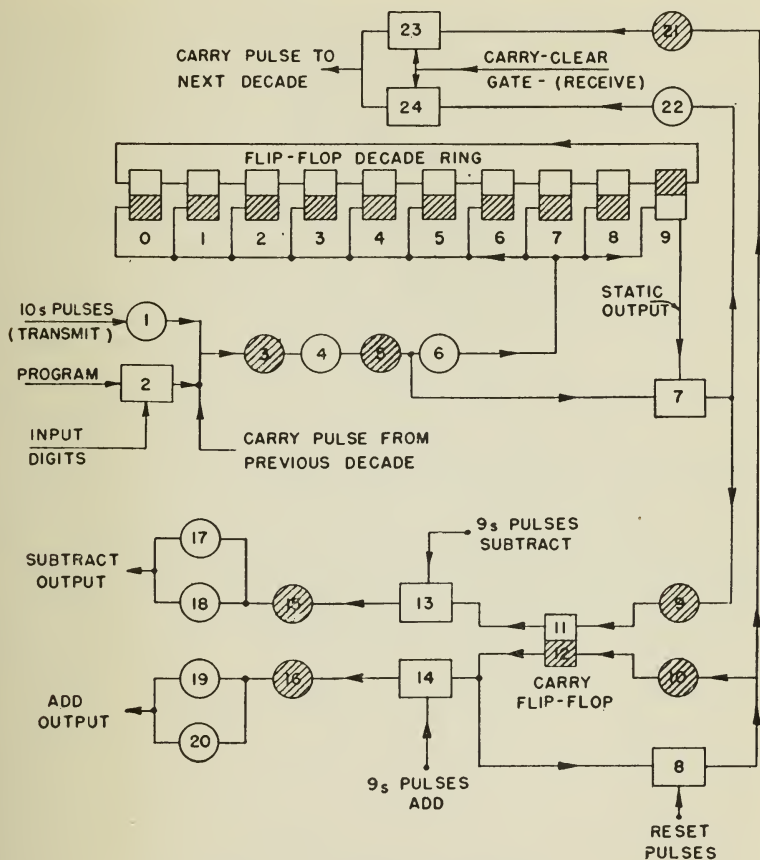

FIG. 4-7.   ENIAC accumulator.

when signaled by tube 8.   Also used on transmit to produce the add pulses and the 9's complement pulses from the 9's pulses.
13, 14. Gates used to choose sum or difference for output.
15, 16. Amplifiers.
17, 18, 19, 20. High-power output buffers to permit the con-

nection of many accumulator transmitters to a single number-transmission bus.

21, 23. Amplifier and gate for the control of a single carry.

22, 24. Same for accumulative carry.

The improvement afforded by this arrangement is the increase in speed realized by having all decades of an accumulator pulsed simultaneously rather than in the sequential manner indicated in the previous example. Theoretically, this addition should require $m + r$ cycles (as compared to mr machine cycles for the example previously discussed) in order to handle the possibility of as many carries of an m-digit number of radix r. The ENIAC actually requires $2r$ machine cycles for any value of m up to the maximum possible value 20. This is caused by the provision of two complete independent carry circuits on each decade, one for ordinary single carry and one for the accumulative carry (in which one carry may originate another, such as the addition of $00 \cdot \cdot \cdot 001$ to $99 \cdot \cdot \cdot 999$). In the block diagram of Fig. 4-7, a single carry would flow through gate 8 and tubes 21 and 23 to the decade of next higher order. If this should create another carry, it will flow through 3, 4, 5, 6 of the next decade to the input of the ring (flipping it from 9 to 0), and gate 7 will cause this same pulse to pass without further delay directly to 22 and 24 and thence to the next higher decade. The timing of the 10's and 9's pulses is as described in the earlier example and constitutes the first half of an addition time. In the second half the carry-over tubes 23 and 24 are opened, and gate 8 is pulsed by a reset pulse. Sufficient time must be allowed to permit an accumulative carry of maximum length and to terminate this operation and initiate the next one. The total time for both halves is 200 microseconds. The timing of the ENIAC as a whole is controlled by 20 pulses generated by a cycling unit in which a master oscillator, a 20-stage ring, and numerous gates and signal shapers all combine to produce a set of standardized pulses which are made available throughout the entire machine.

REFERENCES

1. Aiken, H. H., and G. M. Hopper, The Automatic Sequence Controlled Calculator—I, *Electrical Engineering*, Vol. 65, Nos. 8 and 9, pp. 384–391 (August–September, 1946); The Automatic Sequence Controlled Calculator—II, *ibid.*, No. 10, pp. 449–454 (October, 1946); The Automatic

Sequence Controlled Calculator—III, *ibid.*, No. 11, pp. 522–528 (November, 1946).

2. Burks, A. W., Electronic Computing Circuits of the ENIAC, *Proceedings of the Institute of Radio Engineers*, Vol. 35, pp. 756–767 (1947).

3. Cornelius, E. C., Germanium Crystal Diodes, *Electronics*, Vol. 19, No. 2, pp. 118–123 (1946).

4. Dicke, R. H., A High Speed Coincidence Circuit, *Review of Scientific Instruments*, Vol. 18, pp. 907–914 (1947).

5. Harvard University, Staff of the Computation Laboratory, *A Manual of Operation for the Automatic Sequence Controlled Calculator, Annals of the Computation Laboratory of Harvard University*, Vol. I (Harvard University Press, Cambridge, Mass., 1946).

6. Kurshan, Jerome, Cathode-Follower Gate Circuit, *Review of Scientific Instruments*, Vol. 18, pp. 647–649 (1947).

7. North, H. Q., Properties of Welded Contact Germanium Rectifiers, *Journal of Applied Physics*, Vol. 17, pp. 912–923 (1946).

8. Rossi, Bruno, Method of Registering Multiple Simultaneous Impulses of Several Geiger's Counters, *Nature*, Vol. 125, No. 3156, p. 636 (1930).

9. Stephens, W. E., Crystal Rectifiers, *Electronics*, Vol. 19, No. 7, pp. 112–119 (1946).

10. The Stevens Arnold Co., Inc., South Boston, Mass., *Millisec Relay* (1948). Pamphlet.

11. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *Report on the ENIAC, Technical Report I* (University of Pennsylvania, Philadelphia, Pa., June 1, 1946). 2 vols.

12. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *Functional Description of the EDVAC, Research Division Report* 50-9, a report of development work under contract W-36-034-ORD-7593 with the Ordnance Department, Department of the Army (University of Pennsylvania, Philadelphia, Nov. 1, 1949). 2 vols.

CHAPTER 5

A FUNCTIONAL APPROACH TO MACHINE DESIGN

5-1. Introduction

In the design of a general-purpose digital computing machine, it is necessary to organize elementary potential digital computing components, switches, and gates into a device which can receive input information and operate upon it in a specified manner according to chosen systems of arithmetic and numerical analysis. The operations performed are initiated by information derived either from the data-input system or from the results of the arithmetic computation within the machine. This information is used to select specific commands from a group of operational commands which control the process of computation. The number of operational commands available and the extent of facilities for storage should be large enough that the arithmetic computations within the machine can be performed expeditiously. At the same time a minimum of circuit complexity is desired.

An example of a typical arithmetic problem is presented in the following paragraphs. Several approaches to its solution, each demonstrating a method suitable for use with a machine of different capabilities, are set forth. The exposition is intended to show how a machine which can perform only a few simple arithmetic operations, in accordance with correspondingly simple instructions, can be made to solve a typical problem by following a set of explicit instructions. It is intended to show, also, that the arithmetic techniques to be followed in solving a particular problem efficiently on any given machine must be selected with regard for the capabilities of the machine. Conversely, when a new machine is to be designed, a consideration of its intended uses, *i.e.*, of the kinds of problems it will be asked to solve, is of obvious importance in deciding what arithmetic operations it should be designed to perform.

TABLE 5-1

k	x	y	x^2	y^2	x^2+y^2	$\dfrac{x^2+y^2}{u_{k-1}}$	$2u$	u	$u \cdot u_{k-1}^{-1}$	$2 - u \cdot u_{k-1}^{-1}$	u^{-1}
1	1.0000	0.0000						1.0000			
2	1.0054	0.0029						1.0054			
3	1.0100	0.0065						1.0100			
4	1.0138	0.0079						1.0138			
5	1.0201	0.0134						1.0202			
6	1.0326	0.0201						1.0328			
7	1.0459	0.0300						1.0464			
8	1.0238	1.0634						1.0259			
9	1.0194	0.0802						1.0225			
10	1.0127	0.0889						1.0166			

5-2. An Example

Suppose x and y denote two measurable physical quantities, such as the orthogonal components of a velocity which varies in magnitude and direction but remains at all times in the plane of x and y.   A large number of measurements of corresponding values of x and y have been made.   The problem is to compute the scalar magnitude of the velocity vector for each pair of x and y values, *i.e.*, to compute the square root of the sum of the squares of corresponding x's and y's.   Table 5-1 contains 10 such pairs of values of the measured quantities, in columns headed x and y.   This table contains also the value of u, where

$$u = \sqrt{x^2 + y^2}$$

for each pair of x and y values.   The meaning of each of the blank columns in the table will be made clear shortly.

Suppose the values of u are to be obtained by means of calculations performed on a desk machine which can perform addition, subtraction, multiplication, and division directly but has no automatic device for extracting the square root of a number.   To obtain the square root of $x^2 + y^2$, therefore, some method of approximation will be required.   (Chapter 7 discusses such methods, and Chap. 8 contains a summary of the properties of various commercially available desk calculators; no detailed discussion of either of these topics is of primary importance here.)   Let us also suppose two additional factors: (1) the value of y shown in the table for the first measurement is known to be exactly zero as a consequence of the way the experiment was run, and (2) the measurements were made frequently enough so that succeeding values of u are known to lie close together.   Because succeeding values do lie close together, each computed value of u may be used as the assumed value in the next step of the computation. u_k, the kth value of u, may be found by using the formula

$$u_k \approx \frac{[(x^2 + y^2)/u_{k-1}] + u_{k-1}}{2}$$

The detailed instructions for carrying out the computation expressed by this formula may be given to a machine operator in the following form:

1. For measurement No. 1 copy the number x into the column headed u.

2. For each of the other measurements, in order, multiply the number x by itself; add to the product, without removing it from the machine's register, the product of the number y multiplied by itself. This leaves $x^2 + y^2$ in the product register.

3. Clear the quotient register, leaving the accumulated product in the machine, and divide this product by the entry in the u column computed for the next earlier measurement, or u_{k-1}.

4. Transfer the quotient to the product register and add to it the entry from the u column computed for the next earlier measurement, u_{k-1}.

5. Clear the quotient register and divide the number in the product register by 2; write the result to four decimal places in the column headed u.

5-2-1. Variations in the Method. In following the set of instructions given above the operator had to fill in only the answer column, headed u, in Table 5-1. The use of the other columns is illustrated below in an example of the solution of the same basic problem on a different calculator. Variations in the arithmetic techniques and, consequently, in the program of instructions are required because this calculator is not capable of performing all the operations contained in the foregoing program. Analogous variations are required when the same basic problem is handled on two different large-scale machines.

At the end of step 2 in the calculation described above, the quantity $x^2 + y^2$ had been built up in the product register. The process of building it up depended upon the ability of the machine to multiply and to accumulate products. It might be more convenient to build a machine which would not accumulate products. In this case, if the machine could multiply, add, and divide, but all in *different* components, it might be necessary to write down the quantities x^2, y^2, and $x^2 + y^2$ as they are computed so that they can be reinserted in the proper places in the machine at the desired time. For such a machine, the procedure would be modified in the manner stated before the directions were turned over to the operator. Furthermore, the x^2 column of Table 5-1 would be filled during the computation.

A more serious modification is required if the machine will not carry out some of the arithmetical operations used. Suppose, for example, that the machine will not divide easily. There are two divisions in the process, and they would have to be replaced. The second is division by 2, and it is simply replaced by multiplication by 0.5000. The first is division by the next earlier value of u; this is more difficult. Methods described in Chap. 7 permit an easy running computation of a slowly changing value for u^{-1}, using only the operations of multiplication and subtraction. It is convenient to use the fact that the first x is exactly 1 and the first y is exactly 0. Then the set of instructions or *commands* for a machine incapable of division would have to be changed from the set above to a set similar to these:

1′. For measurement No. 1, write the number 1 in the column headed u and the column headed u^{-1}.

2′. For each of the other measurements, in order, multiply the number x by itself; add to the product, without removing it from the machine's register, the product of the number y multiplied by itself. This leaves $x^2 + y^2$ in the product register.

3′. Transfer the number in the product register to the keyboard, clear the product register, and multiply the number by the entry u^{-1} for the next earlier measurement, u_{k-1}^{-1}.

4′. Add the entry u for the next earlier measurement, u_{k-1}, to the number in the product register.

5′. Transfer the number in the product register to the keyboard, clear the product register, and multiply by 0.5. Write the result to four decimal places in the column headed u.

6′. Clear the machine and insert the number 2.00000000 in the product register and subtract from it the product of u (just calculated) and u_{k-1}^{-1}.

7′. Transfer the number in the product register to a keyboard (the multiplier keyboard if the machine is equipped for automatic multiplication), clear the product register, and compute the product of the number just transferred and the number u_{k-1}^{-1}. Enter the product to four decimal places under u^{-1}.

This method of performing division was first brought to our

attention through its use in the Mark II Calculator.[3] It may
be stated thus:

$$u_k^{-1} \approx (2 - u \cdot u_{k-1}^{-1})u_{k-1}^{-1}$$

but it is not of primary importance in this discussion.

The commands to the operator in this set still make use of
some of the properties which are built into the machine and which
might not be convenient in a machine built on the fast operating
principles being described here.   It might be true, for example,
that a convenient machine would not immediately operate on
the result of any other operation; in this case, the properties of
the product register used repeatedly in the list of commands
could not be used to permit double operations like 6′, and the
single operation 4′ would become more complex.   Although no
machine yet proposed has been reduced to the point that the
result of an elementary arithmetic operation is immediately
removed from the computing component, this is certainly a
possibility which must not be ignored.   If this were the case for
a machine with a human operator and if the machine were capa-
ble of addition, multiplication, and subtraction but not of
division, then a set of fairly stereotyped commands would be
written, and all the spaces in Table 5-1 except a few for measure-
ment No. 1 would be filled.   The following commands or their
equivalent would be required:

  1″. Enter the number 1 in the u column for measurement
    No. 1.
  2″. Enter the number 1 in the u^{-1} column for measurement
    No. 1.

For each other measurement in succession proceed as follows:

  3″. For entry in x^2 column, multiply; multiplicand is entry
    in x column; multiplier is entry in x column.
  4″. For entry in y^2 column, multiply; multiplicand is entry
    in y column; multiplier is entry in y column.
  5″. For entry in $x^2 + y^2$ column, add; augend is entry in x^2
    column; addend is entry in y^2 column.
  6″. For entry in $(x^2 + y^2)/u_{k-1}$ column, multiply; multipli-
    cand is entry in $x^2 + y^2$ column; multiplier is entry in
    u^{-1} column for next earlier measurement, u_{k-1}^{-1}.

7''. For entry in $2u$ column, add; augend is entry in $(x^2 + y^2)/u_{k-1}$ column; addend is entry in u column for next earlier measurement, u_{k-1}.

8''. For entry in u column, multiply; multiplicand is entry in $2u$ column; multiplier is the number 0.5000.

9''. For entry in $u \cdot u_{k-1}^{-1}$ column, multiply; multiplicand is entry in u column; multiplier is entry in u^{-1} column for next earlier measurement, u_{k-1}^{-1}.

10''. For entry in $2 - u \cdot u_{k-1}^{-1}$ column, subtract; minuend is the number 2.00000000; subtrahend is entry in $u \cdot u_{k-1}^{-1}$ column.

11''. For entry in u^{-1} column, multiply; multiplicand is entry in $2 - u \cdot u_{k-1}^{-1}$ column; multiplier is entry in u^{-1} column for the next earlier measurement, u_{k-1}^{-1}.

The stereotyped nature of the commands renders an advantage in clarity and in ease of writing a program. It should be noted that each command except for the unnumbered command ''For each other measurement in succession:'' consists of a prescribed arithmetical operation, addition, subtraction, or multiplication, in which the two entering numbers are specified by their position in Table 5-1 and in which the disposition of the resulting number into a position in Table 5-1 is specified. Burks, Goldstine, and von Neumann[2] have pointed out the advantages of establishing a simple stylized form of this general type no matter what the characteristics of the machine. For the convenient operation of a general-purpose machine, they point out, it is essential that some steps be taken to translate the nonconforming command quoted above to the same stereotype form. This translation of description of all possible operations to prescribed forms has been called the *logic* of the machine by these authors, and the term is now in general use.

5-3. Machine Requirements

The earlier part of this chapter describes a simple calculation in which various postulated types of machinery operated by a human operator are used. The two preceding chapters describe techniques which seem to promise arithmetic calculation with the elementary operations requiring only microseconds for completion. For a machine using these new components, a human

operator intervening at every command would encumber the machine with a slow-moving element whose reaction time is intolerably long. For example, during a second of time the machine is capable of carrying out 1,000 multiplications, but the operator is capable of issuing at most a single command. It is essential that a sequencing mechanism similar to, but faster than, the human operator be built into the machine if its potential speed is to be utilized.

An automatic sequence-controlled calculator is a computing machine into which such a mechanism is built. It follows a pre-scribed set of explicit instructions automatically.

Such a machine is usually much faster and more accurate than a twelve-year-old arithmetic student, but less versatile. Its lack of versatility can be offset by its speed because combinations of a few simple operations can be made to effect results equivalent to those which would be produced by the twelve-year-old in response to less explicit instructions. These combinations of simple operations are therefore components of computation.

A sheet of paper, Table 5-1, was also included as a component used in the calculation described above. This sheet of paper is so inexpensive that it is overlooked frequently as a component, but those experienced in calculation know well that most of the time required in calculation and essentially all the errors of the calculation are connected with the reading of data from and the entering of data on this sheet of paper. The potentialities of speed of computation with the new components and the resulting extent of calculations which can be made in a reasonably short time demand a storage medium which is faster and more accurate than pieces of paper marked and read by a human operator. It is also obvious that when complicated calculations are per-formed in small increments, a voluminous storage outside the arithmetic units is required to store the partial results until they are pieced together into a final solution. The voluminous storage to which quick access must always be available is at present a chief consideration in connection with computing instruments. Many machines have several different storage units, some of which are more quickly accessible than others.

Finally, implicitly, the problem examined above had input data furnished by some source and output data read by some inter-preter. It is true that these input data appeared on the basic

storage medium of the machine (Table 5-1) and that they were read from this basic medium by the final interpreter, but this is not necessarily a general state of affairs. It is likely that a sufficiently fast and accurate storage component for a machine of the speed considered here will be an integral part of the machine, not a detachable part. Thus some convenient attachable medium for input of basic data must be provided, and some convenient detachable medium for output of results must also be provided.

5-4. Stereotyped Commands for the General-purpose Machine

For the sample problem, somewhat similar to some problems in data reduction, a set of 11 commands of a stereotyped nature, designated by 1″ to 11″, and one command not stereotyped located between commands 2″ and 3″ were written above. If these commands are carried out in sequence, the data will be reduced as required. These commands have the advantages that a few commands govern a large number of calculation steps and that the commands themselves are all similar (except for the one nonconforming command); specifically, their similarity is that each specifies an arithmetic operation, and each specifies where the numbers entering into this operation can be found and where to dispose of the number generated by this operation. These commands have disadvantages in that they were written down with no particular regard for the engineering aspects of the computing components and they include the nonconforming command.

A set of 43 commands is listed below. Each is of a stereotyped nature, is written with some regard to ease of engineering achievement, and is so chosen that the sequence of 43 commands controls the same set of calculations as the earlier sequence. In this sequence there are no nonconforming operations, and every operation has a stereotyped form. This form specifies a designator for the operation, a description of the operation, and the location of one number involved in the operation. No operation requires further specification. In an engineering realization, the designator does not turn out to be a part of the command itself but is a statement of where the command is to be placed in the machine. It is assumed that the location of the number involved in the operation is itself a numerical code with 10

assigned values. Thus, every command can be written as a number, a few digits of which state the code number describing the prescribed operation and the remaining significant digits of which state the address code number of a number involved in the operation. In addition, the designator, or the place in the machine at which this command will be placed, must also be specified.

A machine for which commands can be written in this way is clearly a versatile machine. Since the commands are numbers and since the machine must expect to have numbers introduced to it in some fashion, it certainly should not be difficult to enter these coded commands into the machine. Since a small number of commands will control a lengthy calculation, the amount of time spent in writing these commands down for insertion into the machine is not intolerably long. Since only a few different operations are required, the engineering structure of the machine may be expected to be attainable fairly simply. Commands of this sort have been described most completely and lucidly by Burks, Goldstine, and von Neumann.[2] They describe a set of 21 operations which suffice for conveniently carrying out most problems from a large field. The authors describe the machine on which these commands will be carried out specifically enough to give a general idea of the engineering considerations involved.

Bloch, Campbell, and Ellis[1] have described a set of commands more nearly of the type used in the list 1″ to 11″. This publication may be more readily available than the Burks, Goldstine, and von Neumann paper, but the description is not as complete in some respects. On the other hand, a special checking scheme and some other novel ideas are discussed by Bloch, Campbell, and Ellis.

The commands which will be used below are a set taken essentially from the set proposed by Burks, Goldstine, and von Neumann.

5-4-1. Description of the Machine. The input and output of the machine will be ignored in this description. The machine will possess a storage element, an arithmetic element, and a control element. The arithmetic element will contain two main units: the accumulator, which will be denoted by the letter A, and the multiplier register, which will be denoted by the letter R. Addition can take place in the accumulator. It will be assumed

that these elements can handle numbers of the sizes which will occur (although the accomplishment of this assumption may not be perfectly straightforward). The positions in which numbers can be stored will be addressed by numbers; thus, an address is an ordinal number of a position in storage, at which position a number involved in the calculation is to be found or to be placed. It will be assumed that the storage facilities are adequate for the problem in hand and that the size of the arithmetic unit is sufficiently great to carry out the calculations which will later be seen to be involved in modifying the commands. (Specifically the arithmetic unit must handle numbers as large as the largest ordinal number occurring in connection with the storage.) The commands which the machine is capable of following are described below in the order of their appearance in Table 5–2, starting on page 69.

The first operation that the machine will be required to carry out is the clearing of the accumulator and the insertion in the accumulator of the number at a designated address. This operation occurs first in command $1'''$ below. The machine will be built so that the number placed into the accumulator remains unchanged in its storage position. The operation will not affect R.

The second operation is the inverse of that above; it is the operation of storing at a specified address the number which is in the accumulator. This operation destroys any number previously stored at that address, leaves intact the number in A, and does not affect R. This operation occurs first in command $2'''$ below.

The third operation is the operation of destroying any number which is in R and replacing it by the number at a specified address; this operation occurs first in command $3'''$. The number remains unchanged in its storage position, and the operation does not affect A.

The fourth operation causes the number in the accumulator to be destroyed and the product of the number in the R register, as multiplier, and the number at a specified address, as multiplicand, to be inserted in A. This operation occurs first in command $4'''$. The number at the specified address remains unchanged; the number in R will be destroyed during the operation.

The fifth operation retains the number in A and adds to it the

number at a specified address. It occurs first in command 8‴. The addend remains unchanged in its storage position. The operation does not affect R.

The sixth operation is a subtraction operation similar to the one above. The number in A is retained, and the number at a specified address is subtracted from it, the difference being kept in A. The operation occurs first in command 21‴. The number at the specified address remains unchanged at that address, and the operation does not affect R.

The seventh operation occurs first in command 28‴ below. To understand it, it is necessary first to have described the normal sequencing of commands. A program like the one written below is set up in a fixed sequence. The machine automatically carries out each command in the order of this sequence unless a command tells it to depart from this order. (Bloch, Campbell, and Ellis[1] describe a different procedure for specifying the order in which commands are to be carried out.) Thus, a few commands must be used to change the order in which commands are carried out. If these commands were not included, it would be impossible to recycle, and it would be necessary to write one command for every arithmetical operation the machine is to carry out. This seventh command is one used for altering the sequence conditionally; if the number in the accumulator is a negative number, there is no alteration in sequence, and the machine next carries out the next command (29‴), but if the number in A is nonnegative, the machine passes to the command specified in the address portion of command 28‴. 28‴ is called a *conditional transfer* command or *conditional jump*.

The description of this operation and of one or two others below is made more precise by specifying that the commands themselves will be placed in the storage unit of the machine. This is possible because the commands can be written as numbers. It was pointed out above that all that is required is that the operations to be carried out be assigned code numbers and that a complete command be written as a code number specifying an operation followed by one specifying a numerical address. The digits of these two parts may be run together to form a single multidigit number. The description is complete when it is stated that the designator of the command is simply the address in the storage system at which the command is stored. Suc-

cessive commands in the sequence of commands will be stored at successively numbered addresses. Thus, the control unit will be required to start at a specified place and to withdraw commands from storage. After each command is carried out, the control unit passes to the next higher numbered storage cell for the next command unless it has been instructed by the command just carried out to depart from the standard command sequence. Thus recycling and alternate sequences are made possible.

The eighth operation is similar to the second except that the information transferred from A to the storage unit is restricted to be information concerning the address of a command. No matter what is in the accumulator, the portion of the stored command having to do with the operation to be carried out will not be changed; the only change is a change in address. This is an operation which is of fundamental importance in removing the nonconforming command of the set described earlier. The operation first occurs in command $31'''$.

The ninth operation of the set is one which instructs the control to depart from the standard sequence of commands. The nature of its utility is apparent from its use in command $43'''$ below. Nothing in the arithmetic unit or the storage unit is affected by this command, which is called an *unconditional transfer*.

The final operation of the set appears in storage at cell j. It tells the machine to stop calculations. Without this command the machine at the end of a calculation would proceed with meaningless and possibly harmful calculations.

5-4-2. The Coding. A coding which uses the operations described above can be written immediately from the set $1''$ through $11''$. All numbers involved will be assumed to have been inserted in storage. Addresses in storage will correspond to specification of position (measurement number and column heading) on the work sheet referred to earlier. A block of storage addresses must be assigned to each column in which are written data which must be retained permanently; these are the columns headed x, y, and u. The other numbers (which were never written down at all in the program 1 through 5 above) are assigned space as necessary during the time they must be saved; these numbers are then discarded in favor of their successors. Finally, a block of consecutively numbered addresses

TABLE 5-2. COMMANDS

Designator	Operation	Address	Remarks
$1'''$	Clear A and insert the number at	c	
$2'''$	Store number now in A at	d	Completes $2''$
$3'''$	Clear R and insert the number at	a	
$4'''$	Clear A and insert in it the number got by multiplying by the number in R the number at	a	
$5'''$	Store number now in A at	e	Completes $3''$
$6'''$	Clear R and insert the number at	b	
$7'''$	Clear A and insert in it the number got by multiplying by the number in R the number at	b	Completes $4''$
$8'''$	Add to the number in A the number at	e	
$9'''$	Store number now in A at	e	Completes $5''$
$10'''$	Clear R and insert the number at	d	
$11'''$	Clear A and insert in it the number got by multiplying by the number in R the number at	e	Completes $6''$
$12'''$	Add to the number in A the number at	c	
$13'''$	Store the number now in A at	e	Completes $7''$
$14'''$	Clear R and insert the number at	f	
$15'''$	Clear A and insert in it the number got by multiplying by the number in R the number at	e	
$16'''$	Store the number now in A at	$c+1$	Completes $8''$
$17'''$	Clear R and insert in it the number at	d	
$18'''$	Clear A and insert in it the number got by multiplying by the number in R the number at	$c+1$	
$19'''$	Store the number now in A at	e	Completes $9''$

TABLE 5-2. COMMANDS. (*Continued*)

Designator	Operation	Address	Remarks
20'''	Clear A and insert in it the number at	g	
21'''	Subtract from the number in A the number at	e	
22'''	Store the number now in A at	e	Completes 10''
23'''	Clear R and insert in it the number at	d	
24'''	Clear A and insert in it the number got by multiplying by the number in R the number at	e	
25'''	Store the number now at A at	d	Completes 11''
26'''	Clear A and insert the number at	h	
27'''	Subtract from A the number at	i	
28'''	If the number in the accumulator is negative, pass to the next command in this sequence; if it is not negative, jump to the command stored at	j	Checks to see whether calculation is complete
29'''	Clear A and insert the number at	h	
30'''	Add to the number in A the number at	c	
31'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	3'''	
32'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	4'''	
33'''	Store the number now in A at	h	
34'''	Add to the number in A the number at	k	
35'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	6'''	

TABLE 5-2. COMMANDS. (*Continued*)

Designator	Operation	Address	Remarks
36'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	7'''	
37'''	Add to the number in A the number at	l	
38'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	12'''	
39'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	30'''	
40'''	Add to the number in A the number at	c	
41'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	16'''	
42'''	Store the number in A as the address part (leaving the coded operation unchanged) of the command	18'''	
43'''	For the next operation proceed in order after jumping to command	3'''	Repeat command 3''' and all following. This command replaces the unstereotyped command between 2'' and 3'' on the earlier list.

is assigned to the 43 commands 1''' through 43'''. The control unit causes the commands to be carried out in sequence except where deviations are commanded by one of the sequence of commands.

Specifically, in summary, the storage unit is loaded as follows: The numbers x are stored at consecutively numbered addresses, with the first at address a. The numbers y are stored at consecutively numbered addresses with the first at address b. The numbers u_k as computed will be placed in consecutively numbered

addresses with the first at address c.   Suppose that there are n measurements; then the last x is stored at $a + n - 1$, etc.   A block of 43 consecutive addresses is assigned to the coded commands $1'''$ through $43'''$.   Special numbers are stored at address d and eight other addresses designated by letters e through l according to the scheme below:

Address	Number
c	1
d	Latest entry for u^{-1}
e	Latest entry for x^2, $x^2 + y^2$, $2u$, $u \cdot u_{k-1}^{-1}$, and $2 - u \cdot u_{k-1}^{-1}$ in order
f	0.5
g	2
h	The number a, to be replaced by the current address
i	The number $a + n$
j	The coded value of the command *stop calculation*
k	The number $b - a$
l	The number $c - b$

Note that command $1''$ of the early set has been carried out already through these storage operations.

5-5. Plan for General-purpose Machine

The example above indicates a plan for the convenient operation of a general-purpose machine.   The operational requirement is simply that a few written commands can conveniently control a lengthy calculation.   The realization of the requirement suggested is by means of a few stereotyped command forms, each of which specifies an operation and one or more storage positions.   The arithmetic element must be able to carry out the operations contained in these stereotyped commands.   The control element must go through the pertinent commands in order, translate the coded description of the operations to pulses along busses which will cause the arithmetic elements involved to carry out the specified operations, and direct numbers to and from the specified storage locations.

The requirement that a few written commands direct a long calculation is attained by (1) the inclusion of facilities permitting the reuse of commands with the operations unchanged and the addresses changed if desired; and (2) an operation permitting a departure from the normal sequence to a secondary one (con-

sisting of a single *conditional transfer* command in the example), the departure to depend on the calculation.

REFERENCES

1. Bloch, R. M., R. V. D. Campbell, and M. Ellis, Logical Design of the Raytheon Computer, *Mathematical Tables and Other Aids to Computation*, Vol. 3, No. 24, pp. 286–295, 317–322 (October, 1948).
2. Burks, Arthur H., Herman H. Goldstine, and John von Neumann, *Report on the Mathematical and Logical Aspects of an Electronic Computing Instrument*, 2d ed., Part I: *Preliminary Discussion of the Logical Design of an Electronic Computing Instrument* (Institute for Advanced Study, Princeton, N.J., 1947).
3. Harvard University, Staff of the Computation Laboratory, *A Manual of Operation for the Automatic Sequence Controlled Calculator*, *Annals of the Computation Laboratory of Harvard University*, Vol. I (Harvard University Press, Cambridge, Mass., 1946).

CHAPTER 6

ARITHMETIC SYSTEMS

6-1. Introduction

Most modern arithmetics are founded upon a system for representing the positive integers in terms of systematic symbols involving the concept of carry-over. That is, they depend upon a systematic representation of the counting process. These arithmetics are the basis of all automatic computing operations now used.

The basic concepts used in the more usual arithmetics have been described in Carmichael.[3] A detailed exposition will be given here of the concepts of number theory which may be applied directly to an understanding of computer arithmetic systems.

6-2. The Fundamental Counting Systems

The recording of counted numbers in the systems to be studied here will be by means of an infinite ordered sequence of marks, of which the first will be considered to be at the right-hand end and the others to proceed in order to the left. The nature of the sequence will be defined inductively. In each position the mark 0 (or a mark with an equivalent meaning) will be admissible as a mark, and usually one or more other marks will be admissible. For example, in the decimal system the marks admissible in each position are the marks 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. The admissible marks will not depend upon the marks present in other positions. For each position, the admissible marks have an order assigned, and the mark 0 precedes all the rest. (The order of the decimal marks is that in which they are given above.)

The rule for denoting a number is given in the following parts:

1. The number zero is denoted by a sequence each of whose elements is the mark 0.
2. To pass from the representation of any number to that

of the next higher, the first mark which is not the last admissible mark for its position is replaced by the next admissible mark in the set established for its position and all lower order marks are replaced by 0's.

6-3. Examples of Counting Systems

The decimal system of counting follows this rule, with the understanding that a final string of 0's in the sequence, extending to the left, is not written. As noted above, the ordered set of marks which may be used in each position is the set 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. The number 57 is, according to the rule, followed by 58, for 7 in the first (right-hand) position is the first mark which is not the last admissible one in its position and it is replaced by 8, the next higher mark. The number 26,999 is followed by 27,000, for the first three marks are all the last admissible marks in their positions, so the fourth one (which is not maximal) is increased by one and the last three are replaced by 0's. The binary number system admits the marks 0 and 1 at each position, and no others. Thus the counting rule gives the equivalences shown in Table 6-1 between the first 36 decimal and binary numbers. In each case, 0's after the last nonzero mark are omitted except for the number zero.

TABLE 6-1

Decimal	Binary	Decimal	Binary	Decimal	Binary
0	0	12	1,100	24	11,000
1	1	13	1,101	25	11,001
2	10	14	1,110	26	11,010
3	11	15	1,111	27	11,011
4	100	16	10,000	28	11,100
5	101	17	10,001	29	11,101
6	110	18	10,010	30	11,110
7	111	19	10,011	31	11,111
8	1,000	20	10,100	32	100,000
9	1,001	21	10,101	33	100,001
10	1,010	22	10,110	34	100,010
11	1,011	23	10,111	35	100,011

The binary system occurs frequently in considerations of automatic computers, for many of the basic circuits, such as the

flip-flop, naturally assume two stable states and are therefore convenient to use with binary numbers.

In the octal system, any one of eight different marks is admissible at each position, *e.g.*, 0, 1, 2, 3, 4, 5, 6, 7 or the binary equivalents of these same familiar digits, *000, 001, 010, 011, 100, 101, 110, 111*. This system is of some interest wherever binary numbers occur because conversion between binary and octal numbers is trivially easy, and octal numbers make a fairly efficient set for printing and reading. These points are exhibited in greater detail in Sec. 6-6.

The biquinary system was used as a convenient equivalent of the decimal system in some machines produced by the Bell Telephone Laboratories.[1] This application is attributed to Dr. George Stibitz. The system is one in which the decimal system is factored and written as a system in which the odd-numbered positions have five admissible marks and the even-numbered positions have two admissible marks. The admissible marks for the odd positions are 0, 1, 2, 3, and 4, and for the even positions 0 and 5. The equivalence of this system with a decimal system should be obvious, for the sum of a pair of successive marks can be considered to be a decimal mark itself, there being 10 such pairs. For example, the decimal numbers 4 and 8 are represented in the biquinary notation as 04 and 53, respectively. Carry-over occurs at 10 as in the decimal system.

This arithmetic system is much like the tetraquinary system of counting of the Mayan Indians, who had at least two systems of counting, one for general usage and at least one other for dates. These have been described by Morley[4] and by Spinden.[7] The Mayans used a true vigesimal system with a systematic rule of forming the marks. It is said that the Mayans used their vigesimal system prior to the time the decimal system achieved general usage in Europe. In this system there were 20 marks for each position; however, each mark (except zero) consisted of from zero to four dots placed over from zero to three bars. This system was obviously equivalent to a system in which the odd-numbered positions have five marks (the dot configurations) and the even-numbered positions four (the bar configurations). The Mayan zero, a "shell," was used in the formation of 20, 40, etc. Although the Mayans used the system as a true vigesimal number system, they counted as if it were a tetraquinary system:

the number of bars represented the number of hands and feet they had used up in the count, and the number of dots represented the number of odd digits on the next hand or foot.

TABLE 6-2

Decimal	Mayan	Decimal	Mayan
0		10	
1		11	
2		12	
3		13	
4		14	
5		15	
6		16	
7		17	
8		18	
9		19	

The monetary system used in the United Kingdom may be thought of as another, more complicated counting system, satisfying the rule of Sec. 6-2. The mark in the first position gives the number of pence, the mark in the second position gives the number of shillings, and the mark in the third position gives the number of pounds. It is common practice to separate the different positions by a diagonal (/), writing 26/14/11 for 26 pounds, 14 shillings, 11 pence. There are only three positions. Assuming that the penny is the unit (*i.e.*, ignoring the halfpenny), the marks which are admissible in the first position are the decimal numbers from 0 through 11 in order; those admissible in the second position are the decimal numbers 0 through 19 in order, and the infinite ordered set of nonnegative decimal integers is the admissible set for the third position.

6-4. Three Fundamental Theorems

The main principles of behavior of numbers under different number representations may be characterized by three elementary theorems. For all these theorems it will be assumed that there are n_1 different admissible marks for the first position, n_2 for the second, n_3 for the third, etc., with n_i for the ith. The theorems are generally trivially true for any n_i which become infinite, and these cases will be disregarded (*e.g.*, the case of the British monetary system).

Theorem 6-1. *The number of different numbers which can be represented by configurations with marks different from 0 in at most the first k positions is the product*

$$p_k = n_1 n_2 n_3 \cdots n_k \tag{6-1}$$

For proof of Theorem 6-1 see Sec. 6-13-1.

From this theorem, it follows that exactly 1,000 ($10 \times 10 \times 10$) different numbers can be represented with three decimal digits; these are the numbers from 0 through 999, inclusive. Similarly, to represent 1,000 different numbers with binary digits, 10 digits will be required, for 1,000 lies between 2 to the ninth and 2 to the tenth powers (512 and 1,024, respectively, expressed in decimal numbers).

Theorem 6-2. *In the representation of a number x, if p_1 represents the number of marks admissible which precede the mark actually occurring in the first position, if p_2 is the number of marks admissible for the second position which precede the mark in that position, etc., with p_i the number of admissible marks for the ith position preceding the ith mark, then, conforming with the notation for P of Eq. (6-1),*

$$x = p_1 + p_2 P_1 + p_3 P_2 + \cdots + p_i P_{i-1} \tag{6-2}$$

For proof see Section 6-13-2. Theorem 6-2 can be used to write the value in any known arithmetic of a number expressed in any other arithmetic. Thus, for example, the binary number 10,001,001 may be expressed in decimal notation by noting that according to the rule it must equal

$$1 \cdot 2^7 + 0 \cdot 2^6 + 0 \cdot 2^5 + 0 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1$$
$$= 128 + 8 + 1 = 137$$

Similarly, in a ternary system which has the admissible marks 0, 1, and 2, the ternary number 12,021 has an equivalent decimal value

$$1 \cdot 3^4 + 2 \cdot 3^3 + 0 \cdot 3^2 + 2 \cdot 3^1 + 1 = 81 + 54 + 6 + 1 = 142$$

Theorem 6-3. *If x is a nonnegative number and if n_i, p_i, and P_i all have the significance defined above for some system of notation in which x is expressed, then p_1 is the remainder when x is divided by n_1, p_2 is the remainder when this integral quotient (without remainder) is divided by n_2, etc.: explicitly, there exists a set of*

nonnegative integers q_0, q_1, q_2, . . . , q_i such that

$$q_0 = x = n_1 q_1 + p_1; \qquad 0 \leq p_1 < n_1$$
$$q_1 = n_2 q_2 + p_2; \qquad 0 \leq p_2 < n_2$$
$$\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \qquad (6\text{-}3)$$
$$q_i = n_{i+1} q_{i+1} + p_{i+1}; \qquad 0 \leq p_{i+1} < n_{i+1}$$
$$\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots$$

These equations uniquely determine p_i. For proof see Sec. 6-13-3.

Theorem 6-3 furnishes a method for writing the expression for any positive integer in terms of any system of marks, carrying out the arithmetic in any known system. For example, to translate the number 137 to the binary system, where every $n_i = 2$, repeated division by 2 suffices, as shown in Table 6-3.

TABLE 6-3

i	n_i	(q_{i-1})	(p_{i-1})
1	2	137	
2	2	68	1
3	2	34	0
4	2	17	0
5	2	8	1
6	2	4	0
7	2	2	0
8	2	1	0
9	2	0	1

Table 6-3 was calculated as follows: First all the numbers n_i were written in their column; then the number 137 was entered in the center column. The division was performed q_{i-1}/n_i, and the quotient written directly below with the remainder written in the right-hand column. The binary representation is written by copying the right-hand column, the upper mark being placed at the right in copying:

Decimal 137;   binary 10,001,001

A warning is in order at this point: this rule stated in Theorem 6-3 cannot be applied blindly to fractional numbers expressed, for example, with a decimal point.

6-5. Rules of Arithmetic for Numbers with a Single Radix

If all the numbers n_i of a system of notation like those described above are equal, so that $n_i = r$ for every i, then the common value is known as the *radix* of the system. In particular, the ordinary decimal number system is the radix 10 system, the binary system is the radix 2 system, and the Mayan system is the radix 20 system. For automatic calculating machinery, the use of a radix implies that the counters at each position can be identical, so radical numbers are usually used; in cases such as the biquinary arithmetic described by Alt,[1] the use is essentially as decimal, or radix 10, arithmetic.

For any radix arithmetic the basic tables corresponding to the addition and the multiplication tables of decimal arithmetic can be written, and from them the operations of addition, multiplication, subtraction, and division can be carried out.

For binary arithmetic, the tables are given in Tables 6-4 and

TABLE 6-4. ADDITION

Augend Addend	0	1
	Sum	
0	0	1
1	1	10

TABLE 6-5. MULTIPLICATION

Multiplicand Muliplier	0	1
	Product	
0	0	0
1	0	1

6-5. The rules of carry are the same as in arithmetic of any other radix. The construction of the tables is the same.

For octal arithmetic, the tables are given in Tables 6-6 and

TABLE 6-6. ADDITION

Augend Addend	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7	10
2	2	3	4	5	6	7	10	11
3	3	4	5	6	7	10	11	12
4	4	5	6	7	10	11	12	13
5	5	6	7	10	11	12	13	14
6	6	7	10	11	12	13	14	15
7	7	10	11	12	13	14	15	16

ARITHMETIC SYSTEMS

$$100000010110$$
$$11100110$$
$$\overline{111101100010110}$$

with no carries.

For an octal multiplication, consider multiplying 346 by 267. Again, the multiplying steps:

346	346	346
7	6	2
52	44	14
34	30	10
25	22	6
3112	2544	714

The final totalization is

$$3112$$
$$2544$$
$$714$$
$$\overline{122152}$$

Binary division is performed in much the same way as ordinary decimal long division. This may be illustrated in the following example by dividing the binary equivalent of decimal 137 by the binary equivalent of decimal 10:

```
                     begi
                     1101
                1010)10001001
            a    1010
            c    1110
            d    1010
            f     10001
            h      1010
            j       111
```

As shown above, the divisor is placed outside of the long-division sign, and the dividend is placed inside. By a series of successive subtractions, a quotient is formed above the long-division sign. The first step in the division process is to place the divisor beneath the dividend in a position as far to the left as a positive difference will allow. This is shown on line *a*. The first digit 1 of the quotient is placed at *b* in the same column as the lowest order

digit of the divisor. The first digit 1 of the divisor signifies that only one subtraction can be performed before a negative difference is encountered for this position of the divisor. The divisor is then subtracted from the dividend to produce the positive difference (line c). This difference is then compared with the divisor to note that it is smaller in value than the divisor. The next digit in the dividend is brought down to the difference obtained on line c. If the new number on line c is larger than the divisor, the divisor is then placed under this number and the subtraction process repeated. If the new number is less than the divisor, a 0 is placed in the quotient in the same column as the lowest order digit of the divisor in this position. The next digit of the dividend is then brought down to the difference, the divisor shifted one place to the right, and the subtraction process is continued until the quotient is completed to the radical point. Any remainder after the last subtraction is treated in the same manner as in ordinary long division. In the example shown above, the quotient is 1,101 with a remainder of 111 ÷ 1010; this corresponds to the decimal number $13\frac{7}{10}$.

In the above example, it should be noted that no multiplication process was mentioned. This is apparent from the realization that the greatest number of allowable times the divisor goes into the dividend for any partial quotient is one, since this is the highest single-order digit in the binary system. Therefore, the product of the nonzero digit last added to the quotient and the divisor is the same as the divisor alone. This is not true, however, for other radix systems. For example, use is made of the octal multiplication table, displayed above, in the performance of the same division in the octal system as follows:

$$
\begin{array}{r}
15 \\
12\overline{)211} \\
12 \\
\hline
71 \\
62 \\
\hline
7
\end{array}
$$

The quotient, as before, has a decimal equivalent of $13\frac{7}{10}$.

6-6. A Note Concerning Binary-to-octal Conversion

It is true according to Theorem 6-1 that exactly eight numbers can be represented with three binary digits. From this it follows

that triplets of binary digits may be used as marks to represent octal digits, and that conversion from binary to octal notation simply implies changing the notation used for the marks.   Specifically, the notation change is as shown in Table 6-8.   The

<div align="center">

TABLE 6-8

Binary Triplet	Octal Mark
000	0
001	1
010	2
011	3
100	4
101	5
110	6
111	7

</div>

change is made triplet by triplet, beginning at the right.   For example, the binary number 10,001,001 is the octal number 211; the binary 101,100,110 is octal 546, etc.

The binary system has obvious engineering advantages in computer design.   The expression of a number in binary notation requires a comparatively large number of digits, however.   For example, a number with six decimal digits may require 20 binary digits; the same number, in octal notation, may need only seven.

The number of octal digits is thus only slightly greater than the number of decimal digits, the conversion from binary representation is trivially easy, and an inexperienced reader can get good qualitative estimates of the results without worrying about the octal character of the representation.   It can be seen, therefore, that the octal system has certain features which make it very attractive.

6-7. Economy Attained by Radix Choice

The economy to be gained by choice of radix can be estimated as follows: From Theorem 6-1, it is known that the number of numbers expressible with n digits radix r is r^n.   It was noted in Chap. 3 that, using two-state tubes or relays, and for small values of r, each digit radix r requires r triodes or relays (but for larger values of r each digit requires up to $2r$ triodes or relays). Assuming that the complexity is measured according to the number of tubes, it is possible to determine the theoretical optimum value for the radix of the arithmetic system.

Let $N = rn$ represent a fair estimate of the number of tubes required in the system and $M = r^n$ be the largest number of numbers expressible in the system, where r designates the radix and n designates the number of digits. The value of N should be a minimum, subject to M being fixed in value. Then

$$M = r^n \tag{6-4}$$
$$\log_a M = n \log_a r = M' \tag{6-5}$$
$$n = \frac{M'}{\log_a r} \tag{6-6}$$

Substituting the value of n in the expression for N,

$$N = \frac{M'r}{\log_a r} \tag{6-7}$$

Since

$$\log_a r = \log_e r \cdot \log_a e \tag{6-8}$$

then

$$\frac{dN}{dr} = M'\frac{(\log_a r - \log_a e)}{(\log_a r)^2} = 0 \qquad . \tag{6-9}$$

If

$$\frac{M'}{(\log_a r)^2} \neq 0 \tag{6-10}$$

$$\log_a r - \log_a e = 0$$

and

$$r = e = 2.71828 \cdots \tag{6-11}$$

This is the only possible value for a minimum, and therefore N must be a monotonic function of r for $r > e$, and a monotonic function for $r < e$.

TABLE 6-9

r	N
2	39.20
3	38.24
4	39.20
5	42.90
10	60.00

Now, consider some values, with M fixed at a value 10^6, as shown in Table 6-9. Note that N increases with the values of r above 3. This increase must continue, for r has no further critical values. Under these assumptions, the radix 3, on the average, is the most economical choice, closely followed by radices

TABLE 6-10

Radix					
10	8	5	4	3	2
Number Representation					
1	1	1	1	1	1
2	2	2	2	2	10
3	3	3	3	10	11
4	4	4	10	11	100
5	5	10	11	12	101
6	6	11	12	20	110
7	7	12	13	21	111
8	10	13	20	22	1,000
9	11	14	21	100	1,001
10	12	20	22	101	1,010
11	13	21	23	102	1,011
12	14	22	30	110	1,100
13	15	23	31	111	1,101
14	16	24	32	112	1,110
15	17	30	33	120	1,111
16	20	31	100	121	10,000
17	21	32	101	122	10,001
18	22	33	102	200	10,010
19	23	34	103	201	10,011
20	24	40	110	202	10,100
21	25	41	111	210	10,101
22	26	42	112	211	10,110
23	27	43	113	212	10,111
24	30	44	120	220	11,000
25	31	100	121	221	11,001
30	36	110	132	1,010	11,110
35	43	120	203	1,022	100,011
40	50	130	220	1,111	101,000
45	55	140	231	1,200	101,101
50	62	200	302	1,212	110,010
100	144	400	1,210	10,201	1,100,100
1,000	1,750	13,000	33,220	1,101,001	1,111,101,000
10,000	23,420	310,000	2,130,100	111,201,101	10,011,100,010,000

2 and 4. These assumptions are, of course, only approximately valid, and the choice of 2 as a radix is frequently justified on more complete analysis. It should be noted that, even with the optimistic assumption that 10 triodes will yield a reliable decimal ring, radix 10 leads to about one and one-half times the com-

plexity of radix 2, 3, or 4. This is probably significant despite the shallow nature of the argument used here.

For reference, corresponding representations of the same numbers for radices 2, 3, 4, 5, 8, and 10 are tabulated in Table 6-10.

6-8. Negatives and Complements

A desk calculator, in performing subtraction, uses a reversing gear and runs the wheels backward. If the subtraction goes far enough to generate a negative number, a difficulty is encountered. The difficulty, although easily resolved, arises because the indicated numbers continue to decrease, while in standard notation they would suddenly be preceded by a minus sign and begin to increase. The solution lies in noting that the machine has a limited capacity. If the machine has 10-decimal-digit capacity in a register, it can hold numbers as high as $10^{10} - 1$, but it cannot hold 10^{10}. The machine is unable to distinguish between two numbers which differ by any integral multiple of 10^{10}; in terms of the number theorist, it performs arithmetic modulo 10^{10}. Any number is represented on this machine by the number plus or minus an integral multiple of 10^{10} chosen in such a way as to bring the indicated number to a value between the limits 0 and $10^{10} - 1$, inclusive. In particular, the number *negative* 137 would be represented by $-137 + 10^{10} = 9,999,999,863$. Once this condition is recognized, there is no difficulty if the coefficient of the additive 10^{10} is kept in mind; for almost all calculations it is 0 or 1.

For many of the more rapid machines using other types of EPDCC's as computing elements, this system of representation of negative numbers and this system of subtraction are not satisfactory in general, since most of the faster EPDCC's cannot be made to run backward conveniently. Because such components will run in only one direction, the process of subtraction must be carried out in the manner defined in many algebra texts: change the sign and add. Furthermore, the question of changing the sign and adding cannot be treated too lightly, for if numbers are represented by their absolute values preceded by a plus or a minus sign, the result of adding a pair of 1's in the right-hand place depends on the signs of the numbers. Furthermore, if the radix is anything but 2, a binary mark must be used to denote

the sign while all other marks are different from this. This situation may lead to unnecessary complexity in the machine.

The number 9,999,999,863 written above is called the complement of the number 137 with respect to 10^{10}. All the digits to the left of its last nonzero digit are obtained by subtracting the corresponding digits of the original number from 9 (with zero written in front as many times as is necessary). The last nonzero digit is obtained by subtracting the last nonzero digit of the original number from 10, and if there are any zeros to the right of this point in the original number, they are replaced by zeros in the complement. Thus the complement of 5,280 is 9,999,994,720. *The* 9's *complement of a decimal number is the number obtained by replacing each digit of the original number by* 9 *minus that digit*. To perform subtraction, a fast machine could form these complements and add, but this process is still inconvenient. Such a method demands that certain digits be treated in a special manner, and the positions of the specially treated digits vary from one number to another.

The solution of treating all digits alike is so attractive that it has been tried arbitrarily, and it is used in most machines now being considered. The 9's complement of a positive number is used in this system to represent the negative number with the same absolute value. One disadvantage is immediately apparent: the 10-digit numbers 0,000,000,000 and 9,999,999,999 represent the same number. This disadvantage is immediately offset by the fact that the devices in which these complements are to be used may add or they may subtract, but they cannot do both. The number 0,000,000,000 turns out never to be generated in a machine which always adds, and the number 9,999,999,999 is never generated in a machine which always subtracts.

With 9's complements, numbers of 10 decimal digits are expressed modulo $10^{10} - 1$ rather than modulo 10^{10} as above; *i.e.*, numbers are stated with an implicit understanding that an integral multiple of $10^{10} - 1$ (with coefficient invariably 0 or 1 in common practice) has been added to the number originally to be represented. When two such numbers are added and their sum exceeds $10^{10} - 1$, the machine must automatically reduce them by this amount. To do this, it must detect the fact that the number is excessive and then correct. The following statements govern completely:

1. The sum of two 10-digit decimal numbers exceeds $10^{10} - 1$ if and only if there is a carry beyond the 10th place.
2. If there is a carry from the tenth place, the sum modulo $10^{10} - 1$ may be represented with 10 decimal digits by adding this carry to the extreme right-hand digit. This is the so-called *end-around carry*.

The first statement is obvious and requires no comment. The second is an automatic scheme for subtracting 10^{10} from the sum, by not writing the carry in the 11th decimal position, and then adding 1 to the sum, by carrying to the first decimal place. The net result of this operation is to subtract $10^{10} - 1$; this is allowable in the arithmetic modulo $10^{10} - 1$.

Completely analogous rules hold for arithmetic with n digits radix r. They will be stated in a theorem the proof of which is omitted.

Theorem 6-4. *If x is a positive number expressed with n digits radix r, i.e., $0 \leq x \leq r^n - 1$, then $-x$ may be expressed modulo $r^n - 1$ by replacing each of the n digits of x by $r - 1$ minus this digit; to add, modulo $r^n - 1$, two numbers expressed with n digits radix r, proceed according to the rules of arithmetic radix r except that the carry from the nth digit, if it occurs, is to be added to the first digit (end-around carry).*

A particular example illustrates the principle. Two five-digit binary numbers will be added: $10,110 + 01,101$. With ordinary binary arithmetic without end-around carry, the answer is $100,011$. To reduce to modulo $2^5 - 1$, the sixth digit is deleted and added to the first with additional carries resulting to give $00,100$.

In binary arithmetic of n digits it is usually assumed that numbers with the nth digit 1 are negative. Under this assumption the above addition becomes $1,101 - 1,001 = 100$.

6-9. Scale Factors and Radical Points

The foregoing description of the arithmetic of integers in terms of radix r digits can be enlarged to include fractional and other numbers occurring in analysis. The basis of the enlargement is classical, and only the most immediately applicable results will be given here. A complete description can be found in Ritt.[6]

Many of the numbers which are dealt with in scientific computation are derived from measurements of physical quantities. If such a quantity is to be represented in a register of n digits, it must be expressed in terms of a number of not more than n digits, whether or not it has been measured precisely enough to warrant expression in terms of a greater number. For example, to express π in a five-digit radix 10 register, we write 3.1416. This expresses the fact that π lies between 3.14155 and 3.14165, which is entirely correct. 3.1416 is an arbitrarily precise expression for a number the magnitude of which is known more precisely than can be expressed in five decimal digits.

Furthermore, it is the use of the decimal point which permits this convenient representation of the number in five digits. 3.1416 is a fraction. In terms of integers it is $31,416/10^4$. The same scheme may be used with any other radix. If a radical point is placed to the left of the kth digit (from the right) of a number expressed in radix r notation, the number is to be read as the integer, as if there had been no radical point, divided by r^k. Thus

$$\text{Binary } 101.01101 = \text{decimal } \frac{2^7 + 2^5 + 2^3 + 2^2 + 1}{2^5}$$
$$= \text{decimal } 173\!\!\not{3}\!\!/_{32}$$
$$= \text{decimal } 5.406$$

Note that the numerator and the implied denominator must both be translated if the radix is changed and that the rule in Theorem 6-3 cannot be carelessly applied to this case of fractional representation. It is not difficult, however, to devise a multiplication rule similar to the division rule of Theorem 6-3 to handle that portion of the number to the right of the radical point.

For machines with the radical point in a fixed position, and these machines are the simplest to build, the use of scale factors is necessary to obtain the most efficient utilization of the limited number of numbers representable by the machine. For this, the units in which the numbers appear are measured and rechosen in such a way that the largest number of any kind which will appear in the machine is just slightly smaller than the largest number the machine will hold. This is an operation analogous to measuring in inches rather than feet.

Some workers feel that there is a best place for the radical point. In particular, many, including Burks, Goldstine, and von Neumann,[2] feel that this best place is one binary digit to the right of the left-hand end of a binary number. It should be noted that the first left-hand digit does not represent a numerical value but merely serves to indicate whether the number is positive or negative. This placement used with 1's complements gives numbers ranging from -1 to $+1$, and the product of two such numbers lies in the same range. This feature, which does not demand a scale change when the products are formed, is one of the principal advantages of this system.

Other workers have arbitrarily placed the radical point at the right-hand end of the number; *i.e.*, they have built their machines to deal with integers. Thus the basic operations are programmed on the machine in terms of shifting operations. One such machine has been described by Engineering Research Associates, Inc.

A third possible scheme is one in which the position of the radical point is not designated in advance but is placed by the machine in performing each operation. This leads to engineering complexities, but it serves to reduce the burden of converting all numbers involved into a scale best suited for the machine. The Mark II Calculator built at the Computation Laboratory at Harvard University for the Naval Proving Ground at Dahlgren, Va., is an example of such a machine.

6-10. Binary Division—A Special Case

Binary division offers simplifications which have been noted by Burks, Goldstine, and von Neumann.[2] These authors suggest that the division process for binary numbers be modified to proceed according to the following steps in a system of commands:

1. Enter the dividend and the divisor so that like order digits are in the same columns.
2. Shift the divisor to place its first left-hand nonzero digit in the same column as that of the dividend; add as many 0's at the right-hand end of the divisor as necessary to make it equal in order to the dividend.
3. Subtract the divisor from the dividend and register a 1 in the first-order-digit position of the partial quotient register. To establish its proper position, shift this 1 a

number of digit places, to the right or to the left, corresponding to the number of places and direction by which the divisor was shifted in step 2.

4. For the next arithmetic operation, shift the divisor one place to the right.
5. If the remainder (difference) after step 3 is positive, subtract the shifted divisor from this remainder and add 1 to the partial quotient in the next position to the right of the previous 1; if the remainder is negative, add the shifted divisor to it and subtract 1 from the partial quotient in the new position.
6. Shift the divisor and proceed as in step 4 using the new remainder and modifying the quotient by 1 in the next position.
7. Continue until the process of division is completed, at this time adding the finally shifted divisor one additional time in carrying out a corresponding subtraction from the quotient, if necessary, to get a positive remainder.

As an example, a division which occurs as the first step in translating binary 10,001,001 to a decimal number will be performed, keeping all arithmetic in the binary system and using Theorem 6-3. It should be noted that where a negative difference is obtained in Table 6-11, it is represented by its 1's complement plus 1.

Since the last sum in Table 6-11 is positive, the division is complete. The quotient is 1,101 (decimal 13), and the remainder is 111 (decimal 7); these numbers agree with those found earlier.

The justification of this method, mathematically, lies in the trivial proof that the dividend is the sum of remainder and the product of the quotient with the divisor. In terms of machine operation the justification lies in the ease of programming and in the computing speed. The example is included as an applicable division scheme for binary digital computers.

The applicability of this scheme is enhanced by noting that a simple rule may be formulated for carrying out the division without borrows in the quotient. Consider that the quotient is formed from two numbers x and y such that the quotient z is $x - y$ and such that $x + y$, in binary notation, is a series of 1's. That is, the quotient

$$z = x - y \tag{6-12}$$

with

$$x + y = 2^n - 1 \tag{6-13}$$

for some n. It follows from the second expression that

$$z = 2x - 2^n + 1 \tag{6-14}$$

In the above example,

$$\left.\begin{array}{rl} z = & 1,101 \\ x = & 10,110 \\ y = & 1,001 \end{array}\right\} \tag{6-15}$$

where x is formed with 1's in the places where the divisor was added, and y is formed with 1's in the places where the divisor

TABLE 6-11

Operations			Step (Add next step if minus; subtract if plus)	Partial quotient
(a)				
Dividend (decimal 137)	10001001			
Divisor (decimal 10)	1010			
Difference	−010111		−	10000
(b)				
Divisor (shifted)	1010			
Algebraic Sum	111001		+	10000 − 1000 = 1000
(c)				
Divisor (shifted)	1010			
Difference	10001		+	1000 + 100 = 1100
(d)				
Divisor (shifted)	1010			
Difference	−11		−	1100 + 10 = 1110
(e)				
Divisor (shifted)	1010			
Algebraic Sum	111		+	1110 − 1 = 1101

was subtracted. The operation defined in Eq. (6-14) is brought about by shifting all digits one place to the left, adding a 0 at the end (multiplication by 2), deleting the initial 1 (subtraction of 2^n), and then replacing the final 0 by a 1 (addition of 1).

Thus, in the example, step by step,

$$x = 10,110 \to 101,100 \to 01,100 \to 01,101 \to 1,101 = z$$

The negative part of the quotient need never enter the calculation.

6-11. Alternate Arrangements—Linear Mappings

The system of complements described in Sec. 6-8 above is an efficient and convenient method for handling numbers which occur in calculations where the range of numbers extends approximately equally on both sides of zero. The advantages are twofold:

1. Registers with n radix r positions can be utilized fully without use of additional sign symbols.
2. The arithmetic steps are carried out digit by digit as they are in long division, long multiplication, addition, and subtraction with ordinary numbers; the behavior of each digit is independent both of the value of other digits and the sign of the numbers.

In the addition process, the system of arithmetic modulo $r^n - 1$ has the advantage that complements are easily formed and arithmetic can proceed with either an addition or a subtraction process. The first of the two properties of the complement system displayed above can be shown to be shared by linear mappings, and except for some permutations of digits which do not essentially change the results, these are the only mappings which have the advantage of permitting digit-by-digit calculation. The linear mapping is defined as follows:

Definition: *Let the number* x, $0 \leq x \leq r^n - 1$, *be an integer representing the ordinal number of any of* r^n *states assumed by an n-digit radix r register and let* $a(\neq 0)$ *and b be two constants. Then, under a linear mapping, the coefficients a and b and the state x represent the number y with*

$$y = ax + b \tag{6-16}$$

Under this definition it can be shown easily that multiplication of two numbers y_1 and y_2 can be brought about by operating, digit by digit, on the numbers x_1 and x_2 which represent them. Indeed, if

$$y_1 = ax_1 + b \tag{6-17}$$

and

$$y_2 = ax_2 + b \tag{6-18}$$

and if

$$y_1 y_2 = y_3 = ax_3 + b \tag{6-19}$$

then

$$x_3 = ax_1 x_2 + b(x_1 + x_2) + \frac{b^2 - b}{a} \tag{6-20}$$

With a and b constant, this expression can be built up under fixed rule, digit by digit, just as the product in ordinary arithmetic is built up. Similarly, the operations of addition, subtraction, and division may be carried out on a digit-by-digit basis.

Now, if the y values range between α and β so that $\alpha \leq y \leq \beta$, then it is possible to position the numbers a and b in such a way that α corresponds to the value $x = 0$ and β corresponds to the value $r^n - 1$. This possibility permits the use of the linear mapping system to give the registers the greatest possible efficiency. To obtain this result, it is only necessary to write

$$b = \alpha$$
$$a = \frac{\beta - \alpha}{r^n - 1} \tag{6-21}$$

Some simple examples of linear mappings are listed here. In the variant of binary arithmetic described in Sec. 6-10, the representation of the quotient z by the positive part of the quotient x was defined in Eq. (6-14); examination of this equation shows that it is a linear mapping with $a = 2$ and $b = -2^n + 1$. A linear mapping with piecewise constant coefficients can be used to define systems of complements to give arithmetic modulo r^n and modulo $r^n - 1$; in either case, for y nonnegative, the coefficients are $a = 1$ and $b = 0$; in the first case, for y negative, the coefficients are $a = +1$ and $b = -r^n$, and in the second case the coefficients are $a = +1$ and $b = 1 - r^n$.

In general, the full utilization of the range of y numbers appearing will not be the only consideration in choice of a and b. Among other factors which should be considered are ease of subtraction, elimination of round-off bias, and decrease of number of trial subtractions in division.

6-12. Note on Round-off Error

In this chapter the elements of the arithmetic theory useful in applying EPDCC's to computing problems have been out-

lined.   The chapter has shown how by proper choice of radix
an EPDCC with any number of stable states can be applied to a
computing problem.

It must be remembered, however, that many other details must
be taken into account.   Not only are engineering details involved
in connecting the components to carry out the calculation of
crucial importance, but also other mathematical details must be
considered.   In particular, the extremely fast calculating speed
of these machines introduces a new possibility of *round-off
accumulating*.   Although the machines function in the same
way that a desk machine functions, it is no longer feasible to
overcome the round-off error by sheer length of numbers.   The
way in which these errors accumulate might be justifiably
included in this chapter; however, the application of methods for
estimating the accumulation of round-off errors to problems of
numerical calculation is involved, and its exposition is likely to
be long and understood only by computers willing to devote
considerable time to the study.   A detailed treatment of round-off
error is given by Burks, Goldstine, and von Neumann[2] and by
von Neumann and Goldstine.[5]

6-13. Appendix

6-13-1. Proof of Theorem 6-1.   *The number of different
numbers which can be represented by configurations with marks
differing from 0 in at most the first k positions is the product*

$$P_k = n_1 n_2 n_3 \cdots n_k \tag{6-1}$$

Proof: It is a classical proposition that P_k different configura-
tions can be made with the k marks in order, and it is obvious
from the definition of counting (meaning exactly that it can be
proved by mathematical induction) that no two of these con-
figurations represent the same number.   This completes the
proof.

6-13-2. Proof of Theorem 6-2.   *In the representation of a
number x, if p_1 represents the number of marks admissible which
precede the mark actually occurring in the first position, if p_2 is
the number of marks admissible for the second position which
precede the mark in that position, etc., with p_i the number of admis-
sible marks for the ith position preceeding the ith mark, then,*

conforming with the notation for P of Eq. (6-1),

$$x = p_1 + p_2 P_1 + p_3 P_2 + \cdots + p_i P_{i-1} \qquad (6\text{-}2)$$

Proof: The proof of this theorem will be by mathematical induction. It will be shown that the set of values of x for which the statement is true includes the number zero and that if this set includes all nonnegative numbers less than a number N, it also includes N. All such sets clearly (or axiomatically) include all the nonnegative integers, and it will follow that the theorem is true—that the statement is true for every value of x.

First, it is clear that if $x = 0$, the statement is true, for then by definition all the numbers p_i are zero, and the right member of Eq. (6-2) is zero and therefore equal to x. Now, suppose that when x is equal to some particular number N, Eq. (6-2) holds. Then if the representation of the number $N + 1$ is written, one or more of the numbers p_i determined by this new number differ from the corresponding numbers in the representation of N. According to the counting rule, if p_k is the last one which differs, then it is greater by 1 in the representation of $N + 1$ than it was in the representation of N; furthermore, since all p_i with $i < k$ must have been maximal in the representation of N and since they have all been replaced by zero, each such p_i in the representation of $N + 1$ is $n_i - 1$ less than it was in the representation of N. Now, it was assumed that Eq. (6-2) was valid when x was N. When x is replaced by $N + 1$, it is clear that the left member is increased by 1, and it has just been remarked that the right member is increased by the quantity

$$Q_k = P_{k-1} - (n_{k-1} - 1)P_{k-2} - \cdots - (n_i - 1) \quad (6\text{-}22)$$

The theorem will have been established if it can be shown that Q_k is always 1 no matter what value k takes. [For example, for decimal numbers it is true that

$$Q_k = 10^{k-1} - 9(10^{k-2} + 10^{k-3} + \cdots + 1) \qquad (6\text{-}23)$$

and this expression is readily seen to be exactly 1 no matter what value k may have by the theorem giving the sum of terms in a geometric series.] The fact that Q_k is always 1 can be established rigorously by mathematical induction or more intuitively by noting that P_{k-1} is equal to $n_{k-1}P_{k-2}$ by definition and that the various terms of the right member of Eq. (6-22) annihilate one

another in succession until only the final number 1 is left.   This completes the proof of Theorem 6-2.

6-13-3. Proof of Theorem 6-3.   *If x is a nonnegative integer and if n_i, p_i, and P_i all have the significance defined above for some system of notation in which x is expressed, then p_1 is the remainder when x is divided by n_1, p_2 is the remainder when this integral quotient (without remainder) is divided by n_2, etc.: explicitly, there exists a set of nonnegative integers q_0, q_1, q_2, . . . , q_i such that*

$$
\begin{aligned}
q_0 = x = n_1 q_1 + p_1; &\qquad 0 = p_1 n_1 \\
q_1 = n_2 q_2 + p_2; &\qquad 0 = p_2 n_2 \\
\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot &\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\
q_i = n_{i+1} q_{i+1} + p_{i+1}; &\qquad 0 = p_{i+1} n_{i+1} \\
\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot &\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot
\end{aligned}
\qquad (6\text{-}3)
$$

These equations uniquely determine p_i.

Proof: It is clear that Eqs. (6-3) define the p's, for these equations are only a formal method of stating that, say, q_i is the quotient and p_i the remainder when q_{i-1} is divided by n_i.   With this in mind, the following inductive proof will make use of an expression which will be shown for every value of i to define q_i equivalent to the definition in Eq. (6-3) subject to the original restriction $q_0 = x$:

$$
q_i = p_{i+1} + p_{i+2} n_{i+1} + p_{i+3} n_{i+2} n_{i+1} + p_{i+4} n_{i+3} n_{i+2} n_{i+1} + \cdot \cdot \cdot ,
$$
$$
i = 0, 1, \cdot \cdot \cdot \quad (6\text{-}24)
$$

To establish this for $i = 0$, all that is required is to replace i by 0 throughout and to compare the result with Eq. (6-2), which has already been established; if the P's in Eq. (6-2) are written in terms of the n's, the two will be identical.   Next seek to determine whether the set of values of i, for which Eq. (6-24) gives the same value of q_i as Eqs. (6-3), does possess the property that the set contains the next higher integer to any integer it contains. Suppose that for $i = k$, the result from Eq. (6-24) is the same as that from Eqs. (6-3).   Then, for this value of k,

$$
q_k = p_{k+1} + p_{k+2} n_{k+1} + p_{k+3} n_{k+2} n_{k+1} + \cdot \cdot \cdot \qquad (6\text{-}25)
$$

Rewritten, this states

$$
q_k = n_{k+1}(p_{k+2} + p_{k+3} n_{k+2} + \cdot \cdot \cdot) + p_{k+1} \qquad (6\text{-}26)
$$

Now, according to Eqs. (6-3) with $i = k$, this should be

$$q_k = n_{k+1}q_{k+1} + p_{k+1} \tag{6-27}$$

and the expression in the parentheses, which therefore must be q_{k+1}, is indeed the value that Eq. (6-24) gives for q_{k+1} by setting $i = k + 1$. Thus, whenever Eq. (6-24) gives the right answer for any value of i (as it does for $i = 0$), it also gives the right value for the next higher value, and the statement of the theorem is true for these values. This completes a proof by mathematical induction, for the set of values of i for which the statement of the theorem is true must contain all positive integers if it is to fulfill these requirements.

REFERENCES

1. Alt, Franz, A Bell Telephone Laboratories' Computing Machine, *Mathematical Tables and Other Aids to Computation*, Vol. III, pp. 1–13, 69–84 (1948).
2. Burks, A. W., H. H. Goldstine, and John von Neumann, *Preliminary Discussion of the Logical Design of an Electronic Computing Instrument*, Part I, 2d ed. (Institute for Advanced Study, Princeton, N.J., 1947), pp. 23–29.
3. Carmichael, Robert D., *The Theory of Numbers*, (John Wiley & Sons, Inc., New York, 1914), pp. 22–24.
4. Morley, S. G., *An Introduction to the Study of the Mayan Hieroglyphs*, Bulletin 57, Bureau of American Ethnology (The Smithsonian Institution, Washington, D.C., 1915).
5. von Neumann, John, and H. H. Goldstine, Numerical Inverting of Matrices of High Order, *Bulletin of the American Mathematical Society*, Vol. 53, pp. 1021–1099 (1947).
6. Ritt, J. F., *Theory of Functions* (King's Crown Press, New York, 1947), pp. 1–10.
7. Spinden, H. J., *Ancient Civilizations of Mexico and Central America*, Handbook Series No. 3, 3d ed. (American Museum of Natural History, New York, 1928), pp. 118–121.

CHAPTER 7

NUMERICAL ANALYSIS

7-1. Introduction

In this chapter attention will be given to the numerical methods for solving certain of the principal types of equations dealing with algebraic, differential, and integral elements. As stated in the introductory chapter, these methods are regarded as components in the art of computation to the same degree as the physical equipment described in Chaps. 3, 4, 13, and 17. The material dealing with methods of computation for algebraic and differential equations is so extensive that the scope of this chapter will be limited to a brief outline and to the discussion of a few methods which are applicable to the principal types of equations to which a digital computing machine is expected to be applied. For the sake of simplicity, this chapter is organized to treat first with the process of interpolation and interpolation formulas since this procedure forms the basis of many numerical methods. There will then follow brief outlines of methods for numerical differentiation and integration which depend largely upon these interpolation formulas. The numerical solution of algebraic and transcendental equations will next be taken up. Finally a discussion will be given concerning the procedures for numerical solution of total differential equations.

The material for this chapter has been drawn largely from such standard texts as Ford,[7] von Kármán and Biot,[12] Margenau and Murphy,[15] Scarborough,[19] and Whittaker and Robinson,[20] and from such special reports as Bennett, Milne, and Bateman[3] and Bargmann, Montgomery, and von Neumann.[2]

7-2. The Process of Interpolation

The process of interpolation consists essentially in representing a function in a particular form by making use of values which this function assumes for certain given values of the argument, or the independent variable. For example, let $y = f(x)$ take on values

y_0, y_1, y_2, . . . , y_n for respective values of the argument x_0, x_1, x_2, . . . , x_n. Now let $\phi(x)$ represent some arbitrary and, perhaps, simpler function such that it takes on the values of $f(x)$ for the same values of the argument x_0, x_1, x_2, . . . , x_n. If $f(x)$ is now replaced by $\phi(x)$ over the interval of the independent variable, then this procedure is known as *interpolation*, and the function $\phi(x)$ is called the *formula of interpolation*. The mathematical basis for this procedure stems from two powerful theorems by Weierstrass. These theorems may be stated as follows:

1. Every function continuous in the interval (a,b) may be represented in that interval, to any degree of accuracy, by a polynomial $P(x)$ such that the absolute value of the difference between the function and this polynomial may be made smaller than any previously assigned positive quantity ϵ.
2. Every continuous function periodic in 2π may be represented by a finite trigonometric series

$$F(x) = a_0 + a_1 \sin x + a_2 \sin 2x + \cdots + a_n \sin nx \\ + b_1 \cos x + b_2 \cos 2x + \cdots + b_n \cos nx$$

where the absolute value of the difference of the function and this trigonometric series may be made smaller than any previously assigned positive quantity δ.

Let us consider a function $y = f(x)$ for which 10 values of the function are known for 10 values of the independent variable. It will now be instructive to form what is known as a *difference table* for these values of the function and the independent variable. We arrange the values of the independent variable and of the function in the first two columns, respectively, and proceed to take differences between these values of the function, where $\Delta y_0 = y_1 - y_0$, $\Delta y_1 = y_2 - y_1$, . . . , $\Delta y_n = y_{n+1} - y_n$; these are denoted as *first* differences and appear in the third column. Second, third, fourth, and higher differences up to ninth differences may be obtained in the same manner and give rise to the elements appearing in the diagonal difference table (Table 7-1).

Such an arrangement is important because it may be shown (1) that if the values of the independent variable are taken in arithmetic progression, the nth differences of a polynomial of the nth degree are constant, and (2) that if the nth differences of a function are constant for an arithmetic progression of the

argument of the independent variable, this function is a polynomial of degree n. The latter means, essentially, that if an unknown function $y = f(x)$ is tabulated in a difference table and if the fourth differences, for example, are found to be constant, then such a function may be approximated by a fourth-degree polynomial $\phi(x)$ within the region of the limits of data given for the independent variable.

TABLE 7-1

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$	$\Delta^6 y$	$\Delta^7 y$	$\Delta^8 y$	$\Delta^9 y$
x_0	y_0									
		Δy_0								
x_1	y_1		$\Delta^2 y_0$							
		Δy_1		$\Delta^3 y_0$						
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$					
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$				
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$		$\Delta^6 y_0$			
		Δy_3		$\Delta^3 y_2$		$\Delta^5 y_1$		$\Delta^7 y_0$		
x_4	y_4		$\Delta^2 y_3$		$\Delta^4 y_2$		$\Delta^6 y_1$		$\Delta^8 y_0$	
		Δy_4		$\Delta^3 y_3$		$\Delta^5 y_2$		$\Delta^7 y_1$		$\Delta^9 y_0$
x_5	y_5		$\Delta^2 y_4$		$\Delta^4 y_3$		$\Delta^6 y_2$		$\Delta^8 y_1$	
		Δy_5		$\Delta^3 y_4$		$\Delta^5 y_3$		$\Delta^7 y_2$		
x_6	y_6		$\Delta^2 y_5$		$\Delta^4 y_4$		$\Delta^6 y_3$			
		Δy_6		$\Delta^3 y_5$		$\Delta^5 y_4$				
x_7	y_7		$\Delta^2 y_6$		$\Delta^4 y_5$					
		Δy_7		$\Delta^3 y_6$						
x_8	y_8		$\Delta^2 y_7$							
		Δy_8								
x_9	y_9									

where
$$\Delta y_0 = y_1 - y_0$$
$$\Delta y_1 = y_2 - y_1$$
$$\Delta y_2 = y_3 - y_2$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$\Delta y_n = y_{n+1} - y_n,$$

and
$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = y_2 - 2y_1 + y_0$$
$$\Delta^2 y_1 = \Delta y_2 - \Delta y_1 = y_3 - 2y_2 + y_1$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$\Delta^2 y_n = \Delta y_{n+1} - \Delta y_n = y_{n+2} - 2y_{n+1} + y_n$$

The application of machines for tabulating differences of a function has recently been described by Laderman and Abramowitz.[13] The Underwood-Elliott Fisher Sundstrand Accounting Machine, Model D, has been found exceptionally efficient for the construction of difference tables and for computing and

printing the values of a function when initial differences and values of the function are given.

7-2-1. Newton's Interpolation Formula. It will be instructive here to show how a polynomial of degree n may be made to approximate a given function $f(x)$ in the interval defined by the argument. There are a number of interpolation formulas which may be used to define a function $\phi(x)$ as an approximation to a given function $f(x)$. Since the derivation of all these formulas is substantially similar, however, Newton's formula for forward interpolation will be used here as an example. Several remarks will be made concerning the application of this formula and several other related formulas of interpolation.

Let the function $y = f(x)$ take on values $y_0, y_1, y_2, \ldots, y_n$ for respective values of the independent variable $x_0, x_1, x_2, \ldots, x_n$. We are now interested in finding that polynomial $\phi(x)$ which is required to represent $f(x)$ in the interval determined by the range of the independent variable, whose successive values are taken as equidistant. The polynomial $\phi(x)$ of the nth degree may be written as follows:

$$\phi(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1)$$
$$+ a_3(x - x_0)(x - x_1)(x - x_2)$$
$$+ a_4(x - x_0)(x - x_1)(x - x_2)(x - x_3)$$
$$+ \cdots + a_n(x - x_0)(x - x_1)(x - x_2) \cdots (x - x_{n-1}) \quad (7\text{-}1)$$

In order to make $\phi(x_0) = y_0$, $\phi(x_1) = y_1$, $\ldots$, $\phi(x_n) = y_n$, it is necessary to determine the coefficients $a_0, a_1, a_2, \ldots, a_n$. This may be done by substituting the successive values $x_0, x_1, x_2, \ldots, x_n$ into the relation for $\phi(x)$ above, with the added requirement that the value of the arithmetic progression in the independent variable be taken as h such that $x_1 - x_0 = h$, $x_2 - x_0 = 2h, \ldots, x_n - x_0 = nh$. With these substitutions the polynomial $\phi(x)$, known as *Newton's formula for forward interpolation*, may be derived; this is written as follows:

$$\phi_I(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2h^2}(x - x_0)(x - x_1)$$
$$+ \frac{\Delta^3 y_0}{3!h^3}(x - x_0)(x - x_1)(x - x_2)$$
$$+ \frac{\Delta^4 y_0}{4!h^4}(x - x_0)(x - x_1)(x - x_2)(x - x_3) + \cdots$$
$$+ \frac{\Delta^n y_0}{n!h^n}(x - x_0)(x - x_1)(x - x_2) \cdots (x - x_{n-1}) \quad (7\text{-}2)$$

This may be simplified by a change of variable such that

$$\frac{x - x_0}{h} = u$$

From this relation it follows that $(x - x_1)/h = u - 1$, etc.; the generating relationship[19] is then seen to be

$$\frac{x - x_n}{h} = u - n + 1$$

Newton's formula for forward interpolation with this change of variable takes the form

$$\phi_{\mathrm{I}}(x) = y_0 + u \, \Delta y_0 + \frac{u(u - 1)}{2!} \, \Delta^2 y_0 + \frac{u(u - 1)(u - 2)}{3!} \, \Delta^3 y_0$$
$$+ \frac{u(u - 1)(u - 2)(u - 3)}{4!} \, \Delta^4 y_0$$
$$+ \cdots + \frac{u(u - 1)(u - 2) \cdots (u - n + 1)}{n!} \, \Delta^n y_0 \quad (7\text{-}3)$$

The forward-interpolation formula of Newton is so named because this formula treats values of the tabulated function starting with y_0 and proceeding forward to y_n; it does not allow interpolation below or to the left of y_0. For this purpose there exists another form of this formula known as *Newton's formula for backward interpolation*:

$$\phi_{\mathrm{II}}(x) = y_n + u \, \Delta_1 y_n + \frac{u(u + 1)}{2} \, \Delta_2 y_n$$
$$+ \frac{u(u + 1)(u + 2)}{3!} \, \Delta_3 y_n + \frac{u(u + 1)(u + 2)(u + 3)}{4!} \, \Delta_4 y_n$$
$$+ \cdots + \frac{u(u + 1)(u + 2) \cdots (u + n - 1)}{n!} \, \Delta_n y_n \quad (7\text{-}4)$$

where
$$\Delta_1 y_n = y_n - y_{n-1}$$
$$\Delta_2 y_n = y_n - 2y_{n-1} + y_{n-2}$$
$$\cdots \cdots \cdots \cdots \cdots \cdots \cdots$$

7-2-2. Stirling's Interpolation Formula. In the previous section, Newton's formula for forward interpolation $\phi_{\mathrm{I}}(x)$ was related to the array of differences in the diagonal difference table (Table 7-1). It was seen there that this interpolation formula was applicable only to determining the difference of values of the function $y = f(x)$ for a forward progression $y_0, y_1, y_2, \ldots, y_n$. Similarly the Newton formula for backward interpolation allows

differences to be determined for the progression y_n, y_{n-1}, y_{n-2}, . . . , y_1. These formulas are fundamental to the interpolation process; in fact, other interpolation formulas for functions with equidistant increments of the argument may be derived directly from Newton's formulas by simple algebraic transformations. The only probable disadvantage in the use of Newton's formulas is that they quite often do not converge rapidly enough for certain practical applications. Because they obviate this difficulty, interpolation formulas due to Stirling and to Bessel, based on a slight variation of the original tabulation of differences, are important; these may be applied in cases where Newton's formulas do not provide rapid enough convergence.

Let us return to a function $y = f(x)$ with 10 values of the function given for 10 values of the argument. We now arrange these values of the function in a table where the designation of the y_0 position lies at or near the middle of the first column. This is known as a central-difference table (Table 7-2).

TABLE 7-2

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$	$\Delta^6 y$	$\Delta^7 y$	$\Delta^8 y$	$\Delta^9 y$
x_{-4}	y_{-4}									
		Δy_{-4}								
x_{-3}	y_{-3}		$\Delta^2 y_{-4}$							
		Δy_{-3}		$\Delta^3 y_{-4}$						
x_{-2}	y_{-2}		$\Delta^2 y_{-3}$		$\Delta^4 y_{-4}$					
		Δy_{-2}		$\Delta^3 y_{-3}$		$\Delta^5 y_{-4}$				
x_{-1}	y_{-1}		$\Delta^2 y_{-2}$		$\Delta^4 y_{-3}$		$\Delta^6 y_{-4}$			
		Δy_{-1}		$\Delta^3 y_{-2}$		$\Delta^5 y_{-3}$		$\Delta^7 y_{-4}$		
x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$		$\Delta^6 y_{-3}$		$\Delta^8 y_{-4}$	
		Δy_0		$\Delta^3 y_{-1}$		$\Delta^5 y_{-2}$		$\Delta^7 y_{-3}$		$\Delta^9 y_{-4}$
x_1	y_1		$\Delta^2 y_0$		$\Delta^4 y_{-1}$		$\Delta^6 y_{-2}$		$\Delta^8 y_{-3}$	
		Δy_1		$\Delta^3 y_0$		$\Delta^5 y_{-1}$		$\Delta^7 y_{-2}$		
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$		$\Delta^6 y_{-1}$			
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$				
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$					
		Δy_3		$\Delta^3 y_2$						
x_4	y_4		$\Delta^2 y_3$							
		Δy_4								
x_5	y_5									

In this table only the differences lying across the table immediately above and below y_0 (marked by broken lines) are used in

the interpolation formulas of both Stirling and Bessel. Both Stirling's and Bessel's formulas may be derived from a central-difference table; it is found slightly easier, however, to obtain these two formulas by an algebraic transformation of the coefficients of differences appearing in Newton's formula for forward interpolation [Eq. (7-3)]. It is seen that the coefficients in u of the various orders of differences in y_0 are actually the binomial coefficients.[19] By algebraic transformation these coefficients in u may be replaced by the central differences to which we have previously referred. By this procedure, we obtain *Stirling's formula for interpolation:*

$$
\begin{aligned}
y = y_0 &+ u\,\frac{\Delta y_{-1} + \Delta y_0}{2} + \frac{u^2}{2}\Delta^2 y_{-1} + \frac{u(u^2 - 1^2)}{3!}\,\frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} \\
&+ \frac{u^2(u^2 - 1^2)}{4!}\,\Delta^4 y_{-2} + \frac{u(u^2 - 1^2)(u^2 - 2^2)}{5!}\,\frac{\Delta^5 y_{-3} + \Delta^5 y_{-2}}{2} \\
&+ \frac{u(u - 1)(u - 2)}{6!}\,\Delta^6 y_{-3} + \cdots \\
&+ \frac{u(u^2 - 1^2)(u^2 - 2^2)(u^2 - 3^2)\,\cdots\,[u^2 - (n - 1)^2]}{(2n - 1)!} \\
&\qquad\qquad\qquad\qquad \times \frac{\Delta^{2n-1} y_{-n} + \Delta^{2n-1} y_{-(n-1)}}{2} \\
&+ \frac{u^2(u^2 - 1^2)(u^2 - 2^2)(u^2 - 3^2)\,\cdots\,[u^2 - (n - 1)^2]}{(2n)!} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \times \Delta^{2n} y_{-n} \quad (7\text{-}5)
\end{aligned}
$$

7-2-3. Bessel's Interpolation Formula. By a similar procedure, using another substitution for the binomial coefficients in u appearing in Newton's formula for forward interpolation, we obtain *Bessel's formula for interpolation:*

$$
\begin{aligned}
y = \frac{y_0 + y_1}{2} &+ \left(u - \frac{1}{2}\right)\Delta y_0 + \frac{u(u - 1)}{2}\,\frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} \\
&+ \frac{(u - \frac{1}{2})u(u - 1)}{3!}\,\Delta^3 y_{-1} + \frac{u(u - 1)(u + 1)(u - 2)}{4!} \\
&\qquad\qquad\qquad\qquad\qquad \times \frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} \\
&+ \frac{(u - \frac{1}{2})u(u - 1)(u + 1)(u - 2)}{5!}\,\Delta^5 y_{-2} \\
&+ \frac{u(u - 1)(u + 1)(u - 2)(u + 2)(u - 3)}{6!} \\
&\qquad\qquad\qquad\qquad\qquad \times \frac{\Delta^6 y_{-3} + \Delta^6 y_{-2}}{2} + \cdots
\end{aligned}
$$

$$+ \frac{u(u-1)(u+1)(u-2)(u+2) \cdots (u-n)(u+n-1)}{(2n)!}$$

$$\times \frac{\Delta^{2n}y_{-n} + \Delta^{2n}y_{-n+1}}{2}$$

$$+ \frac{(u - \frac{1}{2})u(u-1)(u+1)(u-2)(u+2) \cdots (u-n)(u+n-1)}{(2n+1)!}$$

$$\times \Delta^{2n+1}y_{-n} \quad (7\text{-}6)$$

7-2-4. Lagrange's Formula of Interpolation. The formulas of Newton, Stirling, and Bessel apply only when equidistant values of the argument are given. Lagrange, however, developed an interpolation formula which is applicable to a function for which the values of the argument are not equidistant.

Consider the function $y = f(x)$ for which it is desired to obtain an interpolation polynomial $\phi(x)$ of the nth degree. If this nth-degree polynomial is written as

$$\phi(x) = a_0(x-x_1)(x-x_2)(x-x_3) \cdots (x-x_n)$$
$$+ a_1(x-x_0)(x-x_2)(x-x_3) \cdots (x-x_n)$$
$$+ a_2(x-x_0)(x-x_1)(x-x_3) \cdots (x-x_n) + \cdots$$
$$+ a_n(x-x_0)(x-x_1)(x-x_2) \cdots (x-x_{n-1}) \quad (7\text{-}7)$$

the $n + 1$ constants $a_1, a_2, \ldots, a_n$ must now be determined in order that $\phi(x_0) = y_0$, $\phi(x_1) = y_1$, $\ldots$, $\phi(x_n) = y_n$. By taking $x = x_0$ and $\phi(x_0) = y_0$, a_0 may be evaluated. By a similar procedure using x_1 and $\phi(x_1) = y_1$, a_1 may be evaluated, etc., up to and including the evaluation of a_n. By replacing the values for the $n + 1$ constants by their evaluated operations, *Lagrange's formula of interpolation* is obtained:

$$\phi(x) = \frac{(x-x_1)(x-x_2) \cdots (x-x_n)}{(x_0-x_1)(x_0-x_2) \cdots (x_0-x_n)} y_0$$
$$+ \frac{(x-x_0)(x-x_2) \cdots (x-x_n)}{(x_1-x_0)(x_1-x_2) \cdots (x_1-x_n)} y_1$$
$$+ \frac{(x-x_0)(x-x_1)(x-x_3) \cdots (x-x_n)}{(x_2-x_0)(x_2-x_1)(x_2-x_3) \cdots (x_2-x_n)} y_2$$
$$+ \cdots + \frac{(x-x_0)(x-x_1) \cdots (x-x_{n-1})}{(x_n-x_0)(x_n-x_1) \cdots (x_n-x_{n-1})} y_n \quad (7\text{-}8)$$

Since Eq. (7-8) is a functional relation between the variable y and the variable x, it is allowable to interchange the roles of the dependent and independent variables and obtain a functional

relation between y and the various given values of x as follows:

$$
\begin{aligned}
\theta(y) = {} & \frac{(y - y_1)(y - y_2) \, \cdots \, (y - y_n)}{(y_0 - y_1)(y_0 - y_2) \, \cdots \, (y_0 - y_n)} \, x_0 \\
& + \frac{(y - y_0)(y - y_2) \, \cdots \, (y - y_n)}{(y_1 - y_0)(y_1 - y_2) \, \cdots \, (y_1 - y_n)} \, x_1 \\
& + \frac{(y - y_0)(y - y_1)(y - y_3) \, \cdots \, (y - y_n)}{(y_2 - y_0)(y_2 - y_1)(y_2 - y_3) \, \cdots \, (y_2 - y_n)} \, x_2 \\
& + \, \cdots \, + \frac{(y - y_0)(y - y_1) \, \cdots \, (y - y_{n-1})}{(y_n - y_0)(y_n - y_1) \, \cdots \, (y_n - y_{n-1})} \, x_n \quad (7\text{-}9)
\end{aligned}
$$

This relationship is important in determining the value of the independent variable which corresponds to a given value of the function. Such a procedure is known as *inverse interpolation*.

7-2-5. Aitken's Method of Interpolation.[1] A practical method for finding a numerical value of $f(x)$, for a given value of x, when several values of x and $f(x)$ are known, is Aitken's process of iteration. This method is well adapted to computing machinery. It consists of an iteration of the familiar process of linear interpolation. The basis for the following description was drawn largely from Aitken,[1] Feller,[6] and Milne.[16]

Eliminating all terms above the first degree in Newton's formula, Eq. (7-2), we are left with the expression for linear interpolation, or interpolation by divided differences. Rewriting this expression in determinantal form, for the sake of convenience in computing, and substituting $y_1 - y_0$ for Δy_0 and $x_1 - x_0$ for h, we have

$$
\phi_1^{(1)}(x) = \frac{\begin{vmatrix} y_0 & x_0 - x \\ y_1 & x_1 - x \end{vmatrix}}{x_1 - x_0} \quad (7\text{-}10)
$$

The superscript (1) indicates that the function was determined from the first iterate, and the subscript 1 indicates that it was formed on the difference $y_k - y_0$ for $k = 1$. This process of interpolation by proportional difference is repeated to obtain $\phi_2^{(1)}, \phi_3^{(1)}, \ldots, \phi_n^{(1)}$, where

$$
\phi_k^{(1)} = \frac{\begin{vmatrix} y_0 & x_0 - x \\ y_k & x_k - x \end{vmatrix}}{x_k - x_0} \quad (7\text{-}11)
$$

It may be noted that, for $x = x_k$, $y = f(x) = \phi_k^{(1)}$. In the next iteration, $\phi^{(1)}$ is substituted for y in Eq. (7-11). That is, the same operation performed on the set y_k for $k = 1, \ldots, n$ is iterated on the set $\phi_k^{(1)}$ for $k = 2, \ldots, n$, to obtain

$$\phi_k^{(2)} = \frac{\begin{vmatrix} \phi_1^{(1)} & x_1 - x \\ \phi_k^{(1)} & x_k - x \end{vmatrix}}{x_k - x_1} \tag{7-12}$$

The process of iteration is continued until the computed values of the interpolation function $\phi_k^{(n)}$ agree to the desired number of places.

In using this method to compute specific numerical values, a table of the following type is set up:

x_0	y_0	$\cdots$	$\cdots$	$\cdots$	$x_0 - x$
x_1	y_1	$\phi_1^{(1)}$	$\cdots$	$\cdots$	$x_1 - x$
x_2	y_2	$\phi_2^{(1)}$	$\phi_2^{(2)}$	$\cdots$	$x_2 - x$
x_3	y_3	$\phi_3^{(1)}$	$\phi_3^{(2)}$	$\phi_3^{(3)}$	$x_3 - x$

In practice it is probably most easy to form all the interpolating polynomials of successively higher degrees by systematically following the rules illustrated in Eqs. (7-11) and (7-12). It is evident, however, that the method can be described more generally by stating that each value of $\phi^{(n)}$ is obtained by linear interpolation applied to any two different interpolating polynomials $\phi^{(n-1)}$.

7-2-6. Hermite's Formula for Interpolating Periodic Functions. If a function $y = f(x)$ is periodic in the interval 2π, it is appropriate to use trigonometric interpolation. A formula derived by Hermite, and also independently by Gauss, is applicable to this condition. Similarity to Lagrange's formula, Eq. (7-8), for nonequidistant values of the argument of a nonperiodic function is quite apparent.

$$
\begin{aligned}
y = {} & \frac{\sin (x - x_1) \sin (x - x_2) \cdots \sin (x - x_n)}{\sin (x_0 - x_1) \sin (x_0 - x_2) \cdots \sin (x_0 - x_n)} \, y_0 \\
& + \frac{\sin (x - x_0) \sin (x - x_2) \cdots \sin (x - x_n)}{\sin (x_1 - x_0) \sin (x_1 - x_2) \cdots \sin (x_1 - x_n)} \, y_1 + \cdots \\
& + \frac{\sin (x - x_0) \sin (x - x_1) \cdots \sin (x - x_{n-1})}{\sin (x_n - x_0) \sin (x_n - x_1) \cdots \sin (x_n - x_{n-1})} \, y_n
\end{aligned} \tag{7-13}
$$

In common also with Lagrange's formula, Hermite's trigonometric interpolation formula may be transposed, by interchanging x and y, for inverse interpolation of periodic functions.

7-3. Numerical Differentiation

The process for obtaining the numerical value of a derivative of a given order is connected closely with the use of the interpola-

tion formulas employing differences such as those which we have briefly outlined in Sec. 7-2.   The obvious restrictions also apply in the matter of equidistant values of the argument, wherein the formulas of Newton, Bessel, and Stirling may be applied, and in values of the argument which are not equidistant, wherein the function is represented by formulas of Lagrange for non-periodic functions and of Hermite for periodic functions.

The numerical values of derivatives are determined through the use of differences and interpolation formulas based on these differences.   Taking Newton's formula for forward interpolation as an example:

$$
\begin{aligned}
y = \theta(u) = y_0 + u\,\Delta y_0 &+ \frac{u(u-1)}{2!}\Delta^2 y_0 + \frac{u(u-1)(u-2)}{3!}\,\Delta^3 y_0 \\
&+ \frac{u(u-1)(u-2)(u-3)}{4!}\,\Delta^4 y_0 + \cdots \\
&+ \frac{u(u-1)(u-2)\,\cdots\,(u_n - n + 1)}{n!}\,\Delta^n y_0 \quad (7\text{-}14)
\end{aligned}
$$

Since $u = (x - x_0)/h$, it follows that

$$
\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{1}{h}\frac{dy}{dx}
$$

The values of several of the derivatives of y with respect to x are given below in terms of differences appearing in Newton's formula of interpolation.

$$
\begin{aligned}
\frac{dy}{dx} = \frac{1}{h}\bigg(\Delta y_0 &+ \frac{2u-1}{2!}\,\Delta^2 y_0 + \frac{3u^2 - 6u + 2}{3!}\,\Delta^3 y_0 \\
&+ \frac{4u^3 - 18u^2 + 22u - 6}{4!}\,\Delta^4 y_0 \\
&+ \frac{5u^4 - 40u^3 + 105u^2 - 100u + 24}{5!}\,\Delta^5 y_0 + \cdots \bigg) \quad (7\text{-}15)
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^2 y}{dx^2} = \frac{1}{h^2}\bigg[\Delta^2 y_0 &+ (u-1)\,\Delta^3 y_0 + \frac{12u^2 - 36u + 22}{4!}\,\Delta^4 y_0 \\
&+ \frac{20u^3 - 120u^2 + 210u - 100}{5!}\,\Delta^5 y_0 + \cdots \bigg] \quad (7\text{-}16)
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^3 y}{dx^3} = \frac{1}{h^3}\bigg(\Delta^3 y_0 &+ \frac{24u - 36}{4!}\,\Delta^4 y_0 \\
&+ \frac{60u^2 - 240u + 210}{5!}\,\Delta^5 y_0 + \cdots \bigg) \quad (7\text{-}17)
\end{aligned}
$$

$$\frac{d^4y}{dx^4} = \frac{1}{h^4} \ [\Delta^4 y_0 + (u - 2) \ \Delta^5 y_0 + \cdot \ \cdot \ \cdot] \tag{7-18}$$

$$\frac{d^5y}{dx^5} = \frac{1}{h^5} \ [\Delta^5 y_0 + \cdot \ \cdot \ \cdot] \tag{7-19}$$

It is seen from the method of determining a numerical value for the derivatives of a function that this process is capable of being handled by machine operation where tabular values of a function are given.

7-4. Numerical Integration

This process allows the numerical evaluation of a definite integral. Again here, as in differentiation, the method of determining a numerical value for a definite integral depends upon interpolation formulas discussed in Sec. 7-2. Since numerical integration is based principally on differences, it is capable of being performed by machine where the input data are values of the function and the function differences. In formulas for interpolation of functions with equidistant values of the argument, the increment by which the argument is increased successively was designated by h where $x = x_0 + hu$. In differential form $dx = h \ du$. For n values or intervals of the argument, x_0 is increased by an amount nh. If now we integrate Newton's formula for forward interpolation, Eq. (7-3), between the limits x_0 and $x_0 + nh$ the corresponding limits for u will be 0 and n. The integral of $y \ dx$ may then be expressed in terms of the original differences of the tabulated function and n, the number of equidistant increments in the argument. This results in the following expression for this integral:

$$\begin{aligned}
\int_{x_0}^{x_0+nh} y \ dx = h \Bigg[& ny_0 + \frac{n^2}{2} \Delta y_0 + \left(\frac{n^3}{3} - \frac{n^2}{2} \right) \frac{\Delta^2 y_0}{2!} \\
& + \left(\frac{n^4}{4} - n^3 + n^2 \right) \frac{\Delta^3 y_0}{3!} + \left(\frac{n^5}{5} - \frac{3n^4}{2} + \frac{11n^3}{3} - 3n^2 \right) \frac{\Delta^4 y_0}{4!} \\
& + \left(\frac{n^6}{6} - 2n^5 + \frac{35n^4}{4} - \frac{50n^3}{3} + 12n^2 \right) \frac{\Delta^5 y_0}{5!} \\
& + \left(\frac{n^7}{7} - \frac{15n^6}{6} + 17n^5 - \frac{225n^4}{4} + \frac{274n^3}{3} - 60n^2 \right) \frac{\Delta^6 y_0}{6!} \Bigg] \quad (7\text{-}20)
\end{aligned}$$

This is a general formula from which a number of well-known special formulas for integration may be obtained.[19] It may also

be written in terms of the sums of values of y_n of the function; in this form it is known as *the Newton-Cotes formula.*[15]

Several so-called *rules* derived from the integral result, by allowing n to take on various values such as 1, 2, 3, and 6, are listed below.

7-4-1. The Trapezoidal Rule. If we allow n to take on the value 1 and we neglect all *differences* above the first, the following expression is obtained for the integral of $y\,dx$:

$$\int_{x_0}^{x_0+nh} y\,dx = h\left(\frac{y_0}{2} + y_1 + y_2 + y_3 + \cdots + \frac{y_n}{2}\right) \quad (7\text{-}21)$$

These are just the first terms in the *Euler-Maclaurin* formula.[15]

7-4-2. Simpson's One-third Rule. By putting $n = 2$ and neglecting all *differences* above the second, a very important formula known as Simpson's one-third rule is obtained:

$$\int_{x_0}^{x_0+nh} y\,dx = \frac{h}{3}\,[y_0 + 4(y_1 + y_3 + \cdots + y_{n-1})$$
$$+ 2(y_2 + y_4 + \cdots + y_{n-2}) + y_n] \quad (7\text{-}22)$$

7-4-3. Simpson's Three-eighths Rule. With $n = 3$, and by neglecting all differences above the third, an expression for this interval known as Simpson's three-eighths rule is obtained:

$$\int_{x_0}^{x_0+nh} y\,dx = \frac{3h}{8}\,[y_0 + 3(y_1 + y_2 + y_4 + y_5 + y_7 + y_8$$
$$+ \cdots + y_{n-1}) + 2(y_3 + y_6 + y_9 + \cdots + y_{n-3}) + y_n]$$
$$(7\text{-}23)$$

7-4-4. Weddle's Rule. With $n = 6$ and neglecting all differences above the sixth, we obtain another important formula, probably the most accurate of these four, known as Weddle's rule:

$$\int_{x_0}^{x_0+nh} y\,dx = \frac{3h}{10}\,(y_0 + 5y_1 + y_2 + 6y_3 + y_4 + 5y_5 + 2y_6$$
$$+ 5y_7 + y_8 + 6y_9 + y_{10} + 5y_{11} + 2y_{12} + 2y_{n-6} + 5y_{n-5}$$
$$+ y_{n-4} + 6y_{n-3} + y_{n-2}\ 5y_{n-1} + y_n) \quad (7\text{-}24)$$

It is important to note here that n must be *six or a multiple of six.*

Other methods for obtaining numerical values of definite integrals make use of the Stirling and Bessel interpolation

formulas based on central differences.[19] Since these are similar in character to the ones which have been sketched above, it does not appear appropriate to include them in this chapter. Because of its special character, however, another method for determining the numerical value of an integral, due to Gauss, will be discussed briefly here.

7-4-5. The Method of Gauss. All the previously outlined methods for determining a numerical value for a definite integral are based on equidistant values of the argument. Where these values are not equidistant, however, the determination of a definite value of a numerical integral is best performed by using a formula developed by Gauss.[15] In the development of this formula, Gauss solved the problem of determining, for a given number of values of $f(x)$, how the values of the argument should be distributed in the interval (a,b) so as to give the greatest possible accuracy. The result obtained by Gauss indicated that these values should not be equidistant but should be distributed symmetrically with respect to the mid-point of the interval of integration (a,b). The resulting formula of Gauss may be written

$$I = \int_a^b f(x)\, dx = (b - a)[R_1\phi(u_1) + R_2\phi(u_2) + R_3\phi(u_3) \\ + R_4\phi(u_4) + \cdots + R_n\phi(u_n)] \quad \text{(7-25)}$$

The values of $u_1, u_2, \ldots, u_n$ represent the subdivisions of the interval $u = 0$ to $u = 1$. The values of x corresponding to each of the u's may be obtained from these relations:

$$\begin{aligned} x_1 &= a + (b - a)u_1 \\ x_2 &= a + (b - a)u_2 \\ &\cdots\cdots\cdots\cdots\cdots \\ x_n &= a + (b - a)u_n \end{aligned} \quad \text{(7-26)}$$

Both the quantities u and R must be determined numerically in order that this equation may be applied to a definite problem. The Gauss method gives an *exact* result where the order of the function is less than or equals $2n - 1$. Numerical values of the u's and the R's for $n = 3, 4, 5, 6, 7$ are given in Table 7-3.

7-4-6. An Example Using Three Methods of Numerical Integration. As an example of the application of (1) Simpson's one-third rule, (2) Weddle's rule, and (3) the method of Gauss,

the integral

$$\int_{1.4}^{2.0} \frac{x^6}{1+x^2}\, dx = \int_a^b y\, dx$$

will be evaluated numerically.

Application of Simpson's Rule. Both Simpson's and Weddle's rules demand that a tabulation of the value of $f(x) = x^6/(1+x^2)$ be made. The interval (a,b) has been divided into six equidistant increments of the argument x. Table 7-4 results.

TABLE 7-3.   QUADRATURE METHOD OF GAUSS FOR NUMERICAL INTEGRATION

	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
u_1	0.1127016654	0.0694318442	0.04691007703	0.03376524290	0.02544604383
u_2	0.5	0.3300094782	0.2307653449	0.1693953068	0.1292344072
u_3	0.8872983346	0.6699905218	0.5	0.3806904070	0.2970774243
u_4	. . . .	0.9305681558	0.7692346551	0.6193095930	0.5
u_5	. . . .	. . . .	0.9530899230	0.8306046932	0.7029225757
u_6	. . . .	. . . .	. . . .	0.9662347571	0.8707655928
u_7	. . . .	. . . .	. . . .	. . . .	0.9745539562
R_1	0.2777777777	0.1739274226	0.1184634425	0.0856622462	0.06474248308
R_2	0.4444444444	0.3260725774	0.2393143352	0.1803807865	0.1398526957
R_3	0.2777777777	0.3260725774	0.2844444444	0.2339569673	0.1909150253
R_4	. . . .	0.1739274226	0.2393143352	0.2339569673	0.2089795918
R_5	. . . .	. . . .	0.1184634425	0.1803807865	0.1909150253
R_6	. . . .	. . . .	. . . .	0.0856622462	0.1398526957
R_7	. . . .	. . . .	. . . .	. . . .	0.06474248308

TABLE 7-4

	x	x^2	$1 + x^2$	x^6	$y = f(x)$
x_0	1.4	1.96	2.96	7.5295360	$y_0 =$ 2.5437622
x_1	1.5	2.25	3.25	11.390625	$y_1 =$ 3.5048077
x_2	1.6	2.56	3.56	16.777216	$y_2 =$ 4.7127011
x_3	1.7	2.89	3.89	24.137569	$y_3 =$ 6.2050306
x_4	1.8	3.24	4.24	34.012224	$y_4 =$ 8.0217509
x_5	1.9	3.61	4.61	47.045881	$y_5 =$ 10.2051803
x_6	2.0	4.00	5.00	64.000000	$y_6 =$ 12.8000000

$$
\begin{aligned}
I_S = \int_{1.4}^{2.0} \frac{x^6}{1+x^2}\, dx &= \frac{0.1}{3}\,[2.5437622 + 4(3.5048077 \\
&\quad + 6.2050306 + 10.2051803) \\
&\quad + 2(4.7127011 + 8.0217509) + 12.8000000] \\
&= 4.0157580 \hspace{2cm} (7\text{-}27)
\end{aligned}
$$

Application of Weddle's Rule

$$I_W = \int_{1.4}^{2.0} \frac{x^6}{1+x^2} \, dx = 0.03(2.5437622 + 5 \times 3.5048077$$
$$+ 4.7127011 + 6 \times 6.2050306$$
$$+ 8.0217509 + 5 \times 10.2051803$$
$$+ 12.8000000)$$
$$= 4.0157507 \qquad (7\text{-}28)$$

Application of the Gauss Method

$$I_G = \int_{1.4}^{2.0} \frac{x^6}{1+x^2} \, dx = 0.6[R_1\phi(u_1) + R_2\phi(u_2) + R_3\phi(u_3)]$$

using $n = 3$. We now make use of the values of u_1, u_2, and u_3 in Table 7-4 to obtain the appropriate intervals x_1, x_2, and x_3 of the argument.

$$x_1 = 1.4 + (2.0 - 1.4)u_1 = 1.467620999$$
$$x_2 = 1.4 + (2.0 - 1.4)u_2 = 1.70$$
$$x_3 = 1.4 + (2.0 - 1.4)u_3 = 1.932379000$$

Table 7-5 allows the determination of $\phi(u_1)$, $\phi(u_2)$, and $\phi(u_3)$.

TABLE 7-5

	x	x^2	x^6
x_1	1.467620999	2.153911397	9.992714509
x_2	1.70	2.89	24.13756900
x_3	1.932379000	3.734088599	52.06595692

In this example, $\phi(u) = x^6/(1 + x^2)$; therefore

$$\phi(u_1) = \frac{9.992714509}{3.153911397} = 3.168356130$$

$$\phi(u_2) = \frac{24.13756900}{3.89} = 6.20503059$$

$$\phi(u_3) = \frac{52.06595692}{4.734088599} = 10.99809516$$

Multiplying out the terms $\sum_{i=1}^{3} R_i\phi(u_i)$, using the values of R_1, R_2, and R_3 in Table 7-3,

$$R_1\phi(u_1) = 0.880098923$$
$$R_2\phi(u_2) = 2.757791373$$
$$R_3\phi(u_3) = \underline{3.055026432}$$
$$\sum_{i=1}^{3} R_i\phi(u_i) = 6.692916728$$

$$I_G = \int_a^b f(x)\ dx = (b-a)\sum_{i=1}^{3}[R_i\phi(u_i)]$$

$$= \int_{1.4}^{2.0} \frac{x^6}{1+x^2}\ dx = 0.6 \times 6.692916728$$

$$= 4.0157500 \tag{7-29}$$

In order to check the accuracy of the methods worked out above, the integral has been evaluated to eight significant figures.

$$\int_{1.4}^{2.0} \frac{x^6}{1+x^2}\ dx = \left[\frac{x^5}{5} - \frac{x^3}{3} + x - \tan^{-1}x\right]_{1.4}^{2.0}$$

$$= 4.0157501 \tag{7-30}$$

In this particular example, it is apparent that the method of Gauss is the most accurate, while Simpson's rule is the least accurate. Weddle's rule, however, gives a numerical value close to the true evaluation (within six digits in the eighth place).

7-4-7. Evaluation of Double Integrals. There are many important applications in which the numerical value of a double integral is required and where it is not possible to perform the integration analytically. A definite double integral of a function of two independent variables may be evaluated by extending Simpson's rule to functions of two variables or by a repeated application of ordinary quadrature formulas for one variable. Both of these methods are straightforward in principle but may involve detailed numerical calculations. Graphically the procedure amounts to evaluating the function at selected points on a two-dimensional grid. The intervals along the abscissa and the ordinate may be of the same magnitude or may differ greatly in magnitude, depending upon the manner in which the function changes with each of the independent variables.

7-5. Numerical Solution of Algebraic and Transcendental Equations

The methods for finding roots of a numerical equation are numerous. We are interested in this section in examining those methods for solving definite problems which may be used as components in a computing-machine program. Three methods will be discussed: the Newton-Raphson method, the method of false position, and the method of iteration. These methods apply both to algebraic and to transcendental equations. For two simultaneous algebraic equations, an extension of the Newton-Raphson method and the method of iteration will be discussed, and, finally, the Graeffe root-squaring method for solving algebraic equations. It is realized that the material chosen for this section does not by any means exhaust the discussion of the entire domain of methods which could be applied to particular problems.

The Newton-Raphson method, the method of false position, and the method of iteration for finding roots of algebraic or transcendental equations require that an approximate value of the root be known before these methods may be applied. A graphical procedure is frequently used to find an approximate value of the root of an equation.[14] With such an approximate value given, these methods essentially apply corrections to this approximate value, and, by this means, the successive values obtained converge to the correct root of the equation.

7-5-1. Newton-Raphson Method. This method makes use of an approximate value of the desired root and the derivative of the function. Let us assume that the approximate value of the root of a function $f(x) = 0$ is denoted by the quantity x_0 and that this value must be corrected by a quantity h to give the exact value of the root,

$$x = x_0 + h$$

We now expand the function $f(x_0 + h) = 0$ in Taylor's series and neglect all terms in h higher than the first, leaving the approximate relation

$$f(x_0) + hf'(x_0) = 0$$

The first approximation h_1 to the correction h necessary to give

the desired root is

$$h_1 = -\frac{f(x_0)}{f'(x_0)} \tag{7-31}$$

The root of the equation is then improved to a value

$$x_1 = x_0 + h_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{7-32}$$

This process is continued until the value of the root is obtained to the desired accuracy.

As an example of this method, consider $f(x)$ as the polynomial

$$f(x) = ax^2 + bx + c = 0 \tag{7-33}$$

The derivative of $f(x)$ is

$$f'(x) = 2ax + b$$

Using $x = x_0$ as the first approximation and substituting in Eq. (7-32) yields

$$x_1 = \frac{ax_0^2 - c}{2ax_0 + b} \tag{7-34}$$

A special case of Eq. (7-33) is the condition where $a = 1$, $b = 0$, and $c = -u$, so that

$$x^2 = u$$
$$x = \sqrt{u} \tag{7-35}$$

Substituting into the expression for x_1, the equation is obtained:

$$x_1 = \frac{1}{2}\left(x_0 + \frac{u}{x_0}\right) \tag{7-36}$$

which is a useful and well-known form.

As another example of the use of the Newton-Raphson method consider the transcendental equation

$$f(x) = x - \tfrac{1}{2}\sin x - 1 = 0 \tag{7-37}$$
$$f'(x) = 1 - \tfrac{1}{2}\cos x \tag{7-38}$$

Substituting in (7-32),

$$x_1 = x_0 - \frac{x_0 - \tfrac{1}{2}\sin x_0 - 1}{1 - \tfrac{1}{2}\cos x_0}$$

As a first approximation take

$$x_0 = \frac{\pi}{2}$$

$$x_1 = \frac{\pi}{2} - \frac{(\pi/2) - \frac{1}{2}\sin(\pi/2) - 1}{1 - \frac{1}{2}\cos(\pi/2)} = 1.5$$

$$x_2 = 1.5 - \frac{1.5 - \frac{1}{2}\sin(1.5 \text{ radians}) - 1}{1 - \cos 1.5 \text{ radians}} = 1.498062$$

For this value, $f(1.498062) = 0.000616$.

7-5-2. Method of False Position. To apply this method, it is necessary first to find two values of the argument between which the function $f(x)$ vanishes. Preferably these values should be as close together as possible. With these two values of the argument and the two known values for $f(x)$ at these two points, a first approximation to the root is obtained by linear interpolation between these two values of the function. Essentially the function $f(x)$ in the interval is replaced by a straight line. By successive applications of this method it is possible to find a value of the root to the degree of accuracy required.

7-5-3. Method of Iteration. A method for the evaluation of $f(u)$ which is extremely useful for many special cases is based on the following procedure: Given the equation

$$x = f(u) \tag{7-39}$$

it is possible to choose many functions $g(x,u)$ such that for all values of u

$$g[f(u), u] = k$$

where k is constant. For some convenient choice of $g(x,u)$, consider the equation

$$g(x,u) = k \tag{7-40}$$

The principal criterion of choice is convenience in writing the inverse (in u) of the function $g(x,u)$ with x held fast. For $x = x_0$, an approximate value of $f(u)$, $g(x_0,u)$, may be evaluated and written

$$g(x_0,u) = k + h \tag{7-41}$$

where h is the difference between $g(x,u)$ and $g(x_0,u)$.

The function $f(u)$ may now be expressed as a function of x_0 and h by solving for u in Eq. (7-41) and evaluating $f(u)$:

$$f(u) = F(x_0,h) \tag{7-42}$$

The next approximation for x is then obtained by expanding Eq. (7-42) and retaining only the linear term in h.

In this approximate solution the value for h from Eq. (7-41) is substituted and the result simplified.

As an example of this method, the development of an expression for the square root of u will be made.   For $x = f(u)$ write

$$x = \sqrt{u} \tag{7-43}$$

Squaring both sides:

$$x^2 = u \tag{7-44}$$

$$u - x^2 = 0 \tag{7-45}$$

Equation (7-45) is an identity of the type sought, and it leads to a choice $g(x,u) = u - x^2$ and $k = 0$.   Then, if x_0 is an approximate square root of u, in accordance with Eq. (7-41), write

$$u - x_0^2 = h \tag{7-46}$$

Taking the square root,

$$u^{1/2} = (h + x_0^2)^{1/2} \tag{7-47}$$

and regrouping,

$$u^{1/2} = x_0 \left(1 + \frac{h}{x_0^2}\right)^{1/2} \tag{7-48}$$

The quantity $[1 + (h/x_0^2)]^{1/2}$ may be expanded by the use of Taylor's series.

$$u^{1/2} = x_0 \left[1 + \frac{1}{2}\frac{h}{x_0^2} + \frac{\frac{1}{2}(-\frac{1}{2})}{2}\frac{h^2}{x_0^4} + \cdots\right] \tag{7-49}$$

For the approximation only the linear term in h is considered, so that $f(u) = f(x_0,h)$ becomes

$$u^{1/2} \cong x_0 \left(1 + \frac{1}{2}\frac{h}{x_0^2}\right) \tag{7-50}$$

Substituting in Eq. (7-50) a value for h obtained from Eq. (7-46),

$$u^{1/2} \cong x_0 + \frac{1}{2x_0}(u - x_0^2) \tag{7-51}$$

or

$$u^{1/2} \cong \frac{1}{2}\left(x_0 + \frac{u}{x_0}\right) \tag{7-52}$$

which agrees with Eq. (7-36) above.

A different expression may be developed for $\sqrt{u}$ by use of the same method. From the expression

$$x = \sqrt{u} \tag{7-53}$$

write [to obtain $g(x,u) = k$]

$$g(x,u) = \frac{x^2}{u} \tag{7-54}$$

$$\frac{x^2}{u} = 1 \tag{7-55}$$

Let x_0 be the first approximation to x; then $g(x_0,u) = k + h$ becomes

$$\frac{x_0^2}{u} = 1 + h \tag{7-56}$$

Solve for u and take the square root:

$$u^{1/2} = x_0(1 + h)^{-1/2} \tag{7-57}$$

Again using the Taylor series:

$$\sqrt{u} = x_0 \left[1 - \frac{1}{2}h + \frac{-\frac{1}{2}(-\frac{1}{2})}{2} h^2 + \cdots \right] \tag{7-58}$$

And again retaining only the linear portion of this expression

$$\sqrt{u} \cong x_0(1 - \tfrac{1}{2}h) \tag{7-59}$$

But h may be expressed in terms of u and x_0 according to (7-56), and

$$\sqrt{u} \cong x_0 \left[1 - \frac{1}{2}\left(\frac{x_0^2}{u} - 1\right) \right] \tag{7-60}$$

Collecting terms and regrouping,

$$\sqrt{u} \cong \frac{1}{2}\left(3x_0 - \frac{x_0^3}{u}\right) \tag{7-61}$$

Both Eqs. (7-52) and (7-61) give useful iteration schemes for square root of u.

As a numerical example, it is desired to calculate $\sqrt{2}$ to 10 significant figures using both Eqs. (7-52) and (7-61).

(a) Using Eq. (7-52): Let

$$x_0 = 1$$

then

$$x_1 = 1.5$$

Inserting x_1 as the approximate solution in Eq. (7-52) leads to the improved approximation

$$x_2 = 1.416667$$

then

$$x_3 = 1.414215686$$

and

$$x_4 = 1.414213562$$

which is the correct value of $\sqrt{2}$ to 10 places.

(b) Using Eq. (7-61):   Let

$$x_0 = 1$$

then

$$x_1 = 1.25$$
$$x_2 = 1.386718750$$
$$x_3 = 1.413416937$$
$$x_4 = 1.414212890$$
$$x_5 = 1.414213562$$

which is the correct value of $\sqrt{2}$ to 10 places.

The need for an approximation of the value of the inverse of a quantity is frequently encountered.   The required inverse may also be evaluated by using the previous method.   For example, let

$$x = \frac{1}{v} \tag{7-62}$$

An analysis similar to that for $\sqrt{u}$ may be developed.   For $g(x,v) = k$, write

$$vx = 1 \tag{7-63}$$

Let x_0 be an approximate solution and write

$$vx_0 = 1 + h \tag{7-64}$$

Solving for $1/v$,

$$\frac{1}{v} = \frac{x_0}{1 + h} \tag{7-65}$$

Using the first two terms of the binomial series expansion,

$$\frac{1}{v} = x_0(1 - h) \tag{7-66}$$

Substituting for h

$$\frac{1}{v} = x_0(2 - vx_0) \tag{7-67}$$

which is a simple easily calculable expression improving an approximation to the inverse of v. The application of Eq. (7-67) to the design of calculating machinery was first suggested by Aiken in 1938. It is used in the Harvard Mark II.

The Newton-Raphson method and the method of iteration may be extended to find roots of several simultaneous algebraic and transcendental equations. The extension of both these methods is obvious. The scope of this book does not allow a detailed exposition of these methods, especially since they are well described elsewhere.

7-5-4. The Root-squaring Method of Graeffe. The methods which have been described (Sec. 7-5-1 through 7-5-3) deal only with procedures for finding *real* roots of algebraic and transcendental equations; a further restriction on these methods is the requirement that an approximate value of the root of an equation be known before it is possible to apply any of them.

A process has been developed by Graeffe for finding both the real and the complex roots of polynomials without recourse to initial approximate values. This method is based upon the transformation of a given equation into another equation whose roots are high powers of the roots of the given equation. The first step in this process is to transform a given algebraic equation into one whose roots are the squares of the given equation. The second equation is again transformed into a third equation whose roots are squares of the roots of the second equation and, consequently, are fourth powers of the roots of the given equation. This root-squaring process is applied successively until the ratio of any root to the next larger root is negligible in comparison with unity. The roots of the last transformed equation are then considered as separated. By this method the original equation is transformed into a set of simple equations from which the desired roots may be obtained.

If α_1, α_2, α_3, . . . , α_n are roots of the equation $f(x) = 0$, this equation may be expressed as a polynomial.[4]

$$f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \cdots (x - \alpha_n) \quad (7\text{-}68)$$

If this polynomial is multiplied by $f(-x)$ where

$$f(-x) = (-1)^n (x + \alpha_1)(x + \alpha_2)(x + \alpha_3) \cdots (x + \alpha_n) \quad (7\text{-}69)$$

a new polynomial results:

$$f_1(x) = f(x)f(-x) = (-1)^n(x^2 - \alpha_1^2)(x^2 - \alpha_2^2)(x^2 - \alpha_3^2) \cdots$$
$$(x^2 - \alpha_n^2) \quad (7\text{-}70)$$

This polynomial has roots which are the squares of the roots of Eq. (7-68). Again if Eq. (7-70) is multiplied by $f_1(-x)$, another polynomial $f_2(x)$ is obtained:

$$f_2(x) = f_1(x)f_1(-x) = (x^4 - \alpha_1^4)(x^4 - \alpha_2^4)(x^4 - \alpha_3^4) \cdots$$
$$(x^4 - \alpha_n^4) \quad (7\text{-}71)$$

with roots which are the fourth powers of the roots α_1, α_2, α_3, . . . , α_n of Eq. (7-68). This procedure may be repeated until the roots are *separated* as defined above.

After the rth squaring operation the result may be expressed[4]

$$(x^{2r})^n + b_1(x^{2r})^{n-1} + b_2(x^{2r})^{n-2} + \cdots + b_n = 0 \quad (7\text{-}72)$$

If $y = x^{2r}$,

$$y^n + b_1 y^{n-1} + b_2 y^{n-2} + \cdots + b_n = 0 \quad (7\text{-}73)$$

If the operation of squaring has been carried out so that the roots are properly separated, the approximate absolute values of the roots y_1, y_2, . . . , y_n of Eq. (7-73) and α_1, α_2, . . . , α_n of Eq. (7-68) are

$$\begin{aligned} |y_1| &\sim b_1; & |\alpha_1| &\sim \sqrt[2r]{b_1} \\ |y_2| &\sim \frac{b_2}{b_1}; & |\alpha_2| &\sim \sqrt[2r]{\frac{b_2}{b_1}} \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ |y_n| &\sim \frac{b_u}{b_{n-1}}; & |\alpha_n| &\sim \sqrt[2r]{\frac{b_n}{b_{n-1}}} \end{aligned} \quad (7\text{-}74)$$

The Graeffe method does not permit a determination of the signs of the roots α_1, α_2, α_3, . . . , α_n; the signs must be found either by graphing the function $f(x)$ or by other methods such as the application of Descartes' law of signs.[9]

As an example of the Graeffe method a start will be made to find the roots of the fourth-degree polynomial

$$x^4 - 10x^3 + 35x^2 - 50x + 24 \quad (7\text{-}75)$$

(which are known at the outset to be 1, 2, 3, and 4). On the first squaring, the eighth-degree polynomial is obtained:

$$x^8 - 30x^6 + 273x^4 - 820x^2 + 576 \quad (7\text{-}76)$$

This has roots:

$$x_1 = \sqrt{30} = 5.477$$
$$x_2 = \sqrt{273/30} = 3.017$$
$$x_3 = \sqrt{820/273} = 1.733 \qquad (7\text{-}77)$$
$$x_4 = \sqrt{576/820} = 0.8381$$

The result of the second squaring operation leads to

$$x^{16} - 354x^{12} + 26{,}481x^{16} - 357{,}904x^4 + 331{,}776 \quad (7\text{-}78)$$

The third squaring process gives

$$x^{32} - 72{,}354x^{24} + 448{,}510{,}881x^{16} - 110{,}523{,}752{,}704x^8 \\ + 110{,}075{,}314{,}176 \quad (7\text{-}79)$$

A summary of the first three approximations to the roots α_1, α_2, α_3, and α_4 is given in Table 7-6.

TABLE 7-6

	First approximation	Second approximation	Third approximation
α_1	5.477	4.338	4.049
α_2	3.017	2.941	2.979
α_3	1.733	1.917	1.991
α_4	0.8381	0.9812	0.9995

The root squaring may be continued until the desired accuracy in the roots is obtained. The application of the Newton-Raphson method with these first approximations may, in some cases, be found to converge faster to acceptable roots than the squaring process. It should be noted that the sum of the roots in the third approximation is 10.0185, which differs by less than 0.2 per cent from the value of the coefficient of the cubic term in Eq. (7-75).

7-5-5. Numerical Solution of Simultaneous Equations. The Newton-Raphson method and the method of iteration may be extended to find the real roots of several simultaneous algebraic and transcendental equations. The extension of these methods to several equations is described by Scarborough,[19] by Margenau and Murphy,[15] and by Whittaker and Robinson.[20]

The numerical solution of simultaneous linear equations,

which are of prime importance in physics and in chemistry, may be performed by several methods. The method of determinants is useful[15] where the number of unknowns is less than four or five. Beyond this number the method becomes unduly arduous. Matrix methods for solving simultaneous linear equations are described by Frazer, Duncan, and Collar,[8] by von Kármán and Biot,[12] by Bargmann, Montgomery, and von Neumann,[2] and by von Neumann and Goldstine.[17] These methods are useful where there is a large number of unknowns. In quantum-mechanical problems, in which characteristic roots of secular determinants are required, a polynomial method and an iteration method are described by Margenau and Murphy.[15]

7-6. Numerical Solution of Ordinary Differential Equations

The numerical solution of differential equations may be performed by a number of well-established methods. These methods are well described in the literature,[15,18-20] and extensive samples describing each of these methods are available to those who desire a more detailed understanding of the principal methods. Of the many methods which are available for general and for many special types of equations, four methods have been chosen for treatment in this chapter. These procedures are the methods of Picard, Adams, Runge-Kutta, and Milne. A review of these methods is given by Bennett, Milne and Bateman;[3] numerical examples may be found in Scarborough[19] and in Whittaker and Robinson.[20]

7-6-1. Method of Picard. The method of Picard is essentially a process for obtaining the numerical solution of a differential equation by successive approximations. Let us consider initially a first-order differential equation in x and y. This may be written as follows:

$$\frac{dy}{dx} = f(x,y) \tag{7-80}$$

Transposed to integral form, this equation may be expressed as

$$y = \int f(x,y)\,dx \tag{7-81}$$

or

$$y = \int \frac{dy}{dx}\,dx \tag{7-82}$$

If the initial values for x and y are x_0 and y_0,

$$y = y_0 + \int_{x_0}^{x} f(x,y) \, dx \qquad (7\text{-}83)$$

In order to determine successive values of y as a function of x, this integral equation must be solved. We start by holding y constant at y_0 and actually perform the integration with respect to x. This leads to a value y_1 as a first approximation for y, or

$$y_1 = y_0 + \int_{x_0}^{x} f(x,y_0) \, dx \qquad (7\text{-}84)$$

Using the value of y_1 thus obtained, a second approximation y_2 may be found from the relation

$$y_2 = y_0 + \int_{0}^{x} f(x,y_1) \, dx \qquad (7\text{-}85)$$

In obtaining a numerical value for this integral equation, the region in which the desired solution lies is divided into a number of increments (proceeding from the lower to the upper limit). When a satisfactory value for y is obtained after the first step, the procedure of integration is repeated again in the same manner starting with the values of x and y just determined as initial points. By this means a numerical solution for the integral equation, and hence for the differential equation from which it was derived, can be obtained.

Another method closely associated with the method of iteration, or the method of Picard, may be performed by use of an approximating polynomial similar in form to that which was considered earlier in this chapter in the discussion of Newton's formula for backward interpolation.

The application of the method of iteration can be made to differential equations of the second order. This method, however, requires that the differential equation of the second or higher order be reduced to a system of first-order equations.

7-6-2. Method of Adams. Although the method of Adams is somewhat similar in procedure to the method just described, it depends for its usefulness on the ease with which derivatives of the function can be obtained. If $y = f(x)$ is the desired integral relationship appropriate to a differential equation for which a numerical solution is desired, the method of Adams proceeds by expanding this function in a Taylor series about x_0:

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + (x - x_0)^2\frac{f''(x_0)}{2!} + \cdots$$
$$+ (x - x_0)^n\frac{f^{(n)}(x_0)}{n!} \quad (7\text{-}86)$$

In terms of y, this expansion is equivalent to

$$y = y_0 + y_0'(x - x_0) + \frac{y_0''}{2}(x - x_0)^2 + \cdots + \frac{y_0^{(n)}}{n!}(x - x_0)^n$$
$$(7\text{-}87)$$

In order now to find approximate values for y, namely, y_1, y_2, y_3, y_4, etc., we next determine the derivatives y_0', y_0'', y_0''' from the given differential equation for the initial values $x = x_0$ and $y = y_0$; we take the increments in x, which are h, $2h$, etc., successively in the last equation.

The success of the method of Adams depends on how easily the required derivatives may be determined and how quickly the Taylor series converges. It is pointed out, for example, by Scarborough,[19] that this method is not applicable to starting the computation of a ballistics problem because of the difficulty in calculating successive derivatives of the air-resistance term.

7-6-3. Method of Runge-Kutta. This method differs from either of the methods previously described in that the increments of the function are calculated only once by means of definite relationships which will be given below. In this sense, the numerical solution of a differential equation by this method is in a series of approximations for any single increment. Consider a differential equation of the first order in this form:

$$\frac{dy}{dx} = f(x,y) \quad (7\text{-}88)$$

and let the quantity h be taken as the increment by which the independent variable x is attained. It is required to find the increment by which y changes from the initial value y_0. By the Runge-Kutta method, the formulas which are applied are given below:

$$k_1 = f(x_0,y_0)h$$
$$k_2 = f\left(x_0 + \frac{h}{2},\ y_0 + \frac{k_1}{2}\right)h$$
$$k_3 = f\left(x_0 + \frac{h}{2},\ y_0 + \frac{k_2}{2}\right)h \quad (7\text{-}89)$$
$$k_4 = f(x_0 + h,\ y_0 + k_3)h$$
$$\Delta y = \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

The last equation determines the increment in y as a function of k_1, k_2, k_3, and k_4. By this procedure then, values for x_1 and y_1 determined by the initial conditions x_0 and y_0, and the defining equations above, are

$$x_1 = x_0 + h; \qquad y_1 = y_0 + \Delta y \qquad (7\text{-}90)$$

Successive values y_2, y_3, y_4, etc., are determined in the same way and with the use of the same formulas, with x_1 replacing x_0 and y_1 replacing y_0. It is interesting to note that if the derivative is a function only of x, this method reduces to Simpson's rule.

The method of Runge-Kutta may be extended to simultaneous first-order equations or to higher order equations reduced to a set of first-order differential equations by introducing auxiliary variables for transformation. The principal advantages in this method over the method of Picard or the method of Adams lie in the fact that there is required no process of approximation of lengthy iteration by which each increment in the dependent variable is obtained.

7-6-4. Milne's Methods. An interesting method has recently been developed by Milne for the numerical solution of differential equations. This method allows values of y to be calculated after the first four values of y and y' have been determined by other means, perhaps by one of the several methods described above. With these values, the Milne method can then be applied to determine further successive values of y by using an integral form of Newton's formula for forward interpolation, Eq. (7-3), expressed in terms of y' (see Sec. 7-2-1). For example, the next value of y may be approximated by the use of this formula

$$y_n^{(1)} = y_{n-4} + \frac{4h}{3}\left(2y'_{n-1} - y'_{n-2} + 2y'_{n-3}\right) \qquad (7\text{-}91)$$

This equation is obtained by integrating from x_{n-4} to x_n (an interval of $4h$). The value of $y_n^{(1)}$ thus obtained is then substituted into the original equation

$$\frac{dy}{dx} = f(x,y) \qquad (7\text{-}92)$$

so that the value of y'_n may be determined. With this value of y'_n another value of y_n, namely, $y_n^{(2)}$, is obtained through the use of a second relation obtained from Newton's formula, Eq. (7-3), by

integrating from x_{n-2} to x_n (an interval of $2h$),

$$y_n^{(2)} = y_{n-2} + \frac{h}{3}(y_n' + 4y_{n-1}' + y_{n-2}') \qquad (7\text{-}93)$$

This allows a second value for y_n to be determined and to be compared with $y_n^{(1)}$. If these values agree to the accuracy required, $y_n^{(2)}$ is considered to be correct. It is then substituted in the original equation to get the correct value for y_n'. By this method, succeeding values of y may be determined.

REFERENCES

1. Aitken, A. C., On Interpolation by Iteration of Proportional Parts, without the Use of Differences, *Proceedings of the Edinburgh Mathematical Society*, Series 2, Vol. III, Part I, pp. 56–76 (March, 1932).

2. Bargmann, V., D. Montgomery, and J. von Neumann, *Solution of Linear Systems of High Order*. Report prepared under Contract Nord-9596 between the Bureau of Ordnance, Navy Department, and the Institute for Advanced Study (Oct. 25, 1946).

3. Bennett, A. A., W. E. Milne, and H. Bateman, *Numerical Integration of Differential Equations*, Bulletin 92 (The National Research Council of The National Academy of Sciences, Washington, D.C., 1933).

4. Bocher, M., *Introduction to Higher Algebra* (The Macmillan Company, New York, 1907), pp. 17–18.

5. Crumpler, T. B., and J. H. Yoe, *Chemical Computations and Errors* (John Wiley & Sons, Inc., New York, 1940).

6. Feller, W., On Aitken's Method of Interpolation, *Quarterly of Applied Mathematics*, Vol. I, No. I, pp. 86–87 (April, 1943).

7. Ford, L. R., *Differential Equations* (McGraw-Hill Book Company, Inc., New York, 1933), pp. 113–146, 147–162.

8. Frazer, R. A., W. J. Duncan, and A. R. Collar, *Elementary Matrices and Some Applications to Dynamics and Differential Equations* (Cambridge University Press, New York, 1946), pp. 148–155.

9. Hall, H. S., and S. R. Knight, *Higher Algebra* (Macmillan & Co., Ltd., London, 1946), pp. 459–460.

10. Ince, E. L., *Ordinary Differential Equations* (Longmans, Roberts and Green, London; reprinted by Dover Publications, New York, 1944), pp. 540–547.

11. Jeffreys, H., and B. S. Jeffreys, *Methods of Mathematical Physics* (Cambridge University Press, New York, 1946), pp. 237–284.

12. von Kármán, T., and M. A. Biot, *Mathematical Methods in Engineering* (McGraw-Hill Book Company, Inc., New York, 1940), pp. 196–204.

13. Laderman, J., and M. Abramowitz, Application of Machines to Differencing of Tables, *Journal of the American Statistical Association*, Vol. 41, pp. 233–237 (June, 1946).

14. Levy, H., and E. A. Baggott, *Numerical Studies in Differential Equations*, Vol. I (C. A. Watts & Co., Ltd., London, 1934).

15. Margenau, H., and G. M. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1943).

16. Milne, W. E., *Numerical Calculus* (Princeton University Press, Princeton, N.J., 1949).

17. von Neumann, J., and H. H. Goldstine, Numerical Inverting of Matrices of High Order, *Bulletin of the American Mathematical Society*, Vol. 53, No. 11, pp. 1021–1099 (November, 1947).

18. Runge, C., and König, *Numerisches Rechnen* (Springer-Verlag, Berlin, 1924).

19. Scarborough, J. B., *Numerical Mathematical Analysis* (Johns Hopkins Press, Baltimore, 1930).

20. Whittaker, E. T., and G. Robinson, *The Calculus of Observations* 4th ed. (Blackie & Son, Ltd., Glasgow, 1944).

Part II

COMPUTING SYSTEMS

The following five chapters will deal with present and proposed equipment for performing mathematical computations. Chapter 8 describes the basic functions which the desk calculating machine can perform. Chapter 9 outlines the available types of punch-card equipment and indicates how these may be used to perform the four basic arithmetic operations. A general description of several large-scale digital equipments is given in Chap. 10. Some general remarks on analog computing machinery are made in Chap. 11. A possible form which a digital computer might take, based on the requirement for solving a certain class of mathematical problems, is given in Chap. 12.

The purpose of these five chapters is to define, by example, the field of application of computing systems which are composed of the basic arithmetic and physical components to which this volume is primarily devoted.

CHAPTER 8

DESK CALCULATORS

8-1. Introduction

The commercial desk computing machine is a compact instrument capable of performing the basic arithmetic operations quickly and simply. These devices are not intended to compete with electronic machines of large capacity and high speed which are now in existence or in design. The desk calculator makes available to the individual computer a simple machine capable of straightforward calculations. With this machine, the arithmetic operations of addition, subtraction, multiplication, and division can be performed. Although desk calculators were developed primarily for use in the commercial and accounting fields, they may be used efficiently without modification for many scientific computations. While it is not sensible to discard as hopeless those computations which could be carried out only on large-scale digital computing equipment, it is equally unwise (as Comrie[1] points out) to procure expensive custom-built equipment when one of the simpler mass-produced computing devices will carry out the required computations more economically. For this reason it seems desirable to outline here the capabilities of standard electrically driven desk calculating machines.

For some applications, particularly those which do not require the handling of a large number of digits, key-actuated non-automatic machines are satisfactory. A discussion of machines of this type is beyond the scope of this chapter, which is confined to descriptions of the basic functions of electrically driven desk calculators.

In the United States, five companies manufacture electrically driven desk calculating machines: Friden Calculating Machine Co., Inc., Burroughs Adding Machine Co., Marchant Calculating Machine Co., Monroe Calculating Machine Co., Inc., and Remington Rand, Inc. These machines are all generally similar with keyboard setting, electric drive, automatic multi-

135

plication and division, and a few additional special characteristics. They differ from each other in the ease of carrying out particular operations, in available storage and transfer facilities, but not in the type of results obtainable or in the principle of operation.

8-2. General Description of the Machines

The desk calculator is designed with a keyboard of several columns, each column containing a key for each of the digits from 0 through 9. Near the keyboard are mounted various operating bars, clearing keys, and reversing levers; the number of these and the type depend upon the complexity of the machine. (Some machines have supplementary keyboards for the entry of multipliers or other numbers.) Above this keyboard is a movable carriage with two registers; one is known as the *multiplier register* and the other as the *product register*. The product register usually has a digit capacity twice that of the multiplier register. Various supplementary markers and pointers are available for convenience in indicating decimal points, the position of the carriage relative to the keyboard, etc. The carriage contains knobs for clearing each register, and sometimes provision for setting numbers into the product register manually.

8-3. Operation of the Machines

The operation of desk calculators to carry out the elementary arithmetic processes is briefly as follows:

8-3-1. Addition and Subtraction. Assume first that the augend or minuend is already in the product register. The movable carriage is placed in a position so that the digits in this number are directly above the corresponding digits on the keyboard, and the addend or subtrahend is set digit by digit into various columns of the keyboard. Depression of the plus bar for one cycle of the machine adds the number entered on the keyboard to the number in the product register, leaving the sum in the product register. Depression of the minus bar for one cycle of the machine subtracts the number in the keyboard from the number in the product register and leaves the difference in the product register; negative numbers are represented by complements. For example, if all the dials are cleared and a given number is set into the keyboard and subtraction performed, the machine represents this negative number by indicating, in the

product register, another number to which the given number must be added to give a certain power of 10. In a 20-digit product register this number is usually 10^{20}, which exceeds the capacity of the machine by one digit. In some machines capable of holding 20 digits in the product register, this number may be 10^{14}. This system of representing negative numbers by their complements is commonly used for computations with logarithms in order to keep all mantissas positive; it serves the same purpose on large-scale computing machines. Several desk calculators include a button near the keyboard marked *add;* if this button is depressed, the keyboard is cleared after one machine cycle, and careless prolonged depression of the operating bars will not produce incorrect answers.

8-3-2. Multiplication. To multiply, the multiplicand is set in the keyboard, and the product and multiplier registers are cleared. The machine multiplies by successive additions. With the carriage in its extreme left-hand position and with the *add* button released, the plus bar is depressed for a period of time sufficient to permit the extreme right digit of the multiplier to build up in the multiplier register. In this operation the digit of the multiplier register immediately above the extreme right column of the keyboard increases by one for each cycle the machine operates. Thus, at the time of this first step, in which the right digit of the multiplier is entered into the multiplier register, the multiplicand is added into the product register by a number of times equal to the right digit of the multiplier. Next, the carriage is shifted one digit to the right, and the process is repeated until the 10's digit of the multiplier appears in the proper place in the multiplier register. In this way, iteration of the add-and-shift technique yields the final product in the product register.

Many keyboard machines perform automatic multiplication. In this process the multiplier is set into an auxiliary keyboard, and the multiplicand is set into the upper keyboard. The machine will automatically carry out the process described above.

8-3-3. Division. Normally, division is performed by a process of successive subtractions similar to the multiplication process described above. The dividend is set in the product register well to the left. The divisor is set into the keyboard underneath the dividend. The multiplier register is cleared. A reversing lever

near the keyboard is thrown so that the number in the multiplier register increases when the minus bar is depressed and decreases when the plus bar is depressed. With this arrangement the minus bar is depressed for a time long enough to introduce a series of 9's at the extreme left end of the product register. This indicates that the partial remainder has become negative and that the number in the multiplier register is greater than the quotient sought. The plus bar is depressed for one cycle, making this partial remainder positive again, and the carriage is shifted one digit to the left. With the carriage in this position the process is repeated. This iteration is continued until the carriage is in its extreme left position. At this time the quotient is in the multiplier register, and the remainder is in the product register. This is the process described in terms of elementary computing operations in Chap. 6.

This process is built in for automatic operation in many desk calculators. Provision for automatic division is far more common than automatic multiplication in desk calculators.

8-4. Capacity, Speed, and Capabilities

In describing desk machines the capacity is usually given by a series of three numbers. The first of these numbers indicates the number of columns in the keyboard; the second indicates the number of digits in the multiplier register; the third indicates the number of digits in the product register. Thus, a machine described as $10 \times 11 \times 21$ is one with 10 columns in the keyboard, 11 digits in the multiplier register, and 21 digits in the product register. This is a fairly common size of machine. The speed of calculation is dependent upon the cyclic speed of the machine and upon special features built into machines. Most machines operate at a rate of 500 to 650 cycles per minute, although some models operate at a rate of 1,350 cycles per minute. Some time saving may be obtained if the machine starts an automatic multiplication while other digits are entered. Time may also be saved by automatic short cutting in multiplication; this process shortens the number of steps required in some multiplications by running the machine backward to get high digits on the multiplier register and later correcting left-hand digits for unwanted carry-over. Thus to multiply by 89, the

machine might first multiply by -1, shift and multiply by -10, shift again and multiply by 100 to clear the unwanted 9's.

Since the speeds of various computing machines and their separate components have been mentioned elsewhere in this book, it seems advisable here to note briefly the performance of desk calculators in this respect. In the succeeding section of this chapter the addition speeds of several desk calculators are given. For two of the machines, the Friden and the Monroe specifically, the addition speed is of the order of 500 to 650 counts per minute; for another machine, the Marchant, the addition speed is about 1,350 counts, over twice the speed of either of the first two machines.

In the operations of multiplication and division, which are performed on all these machines by an iterative process of successive addition (or subtraction) and column shift, there is no marked difference in the total times of all three machine operations. To indicate this, there are listed in Table 8-1 a few representative operations involving 10-digit multiplications and divisions with the approximate performance times in each machine.

Table 8-1

	Friden	Marchant	Monroe
Addition speed	500 digits per minute	1350 digits per minute	650 digits per minute
Multiplication: 9,999,999,999 $\times$ 9,999,999,999	12 seconds	10 seconds	11 seconds
Division: $\dfrac{9,999,999,999}{3.141592654}$	9 seconds	8.5 seconds	10 seconds
$\dfrac{9,999,999,999}{2.8182846}$	9 seconds	8 seconds	9.5 seconds

It has been estimated, as a conservative figure, that in an 8-hour day one good operator can produce 400 10-by-10-decimal-digit multiplications or divisions. This includes time for recording the answer of each operation and for making some elementary accuracy checks. One thousand additions may be performed in the course of an 8-hour day by the one calculator. Considerable saving of time can be effected if a program can be

devised in which the numbers in the registers of the machine
are reused without transcription.   For example, in using Aitken's
interpolation method described in Sec. 7-2-5, the computed value
of the determinant appears in the main dials and is divided by
the appropriate factor without transcription.

8-5. Special Features

The manufacturers of calculating equipment have built into
some of their machines several special features designed to
facilitate particular types of computation.   A description of a
few of these features is given here.   It should be emphasized
that an increase in the number of automatic features on the
machine usually represents an increase in the service require-
ments of the device.   The most economical machine which will
carry out an expected computation is generally the best choice for
that computation.   Special features mentioned by the various
companies are as follows:

8-5-1. Burroughs.   The Burroughs Adding Machine Com-
pany lists two Electric Duplex Calculators which are adaptations
of the adding machine.   These calculators have no provision for
automatic multiplication or division; these operations must be
performed by manual positioning in steps across the keyboard to
simulate the action of a movable carriage.   The outstanding
characteristic of these calculators is the provision for using two
accumulating registers, one for the results of detailed calculations,
including the four arithmetic operations, and the other for
accumulating the results occurring stepwise on the first register.
The results from the first register may be added to or subtracted
from those in the total or second register; this transfer resets the
first register for the next operation.

8-5-2. Friden.   The Friden has a 10-key setting for the mul-
tiplier in automatic multiplication.   The machine, however, does
not have complete 10's transmission in the product register;
essentially, the Friden takes complements on 10^{14} and not
on 10^{20}.

The Friden machine is available with both the *split* and the
normal dial clearance.   With this adaptation, the upper dial
may be split into a right and left section at the seventh, eighth
or ninth digit.   The split feature enables the operator to clear

the left upper dial and to accumulate in the right upper dial or to accumulate in the left and clear the right upper dial.

A Multiplier Nonentry Control available only on Friden Models S and ST is an important advantage for calculations such as ab/c. With a special counter control key, the individual or the accumulated multipliers are prevented from entering the lower dial. The product ab, entered in the upper dial, may then be divided by c with the quotient entering the lower dial. In this operation only the results of division are registered and accumulated on the lower dial.

The normal operation of the Friden is approximately 500 counts per minute.

8-5-3. Marchant. The newer models of the Marchant operate at a counting speed of 1,350 counts per minute. In the other types of machines, both multiplicand and multiplier must be set before multiplication begins. In the Marchant the multiplicand is set as usual. The multiplier is set into an auxiliary row of 10 keys. As each key is pressed, its digit is absorbed into the multiplier and the carriage moved to the next position. While this machine is running on one digit, the next may be set. Since the speed of the machine is 1,350 counts per minute, the product is always completed within $\frac{1}{2}$ second after setting the last digit. This feature enables an operator to carry out more multiplications per hour because he does not have to spend as much time waiting for the machine. For the same reason, however, some operators find the machine more tiresome to use for extended periods of time.

The high speed of the Marchant is possible because the numeral wheels are operated by continuously driven gears. These gears are selected by the keyboard setting.

The Marchant has also the split-dial feature mentioned above. Some special results have been obtained by mounting two Marchants together, the result being called a twin Marchant. Both machines can be driven separately, can be turned together in the same direction, or can be turned in opposite directions. An expression of the form ab/c can be evaluated in one operation. Twin machines have been applied to military rectangular-coordinate survey work and to the interpolation of pairs of double-entry tables.

8-5-4. Monroe. Some models of the Monroe Calculator have an automatic feature which makes it possible to form Σa^2 with a single setting of the values a. Some older models of the Monroe incorporate automatic short cutting.

Another useful feature of the Monroe is the ability to call a stored constant multiplier at will instead of setting it when required. This characteristic is useful for computing expressions such as $Kx + y$ where x and y are variables and K is a constant. The storage of this constant in no way limits the use of the keyboard for other operations. By use of a special short cut a constant dividend may also be stored.

The product abc can be taken very easily on some Monroe machines by using a transfer switch which transfers a number from the product register to the keyboard. This transfer facility is also useful in computing such expressions as ab/c.

Models are available with split dials. This feature enables the operator to operate on either half of the dial or on both halves and to clear either half as desired.

In normal operation the Monroe Calculators add digits at the rate of about 650 per minute.

8-5-5. Remington Rand. Remington Rand manufactures two printing calculators, one with a 13-column *list*, or keyboard, and 14-column total, and one with a 10-column list and 10-column total. The principal advantage in these calculators is their provision for printing all steps in the basic arithmetic operations. The unit operation time, including that for introducing the problem and for printing, compares favorably with other commercial calculating machines which do not have the printing feature. The 13-column model provides a special feature for short-cut multiplication by which it is possible to use complements on 9's for digits of the multiplier greater than five. The provision for printing all factors in an arithmetic operation and the result of the operation provides a valuable means by which a check on the individual steps in any calculation, or series of calculations, may be obtained for a permanent record.

8-6. Summary

The desk calculator is a machine which will produce products, quotients, sums, and differences of numbers quickly and conveniently. It will handle augends, minuends, products, divi-

dends, sums, and differences to a capacity of about 20 digits, and multipliers, multiplicands, addends, subtrahends, divisors, and quotients up to a capacity of 10 digits. Such a machine in the hands of a competent operator can produce 400 full-length products or 1,000 sums during an 8-hour working day. The desk calculator is an economical method for performing computations of this order of magnitude. Its speed is low enough to permit continuous exercise of judgment by the operator.

Information concerning desk calculators is readily available from the manufacturers. In addition, the following publications[1-6] are noted as pertaining to the general characteristics and applications of these machines.

REFERENCES

1. Comrie, L. J., The Application of Commercial Calculating Machines to Scientific Computing, *Mathematical Tables and Other Aids to Computation*, Vol. II, No. 16, pp. 149–159 (October, 1946).
2. Hutchinson's Technical and Scientific Encyclopedia, Calculating Machines, Vol. I (The Macmillan Company, New York, 1936), pp. 315–319.
3. Marchant Calculating Machine Company, *Marchant Methods, Index of Marchant Methods and Tables Issued to Nov.* 1946 *Relating to Basic and Statistical Mathematics*, MM-229 Mathematics (Marchant Calculating Machine Company, Oakland, Calif., 1946).
   The following titles of methods applicable to the Marchant Calculator have been abstracted from this reference to indicate the range of usefulness of desk calculators:

   a. *Roots and Powers:*

MM-88	Approximation Method for Extraction of Any Root
MM-95	Square Root by Successive Odd Numbers
MM-222	Fifth Root to 5 Places
MM-302	Square Root by Iteration from Slide-Rule Approximation
Table 68	Cube Root to 5 Places, with Extension to 9 Places

   b. *Algebraic Equations:*

MM-182	A Short Method of Evaluating Determinants and Solving Systems of Linear Equations with Real or Complex Coefficients, by Prescott D. Crout, Ph.D.
MM-183	Notes on Marchant Calculator Application to the Crout Method of Solving Simultaneous Equations (see MM-182)
MM-225	Birge-Vieta Method of Finding a Real Root of Rational Integral Function
MM-226	Setting up an Approximating Polynomial of Degree "n" from Equidistant Tabulated Values of a Function

MM-235 The Nogrady Method of Solving Cubic Equations
MM-289 Evaluation of a Polynomial

c. *Interpolation:*

MM-64 Direct Curvilinear Interpolation with Lagrange Coefficients

MM-152 Direct Curvilinear Interpolation, Assuming Constant Second Differences

MM-189 Direct Interpolation—Straight-Line and Curvilinear

MM-209 Curvilinear Interpolation with Unequal Intervals of Argument. ("Divided difference" method using Lagrange-Newton formula.)

MM-220 Inverse Curvilinear Interpolation—Short-Cut Method Includes Effect of 2nd Differences Only. (Using Bessel's central-difference formula).

MM-221 Inverse Curvilinear Interpolation and Finding of Roots of Tabulated Function. (The Comrie "two-calculator" method, using Bessel's central-difference formula.)

MM-228 7-Place Lagrange 5-Point Interpolation Coefficients for Values of P from 0 to 2, with Argument to 0.001

MM-317 The A. C. Aitken Method of Curvilinear Interpolation. (For equal or unequal intervals of the argument.)

d. *Numerical Integration and Solution of Differential Equations:*

MM-167 Moment of Inertia of Sections Composed of Rectangular Areas. (A systematic work sheet for computing the constants of structural shapes.)

MM-215 Area below Curve for Fractional Portion of Distance between Equidistant Ordinates.

MM-216 Milne Method of Integration of Ordinary Differential Equations. (Complete explanation and systematic work sheets for this popular method. Includes much heretofore unpublished information.)

e. *Statistical Method, and Least Squares:*

MM-45 Pearson Correlation Coefficient. (With formula especially adapted to calculator computation.)

MM-119 Linear "Least Squares" Line of Regression and Coefficient of Regression

MM-242 Summations of X, XY, and XY^2. (A special simplified method.)

MM-245 Hansen-Ahlberg Method of Extending Parabolic Curves. (A method based upon constant second differences for rapidly extrapolating any second-degree function.)

MM-314 Summations for Linear Multiple Correlation

4. Meyer, zur Capellen, W., *Mathematische Instrumente* (Becker and Erler Kom.-Ges., Leipzig, 1944; lithoprinted by Edwards Bros., Inc., Ann Arbor, Mich., 1947), pp. 53–131.

5. Murray, F. J., *The Theory of Mathematical Machines* (King's Crown Press, New York, 1947).

6. Office Machines Research, Inc., *American Office Machines Research Service* (The Office Machines Research, Inc., New York, 1938–1939). Three binders: *Adding and Subtracting*, Index 3.0, Binder II. *Listing Adding Machines*, Index 3.2; *Single Register Listing Adding Machines*, Index 3.21; *Multiplying and Dividing*, Index 4.0; *Calculating Machines*, Index 4.3; Binder III. (This service is no longer available in the United States. Further information on the above subjects may be obtained from the International Office Machines Research Corp., Amsterdam, Holland.)

CHAPTER 9

PUNCHED-CARD COMPUTING SYSTEMS

9-1. Introduction

These systems of computation make use of commercially available punched-card processing equipment. Business operations with such equipment are well known and have been developed extensively in Great Britain and in the United States over the past 50 years. The punched card now in use is an adaptation of the card originally conceived by Hollerith, who, as early as 1889, described tabulating machinery for the punched-card processing of statistical data.

Since the punched-card equipment and its use on scientific problems have been described adequately elsewhere, it is the purpose of this chapter to present merely an outline of a calculating system using punched cards. Although punched-card machinery was primarily developed for business accounting applications, it has been successfully utilized, without major alteration, in extensive computations of a scientific and statistical nature. For scientific calculations, a group of four or five standard business equipments, performing different functions, are usually operated as a computing unit.

There are several advantages to be gained through the use of punched-card equipment in scientific work. In the first place, for many extensive calculations to which it may be applied, this type of equipment provides a speed in excess of that attainable with desk calculating machines. Second, greatly increased accuracy may be attained through the automatic operation with punched-card equipment, even in those computations where the speed is comparable to that of a desk calculator.

The increased accuracy attainable with punched-card equipment stems principally from three sources. First, there is the possibility of automatic transcription of results into tabular form. Where a lengthy transcription is required, it is possible with punched-card equipment to print tables which can be

reproduced by a photo-offset process without any intermediate manual duplication. This operation greatly diminishes the possibility of making a printing error which may go undetected in normal proofreading. Such a system of duplication is now standard with many punched-card installations and has been used extensively in the computation and tabulation of mathematical functions.

A second source of increased accuracy in computations made by punched-card machinery lies in the fact that its automatic operation does not weary of tedious iteration. The lessened dependence on the alertness of the operator is a considerable factor in increasing the reliability of the results even though the computations on punched-card machinery may be performed no faster than on a desk calculator.

A final source of increased accuracy by punched-card devices lies in the possibility of including automatic checking features which are far more reliable than a human proofreader. Use of these devices enables vast reduction in the probability of machine error, and in many cases checking equipment may be used in the same way to reduce the probability of undetected operator error.

In the United States there are two principal manufacturers of punched-card equipment: International Business Machines Corporation and Remington Rand, Incorporated. Descriptions of the equipment now available and the equipments planned for the immediate future are available from these companies; detailed suggestions and instructions concerning the efficient use of machines built by these companies may be obtained from them on request. For scientific computations, at least, installations of International Business Machines Corporation equipment seem more numerous than those of Remington Rand equipment. The description of machines in this chapter will therefore concern itself largely with IBM equipment. Comparable punched-card equipment of the Remington Rand type will also be described where it appears that it may be applied to scientific calculations. It should be noted that it is not the intention of this chapter to provide a sufficient description of any commercial piece of equipment to suggest its acquisition by any activity.

The description of punched-card equipment given here is included for several reasons. One reason is the desire to assist

projects, faced with computing problems, in avoiding undue investment in needlessly complicated computing equipment because of the lack of information concerning the potentialities of punched-card equipment. Second, the manner in which punched-card machinery permits increasing automatic computation is useful in connection with the study of more extensive computing equipment. An examination of the methods of computation already developed for punched-card equipment might be instructive to many activities, even though their computing problems are too extensive for adequate treatment by these machines. Finally, it is of interest to note the development of techniques for the utilization of this automatic equipment in computations. The ingenuity which has been exercised by the various workers in formulating methods with which their problems can be treated on this machinery, and the publication of these methods in the ordinary scientific literature, may indicate that true radical advances in the computing art might follow the acquisition of equipment which is somewhat less powerful than seems necessary at first glance.

9-2. The Punched Card

The primary unit for the automatic operation of these calculating equipments is the punched card. The IBM and the Remington Rand machines operate on punched cards of the same size, 3.250 inches by 7.375 inches, and the same thickness, 0.0067 inch. Generally one corner is cut to facilitate card handling, matching, filing, etc. The IBM and Remington Rand cards do not store the same quantity of numerical and/or alphabetical information; there is also a different distribution of fields and zones on each type of card. It is important to point out here, however, that both IBM and Remington Rand manufacture devices for transcribing data from one type of card to the other.

9-2-1. The IBM Card. The IBM card has 80 vertical columns each with 12 punching positions. The 10 lower positions are assigned to the digits 0 to 9; the top two positions designated x and y, or more commonly referred to as 11 *and* 12 *holes*, are used for special coding, such as the indication of negative numbers or, in combination with one of the digits 1 to 9, for alphabetical representation. For use in specific calculations, a card may be divided into sections, or groups of columns, known as *fields;*

this defines that portion of the card in which information of a certain kind will always appear. In one type of computation, for example, the 80 columns were divided into eight groups of 10; each of these groups in turn was subdivided into one column and three groups of three columns.

The 80 columns are usually punched and read independently, each punched column denoting one digit, one letter of the alphabet, or a special character such as the algebraic sign of a number. For numerical representation, one perforation is required for each digit. For alphabetical representation, two perforations in a single column are used for each letter; one of these is a zone punch (0, 11, or 12) while the other perforation is made in the position identifying one of the digits 1 to 9. Table 9-1 illustrates the coding used to represent the alphabet on the IBM card.

TABLE 9-1

Zone Digit	11	12	0
1	A	J	
2	B	K	S
3	C	L	T
4	D	M	U
5	E	N	V
6	F	O	W
7	G	P	X
8	H	Q	Y
9	I	R	Z

Note that the (0,1) position is not used for the alphabet and is therefore available for special coding.

All IBM computing equipment is capable of operating on data in punched cards. Some machines, however, can also handle data represented as graphite deposits on cards. These marked cards are of different capacity and design but are not radically different functionally. Both sides of a marked card are available for storage with 27 columns of information per side.

The punched card is the fundamental unit of any IBM automatic computation system. Numbers computed in the course of a computation are punched on a card for further use, for transfer from one unit to another, or for storage. The cost of these

cards is about \$1 per 1,000; the cards used as interim storage or as a transfer medium in the course of a computation are not generally used again.

The face of the card may be printed to facilitate identification by visual reading. Some punch equipments print on the card the information which has been punched there; the printed characters usually lie above the column in which the information is punched. Additional equipment is also available for reading information punched on a card by other equipment and for printing some of this information at the edge of the card. Thus, the cards may serve as stencils with no further use of the automatic equipment. This card function is frequently of double convenience, for the stencils are compact, stiff, and easily stored, and they can be automatically reproduced on standard machines. One of the early uses of punched cards as stencils was that by J. D. Elder; he revised and extended the factor stencils originally prepared by D. N. Lehmer[3] and punched the results on cards. This publication is out of print, but copies of the stencils can be prepared quickly and easily by means of a reproducing punch.

9-2-2. Remington Rand Card. The Remington Rand Card has 90 vertical columns each with six punching positions. The card is divided into an upper and lower section, each section extending across the length of the card. There are 45 columns and six rows in each section; this is an adaptation of the Powers card which originally contained a total of 45 columns. The Remington Rand equipment will operate on cards with a capacity of 45 columns or 90 columns. Each column of the card is headed by a 0 position, the five positions vertically beneath this representing, respectively, (1,2), (3,4), (5,6), (7,8), and 9. The odd numbers and 0 are represented by a single punch; the even numbers are represented by a two-hole punch in each column, one punch in the hole for the even number, the other in the nine position. For representing the alphabet, a two- or three-hole punch is required.

This card is the basic unit for the operation of Remington Rand equipment.

9-3. Input

Data are introduced to punched cards through a manually operated punch with a keyboard similar to that of a typewriter.

This keyboard actuates a mechanism which translates a number, letter, or other code into the proper punched designation on the card. For the alphabet, a single key may operate a triple punch simultaneously.

Data may also be introduced to cards by automatic transcription equipment or by punches operating from control panel or plugboard instructions.

9-3-1. IBM Input. Data may be introduced to IBM cards by means of a manually operated punch, a tape-reading punch, or special equipment for automatically transcribing data on cards. Many models of manual punches are available. These punches may produce double punches for alphabetical information from a single stroke on a keyboard; they may print punched information at the top of the column; they may provide for automatic reproduction of punched cards. The larger punches have many other attachments to provide automatic aids to accuracy and convenience, such as tabulating stops, skip bars, automatic ejection and feed, etc.

The tape reader automatically reads punched tape (like teletype tape) and transcribes the data represented by the holes in the tape to a deck of cards. Other equipment can perform the reverse operation, reading the holes punched in the cards and producing a tape on which the same data are represented.

With regard to input equipments, it may be of interest to note that the International Business Machines Corporation has designed and installed special equipment to read wind-tunnel data directly from various gauges and to transcribe these data on card punches. No standard equipment to accomplish this, however, has been advertised as yet.

The existence of the Electric Punched Hole Verifier should be noted in connection with punch equipment. This device provides automatic detection of variation of entering information by separate operators. The original punched cards are set on the machine, and the verifier operator, reading data from the punched source, introduces these data to a keyboard similar to that on the punch. If the punching in a card column does not agree with the key struck on the verifier, an error light is illuminated, and operation of the verifier is automatically stopped. If there are no errors, the verifier card is ejected and stacked. The verifier may be used to verify either alphabetical or numerical punches.

The manually operated punches and the verifier operate at
typewriter speeds; manual punches with automatic reproducing
attachments operate at about 10 columns per second.   The tape
reader operates at about 8 columns per second.   Normally,
manual punching of numerical data from clear manuscript may
be carried out by an experienced operator at the rate of 125 cards
per hour with 80 punches per card.

9-3-2. Remington Rand Input.   The Remington Rand equip-
ment for the input of data is similar in character to the IBM
input equipment.   A typical input mechanism manufactured by
Remington Rand is the 90-column Alphabetical Punch, which
introduces both alphabetical and numerical information on a
card.   With this punch, a verifying attachment is available for
preparing the punched cards for mechanical verification.   The
90-column Alphabetical Punch may include an attachment for
serial numbering of each card which has passed through the
machine.   The serial number is printed along the left edge of
the card; a counter on the front of the machine indicates the
number of cards punched or stamped.

The punch equipment for data input is manually operated and
operates at typewriter speeds.   Other equipments, such as the
Summary Card Punch, the Reproducing Punch, and the Syn-
chro-Matic Punch, operate at speeds above that of a typewriter.
Their function is to read automatically data punched on cards
in order to punch these data on other cards or to print them on
paper tape.

It is interesting to note that the Remington Rand punches
contain an electrical link from the keyboard to the characters on
the carriage.   This is the only such linkage in the Remington
Rand equipment.   All verifying, reading, arithmetic operations,
and printing in the Remington Rand equipment are actuated by
mechanically sensing the card coding.   This is claimed to be an
advantage in favor of accuracy.   The mechanical nature of these
operations, however, prevents the use of orders presented to the
machine by electrical plugboard connections, a function which
gives the IBM equipment its wide application and flexibility.
For introducing numerical data and various special orders, the
use of punched tape, employed by IBM, has made that com-
pany's equipment more readily adaptable to certain types of
scientific calculation.

9-4. Machines for Punched-card Computations

Machines for performing punched-card computations are capable of carrying out the following operations: counting, sorting, consulting tables, addition, subtraction, multiplication, division, and the printing of results. A large number of computational techniques have been developed to utilize these operations. The development of these techniques has, for example, enabled such operations as multiplication to be carried out on machines that are not advertised as multipliers. All the operations enumerated above can be performed on IBM and on Remington Rand equipments. IBM equipments have been, in general, much more widely used on scientific calculation.

9-4-1. IBM Machines. The principal IBM equipments useful in performing scientific calculations are the Sorter, the Collator, the Tabulator, the Electric Multiplier, and the Calculating Punch. These machines usually form a calculating unit when they are used together for scientific work. Table 9-2 is presented as an outline of the operations which can be performed on each of these machines. Detailed information

TABLE 9-2

IBM machine	Count	Sort	Consult tables	Add	Subtract	Multiply	Divide
Sorter	x	x	x				
Collator		x	x				
Tabulator	x			x	x		
Electric Multiplier				x	x	x	
Calculating Punch				x	x	x	x
Electronic Calculating Punch	x			x	x	x	x

concerning these operations is available in published form.[1] In addition, the utilization of this equipment to carry out prob-

lems in scientific computing is described in many papers; a bibliography of some principal papers in English concerning the use of IBM machines in scientific research, statistics, and education has been published by the International Business Machines Corporation; it is available from them. With their permission, the sections entitled Scientific Research and Statistics are included in the Bibliography at the end of this chapter. For convenience, a brief description of the machines and of the operations will be given here in condensed form.

The Sorter. The electric punched-card sorting machine is a machine which separates cards into 13 pockets depending upon the character punched in a chosen column. For normal numerical sorts, one of the 12 pockets receives cards appropriate to one of 12 different punches, and the 13th receives cards with no punch at all in the chosen column. Cards already sorted on one column can be stacked together and resorted on another column so that, finally, a deck of cards can be ordered in accordance with numbers punched across several columns. If a column contains two or more punches, the Sorter normally selects in accordance with the punch lowest on the card in this column. However, it is possible to disable readers at any chosen level; by this means alphabetical sorting is easily brought about on two runs per column through the sorter. The sorter can be procured with 15 counters attached, so that the number of punches detected in each of the 12 column positions, the number of cards unpunched in the chosen column, a total, and a subtotal can be automatically counted. Counting may be done without sorting, if desired. Sorting can be performed at the rate of 450 cards per column per minute; the newest machines operate at 650.

The Electric Punched-card Collator. This machine is designed to carry out somewhat more complicated rearrangements of cards than the sorter. It is particularly useful in connection with consulting tables of functions. The machine has two feeds and four output pockets. It achieves a considerable versatility in its disposition of the cards from the two input decks to the output pockets by virtue of a removable plugboard into which plug wires are manually inserted to provide the desired operations. The operation of the machine depends upon comparison of the information read from three cards. The two leading cards from one feed and the one leading card from the other are read

simultaneously. Comparison circuits permit identification of several cases of relative order of the cards, and the various cases set off driving mechanisms to advance the cards from one of the two feeds, channeling the most advanced card from this feed into one of the four output pockets. As an example, a sorted set of arguments could be introduced to one feed and a function table (containing cards for arguments including those in the sorted set, this function table sorted according to argument) could be introduced in the other. The four pockets could be assigned to:

1. Functional-value cards corresponding to arguments appearing in the first deck.
2. Functional-value cards corresponding to arguments not appearing in the first deck.
3. Cards from the first deck for which functional values were found.
4. Cards in the first deck for which no functional values were found.

(The fourth output pocket would contain cards which might have been misplaced in the original sorting process, for which the cards from the functional deck have been lost or otherwise removed, etc.) In a later operation the Collator can be used to merge the separated decks into their original orders.

The Collator works at a speed of about 240 operations per minute. Each operation may involve the transfer of one or two cards, or none, to an output pocket, depending upon its nature. Almost all Collator operations are on decks which have already been sorted and which are to be dealt into the four pockets with some merging or separation but with no further reordering.

The Tabulator. There are several types of tabulators, of which a standard useful model is the Type 416 Electric Punched Card Accounting machine. The company describes this machine as a combination adding, subtracting, and printing machine. Control of the Tabulator is flexible. It is brought about through wiring set on a removable plugboard.

The input to the Tabulator is from a single feed with reading stations examining the two most advanced cards simultaneously. There are two reading stations: the control station and the arithmetic reading station. The control station causes the device to carry out automatically various arithmetical or printing opera-

tions, such as addition, subtraction, clearing, and printing. The arithmetical operations are all based on addition; access to the adding counters is only through the more advanced reading brush.

The Tabulator, through its provision for automatic printing, is one of the most valuable of all the IBM devices for automatic computation. As pointed out earlier, the possibility of printing the output of a computation without any manual intervention is an important step in increasing the reliability of the output. It has become standard practice at many places to photograph automatically printed numbers for photo-offset printing to avoid typesetting errors.

The output of the Tabulator can be punched on another deck of cards through use of the summary punch. Rearrangement of data on summary cards is possible by virtue of a plugboard on the punch.

The Tabulator operates at a speed of 9,000 cards per hour while tabulating or listing. If its output is to be punched on another deck of cards, the operation is delayed by about 1.2 seconds per card punched.

The Electric Multiplier. This machine multiplies numbers punched on cards and punches the products. One factor is always derived from an individual card, and the other may be obtained either from a different field on the same card or from master cards containing group multipliers. The machine will accept eight-digit multiplicands and multipliers and produce products up to and including 16 digits. The Electric Multiplier contains a summary counter which will accumulate a 16-digit summary of products.

The machine is available in a model which will add and subtract as many as three numbers punched on the same card. Two of these numbers may each contain 12 digits; the third may contain as many as eight digits. Some slightly more complicated operations can be performed combining addition with multiplication.

The speed of the machine depends on the number of digits in the multiplier. Ten-digit products of two eight-digit factors can be accumulated at the rate of 870 per hour; if they are punched, however, the rate is slowed to 730 per hour. For the subtracting function, the speed is about 1,000 cards per hour.

The Calculating Punch. The calculating punch will multiply, divide, add, and subtract. It is capable also of carrying out a short sequence of operations automatically. The machine will multiply two 10-digit factors to yield a 19-digit product. It will accept a 20-digit dividend and a 10-digit divisor to yield a 10-digit quotient. The possibilities of combining operations in this machine are fairly great. The speed is controlled mainly by the time of punching; it also depends upon the size of the numbers involved. Twenty-digit products may be computed at the rate of 1,000 per hour. Division speed is considerably less than this; for a five-digit quotient, the speed is approximately 540 cards per hour. There are two calculating punches, the Type 602 and the Type 602A. The statements made above apply to both machines, but the 602A is a little more versatile.

The Electronic Calculating Punch. During the early part of 1948 the availability of the IBM Electronic Calculating Punch, Type 604, was announced. This machine performs all the basic arithmetic operations by electronic methods. The numerical instructions to the machine are read from IBM cards; the results of the calculations are punched automatically on cards. The Type 604 is much faster and is a more versatile machine than the Calculating Punch, Type 602. It is faster than the 602A but not more versatile.

The Electronic Calculating Punch is a combination of two separate units, a punching unit similar to an IBM Gang Punch and an electronic unit for performing the arithmetic operations. The punch receives operating instructions from two control panels; these panels control factor reading, result punching, checking, and gang punching. The electronic calculating unit has one control panel for instructions relating to the type of arithmetic operation to be performed and to the transfer of numbers from one storage unit to another.

The Type 604 can perform 60 program steps, or operations, per card; a program step includes any one of the four arithmetic operations, or a number transfer. The time usually allotted to a group of program steps is about 80 milliseconds; this is the time required for a card to pass from the first reading station to the punching station.

Numerical information read from cards may be held in the machine for as many arithmetic operations as may be required.

For example, a constant factor may be stored in either the factor or the general storage for successive use. For calculating, the machine has a reading capacity of 21 digits per card. Since the capacity of the electronic part of Type 604 is 50 digits, the total number of digits available for punching is restricted to 29. It is possible to increase the reading capacity of the machine to 37 digits if the number of digits to be punched is reduced to 13.

Group operations of arithmetic processes may be performed with the Type 604; numerical factors may be accumulated from a group of cards and the final results punched on the last card of a group. Data may also be gang-punched on cards at the same time as results of arithmetic calculations are punched on the card. Numerical calculations and the punching of data on cards are automatically checked by the machine.

9-4-2. Remington Rand Machines. The principal Remington Rand machines useful in performing scientific calculations are the Sorter, the Interfiling Reproducing Punch, the Tabulator, and the Printing Multiplier Punch. These machines all operate on a mechanical basis; the sensing mechanism is not a brush contact but is an actual mechanical contact. Table 9-3 represents the distribution of functions among the several Remington Rand machines under consideration. The following is a brief

TABLE 9-3

Remington Rand Machine	Count	Sort	Consult tables	Add	Subtract	Multiply	Divide
Sorter	x	x					
Interfiling Re-producing Punch	x		x				
Tabulator	x			x	x		
Printing Mul-tiplier Punch				x	x	x	x

description of each of the machines listed in the table above:

The Sorter. The Sorter mechanically selects and arranges cards in any required order. The cards may be selected and sorted into 12 receiving magazines. This selection is made by

using 12 selecting pins, one each for the 12 positions of the column. Each of the 12 receiving magazines has an individual counter with a capacity of 9,999. In addition, the sorter has a total counting register which counts the number of cards going into all the receiving magazines. The speed of the counting sorter is 420 cards per minute for 45-column cards, or 250 cards per minute for 90-column cards. Each receiving magazine has a capacity of 350 to 375 cards and the total feeding-magazine capacity is 600 cards.

The Interfiling Reproducing Punch. The function of this machine is to feed two files of cards from two feeding magazines so as to (1) compare the two files, (2) punch the cards fed by the lower magazine subject to an established comparison with the cards in the upper feeding magazine, and (3) interfile selected cards from both files and segregate certain of the cards from both files according to an established comparison. There are four card-feeding magazines, two upper feeding magazines for pattern cards and two lower feeding magazines for cards to be punched. There are five card-receiving magazines, two upper magazines for the segregation of cards from the upper feeding magazine and two lower receiving magazines for segregating the cards from the lower feeding magazine. A third lower magazine is used for the interfiling of cards fed from both the upper and lower feeding magazines.

The speed of operation of the Interfiling Reproducing Punch is 100 cards per minute. Each feeding magazine has a capacity of 600 cards; each receiving magazine has a capacity of 750 cards.

The Tabulator. The principal function of the Tabulator is to translate and print numerical and alphabetical information punched in tabulating cards. The Model 3 90-column Tabulator will perform automatically the operations of addition, subtraction, and printing; it will also produce totals and/or grand totals. Its operation is fully automatic once the cards are inserted into the feeding mechanism. Although there are a number of models in manufacture, probably the one of interest scientifically is the Type 3100 Alphabetic Tabulator which has 100 printing sectors: 50 sectors each equipped with 10 numerical characters and 26 alphabetical characters, and 50 sectors each equipped to print 10 numerical characters only. Such tabulators will operate at

100 cards per minute and have a feeding-magazine capacity of 625 cards and a receiving-magazine capacity of 850 cards. Data are printed on a standard 20-inch carriage, the 100 sectors occupying a length of about $15\frac{5}{8}$ inches across the sheet.

The Printing Multiplier Punch. This machine is capable of performing addition, subtraction, multiplication, and division with a capacity for printing totals of all products up to 14 digits. This machine has a capacity for either 45 or 90 columns interchangeably; results are printed on paper tape or may be punched on cards. The punch is also equipped with both automatic and manual control. A keyboard allows an operator to set up manually one or both factors which are to be handled arithmetically or whose results are to be punched on cards or printed on the recording tape. The use of wiring units provides the machine with flexibility in setting up various types of problems. With this punch, it is possible to print and punch information on a summary card, to print progressive and final totals, and to interject amounts to be subtracted from these totals or added to them with a printed registration of these operations being made. The speed of the machine for six-by-six-digit multiplication is 1,200 calculations per hour. The product capacity is 12 digits, and the total capacity for printing is 14 digits per row.

9-5. Computational Operations on IBM Machines

On page 153, a table showing which IBM machines could be used to carry out each of seven basic operations was presented, and a brief description of these machines followed. Thus, Sec. 9-4-1 analyzed Table 9-2 more or less horizontally. This section will give an equally brief outline of the table vertically, according to the operations. IBM equipments will be discussed here to illustrate these operations.

9-5-1. Counting. The number of cards in which chosen punches appear in a particular column can most rapidly be determined by running the cards through an ordinary counting sorter. For this the sorter would be set to retain the cards in their original order, and the 15 counter dials would give the number of cards for which each of 12 possible punches was detected, the number for which no punch was detected, the subtotal, and the total. The tabulator in the larger machines can be wired to

count, but this operation is uneconomical in time and in the use of expensive machinery.

9-5-2. Sorting. The most rapid and convenient sorting of a badly disarranged deck is brought about through the use of the Sorter. In this device sorting is performed one column at a time at the rate of 450 cards per column per minute. A sort can be made on a number of digits by repeated use of this device. To do this, the initial sort is made on the lowest ranking digit. The cards are gathered from the pockets and stacked in order; they are then re-sorted on the next lowest ranking digit, etc. Thus, if 1,000 numbers from 000 through 999 were to be put in order with the lowest numbered card on top, the following sequence of operations would be carried out:

1. Sort on the right-hand column and gather the cards with those in the zero pocket on top, those in the one pocket next, etc.
2. Re-sort from the middle column and gather the cards with those in the zero pocket on top, the one pocket next, etc.
3. Sort on the left column and gather the cards with those in the zero pocket on top, the one pocket next, etc.

This completes the sort. The total sorting time will have been $6\frac{2}{3}$ minutes for the three passages of 1,000 cards at the rate of 450 cards per minute. For decks which are partially sorted, or for the merging of two sorted decks into a single deck in order, the Collator may be more efficient than the Sorter. The conditions under which efficient sorts can be made on the Collator are fairly apparent from the specifications on the machine, and no guidance to its use more than a familiarity with its operations seems necessary.

Block sorting can be performed on the highest digit if sorting above a given digit is required. In this case, time is saved over the lowest digit sort, since there are fewer cards involved in the sorting operation.

9-5-3. Table Consulting. To consult tables, the easiest procedure is to insert two sorted decks into the Collator. This procedure has already been outlined on pages 154–155 under the Electric Punched-card Collator. The Sorter, also, can be used for consulting tables.

9-5-4. Addition and Subtraction.   Addition and subtraction are most economically carried out on the Tabulator, except where combined with multiplication or division or where results are recorded on the same card.   In this case, the Calculating Punch may be profitably used.   In this operation the Tabulator is a fairly versatile machine and is even capable of carrying out some multiplications efficiently.   A simple automatic control program can be introduced through the control brushes of the tabulator, using either special control cards merged into the final deck through a Collator or the information on the original deck.

9-5-5. Multiplication.   Except for a few simple programs in multiplication which can be carried out effectively on the Tabulator, multiplication requires the use of either the Electric Multiplier or the Calculating Punch.   In these devices the computation is carried out straightforwardly as described in connection with the descriptions of the machines.

9-5-6. Division.   Division must be carried out on the Calculating Punch unless a method of successive approximation is used.   Various schemes of obtaining the reciprocal of a number without actually dividing may be applied to other machines. These schemes vary from table consulting, which may be carried out on the Collator, to use of a successive approximation by a method described in Chap. 7, which involves a somewhat similar table of reciprocals combined with a multiplication operation. Thus, it is true that well-equipped punched-card-machinery installations which do not contain a Calculating Punch may still carry out efficiently those computations involving division.

9-6. Transcription Devices

Transcription of results for reuse in other machines is brought about through reproducing or summary punches of various kinds.   The use of the summary punch with the Tabulator has already been described above.   Other punches permit rapid reproduction of a deck of cards with modification in the position of the punched information, with insertion of gang-punched information, representing several specific functions of these punches.   One device, the Type 513 Electric Card Reproducing Punch, can be used as a gang punch, a reproducing punch, and

as a summary punch. Through temporary wiring on a standard plugboard, considerable rerouting of information is possible in this machine, together with enough rearrangement during the data transcription to permit efficient reentry into the other calculating equipments.

This punch operates at a fixed speed of 100 cards per minute for reproducing and gang-punching operations. Summary punching from an accounting machine requires 1.2 seconds per card.

This machine contains a comparing feature which provides assurance that the punched card agrees with the original from which it was punched.

Other transcription machinery includes ordinary manual punches with duplicating attachments. These will reproduce a card already punched using the same equipment that is used in the manual punching. Some of these machines type the punched information across the top of the card for easy visual reading; this is a service which is not provided by the ordinary reproducing punch. On the other hand, their operation is much slower than the reproducing punch described earlier.

The information punched on a card can be printed across the edge of the card by means of punched-card interpreters. Numeric machines will print the information, punched in any four to five columns, at the rate of 4,500 cards per hour. Another machine will print 60 characters, either numerical or alphabetical, at the rate of 3,600 cards per hour.

Other transcription devices have already been mentioned. These include the printing tabulator and the machines for transcribing information on tape.

9-7. Acquisition of Equipment

Machines for punched-card computation are available on a lease or purchase basis from Remington Rand, Incorporated, and on lease from, but not generally for sale by, the International Business Machines Corporation. This corporation usually gives particular consideration to individual problems occurring at each proposed installation. The corporation maintains offices in most principal cities.

For computation problems which are of too infrequent occur-

rence to justify the installation of an adequate punched-card facility, the International Business Machines Corporation maintains complete machine services at several of its offices. Here machinery is available to carry out many computations, and the corporation accepts these on a job basis. Bids for such services will be made by the corporation on the basis of price per cards used.

Tables and many mathematical functions are available in punched-card editions from several sources. Information concerning these tables may be obtained from the International Business Machines Corporation or the Watson Scientific Computing Laboratory. Information concerning other published tables of mathematical functions is available in periodical form.[3]

9-8. Mathematical Operations to Which IBM Equipment Is Applicable

The International Business Machines Corporation has listed the following operations as typical of mathematical operations to which the equipment has been successfully applied:

1. Construction of mathematical tables.
2. Differencing, to any order desired.
3. Interpolation, using first-, second-, and higher order differences.
4. Numerical integration.
5. Harmonic analysis.
6. Solution of simultaneous equations.
7. Summation of series.
8. Multiplication of series.
9. Matrix and determinant calculations.
10. Method of least squares.
11. Differential equations.
12. Fourier analysis.
13. Successive approximation methods.
14. Curve fitting.

More specific information concerning actual applications which have been made may be obtained by examining the accompanying bibliography[2] on the applications of punched-card machinery to scientific research and statistics.

9-9. IBM Card-programmed Electronic Calculator

Since the foregoing sections of this chapter were written the International Business Machines Corporation has announced a new calculator which combines several IBM units in a sequence-controlled machine well adapted for scientific and engineering work.

It uses the Type 402 or the Type 417 Accounting Machine as the master control unit and printer. The program for a calculation is contained on a deck of cards which are fed into this Accounting Machine. Each card carries an eight-digit instruction which specifies the locations in storage of the factors to be operated upon, the operation to be performed, and the disposition to be made of the result.

For internal storage, the calculator uses the 80 mechanical counters with which the accounting machine is normally equipped. A complete record of a calculation can be printed at the rate of 150 lines per minute.

Cards are used for external storage. A Calculator Punch, Type 521, is used to prepare these cards and also to record on cards the results of the calculations. A Supplemental Storage Unit, Type 941, provides storage for 16 10-digit signed numbers relayed to it from the Type 604 Electronic Calculator.

The Type 604 Electronic Calculator is used to perform the various arithmetic operations. The manufacturer states that additions or subtractions are performed at a speed in excess of 2,000 a second, multiplications or divisions at 86 a second, and that any of these arithmetic operations can be combined.

The units of which this card-programmed computing system is composed may be disconnected from each other and used individually to perform their normal accounting operations.

REFERENCES

1. Eckert, W. J., *Punched Card Methods in Scientific Computation* (Thomas J. Watson Astronomical Computing Bureau, Columbia University, 1940).
2. International Business Machines Corporation, *Bibliography on the Use of IBM Machines in Science, Statistics and Education* (International Business Machines Corporation, New York, 1950). This bibliography consists of technical articles that describe researches in the various fields of science, statistics, and education conducted with the aid of IBM machines, and machine methods useful for such research. It is a revision

of *The Use of IBM Machines in Scientific Research, Statistics, and Education*, which was published in 1947. Reproduction of Part A entitled Science and Part B entitled Statistics is made with the permission of the IBM Corporation.

3. Lehmer, D. N., *Factor Stencils* (Carnegie Institution of Washington, Washington, D.C.; stencils revised and extended by J. D. Elder, 1939).
4. National Academy of Sciences, *Mathematical Tables and Other Aids to Computation*. Quarterly publication of the National Research Council, Washington, D.C.; first publication, January, 1943.

BIBLIOGRAPHY

Science

Aiken, Howard H., and Grace M. Hopper, The Automatic Sequence Controlled Calculator, *Electrical Engineering*, Vol. 65, No. 8–9, pp. 384–391 (August-September, 1946); No. 10, pp. 449–454 (October, 1946); No. 11, pp. 522–528 (November, 1946).

Akushsky, I. J., On Certain Schemes of the Numerical Harmonic Analysis, *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS (N.S.)*, Vol. 52, pp. 471–474 (1946) (Russian).

Akushsky, I. J., The Four-Counter Scheme of Solution of Dirichlet's Problem by Means of Punched-Card Machines, *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS (N.S.)*, Vol. 54, No. 8, pp. 659–662 (1946) (Russian).

Akushsky, I. J., Mathematical Tables on Punched Cards. Functions of One Variable, *Akademiia Nauk SSSR, Otdelenie Tekhnicheskikh Nauk*, No. 11, pp. 1405–1454 (1947) (Russian).

Akushsky, I. J., New Methods for the Calculation of a Sum of Products with a Tabulator, *Doklady Akademiia Nauk SSSR (N.S.)*, Vol. 60, No. 1, pp. 5–8 (1948) (Russian).

Akushsky, I. J., Numerical Solution of the Dirichlet Equation with the Aid of Perforated Card Machines, *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS (N.S.)*, Vol. 52, No. 5, pp. 375–378 (1946) (Russian).

Akushsky, I. J., On the Numerical Solution of Dirichlet Problem on Punched-Card Machines, *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS (N.S.)*, Vol. 54, No. 9, pp. 755–758 (1946) (Russian).

Akushsky, I. J., The Process of Diagonal Summation on a Tabulator and Some of Its Applications, *Akademiia Nauk SSSR, Bulletin, Classe des Sciences Techniques*, No. 5, pp. 475–495 (1947) (Russian).

Akushsky, I. J., Some New Methods of Calculating Sums of Products on a Tabulator, *Akademiia Nauk SSSR, Otdelenie Tekhnicheskikh Nauk*, No. 8, pp. 1193–1228 (1948) (Russian).

Alt, Franz L., Multiplication of Matrices, *Mathematical Tables and Other Aids to Computation*, Vol. 2, No. 13, pp. 12–13 (January, 1946).

Atanasoff, J. V., and A. E. Brandt, Application of Punched Card Equipment to the Analysis of Complex Spectra, *Journal of the Optical Society of America*, Vol. 26, No. 2, pp. 83–88 (February, 1936).

Bailar, John C., Jr., Karl F. Heumann, and Edwin J. Seiferle, The Use of Punched Card Techniques in the Coding of Inorganic Compounds, *Journal of Chemical Education*, Vol. 25, No. 3, pp. 142–143, 176 (March, 1948).

Beevers, C. A., A Machine for the Rapid Summation of Fourier Series, *Proceedings of the Physical Society* (London), Vol. 51, Part 4, pp. 660–663; discussion, pp. 664–667 (July 1, 1939).

Bell, William D., Punched Card Techniques for the Solution of Simultaneous Equations and Other Matrix Operations, *Proceedings, Scientific Computation Forum*, pp. 28–31 (International Business Machines Corporation, New York, 1948).

Bell, William D., A Simplified Punch-Card Approach to the Solution of the Flutter Determinant, *Journal of the Aeronautical Sciences*, Vol. 15, No. 2, pp. 121–122 (February, 1948).

Belzer, Jack, George Gamow, and Geoffrey Keller, Dynamics of Elliptical Galaxies, *Proceedings, Scientific Computation Forum*, pp. 67–69 (International Business Machines Corporation, New York, 1948).

Bergman, Stefan, Construction of a Complete Set of Solutions of a Linear Partial Differential Equation in Two Variables, by Use of Punch Card Machines, *Quarterly of Applied Mathematics*, Vol. 4, No. 3, pp. 233–245 (October, 1946).

Bergman, Stefan, Punch-Card Machine Methods Applied to the Solution of the Torsion Problem, *Quarterly of Applied Mathematics*, Vol. 5, No. 1, pp. 69–81 (April, 1947).

Berkeley, Edmund C., Giant Brains, or Machines that Think (John Wiley and Sons, Inc., New York, 1949). 270 pp.

Blanch, Gertrude, Differencing on the Type 405 Accounting Machine, *Proceedings, Scientific Computation Forum*, pp. 14–22 (International Business Machines Corporation, New York, 1948).

Bower, Ernest Clare, On Subdividing Tables, *Lick Observatory Bulletin*, Vol. 16, No. 455, pp. 143–144, November, 1933.

Bower, Ernest Clare, Systematic Subdivision of Tables, *Lick Observatory Bulletin*, Vol. 17, No. 467, pp. 65–74 (April, 1935).

Bramble, Clinton C., Applications of Punched Card Equipment at the Naval Proving Ground, *Proceedings, Scientific Computation Forum*, pp. 99–100 (International Business Machines Corporation, New York, 1948).

Brinkley, Stuart, R., Jr., and Robert W. Smith, Jr., Calculation of the Equilibrium Composition of Systems of Many Constituents, *Proceedings, Scientific Computation Forum*, pp. 77–82 (International Business Machines Corporation, New York, 1948).

Brouwer, Dirk, On the Accumulation of Errors in Numerical Integration, *Astronomical Journal*, Vol. 46, No. 16, pp. 149–153 (October, 1937).

Brouwer, Dirk, The Reference System with a View to Planetary Dynamics, *Annals of the New York Academy of Sciences*, Vol. 42, Article 2, pp. 133–149 (October 15, 1941).

Comrie, L. J., The Application of Commercial Calculating Machines to Scientific Computing, *Mathematical Tables and Other Aids to Computation*, Vol. 2, No. 16, pp. 149–159 (October, 1946).

Comrie, L. J., The Application of the Hollerith Tabulating Machine to Brown's Tables of the Moon, *Monthly Notices of the Royal Astronomical Society*, Vol. 92, No. 7, pp. 694–707 (May, 1932).

Comrie, L. J., On the Construction of Tables by Interpolation, *Monthly Notices of the Royal Astronomical Society*, Vol. 88, No. 6, pp. 506–523 (April, 1928).

Comrie, L. J., Interpolation and Allied Tables, *Nautical Almanac for* 1937 (London).

Comrie, L. J., Recent Progress in Scientific Computing, *Journal of Scientific Instruments*, Vol. 21, No. 8, pp. 129–135 (August, 1944).

Comrie, L. J., Tables for Interpolation to Tenths and Fifths by the End Figure Process, *Nautical Almanac for* 1931, pp. 828–835 (London).

Cox, E. G., R. J. J. H. Gillot, and G. A. Jeffrey, The Molecular Structure of Thiophthen from X-Ray Crystal Analysis, *Acta Crystallographica*, Vol. 2, Part 6, pp. 356–363 (December, 1949).

Cox, E. G., L. Gross, and G. A. Jeffrey, A Hollerith Punched-Card Method for the Evaluation of Electron Density in Crystal Structure Analysis, *Proceedings of the Leeds Philosophical and Literary Society*, Vol. 5, Part 1, pp. 1–13 (December, 1947).

Cox, E. G., L. Gross, and G. A. Jeffrey, A Hollerith Technique for Computing Three-Dimensional Differential Fourier Syntheses in X-Ray Crystal-Structure Analysis, *Acta Crystallographica*, Vol. 2, Part 6, pp. 351–355 (December, 1949).

Cox, E. G., L. Gross, and G. A. Jeffrey, Use of Punched Card Tabulating Machines for Crystallographic Fourier Synthesis, *Nature*, Vol. 159, No. 4039, pp. 433–434 (March 29, 1947).

Culley, Frank L., Use of Accounting Machines for Mass Transformation from Geographic to Military-Grid Coordinates, *Transactions of the American Geophysical Union* (National Research Council, Washington, D.C., November, 1942), Part 2, pp. 190–197.

Curtiss-Wright Corporation, *Punched Card Method in Structural Calculations*, Report No. V-241-S-4 (Curtiss-Wright Corporation, Airplane Division, Buffalo, N.Y., March 23, 1945).

Domb, C., On Iterative Solutions of Algebraic Equations, *Proceedings of the Cambridge Philosophical Society*, Vol. 45, No. 2, pp. 237–240 (April, 1949).

Donohue, Jerry, and Verner Schomaker, The Use of Punched Cards in Molecular Structure Determinations. III. Structure-Factor Calculations of X-Ray Crystallography, *Acta Crystallographica*, Vol. 2, Part 6, pp. 344–347 (December, 1949).

Dreyer, H. J., *Solution of Systems of Linear Equations by Means of Punched-Card Machines*, No. F-TS-1046-RE (Headquarters Air Materiel Command, Wright Field, Dayton, Ohio, December, 1946).   13 pp.

Duncan, David B., and John F. Kenney, *On the Solution of Normal Equations and Related Topics* (Edwards Bros., Inc., Ann Arbor Mich., 1946).   35 pp.

Dunstan, Lyle A., The General Solution Method of Power Network Analysis, *American Institute of Electrical Engineers Technical Paper* 48–103 (February, 1948).   19 pp.

Dunstan, Lyle A., Machine Computation of Power Network Performance,

American Institute of Electrical Engineers Technical Paper 47–87 (December, 1946).   33 pp.

Dwyer, Paul S., The Solution of Simultaneous Equations, *Psychometrika*, Vol. 6, No. 2, pp. 101–129 (April, 1941).

Dwyer, Paul S., The Use of Tables in the Form of Prepunched Cards, *Proceedings of the Educational Research Forum*; pp. 125–127 (International Business Machines Corporation, New York, August, 1940).

Earthquake Calculations Are Made by Punched Card Machines, *Engineering News Record*, Vol. 141, No. 12, p. 110 (September 16, 1948).

Eckert, W. J., Calculating Machines, *Encyclopaedia Britannica*, 13th Edition (1947), Vol. 4, pp. 548–554.

Eckert, W. J., The Computation of Special Perturbations by the Punched Card Method, *Astronomical Journal*, Vol. 44, No. 20, pp. 177–182 (October, 1935).

Eckert, W. J., Electronic and Electromagnetic Measuring, Computing and Recording Devices, *Centennial Symposia, December*, 1946, Harvard Observatory Monographs, No. 7, pp. 169–178 (1948).

Eckert, W. J., Electrons and Computation, *Scientific Monthly*, Vol. 67, No. 5, pp. 315–323 (November, 1948).

Eckert, W. J., The IBM Pluggable Sequence Relay Calculator, *Mathematical Tables and Other Aids to Computation*, Vol. 3, No. 23, pp. 149–161 (July, 1948).

Eckert, W. J., *Punched Card Methods in Scientific Computation*, (Thomas J. Watson Astronomical Computing Bureau, Columbia University, 1940). 136 pp.   (New edition in preparation.)

Eckert, W. J., Punched-Card Techniques and Their Applications to Scientific Problems, *Journal of Chemical Education*, Vol. 24, No. 2, pp. 54–57, 74 (February, 1947).

Eckert, W. J., and Ralph F. Haupt, The Printing of Mathematical Tables, *Mathematical Tables and Other Aids to Computation*, Vol. 2, No. 17, pp. 197–202 (January, 1947).

Feinstein, Lillian, and Martin Schwarzschild, Automatic Integration of Linear Second-Order Differential Equations by Means of Punched Card Machines, *Review of Scientific Instruments*, Vol. 12, No. 8, pp. 405–408 (August, 1941).

Fenn, George S., Programming and Using the Type 603-405 Combination Machine in the Solution of Differential Equations, *Proceedings, Scientific Computation Forum*, pp. 95–98 (International Business Machines Corporation, New York, 1948).

Ferber, Ben, Planning Engineering Calculations for IBM Equipment, *Proceedings, Scientific Computation Forum*, pp. 52–53 (International Business Machines Corporation, New York, 1948).

Ferris, Lorna, Kanardy Taylor, J. W. Perry, and Maria E. W. Torok, *Bibliography on the Uses of Punched Cards*, (American Chemical Society, Washington, D.C., 1949).   40 pp.

Fletcher, A. J., C. P. Miller, and L. Rosenhead, *An Index of Mathematical Tables* (McGraw-Hill Book Company, Inc., New York, 1946).   450 pp.

Frost, Arthur A., The Approximate Solution of Schrödinger Equations by a

Least Squares Method, *Journal of Chemical Physics*, Vol. 10, No. 4, pp. 240–245 (April, 1942).

George, M. C., An Annotated Bibliography of Some Early Uses of Punched Cards in Meteorology and Climatology, *Bulletin of the American Meteorological Society*, Vol. 26, No. 3, pp. 76–85 (March, 1945).

Gordon, Norman J., Extracting Land-Use Data from Punched Cards, *American City*, Vol. 63, p. 149 (June, 1948).

Grems, M. D., and J. S. Kasper, An Improved Punched-Card Method for Crystal Structure-Factor Calculations, *Acta Crystallographica*, Vol. 2, Part 6, pp. 347–351 (December, 1949).

Grosch, H. R. J., Harmonic Analysis by the Use of Progressive Digiting, *Proceedings of the Research Forum*, pp. 81–84 (International Business Machines Corporation, New York, 1946).

Grosch, H. R. J., *Multiplication of Small Matrices* (International Business Machines Corporation, New York, June 1, 1949). 3 pp.

Grosch, H. R. J., The Use of Optimum Internal Mathematical Tables, *Proceedings, Scientific Computation Forum*, pp. 23–27 (International Business Machines Corporation, New York, 1948).

Grosch, H. R. J. (Editor), Proceedings, Scientific Computation Forum 1948.

Grossman, D. P., The Application of Punched Card Machines to the Solution of a System of Linear Algebraic Equations by the Iteration Method, *Akademiia Nauk SSSR, Otdelenie Tekhnicheskikh Nauk*, No. 8, pp. 1229–1238 (August, 1948) (Russian).

Gull, Cloyd Dake, A Punched Card Method for the Bibliography, Abstracting, and Indexing of Chemical Literature, *Journal of Chemical Education*, Vol. 23, No. 10, pp. 500–507 (October, 1946).

Hainer, R. M., and Gilbert W. King, The Asymmetric Rotor. V. An Analysis of the 3.7-μ Band of H_2S by Punched-Card Techniques, *Journal of Chemical Physics*, Vol. 15, No. 2, pp. 89–91 (February, 1947).

Harrison, Joseph O., Jr., Piecewise Polynomial Approximation for Large-Scale Digital Calculators, *Mathematical Tables and Other Aids to Computation*, Vol. 3, No. 26, pp. 400–407 (April, 1949).

Hartkemeier, Harry Pelle, and Herman E. Miller, Obtaining Differences from Punched Cards, *Journal of the American Statistical Association*, Vol. 37, No. 218, pp. 285–287 (June, 1942).

Hartree, Douglas R., *Calculating Instruments and Machines* (The University of Illinois Press, Urbana, Ill., 1949). 138 pp.

Hartree, Douglas R., Notes on Iterative Processes, *Proceedings of the Cambridge Philosophical Society*, Vol. 45, No. 2, pp. 230–236 (April, 1949).

Harvard Computation Laboratory, A Manual of Operation for the Automatic Sequence Controlled Calculator, *The Annals of the Computation Laboratory of Harvard University*, Vol. 1 (Harvard University Press, Cambridge, Mass., 1946). 561 pp.

Harvard Computation Laboratory, Proceedings of a Symposium on Large-Scale Digital Calculating Machinery, *The Annals of the Computation Laboratory of Harvard University*, Vol. 16 (Harvard University Press, Cambridge, Mass., 1948). 302 pp.

Harvard Computation Laboratory, Tables of the Bessel Functions of the

First Kind, *The Annals of the Computation Laboratory of Harvard University*, Vols. 3–12 (Harvard University Press, Cambridge, Mass., 1947–1949).

Harvard Computation Laboratory, Tables for the Design of Missiles, *The Annals of the Computation Laboratory of Harvard University*, Vol. 17 (Harvard University Press, Cambridge, Mass., 1948). 226 pp.

Harvard Computation Laboratory, Tables of Generalized Sine- and Cosine-Integral Functions, *The Annals of the Computation Laboratory of Harvard University*, Vols. 18–19 (Harvard University Press, Cambridge, Mass., 1949).

Harvard Computation Laboratory, Tables of Inverse Hyperbolic Functions, *The Annals of the Computation Laboratory of Harvard University*, Vol. 20 (Harvard University Press, Cambridge, Mass., 1949).

Harvard Computation Laboratory, Tables of the Modified Hankel Functions of Order One-Third and of Their Derivatives, *The Annals of the Computation Laboratory of Harvard University*, Vol. 2 (Harvard University Press, Cambridge, Mass., 1945). 235 pp.

Hausman, L. F. and M. Schwarzschild, Automatic Integration of Linear Sixth-Order Differential Equations by Means of Punched-Card Machines, *Review of Scientific Instruments*, Vol. 18, No. 12, pp. 877–883 (December, 1947).

Herget, Paul, Integration of the Differential Equation $\frac{d^2P}{dr^2} = P \cdot F(r)$ Using the Type 601 Multiplying Punch, *Proceedings, Scientific Computation Forum*, pp. 39–41 (International Business Machines Corporation, New York, 1948).

Herget, Paul, Numerical Integration with Punched Cards, *Astronomical Journal*, Vol. 52, No. 4, pp. 115–117 (November, 1946).

Herget, Paul, and G. M. Clemence, Optimum-Interval Punched-Card Tables, *Mathematical Tables and Other Aids to Computation*, Vol. 1, No. 6, pp. 173–176 (April, 1944).

Hodgson, M. L., C. J. B. Clews, and W. Cochran, A Punched-Card Modification of the Beevers-Lipson Method of Fourier Synthesis, *Acta Crystallographica*, Vol. 2, Part 2, pp. 113–116 (April, 1949).

Hollander, Frederick H., Punched Card Calculating and Printing Methods in the Nautical Almanac Office, *Proceedings, Scientific Computation Forum*, pp. 83–94 (International Business Machines Corporation, New York, 1948).

Hollerith, H., The Electrical Tabulating Machine, *Journal of the Royal Statistical Society*, Vol. 57, Part 4, pp. 678–682 (December, 1894).

Hollerith, H., An Electric Tabulating System, *School of Mines Quarterly* (Columbia University), Vol. 10, No. 3, pp. 238–255 (April, 1889).

Horton, H. Burke, and R. Tynes Smith, III, A Direct Method for Producing Random Digits in Any Number System, *Annals of Mathematical Statistics*, Vol. 20, No. 1, pp. 82–90 (March, 1949).

Hotelling, Harold, Some New Methods in Matrix Calculation, *Annals of Mathematical Statistics*, Vol. 14, No. 1, pp. 1–34 (March, 1943).

Hotelling, Harold, Further Points on Matrix Calculation and Simultaneous Equations, *Annals of Mathematical Statistics*, Vol. 14, No. 4, pp. 440–441 (December, 1943).

Howe, H. Herbert, Computing Magnetic Observatory Results with Punched Cards, *U.S. Coast and Geodetic Survey*, Document No. T-113 (July 6, 1948).

Hurd, Cuthbert C., IBM Applications in Industrial Statistics, *Proceedings, Scientific Computation Forum*, pp. 45–48 (International Business Machines Corporation, New York, 1948).

International Business Machines Corporation, *IBM Automatic Sequence Controlled Calculator* (International Business Machines Corporation, New York, 1945). 6 pp.

International Business Machines Corporation, *IBM Selective Sequence Electronic Calculator*, Form 52–3927-0 (International Business Machines Corporation, New York, 1948). 16 pp.

International Business Machines Corporation, *IBM Survey Computer*, Form 50-3885-0 (International Business Machines Corporation, New York, October, 1947). 7 pp.

Jennings, Philip D., and George E. Quinan, The Use of Business Machines in Determining the Distribution of Load and Reactive Components in Power Line Networks, *Transactions of the American Institute of Electrical Engineers*, Vol. 65, pp. 1045–1046 (July, 1946).

Johnson, Curt I., *A New Approach to the Design of Dynamically Loaded Extension and Compression Springs, American Society of Mechanical Engineers*, Paper No. 48-SA-23 (1948). 18 pp.

Johnston, Herrick L., Lydia Savedoff and Jack Belzer, *Contributions to the Thermodynamic Functions by a Planck-Einstein Oscillator in One Degree of Freedom* (Office of Naval Research, Department of the Navy, Washington, D.C., July, 1949). 159 pp.

Kimball, Everett, Jr., *A Fundamental Punched Card Method for Technical Computations* (Bureau of the Census, Washington, D.C., undated). 12 pp.

Kimball, Everett, Jr., *A Method of Technical Computations by Punched Card Equipment* (Bureau of the Census, Washington, D.C., undated). 8 pp.

King, Gilbert W., Application of Punched Cards in Physical Chemistry, *Proceedings, Scientific Computation Forum*, pp. 70–73 (International Business Machines Corporation, New York, 1948).

King, Gilbert W., Some Applications of Punched-Card Methods in Research Problems in Chemical Physics, *Journal of Chemical Education*, Vol. 24, No. 2, pp. 61–64 (February, 1947).

King, Gilbert W., The Asymmetric Rotor. IV. An Analysis of the 8.5-μ Band of D_2O by Punched Card Techniques, *Journal of Chemical Physics*, Vol. 15, No. 2, pp. 85–88 (February, 1947).

King, Gilbert W., A Method of Plotting on Standard IBM Equipment, *Mathematical Tables and Other Aids to Computation*, Vol. 3, No. 25, pp. 352–355 (January, 1949).

King, Gilbert W., Punched Card Tables of the Exponential Function, *Review of Scientific Instruments*, Vol. 15, No. 12, pp. 349–350 (December, 1944).

King, Gilbert W., Paul C. Cross, and George B. Thomas, The Asymmetric Rotor. III. Punched-Card Methods of Constructing Band Spectra, *Journal of Chemical Physics*, Vol. 14, No. 1, pp. 35–42 (January, 1946).

Kintas, John, A Survey of the IBM Project at Beech Aircraft Corporation,

Proceedings, Scientific Computation Forum, pp. 54–59 (International Business Machines Corporation, New York, 1948).

Knudsen, Lila F., A Punched Card Technique to Obtain Coefficients of Orthogonal Polynomials, *Journal of the American Statistical Association*, Vol. 37, No. 220, pp. 496–506 (December, 1942).

Kormes, Jennie P., and Mark Kormes, Numerical Solution of Initial Value Problems by Means of Punched-Card Machines, *Review of Scientific Instruments*, Vol. 16, No. 1, pp. 7–9 (January, 1945).

Kormes, Mark, Numerical Solution of the Boundary Value Problem for the Potential Equation by Means of Punched Cards, *Review of Scientific Instruments*, Vol. 14, No. 8, pp. 248–250 (August, 1943).

Kraft, Hans, Aerodynamic Lattice Calculations Using Punched Cards, *Proceedings, Scientific Computation Forum*, pp. 60–66 (International Business Machines Corporation, New York, 1948).

Laderman, Jack, and Milton Abramowitz, Application of Machines to Differencing of Tables, *Journal of the American Statistical Association*, Vol. 41, pp. 233–237 (June, 1946).

Leavens, Dickson H., *The Digit or Group Extension Method of Multiplication* (Harvard Business School, BS475, undated). 6 pp.

Lehmer, D. N., *Factor Stencils*, revised and extended by J. D. Elder (Carnegie Institution, Washington, D.C., 1939). 27 pp. + 2135 stencils.

Leppert, E. L., Jr., An Application of IBM Machines to the Solution of the Flutter Determinant, *Journal of the Aeronautical Sciences*, Vol. 14, No. 3, pp. 171–174 (March, 1947).

Levin, Joseph H., Use of the IBM Relay Calculators for Technical Calculations at Aberdeen Proving Ground, *Proceedings, Scientific Computation Forum*, pp. 101–104 (International Business Machines Corporation, New York, 1948).

Lindberger, N. Arne, Integration of Second Order Linear Differential Equations on the Type 602 Calculating Punch, *Proceedings, Scientific Computation Forum*, pp. 34–38 (International Business Machines Corporation, New York, 1948).

Lorant, Ralph R., *Evaluation of Determinants with Complex Elements by Means of the IBM Punched Card Method*, (International Business Machines Corporation, New York, June 1, 1946). 16 pp.

Lynch, J., and C. E. Johnson, *Programming Principles for the IBM Relay Calculators*, Report No. 705 (Ballistic Research Laboratories, Aberdeen Proving Ground, Aberdeen, Md., October, 1949). 95 pp.

McPherson, J. C., Mathematical Operations with Punched Cards, *Journal of the American Statistical Association*, Vol. 37, No. 218, pp. 275–281 (June, 1942).

McPherson, J. C., On Mechanical Tabulation of Polynomials, *Annals of Mathematical Statistics*, Vol. 12, No. 3, pp. 317–327 (September, 1941).

McPherson, J. C., *Solution of Simultaneous Equations by Punched Card Methods* (International Business Machines Corporation, New York, undated). 5 pp.

Maginniss, Frank J., Some Engineering Applications of IBM Equipment at the General Electric Company, *Proceedings, Scientific Computation Forum*,

pp. 49–51 (International Business Machines Corporation, New York, 1948).

Martin, Louis D., *A Simplified Fine-Pitch Worm Gearing Standard*, Paper No. 48-A-82 (American Society of Mechanical Engineers, New York, 1948).   14 pp.

Massachusetts Institute of Technology, Department of Electrical Engineering, Center of Analysis, *Tables of Supersonic Flow Around Cones*, Technical Report No. 1 (Cambridge, Mass., 1947).   555 pp.

Massachusetts Institute of Technology, Department of Electrical Engineering, Center of Analysis, *Tables of Supersonic Flow Around Yawing Cones*, Technical Report No. 3 (Cambridge, Mass., 1947).   321 pp.

Massachusetts Institute of Technology, Department of Electrical Engineering, Center of Analysis, *Tables of Supersonic Flow Around Cones of Large Yaw*, Technical Report No. 5 (Cambridge, Mass., 1949).   125 pp.

Mayor, Yoland, Applying Statistical Machines to Bibliographical Investigations, *Chimie et Industrie*, Vol. 35, No. 2, pp. 458–466 (February, 1936) (French).

Milne, William Edmund, Numerical Calculus (Princeton University Press, Princeton, N.J., 1949).   393 pp.

Morgan, John A., and D. E. H. Frear, Coding and Sorting Chemical Compounds by Means of Punched Cards, *Journal of Chemical Education*, Vol. 24, No. 2, pp. 58–61 (February, 1947).

Mulligan, J. H., Jr., The Effect of Pole and Zero Locations on the Transient Response of Linear Dynamic Systems, *Proceedings of the Institute of Radio Engineers*, Vol. 37, No. 5, pp. 516–529 (May, 1949); No. 10, p. 1181 (October, 1949).

Munk, Max M., Use of Automatic Calculation Methods in Aeronautical Research, *Aero Digest*, Vol. 50, No. 5, pp. 68–69, 192 (September 1, 1945).

Murray, Francis J., Simultaneous Linear Equations, *Proceedings, Scientific Computation Forum*, pp. 105–106 (International Business Machines Corporation, New York, 1948).

National Bureau of Standards, Bibliography on Automatic Digital Computing Machinery (U. S. Department of Commerce, Washington, D.C., January, 1948).   10 pp.

Nautical Almanac Office, The American Air Almanac, (U.S. Government Printing Office, Washington, D.C., January-April, 1942).   240 pp.

Nautical Almanac Office, Tables of Sunrise, Sunset and Twilight, *Supplement to the American Ephemeris*, 1946 (U.S. Government Printing Office, Washington, D.C., 1945).   196 pp.

Nowacki, W., Die Verwendung von Lochkartenmaschinen zur Fourier- und Patterson-Synthese von Kristallen, *Chimia*, Vol. 2, No. 12, pp. 3–7 (May 21, 1948) (German).

Oakley, Margaret B., and George E. Kimball, Punched Card Calculation of Resonance Energies, *Journal of Chemical Physics*, Vol. 17, No. 8, pp. 706–717 (August, 1949).

Peakes, Gilbert L., Report Indexing by Punch Cards, *Journal of Chemical Education*, Vol. 26, No. 3, pp. 139–146 (March, 1949).

Peebles, Glenn H., A Method for Calculating Airfoil Sections from Specifica-

tions on the Pressure Distributions, *Journal of the Aeronautical Sciences*, Vol. 14, No. 8, pp. 451–456 (August, 1947).

Polachek, Harry, Computation of Shock Wave Refraction on the Selective Sequence Electronic Calculator, *Proceedings, Scientific Computation Forum*, pp. 107–122 (International Business Machines Corporation, New York, 1948).

Rankin, A. W., Calculation of the Multiple-Span Critical Speeds of Flexible Shafts by Means of Punched-Card Machines, Paper No. 45-A-46 (Turbine Generator Engineering Division, General Electric Corporation, Schenectady, N.Y., undated).   10 pp.

Renner, H. W., *Solving Simultaneous Equations through the Use of IBM Electric Punched Card Accounting Machines* (International Business Machines Corporation, Endicott, N.Y., 1946).   6 pp.

Reynolds, William A., A Prepunched Master Deck for the Computation of Square Roots on IBM Electrical Accounting Equipment, *Psychometrika*, Vol. 11, No. 4, pp. 223–237 (1946).

Robinson, Raphael M., *Stencils for Solving $x^2 \equiv a \pmod m$* (University of California Press, Berkeley and Los Angeles, 1940).   14 pp. + 272 stencils.

Sandomire, Marion M., Accumulating Cubes with Punch Cards, *Journal of the American Statistical Association*, Vol. 36, No. 216, pp. 507–514 (December, 1941).

Savedoff, Lydia G., Jack Belzer, and Herrick L. Johnston, Application of Punched Card Methods to the Computation of Thermodynamic Properties of Gases from Spectra, *Proceedings, Scientific Computation Forum*, pp. 74–76 (International Business Machines Corporation, New York, 1948).

Savedoff, Lydia G., Jack Belzer, and Herrick L. Johnston, Computation of Thermodynamic Properties of Gases from Spectra with the Aid of Punched Card Equipment, *Volume of the Third Symposium on Combustion, Flame and Explosion Phenomena*, pp. 595–599 (University of Wisconsin, Madison, Wis., September, 1948).

Scarborough, James Blaine, Numerical Mathematical Analysis (Johns Hopkins Press, Baltimore, Md., 1930).   416 pp.

Schwarzschild, M., On Stellar Rotation. II, *Astrophysical Journal*, Vol. 106, No. 3, pp. 427–456 (November, 1947).

Shaffer, P. A., Jr., and Linus Pauling, The Use of Punched Cards in Molecular Structure Determinations. I. Crystal Structure Calculations, *Journal of Chemical Physics*, Vol. 14, No. 11, pp. 648–658 (November, 1946).

Shaffer, P. A., Jr., Verner Schomaker, and Linus Pauling, The Use of Punched Cards in Molecular Structure Determinations. II. Electron Diffraction Calculations, *Journal of Chemical Physics*, Vol. 14, No. 11, pp. 659–664 (November, 1946).

Shapley, A. H., The Recurrence Tendency and Forecasts of Magnetic Activity, *Transactions of the American Geophysical Union*, Vol. 28, No. 5, pp. 715–721 (October, 1947).

Shimizu, Tatsujiro, and Yoichi Katayama, Solutions of Non-Linear Equations by Punched-Card Methods, *Mathematical Japonicae*, Vol. 1, No. 2, 92–97 (1948).

Snedecor, George W., Uses of Punched Card Equipment in Mathematics,

American Mathematical Monthly, Vol. 35, No. 4, pp. 161–169 (April, 1928).

Sprinkle, Leland W., Two Numerical Methods of Integration Using Predetermined Factors, *Proceedings, Scientific Computation Forum*, pp. 32–33 (International Business Machines Corporation, New York, 1948).

Targoff, Walter P., The Associated Matrices of Bending and Coupled Bending-Torsion Vibrations, *Journal of the Aeronautical Sciences*, Vol. 14, No. 10, pp. 579–582 (October, 1947).

Targoff, Walter P., A Tabulation Method for the Calculation of the Critical Speed of Wing Divergence, *Journal of the Aeronautical Sciences*, Vol. 14, No. 8, pp. 487–488 (August, 1947).

Thomas, George B., and Gilbert W. King, Preparation of Punched-Card Tables of Logarithms, *Review of Scientific Instruments*, Vol. 15, No. 12, p. 350 (December, 1944).

Thomas, L. H., Computation of Statistical Fields for' Atoms and Ions, *Proceedings, Scientific Computation Forum*, pp. 123–128 (International Business Machines Corporation, New York, 1948).

Tucker, Ledyard R., *Factorial Rotation of Axes with Graphs Made by Machine*, Bibliofilm Manuscript (University of Chicago Library, 1941).

Tucker, Ledyard R., A Matrix Multiplier, *Psychometrika*, Vol. 5, No. 4, pp. 289–294 (December, 1940).

Verzuh, Frank M., Evaluation of Higher Order Differences on the Type 602 Calculating Punch, *Proceedings, Scientific Computation Forum*, pp. 9–13 (International Business Machines Corporation, New York, 1948).

Verzuh, Frank M., The Solution of Simultaneous Linear Equations with the Aid of the 602 Calculating Punch, *Mathematical Tables and Other Aids to Computation*, Vol. 3, No. 27, pp. 453–462 (July, 1949).

Von Neumann, John, and H. H. Goldstine, Numerical Inverting of Matrices of High Order, *Bulletin of the American Mathematical Society*, Vol. 53, No. 11, pp. 1021–1099 (November, 1947).

Walker, Robert M., An Analogoue Computer for the Solution of Linear Simultaneous Equations, *Proceedings of the Institute of Radio Engineers*, Vol. 37, No. 12, pp. 1467–1473 (December, 1949).

Whittaker, Edmund, and G. Robinson, *The Calculus of Observations* (Blackie & Son, Ltd., London, 1946).   397 pp.

Whitten, C. A., *Air-Line Distances Between Cities in the United States*, Special Publication No. 238 (U.S. Government Printing Office, Washington, D.C., 1947).   246 pp.

Whitten, C. A., Triangulation-Adjustment by International Business Machines, *American Geophysical Union Transactions of* 1943, Part 1, pp. 31–32 (National Research Council, Washington, D.C., October, 1943).

Womersley, J. R., Scientific Computing in Great Britain, *Mathematical Tables and Other Aids to Computation*, Vol. 2, No. 15, pp. 110–117 (July, 1946).

Statistics

Bartlett, Neil R., Punched Card Technique for Computing Means, Standard Deviations, and the Product Moment Correlation Coefficient, and for Listing Scattergrams, *Science*, Vol. 104, No. 2703, pp. 374–375 (October 18, 1946).

Benjamin, Kurt, An IBM Technique for the Computation of Σx^2 and Σxy, *Psychometrika*, Vol. 10, No. 1, pp. 61–67 (March, 1945).

Benjamin, Kurt, Problems of Multiple-Punching with Hollerith Machines, *Journal of the American Statistical Association*, Vol. 42, No. 237, pp. 46–71 (March, 1947).

Benjamin, Kurt, Statistical Procedures in the Hospital, *Proceedings, Hospital Financial Control Institute*, pp. 58–63 (International Business Machines Corporation, New York, August, 1948).

Berkson, Joseph, A Punch Card Designed to Contain Written Data and Coding, *Journal of the American Statistical Association*, Vol. 36, No. 216, pp. 535–538 (December, 1941).

Berkson, Joseph, A System of Codification of Medical Diagnoses for Application to Punch Cards with Plan of Operation, *American Journal of Public Health*, Vol. 26, pp. 606–612 (June, 1936).

Black, Bertram J., and Edward B. Olds, A Punched Card Method for Presenting, Analyzing and Comparing Many Series of Statistics for Areas, *Journal of the American Statistical Association*, Vol. 41, No. 235, pp. 347–355 (September, 1946).

Black-Schaffer, Bernard, and Paul D. Rosahn, Methods of Analysis of Yale Autopsy Protocols, Including a Code for the Punched Card Study of Syphilis, *Yale Journal of Biology and Medicine*, Vol. 15, pp. 575–586 (March, 1943).

Brandt, A. E., Use of Machine Factoring in Multiple Correlation, *Journal of the American Statistical Association*, Vol. 23, No. 163, pp. 291–295 (September, 1928).

Caddy, E. Reid, Punch Cards Simplify Statistics, *Modern Hospital*, Vol. 52, No. 5, pp. 82–84 (May, 1939).

Castore, George F., and William S. Dye, III, A Simplified Punch Card Method of Determining Sums of Squares and Sums of Products, *Psychometrika*, Vol. 14, No. 3, pp. 243–250 (September, 1949).

Cochran, W. J., Use of IBM Equipment in an Investigation of the "Truncated Normal" Problem, *Proceedings of the Research Forum*, pp. 40–44 (International Business Machines Corporation, New York, August, 1946).

Comrie, L. J., G. B. Hey, and H. G. Hudson, Application of Hollerith Equipment to an Agricultural Investigation, *Supplement to the Journal of the Royal Statistical Society*, Vol. 4, No. 2, pp. 210–224 (1937).

Deming, W. Edwards, and Leon Geoffrey, On Sample Inspection in the Processing of Census Returns, *Journal of the American Statistical Association*, Vol. 36, No. 215, pp. 351–360 (September, 1941).

Deming, W. Edwards, Benjamin J. Tepping, and Leon Geoffrey, Errors in Card Punching, *Journal of the American Statistical Association*, Vol. 37, No. 220, pp. 525–536, December, 1942.

DuBois, Philip H., A Note on the Computation of Biserial r in Item Validation, *Psychometrika*, Vol. 7, No. 2, pp. 143–146 (June, 1942).

DuBois, Philip H., Some Statistical Operations on the Counting Sorter, *Psychometrika*, Vol. 6, No. 6, pp. 383–390 (December, 1941).

Dunlap, Jack W., The Computation of Means, Standard Deviations and Correlations by the Tabulator when the Numbers are Both Positive and

Negative, *Proceedings of the Educational Research Forum*, pp. 16–19 (International Business Machines Corporation, New York, August, 1940).

Dunlap, Jack W., Note on the Computation of Tetrachoric Correlations, *Psychometrika*, Vol. 5, No. 2, pp. 137–140 (June, 1940).

Dunn, Halbert L., Adaptation of New Geometric Code to Multiple Punching in Mechanical Tabulation, *Journal of the American Statistical Association*, Vol. 27, No. 179, pp. 279–286 (September, 1932).

Dunn, Halbert L, and Reed Rockwood, A Record System Suitable for Both Clinical and Statistical Medicine, *Archives of Internal Medicine*, Vol. 41, pp. 499–535 (April, 1928).

Dunn, Halbert L., and Leonard Townsend, Application of Punched Card Methods to Hospital Statistics, *Journal of the American Statistical Association*, Vol. 30, No. 189A, pp. 244–248 (March, 1935).

Dwyer, Paul S., The Computation of Moments with the Use of Cumulative Totals, *Annals of Mathematical Statistics*, Vol. 9, No. 4, pp. 288–304 (December, 1938).

Dwyer, Paul S., Recent Developments in Correlation Technique, *Journal of the American Statistical Association*, Vol. 37, No. 220, pp. 441–460 (December, 1942).

Dwyer, Paul S., Simultaneous Computation of Correlation Coefficients with Missing Variates, *Proceedings of the Research Forum*, pp. 20–27 (International Business Machines Corporation, New York, August, 1946).

Dwyer, Paul S., Summary of Problems in the Computation of Statistical Constants with Tabulating and Sorting Machines, *Proceedings of the Educational Research Forum*, pp. 20–28 (International Business Machines Corporation, New York, August, 1940).

Dwyer, Paul S., and Alan D. Meacham, The Preparation of Correlation Tables on a Tabulator Equipped with Digit Selection, *Journal of the American Statistical Association*, Vol. 32, No. 200, pp. 654–662 (December, 1937).

Eckler, A. Ross, Some Elementary Machine Problems in the Sampling Work of the Census, *Proceedings, Scientific Computation Forum*, pp. 42–44 (International Business Machines Corporation, New York, 1948).

Edwards, Thomas I., The Coding and Tabulation of Medical and Research Data for Statistical Analysis, *Public Health Reports*, Vol. 57, No. 1, pp. 7–20 (January, 1942).

Ellis, Max E., and Arthur J. Riopelle, An Efficient Punched-Card Method of Computing Σx, Σx^2, Σxy and Higher Moments, *Psychometrika*, Vol. 13, No. 2, pp. 79–85 (June, 1948).

Gibbs, Frederic A., and A. M. Grass, Frequency Analysis of Electroencephalograms, *Science*, Vol. 105, No. 2718 pp. 132–134 (January 31, 1947).

Hall, D. H., E. L. Welker, and Isabelle Crawford, Factor Analysis Calculations by Tabulating Machines, *Psychometrika*, Vol. 10, No. 2, pp. 93–125 (June, 1945).

Harman, Harry H., Four Aspects of Factor Analysis, *Proceedings of the Educational Research Forum*, pp. 60–67 (International Business Machines Corporation, New York, August, 1940).

Hartkemeier, Harry Pelle, *Principles of Punch-Card Machine Operation* (The Thomas Y. Crowell Company, New York, 1942). 269 pp.

Hartley, H. O., The Application of Some Commercial Calculating Machines to Certain Statistical Calculations, *Supplement to the Journal of the Royal Statistical Society*, Vol. 8, No. 2, pp. 154–173; discussion, pp. 173–183 (1946).

Hey, G. B., A New Method of Experimental Sampling Illustrated on Certain Non-Normal Populations, *Biometrika*, Vol. 30, Part 1, pp. 68–80 (June, 1938).

Hodges, Fred Jenner, Isadore Lampe, and Kurt Benjamin, Organized Clinical Investigation of Cancer, *Supplement to University Hospital Bulletin* (University of Michigan, University Hospital, February, 1948).

Homeyer, Paul G., Mary A. Clem, and Walter T. Federer, Punched Card and Calculating Machine Methods for Analyzing Lattice Experiments Including Lattice Squares and the Cubic Lattice, Research Bulletin 347, (Iowa State College, Ames, Iowa, April, 1947).

International Business Machines Corporation, *Proceedings of the Educational Research Forum* (International Business Machines Corporation, New York, August, 1940). 127 pp.

International Business Machines Corporation, *Proceedings of the Research Forum* (International Business Machines Corporation, New York, August, 1946). 94 pp.

Kempthorne, O., The Analysis of a Series of Experiments by the Use of Punched Cards, *Supplement to the Journal of the Royal Statistical Society*, Vol. 8, No. 1, pp. 118–127 (1946).

Kempthorne, O., The Use of a Punched-Card System for the Analysis of Survey Data, with Special Reference to the Analysis of the National Farm Survey, *Journal of the Royal Statistical Society*, Vol. 109, Part 3, pp. 284–295 (1946).

Lorge, Irving, The Computation of the Hotelling Canonical Correlation, *Proceedings of the Educational Research Forum*, pp. 68–74 (International Business Machines Corporation, New York, August, 1940).

Martin, Caroline R., *Hospital Medical Statistics* (J. B. Lippincott Company, Philadelphia, 1933). 89 pp.

Martin, Caroline R., The Value of the Work of a Central Medical Statistical Bureau. Paper read before the American Hospital Association, Milwaukee, Sept. 13, 1933. 7 pp.

Meacham, Alan D., *The Tabulating Machine Method as Used for Modern Mathematical Statistics* (Edwards Bros., Inc., Ann Arbor, Mich., 1930). 17 pp.

Meacham, Alan D., The Value of the Collator in Using Prepunched Cards for Obtaining Moments and Product Moments, *Proceedings of the Educational Research Forum*, pp. 9–15 (International Business Machines Corporation, New York, August, 1940).

Mendenhall, R. M., and Richard Warren, Computing Statistical Coefficients from Punched Cards, *Journal of Educational Psychology*, Vol. 21, No. 1, pp. 53–62 (January, 1930).

Millang, Amy, and G. F. Sprague, The Use of Punched Card Equipment in Predicting the Performance of Corn Double Crosses, *American Society of Agronomy Journal*, Vol. 32, No. 10, pp. 815–816 (October, 1940).

Miller, Richard D., Punch Card Simplifies Records, *Modern Hospital*, Vol. 55, No. 1, p. 45 (July, 1940).

Milliman, Wendell A., Mechanical Multiplication by the Use of Tabulating Machines, *Transactions of the Actuarial Society of America*, Vol. 35, Part 2, pp. 253–264 (October, 1934); Vol. 36, Part 1, pp. 77–84 (May, 1935).

Pearl, Raymond, Modern Methods in Handling Hospital Statistics, *Johns Hopkins Hospital Bulletin*, Vol. 32, No. 364, pp. 184–194 (June, 1921).

Reagh, Russell R., A Simplified Code for Multiple Card Punching, *Journal of the American Statistical Association*, Vol. 29, No. 186, pp. 182–183 (June, 1934).

Reid, Mabel, Use of IBM Punch Cards, *Public Health Nursing*, Vol. 39, No. 2, pp. 83–85 (February, 1947).

Rosahn, Paul D., The Punched Card Technique as Applied to Autopsy Protocol Analysis, *Journal of Technical Methods*, Vol. 18, pp. 32–36 (November, 1938).

Royer, Elmer B., A Machine Method for Computing the Biserial Correlation Coefficient in Item Validation, *Psychometrika*, Vol. 6, No. 1, pp. 55–59 (February, 1941).

Royer, Elmer B., and Herbert A. Toops, The Statistics of Geometrically Coded Scores, *Journal of the American Statistical Association*, Vol. 28, No. 182, pp. 192–198 (June, 1933).

Schreiber, Frederic, and Aage Nielsen, Punch Card Code for Classification of Craniocerebral Injuries, *Journal of the Michigan State Medical Society*, Vol. 37, No. 10, pp. 909–912 (October, 1938).

Smith, Bradford B., *The Use of Punched Card Tabulating Equipment in Multiple Correlation Problems* (U.S. Department of Agriculture, Bureau of Agricultural Economics, Washington, D.C., October, 1923). 24 pp.

Taylor, Erwin K., Use of a Single Card Column for Recording Variables with a Range of 30 or Fewer Units, *Proceedings of the Research Forum*, pp. 63–67 (International Business Machines Corporation, New York, August, 1946).

Toops, Herbert A., Annotated Bibliography of Tabulating and Recording Devices, *Ohio College Association Bulletin*, No. 122, pp. 2401–2410 (April 13, 1940).

Toops, Herbert A., Research Possibilities of Addends, *Proceedings of the Research Forum*, pp. 68–74 (International Business Machines Corporation, New York, August, 1946).

Tucker, Ledyard R., *The Centroid Method of Factor Analysis by Punched Cards*, Bibliofilm Manuscript No. 1623 (University of Chicago Library, 1941).

Tucker, Ledyard R., A Note on the Computation of a Table of Intercorrelations, *Psychometrika*, Vol. 13, No. 4, pp. 245–250 (December, 1948).

Tucker, Ledyard R., A Note on a Machine Method for the Quantification of Attributes, *The Prediction of Personal Adjustment* (Paul Horst, Editor), Bulletin No. 48, pp. 347–348 (Social Science Research Council, New York, 1941).

Tucker, Ledyard R., Simplified Card Methods in Factor Analysis, *Proceedings of the Research Forum*, pp. 9–19 (International Business Machines Corporation, New York, August, 1946).

Usilton, Lida J., A Mechanical System for Record Keeping of Morbidity, Treatment-Progress, and Control of Venereal Diseases, *Venereal Disease Information, Vol.* 21, No. 3, pp. 61–67 (March, 1940).

Varossieau, W. W., Use of the U. D. C. in Selecting Data with Mechanical Appliances, *Revue de la Documentation*, Vol. 15, No. 2, pp. 41–46 (1948).

Vickery, C. W., On Drawing a Random Sample from a Set of Punched Cards, *Supplement to the Journal of the Royal Statistical Society*, Vol. 6, No. 1, pp. 62–66 (1939).

Wallace, Henry A., and George W. Snedecor, Correlation and Machine Calculation, *Official Publication of Iowa State College*, Vol. 23, No. 35 (January 28, 1925). 47 pp.

Wangeman, C. P., An Experiment in the Recording of Surgical and Anesthetic Data in Military Service. The Adaptation of Hollerith Punch Cards, Used at Second Army Maneuvers, 1940, *Anesthesiology*, Vol. 2, No. 2, pp. 179–185 (March, 1941).

Warren, Richard, Hollerith Machine Computation of Least Square Trend Lines, *Columbia University Statistical Bureau Document*, No. 2 (1933). 35 pp.

Warren, Richard, and R. M. Mendenhall, The Mendenhall-Warren-Hollerith Correlation Method, *Columbia University Statistical Bureau Document*, No. 1 (1929). 34 pp.

CHAPTER 10

LARGE-SCALE DIGITAL COMPUTING SYSTEMS

10-1. Introduction

Four large-scale digital computing systems using electro-mechanical relays as primary elements in arithmetic, storage, transfer, and input-output are in operation: the Harvard Mark II, the Bell Telephone Laboratories' Models V and VI, and the IBM Pluggable Sequence Relay Calculator. The IBM Automatic Sequence Controlled Calculator, commonly called the Harvard Mark I, employs relays for switching and employs the IBM counter wheel for storage and for the arithmetical operations. Four large-scale electronic digital calculators are at present in operation: the ENIAC, the IBM Selective Sequence Electronic Calculator, the Harvard Mark III, and the BINAC.

This chapter will concern itself with the discussion of some of these computing systems and others which are in design or under development. A brief outline of each computer will be given from the point of view of the types and numbers of components employed in the system. For more detail regarding the functions of those components, the reader is invited to examine the contents of Chaps. 3, 4, and 13 to 17. A list of major design and development projects on large-scale digital computers in this country is presented in Table 10-1, pages 214–215.

10-2. Relay Computing Systems in Operation

During the period 1944 to 1947, four relay computers were put into operation on scientific problems. These computers vary widely in complexity and character. In general terms, however, all of them are automatic sequence-controlled machines. The IBM Pluggable Sequence Relay Calculator, while normally not listed with large-scale digital machines, is included here for purposes of comparison with IBM equipment discussed in Chap. 9.

The general characteristics of relay computers are reviewed in the literature.[21]

10-2-1. Harvard Mark I Calculator. *General.*   Design and construction of this calculator was begun in 1939 at Endicott, New York, and continued until August, 1944, when the machine was presented formally to Harvard University by T. J. Watson, President of the International Business Machines Corporation. Professor Howard H. Aiken of Harvard University, then Commander, USNR, on active duty, B. M. Durfee, F. E. Hamilton, and C. D. Lake were the coinventors of the machine.[1]

The relays, counters, cam contacts, typewriters, card feeds, and card punches employed in the calculator are all standard parts of the tabulator machinery as constructed by the International Business Machines Corporation.[1,3,17,18]

The external outline of the Mark I consists of a 51-foot panel, 8 feet in height; two 6-foot panels extend back at right angles from the front panel.   Along the 51-foot panel are mounted (1) the sequence control, which directs the programming by punched tape, (2) three interpolaters, which are tape-fed units for selecting data required in the interpolation process, (3) functional counters for controlling the interpolation of functions, and for computing logarithmic and trigonometric functions and printing, (4) a multiplying-dividing unit for 23-digit accuracy, (5) 72 storage counters for the intermediate storage of results up to 23 digits capacity, used for arithmetic operations and accumulation, and (6) a storage unit with a capacity of 60 23-digit decimal numbers which can be introduced to the machine by manually setting dial switches.   The mechanical drive for this calculator is a 4-horsepower motor; the total assembly weighs approximately 5 tons.

Input, Output, and Control.   The machine operations are controlled by orders introduced on a coded 24-hole punched tape consisting of three columns of eight holes each.   This input tape also controls the time required to add two numbers, since it is that time which is required for the control tape to move one step forward.   The control tape can be advanced about 200 units per minute, which establishes the fundamental time cycle of the machine as 0.3 second for the addition of two numbers. Three tape-driven interpolaters control the routine of the interpolation process in detail.   One of the three tape-driven interpolaters may also be used to introduce numerical values to the calculator.

The machine processes the results of its calculation through

the use of standard IBM equipment: two punch-card readers, a card punch, and two automatic typewriters.

Arithmetic Operations. The Mark I Calculator can perform arithmetic operations on numbers up to 23 decimal digits, but instructions can be introduced into the machine for its operation to 12 or 46 significant figures. Addition and subtraction are performed with the set of 72 electromechanical adding-storage registers, which can store 23 digit numbers and can combine numbers which are generated by the machine in the course of its operation. Multiplication and division are performed in the multiplication-division unit, and the process followed is that of building up a small table of the first nine multiples of the multiplicand or divisor and using this table in an appropriate manner. Multiplication requires approximately 19 cycles, or approximately 5.7 seconds; division requires about 51 cycles, or 15.3 seconds.

Function tables of $\log x$, 10^x, and $\sin x$ can be constructed using an array of relay elements. Logarithmic tables are computed by using four factors of the form

$$1 + h \cdot 10^{-k} \qquad \text{where} \qquad k = 0(1)3 \text{ and } h = 0(1)9$$

and the power series for $\log (1 + x)$. [The notation $a(b)c$ is used in describing the extent and mesh of the arguments appearing in a table of functions. The first number a in the notation and the last number c are the first and last arguments in the table; the number b in parentheses is the difference between the arguments of successive entries in the table. Thus, for a standard set of five-place logarithms with four-place arguments, the notation is $1.000(0.001)9.999$.] This process reversed gives 10^x. A power-series method is used to compute $\sin x$ or $\cos x$ for values of the argument less than $45°$; other values of trigonometric functions are derived from $\sin x$ or $\cos x$. Calculation of logarithmic and simple trigonometric functions requires about 200 cycles or 1 minute.

Storage and Transfer. Sixty 23-digit decimal numbers can be stored in registers; numbers are set into the registers manually. The 72 registers used in the arithmetic operations provide intermediate storage. Numbers are transferred from one part of the machine to another by timed electrical pulses of 50 volts ampli-

tude on a single bus. It is not possible, therefore, to transfer simultaneously two or more numbers.

10-2-2. Harvard Mark II Calculator. *General.* This calculator was designed and constructed by Harvard University[6] for the Naval Proving Ground, Dahlgren, Va. The design was initiated in early 1945; construction was completed in 1947. The machine is now in operation at the Naval Proving Ground.

The Mark II uses electromechanical relays for the internal storage of numbers, for the transfer of numbers, for performing the basic arithmetical operations, and for sequence control of these processes. The machine handles 10-digit decimal numbers and uses the floating decimal point. The addition time for two 10-digit figures is 125 milliseconds and for two 10-digit multiplications, about 250 milliseconds. About 100 10-digit decimal numbers can be stored in the machine; additional internal or intermediate storage may be supplemented by punching data on tapes for introduction to the machine at a later time. The machine registers its results on punched tape or through the use of page printers in the case of tabular material. It is interesting to note that the Mark II can be operated as a whole on one problem or can be split into two parts and each operated separately.

The Mark II contains about 13,000 electromechanical relays. These relays, specially constructed for the machine, are six-pole high-speed units with double-throw contacts. They operate on 100 volts direct current and take about 6 watts when continuously energized. The operating time of these relays is relatively short, varying from 6 to 10 milliseconds. One-third of the relays are provided with mechanical locking devices which hold them in position without the relays being energized by a steady current. These latches or locks are used in the internal memory of the machine to reduce power requirements and, in the case of a power failure, to provide a nonvolatile storage system.

Input and Output. The calculator has 12 input mechanisms for introducing commands and numbers to the machine. All these mechanisms employ a punched-paper-tape input. *Commands* are introduced to the machine on one group of four input tapes. Another group of four input tapes introduces *numbers* to the machine in any desired preassigned order. The group of four input tapes is used for supplying the calculator with coded tables

of functions.   Eight hundred functional values can be introduced on these punched tapes.

The calculator records the final results of a computation through the use of four page printers or automatic typewriters. Each of these printers can record one complete number (25 characters) in about 4 seconds.   If the results are to be used again by the machine itself, they may also be automatically punched into paper tape.   Four tape punches are available for this purpose and can be made to operate at a maximum rate of about two quantities in every 3 seconds.

Number System and Storage.   A number N is represented by the calculator in terms of another number p and an integer j where

$$N = p \times 10^j, \qquad 1 \leq p \leq 10 \text{ and } -15 \leq j \leq +15$$

In order to represent p, four relays are required for each of the 10 decimal columns and one relay for the algebraic sign; j is represented in the binary system, which requires four relays.   A fifth relay for j is used to denote algebraic sign.   A storage register, therefore, requires 46 relays with 16 additional relays for input and output of information to the register.   Since there are 100 such storage units in the machine, there are 6,200 relays in the complete storage-register system.   The output time for any number in any register is approximately 33 milliseconds.

Six algebraic and transcendental functions are stored permanently within the machine; the reciprocal, the reciprocal square root, the logarithm, the exponential, the cosine, and the arc tangent.   These functions can be calculated for any argument to eight or nine significant figures as long as the argument and the function are within the digit capacity of the machine.   These functions may be computed to this degree of precision within from 5 to 12 seconds.

Arithmetic Operations.   Because of this method of representing numbers in the machine, the process of addition of two numbers has to be accomplished in two steps: The first step is a comparison of the exponents j by the machine, and the difference between the smaller and larger exponent is obtained.   The number characterized by the smaller exponent is then shifted by the amount of the difference between its j value and that of the larger number. After an addition of two numbers has been performed, the sum

may have to be shifted as it is placed into storage.   Negative numbers are represented by complements on 9's; these are obtained in the adder itself, since the storage units do not use the complementary system.   Addition requires about 200 milliseconds; this time is divided approximately into an input time of 33 milliseconds, an operation time of 125 milliseconds, and output time for the sum of about 33 milliseconds.   The Mark II Calculator contains two units for addition.

Multiplication in the Mark II is as described by R. V. D. Campbell[6]:

"In multiplying, the calculator first obtains five multiples of one of the factors, called the multiplicand (mc).   Of these multiples, $1 \times$ (mc), $2 \times$ (mc), $4 \times$ (mc), and $5 \times$ (mc) are obtained by circuits which multiply directly, but $3 \times$ (mc) is obtained by addition.   Each digit of the multiplier, working from right to left, selects the proper multiple of the multiplicand, which multiple, after the appropriate shift, is either added to or subtracted from the sum of the previously chosen multiples. Note that, because only five multiples are used instead of nine, it is necessary to be able to subtract them as well as to add them. The subtraction is carried out by using complements on nine. The exponent of the product is equal either to the sum of the exponents of the factors or to that sum increased by one.   Multiplication requires 0.7 second, regardless of the magnitudes of the factors or of how many non-zero digits they contain.   There are four multipliers in the calculator."

10-2-3. Bell Telephone Laboratories' Relay Computing Systems.   *General.*   The Bell Telephone Laboratories began the design of their computing systems in 1938, using the apparatus and circuit design techniques used generally in dial-telephone systems.   Six types of computers, recently designated by Model numbers I to VI, have been built.   Two Model V computers have been built and one of each of the others, making a total of seven.   All of these except the Model I are in operation.   Models V and VI are in the large-scale computing machine class, and brief descriptions of them are given below.

Of the earlier models, Model I was completed in 1940 and was demonstrated before the Mathematical Society meeting of that year at Dartmouth.   The demonstration consisted of introducing complex numbers into the computer by telegraph to New York

and receiving the answers through a telegraph printer at the meeting.   This machine was recently dismantled when the more versatile Model VI was put in operation.

Model II (1943) and Model IV (1945) are in operation at the Naval Research Laboratory in Washington, D.C.   Model III (1944) is in operation at the Laboratory of the Army Field Forces Board at Fort Bliss, Texas.

Model V.   In 1944 the development of an all-purpose auto- matic computing system[2,25,28] was started.   This computer con- tains more than 9,000 relays and about 50 pieces of teletype equipment, weighs about 10 tons (twice the weight of the Mark I), and covers about 1,000 square feet of floor space.

Two models of the BTL Model V machine have been built. Each of these models consists of two computers with the necessary equipment.   Each computer contains all the components needed for dealing with a computational problem with the exception of tape-reading devices which are used in common between the two computers.   Each contains 15 storage registers; eight sign registers; a calculator for the basic arithmetic operations and square-root extraction; a routine control for directing the execu- tion of instructions to the computer; a BTL register, for *block- forming*, or number-grouping, trigonometric, and logarithmic operations; a table register and table control for receiving and storing numbers read from table tapes; a problem-registering problem control for receiving and storing numbers from sections of the problem tape; a printer-registering printer control for receiving numbers from various parts of the machine and direct- ing the process of printing and perforating; a discriminator, which selects one out of several possible methods of computation and directs the machine to perform one; a recorder table, which contains a printing reperforator, a tape transmitter, and a distributor.

For problems of medium complexity, usually one computer of this system is used for carrying out the computation.   For large problems, the two computers are associated by the problem programming, and both are used simultaneously during the computation.

Input and Output.   The input, similar to that for the other digital machines, consists in introducing both numbers and orders to the machine on a perforated paper tape; the input speed

is approximately that of the computing speed of the machine. About 2 seconds are required to read and transfer into the machine a seven-digit number or a command of average length. Addition can be carried out in 0.3 second, multiplication in about 1 second. Input speed of the numbers, therefore, is slightly slower than the unit time of operations for the machine.

The output of the machine is performed by standard teletype devices called *reperforators*, for punching numerical data on tape, or *page printers*, for printing on paper.

Number System. The essential difference between the Bell Telephone Laboratories system and the other electromechanical calculators is number representation in the biquinary system. In this system, a decimal digit is replaced by two digits; one of these digits is a quinary digit which has one of the values from 0 to 4; the other is a binary digit which has one of the two values 0 or 5 in such a way that the sum of the two values is equal to the value of the decimal digit which it represents. Since there are five binary relays and two quinary relays in this system, the 10 decimal digits, 0 to 9, require the availability of seven relays for their registration, although actually only two relays are activated at any one time. This represents a saving of three relays over a machine operating on the decimal system where 10 relays must be available. Besides the seven relays, two others are required for representing the sign of the number; the negative numbers are not represented by complements but by the number itself preceded by a minus sign.

Like the Harvard Mark II Calculator, the BTL machine uses the floating decimal point. Each number is transformed into a decimal number with seven significant digits multiplied by the appropriate power of 10. The first decimal digit transferred lies to the right of the decimal point, and the exponent of 10 provides that number by which the decimal is to be multiplied. In this machine the exponents include the range -19 to $+19$. These exponents are represented by a positive or negative sign, a units digit and a 10's digit. The 10's digit, 0, and 1, and the sign, $+$ or $-$, are each stored on one of two relays. The units digits, 0 to 9, are stored in the biquinary system on two of seven relays.

Arithmetic Operations. The process of addition is carried out using three groups of relays each of which can store a 10-digit

number in the biquinary system. These groups are so wired that if numbers are introduced to the first two groups, the sum automatically appears in group three. In subtraction the process is conducted by converting the number with the larger exponent into its complement with respect to 9,999,999. The other number is introduced to the second group of relays shifted as required with respect to the exponents. The result, which appears in the third group of relays, is the complement of the desired number, and the machine must convert it back into the number required. Multiplication is performed by repeated addition, and division is performed by repeated subtraction. The extraction of square roots is also performed by repeated subtraction.

The estimated speeds of several arithmetic operations for seven-digit numbers are given below:

Operation	*Time, seconds*
Addition or subtraction..................................	0.3
Multiplication...	1.0
Division..	2.2
Square root...	4.3
Number transfer from one register to another register.......	0.07

These estimates include the time required to introduce a seven-digit number into the calculator and place a seven-digit number (result) in a given register.

Storage. The machine can store 30 seven-digit decimal numbers. A unit of 62 relays, required to store one seven-digit decimal number, is called a *register;* the BTL machine has 44 such registers, 30 for number storage and 14 for performing other functions. Each computer is also equipped with a register called a *BTL register* which serves as a storage medium and performs special operations such as those of forming block numbers and computing trigonometric and logarithmic functions. In this sense, the BTL register is closely related to the permanent function tables in the machine.

These functions, which are available to the machine in storage, include $\sin x$, $\cos x$, and $\tan^{-1} x$. The machine for the Ordnance Department of the Army contains, in addition to these functions, $\log x$ and 10^x.

The setup time has been reduced by provision of more problem

setup positions than can be used at one time. In the Aberdeen installation there are two computers, but there are four problem positions. With this arrangement two problems can be set up on the two idle positions while the computing equipment is actively using the other two. As soon as either one of the two problems has been completed, one of the stand-by problems is automatically connected in. This automatic feature is one of the reasons why the Model V computers can be operated on an unattended basis.

Model VI. The Bell Telephone Laboratories has recently placed in operation in its Murray Hill, N.J., building a new digital system called the Model VI. Remote-control stations are located in the quarters of the computing staffs, and the computing equipment itself is located in another part of the building. A switchboard cable of the type used in telephone central offices is used to provide for the exchange of information between the two locations. The machine contains about 4,300 relays, 86 cold-cathode tubes, and other miscellaneous apparatus. This machine is a little less flexible than the Model V but is simpler in design detail. Each number is introduced into the machine as a three-, six- or 10-digit decimal number times a power of ten ranging from -19 to $+19$. The floating decimal point appears to the right of the most significant digit.

10-2-4. The IBM Pluggable Sequence Relay Calculator. *General.* The IBM Corporation developed two models of this relay calculator[7] during the war; they were used at the Aberdeen Proving Ground during 1944 to 1945. In 1945, three additional machines, with improvements over the first design, were built. One of these machines is now at the Naval Proving Ground, Dahlgren, Va. The other two machines are at the Watson Scientific Computing Laboratory, Columbia University, in the City of New York.

Input and Output. This relay calculator operates on numbers read from punched cards, undertakes a sequence of calculations by relay networks, and punches the result of these calculations on cards. The basic operations of the machine are synchronized with the rotations of a shaft in the card-reading unit. This shaft rotates at a speed of 100 revolutions per minute. The card feeds and punches are synchronized to operate within the time of a single rotation of a shaft. During one shaft rotation, a timing

circuit generates 48 impulses spaced at 0.0125 second; these impulses operate the relays. Another timing circuit generates 14 pulses during a shaft rotation which is used to actuate the reading and punching mechanisms. With this system one impulse for the card-actuating mechanism is equivalent to $3\frac{3}{7}$ relay impulses.

Arithmetic Operations. The operations of addition and subtraction require two impulses and three relay groups, one for the augend, one for the addend, and one for the sum. If successive additions or subtractions involve the same relay, four-group relay impulses are required. Multiplication is performed in a manner similar to that used in the IBM multiplying punch, Type 601 (see Chap. 9). Division is performed by subtracting the divisor or 5 times the divisor from the dividend. In division, the dividend is shifted to the left after the completion of each subtraction operation in one position. The process of subtraction consists in adding the complement of the number to be subtracted. The capacity of this calculator for division allows for six figures, 12 digits in the dividend, six in the divisor, and 10 in the quotient; the time for division is approximately 0.2 second per quotient digit. The extraction of a square root is performed by successive additions of odd numbers to the complement of the original number. In the square-root process the capacity of the calculator is 12 digits in the original number and six digits in the square root.

Storage and Control. The machine has a total of 36 storage and computing registers. The normal capacity of the machine is six-digit decimals, although higher precision than this may be obtained with sequence operation. In conjunction with the punched cards the plugboard facilities connected with the machine allow operations to be performed in parallel rather than the straight sequence operation resulting from punched-tape input. In comparison with the standard IBM Calculating Punch (see Chap. 9 for greater detail), this calculator has a higher operating speed. It is approximately ten times as fast as the calculating punch.

The sequence-control operations can be initiated by a pluggable sequence and also by the punches on the cards. Actually the pluggable sequence on the control panels starts the operations. In this system, there are 96 sequence relays connected to the plugboard so that the impulses through these relays can be used

for control purposes. The relays operate in succession; for a complete sequence of 96 relays to be operated, the time required is two shaft rotations of the basic drive, or 20 milliseconds. In cases where double operations occur, such as multiplication and addition, whose orders are required to be read from the card and whose results are to be printed on the card, the machine may be controlled to operate at 50 cards per minute through the reading and punching units.

10-3. Electronic Computing Systems in Operation

In order that more complicated physical and mathematical problems may be solved in an economical time interval, the unit operation time for automatic computing systems[15,20] must be decreased wherever possible. By using electronic vacuum tubes for switching, storage, and for arithmetic operations, a saving in time in these functions of several orders of magnitude is effected over similar devices using electromechanical relays.

It is probable that, for some time to come, input and output mechanisms using data punched on cards or paper tape or registered as dots on film or as marks on a magnetic medium will have to be employed. Although such physical systems operate at inherently slower speeds than electron vacuum tubes, their use is not as stringent a time-limiting element as would be thought at first. For solving highly complicated mathematical problems, which may represent in reality the interaction of many physical parameters, it is desirable that the computer retain as many intermediate results as possible during the computation. This can be accomplished in an electronic computer if mechanical devices are avoided for use as intermediate storage mediums. For example, there would be no reason for employing high-speed electronic arithmetic circuits if numbers and commands are presented to or abstracted from them by devices whose unit operation time is much slower than that of the electronic vacuum tube. If mechanical devices are limited in their application only to the initial input numbers and commands and to the final registration of completed results and if switching, gating, and arithmetic operations are conducted at speeds which are of the same order as electronic-vacuum-tube operation, the computer may carry out thousands of calculations within itself before arriving at a result worthy of physical representation outside

the machine. In such a calculation the machine may perform a large number of arithmetic operations; it may construct function tables or generate new numbers and orders in sequence; it may also store intermediate results and draw upon them when required. Such a procedure would warrant the complicated electronic circuitry which speed and the consequent reduction of calculating time demand. Although with present input mechanisms it requires a considerable length of time to code, instructions and numbers for input to the machine, the total time for these operations may be made feasible if the machine can carry out calculations a thousand times as fast within itself; in the same way a mechanical output, either a punch or a printer, must not hold up the internal functions of the calculator. For each number punched or printed, there must be at least 1,000 operations performed electronically within the machine. If the machine presents data to the output faster than this, an intermediate-speed storage mechanism is then required for protecting the computer against slow-output intrusions upon its high-speed capabilities. The intermediate storage effectively changes the time scale from that within the computer to that outside of it.

10-3-1. The ENIAC (Electronic Numerical Integrator and Computer). *General.* This computer was developed at the Moore School of Electrical Engineering, University of Pennsylvania.[12,16,23] It was first demonstrated to the public in February, 1946, and was transferred in 1947 to the Ballistic Research Laboratories, Aberdeen Proving Ground.

The ENIAC is used, essentially, as a general-purpose computer. As an example which can be regarded as an illustration of the upper limit of its capabilities, it can be used to solve the system of five simultaneous hyperbolic partial differential equations describing flow around a body of revolution. For each case (*i.e.*, each different mach number and set of shape parameters) this problem takes the machine about an hour.

The ENIAC is the first large-scale machine to make use of electronic circuits for general operations except for input, output, and certain switching functions. There are 40 separate panels arranged in a U configuration with a total of approximately 1,500 electromechanical relays and 18,000 vacuum tubes.

As was indicated by the footnote on page 35, the programming philosophy of the ENIAC has been modified since it was

first installed at the Ballistic Research Laboratories. The machine itself has not been modified except by the addition of a converter. This unit decodes pairs of decimal digits and, for each pair of decimal digits, energizes a unique output cable. There are 100 of these cables.

A signal on any one of the 100 conducting cables which can be energized by the converter initiates a particular operation. The portable function tables which were formerly used to store tables of functions are now used to store sequencing instructions in addition to numerical data, tables of functions, constants, etc. Each instruction consists of a pair of decimal digits which are decoded by the converter unit. The pairs are read one at a time in accordance with the program sequence.

Each output cable from the converter unit initiates an operation which has been determined by the arrangement of the plug connections, and the settings of the switches are called the background wiring. The background wiring is changed only if it is desired to redesign the coding.

To instruct the machine, then, it is necessary only to set the switches on the portable function tables. The operations to be performed and the sequence in which they are to be performed are specified by the switch settings.

Input and Output. The numerical data to be introduced to the machine are first punched on standard IBM cards; these cards contain 80 columns, each column with 12 punched designations (see Chap. 9 for details on IBM equipment). Ten of the punched designations, 0 to 9, are used to represent one decimal digit. Positions 11 and 12 may be used for representing the sign of a digit in one column or a group of digits in blocks of adjacent columns. The 11 and 12 punch holes may then be used for representing a positive or negative sign for each of these numbers; it is seen that 16 such punch positions must therefore be available. The numbers represented by card punches are then read and stored in relays in the *constant transmitter*, which makes these numbers available to the various operating components of the machine when required. In a similar way, the results which have been calculated by the ENIAC are punched on cards using an IBM card punch; the data from these cards can then be printed automatically through the use of an IBM tabulator.

Arithmetic Operations.   All the units which perform numerical operations—addition, subtraction, multiplication, division, and extraction of square root, together with the capacity to obtain numbers from the machine's storage system—operate by using electronic vacuum tubes on the flip-flop principle.   For addition and subtraction, the machine makes use of 20 accumulators. Each one of these accumulators can hold a signed 10-digit decimal number.   An accumulator is capable of receiving a number transmitted to it and adding this number to the number which it already contains.   Also, on one output channel, the accumulator can transmit the number stored at any given time. On another output line the accumulator can transmit the complement of the number it stores.   As described in Chap. 3, the 10-stage (decade) ring counter is designed to count pulses in a one-way progression around the ring.   Subtraction of two numbers, therefore, is most effectively performed with decade ring counters through the use of complements.

The numerical circuits for the multiplier make use of four accumulators, or up to six if products with 10 to 20 digits are required or if the best 10-digit accuracy is required.   The units digits of the partial product of the entire multiplicand by one digit of the multiplier are transmitted to one accumulator while the 10's digits are transmitted to another accumulator.   For example, to multiply 567×234 the machine first multiplies 567×4.   A novel method, which is much faster than the normal multiplication by successive additions, is employed in accumulating this partial product.

The units digit of the product of 7×4 or 8 is transmitted to one accumulator and the 10's digit or 2 to another.   Similarly the units digit of the product of 6×4 or 4 goes to the units accumulator shifted one digit to the left and the 10's digit, again 2, goes to the 10's accumulator also shifted.   As a result of the multiplication of 567×4, therefore, 048 is stored in one accumulator and 222 in the other.   For 567×3, 581 is stored in one accumulator, 112 in the other (shifted with respect to the numbers already accumulated).   Two multiplication tables are employed; one produces the 10's digits of the required product, and the other the units digits.   When the digits of the multiplier have been used up, the 10's and units digits of the partial products

are combined in one of these accumulators to obtain the total product.

The divider is one panel of equipment which controls neighboring accumulators so as to perform division or extract a square root. One accumulator stores the numerator (radicand), one stores the denominator (or twice the square root), one stores the quotient, and finally, a fourth accumulator is used for shifting operation. Division is accomplished by a method similar to that used in some desk machines. The denominator is subtracted from the numerator repeatedly until an overdraft appears. The remainder is then shifted one place, and the denominator is added until this overdraft is wiped out. The quotient accumulator counts the number of cycles in each place. The process of extracting the square root of a number is carried out by a combination of successive subtraction and the addition of odd numbers, analogous to the process used in some desk calculating machines.

Storage. The storage system is composed of units which store numerical tabular data in electronic circuits; it also contains electromechanical switches for storing instructions for a specific problem. There are two additional units, one for initiating computation, for clearing the machine, and for testing; the other is a cycling unit which supplies fundamental signals to the machine and synchronizes their operation. The fundamental unit of time in the ENIAC is $\frac{1}{5000}$ second; during this period a pattern or train of pulses is emitted. One cycle is referred to in the ENIAC as an *addition time*. The constant transmitter, for example, requires one addition time to emit a signed five- or 10-digit number.

The pulse width is 2 microseconds and the machine is operated at about 100 kilocycles per second.

The machine, as now used, has two internal storage systems from which it can withdraw stored numbers called for by a particular program of operations. The basic difference between them is that data can be stored in one of these by the machine itself; data can be stored in the other only by the operator. As indicated under the general description of the machine, the portable function tables (which are simply resistor matrices) are used for the storage of numerical data as well as for sequencing instructions. They comprise the second type of internal storage,

in which data can be inserted only by the operator. The other storage system consists of accumulators.

10-3-2. The IBM Selective Sequence Electronic Calculator. *General.* This calculator, developed by the International Business Machines Corporation,[19] was put into operation in January, 1948. It combines electronic-vacuum-tube techniques and electromechanical relays. Vacuum-tube circuits are employed for arithmetic processes and for high-speed storage, switching, and other control orders. Relays and punch tape form secondary storage units.

Input and Output. The instructions are introduced to the machine either from the standard 80-column IBM cards or from continuous card-stock tapes. The keyboard punch for these tapes may be operated manually by keyboard or may be operated automatically by punched cards. The punch converts a decimal number into its binary equivalent, which is then handled by the calculator. The card-stock tape is approximately the width of the long dimension of the standard IBM card, allowing 80 perforations per line across the tape. This provides ample space for representing a 19-decimal-digit number together with algebraic sign across one line of the tape. The results of machine calculations can be punched on standard IBM cards or can be printed in tabular form. If it is required, both methods may be used. The digits can be punched on standard IBM cards at the rate of 1,600 digits per minute and can be printed in tabular form at the rate of 2,400 digits per minute.

The operating instructions for the machine are read from punched tapes, and groups of commands are delivered at the rate of 50 command groups per second to the machine. Since these commands must be synchronized with the accepted arithmetic transfer and storage operations, a 20-millisecond delivery rate appears entirely feasible in view of the electromechanical devices whose operating characteristics must be tied in with the other machine components. The array of orders and numbers on input tapes does not appear to differ greatly in procedure from other machines which employ punched-tape feeds.

Arithmetic Operations. Addition, subtraction, multiplication, and division are performed by electronic-vacuum-tube counting circuits. The representation of the numbers in the machine is

not specified, but it is stated that this calculator contains more than 100 electronic counting units which can be combined for arithmetic operation to add or subtract two 19-digits in less than 10 milliseconds. The electronic counting units will multiply two 14-digit numbers to give a 28-digit product in 20 milliseconds. For division these units may be combined to allow the division of two 14-digit numbers to give a 14-digit quotient in about 30 milliseconds. Provision for handling decimal points and indicating the algebraic sign of each number is made.

Storage. The internal storage capacity of the machine totals 400,000 digits and is divided among electronic, electromechanical, and punched-tape units. In electronic storage there are eight units, each of which can store one 19-digit decimal number with algebraic sign. About 3,000 digits (150 numbers) can be stored in the electromechanical relay units. Punched-tape units can store up to 396,800 digits (20,000 numbers).

The punched-tape storage system comprises three units, each with a punching unit for the card-stock tapes and each with 10 reading stations. The total of 30 reading stations may be combined in a number of ways to permit their use on individual or specific problems. On a single line of punching on the tape, it is possible to store 19 digits and indicate the algebraic sign of the number.

For the table-consulting function, the machine is provided with a group of 36 reading stations additional to those identified with the storage system. These reading units are used to locate specific information on any of the 36 punched tapes flowing through each reader. Since each of the 36 reading units may have a tape with capacity for 150 19-digit numbers, the total capacity for the tabular system is 5,400 numbers, or 102,600 digits. These units may be combined so that a maximum of 38 digits with algebraic sign may be made out. The access time for this table or for searching the 36 tape units is about 3 seconds. Approximately 160 milliseconds is the time for the tabular unit to find a particular value in an eight-place sine table. In the arrangement of the tabular information, on punched tapes, the values of the argument of the function with variable intervals may be used, which in some cases may allow the use of fewer tabular values. In addition to the punched-tape

storage, intermediate results, which may not be required for reintroduction into the machine for a considerable length of time, may be stored on IBM cards.

10-3-3. Harvard Mark III Calculator. At the time of this writing the Mark III is in operation at the Harvard Computation Laboratory where it was built for the Navy Bureau of Ordnance for eventual installation at the Dahlgren Proving Ground. This machine was made available for inspection by visitors during the Symposium on Large-Scale Digital Calculating Machinery which was held at Harvard on September 13 to 16, 1949, under the joint sponsorship of the Bureau of Ordnance and Harvard University.

The plans for the Mark III computer were discussed by Dr. Aiken at the 1949 IRE National Convention. The Mark III was designed to provide greater speed and reliability, more flexible storage facilities, and greater ease of preparation of input data than were found in the earlier Harvard Computers.[1] It was planned that the Mark III would use a magnetic drum for storing 4,000 orders and that 30 types of basic circuits would be incorporated into plug-in units, with color coding, for ease of maintenance.[2]

A preliminary description of this machine by B. L. Moore will be found in *Proceedings of a Second Symposium on Large-scale Digital Computing Machinery* (Harvard University Press), which is in preparation at the time of this writing.

10-4. Electronic Computing Systems in Design or under Development

The following electronic computing systems, in various stages of development, advanced design, or preliminary consideration, have been outlined here to indicate the trend in component use and the integration of those components into automatic sequence-controlled computers.

10-4-1. The EDVAC (Electronic Discrete Variable Computer). *General.* This computer is the second high-speed electronic digital machine which has been constructed at the Moore School of Electrical Engineering at the University of Pennsylvania.[24] Although the entire calculator has not as yet been operated as a unit, many of the components have already been tested successfully, and the scheme of operation has been

assured. From point of size, the EDVAC occupies much less space than the ENIAC and is constructed with panels 7 feet high occupying some 140 square feet of floor space. The number of electronic vacuum tubes in the EDVAC is only about 3,500, roughly one-sixth the number employed in the ENIAC. Outside of the method for programming, in which a serial rather than a series-parallel arrangement appears to be the goal, the principal difference lies in the use of acoustic delay lines for internal storage of numbers. The EDVAC is designed to handle 44-digit binary numbers.

At the time of this writing, the EDVAC has been accepted by the Ordnance Department, Department of the Army, and has been installed at the Ballistic Research Laboratories at Aberdeen, Md. where it is now undergoing extensive tests.

Input and Output System. As originally designed, binary digits (representing numbers and commands in four-address code) were intended to be read into and out of the machine on magnetized wire. This method was found to be unreliable, and teletype tape will be used as an interim input-output system.

Storage. The EDVAC uses acoustic delay lines for high-speed storage; 126 mercury delay lines will provide storage capacity for 1,024 44-digit binary numbers. A single delay line will circulate eight numbers each of 44 binary digits. The pulse width is 0.3 microseconds (as compared with 2 microseconds in the ENIAC), and the repetition rate is 1 megacycle per second.

Arithmetic Unit. Addition, multiplication, and division are built into the computer; these operations are based on the serial handling of digits. Total addition time, including transfer of numbers, is about 1 millisecond; multiplication and division require about 3 milliseconds. Double precision multiplication and division are built into the arithmetic units, enabling the programmer, at his option, to perform operations yielding 88-digit results. A higher degree of precision can thus be achieved if desired. Checking is provided for by the simultaneous use of two algebraic units with inputs to a comparison unit.

10-4-2. Electronic Computing Instrument (Institute for Advanced Study). *General.* Preliminary information concerning the design of an electronic digital computer has been recorded by Burks, Goldstine, and von Neumann.[5] The planning and

coding of problems for such a computer have been treated by Goldstine and von Neumann.[13,14] It is not intended here to deal with the details of the electronic circuits or of the functioning of any given unit within this machine except as required to understand the principles of operation and the role to be played by each unit within the machine.

In the preliminary exposition describing the design of an electronic computer the authors indicated that a fully automatic general-purpose computing machine requires a storage system, a control unit, arithmetic unit, and input and output units. At the time of this writing, a storage capacity of 1,024 numbers, each of 40 binary digits, has been decided upon as an interim solution to this particular design question, although 4,000 or more would be desirable. The computations which have been considered in order to arrive at this figure include the solution of (1) total differential equations, (2) partial differential equations of the elliptic, parabolic, or hyperbolic types, and (3) problems involving solution by iterative processes including systems of linear algebraic equations.

Input and Output. The numbers and orders to the computer will first be coded on punched paper tape using modified teletype equipment. These coded numbers will then be transferred by a conversion unit to marks on a magnetic wire. Numbers and orders in this form are then introduced to the computer elements. For output, the numbers registered on the magnetic wire by the computer are transferred to punched paper tape. Some of the units for handling the punched paper tape in this input-output system were developed at the National Bureau of Standards.

Storage. The proposed machine will have an electrostatic storage system consisting of a bank of Williams tubes (of the type described in Chap. 14 on pages 366–369). Test operations with cathode-ray tubes as electrostatic storage elements have already been conducted using the Williams technique.

The Arithmetic Unit. The arithmetic units will operate in a parallel manner on binary numbers. At the time of this writing, an accumulator with capacity for 11 binary digits has already been tested, and the 40-stage arithmetic unit is now virtually complete and has already undergone extensive tests.

The bank of Williams tubes operating in parallel as a storage system should make possible a very high total arithmetic speed;

total addition time is expected to be of the order of 90 micro-
seconds for two 40-digit binary numbers (including the time
required to withdraw three numbers from storage and to return
one), total multiplication time approximately 300 to 400
microseconds.

**10-4-3. The UNIVAC (Eckert–Mauchly Computer Corpora-
tion).** *General.* The UNIVAC is a high-speed electronic com-
puter[8] making use of acoustic-delay-line and magnetic-tape
storage. At the time of this writing, practically all production
drawings have been completed, and the first of six UNIVAC
systems is now under construction. All portions of this system
which are not similar to the BINAC (the computer built for
Northrop Aircraft, Inc., by the Eckert–Mauchly Computer
Corp.) have been laboratory tested in one way or another.

The basic pulse rate for all internal circuits in the UNIVAC
is 2.25 million pulses per second. The computer handles char-
acter groups of 12 characters each; one such group is usually
made up of 11 decimal digits and algebraic sign, but alphabetic
letters and punctuation marks are also acceptable characters.
The high-speed storage capacity is 1,000 character groups. Mag-
netic tapes store an additional 10^6 groups under computer control.

Input and Output. The instructions, and numerical or other
data, are introduced to the computer by signals coded on metallic
eight-channel magnetic tapes. More than 10^6 characters
(decimal digits, letters, or other typewriter symbols) may be
recorded in binary form on one reel of such tape 1,200 feet long,
$\frac{1}{2}$ inch wide, and less than 2 mils thick. The reels are approxi-
mately 8 inches in diameter. A keyboard resembling a type-
writer keyboard is used as the input to the apparatus which
records the coded signals on the magnetic tape.

Normally, provision is made for computer connections to 10
servo units, each of which controls one reel of tape. Altogether,
more than 10^7 decimal digits (or other characters) can therefore
be made available to the computer at one time for automatic
use. The tapes may be fed in either direction past the magnetic
heads. Any tape may be used for input or output at any time,
and both input and output functions may be overlapped with
computing. Automatic interlocks are used to ensure that only
one tape functions as input at a given time and only one tape as
output at a given time.

Instructions for rewind are of two kinds. One allows reuse of the same tape in the problem when called for by further instructions, and the other sets an interlock which is not released until an operator changes the tape. The latter feature is of importance in handling large statistical problems where the data exceed 10^7 characters, and also in avoiding reel-changing setup time between problems. Results recorded on magnetic tape can be printed on a special typewriter (with continuous paper feed) which is automatically controlled by the coding on the tape.

Storage. High-speed and low-speed storage units make use of mercury delay lines and magnetic tape, respectively. The acoustic storage unit has 100 channels and an inherent capacity (counting the spaces between digit groups) of 91,000 binary digits.

Of the 10 magnetic tape reels for which servo controls are normally obtainable, the user may allocate as many as he wishes to instructions, input, output, or temporary storage. These allocations may vary within the same problem if desired. Actually, all the instructions needed for a great many different problems could be put on one tape reel.

Approximately 2,000 blocks, of 720 characters each, can be stored on one reel. Each input or output instruction transfers one such block of information at a rate of about 800 12-character groups per second. (This corresponds with a transfer rate of about 70,000 binary digits per second.)

Within the high-speed storage, special transfer instructions are provided for handling two consecutive groups which are stored in a single acoustic channel.

Arithmetic Operations. Addition or subtraction of a number is executed in less than 600 microseconds; multiplication requires about 2,500 microseconds, and division less than 4,000 microseconds. (These times include allowance for obtaining the instruction itself from storage.) Either a 22-digit product may be obtained or an 11-digit product with proper round-off, according to the instruction used.

Supervisory Controls. A typewriter keyboard and automatic printer are directly connected to the UNIVAC computer for supervisory purposes. This permits manual control of computer operations, modification of programs for test purposes, automatic printing out of brief indications of progress in long com-

putations, etc. A panel of selector switches and indicator lamps is included; when a stoppage due to a failure to satisfy the automatic checking circuits occurs, the nature of the failure is indicated, and further diagnosis and remedy are facilitated by use of the selector switches.

Pulse Code. The code for every character contains 7 binary digits. One is the *check digit.* Two others are used for indicating nonnumerical characters and are always zero for decimal digits. The last four correspond, in the case of a decimal digit, to the binary equivalent of the decimal number plus three. Thus, the last four digits of the code for 7 are 1010. In Sec. 13-4-5, this *excess-three code* is used as an example of one system for representing decimal digits in binary notation and some properties of the system are mentioned.

The *check digit* is always chosen to make the sum of the binary digits *odd.* Hence, the complete 7-digit code for 7 is 1001010. When the second or third digit from the left is nonzero, the code represents a nonnumerical character. The check digit is still chosen as before, to make the sum of the binary digits in the 7-pulse group odd. This checking code is used throughout the computer, both in high-speed circuits and on the magnetic tape.

Checking. During all transfers, within the computer and to and from magnetic tape, special circuits examine every character (7-pulse group) to see that it is odd. If it is not, computer operation is arrested and an error indication given. Circuits for arithmetic operations are duplicated, and the outputs of the duplicate units compared. If the outputs are not identical, computation is stopped and an appropriate indication given. Periodically, every character group in the acoustic storage is scanned to check that its digits are odd, so that defects in the storage circuits will be found almost immediately even though the computing routine has not yet had occasion to call upon the defective register. The time required to carry out this periodic check is negligible. The time interval between such periodic checks is short enough to make the probability of making two compensating errors effectively zero. Some other checking methods are used in the tape recording device (UNITYPER) and the automatic printers (UNIPRINTERS).

Alarm-type fuses (Western Electric), similar to those used in

the ENIAC, are used in all distribution lines for d-c power, and a blown fuse will cut off all d-c power.   Problems can be so programmed that a shutdown, either from a blown fuse or from external power failure, does not require starting the problem over from the beginning.

10-4-4. The SDC Raytheon Computer (The Raytheon Manufacturing Company).   The Raytheon Manufacturing Company has started to work on a high-speed digital computer[4] for the Special Devices Center of ONR.   This machine uses mercury delay lines for internal storage and magnetic tape for external storage.

Input and Output.   Problems are introduced to the machine on coded magnetic tape.   The code is generated by means of a manually operated keyboard.   The magnetic tape can store either decimal numbers in binary code or binary numbers in the octal notation.   Through the use of the octal notation, three binary digits identify each digital position.   With the use of this number system the conversions from the binary to the octal and from the octal to the binary are made simple.   Associated with the keyboard device is an input page printer, which prints, for visual observation and record, the information which has been introduced to the machine.

The numerical output of the machine is recorded on magnetic tape and is then distributed to one of two output printers. The output data are normally in decimal form, although binary numbers can be printed in the octal notation.

Storage System and Number Representation.   The SDC Computer is capable of handling several different kinds of numbers. In standard operation it can store numbers to a precision of 35 binary digits with fixed binary point.   The basic arithmetic operations are carried out to a precision of 35 binary digits.   It is also possible, however, to store and manipulate numbers of 70 binary digits as pairs of standard numbers.   Conversions from decimal to binary notation and from binary to decimal notation are performed in the arithmetic unit and not by separate equipment.   This requires that decimal numbers be stored in the internal, or acoustic-delay-line, system.   The storage capacity per number in the storage system, which accepts numbers in binary-coded decimal notation, is eight decimal digits.

The internal storage will have a capacity of 1,152 numbers;

each is represented by 35 binary digits with algebraic sign and is characterized by a check number. Three different types of operation are performed by this storage system: the writing of new numbers into the system, the reading of current information in the storage unit, and the erasing of information to prepare the internal storage for introduction of new numbers. With acoustic storage, numbers are circulated through the delay line in serial order and are regenerated by an electronic circuit which reintroduces them to the acoustic delay line. The pulse repetition rate is 3.7 megacycles per second, and each number requires 42 pulse positions. The time duration of one number is therefore 11.25 microseconds. The acoustic delay lines, however, will store 32 numbers per channel so that a complete cycle through the acoustic delay line requires 360 microseconds. Thirty-six acoustic-delay-line channels are required to provide capacity for 1,152 numbers (32 with the main storage, plus a buffer reservoir housed with each of the four external storage units). The external storage consists of four identical self-controlled units each of which contains one magnetic-tape mechanism capable of storing 10^5 numbers. The tape on any individual external storage unit may be scanned at the rate of about 500 numbers per second. From the design characteristics it may require, in the most unfavorable cases, as long as 2000 seconds to read a given number out of the external storage unit. This premium on time for reading from the external storage may require the rearrangement of data before they are placed into the machine so that the external storage unit presents recorded information at the right time and in the correct sequence for machine use. Each tape unit in the external storage unit has associated with it two 32-number acoustic delay lines, which are employed as buffers between the tape and the main electronic part of the machine.

The transfer and storage of numbers in both the acoustic delay lines and the magnetic-tape storage unit are checked by means of a weighted count. This is a weighted sum of the digits of the number. This weighted sum is computed modulo 16 and is then modified by the addition of 1. A number to be checked and its weighted count cannot therefore be zero simultaneously, so that a null number is not a valid one. By this check, the failure of a gate or other device controlling the transfer will be detected.

Arithmetic Unit. The arithmetic unit of the SDC machine performs the arithmetic processes of addition, subtraction, multiplication, and division. This unit can also generate new numbers or new orders; the ability to generate new numbers differentiates the arithmetic unit from all the other units of the machine whose functions are purely passive. The arithmetic operations are performed by electronic-vacuum-tube techniques. The numbers in the arithmetic unit are handled by using a parallel representation, the various digits of the number being stored in separate flip-flop devices. This requires that the transformation of a number from storage to the arithmetic system be performed by series-to-parallel representation of the digits in each number.

The arithmetic unit can also perform basic operations on either numbers or commands introduced to it. The latter initiate the transfer of numbers from one storage position to another, the shift of a number, and the extraction of any given number of digits from a number.

The process of addition is performed in a straightforward way in this parallel representation. Subtraction uses the same process as addition after the complement of the subtrahend is taken. A convenient method for taking the complement makes use of the ability in the binary notation to replace all the digits in the subtrahend consisting of 1 by 0, and 0 by 1. This computer takes complements on $2 - 2^{-35}$. Multiplication of two 35-digit numbers results in the accumulation of the complete 70-digit product in two of the arithmetic registers. Either 35-digit section, or both sections, of this number can be used for subsequent calculation. In division, both the quotient and the remainder are available in two registers.

All arithmetic operations are checked by means of an arithmetic weighted count so that the failure of any unit of the arithmetic system can be verified.

10-4-5. ERA Computer (Engineering Research Associates, Inc.). Engineering Research Associates, Inc., also has completed and submitted to the Bureau of Standards a preliminary design for a general-purpose electronic binary computer.[10] This design incorporates a magnetic drum[9] for the internal storage of numbers and orders. The general characteristics of the proposed machine include: punched-tape input, arithmetic, and

coding operations[11] to be conducted by electronic vacuum tubes, binary digit representation, magnetic-drum storage, and punched-paper-tape or printed output. The machine is capable of handling 30-digit binary numbers and has a storage capacity of 4,096 30-digit binary numbers.

Input and Output. Numerical data and instructions for machine operation are introduced to the machine by coding on a seven-hole punched paper tape. This tape is prepared by manual operation of a keyboard translating unit which operates a standard seven-hole tape punch. The numerical data occupy six tracks on the tape; the coding for control of input operations occupies the seventh. For introducing a 30-digit number to the magnetic-drum storage, five consecutive lines in each of the six levels or tracks on the punched tape are used. Each track, therefore, supplies successively five digits to the drum. The 30 digits on the magnetic drum may therefore be considered to be divided into six groups of five digits each. The first line or consecutive group of punches in each of the six tracks introduces a digit into the extreme left position of each of the groups of numbers into which the magnetic drum is divided. The second line or group of punches on each of the six tracks supplies digits to the digit positions immediately to the right of those previously filled. This procedure is continued until 30 binary digits are registered in a row across the magnetic drum.

The punched tape is read by a photoelectric scanning mechanism capable of reading 150 lines per second. To introduce 4,096 30-digit binary numbers requires approximately 2.3 minutes.

The seventh track, containing the control code for directions and for checking operations sequentially, is divided into consecutive lines each composing a frame. These five lines are designated *a*, *b*, *c*, *d*, and *e*. Four phototubes read the seventh-track control code. These four tubes follow a binary coding which then identifies the operation directed by that line of tape perforations.

The output section of this computer contains a unit for transmitting data from the magnetic storage to punched paper tape or to an electric typewriter for printing. The electric typewriter then prints characters or numbers in base 10; the perforator punches six digits as a line of data on six- or seven-hole tape.

Storage. This computer has storage for 4,096 30-digit binary numbers. These digits are stored on magnetic tape which is contained on the circumference of a cylindrical drum; the drum is capable of continuous rotation, and the magnetic representations for 0 and 1 are recorded, read, and erased while the drum is rotating. The drum contains two groups of 30 parallel tracks; each track can store 2,048 digits around the circumference of the drum. Each group is therefore equivalent to an array of numbers of 30 columns and 2,048 rows; this array of numbers may be thought of as being wrapped around the circumference of the drum so that the first row is in a position equivalent to row 2,049. Each of the 2,048 rows in this array contains a 30-binary-digit number, and all the digits in a 30-digit number lie in different columns. The location of a number in this storage system, therefore, is performed by locating the row or *box* containing the number. A 12-digit binary number called the *address* characterizes this location. The leftmost binary digit identifies one of the two groups of tracks; the remaining 11 digits determine one of the 2,048 (2^{11}) rows or angular positions of the drum.

There are, additionally, 11 tracks used for number location and two tracks for the timing and control of storage operations. The group of 11 tracks for location contains 2,048 permanently recorded 11-digit binary numbers, each number designating an address of one of the 2,048 rows.

Above each of the 60 tracks is a magnetic head capable of recording, reading, and erasing the information contained on the track. As the drum rotates, the two groups of 30 heads read two 30-digit numbers simultaneously, the sequence of such patterns repeating themselves every drum revolution. Associated with the reading of the information of each track is a group of 60 storage-reading amplifiers for reading signals from the selected group of storage tracks and for causing the number read from a particular position to be transmitted to a specified destination in the computer. Storage-reading amplifiers for three sets of output gates direct the numbers to the X register, to the program-control registers, or to the print or punch register.

Arithmetic Section. The arithmetic section performs the operations of addition, subtraction, multiplication, division, and transfer. This section contains three principal components: the X register (X), the Q register (Q), and the accumulator (A).

The computer performs all its operations on binary numbers in the X and Q registers, capable of handling 30-digit numbers. Negative numbers are represented as complements on $2^{30} - 1$. In the accumulator, however, which has capacity for 60 digits, negative numbers are represented as complements on $2^{60} - 1$. The extreme left digit in each register and in the accumulator is used to designate the sign of the number in the register; a 0 represents a positive number; a negative number is represented by 1.

In the accumulator, the sum, difference, product, or remainder is formed, together with other arithmetic operations. The operation of the accumulator is essentially subtractive. This means that every number transmitted to the accumulator is automatically subtracted (modulo $2^{60} - 1$) from the number previously there. A number is added into A by transferring its complement to A.

A number in A may be shifted to the left by as many places as required from 1 to 59. This can be done by a single command. If a number is shifted to the left five places, for example, the digits dropped from the extreme left end appear in the same relative order in the digit positions at the extreme right which have been left vacant by this shift. In this sense, the shift of position of a number in the accumulator is *circular*.

Since the accumulator contains a 60-digit number and since the operation of the accumulator is essentially subtractive with respect to the number introduced, the accumulator contains input gates which convert the 30-digit numbers transferred from the X or Q registers to the appropriate kind of 60-digit number for induction to the accumulator.

The Q register performs several arithmetic and logical operations. In division, for example, the quotient is formed in Q, and in multiplication the multiplier is held in Q. Since it has bilateral communication with the accumulator, the Q register may also be used as a rapidly accessible storage for one 30-digit number.

The X register is principally used to receive numbers transmitted from the magnetic-drum storage system to the arithmetic section. The X register has no adding or digit-positioning shifting property. The addend, the subtrahend, the multiplicand, or the divisor is held in X during the corresponding arithmetic

operations. The number in the X register or its complement may be subtracted into the accumulator.

The arithmetic section also contains a unit known as the *arithmetic shift counter* for counting the number of places by which numbers are shifted in the Q register or in the accumulator. The arithmetic shift counter counts modulo 60.

Program-control Section. The program-control section directs the execution of instructions required in the performance of a mathematical operation. This program of instructions is introduced to the machine and is contained in the magnetic storage system. The program instructions for each operation are expressed as an aggregate of 30 binary digits. In this form, of course, the digits may be stored and transmitted as numerical quantities. Sequential operations in this coding system may be seen by examining the structure of the 30-digit command number. For example, the right-hand 12 digits of a 30-digit instruction represent y, the execution address. This identifies the box in the magnetic storage system which contains the orders for carrying out a particular instruction. The next 12 digits to the left represent the program address which identifies the box containing the next instruction in the program. The remaining six digits to the left represent the command code. This code directs the machine to execute one of the 39 commands which the machine understands.

Associated with the program-control section are three components which receive the three parts of the 30-digit command instruction. The execution address is received in the execution-address register. The program address instructing the machine on the next operation is received in the program-address register, and the command code of six digits in the command-translator switch. This switch energizes one of its 39 output leads appropriate to a given six-digit code.

The program-control section also contains an end-point counter whose primary function is the counting of the number of repetitive routine operations which have been undertaken in a program.

The machine is capable of undertaking 39 operational commands. Without going into detail with respect to each of the 39 commands, it is considered sufficient here to enumerate the number of commands in each of the principal categories. For example, there are 12 additive commands, seven commands

relating to the transmission of data between the accumulator and the magnetic storage system, five commands relating to the transmission of data into and out of the Q register, six commands relating to the arithmetic processes of the arithmetic sequence control, five test commands, and four commands relating to the print/punch register and the termination of machine operation.

The main program control and the arithmetic sequence control operate at a rate of 400,000 pulses per second. This basic frequency is supplied by a clock-pulse generator. The shifting of digit positions in the Q register and in the accumulator is performed at this rate. The addition cycle in the accumulator is 3 cycles of the basic clock frequency, or 3×7.5 microseconds. A new number may therefore be transmitted to the accumulator every 7.5 microseconds, or a number which has been transmitted to the accumulator may be shifted in position after a time of 7.5 microseconds.

10-5. Applications of Large-scale Digital Machines

Large-scale digital machines can be applied to the solution of a variety of mathematical problems relating to the fields of physics, chemistry, applied mathematics, mechanical and chemical engineering, statistics, astronomy,[41] and the biological sciences. The following is a broad classification of equations whose solutions may be obtained by digital techniques:

10-5-1. Systems of Linear Algebraic Equations. These equations occur frequently in problems relating to vibration, chemical analysis, the control of chemical processes in weather problems, and in statistical analysis. In this class are included techniques[26–30,44] for finding characteristic solutions of matrix equations;[33] these matrix relationships occur in quantum-mechanical calculations and also in classical physics dealing with vibration of independent or coupled systems.

10-5-2. Harmonic Analysis. The determination of coefficients and terms in harmonic series has particular importance in celestial mechanics, in radiation problems, and in the determinations of crystal structure from x-ray-, electron-, and neutron-diffraction data.

10-5-3. Mathematical Tables. Several examples of the use of large-scale digital machines in the preparation of mathematical tables may be found in References 27, 36, 40, and 41.

TABLE 10-1.   LARGE-SCALE DIGITAL COMPUTING-MACHINE PROJECTS IN THE UNITED STATES

Name of machine	Builder or primary contractor	Primary contracting agency	Contributing organizations*	Location or intended location	Status
1. IBM Automatic Sequence Controlled Calculator Harvard Mark I	IBM		Harvard University	Harvard Computation Laboratory, Cambridge, Mass.	In operation
2. IBM Pluggable Sequence Relay Calculator	IBM			Two at BRL, Aberdeen, Md., one at Naval Proving Ground, Dahlgren, Va., two at Watson Scientific Computing Laboratory, Columbia University, New York	In operation
3. ENIAC	Moore School of Electrical Engineering	Army Ordnance		BRL, Aberdeen, Md.	In operation
4. BTL Computer, Model V	BTL	Army Ordnance		One at BRL, Aberdeen, Md., one at NACA, Langley Field, Va.	In operation
5. Harvard Mark II	Harvard University	Navy, BuOrd		Naval Proving Ground, Dahlgren, Va.	In operation
6. IBM Selective Sequence Electronic Calculator	IBM			IBM, New York	In operation
7. BINAC	Eckert-Mauchly Computer Corp.	USAF		Northrop Aircraft, Inc. Hawthorne, Calif.	In operation
8. BTL Computer, Model VI	BTL			BTL, Murray Hill, N.J.	In operation
9. Harvard Mark III	Harvard University	Navy, BuOrd		Naval Proving Ground, Dahlgren, Va.	In operation under test at Harvard
10. Whirlwind I	MIT	Navy, ONR		MIT Servomechanisms Laboratory, Cambridge, Mass.	In operation without final storage and terminal equipment

TABLE 10-1. LARGE-SCALE DIGITAL COMPUTING-MACHINE PROJECTS IN THE UNITED STATES. (*Continued*)

Name of machine	Builder or primary contractor	Primary contracting agency	Contributing organizations	Location or intended location	Status
11. EDVAC	Moore School of Electrical Engineering	Army Ordnance		BRL, Aberdeen, Md.	Under test at Aberdeen
12. SDC Raytheon Computer	Raytheon Mfg. Co.	Navy, ONR (Special Devices Center)	BuAer, USAF	To be designated	Under construction
13. IAS Digital Computer†	Institute for Advanced Study, Princeton University	Army Ordnance	AEC, USN, USAF	Institute for Advanced , Study, Princeton, N.J.	Under construction
14. ORDVAC†	University of Illinois	Army Ordnance		BRL, Aberdeen, Md.	Under construction
15. University of Illinois Computer†	University of Illinois			University of Illinois, Urbana, Ill.	Under construction
16. California Digital Computer	University of Calif.	Navy, ONR	University of California	University of California, Berkeley, Calif.	Under construction
17. "Zephyr"	Institute for Numerical Analysis of NBS at Los Angeles	NBS	USAF	NBS, Los Angeles, Calif.	Under construction
18. NBS "Interim" Computer	NBS	NBS	USAF	NBS, Washington	Under construction
19. GE Computer	GE			General Electric Co. Syracuse, N.Y.	Under construction
20. UNIVAC	Eckert-Mauchly Computer Corp.	NBS	Census Bureau, USAF, Army Map Service	To be designated by the contracting and sponsoring agencies	Under construction

* An organization other than the builder or contractor, furnishing funds or services, is listed as a *contributing* organization.
† The three computers marked thus are essentially alike. The IAS machine is the prototype model.

10-5-4. Propagation Problems. This field includes the solution of total and partial linear differential wave equations where boundary conditions and damping may be important. In the propagation of electromagnetic waves, the solution of the wave equation is important for the design of antennae and waveguides of various geometries and shapes. For example, the solution of the wave equation is required to be obtained in elliptic cylinder coordinates. In this coordinate system the wave equation may be transformed into Mathieu's equation. The mathematical-tables project of the Bureau of Standards has recently calculated the characteristic solutions of Mathieu's equation in regions of useful application to physical problems.

10-5-5. Flow Problems. This field of problems[32,36,38,39] includes the analysis related to the flow of compressible fluids in subsonic, transonic, and supersonic speed ranges. The solution of differential equations describing fluid flow is made increasingly difficult when consideration is given to turbulence and viscous effects in a compressible fluid. Because of the increasing importance of both the transonic and supersonic speed ranges, it is necessary that solutions of these equations be obtained in order to provide an interpretation of experimental phenomena and a guide for future research and development on high-speed flight and on the design and testing of missiles required to operate near or above the speed of sound.

Closely related to the above field of problems is the solution of equations describing the flow of heat, either as a steady-state or a transient phenomenon,[44] in bodies of complicated shapes. Cooling problems connecting the heat flow through a surface with compressible fluid flow along it are closely related to the solution of problems of the utmost importance in the engineering and design of high-velocity missiles.

10-5-6. Problems in Nonlinear Mechanics. Problems occur in the field of electrical and mechanical vibrations, in the study of shock-wave propagation, in stability studies, and in the behavior of control systems such as servomechanisms, torque amplifiers, etc., to which large-scale computers may be applied.

10-5-7. Integral and Integrodifferential Equations. Solutions of differential equations describing boundary-value problems may be performed more readily at times by transforming the differential equation into an integral equation. In certain

cases the integral representation of the differential equation may
assume a very simple form. The integral methods using Fourier
and Laplace transforms will adapt certain types of differential
equations to immediate solution. In dealing with more com-
plicated physical problems, however, a machine calculation in
numerical integration may be required. It may also be conven-
ient to express a physical system as an integrodifferential equa-
tion. Both types may require a sufficient number of points
to be calculated as to necessitate the use of a large-scale digital
machine.

10-5-8. Evaluation of Multiple Integrals. It is frequently
required in problems relating to physics, chemistry, astronomy,
and other branches of science that the numerical value for double
and triple integrals is desired. Methods for evaluating these
integrals, essentially extensions of rules developed for single
integrals,[34,35] may be used for machine calculation. In problems
connected with statistical mechanics, it may be required to
evaluate the incomplete gamma function between $n = 15$ to 30
and infinity. For such an application a machine calculation
might be made starting with certain initial values obtained in
Pearson's tables.

In the evaluation of a triple integral, it may be necessary to
perform 1,000 additions of 10-digit numbers which have been
individually computed to provide a value of the integral at
points on a three-dimensional lattice.

10-5-9. Statistical Problems. Many complex problems arise
in the field of statistics.[31] These problems relate to personnel
location, procurement of raw material, import and export
data, census analysis,[42] actuarial problems,[29] and many others.
Although this field of machine application involves the per-
forming of the basic arithmetic operations on data submitted
to the machine, it is also of prime importance that the resulting
information be in a form which is capable of statistical analysis.
The calculation required to transform the input data into final
results is a straightforward problem on most large calculating
machines since it involves presenting to the machine numerical
data with orders to operate on those data with arithmetic proc-
esses. Input devices from which data may be handled statis-
tically include punched cards, punched tape, and magnetic wire or
tape in the order of increasing speed of access of information.

10-5-10. Life-insurance Problems. This field of application[29] for large-scale calculating machines has been chosen here as an example of the usefulness of such calculations applied to business problems. In the life-insurance business, large-scale machines would be useful on policy transactions where a large number of arithmetic and printing operations may be required dealing with interest, loans, dividends, cash surrender values, and extended insurance. In problems of this kind, the machine must be capable of retaining and consulting tabular data relating to insurance rates for various categories of insurance policies and for various age brackets. In cases where machine calculation is exceedingly rapid, the requirement for built-in tabular data is not so stringent. Uses for such machines include pension-plan calculations, the determination of premiums on and values of policies, and the determination of proposed new dividend scales. It is conceivable that the machine could calculate the required result quickly enough to obviate the necessity of building in great quantities of tabular data. In this case, the dividend rate, age of the insured, type of policy, etc., would be introduced to the machine as raw data, and the necessary calculations on each policy transaction would be performed by the machine.

<div align="center">

REFERENCES

Design Characteristics of Digital Computing Systems

</div>

1. Aiken, H. H., and G. M. Hopper, The Automatic Sequence Controlled Calculator—I, *Electrical Engineering*, Vol. 65, Nos. 8 and 9, pp. 384–391 (August–September, 1946); *The Automatic Sequence Controlled Calculator —II, ibid.*, No. 10, pp. 449–454 (October, 1946); *The Automatic Sequence Controlled Calculator—III, ibid.*, No. 11, pp. 522–528 (November, 1946).
2. Alt, F. L., *A Bell Telephone Laboratories Computing Machine—I, Mathematical Tables and Other Aids to Computation*, Vol. III, No. 21, pp. 1–13 (January, 1948); *A Bell Telephone Laboratories Computing Machine—II, ibid.*, No. 22, pp. 69–84 (April, 1948).
3. Bloch, R. M., Mark I Calculator *in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery (Jan. 7–10, 1947), Annals of the Computation Laboratory of Harvard University*, Vol. XVI (Harvard University Press, Cambridge, Mass., 1948), pp. 23–30.
4. Bloch, R. M., R. V. D. Campbell, and M. Ellis, The Logical Design of the Raytheon Computer, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 24, pp. 286–295 (October, 1948). See also General Design Considerations for the Raytheon Computer, *ibid.*, pp. 317–322.
5. Burks, A. W., H. H. Goldstine, and J. von Neumann, *Preliminary Discussion of the Logical Design of an Electronic Computing Instrument.*

Report prepared under Contract W-33-034-ORD-7481 between the Research and Development Service, Ordnance Department, U.S. Army, and the Institute for Advanced Study, Princeton, N.J. (June 28, 1946).

6. Campbell, R. V. D., Mark II Calculator *in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery (Jan. 7-10, 1947), Annals of the Computation Laboratory of Harvard University,* Vol. XVI (Harvard University Press, Cambridge, Mass., 1948), pp. 69–79.

7. Eckert, W. J., The IBM Pluggable Sequence Relay Calculator, *Mathematical Tables and Other Aids to Computation,* Vol. III, No. 23, pp. 149–161 (July, 1948).

8. The Eckert-Mauchly Computer Corporation, *The UNIVAC System* (1947). Pamphlet prepared by The Eckert-Mauchly Computer Corporation, Philadelphia.

9. Engineering Research Associates, Inc., *Selective Alteration of Digital Data in a Magnetic Drum Computer Memory.* Report prepared for ONR, under Contract N6onr-240 Task I with Engineering Research Associates, Inc., St. Paul, Minn. (Dec. 1, 1947).

10. Engineering Research Associates, Inc., *Summary of Characteristics, Magnetic Drum Storage Computer.* Report prepared for the National Bureau of Standards, Department of Commerce, under Contract CST-10133 with Engineering Research Associates, Inc., St. Paul, Minn., Part I (Nov. 30, 1948).

11. Engineering Research Associates, Inc., *Examples of Coding, Magnetic Drum Binary Computer.* Report prepared for the National Bureau of Standards, Department of Commerce, under Contract CST-10133 with Engineering Research Associates, Inc., St. Paul, Minn., Part II (Nov. 30, 1948).

12. Goldstine, H. H., and Adele Goldstine, The Electronic Numerical Integrator and Computer (ENIAC), *Mathematical Tables and Other Aids to Computation,* Vol. II, No. 15, pp. 97–110 (July, 1946).

13. Goldstine, H. H., and J. von Neumann, *Planning and Coding of Problems for an Electronic Computing Instrument.* Report, *The Mathematical and Logical Aspects of an Electronic Computing Instrument,* Part II, Vol. I, prepared under Contract No. W-36-034-ORD-7481 between the Research and Development Service, U.S. Army Ordnance Department, and the Institute for Advanced Study, Princeton, N.J. (1947).

14. Goldstine, H. H., and J. von Neumann, *Planning and Coding of Problems for an Electronic Computing Instrument.* Report, *The Mathematical and Logical Aspects of an Electronic Computing Instrument,* Part II, Vol. II, prepared under Contract No. W-36-034-ORD-7481 between the Research and Development Service, U.S. Army Ordnance Department, and the Institute for Advanced Study, Princeton, N.J. (1948).

15. Hartree, D. R., *Calculating Machines* (Macmillan & Co., Ltd., London, 1947).

16. Hartree, D. R., The ENIAC, an Electronic Computing Machine, *Nature,* Vol. 158, No. 4015, pp. 500–506 (Oct. 12, 1946).

17. Harvard University, Staff of the Computation Laboratory, *A Manual of Operation for the Automatic Sequence Controlled Calculator, Annals of the*

Computation Laboratory of Harvard University, Vol. I (Harvard University Press, Cambridge, Mass., 1946). This manual contains an extensive bibliography on numerical analysis, pp. 338–396.

18. International Business Machines Corporation, *IBM Automatic Sequence Controlled Calculator* (1945). Pamphlet of the International Business Machines Corporation, New York.

19. International Business Machines Corporation, *IBM Selective Sequence Electronic Calculator* (1948). Brochure of IBM Corporation, New York. IBM Selective Sequence Electronic Calculator Is Dedicated to Science, *Business Machines*, Vol. 30, No. 11, pp. 1–12 (March, 1948).

20. von Neumann, J., *The Principles of Large-Scale Computing Machines* (mimeographed by ORI, Navy Department No. 6553) 16 pp. Address delivered at a meeting of the Mathematical Computing Advisory Panel of the Office of Research and Inventions, May 15, 1946.

21. Stibitz, G. R., *Relay Computers*, Report 171.1 R of the Applied Mathematics Panel of the National Defense Research Committee (February, 1945).

22. United States Air Force, *Scientific Planning Techniques*, Discussion Paper 1-DU, Project SCOOP (Planning Research Division, Director of Program Standards and Cost Control, Washington, D.C., Aug. 5, 1948).

23. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *Report on the ENIAC*, *Technical Report I* (University of Pennsylvania, Philadelphia, June 1, 1946).

24. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *Functional Description of the EDVAC*, *Research Division Report* 50-9, a report of development work under contract W-36-034-ORD-7593 with the Ordnance Department, Department of the Army (University of Pennsylvania, Philadelphia, Nov. 1, 1949). 2 vols.

25. Williams, S. B., Bell Telephone Laboratories Relay Computing System *in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery (Jan. 7-10, 1947)*, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI (Harvard University Press, Cambridge, Mass., 1948), pp. 41–68.

Applications of Digital Machines

26. Alt, F. L., The Multiplication of Matrices, *Mathematical Tables and Other Aids to Computation*, Vol. II, No. 13, pp. 12–13 (January, 1946).

27. Alt, F. L., A Bell Telephone Laboratories Computing Machine—II, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 22, pp. 81–84 (April, 1948).

28. Andrews, E. G., and H. W. Bode, Use of the Relay Digital Computer, *Electrical Engineering*, Vol. 69, No. 2, pp. 158-161 (February, 1950).

29. Berkeley, E. C., Electronic Machinery for Handling Information and Its Uses in Insurance, *Transactions of the Actuarial Society of America*, Vol. 48, Part 1, pp. 36–52 (1947).

30. Bleick, W. E., Calculating Machine Solution of Quadratic and Cubic Equations by the Odd Number Method, *Mathematical Tables and Other*

Aids to Computation, Vol. II, No. 20, pp. 321–324 (October, 1947). Method for possible application to Mark I Calculator or to the ENIAC.

31. Bush, V., As We May Think, *Atlantic Monthly*, Vol. 176, No. 7, pp. 101–108 (July, 1945).   A high-speed (microfilm) selector of the Memex type, suggested in this article, has been constructed by Engineering Research Associates, Inc., for the Department of Agriculture.   Preliminary tests indicate satisfactory performance.

32. Clippinger, R. F., Airflow Problem Planned for the ENIAC, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 23, pp. 206–207 (July, 1948).   Summary of paper presented at the meeting of the Association for Computing Machinery held at Aberdeen Proving Ground, Aberdeen, Md., Dec. 11–12, 1947.

33. Curry, H. B., and W. A. Wyatt, *A Study of Inverse Interpolation on the ENIAC*, Report 615 (Ballistic Research Laboratories, August, 1946).

34. Harvard University, *Evaluation of the Function*

$$S(b_1h) \;=\; \int_0^h \sin\,(x^2 + b^2)^{1/2}\,dx/(x^2 + b^2)^{1/2}$$

Report 2, Computation Laboratory Reports for the Bureau of Ships, Navy Department (June, 1944); *Evaluation of the Function*

$$C(b_1h) \;=\; \int_0^h \cos\,(x^2 + b^2)^{1/2}\,dx/(x^2 + b^2)^{1/2}$$

Report 10 Computation Laboratory Reports for the Bureau of Ships, Navy Department (October, 1944).

35. Harvard University, Computation Laboratory, *Fifteen-Place Table of e^{-x}*, Publication 20, Computation Laboratory for the Bureau of Ships (July, 1945).   Out of Print.

36. Harvard University, Staff of the Computation Laboratory, *Tables for the Design of Missiles, Annals of the Computation Laboratory of Harvard University*, Vol. XVII (Harvard University Press, Cambridge, Mass., 1948).

37. Harvard University, Staff of the Computation Laboratory, *Tables of the Modified Hankel Functions of Order One-Third and of Their Derivatives, Annals of the Computation Laboratory of Harvard University*, Vol. II (Harvard University Press, Cambridge, Mass., 1945); *Tables of the Bessel Functions of the First Kind of Orders Zero and One, ibid.*, Vol. III (1947); *Tables of the Bessel Functions of the First Kind of Orders Two and Three, ibid.*, Vol. IV (1947); *Tables of the Bessel Functions of the First Kind of Orders Four, Five, and Six, ibid.*, Vol. V (1947); *Tables of the Bessel Functions of the First Kind of Orders Seven, Eight, and Nine, ibid.*, Vol. VI (1947); *Tables of the Bessel Functions of the First Kind of Orders Ten, Eleven, and Twelve, ibid.*, Vol. VII (1947); *Tables of the Bessel Functions of the First Kind of Orders Thirteen, Fourteen, and Fifteen, ibid.*, Vol. VIII (1947); *Tables of the Bessel Functions of the First Kind of Orders Sixteen Through Twenty-Seven, ibid.*, Vol. IX (1948); *Proceedings of a Symposium on Large-Scale Digital Calculating Machinery, ibid.*, Vol. XVI (1948).

38. Hicks, B. L., and H. G. Landau, Nonlinear Parabolic Equations, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 23, pp. 207–208 (July, 1948). Summary of paper presented at the meeting of the Association for Computing Machinery held at Aberdeen Proving Ground, Aberdeen, Md., Dec. 11–12, 1947.

39. Holberton, J. V., Laminar Boundary Layer Flow in a Compressible Fluid, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 23, p. 208 (July, 1948). Summary of paper presented at the meeting of the Association for Computing Machinery held at Aberdeen Proving Ground, Aberdeen, Md., Dec. 11–12, 1947.

40. International Business Machines Corporation, *IBM Selective Sequence Electronic Calculator* (1948) p. 3. Pamphlet of the IBM Corporation, New York.

41. Irwin, J. B., The Expected Performance of the EDVAC on Some Astronomical Problems, *Publications of the Astronomical Society of the Pacific*, Vol. 60, No. 355, pp. 235–244 (August, 1944).

42. McPherson, J. L., Applications of Large-Scale High-Speed Computing Machines to Statistical Work, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 22, pp. 121–126 (April, 1948).

43. Mitchell, H. F., Jr., Inversion of a Matrix of Order 38, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 23, pp. 161–166 (July, 1948).

44. Seeger, R. J., On Computational Techniques for Certain Problems in Fluid Dynamics, *in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery (Jan. 7–10, 1947)*, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI (Harvard University Press, Cambridge, Mass., 1948), pp. 157–167.

CHAPTER 11

ANALOG COMPUTING SYSTEMS

11-1. Introduction

Calculating machines of the analog type operate on physical quantities which can be measured, instead of on numbers which can be counted. When a problem expressed in terms of discrete digital quantities is to be solved by an analog machine, these quantities are translated from numerical into physical form, *i.e.*, into light intensities, lengths, angular positions of shafts, electrical currents, and the like. Computing operations are performed on these physical quantities. The computed result also appears as a physical quantity which may be retranslated into digital form.

Let us consider the concept of translation of digital data to physical form in terms of an example. Suppose the input to an analog computer is by manual settings of a continuously variable control dial. The operator turns the dial to the scale division corresponding to the number he wishes to insert. In setting the dial he rotates its shaft a number of degrees from the zero position. The angular position of the shaft is then a physical analog of the number it represents. It may, also, provide the means for setting a continuously variable rheostat and thus establish the value of a current. If so, the number on the dial is represented by the effective value of the variable resistance and by the amount of current flowing through it as well as by the position of the shaft.

From this example it is easy to visualize an analog machine which receives its input information directly from some other machine, without an operator's intervention. Such devices are common. For example, the computer which solves the problem of a shell's trajectory and directs the aiming of an antiaircraft gun may take its inputs directly from the target-sighting apparatus. It should be noted that such a computer

constantly solves the same problem, yielding output values which correspond with continuously changing input values.

In discussing the question of relative freedom from error, as applied to computed results, it is well to start with a common understanding of the terminology. The word *accuracy*, as used herein, denotes conformity to fact. *Precision* denotes sharpness of definition. Thus, the statements $\pi = 3.1416$ and $\pi = 3.141592654$ are equally accurate, but the latter is more precise. The first statement merely expresses the observation that the ratio of the circumference of a circle to its diameter is between 3.14155 and 3.14165 (as noted in Sec. 6-9). This statement is free from error; it conforms to fact. The latter, more precise statement ($\pi = 3.141592654$) likewise conforms to fact.

Suppose we are dealing with a computed result instead of with a direct measurement. Assuming complete freedom from arithmetic error, the precision with which we are justified in expressing the result depends upon the precision with which the input data were expressed and upon the nature of the functional relationship between the input and output data.

In discussing the ability of a machine to yield computed results free from error, both of the terms we have just defined are useful. A computer is accurate which does not make mistakes. A digital computer yields highly precise computed results if it performs its operations accurately, if it carries a large number of digits in the register from which the answer is read, and, at the same time, is able to avoid round-off errors (see Sec. 6-12). Depending upon the nature of the functional relationships involved in a problem, the same computer may yield results which we are justified in expressing with greater or less precision.

An analog computer is accurate if the relationships between its parts are valid analogs of the mathematical relationships in the problem. It yields precise computed results if its parts are made with precision and if the answers can be read with precision. The last part of this statement is particularly obvious, but nevertheless instructive. An analog computer yields an output in the form of a measurable physical quantity of some sort. If a measure of this quantity must be translated to digital form, the question of the precision with which the measurement can be read arises. The computer is similar in this respect to

any other measuring apparatus. The result of the computation can be expressed no more precisely than it can be read. Moreover, the entire operation of the analog computer is based on measurement. Though only the final quantity, the answer, may actually be read and translated into digital expression, other measurements are made by the interaction of cams, gears, currents, etc., at each step in the solution, and the precision with which each of these measurements can be made depends upon the niceties with which the machine is fabricated and with which its components interact.

In the case of the fire-control computer mentioned above, we have a machine from which we do not attempt to read results. We merely feed the results as controlling and directing orders to other equipment. The same problem is being solved continuously; yet the input values are changing. Because this is true, the ability of the machine to follow changes becomes important. If there is an appreciable time lag between the incidence of a change in input data and its effect on the output, the instantaneous value of the computed result will be in error; *i.e.*, the machine will be temporarily inaccurate. Generally speaking, such an analog machine's instantaneous accuracy depends upon its ability to follow these changes, a machine property which is unimportant when we are dealing with a static analog device like the slide rule.

11-2. Components for Analog Computation

11-2-1. Arithmetic Elements. *Addition.* In most mechanical analog machines, addition is performed by differential gears.[11] There are several types of such differential gears, and since these have been described adequately elsewhere in the literature, it will suffice here merely to include the names of several of these types: the bevel-gear differential, the cylindrical-gear differential, the spur-gear differential, and the screw differential. In a differential gear adder the quantities to be added are represented by shaft rotations; the sum of the augend and the addend is also a shaft rotation. For example, if it is desired to add X and Y and find the sum Z, X is the angular rotation of a given shaft from a predetermined zero position. For Y the rotation of a similar shaft on the other differential gear is used. The sum Z is the angular rotation of the shaft on a planetary gear which

bears on both the gears to which the shafts of the quantities X and Y are related.

Subtraction. Subtraction is performed generally by differential gear mechanisms in the same manner as addition, except that the appropriate sign of rotation of the shafts is altered to provide the difference $X - Y$.

Multiplication. There are several methods by which multiplication may be performed with analog equipments. Multiplication may be performed with a slide multiplier, with a

$$\frac{y}{x} = \frac{z}{k_1}$$

$$y = \left(\frac{1}{k_1}\right)^{xz}$$

Fɪɢ. 11-1. Slide multiplier. The bases (k_1 and X) and the altitudes (Z and Y) of the two right triangles are proportional because the triangles are similar. The arm forming the hypotenuse of the larger triangle is pivoted at their common vertex.

nomographic multiplier, with a resolver, or by special adaptation of a wheel-and-disk integrator. Some analog machines contain both slide multipliers and integrators.

The slide multiplier[11,32] is based on the principle of the proportionality of the sides of two similar triangles (see Fig. 11-1). The first triangle has a base of length k_1 and an altitude Z; the second triangle has a base of length X and an altitude Y. By the relationship between similar triangles,

$$\frac{k_1}{X} = \frac{Z}{Y} \qquad \text{or} \qquad Y = \frac{1}{k_1}(XZ) \qquad (11\text{-}1)$$

This type of mechanism has been used by the General Electric Company in the design of their differential analyzer.[22]

Multiplication can also be performed by an integrating mechanism; this is an adaptation of the Kelvin wheel-and-disk

integrator[9,32] (see Fig. 11-2). A sharp-edged wheel resting on a disk is driven by frictional contact with the disk; the axes of rotation of the wheel and of the disk are orthogonal. The radial position of the integrating wheel resting on the disk may be varied by a lead screw through the journal on which the wheel is mounted; this changes the relative speed of rotation of the shafts connected to the integrating wheel and to the disk. In this integrator there are two input shafts and one output shaft.
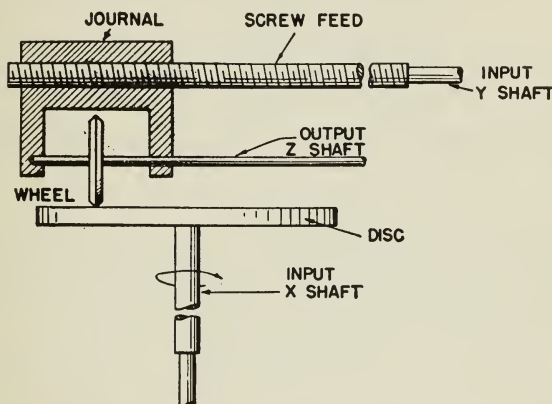
FIG. 11-2. Wheel-and-disk integrator. Rotation of Z, in angular measure from an initial reference position, is proportional to $\int Y\, dX$.

The input motion X is made to rotate the *disk shaft*, the angular rotation of this shaft representing the value of the quantity X at all times. Input motion Y rotates the lead screw controlling the *radial position of the wheel* on the disk. The output of the integrator is the shaft rotation Z of the wheel. If k_2 is a constant of the integrator, then, at any instant, the relationship represented may be expressed as

$$\frac{dZ}{dX} = k_2 Y \tag{11-2}$$

or

$$Z = k_2 \int Y\, dX \tag{11-3}$$

If two integrators are used, their combination can be made to provide the product XY. In the first integrator the variable X is represented by the rotation of the disk shaft, the variable Y by the rotation of the lead screw of the integrating wheel. In the second integrator the positions of X and Y are reversed. The

outputs Z_1 and Z_2 of the two integrators are then added to give the product XY.

$$XY = Z_1 + Z_2 = \int X\, dY + \int Y\, dX \qquad (11\text{-}4)$$

The factor k_2 may be eliminated by adjusting the outputs of each integrator by this constant.

Where the two variables appear within the integral,

$$\int XY\, d\phi$$

two integrators will provide the required product.   Since

$$\int XY\, d\phi \equiv \int X\, d(\int Y\, d\phi) \equiv \int X\, dS \qquad (11\text{-}5)$$

the integrator, with ϕ and Y as input variables, supplies the quantity

$$S = \int Y\, d\phi \qquad (11\text{-}6)$$

The variable S is then introduced to the second integrator as a shaft rotation of the disk, and X is introduced as the lead-screw rotation to the wheel carriage, so that the result

$$R = \int XY\, d\phi \qquad (11\text{-}7)$$

is generated by the output of the second integrator.

For further details on special operations, such as the use of the wheel and disk as an *inverse integrator*, refer to Bush and Caldwell (pp. 316–318).[9]

Division.   Division is performed either by the slide multiplier[32] or by the integrator.[9]   If the slide multiplier is used, the inputs X and Y are interchanged to give

$$X = k_1 \left(\frac{Y}{Z}\right) \qquad . \qquad (11\text{-}8)$$

Division may also be performed with two integrators by taking the reciprocal of one of the variables, Y, for example, and using this reciprocal as the input to the Y (lead-screw) shaft in the first integrator and to the X (disk) shaft in the second integrator:

$$\frac{X}{Y} = \int X\, d\left(\frac{1}{Y}\right) + \int \frac{1}{Y}\, dX \qquad (11\text{-}9)$$

11-2-2. Differentiation and Integration.   Both differentiation and integration are performed with the Kelvin wheel-and-disk

integrator. The method of its use is described in Sec. 11-2-1 above under Multiplication.

11-2-3. Input and Output. Numbers and orders to a differential analyzer may be introduced by means of coded punched tape.[8] The orders for machine setup include those relating to the shaft connections and to the gear ratios required for the solution of a given problem; numbers representing the initial conditions may also be introduced on punched tape, or in the case of earlier units, these may be set into the machine manually.

Where input of a function in graphical form is required, an *input table*[9,22,32] is used. This is usually a flat platform on which the required function is plotted in cartesian coordinates. The input table is equipped with a pointer for following the plotted function; if $f(x)$ is plotted along the ordinate and x along the abscissa, the ordinate position is often controlled manually by a feed-screw mechanism, while the traverse along the abscissa is driven by a motor. The x drive on the input table is taken from the basic x drive of the machine and must always be in synchronism with it so that the correct value of $f(x)$ is introduced to the other elements (adders, multipliers, integrators, etc.) in the proper phase. Automatic curve-following devices have been reported by *MIT*[18] and the University of California at Los Angeles.[32]

The output of a differential analyzer may be punched on paper tape or printed in tabular form by an electrical typewriter. Frequently, it is convenient to plot the output graphically. An *output table*,[9,22,32] similar in design to the input table described above, is used for this type of output and consists of a flat table equipped with a frame for support and manipulation of a stylus. The function $f(x)$, or y, is plotted automatically on the ordinate and the independent variable, x, on the abscissa.

11-2-4. Generation of Functions. Functions in graphical form may be introduced to the machine through the input table. However, when certain simple analytic functions are required as elements in the solution of a differential equation, they are frequently generated within the analyzer itself.

The square of a function may be generated by one integrator. Since $y = x^2$, y may also be expressed as

$$y = \int 2x \, dx \tag{11-10}$$

$2x$ is introduced to the lead screw on the wheel carriage and x to the disk shaft.

The function $y = e^x$ may be generated with one integrator by feeding the output (rotation of the integrating wheel) back into the shaft on the carriage, since

$$y = e^x; \qquad \frac{dy}{dx} = e^x; \qquad y = \int e^x \, dx \qquad \text{and} \qquad y = \int y \, dx$$

(11-11)

By feeding $y = 1/x$ into an integrator $z = \ln x$ can be generated. If n times the output, nz, is fed into another integrator, the function $y = x^n$ is generated, since

$$dy = nx^{n-1} \, dx = nx^n \frac{dx}{x}$$
$$dy = y \, d(n \ln x)$$
$$y = \int y \, d(n \ln x)$$

(11-12)

In this operation, the output y of the integrating wheel is fed back as the input to the lead screw on the wheel carriage.

For a more detailed discussion of these connections, see articles by Amble,[1] Bush,[8] Bush and Caldwell,[9] Hartree,[15] and Shannon.[30]

11-3. The MIT Differential Analyzers

The first large-scale differential analyzer[8] was constructed at the Massachusetts Institute of Technology and put into operation there in 1930. The basic design of this machine followed exploratory work at MIT on numerical methods[5,6] during the period 1925 to 1930. The original differential analyzer operated on mechanical principles. This machine has been duplicated at a number of laboratories in this country and abroad.

In 1945 a new type of differential analyzer, completed and put into operation during the war at MIT, was announced by Bush and Caldwell.[9] The new machine made automatic many of the manual operations of the earlier mechanical model and increased the capacity and precision. It is appropriate here to discuss the principal features of this machine in terms of the types and number of components it contains for computation.

11-3-1. Input and Output. Numbers and orders are introduced to the machine[9] by coding on three punched-paper-tape units. These tapes are designated A, B, and C. The A tape

supplies directions for interconnecting the computing units of the machine. The B tape is punched with gear-ratio data; the C tape contains the initial numerical data with which each of the computing units must be supplied to start the solution of a given equation. Each tape is punched with data identifying the problem under investigation.

The A tape, with the interconnection data, uses a four-digit decimal code to describe the shaft acting as a source (transmitter) and the shaft which is driven (receiver). The first two digits of the code designate the type of shaft to be used; the second two digits specify which one of the computing elements possessing such a shaft is to be chosen. For example, in a hypothetical case, the digit group 2802-3410 might specify that the output shaft (code 28) of the second integrator (code 02) is to drive an input shaft (code 34) on the tenth gearbox (code 10). Each transmitter-receiver connection appears on a *frame* of the coded tape.

The B tape, with its punched code representing the gear-ratio orders, contains numerical gear-ratio data and the numbers identifying the particular unit to which each order and number applies.

The C tape carries coded numbers representing the initial conditions of the problem and the coded identification of the units which will make use of these numerical conditions.

For input of graphical data, a modified form of function unit has been developed. The plotted function is mounted on a drum; the drum is rotated by the input, or in accordance with the independent variable, while the pointer, or index, is kept on the trace of the function by manual rotation of the lead screw which moves the pointer along the direction of the axis of rotation of the drum. Since the output of the function unit must be a source of torque for relaying its information to other elements of the machine, the operator supplies this torque while acting to keep the index on the function.

The output of the differential analyzer is a measurement of the rotation of shafts representing the desired variables. For example, if the dependent variable, y, is the result of integrating a differential equation with x as the independent variable, the output of the y shaft can be made to control the rotation of the screw feed geared to the stylus on the output table while the main

drive which rotates the drum is supplied with torque by the x, or independent-variable, shaft.

For numerical registration of the shaft rotation, a counter unit is servoconnected to each shaft whose output is desired. Shaft rotations may be measured simultaneously on a number of units at any time and without stopping the machine. These data are stored in groups of relay counters. The system can handle up to $\pm 10,000$ revolutions and can record to 0.1 revolution. The measurement of shaft rotation is first corrected for its initial value by a relay matrix capable of performing addition and subtraction; the corrected data are then transferred from the matrix to storage relays prior to introduction to the automatic typewriters. The printing unit is automatic motor-driven IBM equipment. A set of 10 seven-digit numbers (with spaces and algebraic sign) can be printed every 12 seconds.

11-3-2. Integrators. The new differential analyzer[9] has 18 integrators at present; this number of integrators may be increased to 30. In view of the possible expansion to 30 integrators, the machine has already been arranged to operate as three sections, each having its own recording channel, typewriter units, computing matrices, decoding relays, and typewriter storage relays.

The disk of the integrator is constructed of plate glass ground flat; the integrating wheel has a steel hub with a magnesium flange and a hardened-steel tire finished to a sharp edge (with straight sides, 60° included angle). The edge of the wheel bearing on the glass disk is ground to a width of 0.002 inch.

Each integrator has an automatic setting mechanism which consists of a group of rotary selector switches for adjusting the integrator to the required initial position in conformance with coded orders on the C tape. The position of the integrating wheel may be set on the radius of the disk to a precision of 1 part in 150,000. The setting mechanism through its 44 associated relays may also be controlled centrally to read the position of the lead screw on the wheel carriage and transmit the data to a printing station. If the integrating-wheel displacement is to be multiplied or divided by a constant, *e.g.*, 2, 5, or 10, the setting mechanism reads the wheel position and transmits the reading to a relay multiplier-divider unit, which performs the required arithmetic operation. The product or quotient is then trans-

ferred back to the setting mechanism, which automatically adjusts the integrator to the correct position.

11-3-3. Gear Assemblies. A problem on the differential analyzer is set up and ready to operate when each of the required shafts is connected to a source of torque. On the mechanical differential analyzer, the interconnection of these shafts was made manually. In the new machine, clusters of gears are assembled into *gearboxes* provided with electromagnetic shifting devices for change of gear ratio. The gearboxes are of two types: a one-digit decimal box provides any gear ratio between 0.1 and 1.0 in steps of 0.1; a decade decimal box provides any gear ratio between 0.0000 and 1.1110 in steps of 0.0001. The gearboxes are each driven by a standard servomotor; the output of the gearbox is registered by an *angle indicator*.

11-3-4. Machine Control. The machine[9] is controlled through an automatic electrical switchboard which regulates the interconnection of data-transmission elements. Servomechanisms [3,4,7,10,12,17,20] provide the power for driving all mechanical loads. Instead of measuring shaft rotations by selsyns,[24] through magnetic-field variation, an angle-indicator unit operating on the variation of an electric field was found to be feasible; this unit operates at low torque and is capable of being used in conjunction with the standard telephone switching equipment for the transmission of data at low energy levels.

The angle indicator[9] contains two capacitance bridges, each excited with voltage at about 3,000 cycles per second. Each bridge has a pair of fixed and a pair of variable condensers. The angular position is parallel. Rotating between the two plates of the variable condenser are a pair of specially cut cam-shaped disks of dielectric material; these are fastened to the shaft whose rotation is to be measured. As the shaft is rotated, each bridge undergoes a change in capacity such that the output voltages of bridge 1 and bridge 2 are

$$e_1 = kE_1 \sin \phi$$
$$e_2 = kE_2 \cos \phi$$
(11-13)

where e_1 = output of first bridge

e_2 = output of second bridge

E_1 and E_2 = exciting voltages

ϕ = angle of shaft rotation

k = a constant of the system

The outputs e_1 and e_2 may be used to transmit the position of one shaft A to another shaft B. In this case, the two bridges of A are excited with the same alternating voltage E; the output voltages e_1 and e_2 are transmitted through a switchboard and are introduced as input voltages to the angle indicator on shaft B. Since the output voltages from A are

$$e_1 = kE \sin \phi_A$$

and   (11-14)

$$e_2 = kE \cos \phi_A$$

the output of the capacitance bridges associated with shaft B will be

$$e_3 = ke_1 \cos \phi_B$$
$$e_4 = ke_2 \sin \phi_B$$ (11-15)

Combining Eqs. (11-14) and (11-15),

$$e_3 = k^2E \sin \phi_A \cos \phi_B$$
$$e_4 = k^2E \cos \phi_A \sin \phi_B$$ (11-16)

The output voltages e_3 and e_4 from angle indicator B are then subtracted in a transformer network. The transformer output is then

$$e = e_3 - e_4 = k^2E \sin (\phi_A - \phi_B)$$ (11-17)

The voltage e, for small differences in shaft positions, is proportional to the difference between the angular positions of the shafts A and B; this voltage is made the input to a servomotor which supplies the torque to control the shaft B.

The servomotors used are of the repulsion type with two sets of brushes, one to provide positive, the other to provide negative rotation. The torque delivered by the servomotors is controlled by varying the plate current through each of a pair of thyratrons which are connected to the brush pairs through step-down transformers.

The three different sections of the machine are each equipped with an independent variable drive; any one of these drives, however, may be used with all or any part of the computing elements of the entire machine. The power element for the independent variable drive is a variable-speed motor with an associated angle indicator. This angle indicator is associated with a primary element for machine control; it can drive any num-

ber of receiving angle indicators (11 have been so driven). The speed and acceleration of this shaft are controlled automatically.

11-3-5. Relay Equipment. Several thousand relays are in use at present in the new differential analyzer. These relays are of the telephone type. They are used for decoding and executing input-tape orders, controlling the electrical typewriter output, performing addition and subtraction in the registration of shaft rotations, multiplication and division for integrator

TABLE 11-1. TABLE OF EQUIPMENT

The new differential analyzer contains the following equipment:

Independent variable units	3
Integrators and setting mechanisms	18
High-speed numerical counters	10
Numerical recording channels (including special typewriters)	3
Manual function units (these units can also be used for graphical recording)	3
Graphical recording unit	1
Automatic function units (not yet in operation)	4

Gear combinations—all gearboxes can be used separately or with associated adder gears. The details of gearing available are given in the following table.

TABLE 11-2. GEAR TABLE

Type of unit and gearboxes contained	No. of units	Total no. of decade boxes*	Total no. of one-digit boxes†	Total no. of adders
3-input adder unit, with 2 decade boxes and 1 direct drive	2	4	0	4
2-input adder unit, with 1 decade box and 1 one-digit box	6	6	6	6
2-input adder unit, with 1 decade box and 1 direct drive	3	3	0	3
2-input adder unit, with 1 one-digit box and 1 direct drive	3	0	3	3
Nonadding unit, with 1 decade box and 1 one-digit box	3	3	3	0
Grand totals	17	16	12	16

* Decade boxes provide all ratios from 0 to 1.1110, in steps of 0.0001.
† One-digit boxes provide all ratios from 0.1 to 1.0, in steps of 0.1.

positioning, counting of shaft rotations, and for temporary storage.

Tables 11-1 and 11-2 summarize the equipment in the new differential analyzer at MIT. These data were presented by Bush and Caldwell (p. 326).[9]

11-4. The General Electric Differential Analyzer

An improved type of differential analyzer has also been designed and built by the General Electric Company.[22] One model of this analyzer is in current use at the University of California at Los Angeles.[32]

The general mechanical features of the components of this analyzer differ in several respects from those in the new differential analyzer at MIT. For example, multiplication and division are performed generally with a mechanical multiplier by using similar-triangle relations (see Multiplication in Sec. 11-2-1), although the two-integrator method for obtaining products may still be involved.

The integrator output shaft is not in physical contact with the integrating wheel but is controlled by a photoelectric follow-up system. Since it is not connected to the output shaft, the wheel experiences no reaction from the torque drive of that shaft; the only torque on the wheel itself is that imposed upon its own bearings. The integrator wheel is a Polaroid disk with a steel rim and hub; the wheel itself forms part of the optical system. A split beam of light passes normally through the wheel face, one beam traversing one of two additional disks and the other beam traversing the second disk. The two disks are set with their planes of polarization at 90° to one another; they are mounted on a common shaft geared to a follow-up motor. The shaft of this motor simulates the rotation of the integrator wheel.

After passing through the integrator wheel and one of the polarized disks, each beam falls upon a photocell. As the integrator wheel rotates, the illumination on the two photocells is varied. The photocells form two arms of a d-c bridge; the output, or current difference between the photocells, causes one thyratron to conduct more and another thyratron to conduct less than its previous current output. The thyratrons both supply current to a pair of field coils on a split-series servomotor. The

sign of the difference in the current outputs controls the direction of rotation of the motor. Therefore, an unbalance in the light striking each photocell is transmitted by thyratron control so as to correct this unbalance by changing the angular position of the polarized disks. This method of balancing is used to keep the shaft of the servomotor at the same angular position as the shaft on the integrator wheel.

Table 11-3 gives a list of components in the differential analyzer at the University of California at Los Angeles; this material is presented in Reference 32 (p. 2). For the design characteristics

TABLE 11-3. COMPONENTS OF THE DIFFERENTIAL ANALYZER[9] AT THE
UNIVERSITY OF CALIFORNIA AT LOS ANGELES

Adders (also subtract)	20
Multipliers (also divide)	3
Integrators	14
Input tables	4
Independent variable motors	2
Output tables	2

of other differential analyzers, see articles by Beard,[2] Hartree,[13,16] Hartree and Porter,[14] Lennard-Jones, Wilkes, and Bratt,[23] Massey, Wylie, and Buckingham,[25] and Murray.[26]

11-5. Electronic Differential Analyzer

Recently attention has been given to performing differentiation, integration, and arithmetic operations with electronic circuits.[19,21] Analogous mechanical methods have been outlined previously, and these preceded the development of electronic units for differential analysis. The basis for all the electrical analog computers is the mathematical relation between currents and voltages in electric circuits. For example, the voltage drop across a resistor is proportional to the algebraic sum of the currents which are introduced from various sources. Thus, Kirchhoff's law is the basis for an electrical adder.

In general, the order of magnitude of error in an analog computing device is closely associated with the ratio of output energy produced to input energy required. For this reason, the electron tube is a valuable aid to electrical computing because it requires very little input energy. The use of voltage or current amplifiers[29] in an electrical computer is somewhat analogous to the use of torque amplifiers in mechanical computers.

In the electronic differential analyzer the arithmetic operations are performed in the following way: Addition and subtraction can be accomplished through the use of two identical triodes in parallel in a resistance-coupled amplifier: The plate or output potential can be made equal to the sum or difference of the input (or grid) potentials multiplied by the factor

$$\frac{-\mu R_L}{R_L + R_P}$$

where μ = amplification factor of tube

R_P = plate resistance of tube

R_L = load resistance

Multiplication by a constant factor is performed by a single stage of a resistance-coupled amplifier.[28]   With proper adjustment of the operating point, the plate potential is a linear function of the input or grid potential multiplied by the factor $-\mu R_L/(R_L + R_P)$.   It should be noted that this factor can be greater or less than 1.   Thus, the circuit can be used to perform division as well as multiplication by a constant factor.

Multiplication by a variable factor may be performed by introducing a voltage which is proportional to the multiplier across the outside terminals of a potentiometer.   The potentiometer shaft is adjusted to an angle proportional to the multiplicand, and the output voltage is proportional to the product of the two input quantities.   Since both input voltage and shaft angle may be variable quantities, the product of two variable quantities may be formed.

Varying electrical quantities can be generated with an input table similar to the ones used in mechanical differential analyzers. In the case of an electrical machine, the shafts which drive the curve-following stylus are connected to potentiometers and cause an output voltage to vary in accordance with a plotted function. If a function is to be used frequently, a potentiometer can be wound on a card which varies in width in such a manner that the output voltage is the desired function of the shaft rotation. This device is the electrical equivalent of a mechanical cam.

In order to perform differentiations and integrations, the properties of RC circuits may be employed.   An RC circuit in conjunction with an electronic amplifier can be made to have a time constant of about 2 minutes.   Thus, an integration could be

performed for only approximately 3 seconds before a 5 per cent error caused by the leakage of charge from the capacitor is accumulated. By adding a *feedback*[27] circuit in the amplifier, the time constant of the RC circuit can be increased to about 100 minutes, and integration may be performed for 5 minutes before a 5 per cent error is accumulated.

A one-stage conventional resistance-capacity amplifier represents a first-order differential equation with constant coefficients. The plate potential can be shown to be related to the grid potential by a first-order differential equation with a time constant $[R_G + R_L R_P/(R_L + R_P)]C$, where C is the value of the coupling condenser and R_G is the grid resistor in the next stage.

Using n stages of resistance-capacity amplification, equations of the nth order can be simulated, and with the proper choice of circuit constants and tubes, differential equations with real negative coefficients can be set up. By using the principle of feedback the range of coefficients can be extended greatly. Also by using the feedback principle, a differential equation can be set up which has an auxiliary equation with complex roots whose real parts are either positive or negative.

The output potentials of the machine are measured by means of a recording vacuum-tube voltmeter, and it has been suggested that a self-balancing potentiometer of the Brown and Leeds-Northrup type can be used to drive a plotting table.

A commercial electronic differential analyzer is being produced by the Reeves Instrument Corporation, and is called the REAC.[33] The basic machine contains 20 d-c amplifiers, of which seven are integrating amplifiers, seven are inverting amplifiers, and six are summing amplifiers. Servomechanisms are used to carry out multiplication of variable quantities, to resolve vectors, and to introduce arbitrary functions. Provision is made to couple several REAC's together in order to increase the capacity. A single machine will handle differential equations up to the seventh order.

A circuit which can be used as a square or square-root computer element (as well as in various other applications) has been reported by R. D. Campbell.[68] It comprises a d-c amplifier, a feedback loop, and a diode, operated in the region where the slope of the plate-current–filament-current characteristic is relatively steep. The feedback is designed to maintain constant power in

the diode filament.   When the circuit is to be used as a computer element, the input may be an a-c current fed through the filament. The output is the output of the feedback amplifier.   The power generated in the filament is the sum of the mean squares of the input and output currents times the resistance of the filament. This is kept constant.   Assuming the filament resistance is constant, which is a valid assumption so long as the frequency of the input current is not higher than about 10 megacycles,

$$I_i^2 + I_o^2 = K$$

where I_i and I_o are the input and output currents.   Therefore $K - I_o^2$ varies as the mean square of the input current. This relation can be utilized in various computing applications with an accuracy of about 99 per cent.

11-6. Electrical Network Analyzers

Electrical network analyzers for solving algebraic equations and linear and nonlinear differential equations have been developed by manufacturers of electrical equipment, principally the General Electric Company and the Westinghouse Electric Corporation.   They were originally built for studying network problems in connection with the transmission of electrical power and voltage-current characteristics in systems where power loads and power sources may vary rapidly with time.   In recent years, more general use has been found for these machines.   The types of network analyzers include the d-c network analyzer, the a-c network analyzer, and the transient analyzer.   These are described briefly below.

11-6-1. D-C Network Analyzer.   The D-C Network Analyzer has been in service for the solution of industrial problems for the past 25 years.   This analyzer is a system of coupled resistance circuits which can be flexibly arranged to simulate the steady-state conditions existing in an electrical or mechanical problem. The voltage between various points of the resistance lattice and the current through any given group of resistance elements can be determined by direct measurement.   Steady-state solutions for compressible fluid flow and heat flow in addition to solutions of Laplace's equations in two dimensions have been obtained by the D-C Network Analyzer.[36]   In most problems of this type, the voltage across a resistance element simulates the

electrical potential or temperature and the current flowing in the resistance element simulates the fluid flow or heat flow in the medium which is represented by this network. It is obvious that the precision of the results depends upon the precision with which electrical measurements can be made. Thus, a precision of 1 or 2 per cent is considered acceptable.

11-6-2. A-C Network Analyzer. The design and development of the A-C Network Analyzer followed chronologically that of the D-C Network Analyzer. The first large-scale A-C Network Analyzer was constructed in 1929. This analyzer is similar to the d-c type with each resistance element being replaced by a combination of resistance-capacity and -inductance elements. It is more flexible than the D-C Network Analyzer since it can represent phase relationships between currents and voltages in various parts of the network and is therefore capable of handling complex quantities directly. The propagation of electro-magnetic waves has been studied with this analyzer, and during the Second World War it was used to investigate the char-acteristics of cavity resonators. In the field of mathematical physics the A-C Network Analyzer has been used to obtain solutions of the Schrödinger equation for a simple harmonic oscillator and for the rigid rotator in one dimension. Investiga-tions have been made on the effect of the shape of various potential wells on the eigenvalues and eigenfunctions for both the linear harmonic oscillator and the rigid rotator.

11-6-3. Transient Analyzer. This analyzer differs from the d-c and a-c analyzers described above in that it can be used to study nonrepetitive phenomena. It was designed primarily to study electrical transient phenomena resulting from switching and voltage or current discontinuities. The effect of nonlinear elements in physical systems can also be investigated with this analyzer.

The Transient Analyzer[36] is capable of having its circuit parameters varied easily. Therefore it is a powerful tool for use in the synthesis of an optimum physical system from a given set of components of undetermined value. For example, the char-acteristics of a servomotor and its load might be simulated, and the transient response of the over-all system observed as various servoamplifier circuit parameters are adjusted. The transient response at various points in the system may be observed on a

cathode-ray oscilloscope, and a great deal of the design work can be done by bringing the response curves into the desired form. Thus, a minimum of output data needs to be permanently recorded for further analysis. The application of this analyzer to other nonlinear problems in electricity, mechanics, heat, and sound is expected to produce useful and informative results.

11-7. Special Analog Systems

11-7-1. Solutions of Simultaneous Linear Algebraic Equations.
Analog devices on which linear simultaneous algebraic equations
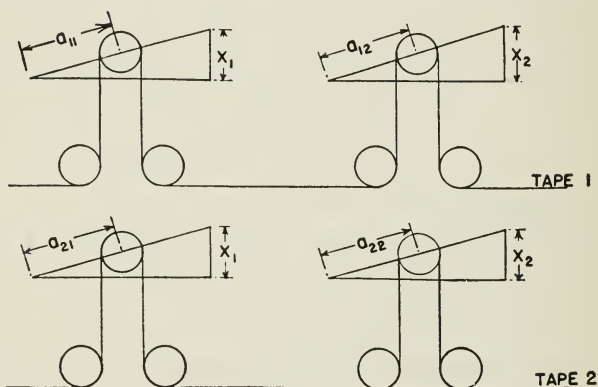
FIG. 11-3. Wilbur's mechanism. The distance between the two ends of each tape is reduced by an amount proportional to $a_{11}x_1 + a_{12}x_2$ or $a_{21}x_1 + a_{22}x_2$.

may be solved have been in existence for at least 50 years, and many types are described in the literature. The more desirable machines have the fewest parts per equation to be solved and therefore have fewest sources of error and are least expensive.

A device has been described by Wilbur[51] in which each equation is represented by a tape. To study the device, consider the two equations:

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

A lever, having a movable roller attached to the lever arm, is provided for each term in the left-hand side of the equation. The levers are arranged in a plane in the same relative position as the terms in the equation. The tapes are constrained by fixed rollers

and are also passed over each movable roller which corresponds to a term in the equation which the tape represents (Fig. 11-3). The arrangement is such that as each lever arm is moved up, the tape is taken up by an amount proportional to the product of the displacement of the roller from the pivot point on the lever arm and the vertical displacement of the lever arm from its normal position. In order to solve the equation given as an example, the displacement of each roller from the pivot point of the lever arm is made proportional to the coefficients a_{11}, a_{12}, a_{21}, and a_{22}. If the x_1 lever arms are displaced from their normal positions through an angle whose sine is proportional to x_1, and the x_2 lever arms are displaced through an angle

|| 11 |12   b_2

21   22   b_2

Φ_1 Φ_2 Φ_0
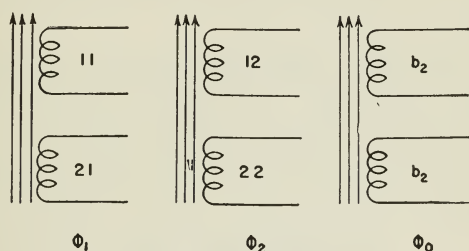
Fig. 11-4. Mallock's circuit.

whose sine is proportional to x_2, the tape corresponding to the first equation is shortened by an amount proportional to $a_{11}x_1 + a_{12}x_2$. Likewise, the other tape is shortened by an amount proportional to $a_{21}x_1 + a_{22}x_2$. If the x_1 and x_2 lever arms are displaced so that the first tape is shortened by an amount b_1 and the second tape is shortened by an amount b_2, the x displacements are proportional to x_1 and x_2, the solution of the equations.

A machine was designed in England in the early 1930's by R. R. M. Mallock[48] which uses alternating voltages and currents and consists of a set of adjustable transformers. This machine is quite elaborate and expensive, partly because of errors in the transformers which must be compensated.

If a scale model of the Mallock machine were applied to the equation of the previous example, three transformers, each having a pair of adjustable coils, would be required (Fig. 11-4).

Let ϕ_1, the total flux in the first transformer, be allowed to represent x_1, and ϕ_2, the total flux in the second transformer,

be allowed to represent x_2. Then if the number of turns in the coil 11 is proportional to the coefficient a_{11}, the potential induced in coil 11 is approximately $a_{11}x_1$. Similarly, the potentials induced in the other coils are approximately $a_{12}x_2$, $a_{21}x_1$, and $a_{22}x_2$, respectively.

The flux ϕ_0 and the turns on the third transformer are adjusted so that the terminal voltages of the two coils are b_1 and b_2, respectively. If ϕ_0 is supplied from an external source and if the three coils in each row are connected in a series loop, the fluxes ϕ_1 and ϕ_2 will adjust themselves so they are in fact proportional to x_1 and x_2.

Precautions must be taken to see that these fluxes do not become excessive in the case of equations with solutions which are large numbers.

Although the Wilbur and Mallock devices are rather complicated, they do converge on the solutions of suitable equations automatically. A more recent machine of this type has been developed at RCA and is described by Goldberg and Brown.[46] The RCA device employs alternating currents, potentiometer-type multipliers, and Kirchhoff adders. Stabilized amplifiers make up for the circuit losses and allow favorable impedance matching. The unknowns are represented by a-c voltages, and these are evaluated by comparing them with the signal supply voltage in a bridge circuit. This technique requires a minimum of precision equipment.

Other equation solvers have also been devised using servomechanisms to drive the multiplying potentiometers, but this seems to be an unnecessary degree of complexity.[39]

A less expensive device of the adjuster type has been described by Berry *et al.*[40] and is now commercially available from the Consolidated Engineering Company. Alternating currents are used for convenience, and potentiometer multipliers are employed. As the name implies, solutions are reached by repeated adjustments of potentiometer shafts in order to carry out the iterative procedure of Gauss and Seidel. Fairly rapid convergence can be realized if the equations can be arranged so that the largest coefficients lie on a diagonal through the center of the determinant of the system of equations. This is a characteristic of the type of equation for which the machine was designed. The commercial model will handle 12 equations and is reputed to

allow solution of simultaneous equations in from a third to a quarter the time required on a desk calculator.

11-7-2. Mechanical Synthesizers and Harmonic Analyzers. It can be shown that the synthesis of a function from its harmonic components and the analysis of an arbitrary function into its harmonic components may both proceed on the basis of equations of the type:[49]

$$.S_p = \sum_{q=1}^{n} d_q \cos \left(pq \frac{\pi}{n} \right)$$

Several machines which can perform this function have been designed and built in this country and are described in the literature.[42,47]

A machine which is a modification of the mechanical synthesizer is called the *Multiharmonigraph*, and its use has been described by Brown and Wheeler.[52] This device will plot a sum of harmonic terms against another sum of harmonic terms. A pair of nonlinear simultaneous equations which have been converted to polar form can also be solved on this machine.

11-8. Applications of Analog Computers

An exhaustive discussion of the applications of analog computers is beyond the scope of this survey, but a glance at the reference list appended[52-71] will give an idea of the large variety of applications which have already been examined. It has been found that new and unpredicted uses are found for existing machines as more workers learn of their characteristics.

References 65 and 67 contain extensive bibliographies on applications of differential analyzers and network analyzers.

REFERENCES

Components and Characteristics of Differential Analyzers

1. Amble, O., On a Principle of Connexion of Bush Integrators, *Journal of Scientific Instruments*, Vol. 23, pp. 284–287 (1946).
2. Beard, R. E., The Construction of a Small-Scale Differential Analyzer and Its Application to the Calculation of Actuarial Functions, *Journal of the Institute of Actuaries*, Vol. 71, pp. 193–227 (1942).
3. Brown, G. S., and A. C. Hall, Dynamic Behavior and Design of Servo-mechanisms, *Transactions of the American Society of Mechanical Engineers*, Vol. 68, pp. 503–524 (1946).

4. Brown, G. S., and D. P. Campbell, Principles of Servomechanisms (John Wiley & Sons, Inc., New York, 1948).

5. Bush, V., F. D. Gage, and H. R. Stewart, A Continuous Integraph, *Journal of the Franklin Institute*, Vol. 203, pp. 63–84 (1927).

6. Bush, V., and H. L. Hazen, Integraph Solution of Differential Equations, *Journal of the Franklin Institute*, Vol. 204, pp. 575–615 (1927).

7. Bush, V., Operational Circuit Analysis (John Wiley & Sons, Inc., New York, 1929).

8. Bush, V., The Differential Analyzer. A New Machine for Solving Differential Equations, *Journal of the Franklin Institute*, Vol. 212, pp. 447–488 (1931).

9. Bush, V., and S. H. Caldwell, A New Type of Differential Analyzer, *Journal of the Franklin Institute*, Vol. 240, No. 4, pp. 255–326 (October, 1945).

10. Gardner, M. F., and J. L. Barnes, Transients in Linear Systems (John Wiley & Sons, Inc., New York, 1942).

11. Greenwood, I. A., Jr., J. V. Holdam, Jr., and D. MacRae, Jr., *Electronic Instruments*, *Radiation Laboratory Series*, Vol. 21 (McGraw-Hill Book Company, Inc., New York, 1948).

12. Harris, H., *The Analysis and Design of Servomechanisms*, Report 454, Office of Scientific Research and Development, Washington, D.C. (January, 1942).

13. Hartree, D. R., The Differential Analyzer, *Nature*, Vol. 135, pp. 940–943 (June 8, 1935).

14. Hartree, D. R., and A. Porter, The Construction of a Model Differential Analyzer, *Memoirs and Proceedings of the Manchester Literary and Philosophical Society*, Vol. 79, pp. 51–72 (July, 1935).

15. Hartree, D. R., The Mechanical Integration of Differential Equations, *Mathematical Gazette*, Vol. 22, pp. 342–363 (1938).

16. Hartree, D. R., Recent Developments in Calculating Machines, *Journal of Scientific Instruments*, Vol. 24, pp. 172–176 (July, 1947).

17. Hazen, H. L., Theory of Servomechanisms, *Journal of the Franklin Institute*, Vol. 218, No. 3, pp. 279–330 (September, 1934).

18. Hazen, H. L., J. J. Jaeger, and G. S. Brown, An Automatic Curve Follower, *Review of Scientific Instruments*, Vol. 7, pp. 353–357 (September, 1936).

19. Jackson, J. M., *Solution of Differential Equations by an Electronic Differential Analyzer*, Admiralty Computing Service Report (Department of Scientific Research and Development, London; reprinted by the Office of Research and Inventions, Navy Department, Washington, D.C., July, 1946).

20. James, H. M., N. B. Nichols, and R. S. Phillips, *Theory of Servomechanisms*, *Radiation Laboratory Series*, Vol. 25 (McGraw-Hill Book Company, Inc., New York, 1947).

21. Koehler, J. S., An Electronic Differential Analyzer, *Journal of Applied Physics*, Vol. 19, pp. 148–155 (1948).

22. Kuehne, H. P., and H. A. Peterson, A New Differential Analyzer, *Transactions of the American Institute of Electrical Engineers*, Vol. 63, pp. 221–228 (May, 1944).

23. Lennard-Jones, J. E., M. V. Wilkes, and J. B. Bratt, The Design of a Small Differential Analyzer, *Proceedings of the Cambridge Philosophical Society*, Vol. 35 (III), pp. 485–493 (July, 1939).

24. Linville, T. M., and J. S. Woodward, Selsyn Instruments for Position Systems, *Electrical Engineering*, Vol. 53, p. 953 (1934).

25. Massey, H. S. W., J. Wylie, and R. A. Buckingham, A Small Scale Differential Analyzer: Its Construction and Operation, *Proceedings of the Royal Irish Academy*, Vol. 45, pp. 1–21 (1938).

26. Murray, Francis J., *The Theory of Mathematical Machines*, 2d rev. ed. (King's Crown Press, New York, 1948).

27. Princeton University, *Bibliography on Feedback Amplifiers, Technical Report* 5 under Contract N6onr-105 Task Order VI between the Office of Naval Research, Navy Department, and Department of Electrical Engineering, Princeton University, Princeton, N.J. (March, 1948).

28. Princeton University, *Circuits for a Linear Negative Resistance and for a Constant Voltage-Ratio Power Amplifier, Technical Report 2* under Contract N6onr-105 Task Order VI between the Office of Naval Research, Navy Department, and Department of Electrical Engineering, Princeton University, Princeton, N.J. (December, 1947).

29. Sack, H. S., R. T. Beyer, G. H. Miller, and J. W. Trischka, Special Magnetic Amplifiers and Their Use in Computing Circuits, *Proceedings of the Institute of Radio Engineers*, Vol. 35, pp. 1375–1382 (November, 1947).

30. Shannon, C. E., Mathematical Theory of the Differential Analyzer, *Journal of Mathematics and Physics* (Massachusetts Institute of Technology), Vol. 20, pp. 337–354 (1941).

31. Stibitz, G. R., A New Class of Computing Aids, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 23, pp. 217–221 (July, 1948); *Function Unit Theory of Computing Devices, Report 1309* (May, 1948). Available on request from Taller and Cooper, Inc., Brooklyn, N.Y.

32. University of California at Los Angeles, Staff of the Department of Engineering, *The Differential Analyzer of the University of California* (1947). Pamphlet published by the University of California, Los Angeles. 25 pp.

33. Zagor, H., Description of a Reeves Electronic Analog Computer (REAC), *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 24, pp. 326–327 (October, 1948). Notes on paper presented by Mr. Zagor before the American Physical Society, Apr. 30, 1948.

Network Analyzers

34. Hazen, H. L., H. R. Schurig, and M. F. Gardner, The MIT Network Analyzer, *Transactions of the American Institute of Electrical Engineers*, Vol. 49, p. 1102 (July, 1930).

35. Kuehne, H. P., and R. G. Lorraine, A New A-C Network Analyzer, *Transactions of the American Institute of Electrical Engineers*, Vol. 57, pp. 67–73 (February, 1938).

36. Peterson, H. A., and C. Concordia, Analyzers for Use in Engineering

and Scientific Problems, *General Electric Review*, Vol. 49, pp. 29–37 (September, 1945).

37. Princeton University, *A Network for Solving Certain Ordinary Non-linear Differential Equations by Analogy, Technical Report 1* under Contract N6ori-105 Task Order VI between the Office of Naval Research, Navy Department, and Department of Electrical Engineering, Princeton University, Princeton, N.J. (October, 1947).

38. Spangenburg, K., and G. Walters, *An Electrical Network for the Study of Electromagnetic Fields, Technical Report 1*, ONR Contract N6-ori-106 Task III (May 15, 1947).

Special Analog Systems

39. Adcock, W. A., An Automatic Simultaneous Equation Computer and Its Use in Solving Secular Equations, *Review of Scientific Instruments*, Vol. 19, pp. 181–187 (March, 1948).

40. Berry, C. E., D. E. Wilcox, S. M. Rock, and H. W. Washburn, A Computer for Solving Linear Simultaneous Equations, *Journal of Applied Physics*, Vol. 17, No. 4, pp. 262–272 (May, 1946).

41. Booth, A. D., Two Calculating Machines for X-ray Crystal Structure Analysis, *Journal of Applied Physics*, Vol. 18, pp. 664–666 (July, 1947).

42. Brown, S. L., A Mechanical Harmonic Synthesizer-Analyzer, *Journal of the Franklin Institute*, Vol. 228, pp. 675–694 (1939).

43. Dietzold, R. L., The Isograph, A Mechanical Root Finder, *Bell Laboratories Record*, Vol. 16, pp. 130–134 (1937).

44. Frame, J. S., Machines for Solving Algebraic Equations, *Mathematical Tables and Other Aids to Computation*, Vol. 1, No. 9, pp. 337–351 (January, 1945).

45. Frost, A. A., and M. Tamres, *A Potentiometric Secular Equation Computer, Journal of Chemical Physics*, Vol. 15, pp. 383–390 (June, 1947).

46. Goldberg, E. A., and G. W. Brown, An Electronic Simultaneous Equation Solver, *Journal of Applied Physics*, Vol. 19, pp. 338–345 (April, 1948).

47. Hagg, G., and T. Laurent, A Machine for the Summation of Fourier Series, *Journal of Scientific Instruments*, Vol. 23, pp. 155–158 (1946).

48. Mallock, R. R. M., An Electrical Calculating Machine, *Proceedings of the Royal Society*, Vol. 140, p. 457 (1933).

49. Murray, F. J., The Theory of Mathematical Machines (King's Crown Press, New York 1948).

50. Ragazzini, J. R., Robert H. Randall, and F. A. Russel, Analysis of Problems in Dynamics by Electronic Circuits, *Proceedings of the Institute of Radio Engineers*, Vol. 35, pp. 444–452 (May, 1947).

51. Wilbur, J. B., The Mechanical Solution of Simultaneous Equations, *Journal of the Franklin Institute*, Vol. 222, pp. 715–724 (1936).

Analog Computer Applications

52. Brown, S. L., and L. L. Wheeler, Use of the Mechanical Multiharmonigraph for Graphing Types of Functions and for Solution of Pairs of

Nonlinear Simultaneous Equations, *Review of Scientific Instruments*, Vol. 13, pp. 493–495 (1942).

53. Brown, S. L., and L. L. Wheeler, The Use of a Mechanical Synthesizer to Solve Trigonometric and Certain Types of Transcendental Equations, and for the Double Summations Involved in Patterson Contours, *Journal of Applied Physics*, Vol. 14, pp. 30–36 (1943).

54. Bush, V., and S. H. Caldwell, Thomas-Fermi Equation Solution by the Differential Analyzer, *Physical Review*, Vol. 38, pp. 1898–1902 (1931).

55. Carter, G. K., Numerical and Network-analyzer Tests of an Equivalent Circuit for the Elastic Field, *Journal of Applied Mechanics*, p. A-162 (September, 1944).

56. Carter, G. K., and G. Kron, A-C Network Analyzer Study of the Schrö-dinger Equation, *Physical Review*, Vol. 67, pp. 44–49 (January, 1945).

57. Clarke, E., and S. B. Crary, Stability Limitations of Long-distance A-C Power Transmission Systems, *Transactions of the American Institute of Electrical Engineers*, Vol. 60, pp. 1051–1059 (1941).

58. Concordia, C., Network and Differential Analyzer Solutions of Torsional Oscillation Problems Involving Nonlinear Springs, *Journal of Applied Mechanics*, Vol. 12, pp. A43–A47 (March, 1945).

59. dePackh, D. C., A Resistor Network for the Approximate Solution of the Laplace Equation, *Review of Scientific Instruments*, Vol. 18, pp. 798–799 (1947).

60. Hartree, D. R., and J. R. Womersley, A Method for the Numerical or Mechanical Solution of Certain Types of Partial Differential Equations, *Proceedings of the Royal Society* (London), Vol. 161A, pp. 353–366 (1937).

61. Hartree, D. R., and A. Porter, The Application of the Differential Analyzer to Transients on a Distortionless Transmission Line, *Journal of the Institution of Electrical Engineers* (London), Vol. 83, pp. 648–656 (1936).

62. Hartree, D. R., The Application of the Differential Analyzer to the Evaluation of Solutions of Partial Differential Equations, *Proceedings of the First Canadian Mathematical Congress, Montreal, 1945* (University of Toronto Press, Toronto, 1946), pp. 327–337.

63. Kron, G., and G. K. Carter, Numerical and Network Analyzer Tests of an Equivalent Circuit for Compressible Fluid Flow, *Journal of the Aeronautical Sciences*, Vol. 12, p. 232 (April, 1945).

64. Levin, J. H., On the Approximate Solution of a Partial Differential Equation on the Differential Analyzer, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 23, pp. 208–209 (July, 1948).

65. Maginniss, F. J., Differential Analyzer Applications, *General Electric Review*, Vol. 48, pp. 54–59 (May, 1945).

66. Massachusetts Institute of Technology, Department of Electrical Engineering, Center of Analysis, *Tables of Supersonic Flow around Yawing Cones, Report 3* under Contract NOrd 9169.

67. Peterson, H. A., and C. Concordia, Analyzers for Use in Engineering and Scientific Problems, *General Electric Review*, Vol. 48, No. 9, pp. 29–37 (September, 1945).

68. Reed Research, Inc., *A Constant Power Circuit and Its Application to Instrumentation and Computation;* Unpublished, Washington, D. C.; Oct. 12, 1949.

69. Schumann, T. E. W., The Principles of a Mechanical Method for Calculating Regression Equations and Multiple Correlation Coefficients and for the Solution of Simultaneous Linear Equations, *Philosophical Magazine*, Vol. 29, Seventh Series, pp. 258–273 (March, 1940).

70. Walters, G., and R. L. Tanner, *Universal Tuning Curves for Reentrant Cavity Resonators, Technical Report 2,* ONR Contract N6-ori-106 Task III (Aug. 15, 1947).

71. Whinnery, J. R., C. Concordia, W. Ridgeway, and G. Kron, Network Analyzer Studies of Electromagnetic Cavity Resonators, *Proceedings of the Institute of Radio Engineers*, Vol. 32, pp. 360–367 (June, 1944).

CHAPTER 12

THE FORM OF A DIGITAL COMPUTER

12-1. Introduction

In the chapters which have preceded this one, the means for performing automatic calculations have been described in general terms. In particular, a detailed description of elementary components has been given, and it has been shown how these components can be combined to perform arithmetic operations rapidly. Also a description of some necessary steps in the integration of components into a machine has been presented. Since these machines depend essentially on the system of arithmetic which is used, and since some systems are more efficiently utilized in computing equipment than others, a discussion of arithmetic systems was included. Since the form of a calculation depends on the details of the means chosen to carry out various simple arithmetical steps (such as extraction of a square root), generalities concerning the arithmetic components which may be used as steps in a calculation have been discussed. In order that the size of a calculation at hand may be estimated in terms of machinery involved, desk calculators, punched-card computing systems, existing and planned digital computers, and existing analog computers have been described; these descriptions were incomplete, for they were intended only to give some idea of the relative extents of the calculations which are suitably carried out on the various machines.

In the chapters which follow, the circuits and the mechanisms of the physical computing and data-handling components which are now available and which are expected shortly to be available are described in sufficient detail to permit rough estimates of applicability and the expense of applying them in machines. Chapter 12 is intended as an aid in choosing components.

12-2. The Field of Application

When a large-scale digital computing machine is to be built, a field of application for this machine presumably exists and is

251

fairly well defined. As a matter of practicability and economy, the machine may be designed to carry out the calculations involved in problems in this field only.

The designation of a machine as a *general-purpose machine* simply indicates that the field from which problems are likely to be extracted is reasonably broad. A *special-purpose machine* is a machine designed to solve problems from a narrower field. However, one can expect to discover that any given machine, even a machine designed for a rather narrow class of problems, can be used effectively to solve problems outside the class for which it was designed. In short, it is practically impossible to design a machine so specialized that it will have value only with respect to the field of application originally intended. Nor is there any computer which is superior to any other computer with respect to every problem. In short, there is no completely general-purpose machine.

The problem of designing a machine is then defined only approximately by the field of calculations to which it is to be applied. In general, the narrower the field, the simpler or the faster the machine. However, there may be some question whether a net gain is achieved by building a special-purpose machine. A conventional general-purpose computer may be capable of sufficient speed, and it may be less expensive to acquire such a machine than to develop a more highly specialized though possibly simpler computer.

The final test of the applicability of a machine to a field of problems lies in an explicit demonstration of a program under which the machine will carry out each of the calculations required. From this program and from the machine's speed and reliability, an estimate of the speed and reliability of the calculation can be made. There is no other way to make this objective estimate.

12-3. The Estimate of Size and Speed

Any calculation must be factored into a set of elementary arithmetic operations before it can be carried out. Once this factorization has been accomplished, it is possible to estimate the size of the calculation. In Chap. 5 this factorization of a simple problem was illustrated, and an explicit program for the particular problem was set forth. If a time limit is placed on the length of the calculation, either arbitrarily or naturally through the

application, then the rate at which the various arithmetic opera-
tions must be carried out is determined, and from this rate the
size of the machine required can be estimated. In particular,
under an assumed rate of multiplication, for example, the number
of multiplying units required can be determined. If this number
exceeds one, the machine is likely to be so complex that its
application will be limited, and its operation as a general-purpose
computer will be difficult. It will then be necessary to examine
the factorization of the calculation program critically to deter-
mine how more than one unit can be utilized. Thus, if each
multiplication depends for one factor upon all preceding multi-
plications, there is no point in increasing the number of multiply-
ing units, for only one multiplication can be carried out at one
time in any event. On the other hand, if the problem consists
of nothing but matrix operations on vectors, one multiplying
unit might be conveniently utilized to compute each of the com-
ponents of a vector.

In general, the extent of the complexities associated with
multiunit computing machines is great. The chief motivation
for increasing the speed of computing units has been the desire
to reduce the number of units in a machine without reducing its
speed.

Specifically, the choice of the components built into a machine
must be based upon the programming of representative problems
from the field to be attacked. This programming will indicate
the amount of storage required, the nature of the input required,
the nature of the output, the rate of calculation in terms of
elementary operations of the various kinds per second, and the
number and nature of the different operations the machine must
carry out. From these numbers an estimate of the feasibility
of the machine can be obtained by a search through the lists of
components; if no components exist which provide the facilities
required, either a more efficient reprogramming is required or,
failing this, a more modest goal must be set for immediate
attainment.

12-3-1. An Example. Suppose it is desired to calculate, by
means of a digital computer, the position of a target tracked by
Raydist[2] equipment. Consideration of this problem will provide
us with an example of the factoring of a calculation into a set of
elementary arithmetic operations for machine solution.

Raydist is a radio navigation and tracking system. It is based on the measurement of the relative phase relationship between continuous-wave signals received at different locations.

When the Raydist system is applied to the problem of tracking a target through three-dimensional space, four fixed stations are required. These may be represented in cartesian coordinates as

$$P_1(x_1,y_1,z_1); \qquad P_2(x_2,y_2,z_2); \qquad P_3(x_3,y_3,z_3); \qquad P_4(x_4,y_4,z_4)$$

The position of the target is $P(x,y,z)$. Comparison circuits in the Raydist equipment give values v_1, v_2, and v_3, which are the differences between the distance from the target to a fixed reference point P_4 and the distance from the target to each of P_1, P_2, and P_3. v_1, v_2, and v_3 are related to the coordinates of the four fixed points and the target by the three quadratic polynomial equations in three unknowns which follow:

$$
\begin{aligned}
v_1 &= \sqrt{(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2} \\
&\qquad - \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2} \\
v_2 &= \sqrt{(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2} \\
&\qquad - \sqrt{(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2} \\
v_3 &= \sqrt{(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2} \\
&\qquad - \sqrt{(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2}
\end{aligned}
\qquad \text{(12-1)}
$$

These equations are equations of hyperboloids of revolution, each with one focus at the point P_4. The other foci are at points P_1, P_2, and P_3, respectively. The problem is to solve for the coordinates (x,y,z) of P, knowing all the other numbers.

A Formulation of the Problem. One formulation of this problem is based on the assumption that the target moves with restricted accelerations and that a fairly continuous record of its position is desired. Thus, a system of sharpening an estimated path of the target will be adequate. The position after any step can be estimated on the basis of its position after the preceding step and its average velocity during the preceding step. Thus, if the calculations of position P occur after equal increments of time, so that the kth calculation occurs at the time $k\,\Delta t$ with Δt constant, and if $p^k(x^k,y^k,z^k)$ denotes the position at the kth time interval, then the coordinates of P^{k+1} can be estimated as

$$
\begin{aligned}
x^{k+1} &= 2x^k - x^{k-1} \\
y^{k+1} &= 2y^k - y^{k-1} \\
z^{k+1} &= 2z^k - z^{k-1}
\end{aligned}
\qquad \text{(12-2)}
$$

The accuracy of this estimate is guaranteed by the size of Δt and the extent to which the acceleration is restricted. It is to be assumed that the initial position P^0 is known accurately.

With this formulation an extension of Newton's method may be applied. For this example Eqs. (12-1) are reduced to polynomial form:

$$A_i x^2 + B_i y^2 + C_i z^2 - 2D_i yz - 2E_i zx - 2F_i xy - 2G_i x \\ - 2H_i y - 2I_i z = V_i, \qquad i = 1, 2, 3 \quad (12\text{-}3)$$

Here all the coefficients are functions of the known numbers in Eqs. (12-1) (the coordinates of P_1, P_2, P_3, P_4 and the numbers v_1, v_2, v_3). The algebra involved in the transition from Eqs. (12-1) to Eqs. (12-3) consists in rearranging Eqs. (12-1) so that a radical is isolated on one side of the equation, squaring, again rearranging so that the only remaining radical is isolated, and again squaring. The process can be carried out using the literal coefficients shown in Eqs. (12-1), and explicit formulas for the coefficients in Eqs. (12-3) in terms of these letters can be exhibited.

Once the coefficients in Eqs. (12-3) are known, the approximate values of x, y, and z to be sharpened can be substituted in the left members. For this substitution the amounts that V_1, V_2, and V_3 differ from the actual values generated by the approximate x, y, and z may be denoted by ΔV_1, ΔV_2, ΔV_3; the problem is to modify x, y, z by changing to values $x + \Delta x$, $y + \Delta y$, $z + \Delta z$ so that changes of the left-hand members of the Eqs. (12-3) induced by the change in x, y, z are just ΔV_1, ΔV_2, ΔV_3. This can be achieved approximately by linearizing the problem through the use of differentials and solving the three equations:

$$\frac{\partial V_i}{\partial x}\, dx + \frac{\partial V_i}{\partial y}\, dy + \frac{\partial V_i}{\partial z}\, dz = dV_i, \qquad i = 1, 2, 3 \quad (12\text{-}4)$$

simultaneously for dx, dy, and dz in terms of dV_1, dV_2, dV_3. If the values ΔV_1, ΔV_2, ΔV_3 are substituted in place of dV_1, dV_2, dV_3, numbers corresponding to dx, dy, dz may be found. These differentials are approximately equal to the actual Δx, Δy, Δz, which must be added to the assumed values of x, y, z in order to account for accelerations during the time interval Δt. Thus, the approximate solution will have been sharpened by the method of Newton.

In summary, this approach to the solution involves the following steps:

1. Read α_1, α_2, and α_3 from Raydist dials. (These are the phase differences between signals as received at different stations.) Multiply by a factor k to convert dial readings from wavelength of radio-frequency carrier to distances v_1, v_2, v_3 in some convenient scale.

2. From v_i computed in step 1 and constants L_i, M_i, . . . , T_i involving the coordinates of P_1, P_2, P_3, P_4, compute A_i, B_i, C_i, etc., the coefficients of Eqs. (12-3).

3. From the values of x^{k-1}, y^{k-1}, z^{k-1}, x^k, y^k, and z^k computed in the two previous calculations, and with Eqs. (12-2), calculate approximate values of x^{k+1}, y^{k+1}, and z^{k+1}.

4. Use the coefficients calculated in **2** and the approximate distances calculated in **3** to compute the nine partial-derivative coefficients of Eqs. (12-4).

5. Assemble the coefficients found in **4** into a Jacobian matrix and compute the value of the nine minor determinants of the nine matrix terms. (NOTE: The Jacobian matrix of a set of differential equations is the matrix formed by the derivative coefficients.)

6. From the terms computed in **4** and the minors computed in **5**, compute the value of the Jacobian determinant of the Eqs. (12-4).

7. Substitute the coefficients of Eqs. (12-3) computed in **2** and the approximate distances computed in **3** into the left-hand sides of Eqs. (12-3). Subtract the value computed from the values of V_i computed for the right-hand sides of the equations. Label the differences ΔV_i.

8. Use the values of ΔV_i computed in **7** as substitutes for dV_i, and the values of the minors and the determinant computed in **5** and **6** to solve Eqs. (12-4) for dx, dy, and dz. Label the results Δx, Δy, and Δz, respectively.

9. Add Δx, Δy, and Δz to x^{k+1}, y^{k+1}, and z^{k+1}, respectively. These sums are the sharpened values of x^{k+1}, y^{k+1}, and z^{k+1}.

Breakdown of the Problem into Elementary Steps. In order to determine the size and speed of the computer required to handle the problem just outlined in a given length of time, it is necessary first to arrange the steps in a chronological order which will allow the computation to proceed directly from the input data

to the final result.   This has been done in the foregoing section.
It is necessary further to count the operations of each type.   An
illustration of how this might be done is presented in tabular form
in Table 12-1, along with a symbolic representation of the opera-
tions listed in the preceding section.

Table 12-1 shows the number and nature of the various opera-
tions required to estimate coordinates and to sharpen the estimate
once.   Slight deviations can be made from the schedule outlined,
and in a binary computer, multiplication by powers of 2 may be
accomplished simply by shifting the multiplicand.   Since
division is required only once, it might be desirable to perform
division by an iterative process involving only addition and
multiplication rather than provide a divider unit for this par-
ticular problem.   A description of such a method appears in the
Annals of the Computation Laboratory of Harvard University[1]
and is repeated on pages 122–123 *supra*.

The time interval Δt is established on the basis of the maneu-
verability or possible acceleration of the target and the precision
with which it is desired to track the target.   The latter factor
also determines the size of the numbers to be handled in the
computer.   For purposes of illustration, let it be assumed that
the coordinates must be calculated every 0.1 second in order to
render valid the approximations involved in the method outlined.

The requirement for 10 solutions per second demands that
automatic input equipment be employed.   The detailed specifica-
tions for automatic input equipment depend upon the output
of the measuring equipment and the arithmetic system used in
the computer.   The primary physical quantities representing
the output of the Raydist system are a pair of audio-frequency
voltages which differ in phase and actuate an integrating syn-
chroindicator.   Each cycle of phase shift might be counted with
a set of electronic counters such as those described in Sec.
13-5.   Undoubtedly other devices would evolve from a careful
study of the problem.   Whether or not the digital data from the
input equipment are fed directly into the arithmetic units of the
computer or into a storage device depends upon whether on-line
or off-line operation is being sought.   In the former case, the
computer solves for points on the trajectory as the experiment is
performed; in the later case, the computations may be performed
at a later time and at a scaled-down rate.

TABLE 12-1

Step No.	Operation	Add or subtract	Multiply or square	Divide
1	$v_i = k\alpha_i$		3	
2	$4v_i^2 = 4(v_i)^2$		6	
	$A_i = 4v_i^2 + L_i$	3		
	$B_i = 4v_i^2 + M_i$	3		
	$C_i = 4v_i^2 + N_i$	3		
	D_i, E_i, and F_i are constants.			
	$G_i = 4M_iv_i^2 + N_i$	3	3	
	$H_i = 4O_iv_i^2 + P_i$	3	3	
	$I_i = 4Q_iv_i^2 + R_i$	3	3	
	$v_i^4 = (v_i^2)^2$		3	
	$V_i = v_i^4 + S_iv_i^2 + T_i$	6	3	
3	Approximate values:			
	$x_0^{k+1} = 2x^k - x^{k-1}$	1	1	
	$y_0^{k+1} = 2y^k - y^{k-1}$	1	1	
	$z_0^{k+1} = 2z^k - z^{k-1}$	1	1	
4	$\dfrac{\partial V_i}{\partial x} = 2(A_ix - E_iz - F_iy - G_i)$	9	12	
	$\dfrac{\partial V_i}{\partial y} = 2(B_iy - D_iz - F_ix - H_i)$	9	12	
	$\dfrac{\partial V_i}{\partial z} = 2(C_iz - D_iy - E_ix - I_i)$	9	12	
5	Let $\begin{vmatrix} \dfrac{\partial V_1}{\partial x} & \dfrac{\partial V_2}{\partial x} & \dfrac{\partial V_3}{\partial x} \\ \dfrac{\partial V_1}{\partial y} & \dfrac{\partial V_2}{\partial y} & \dfrac{\partial V_3}{\partial y} \\ \dfrac{\partial V_1}{\partial z} & \dfrac{\partial V_2}{\partial z} & \dfrac{\partial V_3}{\partial z} \end{vmatrix} = D_J$, determinant of Jacobian matrix. Find the nine minors, where $M_{1,x} = \begin{vmatrix} \dfrac{\partial V_2}{\partial y} & \dfrac{\partial V_3}{\partial y} \\ \dfrac{\partial V_2}{\partial z} & \dfrac{\partial V_3}{\partial z} \end{vmatrix}$	9	18	
6	$D_J = \dfrac{\partial V_1}{\partial x} M_{1,x} - \dfrac{\partial V_1}{\partial y} M_{1,y} + \dfrac{\partial V_1}{\partial z} M_{1,z}$	2	3	
	Find x^2, y^2, z^2, $2yz$, $2zx$, $2xy$, $2x$, $2y$, $2z$		12	

TABLE 12-1. (*Continued*)

Step No.	Operation	Arithmetic processes		
		Add or sub- tract	Multi- ply or square	Di- vide
7	$\Delta V_i = (A_ix^2 + B_iy^2 + C_iz^2 - 2D_iyz - 2E_izx$ $- 2F_ixy - 2G_ix - 2H_iy - 2I_iz) - V_i$	9	27	
8	$\dfrac{1}{D_J} = 1 \div D_J$			1
	$dx = \dfrac{\Delta V_1(M_{1,x}) - \Delta V_2(M_{1,y}) + \Delta V_3(M_{1,z})}{D_J}$	2	4	
	$dy = \dfrac{-\Delta V_1(M_{2,x}) + \Delta V_2(M_{2,y}) - \Delta V_3(M_{2,z})}{D_J}$	2	4	
	$dz = \dfrac{\Delta V_1(M_{3,x}) - \Delta V_2(M_{3,y}) + \Delta V_3(M_{3,z})}{D_J}$	2	4	
9	Improved: $x^{k+1} \simeq x_0^{k+1} + dx$ $y^{k+1} \simeq y_0^{k+1} + dy$ $z^{k+1} \simeq z_0^{k+1} + dz$	1 1 1		
Totals.................................		83	135	1

In the assumed Raydist problem the chief difficulty is in performing the arithmetic operations fast enough. Approximately 135 multiplications are required for each estimate, so if one estimate is adequate and only one multiplier is employed, multiplications must be performed in less than 700 microseconds. It is shown in Chap. 13 that the fastest binary multipliers now known to exist operate in about $5n$ microseconds, where n is the number of binary digits in the numbers to be multiplied per microsecond of pulse width. If a 1-microsecond pulse width and a 30-binary-digit precision are assumed, each multiplication could be carried out in 150 microseconds, which is safely within the required time.

If the calculations are performed on a high-speed general-purpose machine, they will probably be taken sequentially. An examination of the steps listed in Table 12-1 reveals that the quantities to be stored in addition to the three incoming α_i and the k, L_i, M_i, . . . , T_i constants of the Raydist system are

almost all of the 83 sums and the 135 products. Thus, the number of quantities to be stored is approximately 275, and it will be seen in Chap. 14 that this is a relatively small number for present-day storage systems.

The output requirements in this example depend almost entirely upon uses to which the points in the trajectory are to be put. In some cases the position of the target transmitter might be used directly to control some other process. In other cases, a permanent tabular record of the coordinates of the trajectory might be desired. Still another possibility is that the trajectory should be traced by some plotting device for visual inspection. Equipment with which to perform these functions is described in Chap. 15. It is possible that it may be satisfactory to use output data taken at longer time intervals than the input data, since the time interval at which the input data is sampled was determined by the requirements of the mathematical method employed.

12-4. Conclusions

In the example discussed, the crux of the difficulty in performing on-line computing with a general-purpose machine comes in the number of multiplications required. In other problems the input equipment or the storage system may prove to be the source of technical difficulty. In any case, a program similar to the one outlined for the example given in this chapter must be carried through in order to find the machine requirements for the various possible mathematical approaches. In this way, it may be found that some methods are better suited to the computing machinery available than other methods which might have seemed more inviting.

After a preliminary estimate of the relative speeds and capabilities required in various parts of the machine, a complete coding of the problem must be made before an accurate determination of speed can be arrived at.

The question of whether on-line or off-line operation is more suitable can be determined only by applying the speed requirements to those physical components which are available. Likewise the question of whether or not a general-purpose computer or a faster special-purpose computer is required depends ultimately upon whether or not the speeds of the arithmetic units available

are great enough to effect a complete solution in the time allotted. In cases where on-line operation is required because of the nature of the problem, and where computing speeds are not sufficiently high, a certain amount of parallel operation is required, and the designer and the mathematician must devise means of accomplishing various steps simultaneously.

REFERENCES

1. Harvard University, Staff of the Computation Laboratory, *A Manual of Operation for the Automatic Sequence Controlled Calculator, Annals of the Computation Laboratory of Harvard University*, Vol. I (Harvard University Press, Cambridge, Mass., 1946).
2. Hastings, Charles E., Raydist—a Radio Navigation and Tracking System, *Tele-Tech.*, Vol. 6, No. 6, pp. 30–34 (June, 1947).

Part III

PHYSICAL COMPONENTS AND METHODS

The following chapters include a detailed discussion of physical components which have been introduced in Chaps. 3 and 4 of Part I and which have been described in Part II with regard to their integration into computing systems. In Chap. 13 is described the organization of simple arithmetic units using elementary potential digital computing components and switches and gates. Chapter 14 deals with the transfer of data in time and in space. Storage is defined as a transfer of data in time; switches and cables throughout a computer are described as components effecting a transfer in space. Data-conversion equipment, including input, output, and radix-changing devices, is discussed in Chap. 15. Special physical techniques of possible use in automatic calculation are summarized in Chap. 16. Finally, Chap. 17 outlines some of the factors affecting the choice of components.

CHAPTER 13

ARITHMETIC ELEMENTS

13-1. Introduction

Since the first Oriental abacus was constructed, many ingenious units have been devised to perform arithmetic operations. The arithmetic elements to be discussed here, however, are restricted to those of an electronic nature, in which electron tubes and crystal rectifiers are employed to carry out the desired operations. These electronic elements have their counterparts in electromechanical relay systems, mechanical counter wheels, and other mechanical devices. These will not be discussed specifically here. However, specific examples will be given in order to indicate the manner in which high-speed arithmetic elements are built up from individual switches, gates, and counters. These examples will serve to indicate how other more complicated systems may be designed from basic components of any type and to give an idea of the total amount of equipment required to meet various specifications. The differences between serial and parallel arithmetic operations are reviewed so that elements performing these operations may be comprehended. The characteristics of fixed-cycle and variable-cycle operation are also mentioned so that the differences in their requirements may be understood. Examples of coincidence adders and ring-type adders are described. Finally, a discussion has been included of the incorporation of adders into complete arithmetic units. Emphasis has been given to arithmetic units of the binary type, but these may be expanded to compute in radix higher than 2 by an extension of the procedure outlined in this chapter. In the material which follows, it may be assumed that numbers are binary unless otherwise designated.

13-2. Series versus Parallel Operation of Arithmetic Units

An initial choice which must be made in deciding the method of computation in an automatic digital computer is whether

serial or parallel arithmetic elements are to be used.   To indicate
the fundamental difference between serial and parallel arith-
metic units, consider two such elementary adders.   In a serial-
type adder, as shown in Fig. 13-1a, the adder receives simul-
taneously only a single pair of equal-order digits of the addend
and augend.   The lowest order, or right-hand, digits of each
number are added first, and a sum and carry are formed.   This
carry and the second digits are then added, etc., in succession.
Thus, if 0,638 is to be added to 0,092, the arithmetic unit first
receives a 2 and an 8 to add.   The modulo 10 sum 0 is registered
in the adder, and the carry digit 1 is stored to add to the pair
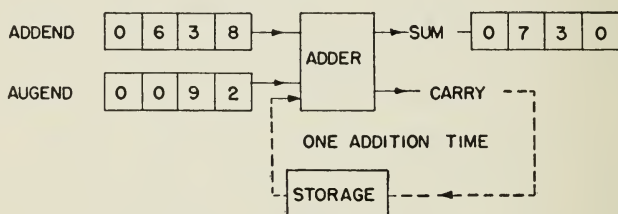
Fig. 13-1a.   Block diagram of serial addition system.

of second-order digits 9 and 3.   These three digits are summed
during the time called the *second addition time* and result in a
sum of 3 and another carry of 1.   This process is continued until
all the sums have been formed, in this case, two more times.
Serial arithmetic might be considered desirable in conjunction
with a cyclic storage medium of large capacity and low access
time.

In a parallel adder all pairs of equal-order digits in the addend
and augend are entered into the adder simultaneously.   Thus,
the number of adder units is equal to the number of digits in the
largest of the numbers to be added.   Each adder unit is identified
with a digit-order number, and it is cascaded with the others in
ascending digit order as illustrated in the block diagram of Fig.
13-1b.   To form the sum of 0,638 and 0,092, of the previous
example, the first-order digits 8 and 2 are paired together and
entered into the right-hand adder unit.   At the same time the
second-order digits, 3 and 9, enter the second adder unit to the
left, etc., for all four pairs of digits.   Additions of all pairs
of digits take place simultaneously.   Carries are formed in
adders I and II.   These carries are passed on to the adders of

next higher order, II and III, respectively, where they are added to the initial sums; this may result in further carries, so that in the extreme case a carry could travel from the lowest order right-hand adder to the highest order left-hand adder.

The major difference between these two methods is that when two numbers having n significant digits are added, a parallel adder requires n times as many individual adder units as a serial adder. The serial adder, however, requires n addition time units in which to find the sum, whereas the parallel adder requires but one addition time unit plus sufficient time in which to allow n
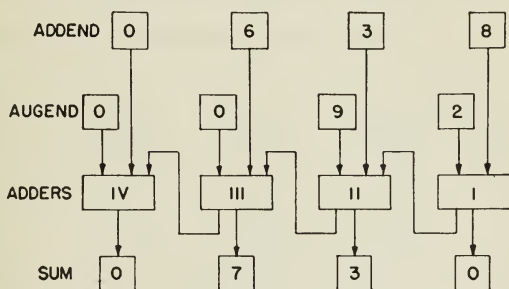
Fig. 13-1b.   Block diagram of parallel addition system.

successive carries. The differences between other arithmetic units which operate serially and in parallel are similar.

It seems quite feasible to build computers in which both serial and parallel representation of numbers is employed.[5] For example, in a machine being built by the Raytheon Company, binary digits are stored serially in acoustic delay lines and are added in a parallel adder. Conversion of numbers from one form to another is fairly simple in a machine in which all pulses are synchronized with a master clock-pulse generator, as illustrated in Figs. 13-2a and b. In each case, one-pulse time-delay circuits delay each digit by a number of pulse times proportional to the number of the flip-flop in which it is received or has originated. Each process is started by a single pulse applied to all gates simultaneously.

The shift register described in Sec. 13-6-3 may also be used as a serial-to-parallel converter and operates in a manner similar to the converters just described.

A distinction between series and parallel operation at the

level of the mathematical process of addition, subtraction, and multiplication also exists. Although it is considered rather complicated to perform several of these processes simultaneously,
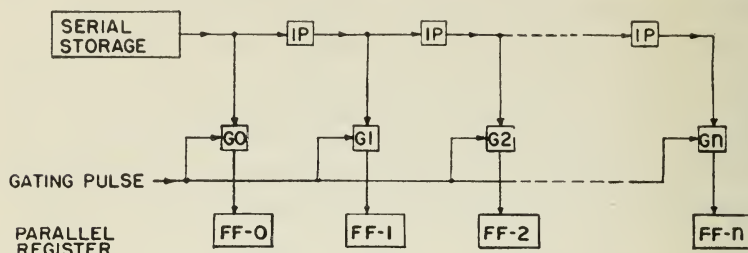
FIG. 13-2a.   Serial-to-parallel converter.
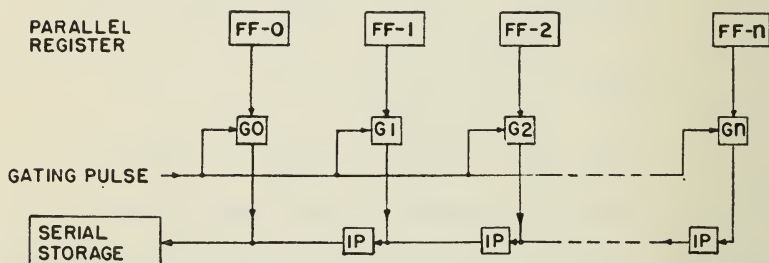
FIG. 13-2b.   Parallel-to-serial converter.

or in parallel, in a high-speed electronic digital computer, it is not uncommon to use punched-card calculating machines in this manner.

13-3. Operation Sequence Control

It is possible to design a computer which requires instructions only to perform certain arithmetical operations in a given sequence, the internal control of the machine automatically initiating one operation upon completion of the previous operation. In the interest of simplicity of equipment, however, most present-day computers require that in the preparation of the problem a particular time be assigned at which each operation shall be initiated. Since an operation such as addition may require more time when one set of numbers is added than when another set is added[3] (depending upon the number of carries generated) and since it is impractical in coding the problem to

predict the time requirement in advance, the maximum amount of time required for the greatest possible number of carries must be allowed. Thus, on the average, the computer which proceeds automatically can perform a given operation faster than one which must wait the maximum time for each operation. These two types of operation are sometimes called *fixed cycle* and *variable cycle*.

As arithmetic units are improved, the time requirements for the two types of operation are becoming more nearly equal and the speed advantage of automatic control is lessened.

13-4. Elementary Coincidence-type Adders

Chapter 4 described an idealized adder element. This adder contained two counters. At the start of an addition, the addend was set up in one counter, called the *addend register;* the augend was set up in the other, the *accumulator.* The accumulator was then made to count. Starting from its initial position, *i.e.,* starting not at zero but at the number representing the augend, it was made to count the addend. When the count was completed, the accumulator contained the sum of augend and addend. Such an adder may be called a *counter-type adder.*

A *coincidence-type adder* differs in that pulses representing two equal-order digits of addend and augend and the previous carry are received by the adder simultaneously, and the correct sum and carry pulses are gated out immediately. Its name is derived from the fact that the output pulses are dependent upon the coincidences of the input pulses. In this section a number of coincidence adders are described, simplified electronic circuits are given, and the arithmetical operations which they can perform are outlined. These examples are patterned after the serial binary-arithmetic equipment developed during the EDVAC study at the University of Pennsylvania[8] and include the following types: (1) binary nonalgebraic adders; (2) a binary algebraic adder; (3) amplitude adders; (4) a coded-decimal adder; (5) matrix and cathode-ray-tube adders.

13-4-1. Two-input Binary Adders. In order to perform an addition of two numbers in any radix system, an addend digit, an augend digit, and a previous-carry digit must be summed. These three digits may be added either as pairs in two separate steps or simultaneously in one step. Two types of adders have

been devised, one for each of these methods of adding three digits.

Since the addition of only two of the three digits referred to constitutes something less than the entire operation, the term *half adder* has been used to signify a device in which two digits are added. A combination of two half adders, which can add three digits in two separate steps, is called a *two-input adder*. A system in which three digits are received simultaneously at the input is called a *three-input adder*. The two-input adder is discussed in this section.

A half adder delivers a proper sum and carry digit for each pair of digits which it receives. Since these are not always the final sum and carry digits, they will be designated as S' and C', respectively. Table 13-1, a binary addition table, shows the sum and carry for each possible combination of two binary digits, A and B, and will serve to illustrate the binary addition process.

TABLE 13-1

Addend A	Augend B	Sum S'	Carry C'
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Suppose we want to design a device capable of performing the binary addition process for each of these combinations of values of A and B. This device, called a half adder, will consist of *and*, *or*, and *not* circuits, as described in Chap. 4, arranged so as to carry out the operations exemplified in Table 13-1. To determine what this arrangement of *and*, *or*, and *not* components should be, let us state a set of rules to be used as functional specifications for the adder. These rules will be derived from an inspection of Table 13-1 and will be stated in the terminology of formal logic.

It may be seen that 1 and 0 are the two possible values of A, B, S', and C', corresponding to the logical concepts *true* and *false*.

The logical concepts *and*, *or*, and *not*, as embodied in com-

ponents of digital computation (*e.g.*, gating circuits), are defined by the following statements in the language of formal logic[4,7] in which the letters A and B themselves denote statements which can be either true or false:

1. *A and B* denotes the statement which is true if and only if both A and B are true. (Applying this definition to a gate circuit, *A and B* represents the condition which obtains when the value 1 appears both at A and at B.)
2. *A or B* denotes the statement which is true if and only if at least one of the statements A, B is true. (That is, *A or B* denotes the condition which obtains in an anti-Rossi circuit when there is a 1 at A, at B, or at both A and B.)
3. *Not A* denotes the statement contradictory to A. If *not A* is true, A is false. [That is, *not A* denotes the condition which obtains when there is a 0 at A. *Not (A and B)* denotes the condition which obtains when the value 1 does not appear at both A and B.]
4. $A = B$ denotes a statement which is true if and only if either A and B are true or A and B are false. (That is, $A = B$ means that if there is a 1 at B, there is also a 1 at A; if there is a 0 at B, there is also a 0 at A.)

On the basis of these definitions, we can derive and state the following rules of binary addition from an inspection of Table 13-1:

$$S' = (A \text{ or } B) \text{ and not } (A \text{ and } B)$$
$$C' = A \text{ and } B$$

Figure 13-3 is a block diagram showing an arrangement of logical *and*, *or*, and *not* components which enforces these rules of binary addition.
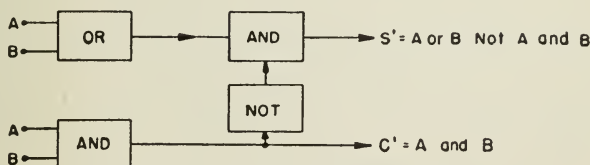
FIG. 13-3. Half adder, block diagram.

A vacuum-tube equivalent of the same arrangement is illustrated in Fig. 13-4.

In Fig. 13-4 the first grids of tubes V_1 and V_2 serve the function of *or* circuits, and the second grids serve as the *and-not* circuits, since their input is received from the *and* circuit output of V_3. V_4 serves as a clock-pulse gate not shown in the block diagram. Positive pulses are required to represent 1 digits in this configuration. V_1 or V_2 will pass a pulse if either receives a 1 pulse from A or B.   If a pulse enters through both A and B, V_3 passes a carry pulse and also cuts off V_1 and V_2 by driving their second grids negative, thereby inhibiting the pulse which would normally
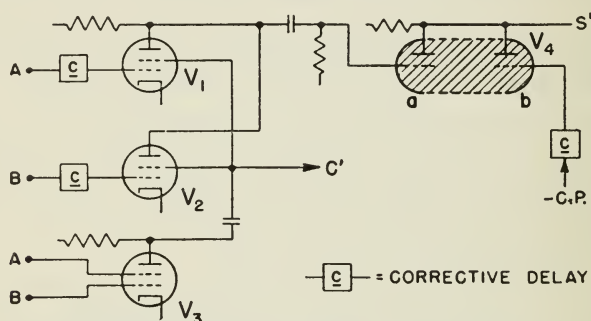
Fig. 13-4.   Half adder, schematic diagram.   The initials C.P. stand for clock pulse.

have come from V_1 and V_2. V_{4a} and V_{4b} constitute a gate through which clock pulses are passed when V_4 is cut off.   This is done to regenerate the pulses, as described in Chap. 4.   The squares marked with $\underline{C}$ are corrective-delay networks, which are required when a tube performs a *not* function requiring pulse inhibition such as V_3 performs with V_1 and V_2 in Fig. 13-4.   In order to inhibit reliably, it is necessary for the negative pulse to reach the second grid before the positive pulse reaches the first grid.   This is accomplished by delaying the signals to V_1 and V_2. Another case where a corrective delay is employed is in the gating circuit which causes clock pulses to replace other pulses.   In this case, the clock pulse should be delayed slightly.   In both cases, the undelayed pulse should be broadened so that it overlaps the delayed pulse.

The next step of the binary addition involves carry-over from the lower order digital addition.   If a synchronous serial system is being employed, the carry pulses can be stored in a delay circuit having a one-pulse time delay.   The development of the second half of a full binary adder stems from the rules of binary arith-

metic illustrated by Table 13-2 in which two digits A and B have been added to produce S' and C', and in which the carry from the previous addition is C''. In the first three columns are shown all eight possible combinations of three binary digits taken two at a time. The fourth and eighth rows are incomplete, because a sum and carry of 1 is an impossible combination in binary arithmetic, just as a sum of 9 together with a carry of 1 is an impossible combination in the decimal system when only two digits are added.

TABLE 13-2

Previous carry C''	Initial sum S'	Initial carry C'	New sum S	New carry C
0	0	0	0	0
0	0	1	0	1
0	1	0	1	0
0̸	1̸	1̸		
1	0	0	1	0
1	0	1	1	1
1	1	0	0	1
1̸	1̸	1̸		

The rules with which this table conforms also may be stated in logical terms thus:

$$S = (S' \text{ or } C'') \text{ and not } (S' \text{ and } C'')$$
$$C = (S' \text{ and } C'') \text{ or } C'$$

These rules are similar to those of a half adder if we substitute S' for A and C'' for B. There is an additional digit C' which must also be recognized by the output carry C. A full adder can be compounded of two half adders and an *or* circuit in the manner illustrated in Fig. 13-5.

Figure 13-6 is a schematic diagram of a serial adder, corresponding to the block diagram shown in Fig. 13-5. Tubes V_1 to V_4 form the first half adder, and V_5 to V_8 form the second. The squares marked C̲ indicate the presence of clock pulses. Their effect is simply to synchronize the times of arrival of pulses at various points in the circuit. The interval between successive clock pulses is the interval between the instants of arrival at the inputs A and B of successive pairs of corresponding digits of the numbers to be added.

Suppose that two numbers $AA'A''$ and $BB'B''$ are to be added. A'' and B'' represent the least significant digits of the two numbers. Suppose that each is 1. Following the notation used in Table 13-2, let C'' designate the carry generated by their addition in the half adder formed by tubes V_1 to V_4. This carry pulse is gated through V_9 and V_{10} into a one-pulse delay circuit repre-
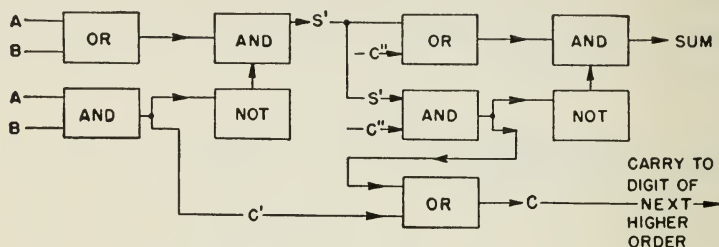
FIG. 13-5. Two-input adder, block diagram. Two half adders (see Fig. 13-3) are combined to form this system for adding three digits in two steps.
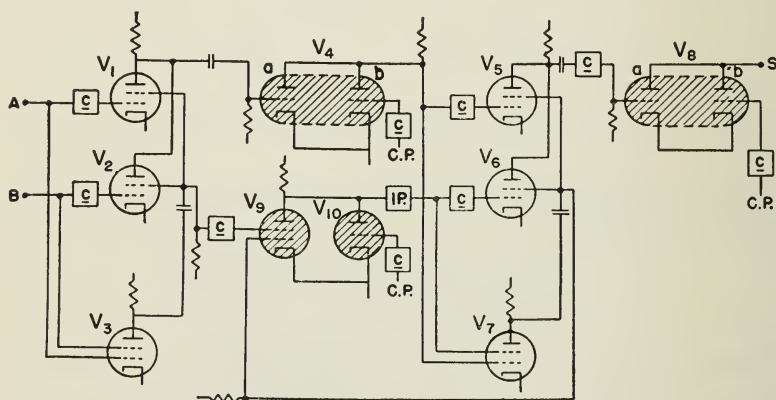
FIG. 13-6. Two-input adder, schematic diagram. The sum digits appear serially at S. Carries from each addition of S' and C'' (see Fig. 13-5) are fed back into the second half adder through the one-pulse delay circuit, $1P$.

sented by the square marked $1P$. Here it will be delayed for one clock-pulse interval.

Now suppose that the next two digits A' and B' are 1 and 0. Then their sum S' is 1. The pulse representing S' arrives at V_5 at the same instant that C'' (which was delayed one pulse time) arrives at V_6.

Tubes V_5 to V_8 constitute another half adder. Since C'' and S' are both 1, the *and* tube V_7 starts to conduct, blocking V_5 and

V_6 and generating a negative pulse which is fed back into V_9. V_9 is an *or* tube. A negative pulse produced by a carry from either half adder will make V_9 generate a carry pulse to be stored in the delay circuit. The pulse generated by V_7, when C''' and S' are applied to its two grids, therefore causes V_9 to store another carry pulse in the delay circuit. (This pulse, after being delayed, reaches V_6 simultaneously with the sum of A and B, the two
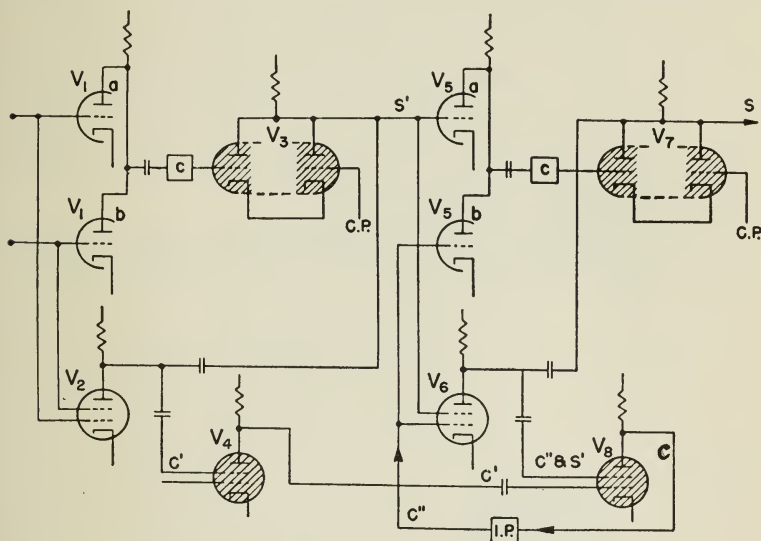
Fig. 13-7. Two-input serial adder using only eight tube envelopes.

digits of next higher order.) It can be shown, by following other possible combinations of input pulses through the circuit, that it will perform all the requirements specified by Table 13-2.

It is worth noting that V_9 can never be pulsed on both grids simultaneously because if S' is 1, C' must be 0 and there is therefore no pulse on the second grid of V_9; if S' is not 1, there can be no pulse on the first grid of V_9.

Since an adder of the type described uses nine logical units, and since there is usually a choice of several circuits for each logical unit, there are many possible variations of the simple coincidence adder.

A two-input adder which is a slight variation of the one described is illustrated in Fig. 13-7. Only eight tube envelopes are used in this circuit as compared to 10 in Fig. 13-6. The

half adders are composed of tubes V_1 to V_3 and V_5 to V_7, respectively. V_4 is simply an inverter for V_8, the *or* circuit. The output of V_8 is the carry C, and it is fed into a one-pulse time-delay circuit, where it is stored. The output of the delay line is connected to the input of the second half adder, so that carry pulses reach the half adder at the same time as the modulo 2 sum of the next pair of digits.

13-4-2. Three-input Binary Adders. The second type of coincidence adder to be considered has three inputs and two outputs and may be operated either from pulses or static voltages. Thus it is more versatile than the one just discussed. Only one rise time is required for the carry pulse, and as a result less time must be allowed for a complete carry when this type of unit is used in a parallel adder.

To derive the logical units of the three-input adder, consider the full binary addition table shown in Table 13-3. The rules

TABLE 13-3

Previous carry C''	Addend A	Augend B	Sum S	Carry C
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

with which this table conforms are

$$S = (A \text{ and } B \text{ and } C'') \text{ or } \{A \text{ or } B \text{ or } C'' \text{ and } not$$
$$[(A \text{ and } B) \text{ or } (A \text{ and } C'') \text{ or } (B \text{ and } C'')]\}$$
$$C = (A \text{ and } B) \text{ or } (A \text{ and } C'') \text{ or } (B \text{ and } C'')$$

The block diagram shown in Fig. 13-8 illustrates a circuit for enforcing these rules in a computer. Many variations of this circuit are possible, but the following may be considered representative.

Figure 13-9 is a circuit in which the digit 1 is represented by a positive pulse. Let us examine its behavior under each of the
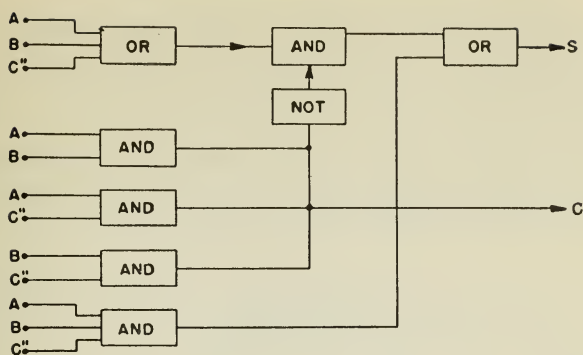
Fig. 13-8. Three-input adder, block diagram.
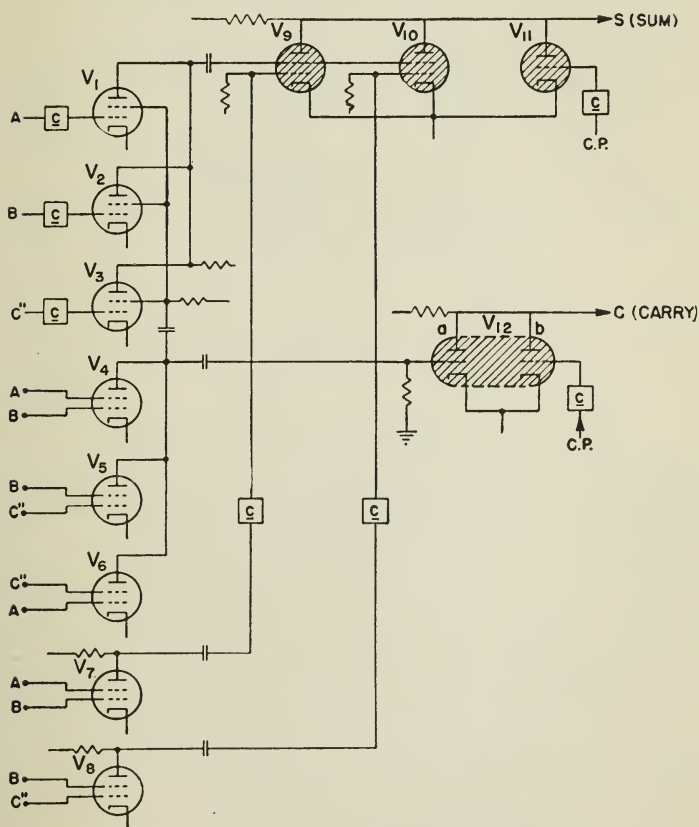
Fig. 13-9. Three-input adder, schematic diagram.

eight conditions represented by the first three columns of the table. Initially the shaded tubes are conducting. If both V_9 and V_{10} are cut off, a positive pulse will appear at S. If V_{12a} is cut off, a positive pulse will appear at C.

The first row in the table can be disposed of immediately. No pulses appear on any of the input leads, and no pulses are produced in either of the outputs S and C.
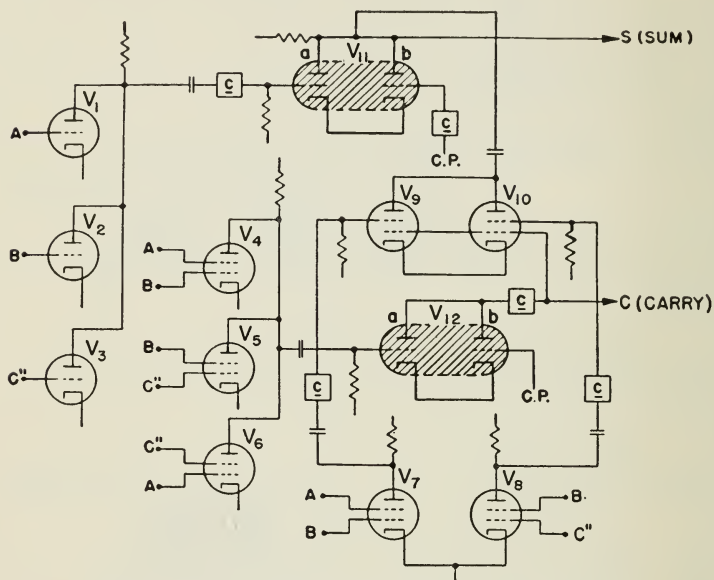
FIG. 13-10. Variation of three-input adder circuit.

Each of the second, third, and fifth rows has a 1 in *A or B or C''*. Either V_1 or V_2 or V_3 will be triggered; V_9 and V_{10} will both be cut off; a positive pulse, or 1, will appear at S.

The fourth, sixth, and seventh rows contain the combinations *A and B*, *B and C''*, and *A and C''*. One of these three combinations will trigger V_4, V_5, or V_6, which will prevent any of V_1, V_2, and V_3 from conducting. There will be no pulse, *i.e.*, 0, at S. At the same time, if any of V_4, V_5, or V_6 conducts, the grid of V_{12a} will drop, the tube will be shut off, and a 1 will appear at C. If both *A and B* or both *B and C''* are 1, then either V_7 or V_8 will conduct also. But either alone, cutting off only one of V_9 and V_{10}, has no effect on the output S.

The only remaining combination is illustrated in row eight,

where all three inputs are 1. The effect of this combination differs from the effect of a combination of two only in that both V_7 and V_8 conduct, cutting off both V_9 and V_{10}, producing a 1 at S.

Figure 13-10 is a slight variation of Fig. 13-9 in which tubes V_1, V_2, and V_3 provide the three single inputs which can cause a 1 digit to be passed through V_{11}. A coincidence of two or three
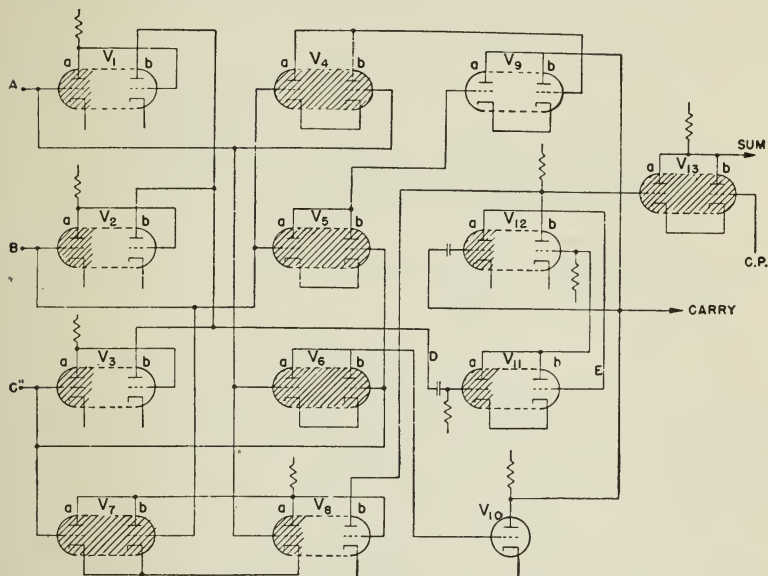
Fig. 13-11. Three-input adder circuit using twin triodes.

input pulses is sensed by the V_4-V_5-V_6 combination, which can cause a carry pulse to pass through V_{12} and also inhibit the sum pulse by shorting V_{11} through V_9 and V_{10}. However, the latter pair can be prevented from shorting out V_{11} in case all three input leads receive input signals, whereupon tubes V_7 and V_8 are both caused to conduct. In this case, V_7 and V_8 cut off V_9 and V_{10} in spite of their positive grids, and the original pulse to V_1, V_2, and V_3 is allowed to act through V_{11}.

Another three-input coincidence-type adder may be derived from triode tubes connected in the parallel, or Rossi, coincidence circuit described in Chap. 4. As an example, in Fig. 13-11 a negative pulse on either input A, B, or C'' cuts off V_{11a} and thereby

causes a sum pulse to appear at the anode of V_{12b}. The three input leads are so connected to the six grids of V_4, V_5, and V_6 that if any two of these receive a pulse, one of these tubes will be wholly (both sides) cut off. This causes at least one of the buffers V_{9a}, V_{9b}, or V_{10} to conduct and thereby cut off V_{12a}. In this circuit V_{11} acts as a logical *D and not* (*D and E*); it will allow a sum pulse to pass if D alone receives a pulse, but will not let it pass if D and E both receive a pulse, as in the case where two or three input pulses are received at A, B, and C. When
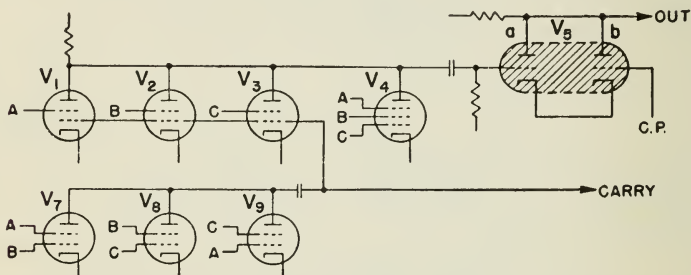
FIG. 13-12.   Three-input adder using triple-control-grid gate tube.

three pulses are received, V_7 and V_{8a} are cut off, thereby transmitting a sum pulse directly through V_{8b}. The carry pulse is passed through V_9 or V_{10} whenever two or three input pulses are received.

A final example is given to illustrate the simplification which results from the use of a triple-control-grid gate tube. As discussed in Sec. 14-3, this type of tube poses a difficult design problem at the present time. In Fig. 13-12 a 1 pulse received on A, B, or C will gate a clock pulse out of V_5 as a sum. Two pulses, however, will be detected by V_7, V_8, or V_9 and used to inhibit a sum pulse through V_1, V_2, V_3 by holding their second grids negative. Two pulses will also place a carry pulse in a delay circuit. Three pulses are detected by V_4, which gates a sum pulse through V_5 despite the inhibiting action of V_7, V_8, and V_9 on V_1, V_2, and V_3.

13-4-3. Algebraic Adders. So far, only adders of positive numbers have been considered. The *algebraic adder* adds or subtracts numbers according to their sign. An example of a method by which this can be accomplished is described and is illustrated in block diagram form in Fig. 13-13.

First consider the rules of binary subtraction. It may be facilitated by handling a borrow as a negative carry to the subtrahend. In other words, if the digit 1 in the subtrahend is to be subtracted from 0 in the minuend, the difference is 1, and a borrow is created. This may be handled by adding 1 to the next digit in the subtrahend. Thus a new subtrahend B' is created which may be handled as indicated in Table 13-4, where $A - B = D$ and negative carry C. From this table the

TABLE 13-4

Original digits			$B + C''$		$A - (B + C'')$	
Minuend	Subtrahend	Previous negative carry	New subtrahend	Intermediate negative carry	Final negative carry	Difference
A	B	C''	B'	C'	C	D
0	0	0	0	0	0	0
0	0	1	1	0	1	1
0	1	0	1	0	1	1
0	1	1	0	1	1	0
1	0	0	0	0	0	1
1	0	1	1	0	0	0
1	1	0	1	0	0	0
1	1	1	0	1	1	1

following rules can be formulated:

$$D = (B' \text{ or } A) \text{ and not } (B' \text{ and } A)$$
$$C = (B' \text{ or } C') \text{ and not } (B' \text{ and } A)$$

In order to form the new subtrahend B' and the intermediate carry C'', a half adder may be employed to add B *and* C''. The difference D may also be obtained from a half adder since the rule pertaining to it states that D is the sum digit of the sum of B' and A. The final carry C is obtained by gating B' *or* C' *and not* $(A \text{ and } B')$ through gate VI. A block diagram of the logical components which might be employed to obtain these conditions is shown in Fig. 13-13.

Figure 13-14 illustrates a circuit arrangement to carry out the rules just stated. Tubes V_1, V_2, V_3, and V_4, comprise a half adder which transmits a clock pulse representing B' to the
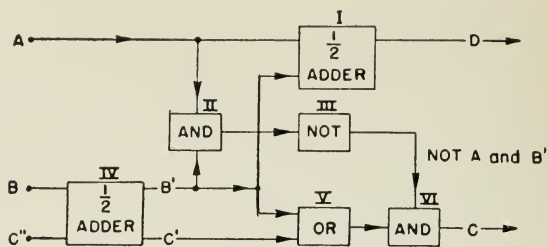
FIG. 13-13.   Subtractor, block diagram.
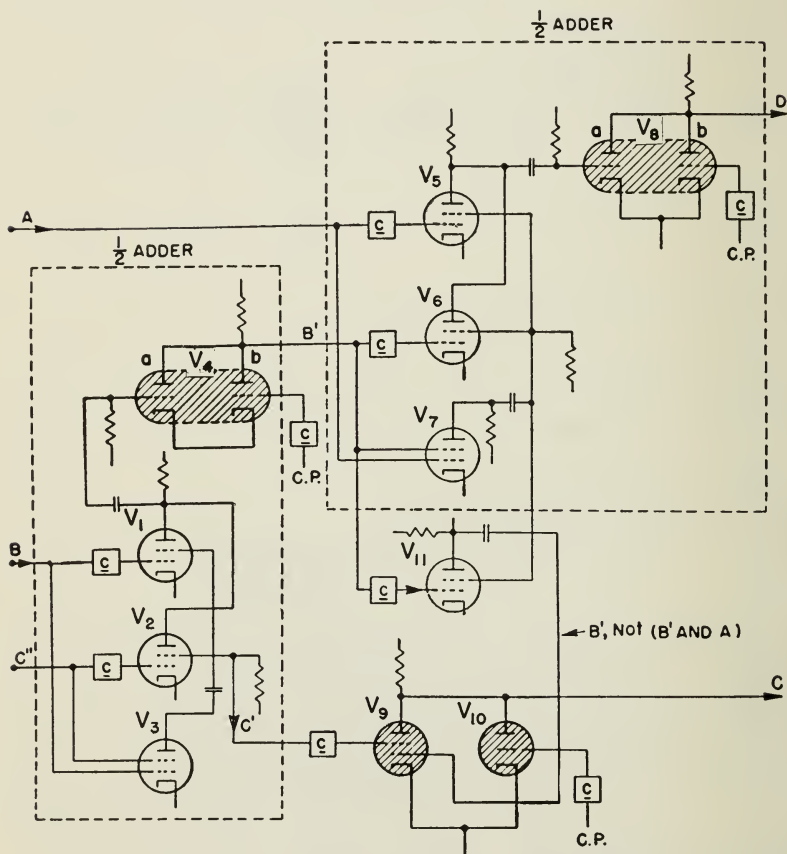
FIG. 13-14.   Subtractor, schematic diagram.

second half adder, made up of V_5, V_6, V_7, and V_8. V_7 also receives pulses from A and forms the logical A *and* B', which is used to suppress the A pulse through V_5 and the B' pulse through V_6 when these two occur simultaneously. V_8 is a clock-pulse gate for the difference digit D from V_5 or V_6 (but not both). V_5 and V_6 are a part of half adder I and also furnish the *not* (A *and* B') pulse in place of blocks II and III in Fig. 13-13.

The analysis of the carry circuit is not quite the same in the schematic as in the block diagram, but it is equivalent. V_{11} receives B' on one grid and B' *and* A on the other. The polarity of the signals is such that the output represents B' *and not* (B' *and* A). V_9 serves an *or* function by receiving this signal on one grid and C' from the first half adder on the other. Thus the output of V_9 is (C' *or* B') *and not* (B' *and* A). The output of V_9 gates a clock pulse through V_{10}, from which it emerges as the final carry C.

Algebraic addition may be performed by either an addition or a subtraction operation, the sign of the sum depending on both the sign and the relative size of the numbers to be added. The rules of algebraic addition are shown in Table 13-5 for two numbers A and B.

TABLE 13-5

Relative sizes of A and B	Sign of A	Sign of B	Operation	Sign of sum S
$|A| > |B|$	+	+	Addition	+
	+	−	Subtraction	+
	−	+	Subtraction	−
	−	−	Addition	−
$|A| < |B|$	+	+	Addition	+
	+	−	Subtraction	−
	−	+	Subtraction	+
	−	−	Addition	−

It may be seen that the operation to be performed is determined by the signs of two numbers. If the signs are alike, the operation is addition; if unlike, subtraction. Moreover, if the signs are unlike, the relative sizes of A and B determine which is to be subtracted from the other and the sign of the sum.

If the sign of each number is represented by a binary digit

occupying the first space preceding the number and coincident in time with a clock pulse, a control circuit containing a half adder may be used on the signs of the two numbers to determine whether they are alike or different.   It should be connected to detect (*A or B*) *and not* (*A and B*) so that a pulse is produced when the signs are different.   This algebraic-sign-control circuit is represented by block I of Fig. 13-15 which illustrates the complete block diagram of an algebraic adder.   By means of the sign-
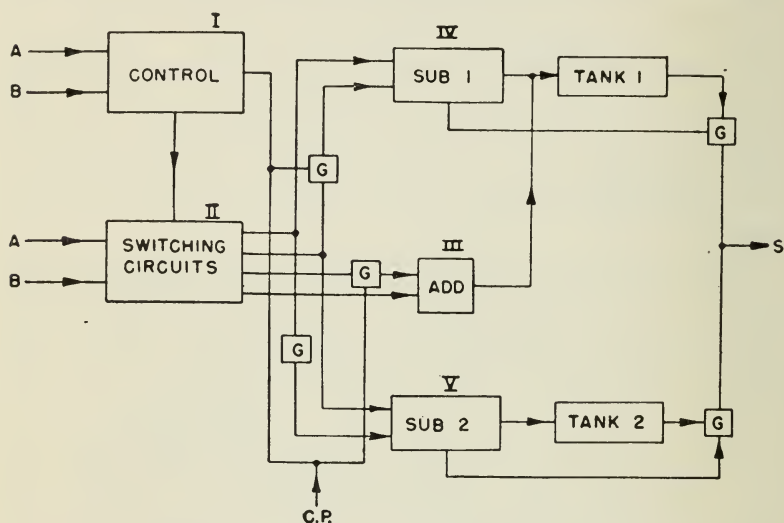
Fig. 13-15.   Algebraic adder, block diagram.

control circuit a pair of switches in block II is set so that the digits are sent either to the adder III or to the two subtracters IV and V, depending upon whether the signs are the same or different.   The three gates shown are used to remove the sign pulses accompanying some of the input digits.   In the case of addition, it is necessary to remove only one of the two sign pulses and pass the other on to the sum because all signs are the same.

   To determine which number is the larger without using any extra time, two subtracters are provided, and both subtractions are carried out simultaneously.   The subtracter which is subtracting a large number from a smaller generates an extra carry pulse at the end of the arithmetic operation.   This pulse is used to suppress the whole result before it is transmitted further.

The sign of the minuend is introduced into each subtracter, but since the difference and the sign are suppressed in the circuit with the incorrect result, the proper sign is received at S along with the correct sum.

13-4-4. Amplitude Adders.   A group of adders can be compounded on the basis of the principle that discrete voltage levels may be created in various circuits by applying integral numbers of pulses simultaneously.   The detecting circuits in the adders
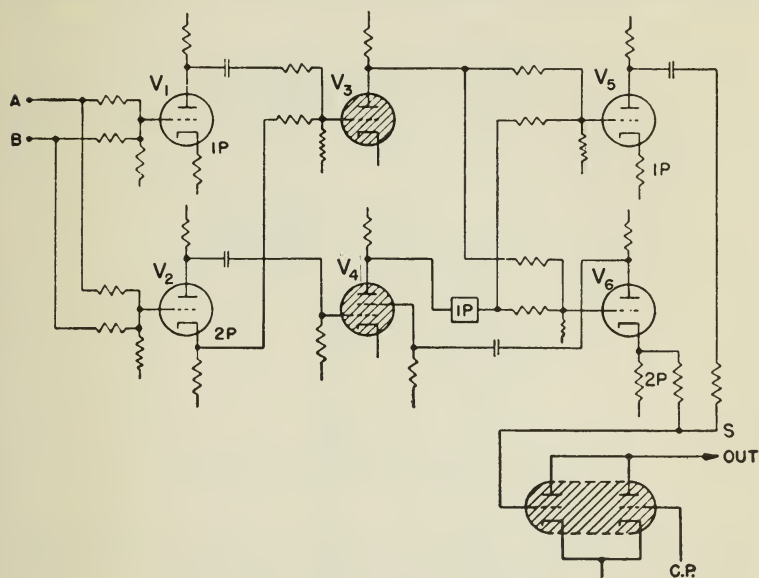
Fig. 13-16.   Two-input amplitude adder.

previously described respond to all voltages which are not below a critical value.   The circuits described here, however, respond to pulses above a specified minimum but not above a critical maximum.   Thus, a certain amount of simplification may be gained.   This simplification, however, is realized only at the expense of greater effort to reduce circuit-component tolerances and maintain pulse amplitudes within prescribed limits.

An amplitude adder compounded out of two amplitude half adders is illustrated in Fig. 13-16. V_1 is biased so that it can be made to conduct by the application of one standard pulse, and V_2 is made to conduct by two standard pulses.   Therefore V_1 is the logical *or* and V_2 is the logical *and* for A and B. V_3 is

conducting and biased so that it will be cut off by one or more pulses.   Therefore it responds to A *or* B from V_1.   When both V_1 and V_2 are caused to conduct, the two out-of-phase output pulses cancel at the grid of V_3, and no pulse passes.   A pulse through V_2 also sends a carry pulse through V_4 and into the one-pulse time-delay line. V_5 and V_6 are identical to V_1 and V_2 and comprise the second half adder, which adds the intermediate sum to the previous carry from the delay circuit.   If V_2 does not
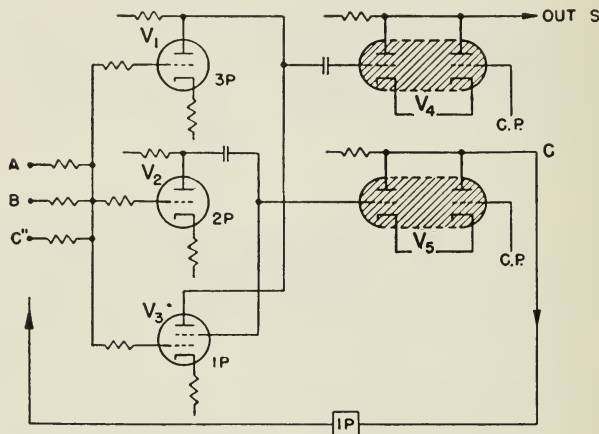
FIG. 13-17.   Three-input amplitude adder.

put a pulse into V_4, there is a possibility that V_6 will pass a carry pulse to V_4 and thence into the delay circuit. V_4 and V_6 never pass pulses simultaneously because if V_2 forms a carry, the intermediate-sum digit is zero, and V_6 does not receive a sum pulse from V_3.   This is because a sum and carry of 1 are mutually exclusive.

Figure 13-17 illustrates a further simplification through the use of amplitude-sensitive elements.   The circuit is the three-input type in which addend, augend, and carry are all received at the input.   In this circuit, V_3 responds to one or more positive pulses to cause a sum pulse to be gated through V_4. V_2 responds to two or more pulses and prevents the sum pulse from leaving V_3 by driving its second grid negative.   At the same time, it causes a carry pulse to be gated through V_5 and into the delay circuit. Three simultaneous input pulses are passed by V_1 as well as by

the other two input tubes. V_1 causes a sum pulse to be passed through V_4, and V_2 again initiates a carry pulse.

Amplitude adders are of particular interest in ternary arithmetic (based on radix 3). Several schemes have been suggested for the physical basis of ternary arithmetic, and one having a nonclassical radix which is easy to deal with physically will be described. The three digits of a ternary system might be taken as $+1$, 0, and -1, and these might be represented by a positive pulse, the absence of a pulse, and a negative pulse, respectively. The significance of the statement that the digits of a ternary system are $+1$, 0, and -1 is that the nth-order digit of a ternary number is $+3^{n-1}$, 0, or -3^{n-1}, depending upon the digits $+1$, 0, or -1 in the nth-order position. Such a representation is illustrated in Table 13-6.

TABLE 13-6

Decimal	Normal ternary	Nonclassical radix
0	000	0, 0, 0
1	001	0, 0, +1
2	002	0, +1, −1
3	010	0, +1, 0
4	011	0, +1, +1
5	012	+1, −1, −1
6	020	+1, −1, 0
7	021	+1, −1, +1
8	022	+1, 0, −1
9	100	+1, 0, 0
10	101	+1, 0, +1
11	102	+1, +1, −1
12	110	+1, +1, 0
13	111	+1, +1, +1

If the positive and negative pulses are held at an equal and standard amplitude, several of the rules of ternary addition may be enforced rather easily. These rules are:

Addend and augend	Sum	Carry
0 + 0	0	0
0 + (+1)	+1	0
0 + (−1)	−1	0
(+1) + (−1)	0	0

To enforce these rules requires only that pulse currents be passed through a common load resistor and the voltage drop taken as the sum.   This arrangement might be called a *pulse-bucking adder*.

The remaining ternary addition rules are:

Addend and augend	Sum	Carry
(+1) + (+1)	−1	+1
(−1) + (−1)	+1	−1

These two rules can be handled by amplitude circuits similar to those just described.   Figure 13-18 illustrates a ternary, amplitude half adder which obeys the rules just stated.

Fig. 13-18.   Ternary, amplitude half adder.

V_1 and V_2 are normally biased to cut off and therefore respond only to positive pulses. Furthermore V_2 is biased to such an extent that it will begin to conduct only if a positive pulse is received at A and B coincidentally. When it does begin to conduct, it cuts off V_1 at its second grid so that only gate V_4 gets a pulse. Likewise in case a single positive pulse appears at the input, only gate tube V_3 receives a pulse. V_3 and V_4 gate clock pulses into the S' and C' output leads. As an example, when V_4 gates a clock pulse corresponding to $(+1) + (+1)$, S' receives a negative pulse for the sum digit of -1, and C' receives a positive clock pulse for a carry digit of $+1$.

In the lower half of the circuit, V_5 and V_6 are normally conducting and respond only to single and double negative pulses, respectively. V_7 and V_8 are pulse inverters, and V_7 receives an inhibitory pulse on its second grid to stop the pulse to clock-pulse gate V_9 when two simultaneous negative pulses are received at A and B. V_8 is also a gate and passes a positive sum and a negative carry when two negative pulses are received at A and B. The problem of isolating the output circuits is not dealt with in this drawing.

13-4-5. Coded-decimal Adders. It is possible to assemble gates and switches to perform addition in any number system which follows a known set of rules. Any coded-decimal system is an example of one such computational scheme. The 10 decimal digits are given a code configuration, and arithmetic operations are carried out according to rules which account for the relationship between the code and decimal arithmetic. There are many possible code configurations, one group of which uses binary digits to represent decimal digits. (Thus there are four binary digits for each digit in the original decimal number.) One such code is known as the *excess-3 code*. In this code, 3 is added to each decimal digit to give an excess-3 value which is then represented by a corresponding four-place binary number. This code has the advantage that the 9's complement used in subtraction may be formed by transposing 1's and 0's. This feature of the excess-3 code is illustrated as an example of what can be done with various codes but does not indicate that it is the best code for all purposes. There are too many possible factors which are interrelated with machine details to be considered in picking a code; therefore, to discuss abstractly the relative merits of various codes is not helpful.

The following rules govern addition in the excess-3 coded-decimal system:

1. A carry in the fourth binary place coincides with a decimal carry. As an example:

Decimal notation:

$$5 + 6 = 11$$

Excess-3 binary notation:

$$1,000 + 1,001 = 10,001$$

2. When two coded-decimal digits in excess-3 code are added the sum is in an excess-6 code since each digit produced an extra 3.

Thus two sets of digits can be added directly and checked to see whether a carry occurred in the fourth place or not. If it did, the sum may be returned to an excess-3 code by neglecting the fifth digit and adding 13. If no carry occurred, 3 should be subtracted from the sum to reduce it from excess-6 to excess-3 code.

The block diagram of a circuit which performs these checks and additions is shown in Fig. 13-19. It requires a minor-cycle pulse generator which generates a pulse in coincidence with every fourth clock pulse.

Two excess-3 coded decimal numbers to be added are introduced serially into half adders I and II at A and B. The delayed carries produced by this addition are brought to gate G_1, which also receives minor-cycle pulses delayed one pulse time. The synchronization between the operand digits and the minor-cycle pulses is such that the latter reaches G_1 at the same time as the fourth-binary-digit carry (if one occurs). Whenever a fourth-place carry takes place, flip-flop FF, which has just been reset by generator I, is set by means of a pulse through G_1. Flip-flop FF in turn opens gate G_3 so that binary 1,101, or decimal 13, is introduced into half adder III in synchronism with the binary sum S_b of addend and augend. If no fourth-place carry occurs, FF opens G_2, and binary 0,011 from generator II is introduced into half adder III.

The second adder receives the sum from the first adder and a $+3$ or -3 from G_2 or G_3, respectively. Since the latter pulses

are delayed four pulse times after the former, the first sum is put through a four-pulse time-delay circuit before entering the second adder.   The second adder performs in a normal manner except for a provision to inhibit a carry pulse from the fourth binary digit if it should occur.   This is accomplished by passing all pulses through G_4, which is open during all but the fifth pulse time when the delayed carry from the fourth place would be
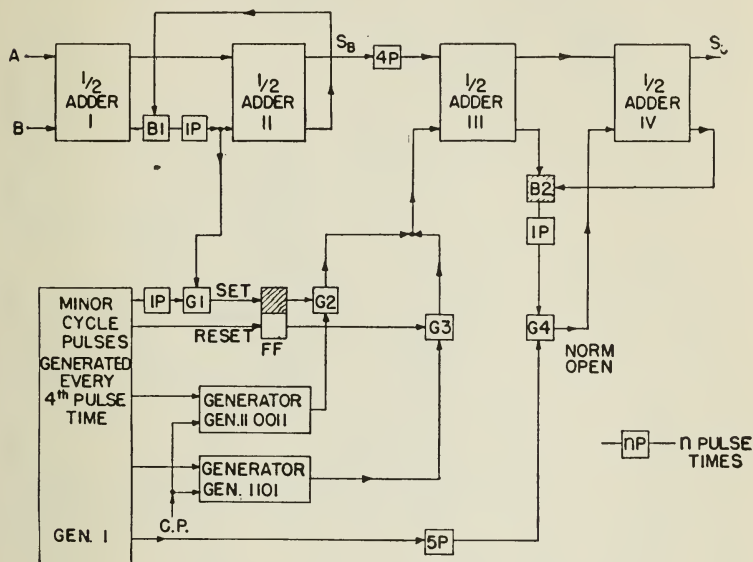
FIG. 13-19.   Excess-3 coded decimal adder.

received.   Then the clock pulse closes the gate and eliminates this undesired carry.

Figure 13-20 illustrates how the block diagram of Fig. 13-19 might be effected with electronic tubes. V_1 corresponds to gate G_1 and V_2 is the flip-flop circuit which is set and reset by the two halves of V_3. V_4 is merely an inverter for reset clock pulses.   The flip-flop controls gate tubes V_{11} and V_{12} corresponding to G_2 and G_3. V_5 and V_7 comprise the $+3$ or 0,011 generator, while V_6, V_8, and V_9 generate the 1,101 for -3.   The sections of V_{10} are inverters which introduce the $+3$ and -3 into gate tubes V_{11} and V_{12}. V_{13} gates a clock pulse into the second adder for each pulse received from V_{11} or V_{12}. V_{14} and V_{15} form a gate,

Fig. 13-20. Coded decimal adder, schematic diagram.

G_4, and a pair of inverters, respectively, which inhibit the fourth carry pulse from the second adder as described before.

13-4-6. Other Types of Adders. Addition may also be performed on a diode matrix similar to the code conversion tables described in Chap. 3. No significant simplification seems to result from the use of such adders.

Some work has been done to develop electronic adder tubes which operate with cathode-ray beams and anodes arranged to produce sum and carry pulses for proper beam deflections. Although such tubes may ultimately result in a more economical and compact unit than arrangements using conventional tubes, the latter are sufficiently flexible and economical to ensure their use for some time to come. An adder tube was developed at RCA during the Second World War.[6]

13-5. Elementary Ring-type Adders

A ring-type adder is essentially a counter which makes sums by counting out the digits of two numbers to be added. Decimal ring-type adders are used in the ENIAC and are described in Sec. 4-5. For binary addition, the counter ring may be reduced to a single flip-flop like the one illustrated in Fig. 3-2 on page 15. It is first cleared, and then the corresponding digits of addend and augend are presented to its input in turn. The sum is produced modulo 2, and another flip-flop or its equivalent is employed to store the carry.

Since the ring-type adders must receive one operand before the other, they are not well suited for use in serial adders of the type described in the previous section. They are, however, a popular form of adder for parallel arithmetic units.

13-5-1. Parallel Ring Accumulator. An *accumulator* is a counter which adds to its contents each number transmitted to it. An accumulator made of binary rings will count both numbers and their complements. Figure 13-21 represents an accumulator of this type. When it is desired to transfer a number from the register to the accumulator, all the G_1 gates are opened, and the transfer is made, leaving the original number stored in the register. If it is desired to transfer the complement of the number in the register to the accumulator, all the G_2 gates are opened.

At least two methods are available for detecting and handling carries in a parallel adder with ring-type elements. In one

method the carry is stored in a flip-flop which has a gate associated with it. The output of each gate is connected to the input of the binary counter to the left. The carry is begun by pulsing all carry gates. Since each carry may initiate another carry, the process must be repeated until the carrying process is finished. If the adder contains n counters, n pulse times normally have to be allowed in order to assure that all carries are effected. As explained in Sec. 13-3, the average number of
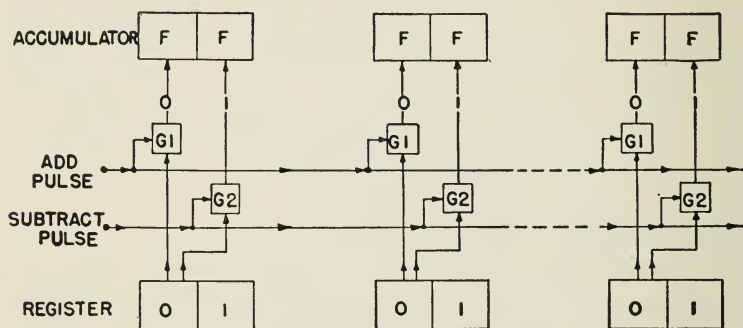
Fig. 13-21. Binary accumulator.

carries is well under n, so, on the average, less time is required for an automatically sequenced computer which can continue as soon as all carries are complete. A circuit connected to each flip-flop stage in such a way that it detects the presence of zeros in all the flip-flop stages can be used to sense the completion of the carry.

A second method of adding allows all carries to be made simultaneously, thereby enhancing the characteristics of nonautomatic sequence computers. An adder of this type is described in the next section.

13-6. Larger Arithmetic Units

It is usually desirable to code operations, other than addition and subtraction, so that they can be handled automatically. These other operations, however, are compounded out of the basic addition operation. Several complete adders and multipliers are described in this section in order to indicate the types of arithmetic units which might be devised.

13-6-1. Serial Arithmetic Units. A serial adder using any one of the coincidence adder units described in Sec. 13-4 is illus-

trated in Fig. 13-22. If a three-input adder is used, provision must be made to return carry pulses to one of the input leads at the next pulse time. The serial adder employs two storage units into which electrical pulses are entered serially, and from which they emerge at some later convenient time. Mediums suitable for this purpose, such as transmission lines or rotating drums, are discussed in Sec. 14-2. The box R contains the amplifiers and clock-pulse gates required to regenerate these pulses after each trip through the storage medium. The combination of storage unit and regenerating unit forms a device for
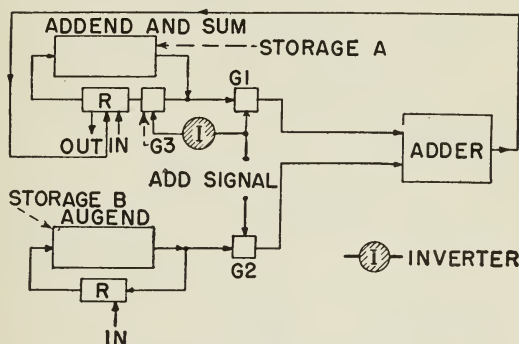
Fig. 13-22. Serial adder.

retaining digital data; such a device is defined as a *register*. The shaded circle designated I, upon receiving a signal to begin the addition operation, provides switching potentials to gates G_1, G_2, and G_3. Obviously, the delay time in the storage section of a serial register must be equal to or greater than the time required for all the serial digits in a binary number.

If addend and augend are first entered into their storage units A and B, respectively, and an *add* signal is applied as the lowest order digits are passing into R, consecutive higher order digits from A and B will be summed in the adder. G_3 is opened by the *add* signal, so addend digits do not continue to circulate through A but are replaced by sum digits, which enter as addend digits leave. Thus addition with a serial adder requires one cycle time and leaves one operand circulating in one register and the sum in the other.

With three registers having circulating times twice that required for introducing the operands, a serial binary multiplier

may be developed from the serial adder described. A device of this sort is illustrated in the block diagram of Fig. 13-23. This multiplier performs additions of the multiplicand, the additions being controlled by successive digits of the multiplier. The partial sums are circulated in storage A. The multiplier and the multiplicand are entered into their respective C and B registers. As the two lowest order digits emerge from the storage registers, a clock pulse is applied to reset flip-flop FF and to

Fig. 13-23. Serial multiplier.

open gate G_5. If the multiplier digit is a 1, it sets FF, which in turn opens G_4, which adds all the digits of the multiplicand into register A. After $2n + 1$ pulse times, the second-order digit of the multiplier is at G_5, and the first-order digit of the partial product in register A is entering the adder. If the second-order digit of the multiplier is a 1, the multiplicand is again added to the contents of the product register. This time, however, the multiplicand is delayed by one pulse time so that a shift of the number by one digit order is effected before addition with the preceding partial product. In this manner the process is repeated, the multiplicand shifting one place for each circulation and being added to the partial product each time a 1 appears in consecutive higher order positions of the multiplier.

Although the serial multiplier requires a minimum of equipment, the total multiplication time is the product of the circulation time times the number of digits in the operands. As an example, it would take 1,800 pulse times to multiply two 30-digit binary numbers in the manner described. Some saving in time might be obtained by rounding off the digits in the product, but one circulation would still be required for each digit in the multiplier.

13-6-2. Parallel Arithmetic Units. The basis of the parallel arithmetic unit is either the elementary ring-type adder of Sec. 13-5 or the coincidence adder of Sec. 13-4. The former type has been studied by the Raytheon Company,[10] and the results of this investigation are used as the basis for a description of a parallel adder with simultaneous carry. A modification of the Raytheon adder which would allow multiplication is outlined briefly in order to indicate the degree of complexity of a parallel multiplier.

Parallel Adder with Simultaneous Carry. An adder in which all carries are accomplished simultaneously has been proposed by the Raytheon Company[10] and is illustrated in Fig. 13-24. The gates labeled G_1 are employed to introduce the addend digits into the accumulator. The presence of a carry is indicated by a 1 in the addend register and a 0 in the accumulator and is detected by G_4, which in turn controls G_2. This latter gate receives a carry pulse which, if the gate is open, is distributed to the delay line following each stage of the register. The carry pulses which pass G_2 are also inverted in I and sent to G_3, which senses the coincidence of a 0 in the accumulator stage and an open G_4 gate. This indicates a double carry and means that a carry pulse should be passed on to the next stage. If G_3 is open, that is exactly what happens, so all carries are made with one carry pulse. Actually, n rise times are required to make n carries, and in the Raytheon circuit, about five pulse times have to be allowed on this account. This is a great deal faster than the serial adder, which requires about n^2 pulse times for addition of two n-digit numbers.

Shifting in a Parallel Ring-type Register. In order to multiply numbers in the conventional manner of summing the partial products, it is necessary to shift the accumulated partial product with respect to the new partial product before the two are summed. A shifting register consisting of one flip-flop for each
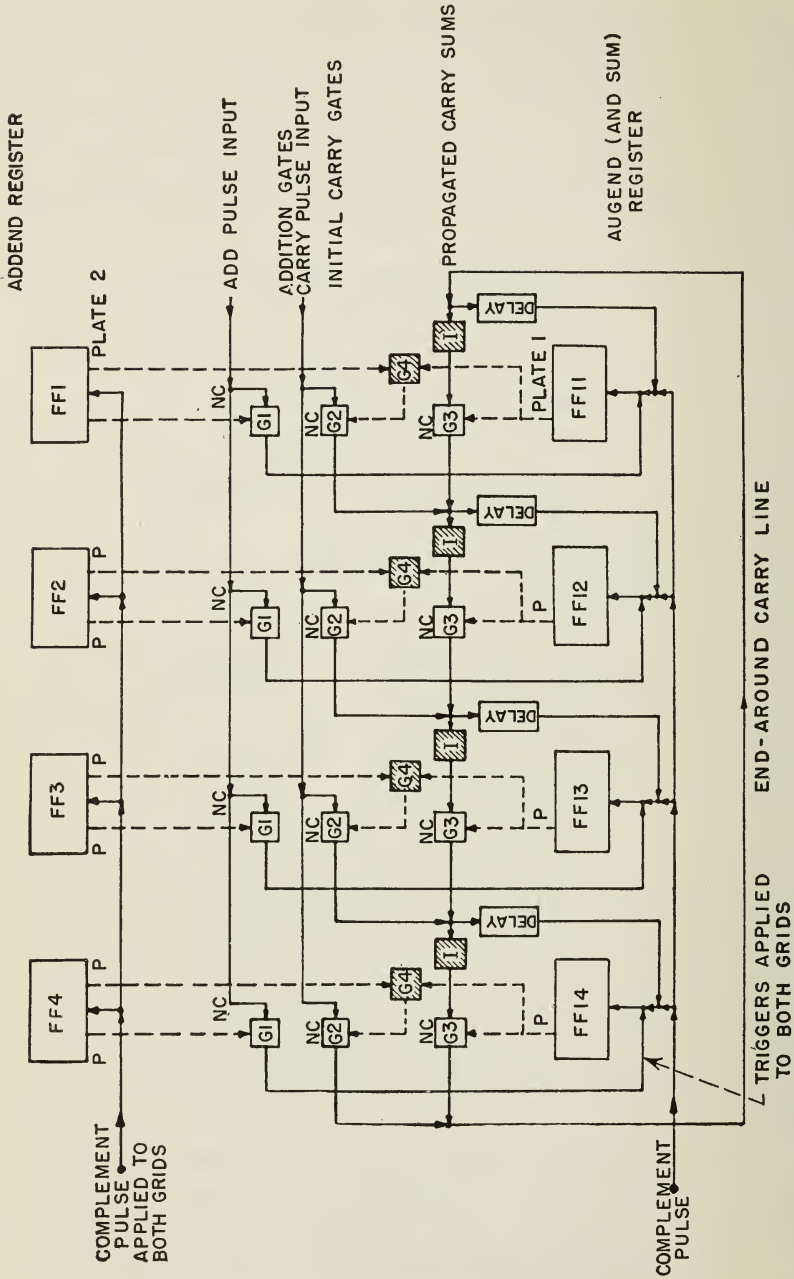
Fig. 13-24. Parallel adder.

digit stored is illustrated in Fig. 13-25. One plate of each flip-flop is connected through a delay circuit to the grid of the opposite tube in the next flip-flop. A bus delivers reset pulses to the remaining grids simultaneously. When a reset pulse is applied to these grids, the flip-flops which are storing a 1 are returned to the 0 condition, and a pulse is introduced into each delay line following a flip-flop in which a 1 was stored. At the end of the delay period, the pulses emerge from the delay lines to set the following flip-flops. Thus, binary numbers stored in
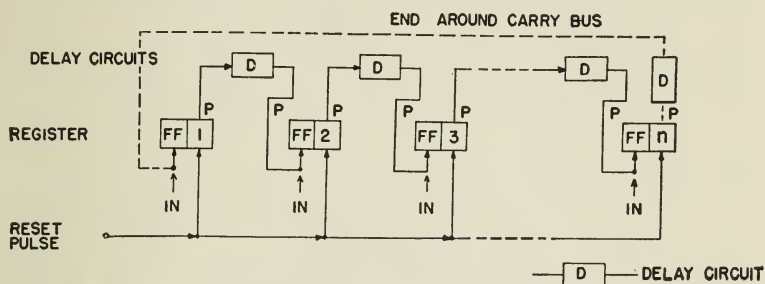
FIG. 13-25. Shifting register.

this register are shifted one space to the right each time a reset pulse is applied.

Another way of shifting is to provide two registers and provision for shifting numbers from one to the other. By providing a set of gates for each type of transfer, the digits may be shifting to the right, to the left, or not at all as the number is passed from one register to the other.

In some operations it is desirable to retain digits which would be shifted out at one end of the register during the shifting process by entering them into the other end of the register as they leave. As an example, in the parallel shift register of Fig. 13-25, the digits shifted out of FF_n could be returned to FF_1 through a delay circuit and the dotted line shown.

Parallel Multiplier. A multiplier might be devised using the parallel adder and the shift register which are described above. A three-digit multiplier is illustrated in Fig. 13-26. The multiplicand is placed in the upper flip-flop register, and the multiplier is placed in a serial register. The product accumulator is twice the length of the operand registers and is also a shift register.
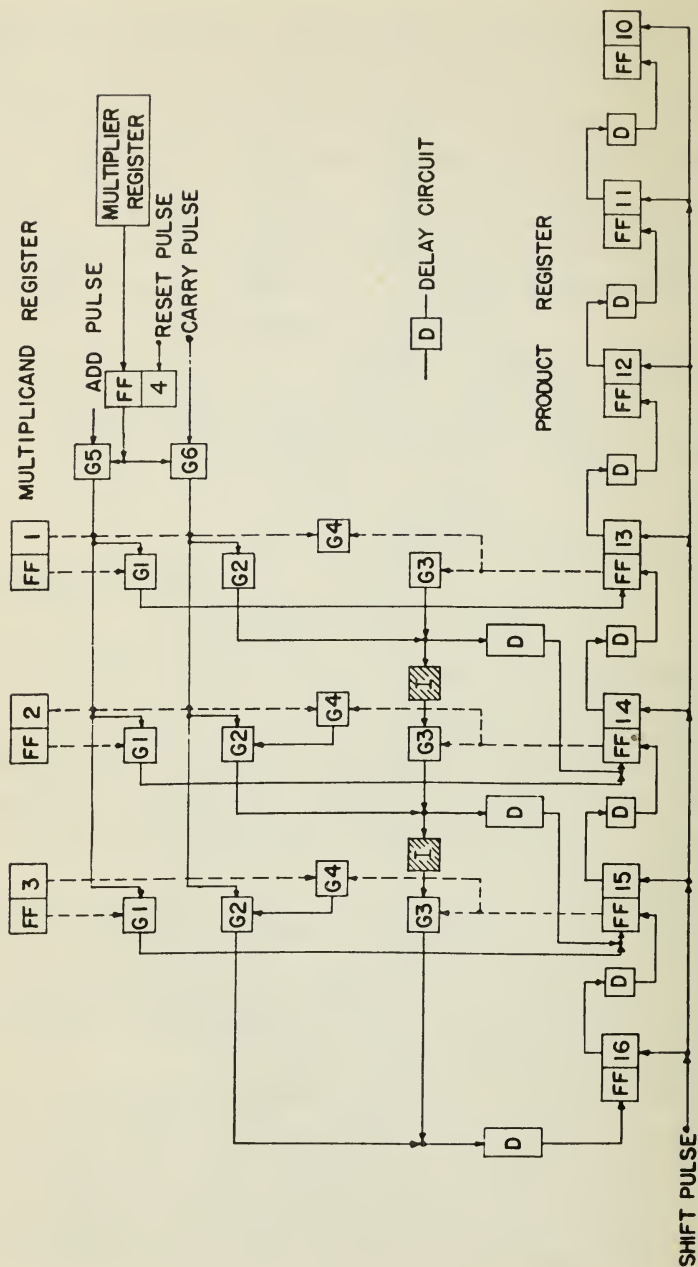
Fig. 13-26.   Parallel multiplier.

When FF_4 is set by 1 pulses from the multiplier register, it opens gates G_5 and G_6 for one addition time (about five pulse times). During this time an add and a carry pulse are applied to adder gates G_1 and G_2 so that the multiplicand is added into the number in the accumulator. The timing of the various steps is such that the lowest order digit of the multiplier is introduced into FF_4 first. If it is a 1, the multiplicand is added into the accumulator. After the addition time, the partial sum in the accumulator is shifted to the right, FF_4 is reset, and the second multiplier digit is applied to FF_4. The process is continued in this manner until all the digits of the multiplier have been used. At that time, the number in the accumulator is the product of the two numbers multiplied. In the example shown, the last digit of the product is also cut off.

It may be seen that the time required to multiply two n-digit numbers is n addition times.

REFERENCES

1. Alt, F. L., A Bell Telephone Laboratories Computing Machine—I, *Mathematical Tables and Other Aids to Computation*, Vol. 3, No. 21, pp. 1–13 (January, 1948).
2. Burks, A. W., Electronic Computing Circuits of the ENIAC, *Proceedings of the Institute of Radio Engineers*, Vol. 35, pp. 756–767 (August, 1947).
3. Burks, A. W., H. H. Goldstine, and J. von Neumann, *Preliminary Discussion of the Logical Design of an Electronic Computing Instrument*. Report prepared under Contract W-33-034-ORD-7481 between the Research and Development Service, Ordnance Department, U.S. Army, and the Institute for Advanced Study, Princeton, N.J. (June 28, 1946).
4. Hilbert, D., and W. Ackerman, *Grundzüge der Theoretische Logik* (J. Springer, Berlin, 1928).
5. Mauchly, J. W., Preparation of Problems for EDVAC Type Machine, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 203–207 (1948).
6. Rajchman, J. A., *R.C.A. Computron, Final Report*, LC PB 40611 (Office of Technical Services, Department of Commerce, March, 1943). 129 pp.
7. Tarski, Andrew, *Introduction to Logic* (Oxford University Press, New York, 1939).
8. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *Report on the ENIAC, Technical Report I* (University of Pennsylvania, Philadelphia, June 1, 1946). 2 vols.
9. University of Pennsylvania, Staff of Moore School of Electrical Engineering, *Progress Report on the EDVAC* (University of Pennsylvania Philadelphia, June 30, 1946). 2 vols.
10. West, C. F., and J. E. DeTurk, Digital Computer for Scientific Applications, *Proceedings of the Institute of Radio Engineers*, Vol. 36, No. 12, pp. 1452–1460 (December, 1948).

TRANSFER MEDIUMS

14-1. Introduction

Computing consists of performing arithmetic operations on numbers. In the strictest sense, numbers are abstractions; arithmetic operations are performed on *representations* of these abstractions or numbers. The numbers themselves are defined by the rules for performing the arithmetic operations. For example, the mark 8 in the decimal system and the mark 100 in the binary system are *representations* of the same *number*. Any computing machine operates, in accordance with applicable rules, upon representations such as these.

In the following discussion *number* is used in a broader sense. It is used to mean an ordered set of digits considered as a unit, regardless of the meaning of the digits. For example, some of a particular set of digits (the first 20, for example) may be the binary representation of a particular number. Some others in the set may represent an address code and still others in the same set an instruction to the machine. We shall call the entire set a number. Thus, we are using *number* as a synonym for what others in the computing-machine field have termed a *word*.

In this chapter we will discuss the means for transferring numbers in space and time during the course of a computation. Any connected series of arithmetic operations implies the necessity for such transfers, and the concept of transfer is basic in any discussion of computing machines. To facilitate discussion and to provide a basis for grouping transfer systems for comparison purposes, the definitions set forth below have been adopted.

Transfer in time alone is necessary when a number is to be reused later at the same location in space. Transfer in both time and space is necessary if the number is to be reused later at a different location in space. As noted below, some types of time transfer are properly spoken of as storage. Therefore, it may be seen that the terms *transfer* and *storage*, as used with

respect to computing machines, are related. *Transfer*, the broader of the two, includes *storage*, because *transfer* includes any change in location with respect to either time or space, or both. The term *storage* is used with respect to a transfer in time only. Moreover, not all transfers in time are properly referred to as storage operations. A number is in *storage* when its representation is introduced into a medium with the intention that it shall remain there until specifically called for and withdrawn. The number is merely being *delayed* (as contrasted with being stored) when it is introduced to some medium from which it is automatically delivered after a fixed interval of time. This distinction is illustrated in the following examples:

Suppose it is desired to transform the serial representation of a 15-digit number to a parallel representation. This transformation may be visualized by imagining 15 digits proceeding in single file down one channel and then turning simultaneously at right angles, to proceed down 15 separate parallel channels. In practice, however, the transformation from serial to parallel expression requires the delaying of each digit by an appropriate interval. The first digit is delayed for the longest time, the second for the next longest, etc. When the delays are properly adjusted, the last digit reaches the entrance to its particular channel just as each of the other 14 digits is ready to be shunted into its particular channel.

Each of the first 14 digits in this case has been transferred in time, but the delaying operation is not thought of as storage. Storage is the term used, rather, when data are made to exist in some interpretable form until called for. An elementary storage is accomplished by means of a pencil and a piece of paper. A legible mark on a piece of paper remains in interpretable form for use whenever needed, so long as the paper is not destroyed and the mark is not erased. The problem in calculating with the aid of pencil and paper is not to delay information but simply to store it for later reuse. In a computing-machine operation it may be necessary to store information in interpretable form for later reference (perhaps for many later references), and it may be necessary also to delay particular data for the purpose of carrying out a particular computation. The example of the serial-to-parallel conversion illustrates the latter.

Transfer systems used for storage or delay purposes are referred

to as either *static* or *dynamic*.   In a *dynamic* system, the data
are in motion, usually cyclically, relative to the medium in which
they are represented.   An example of the dynamic system is the
sonic delay line; in this system, trains of sound waves, represent-
ing digits, travel through a substance in which the velocity of
sound is slow as compared with the velocity of electrons in a
conductor.   The waves are introduced at one end of the delay
line and received at the other.   The same signal may recirculate
indefinitely through the delay line (being retransmitted at one
end each time it is detected at the other), but it is always in
motion relative to the reading and writing elements, *i.e.*, the
receiving and transmitting transducers.

In a *static* system the stored data remain fixed with respect to
the storage medium until transferred by the operation of some
switching device.   A system composed of electromagnetic relays
is an example of a static system.   A rotating magnetic drum also
is a static storage system, according to this definition, because
data recorded on the drum are static with respect to the surface
of the drum.   However, the data on the drum are in motion
relative to the magnetic heads which read and write upon the
surface; hence, in one very important respect, the static mag-
netic-drum system and the dynamic delay-line system are
equivalent.   If we define as a *cyclic* storage system any system
in which each stored datum continually moves in cycles with
respect to the reading and writing stations, then both the rotating
magnetic drum and the delay line are *cyclic* storage systems.
Using the converse as the definition of a *noncyclic* system, a bank
of relays is an example of this type.

If we define the *access time* of a system as the time required
to withdraw a number from storage, the essential difference
between cyclic and noncyclic systems is the following: In a cyclic
system, the average access time is closely related to the time inter-
val between successive appearances of the same storage cell at a
location where its contents can be withdrawn.   In a noncyclic
system, the speed of operation of the switching device is impor-
tant.   The relationships between these factors and other
design considerations are discussed below.   Another difference
between cyclic and noncyclic systems which is worth noting
as an aid in narrowing the definition is the following: In a non-
cyclic system, the transfer of data in space from the storage

medium to some other part of the computer involves noncyclic motion in *space*, *i.e.*, motion of relay contacts, of an electron beam, or of some other physical mechanism. In a cyclic system, this transfer of data to or from storage involves noncyclic motion in *time* while a cyclic motion in space is taking place. This motion in time is the delay which takes place between the instant an order related to a particular storage cell is activated and the instant that cell, in the course of its cyclic motion, arrives at the proper reading or writing station.

Another pair of mutually exclusive terms, one of which may be applied to any storage system, is the pair *erasable* and *nonerasable*. These terms are essentially self-explanatory. For example, data stored as combinations of punched holes in a tape or a card are in nonerasable form; data represented by the condition of magnetization of a magnetic surface are in erasable form.

There is one more pair of terms which will be found useful in discussing storage systems: *volatile* and *nonvolatile* (or *permanent*). These terms refer to the degree of physical stability of a representation of stored information. In a *volatile* storage medium, like a delay line, retransmission of each signal once during each storage cycle period is required. In case of power failure, the stored data are lost. If the data are recorded on magnetic material, they are in *permanent* storage, because the representation of each digit, a particular magnetic condition of a permanent magnet material, will not deteriorate even if the computing machine is turned off; an external force is required to eradicate the data.

14-1-1. Cyclic Storage Systems. Sonic delay lines and magnetic drums are probably the best known of the cyclic storage systems. Other types have been used successfully, *e.g.*, electromagnetic delay lines. These are discussed in detail in other sections of this chapter. The use of radar transmissions to distant objects and the deflection or intensity modulation of electron or ion beams also have been suggested. In theory, at least, any device for sending a signal over a path relatively long as compared with the signal used to represent a digit is applicable in a cyclic, dynamic, storage system.

The basic physical characteristics of any cyclic storage system, with respect to each separate single-digit channel, or path, in the system, are:

1. The storage cycle, which is
   a. The physical length of the dynamic storage path, divided by
   b. The speed of a digit in storage, relative to the receiving and transmitting (or "reading" and "writing") stations.
2. The number of digits which can be carried on the storage path at one time, which is
   a. The physical length of the dynamic storage path, divided by
   b. The physical length of the signal representing a single digit in the storage medium.
3. The radix, which is the number of elementary states distinguishable in the medium.

Primarily upon these characteristics depends the quantity of information which can be stored and the speed with which the computing mechanism can have access to each stored datum.

For purposes of illustration, let us consider a magnetic-drum storage system. A binary digit is represented on the surface of the drum by the magnetic state of a particular area. If this area is magnetized in one direction, the digit it represents is 1. If the area is magnetized in the other direction, the digit is 0.

In practice, 1,600 inches per second has been found to be a reasonable drum surface velocity (corresponding to a storage concentration of 80 digits per inch). A small drum, capable of storing 1,024 digits in each single-digit path, or track, has a storage cycle of about 8 milliseconds. (The physical length of the storage path is actually a little more than 1,024 times 0.0125 inch—which is the physical dimension of the digit, measured in the direction of motion—because a small amount of dead space must be left on the drum.) In practice, a drum 4.3 inches in diameter will carry 1,024 binary digits per track.

Let us apply the same considerations to a mercury storage system. The velocity of sound in mercury is about 5.7×10^4 inches per second (1.45×10^5 centimeters per second[48]). For a storage path 14 inches long, the storage cycle time will be

$$\frac{14}{5.7} \times 10^{-4} \cong 250 \text{ microseconds}$$

Suppose the digits are introduced to storage at the rate of four digits per microsecond, which is the rate at which the BINAC,

built by the Eckert-Mauchly Computer Corporation, operates.
The physical length of the signal representing each digit in
storage is therefore

$$5.7 \times 10^4 \times 0.25 \times 10^{-6} \cong 0.014 \text{ inch}$$

The number of digits carried on the storage path at one time is

$$\frac{14}{0.014} = 1,000$$

Mercury delay lines are suitable for the storage of a few
thousand digits. The problem of controlling the temperature
(in order to maintain constant velocity) is not difficult to solve
in a delay line of reasonable length, *e.g.*, of the order of a thousand
digits.

It is not practical to store a very large number of digits in a
mercury system because of the number of vacuum tubes required.
For purposes of estimating engineering complexity, it may be said
that approximately 11 tubes are required to generate signals and
control and read each delay line in the EDVAC. At the rate of
11 tubes per thousand-digit line, 11,000 tubes would be required
for a million digits. In terms of cost and complexity this would
be a high price. On the other hand, such a large storage system,
with the access speed of a thousand-digit delay line, is not likely
to be required. (The engineering aspects of the mercury storage
design problem are discussed in more detail in Sec. 14-2-6.)

In a cyclic storage system, the average time required to locate
and read from or write into any given storage cell may be related
directly to the storage cycle period. This average access time
can be reduced by adding more reading and writing stations.
For example, several magnetic recording heads can be placed at
equal intervals around a magnetic drum to reduce the average
access time without reducing the storage capacity or speeding
up the drum. However, the price paid for this reduction in
average access time is high, in terms of increased equipment
complexity. Applying this same technique to a mercury delay
line is not practical because the additional stations introduce
serious attenuation.

Where greater storage is required (*e.g.*, storage for 1 million
digits) the mercury delay line is not practical, without serious
modification to increase its capacity by eliminating the effect

of temperature variation on sound velocity, because of the number of vacuum tubes required. At the rate of 11 tubes per 1,000 digits, 11,000 tubes are required for 1 million digits. On the other hand, mercury delay lines are well suited for storing a few thousand binary digits.

The following seem to offer some promise of immediate and future applicability in cyclic storage systems in computing machines:

1. Sonic delay lines, liquid medium.
2. Sonic delay lines, solid medium.
3. Magnetic recording.
4. Electromagnetic delay lines.
5. Punched tape, read electrically.
6. Punched tape, read photoelectrically. (Note possibility of punching holes with electric arc.)
7. Photographic tape, read photoelectrically.
8. Phosphor drums.
9. Electrostatic storage tubes, read cyclically.

14-1-2. Noncyclic Storage Systems. The important considerations pertaining to noncyclic storage systems are:

1. Radix—the number of distinguishable states which the medium may assume.
2. The complication of the switching equipment—electromagnetic, electron beam, or other.
3. The complication of equipment required to set the storage medium in its various radix states and of the equipment required to detect these states (write and read).

The types of noncyclic static storage which now seem to be of interest are:

1. Relay—electromagnetic or electron tube.
2. Electrostatic storage—cathode-ray tube or Selectron.

As indicated above, in meeting operational requirements with noncyclic storage systems the main difficulty lies in providing switching facilities. If the choice of items of intelligence stored or to be stored is by means other than physical motion of the storage medium, and if items corresponding to even a few thou-

sand binary digits are to be stored, it is clear that the problem of switching between the storage cells may become awkward.

A convenient, well-known, rapid means of switching utilizes the deflection of beams of electrons or the blocking of beams of electrons by electrostatic fields. This means gives promise of becoming useful in connection with computers; this is particularly true since other ingenious techniques have been developed to utilize this switching when the number of leads is small enough to compare favorably with the number required for dynamic storage. This system is described more fully in Secs. 14-2-8 and 14-2-9 on electrostatic storage schemes.

In cyclic storage, as opposed to noncyclic storage, the switching function may take place mainly on a time basis. This immediately raises a host of problems in connection with synchronization, and it implies some sort of periodicity with consequent limitations on the speed with which the system can function. The same objections apply to systems in which switching is on a space displacement basis. Maximum velocities attainable imply a minimum consulting period. Thus, unless the information is to occur and to be used in an orderly and predictable manner, any cyclic storage system is wasteful of time. Noncyclic systems are intrinsically much better suited for consulting where the addresses of the information stored are likely to be somewhat random.

It would be possible to use as noncyclic storage most of the cyclic storage schemes which do not involve transmission of intelligence as wave motion or some similar motion through a static medium. Punched tape, for instance, could be used in a static noncyclic storage application, but this would involve a tremendous increase in the number of reading stations as well as complications in switching.

Early noncyclic, static storage systems frequently used an array of switches. Both mechanical relay systems, with some scheme for stabilizing and holding each relay in position, and vacuum-tube relays arranged in flip-flop circuits were used. In either case, storage of binary digits is simple and straightforward. There are no difficulties other than switching difficulties.

In exactly the same way, punched tape, photographic tape, and other similar systems could be used instead of relays. Again the only difficulty would be in switching. Magnetic storage also could be used in a noncyclic application (though radical

modifications in the system of reading would be required) because the state of magnetism in an object is easily read by means of a coil moving relative to the magnetic medium.

The use of magnetic material seems almost automatically to imply the idea of motion, primarily in cyclic systems. However, three different schemes are available for reading the state of magnetism of a body without moving the body itself: (1) a coil on a magnetic core, or a *magnetic head,* may be moved past the magnetized object and thus furnish the same relative motion which would be used to detect the magnetic condition of the body if the body itself were moved; (2) the nonlinearity of the magnetization curve may be used, so that an increase in magnetomotive force in a particular direction changes the flux substantially only if the substance is not saturated in that direction; (3) a cathode-ray beam may be deflected by the magnetic field of the substance. Of these schemes only the second seems to show any promise at all, and no well-developed efforts to utilize this scheme are known to have been carried out. However, much work has been done recently on saturable-core reactors, and use of some types of these in magnetic storage is feasible.

14-2. Transfer in Time

Examples of the need for transfer of numbers in time are given in Secs. 5-3 and 14-1. Several time storage mediums are discussed in the following material, and for convenience of comparison they are considered in application to:

1. The high-speed storage section of a computer.
2. Permanent storage.
3. Supplementary storage.

In a given computer there may be some overlapping of these types of storage, but each has its own distinctive features.

The high-speed storage section of a computer includes the registers of the arithmetic unit, and its requirements are intimately tied in with the type of arithmetic chosen. The most important criteria for high-speed storage mediums are (1) their access time; (2) the flexibility with which they can be switched from one unit to another; and (3) their simplicity and reliability.

Permanent storage is desirable if the computer is to operate with the aid of tables, repetitive program routines, repetitious

similar input data, and output data from one problem as input data for another problem. The choice of permanent storage medium determines certain characteristics of the input-output equipment and usually is chosen with that point in view. Input-output systems are hereinafter discussed as an important type of permanent storage system. The principal criteria for judging a permanent storage medium are (1) the compactness of the medium, or the number of digits stored per unit cost and per unit volume; (2) the life expectancy of the secondary medium on the shelf and in operation; (3) the rates at which data may be written in and read from the medium; (4) the ease of handling the medium mechanically by the machine, and by human hand when being removed from or loaded into the machine; (5) the ease with which the data on the medium can be changed for additions, deletions, or corrections; (6) the facility with which the data can be reproduced as a new permanent storage medium; (7) the suitability for convenient inspection and checking; and (8) the reliability of the medium and its associated equipment.

There is usually a need for supplementary storage having characteristics which are a combination of those of high-speed storage and of permanent storage. This is the result of the expense in complexity of extending high-speed storage to cover all the computing requirements and the relatively poor accessibility of numbers in permanent storage when the numbers required must be read in a random order. Criteria for good supplementary storage mediums are (1) the flexibility and speed of switching for reading from and writing on the medium; (2) the capacity of the medium in number of digits stored as a function of volume, cost, and complexity; (3) erasability and reusability of the medium; (4) reliability of the medium and its associated equipment. Since access time and capacity of a supplementary storage system are closely related, (1) and (2) are necessarily considered together. It is imperative that supplementary storage systems be erasable and reusable.

14-2-1. Relay Storage. The only forms of high-speed relay storage employed in present-day large-scale digital computers are the electromechanical[1,2] or the electron-tube relay. In the Harvard Mark II and the Bell Telephone Laboratories machines, electromechanical relays are employed, whereas in the ENIAC, electron-tube relays, or flip-flop circuits, are employed. The

latter type is also employed in the EDVAC and other electronic computers. Although they are relatively large and expensive, relays are also used in the first three computers as supplementary storage mediums. The present tendency is not to use relays as supplementary storage for large amounts of data. They are well suited, however, as the high-speed storage medium in conjunction with arithmetic units.

Admissible Radix. The operation of the electromechanical relay as a switch is described in Chap. 4. In present-day computers, two stable positions of this type of relay are utilized so that these relays store digits modulo 2. It is possible, however, to establish three stable states in a relay equipped with centering springs and a pair of coils which deflect the armature in opposite directions. As a variation of this configuration, a permanent magnet might be made for the armature so that its deflection is dependent upon the direction of current flow through a single coil. Thus it is possible to construct an electromechanical relay which could be used to store radix 3 numbers. The conventional electron-tube flip-flop circuit has but two stable states.

Access Time. The *pickup*, or setting, time for commercial electromechanical relays used in computers varies from 1 to about 10 milliseconds. The relays used in the Mark II equipment operate in from 6 to 10 milliseconds. The wire-brush type of relay used by the IBM Corporation contains up to 12 contacts and operates in 2 milliseconds. Another relay discussed in Chap. 3 has been made to operate in 1 millisecond. The total access time for digits in the Mark II is $16\frac{2}{3}$ milliseconds. This is longer than the pickup time of the relays and includes factors of safety for the synchronization between units and current-switching equipment described later in this section. It seems likely that 10 milliseconds is about the minimum access time which can be expected when electromechanical relays are used.

The time duration of a digit in a flip-flop circuit may be very short but is limited by the figure of merit of the tubes. Flip-flop circuits cycle satisfactorily in 0.1 microsecond and may be expected to have a reliable access time of 0.2 microsecond.

Mode of Operation. A desirable feature associated with the electromechanical relay is the simplicity of the auxiliary equipment required to generate digital signals. A relay may be set

by a d-c pulse, the wave shape of which is not critical. Reading information out of a relay requires merely the detection of the position of the contacts and may be done instantaneously. The armature may be held either by the current flowing through the coil or by a mechanical latch which is released by a second coil. In the latter case, the storage unit is said to be nonvolatile because it can sustain loss of power without losing all the information stored.

If it is desirable to have the contacts of the relay carry large amounts of current, a synchronous system may be employed in which the relay contacts are opened and closed before current is applied to them. This is done in the Harvard Mark II machine, in which pulses of direct current are generated by large cam-operated switches and distributed to the relays as signal pulses. Thus, the only current-making and -breaking contacts are at one point where they are available for inspection and service. This form of operation, however, has the disadvantage that some time is wasted between the relay operation time and signal-current time.

The electromechanical relay is versatile in its application as an arithmetic component because it has a pair of low-impedance input or actuating leads which may be kept isolated completely from those of the output.

Reliability. Reliability in computing devices is of the utmost importance, and therefore a serious problem arises with the use of electromechanical relays. Interference with one or more computing operations may result if a particle of dust happens to rest on the contact faces, since electrical contact is made between two points of microscopic size. Since there is no direct way to detect such a failure either before or after it occurs, the use of checking circuits is desirable. Fortunately, it is relatively easy to add contacts to an electromechanical relay for checking purposes. This was done on the Bell Telephone Laboratories machines, and their ability to resume operation quickly after a failure is one of their outstanding features.

A comparison between the electromagnetic relay and the electron tube reveals that the former has a longer life expectancy in operating hours but will operate fewer times than the latter. Furthermore, the electromechanical relay is subject to failure from vibration and dust on the contact faces and is apt to require

more delicate care than a vacuum tube. The relay usually operates on less power than the filaments of a vacuum tube, but the latter can operate on alternating current taken directly from the power mains through step-down transformers. Most fast-acting relays operate on direct current; hence when the rectification conversion loss is taken into account, their power requirements are comparable to those of electron tubes. The two are also nearly the same in cost and size. The greatest difference in characteristics is in the maximum speed of operation; the electron-tube relay is at least 10^4 times faster than the electromechanical relay. Therefore, in situations where computing may be done serially, a single tube can perform the same operations in the same time as a large number of relays operating simultaneously in parallel.

Size and Cost. The cost of relays of the type used in the high-speed storage section of a computer varies between \$1 and \$5 per digit stored, depending upon the amount of auxiliary equipment which is used. The required volume per digit stored is about 3 to 6 cubic inches.

In small computing devices, relays are particularly suitable because of their availability and versatility. On the other hand, they have also been used quite lavishly in large computing machines such as the Harvard Mark II computer, which employs a total of about 13,000 relays. These relays are made by Autocall, and most of them are of rugged construction. They have six-pole, double-throw contact arrangements with large silvered contacts, consume about 6 watts at 100 volts direct current, and operate in 6 to 10 milliseconds. About 5,000 of these relays of the latch type, Model HHA-JD, are used in the internal storage. There are about 9,000 relays in the largest Bell Telephone Laboratories machines. Most commercially available electromagnetic relays are described in a book compiled at MIT,[1] and flip-flop circuits are amply covered in the references for Chap. 3.

14-2-2. Punched-tape Storage. A common storage medium for digital information is found in tape which is opaque to light or is physically strong enough to operate mechanical feelers and has a low electrical conductance. Information is stored by means of an array of holes which are placed in the tape and located in reference to sprocket perforations.[3-7] Tape is used for storage in the Bell Telephone Laboratories and the Harvard

machines as well as in many special-purpose devices. Because it is nonerasable, it is most suitable as a medium for permanent storage. It may be used in rolls or, where iteration is desired, in endless loops.

Admissible Radix and Arrangement. An area on a tape of the sort discussed above may have two possible elementary states, punched and not punched. The admissible radix is therefore 2. Several adjacent columns, or data channels, may be punched along a length of tape.

Access Time. The limiting factor which determines the minimum spacing between holes may be the tendency of the tape to shrink or stretch, the precision with which the holes can be punched in the tape, with respect to some reference such as a row of sprocket perforations, or the minimum spacing at which holes can be reliably sensed by the reading equipment. The minimum time interval between digits is directly proportional to the minimum space between holes and inversely proportional to the maximum linear velocity at which the tape can safely be moved with accurate guidance.

Two types of reading equipment are in present use: electrical sensing and photoelectric sensing. Electrical sensing of holes in punched tape is accomplished by passing the tape over an electrically charged metallic platen, over which are placed metal wipers in position to make contact with the platen through holes in the tape, or by arranging fingers which slip through the holes in the tape and operate switches. The top speed to be expected with electrical sensing is about 100 digits per column per second. The output is in the form of electrical pulses, which are convenient for computer use.

The top speed at which data can be recorded on standard teletype tape by the Western Union 10B Reperforator is 10 operations per second. Since as many as five holes may be punched at once, this corresponds to a speed of 10×5 or 50 digits per second. The Model 12A transmitter will read 75 digits per second.

A faster method of sensing holes in punched tape is by means of photoelectric cells. The tape is driven past a mask which contains optical apertures of proper size and spacing for one hole in each row. These apertures are illuminated through the holes in the punched tape by a constant source of light. Light which

passes through the holes and the apertures is gathered in an optical system which is arranged to focus the light on photocells. An electrical pulse is generated in a photocell each time a punched hole passes over an aperture in the mask. With this type of system, the British have succeeded in sensing holes in punched tape at a rate of 5,000 digits per second per column. A five-hole tape was used, so this corresponds to a sensing rate of 25,000 binary digits per second.

Mode of Operation. It is sometimes convenient to use tape as a cyclical storage medium by forming it into a continuous loop. In the Harvard Mark II equipment these loops are used to store tables of functions and other data which are to be used repeatedly. Punched tape is also used as a supplementary storage medium in the Bell Telephone Laboratories machine. By punching data into the tape at one station and reading them again at another, a short-time storage unit is made available. This feature proved to be much more useful than originally anticipated. A practical limit to the number of digits which may be stored by this method is the mechanical difficulty of handling a long loop of tape. A cyclical series of digits can be created in this way, however, with no complication in reading and recirculating equipment, because punched tape is provided with sprocket holes which automatically serve to maintain synchronism between the tape and the read-write equipment. In order to shorten the access time of the system, it is quite feasible to use independent read and write heads at various locations around the loop, thereby adding only the complication of additional read-write circuits and switches. Presently available tape-preparing equipment employs mechanical punches actuated by electromagnets and operated from a keyboard similar to that of a typewriter. The electromagnet circuits have been modified to operate from the outputs of computers but are inherently limited in speed by their physical inertia. The possibility of perforating tape with an electric arc is of interest because of the higher perforating speeds which might be attained.

If it is desired to correct or change data punched on tape, it is necessary to cut the tape, add or remove sections, and splice it together again. This process is tedious but is simplified by the fact that the recorded data can be inspected visually and directly. This factor results in a simplification of the checking process when errors are found to exist in the storage medium.

Reliability. An important advantage of punched tape at the present time is the fact that tape-preparing and -reading equipment has been in commercial use for many years and has been refined to a high state of reliability. The wearing characteristics of this tape depend upon the method of reading, the amount of acceleration required in driving it, and the type of drive mechanism. Photoelectric sensing can be effected without use of fingers or pins pressing against the tape, thereby eliminating one source of wear. In any case, a paper tape has a fairly short life owing to wear on the sprocket holes even when it is driven at constant velocity. On the other hand, standard commercial equipment is available with which to reperforate duplicates of the master tape for spares so that they can be replaced as they become worn.

Teletype tape is reliable to the extent that it can operate under a reasonable range of temperatures and is unaffected by small amounts of dust. Since sprocket or reference holes are punched in the tape as the data are added, the synchronization problem is minimized. The Bell Telephone Laboratories machines use a six-hole tape and a code in which two of the six holes are used for each digit. As a check against additional holes which might appear in the tape as a result of wear, a circuit is provided which stops the machine in case more or fewer than two holes are detected simultaneously.

Size and Cost. There are so many materials from which punch tape can be made and so many varieties are now available that it is beyond the scope of this book to discuss each separately. However, the common five-hole teletype transmitting tape is described as a representative sample, and some data concerning other tapes are listed in tabular form at the end of Chap. 17. The capacity of teletype tape is about 3×10^7 binary digits or punches per cubic foot, and the cost of the tape is trivial: about one dollar per 3×10^6 binary digits.

Punched tape is commercially used in several different forms. The Wheatstone tape used for Morse code representation is about $\frac{1}{2}$ inch wide and accommodates two rows of dots and a row of sprocket perforations. This tape costs about 20 cents per roll and may be procured through the Link Paper Co. in New York.

Teletype tapes are available in five-, six-, and seven-hole widths and may be procured at about 25 cents per roll through

manufacturers, the names of whom will be supplied by the Teletype Corporation upon request. The five-hole tape which is common to office equipment may be procured at stationery stores at about 75 cents per 1,000-foot roll. A tape of greater durability than ordinary paper tape is identified as gray fiber tape and has been procured through Western Union by special arrangement.

There are also opaque tapes, such as film leaders, available from the Eastman Kodak Company and other film manufacturers, which might be used for punched-tape application. Thirty-five-millimeter film leader costs about 2 cents per foot when purchased in quantity.

Commercial Tape-punching and -reading Equipment. The Teletype Corporation makes printers, tape readers, and tape punches of several styles. This equipment is highly refined and has been in reliable use for many years. The following equipment is incorporated into the Harvard Mark II computer:

1. Teletype Model 15, Page Printers.
2. Teletype Model 10B, Reperforators.
3. Teletype Model 12A, Transmitters (five-hole).
4. Teletype Model 11A, Transmitters (six-hole).
5. Teletype Model 101, Motor-Driven Reperforators (both five- and six-hole).
6. Teletype Model 3B, Distributor-Transmitter (five-hole).
7. Teletype Model 10A, Distributor-Transmitter (six-hole).
8. Distributors, stepping switches (or rotary switches), relays, and associated gear.

The operating speeds of some of these pieces of equipment are of interest. Ordinarily, the printers operate at between 390 and 460 operations per minute, although they will operate as fast as 600 operations per minute. The transmitters normally operate at 900 operations per minute on a five-wire system. This equipment is used by Western Electric and Western Union.

Equipment for perforating and reading the two-hole Wheatstone tape used in automatic telegraphy is available through the McElroy, Kleinschmidt, and Creedy Companies. The McElroy mechanical tape-reading equipment operates at speeds up to 500 words per minute. Two pairs of binary spaces are used to represent a Morse code letter, so there are approximately 25 pairs of binary spaces per word.

The IBM Corporation also uses tape-perforating and -reading equipment which is available on a special rental basis. Much of this equipment is adapted to transferring data from punched cards to tape and back again.

Drive Mechanisms. In the Bell Telephone Laboratories machines a variable-speed tape drive is used, and tables are indexed by sections. Therefore, when searching for an entry in a large table, the section may be designated, and the tape drive drives the tape at high speed until the desired section is reached, whereupon the speed is decreased to normal reading speed so that the entry may be withdrawn. This feature complicates the drive mechanism considerably, so that the cost of punched-tape equipment depends to a considerable extent on the drive mechanism employed.

14-2-3. Photographic Storage. Storage of digital data on photographic strip film is similar to storage on punched tape in many respects and may, therefore, be classified with it in the group of nonerasable storage mediums.[8–12] Data may be represented on photographic film by areas of different opacity which are applied photographically to a film base and which are detected photoelectrically. Work is being carried out on the utilization of film as a nonerasable storage system by O'Neal and Tyler of the Eastman Kodak Company. This has been described in the literature.[10–12] Most of the information to be presented on this subject has come from these studies.

Recording of digital data may be effected by illuminating the film through a mask which restricts the images of a bank of discharge tubes or the screens of cathode-ray oscilloscope tubes. The discharge or c.r.o. tubes act as light sources which may be modulated by the signals to be recorded. In accordance with the coded information supplied to the modulators of the light sources, beams of light are varied in intensity to produce a horizontal row of clear and opaque blocks. After one row has been exposed, the film can be advanced and another row of data composed and photographed. The film may be moved continuously if the flashing time of the light source is made short compared to the time required for the film to travel the aperture distance of the mask.

Admissible Radix. Although it is possible to record and distinguish between several degrees of opacity on film, present

efforts are being concentrated on a two-state representation of data in which the light is either passed or cut off. This on-off technique has been considered necessary for perfect dependability of operation.

Access Time. The digit area or block area must be considerably larger than the limit set by the resolving abilities of the film emulsion. A more important operational criterion is the requirement for optical reliability when the film is slightly dusty or somewhat scratched. Coded areas 0.010 by 0.020 inch are of the magnitude considered feasible. It is probable that dots of this size may be recorded at a rate of 10^4 elements per second per channel. If 50 channels are placed side by side on a 35-millimeter film, this rate would result in a total recording speed of 5×10^5 digits per second.

Reading is accomplished in a manner similar to that described for the photoelectric reading of punched tape. A bank of photocells is located on one side of a slit which is illuminated from the opposite side through the film. An optical system is arranged to project light from each block section of the slit to a corresponding photocell. Thus, opaque areas of the film interrupt the light to corresponding photocells, and transparent areas allow light to pass through. Electric currents generated in the photocells from varying illumination are then amplified to form the output signals. Photoelectric reading equipment is limited in speed only by the rise time of the photocell and its amplifier and has an inherent speed of at least 10^6 centimeters per second; furthermore it is capable of reading independently of scanning velocity. The maximum reading speed achievable is limited by the accuracy with which the film can be guided past the reading stations and is in the order of 10^4 elements per second per channel, not 10^6 centimeters per second.

Mode of Operation. The minimum requirements for film recording include a light source which can be modulated at a high repetition rate, an optical system capable of fine resolution, and a film-advancing mechanism capable of guiding the film accurately. Flash tubes of the type used in commercial strobotacs are being investigated for their applicability to this field as light sources.[8] Cathode-ray tubes with low-persistence screens are also being studied.[11] A single c.r.o. tube is used to generate a row of marks across the film by stepping the electron

stream across the face of the tube, the beam being modulated so that spots of light appear at the desired locations behind a mask on the tube face.

A straightforward method of recording a binary number and its inverse simultaneously, using a c.r.o. tube, has been demonstrated. If a mask with two horizontal rows of apertures is placed over the face of the c.r.o. tube, the vertical sweep may be used to choose one of each pair of spots in a given column, and the horizontal deflection may be used to scan the spot of light over the columns in sequence. Thus, when the light spot is
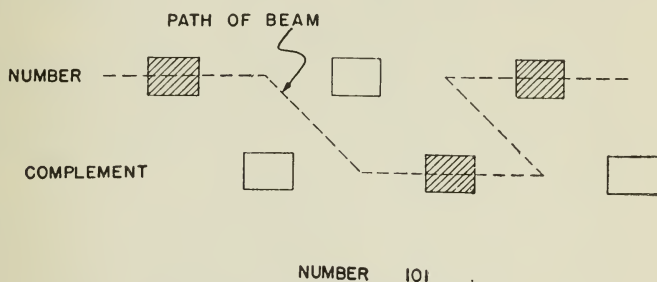
FIG. 14-1. Binary optical recorder.

recorded in the number, a dark spot is recorded in the corresponding column of the inverse row, and vice versa. Several mask patterns are possible, and one is illustrated in Fig. 14-1.

Film, like punched tape, may be used in rolls and in endless loops. Since it is essentially a permanent storage medium, it is most useful for storage of tables and as an output medium.

Reliability. Dependable driving equipment for standard-width strip film has been developed for the motion-picture industry.

The life expectancy of the film depends upon the method by which it is driven. Ordinary film has a reasonably long life when driven with sprocket wheels. The life expectancy is increased if the film is driven with rubber rollers instead. When the sprocket holes are not used, however, a synchronizing channel must be recorded on the film. As the film ages, copies may be made by standard commercial copying procedures. Minor changes may be made by the usual film-cutting and -splicing techniques. Film may be inspected visually with the aid of film-viewing apparatus.

The reliability of this medium depends largely on the reliability of the handling equipment and on the degree to which the film can be kept clean and unscratched. A positive signal for each digit and a check on reading equipment may be had by recording the inverse of each number below the number. This allows a positive signal for both 0's and 1's and provides the basis for a good checking system, for in the film reader each digit of the number and its inverse can be compared. Failure of the recording equipment to record a digit or failure of one reading section is shown in the comparison process.

Size and Cost. Data may be stored on photographic film to a density of 10^9 binary digits per cubic foot and at a cost of 10^6 binary digits per dollar. Thirty-five-millimeter positive film is available at about \$3.75 per 200-foot roll, and commercial processing averages about \$3 more per 200-foot roll, depending upon the quantity. Films with faster emulsion speeds cost two or three times as much as positive film.

A disadvantage of film as a storage medium is the complexity of photographic processing. Automatic equipment is available which speeds up the conventional developing procedure and which continuously turns out strip film at a rate of 6 feet per minute. An even faster process may result from work now being done by the Haloid Corporation in connection with Xerography, an electrostatic and thermal dye-deposition process.

It is not known that any equipment of the type described is now commercially available, and its future cost is difficult to estimate. One of the largest factors in cost is the type of film drive required. In some cases a servodrive is desired which will search for and locate values in a table. Driving the film at a variable velocity introduces the problems of film-roll acceleration, servo stability, and the provision of a reference sensing equipment with a link to the drive mechanism.

The use of film as the medium on which the final output of a computation may be recorded is of interest and is described in Chap. 15. Film can be exposed at a high speed and can record charts and graphs as well as numbers and coded data. Enlarged prints can be made on paper for visual inspection.

14-2-4. Magnetic Storage. Magnetic recording[17] is a technique which presents attractive possibilities for both the internal storage and the input-output functions of computing devices.

The alterable physical state requisite to any storage medium is in this case the intensity of residual magnetization (magnetic moment per unit volume) in a small volume of magnetic material. The techniques used in the magnetic recording of speech and music have been treated extensively in the literature.[19-39] These sources should be consulted for background material and fundamental principles. The material in this section will attempt to indicate factors peculiar to the magnetic recording of digital information.

The Magnetic Recording Process. The magnetic recording process briefly is as follows: To record information, the medium, or carrier, is caused to move past the gap of a recording head, which is a suitably designed electromagnet energized by the signal current. Each succeeding element of the carrier is brought into a definite magnetized state, in which it tends to remain upon leaving the recording field. The recorded pattern of remanent magnetization along the carrier is related, in a manner determined by the system characteristics, to the time variation of the signal current. This pattern is read by drawing the carrier past the gap of a reproducing head, which may be similar to the recording head. Some of the magnetic flux from the section of the carrier in close proximity to the gap follows the low-reluctance path through the core of the head. Time variation of this flux generates an output voltage in the coil surrounding the core.

Reading can be accomplished by means of saturable-core reactors, also, as indicated on page 310.

In the recording and reading of digital information, this output voltage need not be a replica of the recording signal voltage. It is sufficient that the system recognize in a reliable manner the value of the digit stored at a given location on the medium. For practical reasons, this type of recording is generally limited to binary digits, where the system need recognize only two possible values of digit.

The magnetizable medium, or carrier, may take a variety of physical forms. It may be a homogeneous metallic wire or tape. The magnetic material may be an alloy of appropriate magnetic properties, plated on the surface of a ductile, nonmagnetic wire or tape. One increasingly popular form of carrier is the coated nonmetallic tape. A suspension of a powdered magnetic iron oxide (magnetite) is applied in a thin layer to a paper-tape base.

In another form the powder is dispersed uniformly through a plastic tape. In internal storage applications it is generally desirable to record on the surface of a rigid member such as a drum or cylinder. In such cases the rigid member may be made of a magnetic alloy, the magnetic material may be plated or coated on the surface, or a magnetic tape may be bonded to the surface.

The choice of physical form of the carrier depends mainly on the requirements imposed by the intended application. At one extreme, for example, there is the simple input application, in which information is recorded at a relatively low speed and at a later time read out at a higher speed in synchronism with an internal storage unit. There is no accessibility-speed or searching requirement, as information is read out only in recorded sequence. The record must be capable of permanent storage. A large quantity of information must be stored on a single record. The erasability feature is not essential. For this type of application, the medium is best handled in the flexible form of fine wire or thin tape wound on a pair of spools or reels.

The rapid internal-storage type of application lies at the other extreme. Here the information may have to be recorded, read, and erased at frequent intervals. It must be possible to locate quickly any specified storage cell for transferring information to or from the cell. A reasonable arrangement for this type of service is a rapidly rotating rigid drum with the magnetic medium on the cylindrical surface.

Information may be recorded on a number of parallel tracks located side by side along the length of the drum. In this way, the time required for a selected cell to reach the reading head will not exceed one revolution period of the drum. If several reading heads are spaced around the circumference of each track, this maximum waiting time is correspondingly reduced.

Between these two extremes many variations in physical form are possible. For example, a supplementary storage system of moderate reference speed might use as a medium an endless loop of tape containing two or more parallel tracks of information.

The styles of head most commonly employed in magnetic recording are diagrammed in Fig. 14-2. In Fig. 14-2a the pole pieces are in line with each other and are in contact with either side of the carrier. This magnetizes the medium in a direction

normal to its surface and is known as *perpendicular recording*. In Fig. 14-2*b*, the pole pieces are not in line but are displaced slightly with respect to each other. This arrangement tends to concentrate or focus the flux in the tape into a narrow region, so that wider pole tips of lower reluctance may be used. The resultant magnetization in the medium is still essentially perpendicular, although it may have a small longitudinal component. In Fig. 14-2*c*, the pole pieces are still further displaced to provide *longitudinal recording*, in which the magnetization is
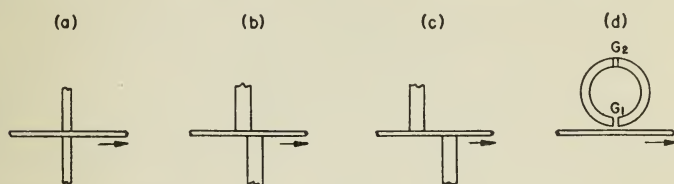
	(a)	(b)	(c)	(d)
MAGNETIZATION	PERPENDICULAR	PERPENDICULAR	LONGITUDINAL	LONGITUDINAL
FORM OF CARRIER	TAPE	TAPE	TAPE, WIRE	TAPE, WIRE, DISC, DRUM
CLEARANCE	CONTACT	CONTACT	CONTACT	CONTACT, NON-CONTACT

Fig. 14-2. Magnetic-recording-head styles.

parallel to the direction of motion. Figure 14-2*d* shows a ring-type head in contact with one side of the carrier, also for longitudinal recording. This is an increasingly popular form of head.

It is evident that round wire is satisfactory only for longitudinal recording. It is not suitable for perpendicular recording because there is no way of returning the wire to its original axial orientation for playback. Tape is satisfactory for either method.

Efficiency of magnetic coupling between the carrier and the head, and also sharpness of resolution, are more easily achieved if the pole pieces are maintained in contact with the carrier. Contact recording, however, is feasible only if the linear speed of the carrier is relatively low, for abrasion at the head due to high-speed passage of the carrier is sufficient to wear it away rapidly. This method is used almost exclusively in the recording of sound, where speeds fall in the range of 6 or 8 inches per second to several feet per second. Digital recording, on the other hand, generally calls for considerably higher linear speeds. Particularly for

rapid internal storage systems, speeds upward of 100 inches per second must be used, and 2,000 inches per second is conceivable. At such high speeds, abrasion of the head and the carrier makes contact recording quite impractical, and it becomes necessary to maintain a small head-to-carrier clearance. Head styles *a, b, c,* and *d* in Fig. 14-2 may all be operated in contact with the carrier, but of these only the ring type of longitudinally magnetizing head *d* is considered suitable for noncontact recording.

For rapid internal storage, then, we may confine our discussion to the case of noncontact ring-shaped heads and longitudinally recorded tracks around the periphery of a cylindrical drum.

Transfer of Data. Transfer of information to and from the drum consists in general of three processes: erase, record, read.

The purposes of erasure are to obliterate previously recorded material and to bring the medium into the initial magnetic state required by the recording process. The required initial state may be either complete demagnetization or longitudinal saturation. (A possible alternative to demagnetization is transverse saturation.)

Erasure by saturation, generally referred to as *d-c erase,* is accomplished by subjecting the moving carrier to a strong constant magnetic field in the longitudinal direction. The d-c erase head may be either a d-c electromagnet or a permanent magnet of suitable design.

Erasure by demagnetization, or *a-c erase,* is accomplished by the familiar method of subjecting each element of the carrier to an alternating magnetic field, the amplitude of which diminishes to zero over a number of cycles. This may be done with an a-c erasing head with a field which tapers off suitably on the trailing edge. If the demagnetization of a given element must be completed within a short distance of travel, then the erasing field must alternate at a high frequency. This may limit the depth of penetration of the erasing field so that previous signals are not entirely removed. Thickness and electrical resistivity of the magnetic coating of the carrier must be chosen to take this effect into account.

The nature of the recording and reproducing, or reading, processes may be illustrated by first considering a very simple type of recorded signal and the signal forms associated with it. Suppose the carrier has previously been demagnetized by a-c

erasure; then if the carrier is held stationary with respect to the recording head while a signal is imprinted by passing a direct current briefly through the head winding, the resulting typical distribution of magnetization J along the carrier will be as shown in Fig. 14-3a.

If the carrier bearing such a signal is moved with constant linear speed past the gap of the reading head, the shape of the curve of time variation of the flux threading the winding will be
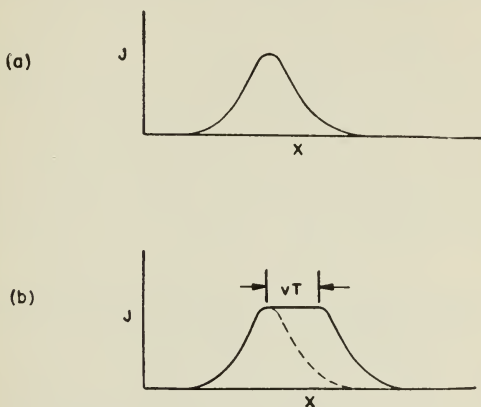
Fig. 14-3.   Magnetization J plotted against linear displacement along the carrier.   In (a) is shown a distribution produced by applying a d-c signal to the winding while the carrier is held stationary.   In (b) is shown the effect which would be produced by applying a signal while the carrier is in motion.

roughly similar to one of the humps shown in Fig. 14-4a.   The electromotive force developed in the winding is equal to the product of the time derivative of this flux times the number of turns, and it will therefore have a shape like that shown in Fig. 14-4b.   This voltage may be further differentiated by means of a simple resistance-capacitance network to give a single sharp pulse with definite polarity of the sort shown in Fig. 14-4c.

Suppose now that the signal is recorded dynamically rather than statically, *i.e.*, with the medium in motion.   If the magnetizing field were able to rise abruptly and remain constant for a brief interval T, then drop sharply to zero, the pattern of Fig. 14-3a might be expected to "stretch," as in Fig. 14-3b. The flux pulse would display a flat top of length equal to the pulse duration times the carrier velocity.   The second derivative

of this signal would contain a pair of pulses, which for most purposes would be undesirable.

In practice, the magnetizing field will rise and fall in a manner determined by the input signal and the circuit and head constants. If the field is caused to fall from a fairly steep portion of the rise curve, the shape of Fig. 14-3a can be approximated to a reasonable degree in the dynamic case. Indeed, if a system is capable of yielding the flat-topped voltage pulse of Fig. 14-3b, then the
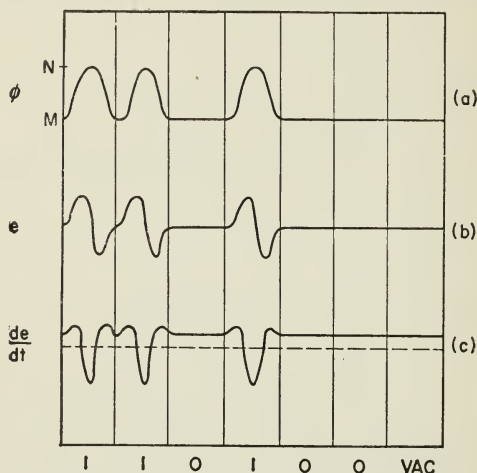
Fig. 14-4.   Effects produced by motion of carrier on which a representation of the digits 110100 has been recorded. The signal shown is called a *two-level return* signal because the flux ϕ varies between *two* levels (one of which represents the digit 1 and the other 0) and because it *returns* to the 0-level between 1-pulses.

information represented by the pulse could evidently have been recorded in a shorter length of track.

There are several ways in which binary digits can be represented by a recorded signal. Three classes of signal are illustrated in Figs. 14-4, 14-5, and 14-6. These represent the two-level return signal, three-level return signal, and the two-level nonreturn signal, which are discussed below. The horizontal coordinate in all of these represents time. The distance between adjacent vertical lines represents the time interval, or cell, allotted to one binary digit. A series of seven cells is shown. These contain the digit sequence 1, 1, 0, 1, 0, 0, with no signal entered in the seventh cell.

Figure 14-4*a* shows the wave form of the flux in the core of the reading head, for the case where the flux remains at a given level *M* to indicate either a 0 or no signal, but shifts momentarily to another level *N* to indicate a 1. To a first approximation this time plot of ϕ resembles the space plot of the magnetization intensity *J* along the carrier. The two levels *M* and *N* between which the flux shifts may correspond to saturation of the carrier in the positive and negative directions, respectively. That is, the carrier is previously d-c erased to saturation in one direction, and a 1 is recorded by pulsing to saturation in the opposite direction. Alternatively, the carrier may be a-c erased, in which case the level *M* corresponds to demagnetized carrier and zero flux, with *N* corresponding again to saturation. (Although this would appear to provide about half the signal voltage, there is the possibility that the noise in a suitably demagnetized medium may be sufficiently lower than the noise for the saturated state to result in a net gain in signal-to-noise ratio.)

The emf induced in the winding of the reading head is shown in Fig. 14-4*b*. If this voltage is differentiated by passing it through a suitable resistance-capacitance coupling network, the derivative voltage will be of the form shown in Fig. 14-4*c*. The portion of this voltage which falls below the threshold indicated by the dotted line is a sharp pulse, the presence or absence of which denotes whether the digit being read is a 1 or a 0.

Since only two states, or levels of magnetization, of the medium are employed, one of the two kinds of binary digit must necessarily be represented in the same manner as the absence of a digit. In Fig. 14-4*a*, 0 is indicated in the same manner as no signal. In some applications this may be considered a serious limitation.

In the *three-level return* system, a definite signal is used to indicate either kind of digit, with absence of signal reserved to denote a vacant cell. When three-level recording is used, the carrier is initially demagnetized by a-c erasure. A 1 is recorded by saturating the carrier in one direction with a pulse, and a 0 is recorded by saturating it in the opposite direction. Definite positive and negative pulses are thus available for indicating 1's and 0's to the output circuits. This method is illustrated in Fig. 14-5.

In the two cases described above, the reading-head flux returns

to the no-signal condition at the end of each unit interval. Figure 14-6 illustrates a method of recording known as the *two-level nonreturn* system, in which the flux shifts from one level to the other only when the new digit is different from the one in the preceding cell.

Figure 14-6a is the time plot of the recording-head flux ϕ, which, again, is similar to the space plot of the magnetization J along the carrier. The recording head is driven by a flip-flop circuit in such a manner that the magnetizing field saturates the
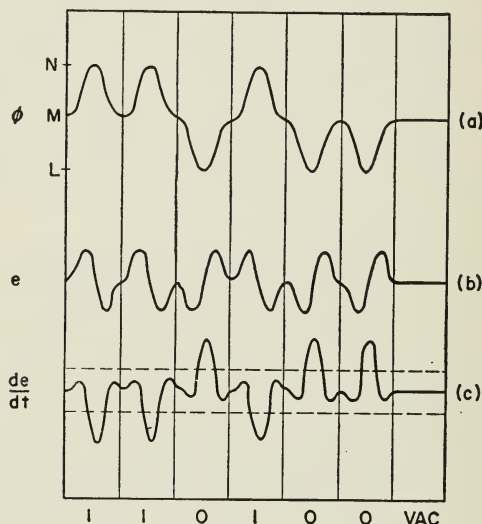
FIG. 14-5.   Three-level return signal.

carrier in the positive direction for one condition of the flip-flop and in the negative direction for the other. A 1 pulse at the beginning of a cell causes the flip-flop to shift to the positive condition, unless it is already there; a 0 pulse, to the negative condition, unless it is already there.

If the time required for the transition, as seen at the reading head, is taken as the minimum practical length of unit interval, then ϕ will vary with time somewhat as shown in Fig. 14-6a. The voltage e_1 induced in the reading head will vary as the time derivative of ϕ, as shown in Fig. 14-6b. A positive pulse denotes the beginning of a series of 1's and a negative pulse the beginning of a series of 0's.

When e_1 swings beyond either the positive or the negative threshold level indicated by the dotted lines, it causes a reading flip-flop to shift to the corresponding positive or negative condition, where it remains until stimulated by a pulse of opposite sign. The voltage e_2 appearing at one of the plates of the flip-flop pair is shown in Fig. 14-6c.
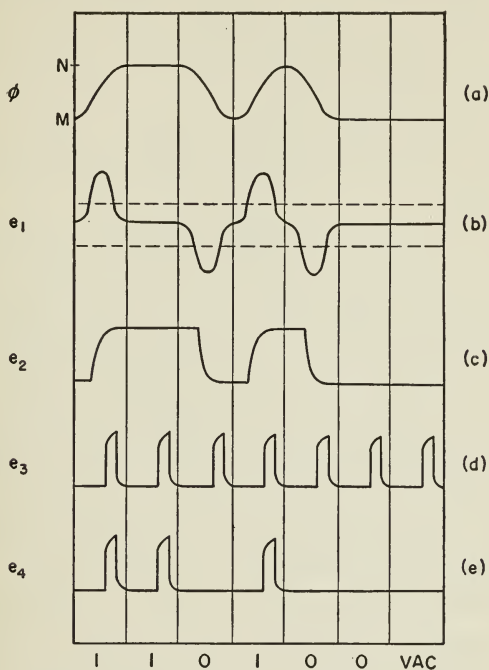
FIG. 14-6. Two-level nonreturn signal.

Figure 14-6d shows a clock pulse e_3 of short duration occurring near the end of each interval. If e_2 and e_3 are applied to two grids of a gating tube, the clock pulses will be transmitted only when e_3 is in the positive state. The gate output voltage e_4 is shown in Fig. 14-6e, where the presence of a pulse denotes a 1 and the absence denotes a 0.

Comparison of Figs. 14-4 and 14-6 reveals some advantages of nonreturn recording. If the cell repetition rate is taken to be the same in both cases, it is evident that the maximum frequency which must be handled by the recording and reading heads in the nonreturn case is of the order of half the correspond-

ing maximum frequency in the return case. Furthermore, if the length of carrier per cell is the same in both cases, the shortest region of unidirectional magnetization is about twice as long as the shortest region in the return case. This implies that if the heads and carrier in a given system are capable of storing a definite maximum number of digits per inch of track and of transferring at a rate of so many digits per second on a return-signal basis, then the same system should store effectively twice as many digits per inch and transfer twice as many digits per second on a nonreturn basis.

Transfer Schemes. The specific manner in which digits are transferred to and from tracks on a storage drum is subject to a degree of choice.

Consider the type of recording designated as *selective mark insertion*, in which the two magnetic states of the carrier are saturation in the positive and negative directions, on either a return or nonreturn basis (Figs. 14-4 and 14-6). Suppose there is no separate erasing head. If it is desired to write a digit into a given cell, the recording head is pulsed in the appropriate direction at the time this cell is passing the recording gap. If synchronization and pulse shape are correct, the new digit appearing in the cell will be independent of what was previously stored in the cell. To read a given cell, the output of the reading head (which may also be the recording head) is gated through at the time the desired cell is passing the gap of this head. Associated with each discrete digit position around the periphery of the drum is a number, the binary representation of which is permanently recorded on the drum in a group of *address* tracks. A given cell is selected by referring to these address numbers by means of coincidence circuits.

This method, which may be termed selective mark insertion, entails the considerable problem of entering a digit into a cell in such a precise manner that the previously stored digit is completely eradicated. The principal advantage of the method over continuous erasure is that the stored information is fixed, or nonvolatile; *i.e.*, failure of power or of the writing system does not destroy information stored on the drum. In fact, magnetic recording in this fashion appears to be *the only means of storage which is both erasable and nonvolatile* and at the present time capable of reasonable speed at reasonable cost.

Another method of transferring data to and from the drum is to apply the *continuous erase-rewrite* recirculating techniques similar to those used in acoustic-delay-line storage systems.[44,45] The carrier travels successively past the reading head, an erasing head which continuously removes previously recorded data, and the recording head. Signals from the reading head are reshaped and fed to the recording head. By means of suitable interposed gates, it is possible to enter a new digit by impressing it on the recording-head circuit at the proper time in place of the recirculating digit coming from the reading head. Recorded digits may be selectively read out by gating the reading-circuit output into the output storage register at the appropriate time. Since any delay in the rewriting circuit, even if introduced deliberately, is apt to be considerably shorter than the transit time of the carrier between the reading and recording heads, the stored digital information is repeatedly transferred to new positions on the track, and the entire pattern keeps shifting about the drum. Location can be handled by a counting track as in the spot insertion technique, but it is evidently necessary to erase and rewrite the counting track in the same manner as the digital tracks. A second control track containing permanently recorded timing pulses may serve as a source of clock pulses for synchronizing the shaping, gating, and rewriting circuits with the drum rotation. Here, as in delay-line storage systems, the stored information is volatile if the erasing is done with permanent magnets.

Track Schemes. Magnetic-drum storage is readily applicable to parallel bus computer systems, with one or more tracks on the drum serving each place or column of a multidigit number. The locating tracks and circuits can be common to all columns. If one track is used per column, then the system can store as many numbers as there are digits per track. However, the location of a given cell for reading or writing may require a waiting time of one whole revolution of the drum. The access time may be reduced in several ways. One method is to use more tracks per column, with a drum of smaller diameter. This may be uneconomical of circuit components. A more economical method of reducing the waiting time is to install several reading heads spaced around the circumference of each track. This, of course, helps only the reading problem and does not reduce the

time to reach a given cell for writing.   Installation of a number of dual-purpose heads around each track appears feasible, but the circuit economy is again questionable.

Magnetic Design Considerations.   In any magnetic storage system the speed of operation and over-all size are largely determined by the physical and magnetic properties of the magnetic medium and the head structures.

The principal properties desirable in a magnetic recording medium and in the head structures for rapid and compact digital storage will be described briefly.   These properties as applied to magnetic recording in general have been discussed in a number of papers.[21,39]

The value of the maximum energy product in a medium (BH_{max}) should be large.   This quantity is indicative of the energy which may be stored in a given sized volume of the medium and is, therefore, an index of the relative signal strength obtainable from the record when wavelengths are short, as required in the practical case.   This statement is explained by the following considerations:

The coercive force H_c must be large to enable a given element of the medium to be magnetized independently of adjacent elements.   High coercive force makes possible improved definition, or resolution of the recorded pattern into smaller elements.

However, if a magnetized element is shortened in the direction of magnetization, its demagnetizing factor becomes larger, and the residual field available for generation of signals in the reading head diminishes.   This variation of residual field with length of magnetized area is less if the ratio of coercive force to remanence, H_c/B_r, is large.   The role of self-demagnetization in the magnetic recording process has been treated by Wooldridge,[39] Camras,[21] and Kornei.[28]

Since the voltage generated in the reading coil is proportional to B_r, both B_r and H_c are as large as possible in the best recording mediums.

If the recording signal contains high-frequency components, the depth of penetration of the signal into the medium is limited by *skin effects*.   If the electrical conductivity of the medium is low, the depth of penetration is not so severely limited.   This is an important advantage of the magnetic oxide coatings over solid metallic mediums.   If the magnetic coating is kept suffi-

ciently thin, there will be no great differences in effective depth, or thickness, of recorded signals at different frequencies, and consequently no great difference in flux available for generation of signal voltage. If a-c erasure is contemplated, these same factors must be recognized in order that the erasing field may be capable of penetrating to the deepest level of recorded signal.

In noncontact longitudinal recording, the fringing flux in the vicinity of the gap of the ring-shaped head penetrates the carrier and records the pattern of information. The geometry of the core and the gap must be carefully chosen so that the fringing field exhibits the desired properties. The longitudinal component of this field must be as large as possible at the inter- section of the equatorial plane of the gap and the surface of the carrier, at the specified head-to-carrier clearance. For resolving power, the flux density must drop off rapidly on either side of the equatorial plane. Both the longitudinal extent of the flux pattern and its maximum intensity must be as insensitive as possible to variations in head-to-carrier clearance.

The magnetic properties of the head cores must be similar in many respects to those of communication-type transformers capable of passing wide bands at high frequencies. The initial permeability should be high in reading heads, because of the very small magnetomotive forces involved. For recording it is important that the maximum permeability be high enough to prevent saturation. The coercive force should be small in both cases to eliminate residual magnetization effects. Alloys such as 4-79 Permalloy, Supermalloy, and Mumetal are representa- tive.[19] Cores must be built up of suitably insulated thin lamina- tions for proper high-frequency response.

The recording-head winding must supply a large number of ampere turns. At the same time, the ratio of inductance to resistance in the circuit must be kept low for rapid response. This indicates the desirability of a relatively small number of turns of large wire, operated at high currents. In the reading head, large output voltage is the critical quantity, and therefore a large number of turns is indicated. In both types of winding, resonant frequencies within the pass band of the system must be avoided.

Ring-shaped heads are frequently made up of two equal halves, for ease of fabrication. The additional air gap (see G_2, dotted in

Fig. 14-2d) appearing in the magnetic circuit tends to reduce the inductance, the flux due to a given magnetomotive force, and the remanent induction.

Description of a Typical System. Conservative magnetic-drum design constants on which ERA has standardized are the following:[23]

Digit spacing: 80 digits per inch, on each track around the drum.

Track spacing: eight tracks per inch along the drum surface, in the direction parallel to the axis.

Surface speed: 1,600 inches per second.

Maximum diameter: 34 inches.

These same specifications were used in the magnetic-drum example of Sec. 14-1-1; reliable storage systems conforming with them have been developed and tested. It is probable that reliable systems for the storage of from a few thousand binary digits to 2 million or more on a single drum can be constructed.

The most suitable recording medium[25] known to the authors at the time of this writing is magnetic iron oxide applied in suspension with a spray gun and protected by a thin layer of hard lacquer. A clearance of 0.002 inch is maintained between the heads and the drum surface.

Typically, numbers are stored on the drum in parallel representation; *i.e.*, every digit of a number appears in a separate track, and all digits of the same number pass under their respective heads simultaneously.

As another example, suppose we wish to store 10,000 numbers ranging from 000,000,000 to 999,999,999. In the binary system we need 30 digits to express the number 999,999,999 because $2^{30} = 1,073,741,824$. Therefore we need the equivalent of 30 parallel tracks around a drum, each track long enough to contain 10,000 digits. (The corresponding considerations pertinent to a binary-coded decimal representation may be derived from the rules set forth in Sec. 13-4-5.) For the sake of ease of reference, as will become clear shortly, it is convenient to design a drum so that the number of digits in each track around it is a power of 2, though this is by no means necessary. In this example, we shall assume a drum containing 8,192 (*i.e.*, 2^{13}) binary digits per track. 10,000 is between 2^{13} and 2^{14}, so we need two groups, each of 30

parallel tracks, around the drum to accommodate the 10,000 numbers.

In order to locate any number in storage, as was explained in Sec. 5-4-2, there must be associated with it an address which specifies uniquely the location in the storage medium where that number may be found. We need 8,192 addresses to identify the 8,192 angular positions around the drum and two addresses to identify the two 30-track groups. Provision must be made on the drum for the addresses of the 8,192 angular positions expressed in binary digits; 13 channels will be required. Finally, one channel is required for timing. So the drum must have 74 tracks, with each of which will be associated one magnetic head.

Such a drum, which is actually capable of storing 16,384 (*i.e.*, 2^{14}) numbers of 30 binary digits each—though our assumed requirement specified only 10,000—will be about 34 inches in diameter and 10 inches long. The storage cycle time will be about 64 milliseconds. About 660 vacuum tubes will be required in the associated reading and writing circuits, or about 1.3 tubes per 1,000 digits stored. This example was chosen to show the economy of magnetic-drum storage for large quantities of data. In terms of tubes per 1,000 digits, the magnetic drum is more economical than the mercury delay line, as may be noted by comparing the figure given above with the figure of 11 tubes per 1,000 digits for a particular mercury delay line machine mentioned in Sec. 14-1-1. Another major advantage of the drum is nonvolatility of the recording. On the other hand, the storage cycle period is relatively long. As was pointed out in Sec. 12-2, the choice of a particular machine must be made with all the problem requirements in mind.

Magnetic drums can be built to store many more or fewer digits than the drum postulated above. Table 14-1, a condensation of a table prepared by ERA,[23] shows a few possible combinations of design parameters and the corresponding cost of each system in tubes per 1,000 digits.

Reliability. The life expectancy of magnetic wire and tape depends upon whether or not the heads are placed in contact with the medium. Home recording equipment built by the Brush Development Co. employs coated paper tape and contact heads, and the tape is specified to be good for 1,000 playings. Plastic tape and plated tape should wear better than coated tape,

and if the heads are not placed in contact with the medium, its life is limited only by the wear resulting from passing it from one reel to another.

<div align="center">TABLE 14-1</div>

Storage capacity		Storage cycle time, milliseconds	Drum dimensions, inches		Number of magnetic heads	Number of track groups	Number of tubes	Tubes per 1,000 digits
Number of binary digits	Corresponding number of 30-digit numbers		Diameter	Length				
61,440	2,048	8	4.3	10	71	2	630	10
122,880	4,096	8	4.3	18	131	4	820	6.7
		16	8.5	10	72	2	640	5.2
245,760	8,192	8	4.3	33	251	8	1,200	4.9
		16	8.5	18	132	4	830	3.4
		32	17	10	73	2	650	2.6
491,520	16,384	16	8.5	33	252	8	1,210	2.5
		32	17	18	133	4	840	1.7
		64	34	10	74	2	660	1.3
983,040	32,768	32	17	33	253	8	1,230	1.3
		64	34	18	134	4	850	0.87
1,966,080	65,536	64	34	33	254	8	1,240	0.63

Magnetic storage mediums are reliable and do not wear rapidly when noncontact heads are used. The record is nonvolatile and insensitive to temperature, humidity, and small amounts of scratching or dust.

When the reading or recording heads are maintained in contact with the medium, there is a tendency to scrape up magnetic material and fill the air gap. The reliability under such circumstances therefore is rather unpredictable.

The signal level at the output or reading head is low and requires voltage amplification in order to obtain a reasonable pulse size. The impedance of a writing head is low, so power amplification is required at the input to the system.

Size and Cost. The capacity of magnetic wire and of coated paper tape is in the order of 10^{10} digits per cubic foot and 3×10^9 digits per cubic foot, respectively. The cost of storing digits in the two mediums is about 2.5×10^6 digits per dollar on wire and 10^7 digits per dollar on tape. Read-write speeds are about 10^4 digits per second for wire and 10^5 digits per second for coated paper tape.

To make supplementary copies of a magnetic wire or tape record, it is feasible to connect a reading unit of a computer to the recording unit and run them off. Magnetic reading equipment is required to inspect the record, but checking equipment should not prove a particularly expensive addition to a large computer. An average of about six tube envelopes is required to read from and write on each channel of digital magnetic record.

Conclusions. Magnetic mediums are particularly suitable in situations where alterations of the recorded data are desirable from time to time, since new data may be recorded over old. It is also convenient to be able to correct errors at the time that records are being made.

A double-ended medium is easier to handle than a loop, and tape is somewhat easier to handle than wire. Tape is more versatile than wire since only longitudinal recording can be used on the latter.

Magnetic materials are also considered an excellent medium for a permanent storage system such as might be used at the input and output of a digital computer. Two possible configurations are of the most interest: double-ended wire or tape and loops of wire or tape. The former is most convenient for data which are to be read in or out only once, and the latter is useful for tables or other data which must be referred to frequently and possibly in random order.

14-2-5. Phosphor Storage. The persistence of light radiation from phosphorescent materials has been considered as the basis for digital storage; O'Neal and Tyler of Eastman Kodak Company[11] and Moore of Harvard[41] have reported their early findings on this subject. A light-spot pattern may be placed on a phosphorescent medium at a given time and place and be detected at a later time and different place. Since phosphorescent storage of light is volatile, the light spots must be regenerated periodically. Therefore, phosphors are considered suitable as cyclic

dynamic storage mediums. They are best used when in the form of a coating upon a rotating disk or drum similar to the magnetic drum described in the preceding section.

When a phosphorescent material is energized by electromagnetic radiation, it continues to radiate energy, usually of a longer wavelength, for some time after the excitation energy has been removed. The decay of energy from a phosphor is exponential in character, and therefore the radiated energy must be utilized before the elapse of a time corresponding to a certain minimum value of radiation. It is also desirable to erase the record completely before recording on the same area again. This is to ensure a uniform degree of discrimination between areas in an excited state and areas in a nonexcited state. A method of erasing is to irradiate the phosphor with infrared. Experiments have shown that thorough erasure by this procedure may be obtained only at the expense of overheating the carrier medium.

Phosphors were developed during the Second World War which have the property that they store a considerable amount of energy when irradiated with ultraviolet light and release visible light when stimulated by infrared. Further irradiation by infrared completely removes the stored energy. The process is cyclic, and the phosphor is not consumed in being returned to its original energy-free state. These phosphors are being studied at the Eastman Kodak Company.

Because the energy level of a phosphor decreases with time, only two states, energized and nonenergized, are readily usable in a phosphor storage device. An important factor in determining the usefulness of phosphors as a storage medium is the resolution which can be obtained. From this factor, the capacity and access time of a drum of reasonable size can be computed. However, this information has not been announced.

The writing equipment for a phosphor storage system consists of a visible or an ultraviolet light source which is capable of being modulated at a high frequency. The light source is projected as a small spot on the periphery of a phosphor-coated disk or drum. Electrical pulses corresponding to digits of information are supplied to modulate the light source as the disk or drum is rotated. Energized spots are then produced sequentially around the periphery to correspond to the pulsations of light.

The reading equipment includes an optical system to project

on a photocell the light from an area on the drum surface. If the phosphor used must first be stimulated by infrared radiation, a source of infrared is located ahead of the phototube.

The reliability of a phosphor storage system is difficult to ascertain. There are engineering problems of light shielding, heat dissipation from the system, and dust elimination from the optical equipment. On the other hand, the spacing between reading and writing heads and the surface of the drum is probably not as critical for the phosphor drum as it is for the magnetic drum.

At the present time, the phosphor storage system does not seem to compete seriously with other better developed devices.

14-2-6. Sonic-delay-line Storage. One of the most adaptable means of high-speed dynamic storage is the sonic delay line.[42-50] Binary digits have been stored in these lines with a digit duration in the order of microseconds. Sonic delay lines have been selected as the storage medium and as part of the arithmetic unit of several large computing machines, including the EDVAC, the BINAC, and the proposed UNIVAC. The first of these machines was built by the Moore School for the U.S. Army Ordnance Department; the second was built by the Eckert-Mauchly Computer Corporation for Northrop Aircraft, Incorporated; and the third is undergoing final development and construction at the Eckert-Mauchly Corporation at the time of this writing.

Mediums Available and Modes of Transmission. Sonic delay lines function by transmitting sound pulses (usually at ultrasonic frequencies) through a chosen medium. The general nature of this medium is of prime importance and deserves first consideration in a description of a sonic storage system. Gaseous mediums seem not particularly useful for this application because of high attenuations and other obvious difficulties. Either solid or liquid mediums seem to offer some promise, although at the moment liquid mediums seem more convenient. This convenience arises from the fact that in liquids only one type of wave motion is important, the compressional mode. Since there can be no elastic shearing force in liquid, the only transverse waves possible are surface waves. On the other hand, solids exhibit pronounced tendencies to transmit transverse waves as well as numerous other modes, and all these modes propagate with different velocities.

The compressional wave in many liquids propagates with a velocity in the neighborhood of 5 feet per millisecond, more or less, depending upon the medium. In particular, in mercury, which is the most popular medium for delay lines for computing devices, the propagation time is just slightly less than 17.5 microseconds per inch at normal room temperature. The delay time for compressional waves in quartz is about 6.3 microseconds per inch. Compressional waves are likely to be propagated one and a half to two times as rapidly as transverse waves, depending upon the elastic properties of the substance. In fused quartz, for example, the velocity of the transverse wave is about 63 per cent of the velocity of the compressional wave, so that the delay time is about 10 microseconds per inch.

Four mediums have been seriously considered for delay lines. These are fused quartz and magnesium alloys for solid delay lines, and mercury and aqueous solutions of ethylene glycol or other solutes for liquid delay lines. The transducer for converting sonic to electrical energy is a piezoelectric crystal, usually quartz.

The choice of the medium to be used is usually based upon one or more of the following factors, the importance of which must be estimated in connection with the particular problem at hand:

1. Required delay.
2. Allowable duration of digits.
3. Attenuation, and minimum allowable signal-to-noise ratio.
4. Discrimination against spurious signals and echoes.
5. Dependence of delay on temperature, pressure, frequency, etc.
6. Cost, availability of materials, ease of manufacturing, etc.
7. Reliability, ruggedness, etc.

General Structure of Delay Lines. The general structure of delay lines is described in the literature.[42,43] The structure usually consists of two quartz-crystal transducers separated by the transmitting medium. The resonant frequency of these transducers may be from 5 to 30 megacycles. One is a transmitting transducer; it receives electrical signals and through piezoelectric contractions transforms them to sonic signals in the medium. The other is a receiving crystal; it vibrates under the influence of the sonic signals from the medium and generates

electrical signals. Both crystals are cut with orientation of anisotropic axes so that the mode of vibration will introduce waves of the type chosen—compressional waves in liquid and, possibly, transverse waves in solids. In order to gain good directional properties of transmission, the crystal faces are plane and several wavelengths in dimension. The crystals are carefully aligned so that the faces of the transmitting and receiving crystals are as nearly parallel as they can be made, and so that the perpendicular from the center of one face strikes the center of the other. In some cases reflectors are used to reduce the over-all length of the delay lines while maintaining a given acoustic length. The amount of delay depends upon the velocity of sound waves through the medium and upon the distance between the transducers.

In the past, for computing purposes, binary digits have been used; a burst of signal persisting a fraction of a microsecond is used to denote the digit 1, and no signal at all is used to denote the digit 0. An alternate scheme might be to use two different frequencies to represent the two digits 0 and 1, or r different frequencies to represent different digits radix r.

Acoustical Match. The choice of mercury as the medium for sound transmission in liquid delay lines is dictated mainly by consideration of acoustical impedance. Where sound is being transmitted across a boundary from one substance into another, reflection will normally take place. The fraction of energy reflected and the fraction transmitted through the substance as a ray strikes the boundary perpendicularly depend upon the degree to which the acoustical impedances of the two substances match. Acoustical impedance is a parameter which depends upon the elastic properties of the substances. The only known liquid with an acoustical impedance comparable to that of quartz is mercury. This implies that the vibrational energy of the transmitting transducer is transferred efficiently to the mercury and that most of the energy transmitted to the receiving transducer is absorbed by this transducer and converted to electrical energy.

The efficiency of transfer of energy is important in several ways. In the first place, there is the obvious desire to transmit strong signals through the medium; this, however, is probably not as important a consideration as two others. The energy

not absorbed by the receiving crystal is reflected back to the transmitting crystal, and it is again reflected by the transmitting crystal to the receiving crystal; thus, if there is not a good impedance match between the transducer and the medium, the possibilities of echoes are enhanced, and signals may be modified seriously by these echoes. Finally, transducers normally are very sharply resonant, and some steps must be taken to decrease the Q of the crystals in order to increase the obtainable bandwidth. The most efficient way of decreasing the Q is by loading the crystal mechanically with the transmitting medium itself; in this way the Q is decreased by virtue of the fact that the energy is transferred at a high efficiency to the medium.

Fused quartz is a popular choice for solid delay lines because of the absence of many objectionable features found in other solids. Its acoustical match to quartz crystals is extremely good. Of the other two types of delay lines already mentioned, the match between magnesium alloys and quartz is so poor, and that between aqueous solutions and quartz is so poor, as to make these delay lines probably useless in computing machine applications. Magnesium alloys, however, have a somewhat better acoustical match with ADP* crystals, and the match between the solid delay lines and the quartz crystals can conceivably be improved by judicious choice of cement between the crystal and the transmission medium.

Temperature Variations. The velocity of sound in mercury at frequencies of a few megacycles will change about 1 part in 3,000 for each degree centigrade change in temperature. This means that if 3,000 binary digits are stored in the delay line, and if no log is kept of the digits emerging, and if the temperature in the delay line is not known to within 1°C, then even though the individual digits emerging can be read perfectly, it is impossible to tell which digit is being read at a given time. This implies practical limitations on the length of mercury delay lines. Other substances show a similar dependence of velocity on temperature. Fused quartz, used with transverse waves, has a temperature dependence less than half that of mercury; this implies that more than twice as many digits can be stored in quartz as in mercury before temperature troubles arise, assuming that the temperature control of each substance is comparable.

* Ammonium dihydrogen phosphate, $NH_4H_2PO_4$.

This assumption is probably not completely justified. Pure water shows a zero dependence of velocity on temperature for temperatures around 74°C. For lower temperatures, the dependence is fairly marked; however, some solutes, in particular alcohol and ethylene glycol, can be used to reduce the temperature at which this zero temperature coefficient occurs. This is the consideration which has led to the adoption of aqueous solutions in sonic delay lines. However, in the case of most computing applications, where the pulse repetition rate is so great that echoes could cause difficulty, this temperature consideration is overruled by the interference consideration.

Attenuation. The theory of transmission of sound through liquids predicts an attenuation of sonic energy proportional to the square of the frequency. This condition is confirmed for frequencies of the order of a megacycle. For mercury, attenuation is about 5 decibels per millisecond at 10 megacycles. It is therefore not a serious consideration at this frequency and reported results of propagation measurements indicate no significant change in velocity or frequency-free absorption coefficient with frequency.[48] In solids, the attenuation is nearly proportional to the frequency. At 10 megacycles the attenuations in fused quartz and in mercury are comparable. At 30 megacycles, fused quartz should still be usable without serious attenuation.

Beam Characteristics of Crystals. For frequencies so high that the wavelength of the transmitted sound is short compared with crystal dimensions, plane crystals radiate narrow, highly directive beams. The width of the beam is governed by the usual equations of transmission of plane waves through windows. In particular, with crystals of reasonable size and with transmitted frequencies of a few megacycles, the beam width is sufficiently narrow to require careful aiming to give efficient transmission of sound through the medium. For transmission through solids, where the wavelength is longer, this directivity is reduced. This is one consideration which is of some importance in deciding between liquid and solid delay lines.

Pulse Duration. The shortest pulse which can be successfully transmitted is a function of several variables. In the first place, it must be realized that the pulse is a modulation of a carrier and that, as such, it contains essential side bands with band-

TABLE 14-2. SUMMARY OF IMPORTANT SONIC PROPERTIES OF DELAY-LINE MEDIUMS

Property	Magnesium alloy (Solid)	Fused quartz (Solid)	Mercury (Liquid)	Alcohol solution* (Liquid)
Match with crystal	Fair (with ADP)	Excellent	Good	Unsatisfactory
Dependence of velocity on temperature	About 1 part in 2,500 per degree centigrade	About 1 part in 9,000 per degree centigrade	About 1 part in 3,000 per degree centigrade	Zero at normal temperature
Dependence of velocity on frequency	Too small to be noted			
Dependence of attenuation on frequency	Unknown	Linear	Quadratic	Quadratic
Minimum binary-digit duration	Less than 5×10^{-7} second			Unknown
Velocity	Transverse: 300 centimeters per millisecond Longitudinal: 550 to 580 centimeters per millisecond	Transverse: 380 centimeters per millisecond Longitudinal: 590 centimeters per millisecond	145 centimeters per millisecond	
Attenuation	10 decibels per millisecond at 10 megacycles	10 decibels per millisecond at 15 megacycles	5 decibels per millisecond at 10 megacycles	

* Other solutions also give zero temperature dependence with poor transducer match.

width depending on the pulse shape and length. The sharply resonant properties of the crystal transducer limit the bandwidth which it will accept; in order to widen this bandwidth, loading of the crystal is necessary. It has been pointed out that this loading is most effectively brought about by a good match between the crystal and the medium. In fused quartz, it seems likely that a $6\frac{1}{2}$-megacycle bandwidth can be attained easily for a 30-megacycle carrier. This might permit pulses spaced 6 to the microsecond. Four pulses per microsecond have already been attained in mercury; at least 2 per microsecond have been
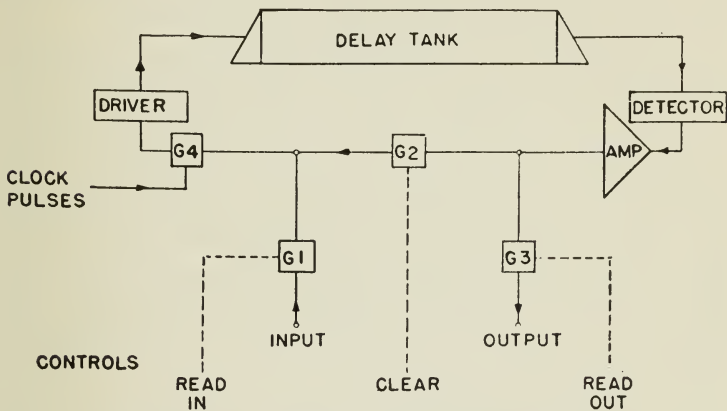
FIG. 14-7.   Acoustic-delay-line storage system.

attained in solid delay lines. Roughly, the same consideration will hold for either frequency-shift or on-off indications of digits.

Complete Storage System. The read-write equipment associated with an acoustic-delay-line type of storage system consists of a driver which excites the input crystal, an amplifier at the output to compensate for the attenuation through the delay line, and a set of gates with which to control the course of the stream of pulses. These components are illustrated in Fig. 14-7. Pulses may be introduced into the tank when G_1 is open and will continue to circulate until G_2 is opened, whereupon the stream will be terminated at G_2. All the circulating pulses or any portion thereof may be read by opening G_3 when the desired pulses are leaving the delay tank.

The components described require about 11 tubes, and the

complexity of the equipment is independent of pulse-stream length. On the other hand, the degree of temperature control required does depend upon the pulse-stream length. In order to keep the pulses emanating from the delay line synchronized with clock pulses, the temperature of the medium must be controlled closely so that the actual delay time remains nearly constant. Another approach to the problem is to mount the delay lines together so that they all have the same temperature within an allowable tolerance and to use one of the delay tanks to control the spacing of the clock pulses.[50] Thus, the delay in a tank would correspond to a constant number of pulses even though the actual delay time varied with temperature.

In the BINAC there are 18 separate delay lines or channels in the same tank of mercury. This is a cylindrical tank with circular cross section about 3 inches in diameter. The capacity of each channel is 1,344 pulse positions, including 960 digits and the algebraic signs and spaces between digit groups. At the effective carrier frequency used in the BINAC (about 12 megacycles) a transducer of the order of $\frac{1}{2}$ inch across becomes such a sharp-beamed acoustic radiator that no shielding in the tank is actually necessary to isolate each beam through the mercury from the others, although shielding has been used as an additional precaution.

The possibility of adding reading heads along the path of a stream of pulses within an acoustic delay line has been considered, but no results are known to have been announced. The attenuation introduced by such units is large if the acoustic match is good; on the other hand, if the acoustic match is poor, there will be echoes in the tank and the true signals will be difficult to distinguish.

Conclusions. To date, more work has been done on mercury delay lines than on other types. The cost of such systems of storage is estimated to be about $1 for 10 binary digits, and the space required is approximately 1 cubic foot for 2×10^4 binary digits.

Mercury delay lines have been used for some time in several applications, including radar equipment and training devices. They are reasonably reliable. Their only serious defect is the variation of velocity with temperature, but this may be handled in a straightforward manner.

14-2-7. Electromagnetic Transmission-line Storage. An obvious way in which to delay and store electrical pulses is to introduce them directly into one end of a transmission line and to read them at a later time as they emerge from the other end.[51-59] Transmission lines may be used as cyclic or noncyclic storage mediums by regenerating and recirculating the pulses at the receiving end of the line or by using the pulses as they emerge after a fixed delay time.

General. The major difficulty with transmission-line storage is that an unreasonable length of line is required to store a large number of digits. This is true because the phase velocity of electrical pulses along ordinary transmission lines is of the order of magnitude of the velocity of light. Since the lines must necessarily be long in order to be of use as storage mediums, considerable power is dissipated in the resistance of the conductors in the form of ohmic, eddy-current, and hysteresis losses. A number of attempts, therefore, have been made to decrease the phase velocity along various types of transmission lines in order to make them more suitable for use as storage mediums.

Several features of the pass band are important if a short pulse is to be passed through a transmission line and remain distinguishable at the far end. If a pulse has a width p, the line must pass at least the fundamental frequency $1/p$. Kallman shows[54] that nonlinear phase shift of the major frequency components, or phase distortion, of the pulse is more detrimental than amplitude attenuation, or frequency distortion. Phase distortion of component frequencies tends to spread out the pulse in time and causes it to be indistinct. In some situations where a lengthened pulse is desired, this property of a delay circuit can be used to advantage, but when delay lines are used for pulse storage, phase distortion must be held within rather close limits in order to get a good pulse-packing factor.

It is desirable that the pass band of a delay line include the pulse-repetition-rate frequencies so that d-c pulses may be passed directly. Waveguides, however, cut off low frequencies but may be used to store a pulsed carrier which lies within the pass band. As a result, the auxiliary equipment necessary for use with waveguide storage is more complicated than for wire lines.

Some thought has been given to the possibility of storing digits

of radix higher than 2 by representing each with a corresponding frequency. This possibility is described in more detail in Sec. 16-2, but it might be noted that a wide-band transmission line would be suitable for the storage of these frequency-shift pulses.

Three familiar forms of transmission line have been considered for computer storage. They are the waveguide, the wire transmission line, and the artificial line with lumped parameters.

Waveguide Storage. If the techniques were available to utilize pulse repetition rates much greater than 10^6 pulses per second, advantage might be taken of the bandwidth of waveguides since it is much greater than that of other types of transmission lines.

A practical limitation on the bandwidth of frequencies which can be sustained in a waveguide is determined by the disturbances to the structure caused by the input and output coupling mechanisms. By careful design of the couplers, it is possible to make the standing-wave ratio caused by the coupling less than 1.15 over a frequency deviation range of 10 to 15 per cent. As an example, if a carrier frequency of 3,000 megacycles is used, and a standing-wave ratio of less than 1.15 must be maintained, a maximum bandwidth of 900 megacycles is possible.

A waveguide with a bandwidth of 900 megacycles may be used to store double-side-band pulsed carriers at a repetition rate in the order of 10^8 pulses per second, which is 100 times as great as for the example cited using a sonic delay line. However, the capacity of the waveguide is only 1 digit for each 10-foot length, and the attenuation problem in a waveguide long enough to store 10^3 digits is severe. In a 1.5- by 3-inch waveguide the attenuation is approximately 1 decibel per 100 feet, so the total attenuation for the storage of 10^3 digits is about 100 decibels. Amplification over a bandwidth of 900 megacycles to restore the signals to their original amplitude is not feasible with conventional tubes, but it is possible with traveling-wave tubes.

Some work has already been done on low-velocity waveguides. For example, in the linear accelerator being built at Stanford University, waves are slowed down in a waveguide by means of a series of irises placed inside the guide. Although equidistant irises are interesting, in that they are effective in slowing down a wave, they cause the guide to act like a band-pass filter so that attenuation and time delay vary considerably with frequency. This results in pulse distortion; hence a low-velocity waveguide

of this type is of limited usefulness as a wide-band storage medium.

Transmission Lines with Distributed Constants. A considerable amount of work has been done on wire transmission lines to reduce phase velocity. As a result, solenoidal lines have been developed for use in equipments including synchroscopes and pulse-forming networks.[51,53,54,58,59]

In a wire transmission line at relatively low frequencies, if losses are neglected, the time required for a pulse to travel along a unit length of line is represented by the equation

$$T = \sqrt{LC}$$

where T = delay time per unit length

L = series inductance per unit length

C = shunt capacity per unit length

The first attempts to increase the delay time resulted in a line with one wire wound in the form of a solenoid inside of or surrounding a conductor of large surface area. The effect of the solenoidal winding is to increase the series inductance of the line owing to the increased self-inductance and the introduction of mutual inductance between turns. The other conductor, formed in a semicylindrical shape, increases the capacity from one turn to the next. In order that it may not present a shorted turn in parallel with the turns of the solenoid, the semicylindrical conductor has an insulated longitudinal strip. The resulting structure is quite complicated to analyze; however, Rudenberg has shown[56] that waves follow the spiral with approximately the velocity of light so that the delay time may be represented by the equation

$$T \approx \frac{l_w}{c}$$

where T = time delay per unit length of line

l_w = length of wire per unit length of line

c = velocity of light

This expression results in a good approximate design formula for a spiral delay line at low frequencies.

When pulses of short duration are transmitted through a delay line of the solenoidal type, it is found that the various frequency components do not reach the receiving end at the same time,

owing to phase distortion.   This effect is caused in part by the high-frequency path through the turn-to-turn shunt capacity along the solenoid and in part by a decrease in effective series inductance as the wavelengths of frequency components approach the dimensions of a turn.   The latter effect results from the fact that when the phase angle between the fields surrounding adjacent turns is reduced, the effective mutual inductance between turns is also reduced.

One of the first low-velocity transmission lines being manufactured is a solenoidal delay line built by Federal Telephone and Radio Corp. and reported by Zimmerman.[59]   It is the RG-64U cable and has a delay of 0.042 microsecond per foot and an attenuation of about 0.14 decibel per foot at 5 megacycles. Therefore, the attenuation is about 3.3 decibels per microsecond delay.   However, this line is not equalized and causes a decrease in delay of approximately 0.2 per cent per megacycle in the 5-megacycle range.   The number of microsecond pulses which can be stored in a line of this type must therefore be limited to about 40 pulses to prevent undue phase distortion.   This cable was designed as a high-impedance cable rather than a delay line; hence presumably greater delay can be achieved if desired.

A spiral delay line reported by Kallmann,[54] in which various shapes and sizes of conductors were placed adjacent to the coil in order to increase its distributed capacity, has a velocity as low as $c/4,300$; however, this is attained at the expense of high attenuation and a narrow bandwidth.   It seems reasonable to expect that velocities of $c/1,000$ might be obtained with a bandwidth of 10 to 20 megacycles.   This corresponds to a delay of 1 microsecond per foot of line and two to four pulses per foot.

Kallmann has reported considerable work to equalize the phase shift in delay lines by means of strips of conductors, which are placed between the spiral winding and the cylindrical conductor and insulated from both.   When these conductors are arranged correctly, they introduce an equalizing effect which reduces the variation of phase velocity with frequency over a band of the order of 5 megacycles.   Other methods of improving the delay characteristic, such as increasing the diameter of the coils and separating the windings, were also investigated.   When the windings are separated, the line might be considered to be a

semilumped-parameter line, similar to the lumped-parameter lines discussed in the next section.

Artificial Transmission Lines with Lumped Parameters.   Delay lines with lumped parameters have been in use for many years as low-pass filters and artificial lines for laboratory experiments. A convenient form is the iterative *pi* or *T* with inductances in the series arms and capacity in the shunt legs.   Actually the distinction is unclear between spiral lines with pieces of foil to increase the shunt capacitance and lines with lumped parameters. In either case, there is mutual inductance between adjacent turns and shunt capacities across various groups of series turns.   A series of simple pi networks does not have a flat pass band, but it has been found that a mutual inductance M between adjacent inductance elements L, adjusted so that $M = 0.12L$, will equalize the attenuation and phase shift over the pass band.   Lumped-parameter lines have been built in the laboratory, but they are rather difficult to duplicate with manufacturing processes. Brillouin[52] estimates that the velocity in a lumped-parameter line may be reduced to $c/10,000$, with a bandwidth of about 5 megacycles.   This corresponds to a delay of about 10 microseconds per foot or a packing factor of about 20 pulses per foot of line.

 Transmission lines with lumped constants or semilumped constants may be built with less attenuation per unit delay time than the continuous-wire lines which have been produced to date. Semilumped lines can be equalized and probably made suitable for bandwidths up to 15 or 20 megacycles.   Because of their attenuation, they too are quite limited with regard to the number of digits which they can store.   One of Kallmann's best lines had a transmission loss of 3 decibels per microsecond at 5 megacycles.   This corresponds to a loss of about 6 decibels per digit at 5 megacycles, so approximately one stage of amplification is required for every two or three digits stored.

Conclusions.   Not enough use has been made of electromagnetic delay lines in computer applications to prove their dependability, but experience with delay lines in other applications indicates that they should be quite dependable.   In systems presently contemplated they are designated for applications in which less than 40 to 100 digits are to be stored or where short adjusting

delays or pulse shaping are desired. A comparison of the various types of delay line discussed above is shown in Table 14-3.

Electromagnetic transmission delay lines are being manufactured commercially by the Federal Telephone and Radio Company, General Electric, and the James Millen Company.

TABLE 14-3

Type of transmission medium	Maximum delay, microseconds per foot	Attenuation, decibels per foot	Pass band, megacycles	Maximum number of pulses per foot	Attenuation, decibels per pulse
1.5- by 3-inch waveguide............	0.001	1.2	450	0.20	6
RG-64U cable.......	0.042	0.14	5	0.08	2
	0.042	0.45	10	0.16	3
Equalized delay lines (estimated).......	1	.....	10–20	2–4	
Lumped-parameter..	10.0	50	5	20	2.5
Semilumped-parameter.............	......	.....	20		

14-2-8. Deflection-type Cathode-ray-tube Storage. One of the most promising means for storing digital information is the acquisition and retention of a charge by an electric capacitor. In connection with television work and other developments, methods have been devised for constructing small storage capacitors fairly cheaply, compactly, and in such a way that they can be charged and discharged rapidly and simply by means of a low-inertia electron beam.[67] Storage tubes are now being developed by the Naval Research Laboratory,[63,64] by the Servomechanisms Laboratory at MIT,[60–62] by the University of Manchester, England,[65] by RCA[68–70] and by the Raytheon Manufacturing Co. There are probably other storage-tube developments in progress, for some work was carried out during the Second World War on tubes which might have had application to computing equipment, and it is presumed that some of this work is continuing.

The National Bureau of Standards is conducting experiments to determine what can be done with standard components,

confining its interests to storage techniques rather than tube development.

Dielectric Storage Principle. Storage in these tubes depends upon electric charges placed at various points on a dielectric plate. As an example of such a device, a tube shaped like an ordinary cathode-ray oscilloscope tube might be visualized. The dielectric plate is placed at the large end of the tube. Means, to be described later, are furnished whereby regions on the surface of this plate may be charged to various potentials under the influence of the cathode ray directed to these regions in conjunction with various fields introduced at these regions by additional elements. Storage of many digits is obtained by charging each of the chosen regions, one for each digit, independently of the rest. If a grid with 32 divisions on each side is chosen as a basic pattern for the charges, then this tube will store 32 squared, or 1,024, independent digits. The radix of these digits will depend upon the number of different distinguishable potentials which can be introduced to the dielectric plate. In the utilization of the principles in a manner to be described, this number is definitely limited to two potentials; for other utilizations of the same principle it might be increased.

Secondary Emission Characteristics. Figure 14-8 shows, as a function of *target* potential, the ratio of the rate at which electrons are leaving a target spot on a dielectric sheet to the rate at which electrons in a cathode ray are bombarding this spot. This ratio takes various values because of capture of electrons by the dielectric sheet, or *target*, and because of secondary emission from the target; these are functions of the field in the tube and of the energy of the impinging electrons, and these in turn are functions of the potential of the target relative to the cathode and other tube elements. Consider, then, a cathode-ray gun bombarding a secondary emitter at the end of a tube shaped like a conventional cathode-ray tube. Near the target end of the tube is inserted some sort of collector element which will catch any electrons which are drifting in the tube after reflection from the target and will also catch the secondary electrons emitted randomly from the target. As ordinate of the curve is plotted the ratio of the number of electrons received by the collector to the number of electrons in the cathode ray (for convenience this is called the *effective secondary-emission ratio*); as abscissa is plotted the

potential of the secondary-emitter target relative to the potential of the cathode.   A qualitative explanation of the shape of this curve is as follows:

When the secondary-emitter target has a potential just higher than the cathode potential, electrons arrive with insufficient energy to knock out an appreciable number of secondary electrons, and the main field between the cathode and the secondary emitter holds the received electrons at the secondary emitter;
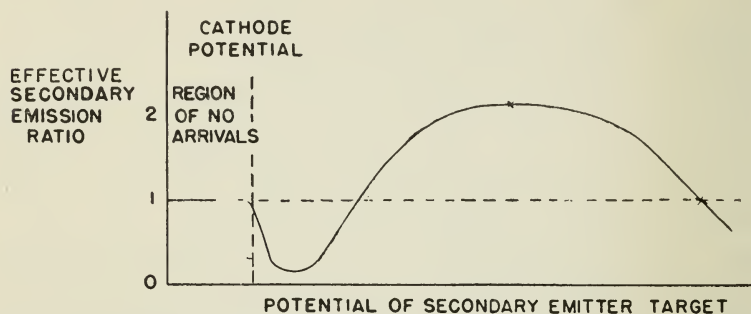
FIG. 14-8.   Effective secondary-emission characteristics.   The parameter plotted as ordinate is the number of electrons collected by the screen per primary electron, *i.e.*, per electron in the cathode beam.   This ratio is plotted as a function of the potential of the target bombarded by the cathode ray.   The number of electrons collected includes both the number due to *secondary emission* alone and the number due to electron optical effects and reflections from the target as well.   If I_2 stands for the secondary-emission current, I_1 for the primary (cathode-beam) current, and I_r for the effects of reflection, the ratio plotted is $(I_2 + I_r)/I_1$.   This should not be confused with the conventional plot of the ratio I_2/I_1, which is the true *secondary-emission ratio.*

thus the electrons are captured, and there is a point on the curve where the ratio is practically zero just above the cathode potential.   When the target is lower in potential than the cathode, the original cathode beam is turned back by the dielectric and captured by the collector, giving an effective secondary-emission ratio of 1.   As the target voltage increases, the energy of the impinging electrons is increased, and they knock out more secondary electrons.   The maximum of the effective secondary-emission ratio may be well above 1; *i.e.*, many more electrons may be collected than are received by the target.   However, as the potential of the target increases further, the energy required

for electrons to escape from it also increases, and the curve may fall off.

By placing a screen close to the front of the target and maintaining it at a constant potential, severe and desirable modification of the high-potential end of this curve may be achieved. Here the assumed structure will be an electron gun (as before) with a comparatively long path to a conductive screen, which serves as a collector for scattered electrons and through which practically all the beam electrons pass. Fairly close behind this screen, the dielectric secondary-emitting target is placed. Figures 14-9 and 14-10 show how the curve is modified by this screen when it is at different potentials. The screen has little effect on the part of the curve corresponding to target potentials much below the screen potential, but it profoundly affects the higher potential parts. In Fig. 14-9, the screen serves to catch rejected electrons in the low-potential part of the curve, but it does not affect the energy with which the primary beam impinges on the target and, therefore, does not affect the secondary-emission characteristics at this point. However, when the potential of the target is higher than the screen potential, the steep potential gradient between the target and the screen increases the energy required for secondary escape from the target, and the secondary-emission ratio is reduced practically to zero, as shown. The same effect is illustrated in Fig. 14-10 with the screen potential so low that the secondary-emission ratio is held at a value below 1 for all target potentials above the cathode potential.

The importance of the unity value for the secondary-emission ratio is that it separates the region in which the potential of the target will increase from the regions in which this potential will decrease when the primary beam impinges on it. When this ratio is 1, the number of electrons leaving the target equals the number arriving, and there will be no change in target potential. When this ratio is greater than 1, the number of electrons leaving the target is greater than the number arriving; under this condition, the target loses negative charge, and its potential increases. When this ratio is less than 1, the number of electrons arriving at the target exceeds the number leaving; the target accumulates negative charge, and its potential decreases. Thus, in Fig. 14-9, the dielectric target retains its potential if this potential is below that of the cathode. (Here minor overshoots of the secondary-

emission characteristics due to variations in electron energies are
ignored; they simply shift the potentials being discussed here
by an insignificant amount, if the dispersion of energies is small.)
For potentials between the cathode potential and a *critical
potential*, at which the curve crosses the unity ratio value in an
upward direction, the potential of the target will drop. Thus,
a target potential anywhere between the cathode potential and
the critical potential will decrease under irradiation by the elec-
tron beam until it reaches the cathode potential. Similarly, if
the potential of the target is between the critical value and that
of the screen, the target will increase to the screen potential
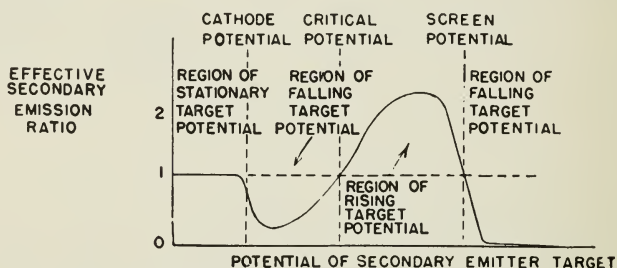
FIG. 14-9.   Ratio of collected electrons to electrons in cathode ray plotted
as a function of target potential.   Screen potential above critical value.

under irradiation; and if the target is above the screen potential,
it will fall to the potential value of the screen. Thus, there are
only two stable target potentials not lower than the cathode
potential for the configuration indicated in Fig. 14-9; one of
these is the cathode potential, and the other is the screen poten-
tial. In Fig. 14-10, the only stable potential not less than the
cathode potential is the cathode potential itself.

The structure of a tube to fulfill these conditions is such that
the target cannot long maintain a potential below that of the
cathode because of leakage. Actually, the screen itself becomes
a cathode under such conditions, and the effective secondary-
emission ratio becomes just larger than 1, tending to move the
target potential toward the cathode potential; however, the
screen is usually an unproductive cathode, so this effect is not
shown on the curves.

Writing Principles. There are many parameters available
for use to represent digital values by this configuration, and
various choices have been made by different workers. The effect

desired, as has been stated already, is the ability to charge discrete regions of the dielectric target independently to potentials chosen from a set of two or more, and to determine which of the set of potentials was placed at each region. Inasmuch as the target is chosen as a nonconductor, the potentials at the different spots will remain essentially unchanged for some while, and it is probable that the regeneration of each spot potential will not require any substantial fraction of operating time of the tube. The parameters available are (1) potential of the dielectric target; (2) screen potential; (3) cathode potential.
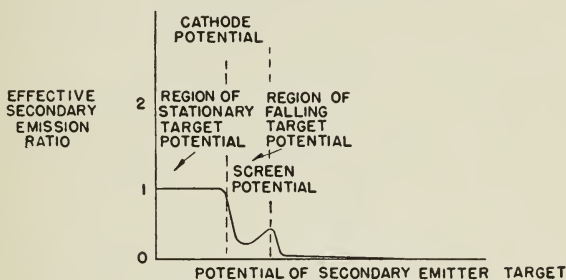
Fig. 14-10. Ratio of collected electrons to electrons in cathode ray, plotted as a function of target potential. Screen potential below critical value.

One method of utilizing the secondary-emission characteristics represented in Fig. 14-9 will now be described. Here the screen potential and the cathode potential are held fast. The effects of these restrictions are represented by the curve in Fig. 14-9, with the cathode potential and the screen potential held constant. The potential of the dielectric is the variable; its two stable points, one at the cathode potential and one at the screen potential, are used to represent the binary digits 0 and 1. One additional element is added to the tube to change the potential of the dielectric; it may be a capacitor plate placed outside the tube so that the secondary-emitting dielectric lies between it and the screen. Once one of the stable potentials is attained by the dielectric, this potential will be held while the target is being irradiated and for a considerable time after irradiation has ceased. Thus, if the target is at the cathode potential, it will remain at the cathode potential under bombardment by the cathode beam. However, the target potential can be increased momentarily by introducing a strong positive charge on the

additional capacitor plate outside the tube. The capacitor action between this element and the screen will introduce a potential at each point of the dielectric between them, and if the potential of the outside plate is made sufficiently high, the potential of the point on which the electron beam is impinging will rise to a value greater than the critical potential shown in Fig. 14-9. If this happens while the secondary emitter is being bombarded, the potential of the secondary emitter will drift upward to the screen potential, and it will remain there if the charge is gradually removed from the outside plate. Thus, a charge can be placed on the outside plate to force the potential of the irradiated dielectric from the value corresponding to the digit 0 to the value corresponding to the digit 1. If the potential had had this value in the beginning, the added potential resulting from the charge on the outside plate would have been compensated for by decreased secondary emission as shown in the curve; in either case, whether 0 or 1 was originally stored, the final result will be that the digit 1 is stored. In the same way, a negative charge introduced to the outside plate may be used to force the dielectric potential to a value below the critical potential, whether the original potential was at the 1 or 0 level. When this charge is gradually removed, the potential of the irradiated dielectric will stabilize at the cathode potential, and the digit 0 will have been stored. There will be no change in the information stored at the points of the dielectric which are not being bombarded by cathode rays during potential changes introduced by the charges placed on the outside plate, for although their potential may be forced above or below the critical potential, there is no change in charge on them during this period, for no electrons impinge on them and no secondary electrons are emitted.

A second scheme for storing and changing data involves variations of screen potential. In this method, the only stable point used on the secondary-emission curve is the point at the screen potential in Fig. 14-9, and care is taken never to permit the potential of the secondary emitter to drop below the critical potential. Several different potentials can be assumed by the secondary emitter corresponding to several different screen potentials. Thus, a screen potential can be chosen and a point

on the dielectric bombarded. This point will attain the chosen screen potential.

A third system involves change of cathode potential with the screen potential held fast. The stable configuration is the curve of Fig. 14-9 with the digits 0 and 1 recorded at the two stable points, where the dielectric assumes the cathode and screen potentials. To charge a spot originally at either of these stable potentials to the higher stable potential, the cathode potential is decreased to a value so low that its critical potential falls below the cathode potential in the stable condition, and then the electron beam is aimed at the point to be charged. Under these conditions the potential of the point to be charged is above the critical potential for this electron beam, and this point takes the potential equal to the screen potential. After sufficient time, the cathode potential may be returned to its stable value, and the potential of the point being bombarded remains at the screen potential. To charge a spot originally at either of the stable potentials to the lower stable potential, the cathode potential is increased to a point below the stable critical potential but so high that the secondary-emission curve takes the form of Fig. 14-10, rather than that of Fig. 14-9. The cathode ray is then aimed at the point in question. The potential of this point will drop to that of the cathode, and if the cathode is gradually returned to its original stable potential, the potential of the points of the dielectric will follow.

In the description above, the principal idealization, which is not entirely valid, lies in assuming that each region of the dielectric can be charged independently of its neighboring regions. The description of the stable potentials under the curves shown would be valid for a target if the whole target were held at a single potential; however, if there are two potentials stored at different places on the target, it is clear that an infinite gradient cannot exist between them and that leakage in the target will produce a boundary region with potentials varying continuously from one of the stable potentials to the other. The existence of this boundary region was ignored in the discussion, but its behavior is of extreme importance. There is further influence of one region upon another by virtue of the small field generated by the charges at the various regions.

Reading Principles. Several reading principles have been suggested for use in electronic static storage tubes. Of these, the most popular seems to be reading by means of the capture of secondary-emission electrons. This can be brought about in several ways, particularly for binary storage. The means chosen for any particular tube depends basically, of course, upon the choice of the method of using the stable points on the secondary-emission curves for storage.

For the first described method of utilizing the secondary-emission characteristics, the screen and cathode potentials are held constant at all times. The two stable potentials of the dielectric, the cathode potential and the screen potential, are used to store the binary digits 0 and 1. An additional capacitor plate is provided, and a charge introduced on it will change the potential of the dielectric. If one point of the dielectric is irradiated and is at a stable potential, the number of electrons collected is exactly equal to the number of electrons furnished by the electron gun. On the other hand, if the potential of the target is decreased slightly by means of a charge introduced on the condenser plate, and if the stable potential of the dielectric before this charge was introduced was the screen potential, the displacement in potential will increase the secondary-emission ratio of the dielectric, as shown in Fig. 14-9. Momentarily the number of electrons collected will increase. Thus, the flow of current to the collector element will increase under such a displacement of potential if the stable potential was at its higher value. For the lower stable potential, however, this displacement of potential will have a negligible effect on the secondary-emission ratio, as shown in Fig. 14-9. Thus, displacement in this fashion gives rise to two signals which differ depending upon the original potential of the target; this gives an adequate reading system, and fairly good signal-to-noise ratio has been obtained.

For the second storage scheme described, where only the higher stable potential is used and the value of this potential is varied by changing the screen potential, a somewhat different reading system may be used. If only two values are stored, one corresponding to screen potential E_0, and one corresponding to screen potential E_1, a reading may be made by setting the screen to a potential between these two values. If this is done, the characteristics of Fig. 14-9 are still valid. At the time of reading,

if the higher potential is stored, the target potential is at a point above the stable point on the curve, and the secondary-emission ratio is less than 1. On the other hand, if the stored potential is the lower stable potential, the secondary-emission ratio is greater than 1. It is to be noted that this reading scheme destroys the information stored, whereas the scheme described earlier leaves the stored information intact.

For the third system of storage described above, in which the cathode potential is varied in order to bring about storage at the two stable potentials shown in Fig. 14-9, reading can also be brought about by variation of the cathode potential. The potential of the cathode is lowered to such an extent that the previous cathode potential (*i.e.*, the lower stable storage potential) falls to approximately the position of the minimum value of the secondary-emission ratio on Fig. 14-9. The screen potential remains unchanged. When this is done, the secondary-emission ratio is decreased almost to zero for the lower stable storage potential, and it remains at approximately 1 for the higher stable storage potential. This permits reading without destruction of the stored information.

Other schemes for reading stored information have been suggested. An example is an ingenious scheme tried by Rajchman in connection with the Selectron, which is described in Sec. 14-2-9. He used anodized aluminum for the dielectric, with a hole drilled in the aluminum at the center of each of the regions into which the grids divided it. An anode element was placed behind this dielectric screen, and the reading principle was based on the grid action of the dielectric screen itself. With the beam directed at a region of the screen, the screen functioned as a blocking grid when at its lower stable potential in this region and as a passing grid at its higher stable potential. This scheme was found to operate satisfactorily, with suitable choice of parameters, but the use of discrete metallic elements has led to simpler and more reliable operation, as will be explained later (see page 371).

A number of workers have suggested that continuous radiation on the dielectric be provided to hold it at a stable potential, thereby overcoming leakage difficulties. Where stability is attained in this fashion, patterns have been stored for many hours without noticeable deterioration.

MIT Storage Tube. A number of electrostatic storage tubes having storage capacities from 256 to 1,024 digits have been constructed and used successfully at the MIT Servomechanisms Laboratory. A bank of such tubes will be used as the high-speed storage system in the Whirlwind I computer.

The MIT tube[60–62] has two cathode-ray guns, a high-velocity gun for use in recording and reading data, and a low-velocity gun to supply a diffused electron flood for the purpose of maintaining the charge pattern, or array of stored data.

The storage surface is a mosaic of conducting beryllium, sputtered on a dielectric surface. There is a collector screen between the guns and the storage surface. A signal plate beyond the beryllium storage surface is insulated from it by the dielectric sheet.

The potentials of the guns and the screen are not varied. The signal plate is switched between screen potential (ground) and a positive potential of about 100 volts.

Writing Principle. The two stable conditions corresponding to the digits 0 and 1 are zero charge and negative charge on the storage surface. The recording of either one or the other of these stable conditions upon a particular area on the surface is accomplished by pulsing the high-velocity gun with the signal plate at either ground potential or at positive potential. Regardless of the potential of the signal plate, the irradiated area on the storage surface will assume the potential of the collector screen (ground). However, if the signal plate is at positive potential at the instant of irradiation, then, when the signal plate potential is returned to ground, the irradiated area becomes negative (representing the digit 1).

Holding. During the time when the high-velocity gun is not being pulsed, the low-velocity holding gun produces a continuous stream of electrons. If a spot of zero potential on the storage surface (corresponding to the digit 0) becomes slightly positive, it will collect electrons from the holding-gun beam until it returns to collector potential. If the same spot becomes slightly negative, electrons will be lost from the surface until it again reaches the collector potential. Hence, the spot at zero potential is in a stable condition.

A negative spot (corresponding to the digit 1) is stable also. Since it tends to be at the same potential as the holding-gun

cathode, electrons will ordinarily be decelerated as they approach and will arrive with low or zero velocity. If a negative spot becomes slightly more positive than the holding-gun cathode, the storage surface will collect electrons faster than they are being emitted from it and will become more negative.

On the other hand, if the negative spot becomes more negative than the holding-gun cathode, primary electrons will be reflected. Because of leakage, the negative spot will gradually become more positive.

Reading. Reading is accomplished by examining the change in the voltage of the signal plate when a particular spot on the storage surface is irradiated by the high-velocity gun. Irradiation of a spot at ground potential (representing a 0) produces no change and no output. When a negative spot is irradiated, secondary electrons are given off faster than primary electrons are striking it; the spot becomes more positive. This effect produces a positive output signal on the signal plate which is capacitively coupled to the storage surface.

The NRL Tube. The storage tube developed by Dr. Haeff at Naval Research Laboratory[63,64] resembles a conventional cathode-ray tube and employs the screen-voltage writing technique described above. To place a negative charge upon the target, a large positive pulse is applied to a control grid located near the electron gun at the same time a negative pulse is applied to the collector screen. The positive voltage on the control grid increases the intensity of the beam, and the negative voltage on the screen decreases the velocity of the stream. Thus, an intense beam of slow electrons deposits a negative charge on the dielectric target, reducing its potential to that of the cathode. In order to place a positive charge upon the target, positive pulses are simultaneously applied to both control grid and collector screen. A high-intensity, high-velocity electron beam results and causes secondary emission from the point on the target being bombarded. The net flow of electrons away from the target brings its potential up to that of the screen, and it tends to stay there after the pulses have been removed.

A diffused holding beam is irradiated over the target to maintain the established charge pattern. The potential of the holding-beam gun is such that positively charged spots are above the critical potential and negatively charged spots are below the

critical potential. The effect of the electron spray from the holding gun is to hold the negative spots at cathode potential and the positive spots at screen potential. In order to keep the velocity of this beam nearly constant, the pulses applied to the collector screen are also applied to the gun from which the holding beam emanates. Thus, the relative potential remains constant.

In a present model of this tube, the same gun is used for reading and writing. In order to minimize noise caused by modulating pulses and switching operations during the reading process, the reading beam is modulated at an intermediate frequency. Since the output signal from the collector screen is modulated, it can be passed through a tuned amplifier which rejects the low frequencies, including the noise. The output of the intermediate-frequency amplifier is rectified and introduced to a video amplifier where 1's and 0's appear as positive pulses of different amplitudes. These may be balanced with negative pulses of such an amplitude that a negative pulse will be received for a negative spot on the target, and a positive pulse will be received when a positive spot is probed.

University of Manchester Tube. Another deflection-type electrostatic storage tube is being developed at the University of Manchester, England, and has been reported by Williams and Kilburn.[65] Their work has been done on commercial cathode-ray tubes, which are modified by the addition to the large end of the envelope of a metallic plate which serves as a signal plate. The glass and the phosphorescent screen coating become the dielectric target. In the tube described, the deflection plates, the internal conductive coating, and the first and third anodes are all at ground potential; the cathode, the grid, and the focus electrode are given a negative potential in the order of $-2,000$ volts. Before an electron beam is initiated, the inside of the target tends to assume ground potential because of leakage to the grounded coatings on the inside of the envelope. Since there is no collector screen, drifting electrons are caught by the anodes or the deflecting structure, and the secondary-emission characteristics are similar to those illustrated in Fig. 14-8; the cathode potential is near the middle of the range in which the secondary-emission ratio is greater than 1. Thus when a beam is initiated and aimed at a spot on the screen, there is a net

flow of electrons away from the spot for a short period of time until equilibrium conditions are reached again.

In order to understand the method employed for representing binary digits on the target of this tube, it is necessary to consider the charge distribution and the charge-time history in more detail than was required for the tubes previously described. As electrons leave the spot bombarded, the spot becomes positive in potential so that the secondary electrons emitted from the target are attracted back to it by a field which is increasing for a
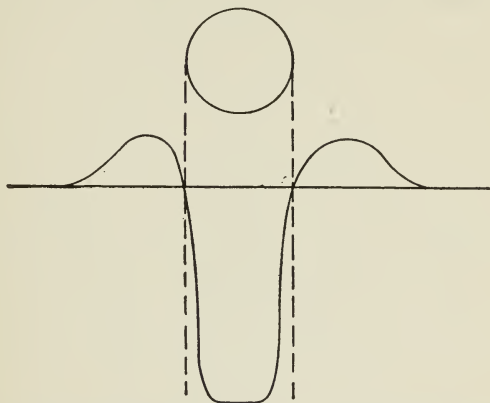
Fig. 14-11. An electrostatic well. The curve illustrates voltage distribution along a line in the plane of the tube face. The downward direction is positive.

short time after the primary beam is initiated. Thus, more and more electrons return to the target, and equilibrium is reached when the number escaping the target is equal to the number in the primary stream. This is the condition of unity secondary-emission ratio, described earlier. The electrons which leave the target with higher velocities have time to acquire a velocity parallel to the face of the target so that some of those which return lie around the rim of the original spot area. Thus the potential along a line through the center of the spot which has been bombarded is illustrated by Fig. 14-11. Such a charge distribution has been called a *well*. Experiments have shown that if the bombardment does not continue for more than 400 microseconds, the target is not affected at distances greater than a spot diameter from the center of the spot.

When a spot on a target near the rim of an existing well is bombarded, there is a tendency for secondary electrons to fill the original well.   The original well is not filled as fast as the second well is formed, however; hence after a short period of time the potential along a line through the centers of the holes is as shown in Fig. 14-12.

Two spots placed side by side on a target will be mutually independent if their spacing is more than a critical value, about 1.33 diameters, but they will affect each other if their spacing is less than this critical distance.   Use is made of this phenomenon in order to make two kinds of distinguishing marks on the target.
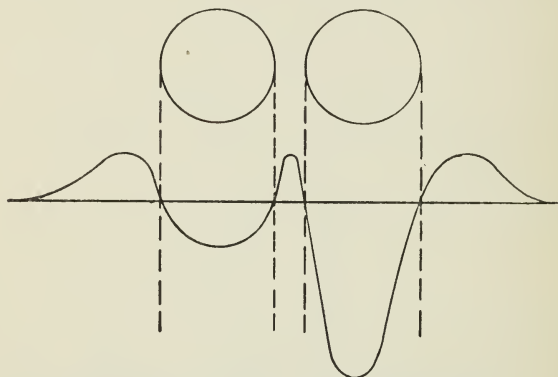
Fig. 14-12.   Double electrostatic well.

One mark is a full well, which is made by bombarding a single spot for a short time.   The other mark is a well which is modified by bombarding another spot in the near vicinity of the storage spot.   These marks will be maintained for a few tenths of a second, depending on surface leakage.   Therefore, it is necessary to read and rewrite the marks continually.   Several methods have been tested for reading and differentiating between marks of this type.   Reading is based on the fact that when an electron beam is aimed at a full well and initiated, the presence of the negative cloud of electrons in the primary beam and in the secondary-electron stream induces an instantaneous flow of current to the metallic plate on the outside of the tube.   This plate is connected to the input of an amplifier.   An output wave form is obtained which is dependent upon the transient response of the amplifier to the negative pulse received at the input.

When a beam is turned on a storage spot where the well has been modified by the presence of a second well nearby, there is a redistribution of charge on the face of the tube, which induces a current in the input circuit of the amplifier, this current being of opposite polarity to that of the current induced by the electron cloud. If parameters are adjusted properly, the pulse of positive polarity induced by the shifting charge may be made large enough to counterbalance the negative pulse induced by the electron cloud. Thus the type of mark stored determines the polarity of pulse received at the plate when the spot is bombarded by an electron beam.

A regenerative system using a gate and a pulse generator can be used to renew each mark periodically after it is read. It is proposed to scan this tube in a manner similar to that used with a television tube, so that regeneration can be accomplished. In addition, a modification of this sequence can be made in order to read or write a digit or line of digits immediately after orders to do so are generated by the computer. If scanning is from left to right, the left-hand is modified by the right-hand well when the double mark is written. When the mark is regenerated, the right-hand well is modified by the left-hand one momentarily, but finally as the beam of electrons moves to the second well, the left-hand well is again modified by the right-hand well.

A space approximately two by three spot diameters (d) on the target is required to store each mark, so if $d = 1$ millimeter, a 6-inch tube will store about 600 digits. Improving the focus of the beam increases the storage capacity as the inverse square of the spot diameter.

Other Electrostatic Storage Devices. Zworykin and Morton[67] have described a type of storage tube in which one electron beam writes a potential-variation pattern on one side of a two-sided mosaic while another electron beam scans the other side to release electrons, which produce a variable-current signal in a collector anode. Such a tube might be modified to include an electron multiplier for amplifying the collected current. A storage tube called the *Krawinkel*[66] was devised in Germany during the war. Apparently it uses an Iconoscope type of mosaic for the storage medium.

That the ultimate storage potentialities of such devices may be far in excess of 4,096 digits per tube is suggested by the present

commercial-television picture standards. The RMA standard picture contains some 360,000 elements (525 lines, 3 by 4 aspect ratio), and satisfactory experimental pictures have had over a million elements. Although the elements are by no means completely resolved, these figures indicate what may be expected in future refinements of electronic image and storage devices.

14-2-9. The Selectron. An electrostatic storage tube called the *Selectron*, the latest model of which has been designated SE256, is at the time of this writing in an advanced stage of development at the RCA Laboratories in Princeton, N.J.[63-65] This tube has a storage capacity of 256 binary digits. It is 3 inches in diameter, 7 inches long, and utilizes a 40-lead stem. Access time to any element is 20 microseconds and is independent of all previous accesses to other elements.

A number of these tubes have been made already for experimental use; uniformity of selection and control characteristics have been observed in all the tubes.

Principle of Operation. The Selectron utilizes the fact that an insulated secondary-electron emitter can be made to "float" at either of two stable potentials under certain conditions of primary-electron bombardment and secondary-electron collection. Secondary electrons are received by a collector electrode placed near the insulated secondary emitter in a manner similar to the collector screen of the deflection-type storage tubes. While a surface element is being irradiated, its potential is locked at one of these stable values.

As originally conceived, the Selectron consisted essentially of a cylindrical dielectric target with the cathode at its axis and the electron stream flowing radially from the cathode. There was a grid, made of 64 vertical metallic bars and 65 circular horizontal metallic rings, just inside the target. The 64², or 4,096, windows in this cylindrical grid divided the target into 4,096 discrete cells. The size of the windows was such that electrons could pass through a window only when the four metallic conductors bounding that window were *all* more positive than some particular voltage.

The physical form of the tube has been altered materially since it was first conceived. As a result of Rajchman's experimental work, directed primarily toward improvements in reliability, he has adopted a flat target (in place of the original

cylindrical form) and eight elongated cathodes. In the SE256 there are two parallel target planes on opposite sides of the plane of the cathodes. A detailed description of the engineering features of this tube has been published.[65]

Regardless of the physical form, the operation of the Selectron depends upon the effects obtained by energizing the four conductors bounding a particular window. One of its greatest advantages is derived from the fact that the means for directing electrons into a selected cell does not depend upon the accurate reproduction of deflection voltages.

In the tube's quiescent operating state all the selecting bars are at the potential of the cathodes (zero). Electrons emitted from the cathodes are focused into 256 beams by the combined action of these bars (at zero potential) and a collector plate which is held at some positive potential such as 180 volts. These beams are focused on 256 nickel-plated steel eyelets, which are floating electrically and are fixed in mica sheets. There are only two stable voltage conditions at which an eyelet can be held, because of the secondary-emission phenomena. Every one of the 256 eyelets will remain in one or the other of these states so long as the tube is in its quiescent operating state.

To write into a particular element, current is interrupted everywhere except to it. Then a signal is applied to a perforated metal plate, called the *writing plate*, which is located (with respect to the cathodes) beyond the mica plate in which the eyelets are set. This plate is capacitively coupled to all the eyelets. The voltage on the particular eyelet will be left at either one or the other of its two stable levels depending upon the presence or absence of an additional pulse on the selecting bars. This pulse cuts off the current to the selected eyelet during the decay of a pulse applied to the writing plate. (This pulse is positive for either polarity writing.) The eyelet will be left at the higher level if only a single positive pulse is applied and at the lower level if the initial positive pulse is followed by an additional pulse to one or more of the four bars bounding its particular window.

Reading is effected by obtaining an indication as to whether current is passing through the central hole in the eyelet. When an eyelet is positive, some of the electrons directed at it go through the hole. When it is negative, electrons are repelled and do not go through it. An auxiliary electrode, called the *reading*

plate, normally biased negatively, cuts the potentially large current from all or a large proportion of the eyelets to the output electrode. To read, an element is selected, then the reading electrode is pulsed positively. There is, or there is not, a resulting pulse of current in the output electrode, depending on the potential of the interrogated eyelet. The output electrode is a set of wires well shielded from capacity pickup by a Faraday cage, which also serves to support a fluorescent screen for monitoring purposes.

The Selectron Aiming Principle. It is of interest to note that the aiming method described is not the only one which might be applied to a system of bars and windows. Therefore, a more general discussion of the Selectron aiming principle is included here, as an appendix to the foregoing material. Consider a circular picket fence with 2^n pickets (n integral and at least 2) and 2^n spaces between pickets. Let these pickets be numbered in their cyclic order from 0 to $2^n - 1$ with a number i. Assign to the space following a picket the same number assigned to a picket.

Now choose two positive integers p and q, such that

$$p + q = n$$

Assign to each even-numbered picket a number j (varying from picket to picket) with $0 \leq j \leq 2^p - 1$; to each odd-numbered picket, assign a number k with $0 \leq k \leq 2^q - 1$. It is clear that for any given space between two pickets there are 2^p ways the adjacent even-numbered picket can be labeled (because there are 2^p different numbers j); and 2^q ways the adjacent odd-numbered picket can be labeled (because there are 2^q different numbers k). Then, for any space there are

$$2^p \times 2^q = 2^{p+q} = 2^n$$

ways the two adjacent pickets can be labeled. This is exactly the number of spaces available to be labeled. The Selectron principle of aiming is based on a labeling scheme in which every possible labeling of pairs of successive pickets occurs exactly once.

One way in which this can be done is as follows: First, let us assign the numbers 0 through $2^p - 1$ cyclically to the pickets

with even numbers; *i.e.*, assign numbers j to the pickets with even numbers i as shown in Table 14-4.

TABLE 14-4

i	j
0	0
2	1
4	2
.	.
.	.
.	.
$2 \times 2^p - 2$	$2^p - 1$
2×2^p	0
$2 \times 2^p + 2$	1
.	.
.	.
.	.
$4 \times 2^p - 2$	$2^p - 1$
4×2^p	0
.	.
.	.
.	.
$6 \times 2^p - 2$	$2^p - 1$
6×2^p	0
.	.
.	.
.	.
$2^n - 2 = 2^q \times 2^p - 2$	$2^p - 1$

It is seen that for each j cycle the number i increases by $2 \times 2^p = 2^{p+1}$. Then, the number of j cycles must be

$$\frac{2^n}{2^{p+1}} = 2^{q-1}$$

To the odd-numbered pickets appearing in each of these cycles let us assign two values of k alternately, using a different pair of values in each j cycle. In the first j cycle the values $k = 0$ and $k = 1$ are assigned alternately; in the second the values $k = 2$ and $k = 3$, etc., until the (2^{q-1})th, which has the values $k = 2^q - 2$ and $k = 2^q - 1$ assigned alternately. The whole assignment is given in Table 14-5, a being the j-cycle number, from 0 to $2^{q-1} - 1$.

It is clear that this scheme has the desired property. Indeed, any chosen value of k occurs in exactly one j cycle, and at some point in this cycle it is adjacent to each j value. Thus, *given*

any admissible j value and any admissible k value, there is one space with these labels on its pickets.

TABLE 14-5

a	i	j	k
0	0	0	—
0	1	—	0
0	2	1	—
0	3	—	1
0	4	2	—
0	5	—	0
0	6	3	—
0	7	—	1
.	.	.	.
.	.	.	.
.	.	.	.
0	$2 \times 2^p - 2$	$2^p - 1$	—
0	$2 \times 2^p - 1$	—	1
1	2×2^p	0	—
1	$2 \times 2^p + 1$	—	2
1	$2 \times 2^p + 2$	1	—
1	$2 \times 2^p + 3$	—	3
1	$2 \times 2^p + 4$	2	—
1	$2 \times 2^p + 5$	—	2
.	.	.	.
.	.	.	.
.	.	.	.
$2^{q-1} - 1$	$2^n - 4$	$2^p - 2$	—
$2^{q-1} - 1$	$2^n - 3$	—	$2^q - 2$
$2^{q-1} - 1$	$2^n - 2$	$2^p - 1$	—
$2^{q-1} - 1$	$2^n - 1$	—	$2^q - 1$

14-3. Transfer in Space

It is evident that data must be transferred from one unit in a computer to another while a problem is being solved. In electronic digital computers, data are transformed into electrical pulses which are transferred along busses and cables and are initiated by the action of buffers and gates. All of these four components are discussed in previous sections and are considered here only to the extent that the problem of transferring data in space affects the requirements which must be met in their design.

14-3-1. General Requirements. The operational require-
ments for transferring pulses in a computer are:
1. Freedom from transfer failure due to:

   a. Undesired inductively or capacitively coupled signals.
   b. Cross modulation.
   c. Spurious noise.
   d. Variations in components caused by aging, atmospheric
   . conditions, or from any other predictable cause.
2. Limited phase, frequency, and amplitude distortion.
3. A minimum of time delay.
4. Limited reflection of pulses caused by discontinuities and
   mismatches.

In addition, the devices employed should be economical and
easy to service.

The designers of the ENIAC have taken adequate precautions
to observe requirement 1. The average signal level is between
25 and 50 volts. A normal variation in component characteris-
tics is not harmful. To guard against high cross-coupling volt-
ages, the impedance levels of the transmission lines have been
kept low, and connections between units of the machine have
been made through shielded microphone cables. The precau-
tions to be observed in tube and wiring layout inside units are
similar to those of video amplifier design.[67,71,72,75,76] Compo-
nents should be arranged to minimize lead length, and critical
leads should be spaced apart or transposed frequently. Imped-
ance levels should be kept as low as possible. In the ENIAC,
circuits have been designed so that wide variations in parameters,
and particularly vacuum-tube variations, do not interfere with
the positive, dependable operation of the unit.

Requirements 2 and 3 are familiar to designers of radio-fre-
quency amplifiers and transmission networks and are, essentially,
the problems discussed in Sec. 14-2-7 regarding time delay and
distortion in transmission lines. Distortion in the transfer
circuits must be limited sufficiently so that any combination of
pulses is perfectly distinguishable at the receiving unit. Time
delay, on the other hand, is somewhat less objectionable than
distortion since it can be compensated by adding equivalent
delays in other circuits in the unit. However, in the interest of
simplicity, it is well to eliminate the necessity for corrective-

delay circuits wherever possible. In circuits which require a high impedance level, shunt capacity should be minimized, and if circuits are frequency-selective within the frequency range of interest, frequency compensation may be required. A familiar example is that of inductive compensation in the plate load of a vacuum-tube amplifier, as illustrated in Fig. 14-13. The high-frequency components of a pulse are partially shunted by capacitor C, which represents the effective sum of all the capacities in parallel with the plate resistor. A properly chosen value

$$C = C_{OUT} + C_{STRAY} + C_{IN}$$
$$L = .42 \, R^2 C$$
$R_D =$ DAMPING RESISTANCE TO PREVENT OVERSHOOT
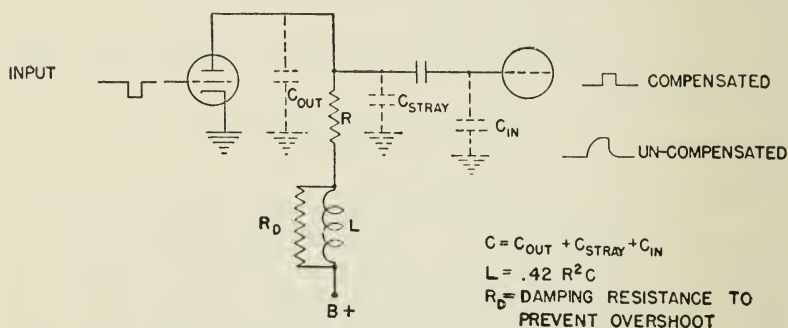
FIG. 14-13.   Inductive compensation of plate capacitance.

of inductance L will compensate for some of this shunting by holding the plate load impedance to a constant value for a greater portion of the frequency range. More sophisticated forms of phase and frequency compensation are discussed in the literature.[71]

Another source of frequency discrimination is in the interstage coupling capacitors, which affect the low-frequency response of the units. To overcome this difficulty, many of the circuits in the ENIAC are direct-coupled. Direct coupling introduces another difficulty, however. In the ENIAC, largely because of the direct coupling between stages, there are 78 different d-c voltages which must be held constant within a few per cent. In some video amplifiers, d-c restoring circuits are employed to avoid the selective characteristics of a-c coupling without the necessary complications of d-c coupling. Undoubtedly these circuits will find use in future high-speed computers.

In the ENIAC it was not necessary to terminate transmission cables in their characteristic impedances in order to reduce reflections, but as computers using shorter pulses are designed,

it becomes necessary to satisfy this requirement. Moreover, at the time of this writing, it has been found desirable to terminate the lines in the ENIAC in impedance-matching load boxes for the sake of greater reliability, although the machine will operate without them.

14-3-2. Gates and Buffers. Gates and buffers have been discussed in Chap. 4, and as such the multicontrol grid, the parallel load, and the crystal gates were described.[73] In applica-
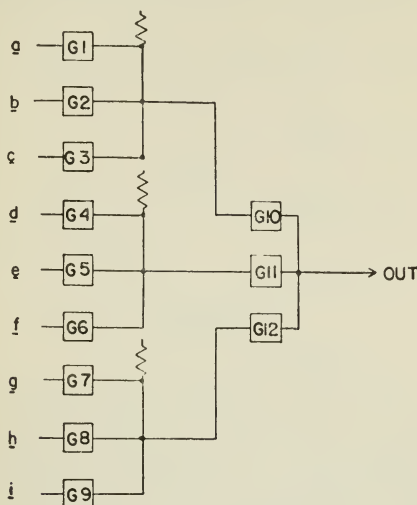
Fig. 14-14.   Multiple-gate circuit.

tions where it is necessary to use gating or buffer circuits having a multiplicity of inputs, the accumulated shunt capacity at the plate circuits of a number of triodes connected in parallel limits the leading-edge rise time of pulses to be passed. A circuit known as the *christmas-tree circuit* and illustrated in Fig. 14-14 is useful in reducing this output capacity. A block diagram of the circuit resembles a tournament play-off diagram. For example, circuit a consists of two-input gates for a minimum of capacity loading. A larger number of elements can be used in parallel at the expense of circuit response and/or gain. A similar circuit such as Fig. 14-15 might also be devised using multiple-grid tubes.

A miniature pentode tube which is suitable for use as a gate is the 6AS6 manufactured by Western Electric. The control grid

and the suppressor grid are used as the independent control elements. Some combinations of electrode voltages yield a suppressor-grid transconductance which is almost half that of the control grid and nearly equal cutoff voltages for the two grids. In most pentodes, the control-grid transconductance is many times the suppressor-grid transconductance. It does not seem feasible to increase the number of control grids beyond two, because in order to make grids of nearly equal sensitivity, it is necessary to decrease the sensitivity of all grids to that of the one which is spaced farthest from the cathode. The result is a tube of very low transconductance for each control grid.
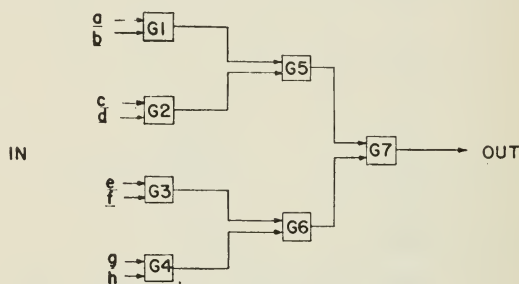
Fig. 14-15.   Multiple-gate circuit.

A study has been made at Harvard[74] of multiple-grid tubes for use as multiple-input gates. A fundamental difference in requirements has been established between gates which are to handle large voltage pulses and those which receive small voltage pulses. For the former case, the grid is driven from anode cutoff to anode saturation by the pulses. Sharp cutoff with high average transconductance is desired in order that output pulses of steep wave front may be obtained. In multiple gates which receive small input pulses, the *ratio* between on and off transconductance should be large. When a large number of gate tubes are used with a common plate load resistor, there is a tendency for the spurious noise signals on grids biased to cutoff to add up to an amount comparable to an output pulse. For this reason a high ratio between on and off transconductance is desired. A diode gate is shown in Fig. 14-16 which is a slight variation of the one illustrated in Fig. 4-2b and described on page 41. Resistor R_1 is much greater than R_2, with the result that input pulses are effectively shorted through one diode or

the other, depending upon pulse polarity. Application to R_2 of a gate voltage of the polarity indicated places a bias on the two diodes so that input pulses are not shunted and appear at the output. Actually a small voltage gets through this type of gate whether it is opened or not, and the ratio of voltage in the on and off conditions is limited to the ratio of backward to forward resistance through the crystal rectifiers.

High-speed Coincidence Devices. A variation of the Rossi circuit described in Sec. 4-3-3 and illustrated in Fig. 4-1*b* on
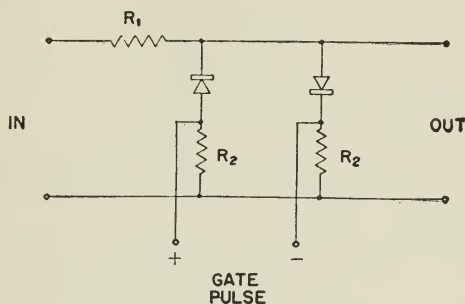
FIG. 14-16. Diode gate.

page 38 has been developed in connection with nuclear experimentation.[77] This reference is noteworthy in this text because it describes by far the fastest use of a coincidence or *and* circuit known to the authors. The resolving power reported is of the order of 10^{-8} to 10^{-9} seconds. In this device a tube equivalent to the Type 6AC7, having high mutual conductance and low capacity, is used.

A method of detecting coincidences of very short duration and observing them on a cathode-ray screen also has been devised and reported.[79]

REFERENCES

Relay Storage

1. M.I.T. Radiation Laboratory Staff, *Relay Data Including Shock and Vibration Measurements, Radiation Laboratory Report 747* (Aug. 1, 1945).
2. Rosslyn, J., *Electric Relays* (Chemical Publishing Company, Inc., Brooklyn, 1921).

Punched-tape Storage

3. Alexander, Samuel N., Input and Output Devices for Electronic Digital Calculating Machinery, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 248–253 (1948).

4. Lake, Ross A., Teletype Reperforator Transmitter for Automatic Switching Systems, *Bell Laboratories Record*, Vol. 26, No. 3, pp. 106–110 (March, 1948).

5. Miller, Frederick G., Application of Printing Telegraph Equipment to Large-Scale Calculating Machinery, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 213–222 (1948).

6. Sparks, S., and R. G. Kreer, Tape Relay System for Radio-Telegraph Operation, *RCA Review*, Vol. 8, No. 3, pp. 393–426 (September, 1947).

7. Williams, S. B., Bell Telephone Laboratories Relay Computing System, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 41–68 (1948).

Photographic Storage

8. Carlson, F. E., and D. A. Prichard, Flash Tubes, *Illumination Engineering*, Vol. 42, pp. 235–248 (February, 1947).

9. Mees, C. E. Kenneth, *The Theory of the Photographic Process* (The Macmillan Company, New York, 1942), pp. 838.

10. O'Neal, R. D., Photographic Methods for Handling Input and Output Data, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 260–266 (1948).

11. O'Neal, R. D., and A. W. Tyler, *Photographic Digital Reader-Recorder*, *Progress Report 1* (Eastman Kodak Co., Rochester, N.Y., 1948).

12. Tyler, A. W., Optical and Photographic Storage Techniques, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 146–150 (1948).

Magnetic Storage

13. Aldous, D. W., A Bibliography of Magnetic Recording, *Electronic Engineering*, Vol. 19, pp. 390–391 (December, 1947).

14. Barrett, A. E., and C. J. F. Tweed, Some Aspects of Magnetic Recording, *Journal of the Institution of Electrical Engineers (London)*, Vol. 82, No. 495, pp. 265–285 (March, 1938).

15. Begun, S. J., Recent Developments in Magnetic Sound Recording, *Journal of the Society of Motion Picture Engineers*, Vol. 48, pp. 1–13 (1947).

16. Begun, S. J., Magnetic Recording and Some of Its Applications in the Broadcast Field, *Proceedings of the Institute of Radio Engineers*, Vol. 29, No. 8, pp. 423–433 (August, 1941).

17. Begun, S. J., *Magnetic Recording* (Murray Hill Books, Inc., New York, 1949).

18. Boenning, W. R., *Magnetic Recording and Reproduction of Pulses* (Engineering Research Associates, Inc., Arlington, Va., 1947).

19. Bozorth, R. M., Magnetism, *Reviews of Modern Physics*, Vol. 29, No. 1, pp. 29–86 (January, 1947).

20. Camras, M., Recent Developments in Magnetic Recording for Motion Picture Film, *Journal of the Acoustical Society of America*, Vol. 19, pp. 322–325 (March, 1947).

21. Camras, M., Theoretical Response from a Magnetic-Wire Record,

Proceedings of the Institute of Radio Engineers, Vol. 34, pp. 597–602 (August, 1946).

22. Cohen, A. A., and W. R. Keye, *Selective Alteration of Digital Data in a Magnetic Drum Computer Memory*, unpublished (Engineering Research Associates, Inc., St. Paul, Minn., 1948.) Pamphlet.

23. Coombs, J. M., *Storage of Numbers on Magnetic Tape* (Engineering Research Associates, Inc., St. Paul, Minn., 1947). Pamphlet.

24. Hansel, C. W., *et al.*, *Report on the Magnetophone*, OTS Report PB 1346 (1945).

25. Hickman, C. N., Sound Recording on Magnetic Tape, *Bell System Technical Journal*, Vol. 16, pp. 165–177 (1937).

26. Holmes, L. C., and D. L. Clark, Supersonic Bias for Magnetic Recording, *Electronics*, Vol. 18, pp. 126–136 (July, 1945).

27. Korhone, E. D., L. W. Reid, and W. L. Anderson, *Preparation and Characteristics of Magnetic Recording Surfaces*, pamphlet prepared by Engineering Research Associates, Inc., under Bureau of Ships contract NObsr 42001 (Jan. 7, 1949).

28. Kornei, Otto, Survey of Magnetic Storage, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 223–237 (1948).

29. Long, T. H., and G. D. McMullen, *B-H* Curve Tracer for Magnetic Recording Wire, *Transactions of the American Institute of Electrical Engineers*, Vol. 65, pp. 146–149 (1946).

30. Long, T. H., A New Wire Recorder Head Design, *Transactions of the American Institute of Electrical Engineers*, Vol. 65, pp. 216–219 (1946).

31. Massachusetts Institute of Technology, Staff of the Department of Electrical Engineering, *Magnetic Circuits and Transformers* (John Wiley & Sons, Inc., New York, 1943), pp. 3–40.

32. Menard, J. Z., *High Frequency Magnetophone Magnetic Sound Recorders*, OTS Report PB 12659 (1946).

33. Moore, B. L., Magnetic and Phosphor Coated Discs *in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery* (Harvard University Press, Cambridge, Mass., 1948), pp. 130–132.

34. Sears, A. W., Wire Recorder Wow, *Journal of the Acoustical Society of America*, Vol. 19, No. 1, pp. 172–178 (January, 1947).

35. Selby, M. C., Investigation of Magnetic Tape Recorders, *Electronics*, Vol. 17, pp. 133–135 (May, 1944).

36. Shaper, H. B., FM Magnetic Tape Transient Recorder, *Proceedings of the Institute of Radio Engineers*, Vol. 33, pp. 753–760 (November, 1945).

37. Sheppard, C. B., Transfer between External and Internal Memory *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 267–273 (1948).

38. Toomim, H., and D. Wildfeuer, Mechanism of Supersonic Frequencies as Applied to Magnetic Recording, *Proceedings of the Institute of Radio Engineers*, Vol. 32, pp. 664–668 (November, 1944).

39. Wooldridge, D. E., Signal and Noise Levels in Magnetic Tape Recording, *Transactions of the American Institute of Electrical Engineers*, Vol. 65, pp. 343–352 (1946).

Phosphor Storage

40. du Pont, E. I., de Nemours & Co., Inc., *Bibliography—Phosphorescence and Phosphorescent Sulfides*, special report prepared in connection with research under Contract No. W36-039-SC-32256 (Feb. 16, 1948).
41. Moore, Benjamin L., Magnetic and Phosphor Coated Discs, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 130–132 (1948).

Sonic-delay-line Storage

42. Arenberg, D. L., Ultrasonic Solid Delay Lines, *Journal of the Acoustical Society of America*, Vol. 20, pp. 1–26 (January, 1948).
43. Hunting, H. B., A. G. Emslies, and V. W. Hughes, Ultrasonic Delay Lines, Parts I and II, *Journal of the Franklin Institute*, Vol. 245, pp. 1–24, 101–115 (January and February, 1948).
44. Knott, G. C., Reflexion and Refraction of Elastic Waves with Seismological Application, *Philosophical Magazine*, Vol. 48, Series 5, pp. 64–97 (July, 1899).
45. Macelwane and Sohon, *Introduction to Theoretical Seismology* (John Wiley & Sons, Inc., New York, 1932, 1936).   2 vols.
46. Metz, F. A., Jr., and W. M. A. Andersen, Improved Ultrasonic Delay Lines, *Electronics*, Vol. 22, No. 7, pp. 96–100 (July, 1949).
47. Ridenour, L. N., *Radar System Engineering* (McGraw-Hill Book Company, Inc., New York, 1947), pp. 667–677.
48. Ringo, G. R., J. W. Fitzgerald, and B. G. Hurdle, Propagation of U.H.F. Sound in Mercury, *Physical Review*, Vol. 72, No. 1, pp. 87–88 (July, 1947).
49. Sharpless, T. Kite, Mercury Delay Lines as a Memory Unit, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 103–109 (1948).
50. Sharpless, T. Kite, Design of Mercury Delay Lines, *Electronics*, Vol. 20, No. 11, pp. 134–138 (November, 1947).

Electromagnetic-delay-line Storage

51. Blewett, J. P., and J. H. Rubel, Video Delay Lines, *Proceedings of the Institute of Radio Engineers*, Vol. 35, No. 12, pp. 1580–1584 (December, 1947).
52. Brillouin, Leon N., Electromagnetic Delay Lines, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 110–124 (1948).
53. Kallmann, Heinz E., High-Impedance Cable, *Proceedings of the Institute of Radio Engineers*, Vol. 34, No. 6, pp. 348–351 (June, 1946).
54. Kallmann, Heinz E., Equalized Delay Lines, *Proceedings of the Institute of Radio Engineers*, Vol. 34, No. 9, pp. 646–657 (September, 1946).
55. Montgomery, Dicke, and Purcell, *Principles of Microwave Circuits*, *Radiation Laboratory Series* (McGraw-Hill Book Company, Inc., New York, 1948), pp. 203–206.
56. Rudenberg, Rheinhold, Electromagnetic Waves in Transformer Coils Treated by Maxwell's Equations, *Journal of Applied Physics*, Vol. XII, pp. 219–229 (March, 1941).

57. Slater, J. C., *Microwave Transmission* (McGraw-Hill Book Company, Inc., New York, 1942).

58. Trevor, J. B., Jr., Artificial Delay Line Design, *Electronics*, Vol. 18, No. 6, p. 135 (June, 1945).

59. Zimmerman, K. H., Spiral Delay Lines, *Electrical Communications*, Vol. 23, No. 3, pp. 327–328 (September, 1946).

Electrostatic Storage

60. Forrester, J. W., *Electrostatic Storage Tubes, MIT Servomechanisms Laboratory Report No. 6345-R-110* (Jan. 13, 1947).

61. Forrester, J. W., *Electrostatic Storage Tubes for Digital Computers and other Information-Processing Systems, MIT Servomechanisms Laboratory Report No. R-153* (Jan. 26, 1949).

62. Forrester, J. W., High-Speed Electrostatic Storage, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 125–129 (1948).

63. Haeff, A. V., The Memory Tube and Its Application to Electronic Computation, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 24, pp. 281–286 (October, 1948).

64. Haeff, A. V., A Memory Tube, *Electronics*, Vol. 20, No. 9, pp. 80–83 (September, 1947).

65. Williams, F. C., and T. Kilburn, A Storage System for Use with Binary-Digital Computing Machines, *Proceedings of the Institution of Electrical Engineers*, Part III, pp. 81–100 (March, 1949).

66. Unsigned, An Image Storage Tube, *Electronics*, Vol. 21, No. 5, pp. 132–134 (May, 1948).

67. Zworykin, V. K., and G. A. Morton, *Television* (John Wiley & Sons, Inc., New York, 1940), p. 326.

The Selectron

68. Rajchman, Jan A., The Selectron—A Tube for Selective Electrostatic Storage, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 133–135 (1948).

69. Rajchman, Jan A., The Selectron—A Tube for Selective Electrostatic Storage, *Mathematical Tables and Other Aids to Computation*, Vol. 2, pp. 359–361 (October, 1947).

70. Rajchman, Jan A., The Selectron—A Tube for Selective Electrostatic Storage, *RCA Laboratories Division News*, Vol. 6, No. 2 (August, 1949). A paper presented at the IRE Tube Conference at Princeton University in June, 1949, and also at the Symposium on Large-Scale Digital Computing Machinery at Harvard University in September, 1949.

Transfer in Space

71. Bode, Hendrick W., *Network Analysis and Feedback Amplifier Design* (D. Van Nostrand Company, Inc., New York, 1945).

72. Hansen, W. W., Maximum Gain Bandwidth Product in Amplifiers, *Journal of Applied Physics*, Vol. 16, pp. 528–534 (September, 1945).

73. Page, C. H., Digital Computer Switching Circuits, *Electronics*, Vol. 21, No. 9, pp. 110–118 (September, 1948).

74. Rubinoff, Morris, An Input Device Using Multiple Gates, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 254–259 (1948).

75. Terman, F. E., *Radio Engineering* (McGraw-Hill Book Company, Inc., New York, 1947).

76. Wheeler, H. A., Wide Band Amplifiers For Television, *Proceedings of the Institute of Radio Engineers*, Vol. 27, pp. 429–438 (July, 1939).

Coincidence Circuits

77. Bay, Z., and G. Papp, Coincidence Device of $10^{-8} - 10^{-9}$ Second Resolving Power, *Review of Scientific Instruments*, Vol. 19, No. 9, pp. 565–567 (September, 1948).

78. Dickie, R. H., A High Speed Coincidence Circuit, *Review of Scientific Instruments*, Vol. 18, pp. 907–914 (December, 1947).

79. Hofstadter, R., and J. McIntyre, Note on the Detection of Coincidences and Short Time Intervals, *Review of Scientific Instruments*, Vol. 21, No. 1, pp. 52–53 (January, 1950).

CHAPTER 15

DATA-CONVERSION EQUIPMENT

15-1. Introduction

The essence of the computing problem may be said to be the expedient handling of data from origin to final state. However, the steps between the original sensing of values and their final presentation after a calculation are sometimes complex and may require considerable time. Moreover, it may be desirable to accept original data in one form, say as decimal digits, perform computations on these data expressed in another form, say as binary digits, and produce results in still a different form, say as graphs.

Let us consider all the steps in the conversion of data from their primary state into an appropriate form for output from a computer.

The sources of data which are suitable for automatic computing and the channels through which the data may flow to the computer are highly diverse. For example, machines used by the Bureau of the Census handle data which were passed by word of mouth to a census taker and jotted down by him in a notebook. The data for a Fourier analysis of a complex wave form may come from an electromechanical pickup mounted on an aircraft engine. Number-theory problems often originate in a mathematician's mind.

Data may accumulate rather slowly from a number of different sources, as in the census example, in which case the data must be stored and transported to a central location. Data may emanate from one source and accumulate rapidly, as in the engine-vibration example. Here the computations may be performed at the time and place where the data are being sensed, and intermediate storage of data may be avoided. The latter arrangement has been termed on-line operation.

It may be desirable to convert the data from analog form to digital form, or vice versa. For example, it may be necessary

to use as one of the problem inputs the instantaneous value, at 1-second intervals, of an output voltage which varies with time. The instantaneous values of this voltage, appearing as physical voltages and not as *numbers* of volts, are analog data. Measuring a physical quantity and expressing it as a *number* (of volts, inches, degrees, etc.) constitute conversion from analog to digital form. Analog computers operate directly on analog data; digital computers operate on numbers, as was explained in Sec. 11–1.

When data are introduced into a high-speed machine, it may be desirable to enter them at a faster rate than operators can handle. Therefore, speed-changing devices may be required to introduce data rapidly. Likewise, computed results may be produced more rapidly than the output medium can assimilate them. In this case, the results must accumulate in some intermediate form until they can be written in final form. Finally, when data are written, they should be in a form which will be convenient for the user.

15-2. Analog–to–digital Converters

An analog-to-digital converter is a device which accepts instantaneous values of continuously variable quantities and expresses them in discrete numerical form.

15-2-1. Elementary Types. An example of an analog-to-digital converter is the crystal-controlled pulse generator and associated ring counter of Sec. 3-6, which measures and expresses time intervals in discrete numbers of electrical pulses.

Another example is a device for converting speed of rotation to digital form A cam-operated electrical contactor, fastened to a revolving shaft, opens and closes its contact N times for each shaft rotation, N being the number of raised segments on the cam. Incidentally, this converter has the disadvantage that positive and negative directions of rotation are represented in the same manner, and there is no zero reference. In other words, it counts *modulo N* but not *radix N*.

15-2-2. Meters and Counters. In the familiar watt-hour meter[2] there is a set of gears with a ten-to-one reduction ratio, a pointer, and a dial numbered from 0 to 9 for each decimal digit to be registered. This device is a measuring instrument, as distinguished from a counter. The dials rotate smoothly, at a

rate proportional to the instantaneous power usage; the device
measures the time integral of power usage.   At regular intervals
the meter reader looks at the dials and assigns a discrete value to
each dial setting.   In assigning discrete values to these settings,
he is performing a conversion of data from analog to digital form.
In fact, the meter reader himself satisfies the definition stated in
Sec. 15-2.

Meters (which yield analog data) may be distinguished from
counters (which yield digital data) on this basis: a counter is
arbitrarily precise in its representation of a quantity in digital
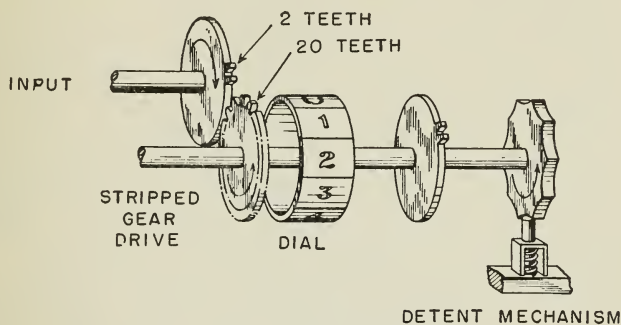
FIG. 15-1.   Stripped-gear counter.

form, and no particular skill is required to read it; on the other
hand, some skill is required to read a meter, because the meter
measures and expresses a quantity, or, at least, the last digit of
its measured value, in analog form.   The higher the degree of
precision required in the expression of a particular meter reading,
the greater the observer's skill must be and the less likely two
different observers are to read identical digital values from the
same dial or scale and pointer configurations.

An example of a typical rotation counter is shown in Fig. 15-1.
This counter has numbered dials which move in discrete steps
from one digit to the next.   This counter employs a driving
pinion connected to an input shaft stripped of all but n/r of its
teeth, where n is the original number of teeth and r is the radix
of the number system being used.   The stripped pinion drives a
gear having n teeth, possibly through an idler gear, and advances
it $1/r$ revolutions for each revolution of the stripped driver.
Thus on a radix 10 counter, in which the position of the full-
toothed gear is represented by the digits 0 to 9 around a dial, a

new decimal digit appears at an aperture after each revolution of the input shaft. A single decimal counter is illustrated in Fig. 15-1. A group of such dials, cascaded so that the output of one rotates the stripped gear of the next, form a decimal counter on which as many digits as desired may be represented. A detent mechanism may be required to hold the full-toothed gears in their discrete positions when they are not meshed by the teeth on the partially stripped gears. The dials on these counters are commonly placed on the outside of short cylinders, which are located behind a window aperture and arranged so that only one digit on each cylinder is in view at any time.

Mechanical counters are manufactured by the Veeder-Root Company. Various models of this device count rotations, reciprocating displacements, and electrical pulses. The output of these counters is normally a set of numbers to be read visually. However, several models are available with printing wheels as an optional feature. The top speed recommended for any Veeder-Root Counter is about 42 counts per second. Most of them operate at about 15 counts per second maximum.

Mechanical counters have also been equipped with electrical *pick-offs* at each discrete position in order to generate electrical signals representing shaft position. An analog-to-digital converter which consists of a group of rotary switches geared to a rotating shaft is used in the MIT differential analyzer.[1] These switches in conjunction with conversion circuits operate an electric typewriter and type numbers which indicate the shaft location. Another electrically read device is used in the Harvard Mark I and is known as an electromechanical counter wheel. Each counter wheel has 10 discrete positions and a set of electrical contacts for each position. The wheels are driven to their prescribed position by means of a common drive motor through a magnetic clutch. The clutch actuation is controlled by a number-input mechanism which causes each wheel to be driven to the digit to be represented.

15-2-3. Film Converter. Another rotational analog-to-digital converter has been proposed which can be arranged to convert shaft rotations to numbers of radix 2. This converter employs a continuous strip of film on which transparent and opaque spots are photographed. These spots are arranged in columns across the width of the film to represent binary numbers, which increase

by one unit for each successive column along the length of the film. Thus, a 15-digit binary number is represented by a column of 15 code areas across the film and provides an opportunity to employ up to 2^{15} different numbers for the shaft rotation involved. The film drive is of the intermittent type used in motion-picture projectors and thus fulfills the definition of a continuous-to-discrete-value converter. Common types of intermittent drives are the Geneva movement and the Powers pin-cross and the
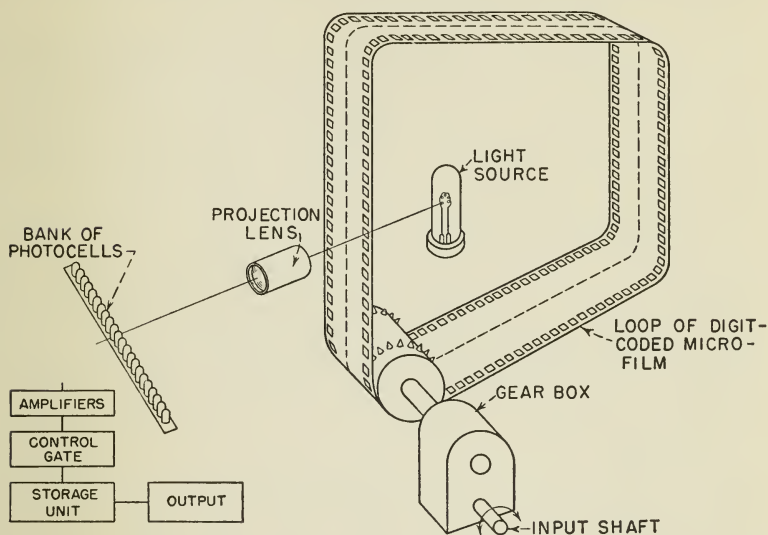
Fig. 15-2. A scheme for converting analog to digital data. The micro-film drive moves smoothly while coded areas on the film are sensed photo-electrically.

star-wheel drives.[3] A given amount of input shaft rotation causes the film to move forward one discrete frame or, in this case, one column of code areas. If a photocell is used in conjunction with each of the 15 rows of spots in a column and if a light source is arranged to shine through the film and on the respective photocells, as described in Sec. 14-2-3, a radix 2 digital representation of the input shaft position is imposed upon the photocells.

A mechanical film drive requires more than a millisecond to step the film from one code column to the next. In order to increase the maximum film speed and to eliminate ambiguity

during the advance time, an electronic synchronizing scheme could be used instead of the intermittent film drive. A row of synchronizing dots on the film might control gates which allow the photocells to be sensed when the code pattern is centered in the optical system, but at no other time. Thus, each number is held until it is possible to sense the next one correctly. A converter of the type described has been proposed by ERA and is illustrated in Fig. 15-2. It is thought that 50 binary digits can be stored across the width of a 35-millimeter film.

15-2-4. Pulse Code Modulation. A wide-band communication system which has desirable signal-to-noise characteristics has been devised using a binary representation of analog signals.[4–11] This system is called pulse code modulation, or p.c.m., and is of interest because of the converters which are used and because of the possibility of transmitting analog data to a computer in digital form over a telemeter link.

Three general types of analog-to-digital converters have been devised for p.c.m. In all three schemes, the instantaneous amplitude of the analog signal is measured or sampled periodically. The sampling rate is made slightly more than twice the highest frequency component which is desired to be sent. Higher frequency components are filtered out of the signal previous to the sampling procedure.

In the first type of p.c.m. converter, the number of voltage units in an amplitude sample are counted successively a unit at a time until a residue of less than one unit remains.[4,7] In the second type of p.c.m. converter, the amplitude of the sample is compared with a standard voltage which corresponds in magnitude to the highest order digit, and the standard voltage is subtracted if it is smaller than the sample. The remainder of the sample is then compared with a voltage corresponding to the next highest order digit, and another subtraction is made if the sample is larger. This process is continued until the lowest order, or unit, digit is compared with the remainder and is determined to be either larger or smaller.[6] In the third type of p.c.m. converter, the sample amplitude is measured, and the corresponding digit pulses are produced without intermediate counting or subtracting.[8,10,11] The speeds of these three systems increase in the order of listing given above.

Examples of the first and third converters are given in order

to illustrate their operation and circuitry. The first is perhaps the most simple of the p.c.m. conversion schemes and will serve as a basis for comparison in describing the third type, which is the fastest.
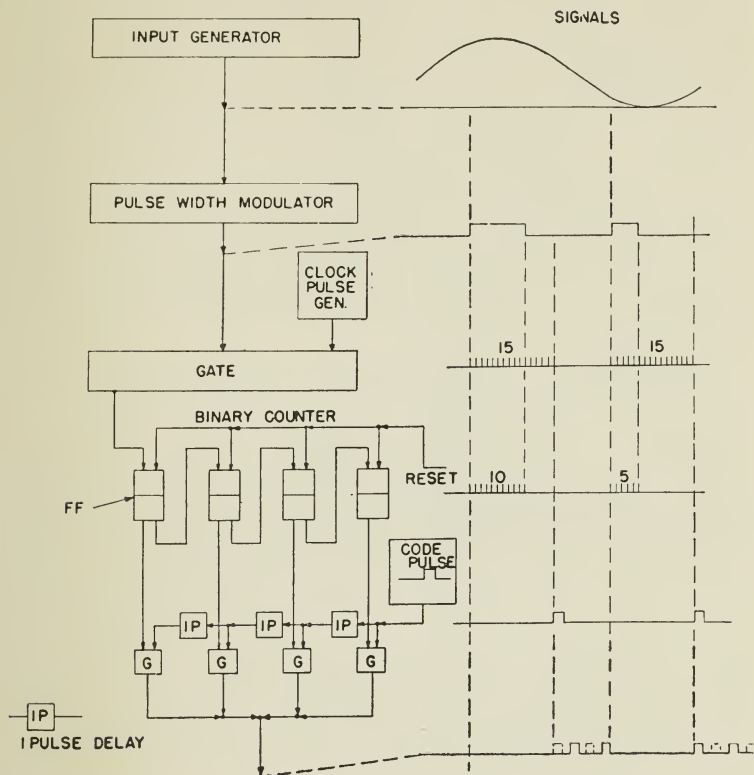
FIG. 15-3. Pulse code converter. The instantaneous amplitude of a signal produced by the input generator is periodically sampled and converted into a proportionate number of clock pulses by the modulator and gate and into digital representation by the counter. This number of pulses must be between 0 and 15 in the equipment illustrated. Here two samples of the signal are being taken; the amplitudes at the points where the signal is being sampled, expressed in numbers of pulses, are shown as 10 and 5 (with respect to some bias voltage).

The first scheme was patented by A. H. Reeves and has been investigated by Bell Telephone Laboratories and by the Federal Radio and Telephone Corp.[4,7] A sample of the voltage representing the data to be converted is first introduced into a sampling

pulse-width modulator which measures the amplitude of the incoming voltage with respect to a reference bias voltage and gates out a pulse having a width proportional to the measured amplitude. The variable-width pulses control a gate in such a way that a number of clock pulses nearly proportional to the width of the pulse is passed through. If the analog data are to be represented by four binary pulses, they are represented to within 1 part in 2^4, or 16. Therefore the clock-pulse rate is set so that pulses of maximum width consume the same amount of time as 16 clock intervals, and each amplitude sample is replaced with from 0 to 15 clock pulses which represent the amplitude of the sample. The series of clock pulses is counted in a four-stage binary counter of the type described in Sec. 13-5 on ring counters. A block diagram of this system is illustrated in Fig. 15-3.

The third type of p.c.m. conversion has been accomplished at the Bell Telephone Laboratories[8,10,11] using a perforated-mask type of converter. The mask intercepts an electron stream in a tube similar to a conventional cathode-ray tube. The beam passes between two perpendicular pairs of electrostatic deflecting plates which serve to locate the beam on the mask. A positive-potential pulse plate is mounted at the large end of the tube perpendicular to the longitudinal axis, and the mask is placed in front of it. Thus, the electron stream may or may not reach the collector plate, depending upon whether it is deflected to pass through a hole in the conducting mask or not. The collector plate might be replaced with a photoelectric cell which detects the presence of a light spot anywhere on a conventional c.r.o. screen.

The analog voltage is sampled periodically, and a voltage which is proportional in amplitude to the sample is applied to one set of deflection plates. Simultaneously, a sweep voltage is applied to the other pair of deflection plates so that the electron beam scans across some portion of the mask. The mask has holes cut in it similar to the pattern illustrated in Fig. 15-4 so that if the sample amplitude controls the vertical location of the beam, the electrons reaching the collector will be modulated by the mask in a manner dependent upon the code pattern. The code pattern illustrated represents the binary numbers from 0 to 15, and consequently the electrical pulses on the collector

plate produce a serial binary number which is roughly proportional to the analog sample. Auxiliary equipment is used to assure that the electron beam moves along the center of a row of code areas and not along the edge.

This device has been used to sample 12 channels of voice frequencies up to 3,400 cycles per second simultaneously utilizing a mask which is seven binary digits wide. Electrical pulses 1.5 microseconds wide are transmitted at a rate of 672 kilocycles per second.

15-3. Digital-to-analog Converters

After the required computations have been performed in a digital computer, it is sometimes desired to transform the digital results into analog form for ease of viewing and assimilation.

Fig. 15-4. An analog-to-binary-converter screen.

Three types of digital-to-analog conversion are discussed here: conversion of pulses to voltage amplitudes, conversion to rotary motion, and conversion directly to a graphical representation.

15-3-1. Pulse Converters. A series of electrical pulses may be transformed into a d-c voltage by allowing them to gate standardized pulses into a capacitor having very low leakage. The voltage across the terminals of the capacitor at any instant is proportional to the number of pulses gated into the capacitor since the last time it was fully discharged. This transformation method was the basis of some of the early measuring circuits associated with Geiger-Müller counters and is also used in frequency dividers.

A digital-to-analog converter employed in the Bell Telephone Laboratories p.c.m. system makes use of the exponential decay characteristics of an RC circuit. The digits of a binary number are represented by standardized electrical pulses, separated

evenly in time and arranged so that the lowest order digits appear first. They are applied to a capacitor circuit with the RC constant so adjusted that half the charge leaks off between pulse times. The total charge after all the digits in a number have been inserted is proportional to that number.

For example, consider a converter in which four-digit binary numbers are introduced in pulse times 1 to 4 and are removed at pulse time 5. If the binary numbers 1,010 and 1,111 are converted in this fashion and inserted pulses are 80 volts in amplitude, the voltage on the capacitor as a function of pulse times is shown in Table 15-1.

TABLE 15-1

Pulse time	Binary number = 1,010 Decimal number = 10		Binary number = 1,111 Decimal number = 15	
	Binary digit	Capacitor volts	Binary digit	Capacitor volts
1	0	0 + 0 = 0	1	0 + 80 = 80
2	1	0 + 80 = 80	1	40 + 80 = 120
3	0	40 + 0 = 40	1	60 + 80 = 140
4	1	20 + 80 = 100	1	70 + 80 = 150
5	. .	50	. .	75

It can be seen from the above example that the decimal conversion factor is 5 volts per unit at the pulse time. The converter described was proposed by Shannon at Bell Telephone Laboratories. An important modification has been developed by Rack[5] which makes timing less critical.

15-3-2. Rotation Converters. By a slightly different arrangement the rotational analog-to-digital converters previously discussed can be operated in an inverse manner to produce a rotational output for a digital input. A servomechanism[12-15] may be coupled to the film converter or electromechanical wheel counter in such a way that the motor drives the shaft until a desired number is located.

As an example, consider the film converter of Fig. 15-2. Numbers from 0 to $2^4 - 1$ are represented successively in binary form as transparent and opaque dots on a four-row field as in

Fig. 15-4. The fourth-order or left-hand row of digits is half 0's and half 1's (eight of each). The photocell output for that row is compared with a signal representing the desired digit, and the servo drives the film in the correct direction so that the fourth-order digits are caused to match. After the servo has driven the film to the nearest point where fourth-order digits match, a comparison of third-order digits is made. Within the limits in which fourth-order digits are matched, half the third-order digits are 1's and half are 0's (four of each). Again, the desired third-order digit is compared with the third-order digit on which the film is set, and the film is either left where it is or driven to the first point where third-order digits coincide. This process is continued until the film is driven to a point where its digits coincide with those of the number desired.

Another example utilizing the principles involved in the film converter can be found in the MIT differential analyzer.[1] This is a tape-operated device which sets the displacement of the integrator disk to within 1 part in 150,000 of the maximum displacement after a punched tape is prepared by an operator. The conversion from the digital value on the tape to shaft rotation is accomplished with the aid of 44 relays and a set of rotary selector switches on the output shaft. The shaft is driven by a servomotor which, in turn, is controlled by the switches. First the relays are set by electrical contactors actuated by the holes in the punched tape; then the servomotor is caused to rotate in the direction to close corresponding switches on the integrator shaft. The motor continues to drive the integrator shaft until a switch corresponding to each digit is closed, whereupon it is caused to stop.

15-3-3. Use of the Selectron as a Digital-to-analog Converter. As originally conceived (see pages 370–371) the Selectron had a 64-by-64 grid, on a cylindrical surface. By deforming such a grid into a plane, it would be possible to produce a device for converting digital data directly into graphic form. The electron beam would be directed upon a spot on the screen opposite a particular window in the grid for each digital value; the composite of the spots would form a graph.

For example, if some digital $F(x)$ is to be plotted as a function of digital quantities, represented by the letter x, the x quantities are caused to operate the gates associated with the vertical bars,

and the $F(x)$ quantities are simultaneously applied to the horizontal bars.   Thus, in a Selectron system having 2^5 vertical bars, values of x ranging from 0 to $2^5 - 1$ might be plotted.   Then, all values of x between 0 and 1 would cause the left-hand set of bars to be made positive; all values between 1 and 2 would cause the next pair to the right to become positive, etc., so that values between 30 and 31 would cause the right-hand pair of bars to become positive.   A similar scheme would be used to introduce $F(x)$ values on the horizontal bars, and quantities would be plotted to within an accuracy of about 3 per cent.   If x and $F(x)$ values are directed to the vertical and horizontal bars simultaneously, a Selectron window is opened at a point on the grid which represents the value of these two quantities.   Numbers to be plotted can be brought into the range of the plotting device by a previous multiplication or division by a scale factor as required.   A photosensitive screen placed behind the Selectron grid, in place of the dielectric storage plate, would cause a visual spot to be obtained for each x versus $F(x)$ entered.

15-4. Conversion of Data from Written to Automatic Machine Form

Data for automatic computation often are presented in a written form not easily assimilated by a computer.   They must be converted therefore into some new form for introduction into an automatic machine.

The most common written form for recording data is the familiar *numerical tabular form* used in accountants' records, engineering data sheets, and tables of mathematical functions.   Data may be recorded also as a set of plotted points, which represent in analog form a function of one parameter in terms of discrete values of another parameter.   As a variation of this form, a set of lines may be drawn from the plotted points perpendicular to one coordinate axis.   These two presentations of plotted data will be referred to as *sampled analog data*.

The numerical tabular form of written data may be introduced into a computing machine by a human operator who reads the data and manipulates dials or a keyboard in a manner which corresponds with his mental evaluations of the data.   Sampled analog data may be transferred by an operator in one of two different ways.   First, the various points of a plotted set may

be measured from a reference with a scale or ruler. Numerical measure thus is afforded which may be introduced into a machine in a manner similar to numerical tabular data. A second method of transferring plotted data is to superimpose a sighting device upon the points, thereby producing shaft displacements which may be used in the machine as measures of the function values. This method permits the transfer of data without requiring the operator actually to sense the values mentally.

Data in the form of continuous curves may be evaluated in a manner similar to the methods used for sampled data. Automatic curve-tracing devices have been made which utilize photoelectric-cell sensing to control a motor drive for the curve follower.

In this section are included descriptions of equipment which has been devised to perform data conversion from a written form into a form suitable for introduction into an automatic computing machine. In addition, a description is included of a proposed device with which data may be entered into a computer from vocal commands.

15-4-1. Conversion from Written to Automatic Digital Form. The automatic digital forms into which data may be transcribed are discussed in Sec. 14-2 under the topic of storage mediums. The ways in which data have been transcribed from written form to these automatic digital forms are numerous, but all employ the idea of the keyboard or the dial. It is important that the converting process allow the operator to keep his eyes on the data being read.

Keyboards. There are several types of keyboards in common use. The simplest of these is the single key used by the telegraph operator. The semiautomatic key for high-speed telegraphy has a neutral position, a position for dots, and a position for dashes so that it can be used to generate binary numbers serially.

The Sundstrand keyboard is used on some desk calculators. It employs only 10 keys representing the 10 decimal digits. These keys are depressed successively, and a number is formed in a small register window where it can be checked before the operation begins or before the number is transferred to the storage medium. This keyboard can be operated quite rapidly by the touch system.

The Burroughs keyboard, in which $10n$ keys are required to

set up an n-digit number, is common on desk computers. It may be operated by the touch system, but not as rapidly as the Sundstrand. The desirable feature of this keyboard is that after all the numbers are set on the keyboard, they may be checked visually, and any errors may be corrected easily before the computation with the number begins.

A familiar keyboard is that of the ordinary typewriter, which may be operated at speeds of about five letters per second. It has the desirable feature that many operators are already proficient in its use. A keyboard based on the standard typewriter keyboard is being used as the input device for the UNIVAC. If a typewritten record is made in conjunction with the data-conversion process, it may be checked for errors, which may then be corrected before computations begin.

Dvorak[18] has suggested rearranging the positions of the keys on the typewriter keyboard so that easier finger motions are required in typing the frequently used letters and combinations of letters. Machines have been equipped with Dvorak keyboards and used successfully in speed tests and other demonstrations. It is said to be difficult to retrain an experienced operator to use this keyboard, but novices can probably learn to use it at least as easily as a conventional keyboard. Typewriters equipped with this keyboard are available.

The keyboard on the teletypewriter basically is similar to that of the typewriter but has three rows of keys rather than four. All letters are upper-case, and numbers are accessible by use of the shift key. The machine has a top speed limit of about six letters per second. Teletypewriter equipment is used in the input equipment for the Mark II and Bell Telephone Laboratories machines.

Another keyboard which has been devised for high-speed typing is manufactured by the Stenotype Company. This keyboard contains 22 keys which represent the common sounds of the language used. The operator uses the fingers of both hands and by depressing several keys simultaneously produces the associated sounds in the desired order. Stenotype operators can attain a speed of 250 words per minute. Part of this increase in speed results from the reduced number of keys which must be struck for each word written, and part of it results from the fact that several keys may be depressed simultaneously. Although

this machine is not particularly suitable for use in a computer, the possibility of using the principle of making up numbers and orders out of combinations of keys might well increase the speed of transcription from written to automatic machine forms.

Keyboard Coupling Devices. Manually operated keyboards are usually employed to record data on storage mediums for future use in a computer. The actual coupling between the keys and the recorder unit then depends to a large extent upon the type of keyboard and the type of storage employed. There are possible a large number of coupling devices, some of which are suggested in this section. The teletypewriter is described as an example which employs mechanical coupling, electric switches, and an encoding commutator.

When storage mediums such as punched tape and punched cards are used, the coupling between keyboard and punches may be entirely through mechanical linkages. These linkages are arranged to set the proper punches, which are usually driven through the paper by an auxiliary power source.

When an electrical output is desired from the keyboard, the keys may operate switches directly or may operate small shutters which control a supply of light to a group of tiny photocells. The latter principle is applied in Teleregister Corporation equipment.[16] Photocells have the advantage over switches that they do not have contacts which may become dirty and thus cause intermittent operation. Other couplings, which might be devised to overcome this rather fundamental difficulty with small switches, are electrostatic and inductive devices which can be coupled and uncoupled with a keyboard.

The teletype machine is a combination keyboard and printer.[20] The keyboard section contains five switches, and each key causes one or more of these switches to be closed when it is depressed. The coupling from key to switch is through levers, cams, and five sets of push rods, which run the width of the keyboard. When a teletypewriter is used in conjunction with a tape perforator, the five switches control five pin-setting devices. An electromagnet in the punching machine causes the set pins to be driven through the paper tape. Thus, a unique hole pattern is produced in the tape for each key depressed.

When the output of the teletype machine is to be transmitted by a single wire, the five switches are connected to the central

five positions on a seven-place electromechanical commutator. When a type key is depressed, corresponding switches are closed and latched, and the commutator arm begins a single rotation. In its first position a start pulse is placed on the single wire; in the next five positions positive or negative pulses corresponding to the particular key depressed are generated; and in the seventh position a stop pulse is sent out. At the completion of the cycle the switches are released, and the machine is ready for the next key to be depressed. Thus, for each letter typed on the keyboard, a string of seven positive and negative pulses are emitted serially from the single-wire transmitter.

The receiver unit of a teletypewriter consists of a mechanical amplifier which increases the mechanical power available from the relatively weak incoming signals so that they can be separated and used to swing a type bar corresponding to the key depressed.

When a teletypewriter is set to send or receive at one time, the receiver unit may be electrically connected to the output of the keyboard so that a copy of the material being sent out is typed at the sending station. This arrangement aids the operator in keeping track of lines and in locating errors. Since the keyboard and the printer are located together, one unit suffices for sending and receiving.

The IBM Corporation produces the Electromatic Typewriter, which is available on a rental basis. Although this machine was not originally intended for transmitting information, the Globe Wireless Company has modified these machines so that they are suitable for this purpose. In the IBM machine the keys operate combinations of six switches simultaneously. The Globe Wireless Company has added an electronic conversion unit which transforms the information in six static switch settings into a serial string of six pulses. When the key is struck, it also causes a type bar to strike the paper in the manner of a conventional mechanical typewriter, so that a record of the message sent is also made at the transmitter.

A separate machine is required as a receiver in the Globe Wireless equipment. This machine has an electronic converter called a *permutation unit*, which distributes the six serial pulses into six channels. Each of these six channels controls an electromagnet, which in turn operates code levers. These code levers operate

the type bars which actually print the characters corresponding to the serial pulses received.

Another tape-punching typewriter designed for special uses in offices is made by the Justowriter Company. It, also, is available on a rental basis.

Dials. A second conceivable manner in which an operator might transcribe data from numerical form to automatic digital form is by setting the numbers on dials with discrete positions on them. These positions might be determined by a detent mechanism arranged to hold the dial on positions corresponding to whole numbers. A variation of this idea is illustrated by the telephone dial, with which discrete numbers are registered by means of discrete holes in the finger ring.

Speech Converters. One of the most common forms of error which operators make in keying or dialing numbers is to transpose digits. Furthermore, a certain amount of skill is required in order to operate a keyboard speedily and accurately. An input device which allows the operator to concentrate on the data which he is entering into the machine and which requires little practice to use might be devised from a speech converter. The operator merely reads the digits of a number into a microphone, and the digit names are recognized by selective filters in an amplifier. On the basis of this recognition, gates are opened which convert the sound of the spoken digit into automatic digital form. The reproduction of speech from elementary sounds was studied at Bell Laboratories prior to 1938[17] and has recently been reconsidered in England.[19]

15-4-2. Conversion from Written to Automatic Analog Form. A simple way in which to convert numerical values into analog form suitable for automatic machine operation is by means of a dial mechanism or keyboard. Plotted data may be converted for introduction to a computer either by measuring and entering from numerical values or by sensing with a curve-following device connected to a displacement mechanism.

Dials, Scales, and Keyboards. A dial may be used to establish an angle of rotation on a shaft, to set a potentiometer, or to rotate a synchro.[20] A scale is customarily used for measuring linear distances. The familiar slide rule, for example, has linear scales along its edges, and numbers may be converted from

digital to analog form and back again by the operator several times in the course of an arithmetic operation.

Similarly, analog values could be set with a keyboard in a number of different ways.   For instance, each key might operate a lever arm having a length proportional to the key number. Also, each key might operate an electrical contactor and thereby switch impedances or voltages in a manner designed to create appropriate analog values.   Another possibility is to have keys set stops in some device which moves until it hits a stop, thus converting key settings to linear or angular displacements.

A keyboard is used originally to punch tape with which the integrator disks are set on the MIT differential analyzer as described in Sec. 15-3.

Curve Followers.   An operator can transfer discrete or continuous values plotted on paper into automatic form by means of a curve follower.   In essence, the operator controls a cross hair or similar sighting device so that it moves from one to another of a set of discrete plotted points or along the curve of a continuous record.   The position of the tracing element with respect to a selected reference point corresponding to the origin of the curve becomes the analog output of the machine, which may be in the form of shaft rotations or linear displacements.   If a curve is plotted in rectangular coordinates designated by x and y, the medium on which the curve is drawn may be moved by a motor drive in the x direction.   The operator then controls the y position of his sight and possibly the speed and direction of the x drive.   Such a scheme is easier for the operator to use than one in which he controls both axes completely and simultaneously.   Equipment requiring manually operated drives similar to the one just described is common in gun-laying and bombing computers.

The output of a curve follower usually consists of a pair of shaft rotations which are proportional to two plotted quantities. Sometimes one shaft will be driven by another part of the machine, in which case only the operator-controlled shaft will constitute an output.   As a check upon the accuracy of the operator, a pencil can be attached to a curve follower so that the distance between the recorded curve and the curve traced out by the operator can be determined.

Automatic curve followers have been considered for some time,

and one was built by Hazen, Jaeger, and Brown at MIT and reported in 1936.[22] It is capable of following a single-valued function of one variable within the range where its first derivative remains finite. Their machine was successful and is the fore-runner of a similar automatic curve follower which was developed in Germany at the Technische Hochschule, Darmstadt.[23] It operates on the basis of a photocell which rides along one edge of a black line on a light background. An image of an illuminated section of the line is focused on the photocell through a circular aperture which passes a spot of light having a diameter equal to about one-third the width of the line. The equilibrium position of the photocell is established at the place where half the image is black and half the image is white. With this arrangement an increase in illumination, which results when the photocell is moved away from the center of the line, energizes a servomechanism, which drives the photocell back toward the center of the line, and a decrease in illumination causes the photocell to be driven away from the center of the line. Since the servodrive is along the y axis, the sight is actually driven in the $+y$ direction in one case, and in the $-y$ direction in the other. Therefore, it is essential that the sight remain on the side of the curve on which it will always be led back toward the center of the line.

The accuracy of this curve follower is a function of the amount of hunting which occurs in the servosystem. The amount of hunting depends in a rather complicated manner upon the circuit parameters and the input functions and is described in the theory of servomechanisms.[12,14,15] In general, the inertia of the system should be as low as possible to reduce hunting amplitude. Increased damping decreases the hunting but also increases servo lag and consequently error.

With a given servosystem and a limit set on the error which can be tolerated, the velocity of the servo movement must be restricted. This velocity is equal to the product of the x-direction velocity of the curve times the slope of the curve. In the German machine a speed of 3 millimeters per second in the x direction is a safe drive speed. The error caused by the aperture hunting a distance d normal to the curve and with the curve making an angle α from the y axis is

$$E = \frac{d}{\cos \alpha}$$

and on the German machine a value of $d = 0.1$ millimeter was attained.

Other automatic curve followers have been considered in which a second photocell is used to detect the slope of the curve. The main servo-circuit parameters thereby may be controlled in such a way that the curve may be followed efficiently over the wide range of required servo velocities resulting from the wide range of slopes on various parts of the curves to be followed. An application of modern servomechanism circuit techniques to this problem would probably be well rewarded.

The automatic curve follower just described, like the MIT device, is limited to sections of curves in which the first derivative or slope is finite, since the control circuit is unstable when its orientation from the center of the curve is reversed. Manually operated followers, however, are not limited in this respect. When a curve follower is guided by a human operator, any of several curves plotted on the same medium can be chosen and even changed during the course of the problem if data are changing slowly. Therefore, functions of more than one variable can be followed manually. Use of an additional parameter is not feasible with an automatic curve follower, since it is unable to distinguish one curve from another or to interpolate between curves.

15-5. Transcription of Data from Automatic Machine to Written Form

After a computation is completed, it may be desirable to study the results, and therefore the output of the machine must be produced in a written form. This is not particularly necessary in a gun-director computer or similar tactical device, where the results are used as they are computed, but it is important in solving the problems of science. The written values may take any of several forms, the most convenient depending upon the nature of the data and their eventual use. If large amounts of data are to be listed in tabular form, it is well to list tabular summaries of the data also. In some cases it will be most convenient to have the output data plotted in graphical form. In this section is discussed the conversion of quantities from machines to written numerical or to written analog forms. Equipment which performs this function is referred to as output equipment.

15-5-1. Transcription from Automatic Digital to Written Numerical Form. The mechanical, electronic, and photographic printers which have been devised for numerical transcription are discussed in this section.

Electric Typewriters. There are several commercially available electric typewriters which are of interest as output devices for printing results in numerical form.

The teletype equipment,[20] described in Sec. 15-4-1, is a combination keyboard and printer. It can be used to type numbers and letters from groups of synchronized pulses. Similarly, the modified IBM Electromatic Typewriter and the Justowriter equipment can be operated from punched tape or other sources of electrical input signals. RCA Communications, Inc., also produce a number of tape-operated typewriters.[28]

In a conventional typewriter, when a key is depressed, an arm with an individual type face on it is caused to fly up and hit the paper, thereby printing a letter or a number. There is another group of typewriters in which the font of characters is placed on a small wheel located behind the paper. When a key is struck, the wheel is caused to rotate until the corresponding type lies behind the paper, which is then struck by a hammer from the side opposite the font. This is the principle of the early Hammond typewriter, a Dow Jones machine, a Western Union machine, and the Vari-typer. The action is simple, and it is rather easy to change type fonts. Typewriters are essentially serial printers since only one letter is printed at a time.

Multibar Printers. A somewhat faster device is known as the multibar, or gang, printer in which all the type is arranged along a bar, and several like bars are mounted side by side. In one version of the multibar printer, each bar is displaced longitudinally so that the correct figures are aligned to form an anvil behind the paper. Letters and numbers are printed by means of a hammer which hits the paper and presses it against the bars. As an example, a printer having three bars each containing the digits 0 to 9 can print three-digit decimal numbers. If the number 281 is to be printed, the left-hand bar rises three places to the digit 2, the second rises to the digit 8, and the third rises so that the digit 1 is under the hammer. Since these bars may be raised simultaneously, the total time required for printing is the time required to set a bar and to strike it with the hammer.

A mechanism of this type is found in the ordinary desk adding machine.

In a Remington Rand multibar printer, each letter is mounted at the end of a rod which is free to slide longitudinally.   A stack of these rods is arranged in a carriage which slides up and down like the bar in the multibar printer.   In this type, the hammer hits the correct type rod and presses it against the paper.

Gang printers which are now available include the Remington Rand and IBM printers, which attain a printing speed of over 100 characters per second.   These machines print letters as well as numbers.   An IBM printer which handles numbers only will print at a rate of about 200 digits per second.   It has been estimated that gang printers with bars actuated by pneumatic or hydraulic instead of electromechanical means could be made to operate at a rate of several thousand characters per second.

A gang printer in which the type is located on rotating drums rather than bars has been suggested.   By rotating the drums at a constant velocity, the need to accelerate the type is eliminated.   Printing is accomplished by causing a hammer to strike the paper against the wheel as the proper letter is passing by. Since the hammer has to move only a short distance, it can be brought in place very quickly so that the operating speed of such a gang printer should prove to be quite high.

Flash Photographic Printers.   If extremely high speeds are required in an output printer, photographic techniques can be exploited.   They are limited in top speed by the maximum rate at which reasonably intense light sources can be modulated.   At the present time this limit is around 1,000 flashes per second with small Edgerton-type flash tubes employed in commercial stroboscopes and mentioned on page 320.   A compromise between flashing rate and illumination is necessary to prevent overheating when a flash tube is used repetitiously.   Several models of this tube are produced by Sylvania Electric Products, Inc., and the General Electric Co.   The latter is experimenting with a flash tube which promises to be smaller and faster than the coil-type lamps.

The Numeroscope.   A high-speed, photographic recording device called the *Numeroscope*[25,26] for use with a computer has been developed at Harvard.   Each digit in a number is formed on the face of a cathode-ray tube, and a picture is taken of all

the tubes in a row. In the system proposed, a 1-inch cathode-ray tube will be provided for each of 20 decimal digits. A group of 10 generators produce the electrical wave forms required to produce each Arabic numeral on a cathode-ray-tube screen, and these signals are gated to the proper tubes by the output of the computer. The 20 tubes are arranged side by side in a light-proof box with a shutterless camera which is aimed and masked to photograph a row of numbers across the cathode-ray tubes. After a frame of numbers is photographed and has disappeared from the screens, the film is advanced a short distance in readiness for the next exposure.

Most of the work reported on the Numeroscope consisted in the development of the electronic circuits required to form the figures. It was decided after an analysis that it would be very tedious to generate the figures with a combination of the required sinusoidal components; hence nonlinear circuits were devised with which to operate on sine waves of 2,000 and 5,500 cycles per second. Each figure is broken down into straight lines and sections of ellipses which are generated in turn and correctly positioned on the face of the tube. For example, the figure 8 is a small ellipse which is alternately moved up and down so that it appears as one ellipse on top of another. Half-wave-rectified signals are used to produce the figure 7, one leg of which is produced during one half of the cycle and the other leg during the alternate half.

The smallest figure which can be photographed is a function of the graininess of the film, and the exposure time is a function of the emulsion speed. The speed of this device in number of digits which can be photographed in a second can be determined from a knowledge of these two factors. Recent information[27] indicates that writing speeds of 270 inches per microsecond can be photographed on a P-11 screen with high accelerating potentials using an $f/1.5$ lens and E. K. Linagraph Pan (5244).

ERA Rapid Printer. At the time of this writing, ERA is developing a high-speed printing device under contract with the National Bureau of Standards. In this device, as in the Numeroscope, characters are produced on a cathode-ray tube screen by means of combinations of signals applied to the deflection plates. A technique based on magnetic recording principles is used to generate the signals.

The Monoscope. A piece of equipment commonly used in television transmitting stations is known as a *monoscope*.[24,29] This device is a small Iconoscope television transmitting tube with a picture or a diagram fixed on the inside of the envelope. A plate with a fairly high secondary-emission ratio, such as aluminum with its normal coat of oxide, is used in place of the mosaic. A figure is painted on the plate with a material, such as carbon ink, having a contrasting secondary-emission ratio. As the plate is scanned with an electron beam, a secondary-emission-current pattern of the diagram is produced. A device similar to the Harvard Numeroscope has been contemplated using a bank of 10 monoscopes each having a decimal digit built in. All could be operated from common sweep generators as could the Kinescope tubes on which the numbers are made visible. The output register would serve to key the proper monoscope into a row of Kinescope tubes arranged to be photographed as in the Numeroscope.

Other Digital Recorders. The use of the Selectron as a digital-to-analog converter was suggested on page 395. An image plotted on a Selectron used as a converter could be photographed in order to make a permanent record of computer results. In order to keep the trace on the screen for a time long enough to be photographed, it might be necessary to introduce the same set of values cyclically for a short period of time.

Electrical pulse-operated counter wheels also may be photographed to provide a permanent digital record.

15-5-2. Transcription from Automatic Analog to Written Form. It is often desirable to record the output of an analog machine for future reference and study. The output of the most common analog devices is a voltage or a shaft rotation. Various ways of recording these outputs are discussed in this section.

Electrical Output Recorders. An important type of electrical analog device is the d-c or a-c network analyzer. The output of a network analyzer is the current, power, or voltage at a number of points in the network, the number being dependent upon the type and size of the network under analysis. As an aid to reading and transcribing to paper a large number of voltages, a recording voltmeter, such as the Brown Recorder, may be employed. These devices contain servosystems which position a pen on a linear track perpendicular to the direction in which a

long sheet of paper can be moved in discrete steps.[30] The zero-voltage position of the pen and the scale factor may be adjusted so that any voltage of sufficiently high level can be recorded. If voltages are to be measured at 50 different places in the network analyzed, a probe may be applied to each of these in turn and the voltages recorded along 50 different axes on the paper. If some parameter is changed, the paper can be rolled back and the entire process repeated. Each set of dots, when connected or otherwise identified, results in a plot of values along some route through the network.

The output of the Westinghouse Transient Analyzer is a pair of cyclical voltages, displayed on the screen of a cathode-ray oscilloscope to represent the wave form of a transient quantity. One of the cyclic outputs is usually a saw-tooth voltage to represent a time base. Still photographs are taken of the output curves for permanent records.

The output of an electronic differential analyzer is also one or more voltages which vary in time. A display of an output voltage along one axis of a cathode-ray screen can be photographed with a shutterless camera to produce a continuous curve as a function of time. Similarly, output voltages may be recorded on an electromagnetic oscillograph.

Mechanical Curve Tracers. The output of a mechanical differential analyzer usually consists of the rotation angle of two or more shafts. A plot of one output as a function of another (usually machine time) can be obtained by arranging a pair of drive systems so that one output drives a marking instrument, such as a pencil, in one direction and the second output drives either a paper or the pencil in a direction perpendicular to the first. Thus, the pencil traces values on the paper in cartesian coordinates. A convenient mechanical arrangement is to place the paper over a cylindrical drum and have one drive cause the drum to rotate around its axis. When polar plots are desired, the pencil is moved in and out from the origin, and the paper is caused to rotate about the origin.

Since the output mechanism of the integrators used in a differential analyzer will slip if loaded heavily, a servosystem is usually required to drive the curve-plotting device. The requirements for the servosystem are low static and dynamic error and good stability, or freedom from overshooting. The

marking device may be a pencil, a pen, or an electric scribe, used with paper which is marked when current is passed through it. A paper of this type is called *Teledeltos paper*.

Counter-wheel Printers.   It may be desirable to sample the output of an analog quantity occasionally and write its numerical value.   The first step in this process is to convert the continuous quantity to digital form and then to print the digital quantity by one of the methods described in Sec. 15-5-1.   In the first MIT differential analyzer, for example, the output was on a mechanical counter with a printing wheel on the dials.   At discrete intervals, a paper was pressed against these continuously rotating wheels, and a slightly blurred print of the representative numbers was obtained.   The top speed of the commercial continuous-input, Veeder-Root counters with printing wheels is 17 counts per second.   In the later model of the MIT differential analyzer, a bank of rotary counter switches is used to operate an electric typewriter, and a record of shaft position is typed from time to time.

15-6. Radix Converters

It may be recalled from descriptions of various arithmetic elements that circuits are more simple for arithmetic operations in the binary system than in the decimal system.   Furthermore, most of the digital computing elements which have been used to date are employed in an on-off, or binary, fashion.   Binary numbers, however, have more than three times as many digits as decimal numbers and are much less convenient for human operators to deal with.   Therefore, it is usually desirable to have operators enter decimal numbers into a machine.   Some thought has been given to the use of octal numbers at the input of a machine because of the ease of converting them to binary form. In any case, unless the operator inserts binary numbers, a change in radix takes place at some level in a digital computer.[31]

In the Harvard Mark I and the ENIAC, decimal arithmetic is performed, and the only radix conversion is at the switches and flip-flops, which represent the decimal digits in a binary manner. In the Harvard Mark II, the UNIVAC, and other computers, each decimal digit is represented by a group of four binary digits. Such a system is known as a *coded-decimal system*.   A method of

conversion of decimal numbers to a binary-coded system is described on page 289.

This section includes a discussion of three types of radix conversion used in present or contemplated computers. They are the decimal-to-binary and binary-to-decimal conversion, using binary equipment, and the coded-decimal conversion.

15-6-1. Decimal-to-binary Conversion Using Binary-arithmetic Elements. In the conversion of decimal numbers to binary numbers, the process varies somewhat depending upon whether the numbers are whole numbers with no fractional part, are fractional numbers with no whole part, or are a combination of the two. However, the general idea of the process is illustrated by a single case.

To convert whole decimal numbers to binary form using binary equipment, one might first find the binary equivalent of each digit. This process can be performed in a matrix similar to the one illustrated in Fig. 4-3b on page 43 and described as an octal-to-binary converter. Each digit might then be multiplied by the binary notation for 10 a number of times corresponding to the power of 10 with which it was associated in the original decimal number, and all the products summed. For example, to convert 132 to binary form, the digits are first converted to 0,001, 0,011, and 0,010, respectively. The first of these is next multiplied twice by binary 10, or 1,010, the second once, and the third not at all. The resulting numbers are 1,100,100, 11,110, and 0,010, which are added to produce 10,000,100, which is the binary expression for 132.

Since the only two arithmetic steps required to accomplish this conversion are multiplication by 1,010 and addition, a complete arithmetic unit is not actually required to perform this work. Therefore, either the automatic conversion of input data can be accomplished in the arithmetic unit of the machine at the expense of some computing time or it can be accomplished in a preliminary conversion unit at the expense of some extra equipment. In a modern computer, the time lost in the first case usually will be negligible.

If the conversion is accomplished in a separate unit previous to the time when the problem is to be solved, there is a possibility of checking input data after conversion, thereby checking the conversion process as well as the operator.

15-6-2. Binary-to-decimal Conversion Using Binary-arithmetic Elements. To convert a whole binary number to the decimal system, it is necessary only to divide successively by 1,010, the binary equivalent of 10, and to convert the remainders to decimal digits in a matrix which is the inverse of the one required to convert from decimal to binary form. As an illustration, consider the transformation of the binary number 10,000,100 into the decimal system. The form of the division required is as shown:

$$
\begin{array}{r}
1101 \\
1010\overline{)10000100} \\
\underline{1010} \\
1101 \\
\underline{1010} \\
1100 \\
\underline{1010} \\
0010 \qquad \text{Remainder}
\end{array}
$$

$$
\begin{array}{r}
1 \\
1010\overline{)1101} \\
\underline{1010} \\
0011 \qquad \text{Remainder}
\end{array}
$$

$$
\begin{array}{r}
0 \\
1010\overline{)1} \\
\underline{0} \\
0001 \qquad \text{Remainder}
\end{array}
$$

The three remainders are arranged in inverse order and converted to decimal notation with another matrix circuit similar to the one illustrated in Fig. 4-3a on page 42. Thus 0,001, 0,011, 0,010 becomes 132, which may be recognized as the original decimal number.

Again, only a limited amount of arithmetic equipment is required to perform the conversion, so there is a choice to make between programming the conversion through the main arithmetic unit or providing a converter in the output of the computer. Since the problem is again one of equipment economy versus economy of time, the method of solution chosen will depend on

both the type of problem which the computer is expected to solve and on the computer development budget.

15-6-3. Coded-decimal Converters. When a coded-decimal system is used, the conversion process is shorter by one step than a binary-decimal conversion. Since each decimal digit is
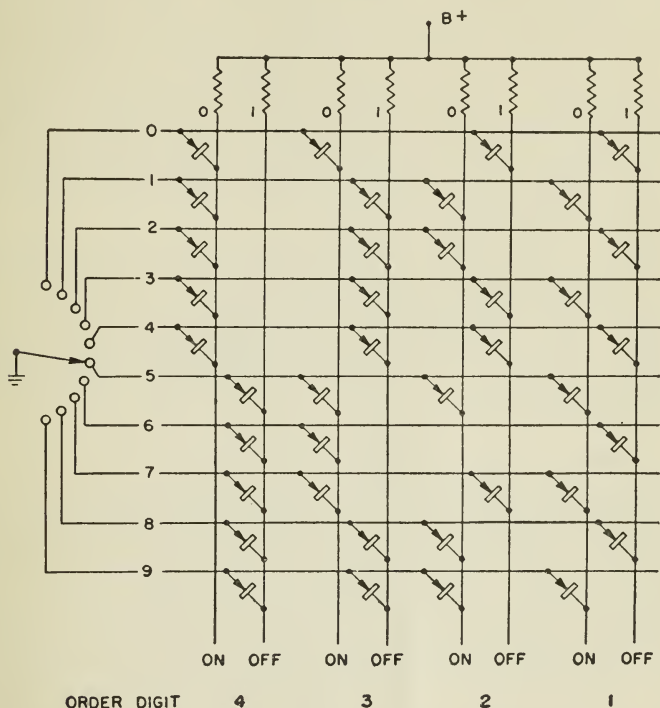
FIG. 15-5. Diode matrix for converting from decimal to *excess-3* coded decimal notation.

represented by a four-digit binary number, the conversion may be made on a simple matrix arranged to give the coded-decimal equivalent of each digit in a decimal number. A converter of this type is illustrated in Fig. 15-5.

The switch at the left-hand side of Fig. 15-5 is set on the decimal digit 5. According to the diode current-flow convention adopted in this text, this position of the input switch causes the positive voltage supply to be grounded through the resistor labeled 1, of the left-hand pair at the top of the diagram. This

signifies that the fourth-order binary digit of the number into which 5 is to be converted is 1.   Of the second pair of resistors at the top of the diagram, representing the third-order binary digit, the one labeled 0 is grounded.   Hence, the third-order digit is 0. Similarly, in both remaining pairs the 0 side is grounded.   The binary representation is therefore 1000.   This is the binary representation for 8 and is therefore the proper representation for the decimal number 5 in an *excess-3* code.   A study of the matrix shown in Fig. 15-5 will demonstrate to the reader that if any of the other input switch positions is selected a different binary number will be represented and each of these binary numbers will be the proper representation of the number on which the 10-position switch is set, plus 3.

It may be of interest to compare this matrix with those shown in Figs. 4-3*a* and 4-3*b* on pages 42 and 43.

There are other systems for coding decimal numbers in binary notation.   The *excess-3* code has been used as an example of one of these methods.   It is the method of number representation used in the UNIVAC.

15-7. Speed-changing Devices

From a consideration of the numerous forms of data which are suitable for computations, it is evident that data may be generated at vastly different speeds.   Likewise, in a computer, certain operations can be performed faster than others.   The most efficient machine design for the handling of a given problem is one which allows a continuous flow of data from beginning to end.   This may require parallel operation of some of the slower units so that data may be treated as rapidly as in a faster unit. It may also require the use of intermediate storage mediums on which to store the data until time for the next operation.   Data may be accumulated on some of these mediums at one rate of speed and released at another.[32-34]

A situation in which speed changing is obviously needed is the entry of manually prepared data into a high-speed machine. If several operators can prepare storage mediums for relatively high-speed automatic entry of data into the computer, a great deal of machine utilization time can be saved.   Likewise, if the computation results are temporarily stored on an intermediate medium which can be removed from the machine, the data may

be transcribed manually or automatically at a suitable rate without interfering with machine operation on a new problem. Several transcribers can then operate on the mediums simultaneously, thus providing compensation for the difference in speed between the output of data and printing. Also, use of intermediate storage mediums allows problem preparation at various locations; the problems may then be sent to a central location for actual computation.

15-7-1. Punched Cards and Punched Tape. Cards and tape may be punched by several operators simultaneously and then fed into an automatic device at top automatic speed. Their top speed, however, is somewhat low in comparison with the operating speed of an electronic computer. For example, punched cards are fed into the ENIAC only at the rate of one every half second, but arithmetic operations can be performed on the data in a few milliseconds. In problems which require a large number of arithmetic steps after each unit of data is inserted, this mismatch in speed might not be a great handicap. (As indicated in Sec. 10-3-1, the ENIAC, as currently used, is not limited by its card-handling speed.)

In the Bell Telephone Laboratories machines which use teletype tape as the input medium, a seven-digit decimal number, or command, may be entered into the machine in 0.2 second, added in 0.3 second, and multiplied in 1.0 second. It has been found from experience that numbers usually reach the machine at about the proper rate, but the input of commands is somewhat slow. This indicates roughly that input time should be about one-fifth of a multiplication time.

15-7-2. Magnetic Mediums. Several characteristics of magnetic mediums make them desirable for use with speed changers. Data may be recorded on them at low speeds and removed at fairly high speeds, and they can be erased and used over and over again. The original cost and volume per digit is relatively low, and it is easy to correct errors made by operators in preparing magnetic mediums. One disadvantage in their use is the extra equipment required to read from magnetic mediums when they are moved at low speeds.

It is estimated that a 10-digit decimal number could be entered into a machine from magnetic mediums in 2 milliseconds and that numbers could be recorded on the tape at the same rate.

15-7-3. Photographic Film. Like magnetic mediums, photographic film is another medium which can be recorded on at either a fairly high speed or a low speed and can be read over a similar range of speeds. Also, it is commercially available in several widths; driving equipment is well developed; and data can be stored compactly. Since it can be used only once and since it is fairly expensive to supply and process, its use seems to be most suitable in places where it will be referred to repeatedly. Thus, film might be used to provide rapid access to tables of functions. Film is also interesting as an output medium because decimal numbers written on it at high speeds can be printed on photographic paper in the form of tables of output data. It is also versatile in that continuous curves as well as numbers may be recorded.

REFERENCES

Mechanical Counters

1. Bush, V., A New Type of Differential Analyzer, *Journal of the Franklin Institute*, Vol. 240, No. 4, pp. 292–295 (October, 1945).
2. Murray, F. J., *The Theory of Mathematical Machines* (King's Crown Press, New York, 1948), pp. I-1–I-9.

Film-type Converters

3. Jones, F. D., *Ingenious Mechanisms for Designers and Inventors*, (Industrial Press, New York, 1936), Vol. I, pp. 71–77, Vol. II, pp. 164–188.

Pulse Code Modulation

4. Black, H. S., and J. O. Edson, Pulse Code Modulation, *Transactions of the American Institute of Electrical Engineers*, Vol. 66, pp. 895–899 (1947).
5. Carbrey, R. L., Decoding in P.C.M., *Bell Laboratories Record*, Vol. XXVI, No. 11, pp. 451–455 (November, 1948).
6. Goodall, W. M., Telephony by Pulse Code Modulation, *Bell System Technical Journal*, Vol. XXVI, No. 3, pp. 395–409 (July, 1947).
7. Greig, D. D., Pulse-Count Modulation, *Electrical Communications*, Vol. 24, No. 3, pp. 287–296 (September, 1946).
8. Meachan, L. A., and E. Peterson, An Experimental Multi-Channel Pulse Code Modulation System of Toll Quality, *Bell System Technical Journal*, Vol. XXVII, pp. 1–43 (January, 1948).
9. Oliver, B. M., J. R. Pierce, and C. E. Shannon, The Philosophy of Pulse Code Modulation, *Proceedings of the Institute of Radio Engineers*, Vol. 36, No. 11, pp. 1324–1331 (November, 1948).
10. Sears, R. W., Electric Beam Deflection Tube for Pulse Code Modulation, *Bell System Technical Journal*, Vol. XXVII, pp. 44–57 (January, 1948).

11. Unsigned, Coded Pulse Modulation Minimizes Noise, *Electronics*, Vol. 20, No. 12, pp. 126–131 (December, 1947).

Rotation Converters

12. Brown, G. S., and D. P. Campbell, *Principles of Servomechanisms* (John Wiley & Sons, Inc., New York, 1948).
13. Bush, V., A New Type of Differential Analyzer, *Journal of the Franklin Institute*, Vol. 240, No. 4, pp. 292–295 (October, 1945).
14. Hall, A. C., Applications of Circuit Theory to the Design of Servomechanisms, *Journal of the Franklin Institute*, Vol. 242, No. 2, pp. 279–307 (October, 1946).
15. Sobczyk, Andrew, Stabilization of Carrier Frequency Servomechanisms, *Journal of the Franklin Institute*, Vol. 246, No. 1, Nos. 2 and 3; Part I: pp. 21–43; Part II: pp. 95–121; Part III: pp. 187–213 (July, August, September, 1948).

Conversion of Data from Written to Automatic Digital Form

16. Bush, G. L., Photoelectric Transmitting Typewriter, *Electronics*, Vol. 20, No. 4, pp. 96–99 (April, 1947).
17. Dudley, H., Remaking Speech, *Journal of the Acoustical Society of America*, Vol. 11, p. 169 (May, 1939).
18. Dvorak, A., There Is a Better Typewriter Keyboard, *National Business Education Quarterly*, Vol. XII, No. 2, pp. 51–58 and 66 (December, 1943).
19. Hasley, R. J., and J. Swaffield, Analysis Synthesis Telephony with Special Reference to the Vocoder, *Journal of the Institution of Electrical Engineers*, Vol. 95, Part III, pp. 391–410 (September, 1948).
20. Watson, E. F., Fundamentals of Teletypewriters Used in the Bell System, *Bell System Technical Journal*, Vol. 17, pp. 620–639 (October, 1938).

Conversion of Data from Written to Automatic Analog Form

21. Greenwood, Holdam, and MacRae, *Electronic Instruments*, *Radiation Laboratory Series* (McGraw-Hill Book Company, Inc., New York, 1948), pp. 362–372.
22. Hazen, H. L., J. J. Jaeger, and G. S. Brown, Automatic Curve Follower, *Review of Scientific Instruments*, Vol. 7, pp. 353–357 (September, 1936).
23. Walther, A., *Development of Mathematical Instruments in Germany 1939–1945* (translated by British Naval Gunnery Mission). Pamphlet.

Transcription from Automatic Digital to Written Numerical Form

24. Burnett, C. E., The Monoscope, *RCA Review*, Vol. 2, pp. 414–420 (April, 1938).
25. Fuller, Harrison W., The Numeroscope, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 238–247 (1948).
26. Fuller, H. W., Numeroscope for Cathode Ray Printing, *Electronics*, Vol. 21, No. 2, pp. 98–102 (February, 1948).

27. Mansberg, H. P., A New, Versatile Camera for the Cathode Ray Oscillograph, *Oscillographer*, Vol. 10, No. 4, pp. 6–8 (October–December, 1948).

28. Sparks, S., and R. G. Kreer, Tape Relay System for Radio-Telegraph Operation, *RCA Review*, Vol. 8, No. 3, pp. 393–426 (September, 1947).

29. Zworykin, V. K., and G. A. Morton, *Television* (John Wiley & Sons, Inc., New York, 1940), p. 327.

Transcription from Automatic Analog to Written Form

30. Kane, J. L., *Conversion of Brown Recorder*, unpublished report (Douglas Aircraft Co., Inc., El Segundo, Calif., September, 1947).

Radix Converters

31. Mauchly, John W., *Theory and Techniques for the Design of Electronic Digital Computers*, Vol. III (Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, June, 1948), pp. 25-1–25-8.

Speed-changing Devices

32. Alexander, S. N., Input and Output Devices for Electronic Digital Computing Machinery, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 248–253 (1948).

33. Alt, F. L., A Bell Telephone Laboratories Computing Machine, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 21, pp. 1–13 (January, 1948).

34. Sheppard, C. Bradford, Transfer between External and Internal Memory, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 267–273 (1948).

SPECIAL TECHNIQUES AND EQUIPMENT FOR POSSIBLE USE IN COMPUTING SYSTEMS

16-1. Introduction

As problems of greater complexity are presented for solution on high-speed computers, the question arises as to what speeds and improvements in computing techniques can be expected in the future.   To a large extent the answer to this question depends upon how well the components and techniques of other fields of endeavor can be utilized in computing systems.   Although it is difficult to predict the exact trend of future development in automatic computing, it seems clear that higher pulse rates are desirable and that greater simplicity, reliability, and tube economy will be sought.   Attempts are also being made to find dependable elementary potential digital computing components and storage mediums which have more than two stable states so that arithmetic operations with radices higher than 2 can be performed in a straightforward manner.   Success in these attempts will increase the capacity of storage mediums and may possibly increase the speed of computation.

16-2. New Techniques

Digits are stored in sonic delay lines as pulses of a high-frequency carrier voltage.   Likewise, if low-velocity waveguides are used as a storage medium, an ultrahigh-frequency carrier voltage is required.   It may be convenient to transfer digits through other sections of a computer as high-frequency pulses. This practice is followed in television and in radar receivers, where pulses are partially amplified in intermediate-frequency amplifiers rather than totally in video amplifiers, since the inter-stage coupling problem is less complicated for an amplifier of 30 ± 5 megacycles than for an amplifier with half the band pass yet requiring d-c amplification.   The cost advantage of high-frequency components in comparison with video circuit

components may be increased when printed-circuit techniques are developed further. Also, high-frequency circuits are more convenient when band-pass characteristics are desired.

The field of wide-band-amplifier design has recently had some special developments, of which the *distributed amplifier* is an example. This type of amplifier has many operational features in common with the traveling-wave tube, which is discussed below. However, it is different in that it uses standard components. When amplifier tubes are cascaded in the conventional manner, the shunt capacities of parallel tubes increase in the same ratio as the gain. Therefore, it is not possible to realize a gain-bandwidth product which is any greater than that of a single tube. In other words, as more tubes are added, gain is increased at the expense of bandwidth. In distributed amplifiers, the tubes are separated by inductances, and the resulting system simulates a transmission line. As a result, the gains of the tubes are increased and the distributed capacities are not. Therefore, the bandwidth of the whole amplifier can approach the figure of merit, or unit gain bandwidth, of a single tube.

A development-type amplifier employing 14 tubes has a gain of 18 decibels over the frequency band of 0 to 200 megacycles. An extension of this technique may be applicable to computing circuits.

The degree of compactness with which large numbers can be represented is a function of the radix of the number system being used. The largest number n which can be represented with a digits of radix r is $r^a - 1$. Since a given number can be represented by a smaller number of digits in a larger radix system, there is a resulting saving in time when the higher radix number is stored in serial storage, and a resulting saving in number of channels when the number is stored in parallel channels. A saving of 58 per cent can be gained in going from a binary to a ternary system. A smaller percentage gain is realized in going from a radix 3 to a radix 4 system.

One means of increasing the radix of computing components is to represent individual digits by corresponding frequencies. For engineering convenience, those frequencies might be chosen which could be stored directly in the storage medium. This system is also interesting because some of the elementary arithmetic operations may be accomplished by beating frequencies

together. It also has the advantage that components which are essentially video amplifiers in on-off pulse techniques may be replaced with elements which are similar to intermediate- or high-frequency amplifiers. A certain amount of background development has already been done in frequency-shift telegraphy.[1,3,4]

It is interesting to note that the increase in access time and storage capacity of a medium by use of higher radix frequency-shift methods is at the expense of greater bandwidth in the storage medium and the computing circuits. Recent developments indicate that it may be easier to develop circuits with greater bandwidth than to increase the cycling rate of an EPDCC such as a flip-flop circuit.

16-3. Special Electronic Tubes and Crystals

The electronic elements described in this section are of interest as high-frequency amplifiers and as means for increasing the simplicity, economy, and dependability of circuits.

16-3-1. The Traveling-wave Tube. An electronic tube has recently been developed at the Bell Telephone Laboratories[6,7] which provides amplification over a very wide band of frequencies. The amplification results from the interaction between an electron beam and a slowly moving electromagnetic field. The field is produced by a helical transmission line; the phase velocity of the field is reduced to about one-thirteenth the velocity of light. The property of helical delay lines to reduce phase velocity was discussed on page 351. The electron beam in the tube is passed along the axis of the helix at a velocity slightly higher than that of the field, and energy is transferred from the electron stream to the traveling field. In an early model[5] an input signal was amplified 23 decibels as it traversed the length of the helix. The helical line replaces the resonant circuits in conventional high-frequency amplifiers, and, as a result, the bandwidth for a 3-decibel decrease in amplification is 800 megacycles at a mean frequency of 3,000 megacycles.

16-3-2. Klystrons. Another high-frequency tube, the *klystron*, has been used extensively as a local oscillator in microwave radio and radar receivers.[8,9,11,12] In the klystron an electromagnetic field is produced in a resonant cavity; the velocity of an electron beam traversing the cavity is modulated by *bunching*

electrodes so that the beam of electrons is bunched to form current pulses. The dimensions are of the order of a wavelength at the frequencies handled, so that conventional klystrons become inconveniently large at frequencies much below 1,000 megacycles. Klystrons may be used as voltage amplifiers, power amplifiers, and oscillators with relatively high efficiency and power-handling capacity. Some recent work on klystrons used as generators of very short pulses has resulted in pulses of adjustable duration as low as 0.2 millimicrosecond.[10] The experiments were performed at a repetition rate of 210 megacycles with a klystron in which the resonant cavity had been replaced with a nonresonant line. With present-day circuits and production methods it is difficult to utilize pulses of duration shorter than 0.1 microsecond in computers, but the full capabilities of the klystron pulse generator may become useful in computers as pulse-handling techniques are improved.

16-3-3. Miniaturized Tubes. Counterparts for most of the common radio receiving tubes are now manufactured in miniature form.[14] These miniature tubes occupy about three-tenths the volume of conventional tubes of similar type and make possible a considerable reduction in size and weight of electronic equipment. Miniature tubes seem to be dependable and as well suited to computer systems as the larger tubes.

Another group of tubes called subminiatures has been developed. The diameter of these tubes is approximately that of an ordinary lead pencil, and flexible leads are used rather than base pins. The tubes are available in both the filament and the cathode types. A typical application is found in the modern hearing aid.

Manufacturers' catalogs of miniature tubes are available,[15] and a complete list of subminiature tubes has been compiled by the Signal Corps.[13] Miniature tubes are manufactured by RCA, Raytheon, and Sylvania; subminiature tubes are manufactured by Raytheon, Sylvania, and Sonotone.

16-3-4. Transistor. A new type of amplifier called the *transistor* was recently announced by the Bell Telephone Laboratories.[16-18] It seems likely that this device will simplify computer circuits considerably. The transistor is similar to the crystal rectifier which was described in Chap. 3 and consists of a block of germanium crystal with which two cat whiskers of

2-mil tungsten wire are in contact. These two wires are spaced 2 mils apart and form one side of the input and the output circuits, respectively. The germanium crystal is at the common connection between input and output. A gain of 20 decibels at 25 per cent efficiency was measured on the model announced. The transistor has an input impedance which is only 1 per cent of its output impedance, so that interstage coupling must necessarily differ from that used in vacuum-tube circuits. Transit time establishes a top frequency limit of about 10 megacycles for the transistor.

The advantages of this element are its small size, low power drain, sturdy construction, long life, and cool operation. It will probably be competitive with the electron tube in total cost per stage.

16-4. Amplitude- and Frequency-discrimination Circuits

The action of the gates, flip-flops, and relays which have been discussed previously depends upon the ability of these units to discriminate between low-level spurious voltages and high-level signals.[22] The units heretofore described discriminate on the basis of voltage amplitude. Equipment is also available which discriminates on the basis of frequency.

The possibility of representing digits by discrete frequencies is introduced in the second section of this chapter. In such a system, it may be desirable to convert the carrier frequencies into corresponding d-c voltage levels for arithmetic operations such as those described in Sec. 13-4-4 on amplitude adders. Conversion from frequency to voltage amplitude is accomplished by means of a frequency discriminator.

Frequency discriminators were first used in conjunction with automatic frequency controls and later in radio receivers using frequency modulation. There are many competing methods of frequency discrimination.[19-21,26] The better known of these are the Seeley-Foster and the Armstrong methods. Recent variations of the Seeley-Foster circuit are the ratio detectors,[23,27] and a variable bandwidth discriminator developed at MIT.[28]

A discriminator which is used at microwave frequencies is described by Pound.[24,25] This discriminator makes use of two waveguide arrangements known as *magic T*'s and a cavity which is resonant at the mid-frequency of the frequency-modulation

signal.   The cavity is not loaded, so that it may have a high Q. The output of this discriminator is the difference in rectified voltage at two points in the waveguide.   The relative phase of voltage received at these two points is such that the voltage difference is a function of the applied frequency.

16-5. Frequency-stabilizing Systems

Computers using cyclic storage mediums require pulse-repetition-rate stabilization in order to preserve synchronization. Accurately held high frequencies may be used to time lower frequency pulse generators.   Several circuits have been developed which perform the function of maintaining microwave oscillators at nearly constant frequency.   One circuit, developed by Pound at the MIT Radiation Laboratory,[31,32] uses the microwave discriminator described in the previous section as a frequency-sensing device.   The output of this discriminator is amplified and applied to an electrode of the oscillator through a negative-feedback loop, and the oscillator frequency is held near the frequency of the resonant cavity in the discriminator.

Another such device, developed at RCA,[30] employs the spectral line of a gas as the frequency reference with which the oscillator frequency is compared.   The voltage output of a comparison circuit is used in a negative-feedback, or servo, loop to control the frequency of the oscillator.   Multivibrator step-down circuits can be used to obtain lower frequencies from these high-frequency standards.[29]

16-6. Unconventional Circuit Fabrication

Interest in subminiature electronic equipment stems from its extensive and successful use in certain war weapons.   One of these weapons, the *VT fuse*, was made in mass production using techniques specially developed for rapid fabrication.   Mass-production methods known commonly now as *printed-circuit techniques* were used in making fuses for trench-mortar shells.[34]   Interest in these methods has been displayed by the electronic manufacturing industry, and subsequently extensive development work to expand and improve these techniques was undertaken.

16-6-1. Methods of Printing Circuits.   Brunetti and Curtis[34] have listed six general methods for the mass production of printed circuits:

1. *Painting.* A special conducting paint is applied to a prepared ceramic surface placed under a stencil. It is possible to wire the circuit components as well as to create the components with conducting paint. For example, a certain deposited thickness, width, and length of a special resistor paint may be made to represent a given value of resistance, while a change of any of the dimensions or of the paint itself may be made to represent a value different from the first. It is possible to print inductances and some capacitors by an extension of this method.
2. *Spraying.* This method is similar to the painting technique discussed above with the exception that the coatings are sprayed on the ceramic plate placed under the circuit stencil.
3. *Chemical deposition.* It is possible to construct an electronic circuit by electrolysis. The wiring is in the form of a metallic deposit on the special ceramic plate.
4. *Vacuum process.* Less used but interesting techniques for printing circuits are those of cathode sputtering and metallic evaporation.
5. *Die-stamping.* This method allows the interconnecting wires of the circuit and some high-frequency inductances to be die-punched out of a solid sheet.
6. *Dusting.* The metallic conductors may be dusted with a small brush onto a ceramic plate under a stencil and then fused on the surface.

16-6-2. Examples of Commercially Available Units. A typical example of the commercial application of the printed circuit is to be found in a subminiature audio amplifier called the *Ampec*. This unit is available in production quantities from the Centralab Division of the Globe Union Co., Inc., Milwaukee, Wisconsin. The Ampec is a three-tube amplifier consisting of two voltage amplifiers followed by a power amplifier. The unit does not include the volume control, output transformer, or battery power unit; it is fabricated with printed-circuit techniques entirely and measures 2.250 inches long, 1.15 inches wide, and weighs 0.63 ounce. This unit has a voltage gain of 4,000 when used with a 30-volt plate battery and a 1.5-volt filament cell. The plate-battery drain is approximately 1 milliampere and the filament drain is 40 milliamperes. Although the unit is somewhat micro-

phonic, it is used extensively as the basic unit for several commercial hearing aids.

The manufacturers of the Ampec have also made available printed-circuit-component combinations, such as interstage coupling units, that save space in the more conventional radio applications.

Another printed-circuit technique which has been greatly emphasized is the die-stamping process. The Franklin Corporation manufactures an inductance loop stamped out of sheet copper which serves as a high-Q tuned element of the table-model radio receivers.

16-6-3. Features of Printed Circuits. The major features of printed-circuit techniques are (1) size reduction, (2) uniformity, and (3) economy for large numbers of units.

Size reduction has been made possible by printed circuits to the extent that the volume occupied by the wiring and supporting plate is only a small percentage of the total volume of the unit. The development of truly subminiature units is dependent not upon further size reduction in circuitry but rather upon the development of smaller inductors, transducers, and power supplies.

The application of the printed-circuit technique to the fabrication of electronic equipment has placed the manufacture of this equipment in the true mass-production class, where machine operations can be accurately controlled and the human errors generally encountered in wiring work can be minimized.

A great advantage in the use of printed circuits is the considerable reduction in man-hours required to fabricate the units. A further saving is made in the man-hours of inspection time required to check the work. Obviously, the printed-circuit technique requires specialized and complicated manufacturing equipment to fabricate printed-circuit assemblies. The expense involved for such equipment is not justified for small quantities; large production schedules are necessary before the techniques can be profitably applied.

16-6-4. Plug-in and Potted Circuits. The fabrication of printed-circuit subassemblies suggests the use of plug-in units for these assemblies. Tests made at the Bureau of Standards showed that it was possible to *pot* printed circuits in a special casting resin in such fashion as to permit the plugging in of the

complete subassemblies. By the use of special resins, the frequency of operation of the printed circuit may be extended into the ultrahigh-frequency region without a noticeable decrease in efficiency owing to the enclosure in the casting resin. The use of casting resins, however, requires critically controlled conditions, and consequently mass production of potted plug-in units depends upon the development of complex process controls.

REFERENCES

Improved Techniques

1. Darey, J. R., and A. L. Matte, Frequency Shift Telegraphy—Radio and Wire Applications, *Bell System Technical Journal*, Vol. 27, No. 2, pp. 265–304 (April, 1948).
2. Ginzton, Edward L., *et al.*, Distributed Amplification, *Proceedings of the Institute of Radio Engineers*, Vol. 136, No. 8, pp. 956–969 (August, 1948).
3. Hatfield, L. E., Frequency Shift Radio Transmission, *Proceedings of the Institute of Radio Engineers*, Vol. 36, No. 1, pp. 116–120 (January, 1948).
4. Vanderlippe, R. A., Frequency Shift Radio Teletype in World War II, *Bell Laboratories Record*, Vol. XXV, No. 4, pp. 442–446 (December, 1947).

Traveling-wave Tubes

5. Kompfner, Rundolf, The Traveling Wave Tube as Amplifier at Microwaves, *Proceedings of the Institute of Radio Engineers*, Vol. 35, No. 2, pp. 124–127 (February, 1947).
6. Pierce, J. R., and L. M. Field, Traveling Wave Tubes, *Proceedings of the Institute of Radio Engineers*, Vol. 35, No. 2, pp. 108–110 (February, 1947).
7. Pierce, J. R., Theory of the Beam Type Traveling Wave Tube, *Proceedings of the Institute of Radio Engineers*, Vol. 35, No. 2, pp. 111–123 (February, 1947).

Klystrons

8. Hamilton, D. R., J. K. Knipp, and J. B. H. Kuper, Klystrons and Microwave Triodes, *Radiation Laboratory Series*, Vol. 7 (McGraw-Hill Book Company, Inc., New York, 1948).
9. Harrison, A. E., Klystron Oscillators, *Electronics*, Vol. 17, No. 11, pp. 100–107 (November, 1944).
10. Hasted, J. B., Milli-Microsecond Pulse Generation by Electron Bunching, *Proceedings of the Physical Society*, Vol. 60, p. 397 (Apr. 1, 1948).
11. Pierce, J. R., and W. G. Shepherd, Reflex Oscillators, *Bell System Technical Journal*, Vol. 26, No. 3, pp. 460–682 (June, 1947).
12. Varian, R. H., and S. F. Varian, A High Frequency Oscillator and Amplifier, *Journal of Applied Physics*, Vol. 10, No. 5, pp. 321–327 (May, 1939).

Miniaturized Tubes

13. Evans Signal Laboratory, *Handbook of Subminiature Tubes* (Signal Corps Engineering Laboratories, Bradley Beach, N.J., July, 1947).
14. Green, N. H., Miniature Tubes in War and Peace, *RCA Review*, Vol. 8, No. 2, pp. 331–341 (June, 1947).
15. Radio Corporation of America, *RCA Quick Reference Chart Miniature Tubes* (RCA Tube Department, Harrison, N.J.).

Transistor

16. Bardeen, J., and W. H. Brattain, The Transistor, A Semi-Conductor Triode, *Physical Review*, Vol. 74, No. 2, pp. 230–231 (July 15, 1948).
17. Becker, J. A., Transistor Characteristics, *Electronics*, Vol. 22, No. 1, p. 132 (January, 1949).
18. Fink, D. G., and F. H. Rockett, The Transistor—A Crystal Triode, *Electronics*, Vol. 21, No. 9, pp. 68–71 (September, 1948).

Amplitude- and Frequency-discrimination Circuits

19. Barco, A. A., C. M. Kimball, and S. W. Seeley, Generation and Detection of FM Waves, *RCA Review*, Vol. VI, No. 3, pp. 269–286 (January, 1942).
20. Crosby, M. G., Communication by Phase Modulation, *Proceedings of the Institute of Radio Engineers*, Vol. 27, No. 2, pp. 126–131 (Feburary, 1939).
21. Foster, D. C., and S. W. Seeley, Automatic Tuning, Simplified Circuits, and Design Practice, *Proceedings of the Institute of Radio Engineers*, Vol. 25, No. 3, p. 289 (March, 1937).
22. Lewis, W. B., Electrical Counting (The Macmillan Company, New York, 1943), pp. 71–76.
23. Maurice, D., and R. J. H. Slaughter, F. M. Comparison Test between Phase Discriminator and Ratio Detectors, *Wireless World*, Vol. 54, pp. 103–106 (March, 1948).
24. Pound, R. V., *Microwave Mixers, Radiation Laboratory Series*, Vol. 16 (McGraw-Hill Book Company, Inc., New York, 1948), pp. 190–231.
25. Roder, Hans, Theory of the Discriminator Circuit for Automatic Frequency Control, *Proceedings of the Institute of Radio Engineers*, Vol. 29, No. 5, pp. 590–596 (May, 1938).
26. Seeley, S. W., Balanced Phase Shift Discriminators, *RCA Licencee Bulletin LB-611* (Mar. 15, 1946), *RCA Licencee Bulletin LB-645* (Sept. 15, 1945).
27. Terman, F. E., *Radio Engineers' Handbook* (McGraw-Hill Book Company, Inc., New York, 1943), pp. 585–588.
28. Tuller, W. G., and T. P. Cheatham, Jr., Adjustable Band with FM Discriminator, *Electronics*, Vol. 20, No. 9, pp. 117–119 (September, 1947).

Frequency-stabilizing Circuits

29. Davis, K. H., Multivibrator Stepdown by Fractional Ratios, *Bell Laboratories Record*, Vol. XXVI, No. 3, pp. 114–118 (March, 1948).

30. Hershberger, W. D., and L. W. Norton, Frequency Stabilization with Microwave Spectral Lines, *RCA Review*, Vol. 9, No. 1, pp. 38–49 (March, 1948).

31. Pound, R. V., Frequency Stabilization of Microwave Oscillators, *Proceedings of the Institute of Radio Engineers*, Vol. 35, No. 12, pp. 1405–1415 (December, 1947).

32. Tuller, W. G., Recent Developments in Frequency Stabilization of Microwave Oscillators, *Proceedings of the Institute of Radio Engineers*, Vol. 36, No. 6, pp. 794–800 (June, 1948).

Unconventional Circuit Fabrication

33. Brunetti, C., and W. S. Hinman, Radio Proximity Fuse Design, *Journal of Research of the National Bureau of Standards*, RP 1723, Vol. 37, pp. 1–13 (July, 1946).

34. Brunetti, Cledo, and Roger W. Curtis, Printed Circuit Techniques, *Proceedings of the Institute of Radio Engineers*, Vol. 36, No. 1, pp. 121–165 (January, 1948).

35. Greenwood, Holdam, and MacRae, *Electronic Instruments*, *Radiation Laboratory Series*, Vol. 21 (McGraw-Hill Book Company, Inc., New York, 1948), pp. 700–704.

CHAPTER 17

FACTORS AFFECTING CHOICE OF EQUIPMENT

17-1. Introduction

This chapter is concerned with the means by which computing equipment may be made accurate and reliable; there is also presented here a summary of the capacity and cost of storage mediums and an approximation of the storage, reading, and writing times to be expected. Although there are other factors affecting the choice of equipment, these cannot be summarized readily and are implicit in the foregoing discussion of equipment characteristics. Furthermore, the high-speed computing art is not in a mature enough state to allow an extensive comparison of equipment which has been thoroughly tested. However, the following material should indicate the orders of magnitude involved, and the degree of complexity required to meet given equipment specifications.

17-2. Reliability of Components

The operating time of a large-scale computer is very valuable; occasional failures, therefore, are expensive as well as annoying. There are at least two approaches to the problem of reducing equipment failure during operation. One method is to make each component so reliable that it will last longer than the period decided upon as the useful life of the equipment. Although it is difficult to guarantee absolute reliability, if the statistical life expectancy of the component is projected far enough into the future, the actual rate of failure should prove fairly low during, the early years of machine operation. As a second approach, means may be taken to extend as far as possible the operating period of equipment considered to have a limited life expectancy and to detect incipient failures before they occur.

It is not enough to life-test one or two models of a component under normal operating conditions to determine whether or not it is suitable for use in a computer. It is necessary as well to

ascertain that when manufactured parts are used together, any part having a value within the manufacturing tolerance range will work reliably with any combination of other components which are also within the prescribed range. Since it is impractical to test all possible combinations of values during design, circuits should operate with large factors of safety for voltages, pulse shapes, and repetition rates so that reasonable variations in components will not affect operation adversely. Furthermore, circuits should be designed so that failure of one unit does not overload other units and also cause their eventual failure. This condition may be difficult to carry out completely, but, at least, interdependent parts should be assembled in units so that all the parts which may be overloaded by the failure of one part are located together and can be replaced by exchanging units. Interchangeable chassis are also desirable so that repairing and periodic testing of units can be accomplished simultaneously with normal operation.

There are important factors in the philosophy of circuit design which also affect reliability. For example, time allowance between operations should be adequate to ensure correct sequences even when individual circuits lose their original adjustment. Allowances should be made for variations in signal voltages and in power-supply voltage. Direct coupling is often required between stages so that coupling capacitors will not become charged by large pulses and hold subsequent stages conducting longer than desired. Some means of detecting an instantaneous failure of the primary power source is required, since such a failure might otherwise cause an undetected error. The use of latch-type relays is desirable in a relay computer so that values will be held during the period between a failure and its repair. Finally, it should be recognized that many manufacturers do not supply life-expectancy data on mechanical apparatus and that expensive tests may be required to get this information.

In order to extend the life of individual components, thereby increasing operational reliability, it is necessary to follow recognized design practices in each particular field. Mechanical devices must be designed for good wearing qualities and with adequate factors of safety in strength. Electrical apparatus, likewise, must be used at voltages and currents well under rated

values. Cooling equipment must be provided wherever required in order to maintain safe temperatures in critical parts, and air conditioning is advisable for relay computers in order to minimize trouble from dust accumulations.

In order to present a comprehensive view of the types of component failure which limit the reliability of large-scale computer operation, a number of specific examples of component failure are presented. Some of them have been taken from the operational history of the ENIAC.

17-2-1. Vacuum Tubes and Crystals. It has been reported that in the ENIAC the most common source of trouble is in the 18,000 vacuum tubes. Every tube in the machine undergoes a routine operational check of some sort each week. The practice, for example, is to test each of the several hundred three-tube amplifiers each week as a unit and to replace all three tubes if one of these amplifiers is not operating satisfactorily. During the first eleven months of 1949 about 2,000 tubes were replaced each month. About half of these were actually bad at the time they were replaced.

An attempt was made to increase the tube life in the ENIAC by operating 6.3-volt filaments at 5.95 volts. Later tests indicated that it may be better to use a value between 6 volts and 6.3 volts. The filaments are left turned on for as long a period of time as possible, since several tube failures usually result from turning the equipment on. A large number of failures have resulted from open heaters, damaged cathodes, and near shorts in the tube framework.[3] It was found from examination of all the tubes which failed that the failure rate was considerably higher in tubes with folded filaments than in tubes with twisted filaments. Experience also indicates that low-voltage filaments last longer than high-voltage filaments of the same power rating. For this reason manufacturers provide series filaments in twin triodes of the industrial variety which are rated for a life of 10,000 hours.[4]

High plate-load resistances are used in the ENIAC in order to minimize the effect of changes in tube plate resistance. The average tube has lost emission noticeably after approximately 9,000 hours of operation, but actual failure by this cause can be minimized by replacing tubes periodically. It has been suggested that provision be made in a computer to vary the voltage of the

power supplies over a fairly wide range periodically. In this manner a large percentage of the incipient failures in tubes and other components could be made to occur.

In the ENIAC the operating frequency is varied, but not the voltages.

The practice of replacing a large number of tubes at once has the disadvantage that it may complicate subsequent trouble shooting inasmuch as most of the tube failures occur either within the first 250 hours of operation or after the 5,000th. The present ENIAC operating staff is not in sympathy with a policy of replacing a large number of tubes at once.

Characteristics other than those affecting the life expectancy of electron tubes are equally important in considering the reliability of these components. For example, in applications where low-level voltages are to be amplified, as in the amplifiers for the output of magnetic recording pickups, the microphonic characteristics of tubes must be considered by the designer.

Crystal diodes have been tested by the Eckert-Mauchly Computer Corporation under conditions typical of usage in electronic computers. The results of these tests have shown that the manufacturers' tests under other conditions are not necessarily valid for computer use.

17-2-2. Capacitors. The capacitor is an electronic component which is subject to deterioration and eventual failure.[1] The length of its life depends upon the applied voltage, the temperature, and the degree to which it is suited to its application. A statistical rule has been determined which shows that the minimum life to be expected of a group of capacitors before a large number fails is proportional to some power of the ratio of the applied voltage and the rated voltage. This exponent varies between 4 and 6 for conventional capacitors. The minimum life expectancy of capacitors is also a function of ambient temperature which should be carefully estimated before voltage ratings are established.

17-2-3. Resistors. The life expectancy of a resistor with a given power rating depends upon the actual power dissipated and the ambient temperature. The small resistors used in the ENIAC are operated at approximately one-half rated power and at a temperature of 110°F, as compared to about 100°F, on which the rating is based.[2] Wire-wound resistors are operated

at about one-fifth rated value in order to decrease their operating temperatures, and consequently the temperature to which nearby equipment is subjected.

About 70,000 resistors are employed in the ENIAC, and not more than five were known to have caused a machine failure in a 9,000-hour period of operation.

17-2-4. Fuses and Indicators. An individual fuse was used for each power-supply voltage in the ENIAC. The fuses are made by Western Electric and have a small switch which closes when a fuse opens. These switches are used to turn off all high voltages when one fuse goes out. In the IBM Automatic Sequence Controlled Calculator, neon lights are used to indicate shorted capacitors, open resistors, and low emission in tubes.

17-3. Possibilities for Checking

The ultimate goal in computer design is complete intrinsic reliability. Checking, or the detection and location of errors, should therefore be considered a means and not an end in itself. Usable solutions to problems are obtained only when the equipment is operating properly, and no amount of checking equipment will procure a correct solution. Errors in computation may be machine errors or operator errors, and checks are required for each. The amount of checking equipment and the frequency with which checks are programmed into the machine depend primarily upon the reliability expected of the machine. When errors are detected in the results, it is necessary to locate the source of these errors and to eliminate their causes. The detection and location of errors are discussed in this section with respect to both machine and human errors.

17-3-1. Machine Errors. A machine error may be caused by either of two types of equipment failure: (1) an intermittent, or (2) a persistent failure. The detection of components which fail intermittently in conventional equipment may be very difficult. In relay computers, for example, approximately 85 per cent of the failures are intermittent. The majority of the failures in the ENIAC also have been of the intermittent type.

Efficiency of operation of the ENIAC over any given time interval is closely related to the number of *intermittent* failures occurring during that interval.

It has been estimated that the ENIAC is actually in useful operation on problems about 50 per cent of the time and that, allowing for a complete repetition of each problem, it therefore turns out about one-fourth as many useful answers as could an ideal, perfectly reliable, machine with the same basic mathematical characteristics, which required no maintenance whatever. This is an excellent performance record.

In the Bell Telephone Laboratories machines, continuous automatic checking is done at the level of the individual relay. This is called *low-level checking*. In the ENIAC, checks are made of solutions and partial solutions to problems, and this is termed *high-level checking*. When a failure occurs in a Bell Telephone Laboratories machine, all digital results are held in their respective relays until the trouble is cleared, whereupon the computer continues to operate with the results obtained previously. Operation of the ENIAC is so fast, however, that it was not felt worth while to provide similar holding equipment, and results are recalculated after a failure.

Several types of high-level checks have been worked out for computing equipments,[7] and they will be described briefly as follows:

1. If a problem can be solved in two different ways in a computer, agreement between two answers so obtained constitutes a high-level check on the correctness of the solution.

2. A check on the equipment can be made by getting answers to the problem in some way other than using the same equipment. For example, part of a solution could be worked out on a desk calculator, or perhaps solutions to restricted cases might be found and checked against the machine solution.

3. In physical problems, the results are usually continuous, and the machine can be programmed to check high-order differences and cause an alarm to signal the operator if these vary too much.

4. Special problems with known results may be designed to test certain sections of the machine. In this manner the machine may be stopped at various stages of the solution and the machine results compared with the known results.

The machine operation at which the first error is made can be located in this manner.

5. To check for intermittent errors, it may be sufficient to repeat the same problem two or more times. The results of the two runs may be checked in a comparator and the presence of any discrepancy detected.

The ENIAC has indicator lights to represent the digits in each register. There is also provision for supplying the computer with one clock pulse at a time. Thus, an operator can run off the problem at low speed and observe the effect of each pulse on the quantities in the register, and the sources of error can be located quite accurately. This step-by-step operation is not feasible in a machine which uses dynamic storage mediums, since these generally depend upon a constant high-frequency pulse rate in order to maintain synchronism.

The advantages of a machine which handles data in a serial fashion have been cited previously. Since there is less equipment, errors due to a failure are likely to show up soon after the failure occurs, and the number of steps required to locate the source of the trouble thereafter is usually small. On the other hand, if the serial machine has dynamic storage, it cannot be slowed down readily for the step-by-step check just described. In any case, a systematic program of periodic checks is desirable so that errors which are not detected by low-level checks will be found within a reasonable time after they occur.

Perhaps the best example of a low-level check is that incorporated in the Bell Telephone Laboratories machines, which employ a biquinary representation of numbers.[8,9] In this representation, one group of five relays indicates the digits from 0 through 4, and a second set of two relays represents the digits 0 and 5. Thus each decimal is represented by one relay in each set, which allows a simple check to be made on the operation of the relays. Each one has extra contacts which are connected in an alarm circuit so that they indicate any situation in which more or less than one relay in each pair of sets is actuated simultaneously. Thus, the failure of one relay to operate because of dust between the contacts of another relay is detected. The operation of a relay which is not supposed to operate is also detected. These situations cause the equipment to stop operat-

ing until the trouble is cleared. The biquinary check fails when two compensating failures occur simultaneously, but this possibility proves to be very remote.

Another possibility for low-level checking is to provide two identical accumulators which operate on the same digits simultaneously. Anticoincidence circuits note any difference in condition between the two accumulators and stop the computation if such differences occur. This duplication might be fairly economical if the extra accumulator were used for shifting to the right or to the left in the accumulator, so that the shift is accomplished by transferring the number from accumulator 1 to accumulator 2 and readjusting accumulator 1 to agree with accumulator 2.

Duplicate systems and anticoincidence circuits may also be used at other critical units in order to detect incorrect operation.

In the computer proposed by the Raytheon Company,[6] self-checking and diagnostic equipment is provided throughout. It is estimated that the proposed checking system demands the use of about 20 per cent additional equipment. The checking covers all storage and transfer of words, all arithmetic operations, all selections made by the memory, and most other machine processes.

For example, to check the storage and transfer of numbers, R. M. Bloch has devised a system employing a check number associated with each number. This check number is merely a weighted sum of the binary digits in the number. When a new number is constructed in the arithmetic unit, or is introduced into the machine, it is provided with a check number, which is then stored with it. When the original number is withdrawn from storage, a new weighted sum is computed and compared with the old one which was stored with it. Any discrepancy between the new and old weighted sums indicates an error made during storage or transfer. These check numbers can also be used to check arithmetic operations.

17-3-2. Human Errors. Perhaps the ultimate computing machine will take data from measuring devices, automatically convert them to the desired form for the computer, perform the required computations, and deliver the written results in the optimum written form. Even though the operator is eliminated as an intermediate handler of the data, instructions must still

be prepared for the computer so that it will perform the proper operations These instructions are first formed in terms of the mathematical steps desired and finally are written on a permanent storage medium for automatic introduction into the computer. The intermediate process of expressing the mathematical steps in signals to the computing machine is known as *coding*. Coding may become a rather long and tedious process for some problems, and it may be advantageous to employ small computers to do the detail work in coding problems for the larger machines. In any case, it is possible to compare one coded medium with another, but the original assignment of instructions must be checked either mentally or by experimenting on a sample problem whose solution is known.

Two methods have been proposed for checking the work of the operator in writing the data and instructions on the input mediums.[5] As an example, consider the preparation of tapes with a keyboard. It is possible for two operators to prepare tapes at the same time. These two tapes are compared automatically, and only those marks which are identical on the two are admitted to the machine. Any difference in the two tapes stops the machine and signals the two operators, who must resolve the difference before entry of data into the machine can continue.

In a slight variation of this idea, one tape is made up in advance, and an operator simulates the preparing of the tape a second time. As he operates the keyboard the second time, a typed record of his work is prepared for visual inspection. His typing is also automatically compared with the data on the original tape, and all identical data are inserted in the final tape. A discrepancy in the two stops the machine and indicates the error to the operator. Since he has the manuscript before him, the operator can compare it with what he has just typed and immediately resolve the difference.

17-3-3. Location of Source of Errors. The sources of errors in computing equipment may be located with the aid of several of the devices which have been discussed in the sections on reliability and checking machine errors. For example, indicating fuses, neon lamps in resistor and capacitor circuits, and neon lamps to indicate the quantities in registers are all valuable aids to a technician who is thoroughly acquainted with the functioning

of the computer.   After troubles have been traced to one unit, it can be removed from the computer and subjected to diagnostic circuit tests if required.

17-4. Size, Cost, and Speed

Table 17-1 shows the number of digits which can be retained in various storage mediums for a unit cost and in a unit volume.

TABLE 17-1

Medium	Digits per dollar	Digits per cubic foot	Maximum convenient speed, digits per second		Remarks
			Recording	Reading	
Electromechanical relays*.........	0.2–1	300	100	100	
Electronic relays*.	0.2–1	300	10^7	10^7	
Punched cards....	2.5×10^5	9×10^6	2.5×10^3	5×10^3	10^3 binary digits per card, 0.4 cents per card
Teletype punched tape	3×10^6	3×10^7	50	50	Mechanical sensing, 0.05 cents per foot
Teletype punched tape	3×10^6	3×10^7	50	7×10^3	Optical sensing at 10 feet per second, 0.05 cents per foot
Thirty-five-milli-meter film, punched	6×10^4	3×10^7	100	10^4	Mechanically punched, optically read at 10 feet per second, 2 cents per foot
Wide punched tape.	10^5	2×10^7	6×10^2	6×10^4	Mechanically punched, optically read at 10 feet per second
Thirty-five-milli-meter photo-graphic film	10^6	10^9	6×10^5	6×10^5	Optical sensing of 0.01-by 0.02-inch spots, 10 feet per second, 5 cents per foot
Magnetic wire......	2.5×10^6	10^{10}	10^4	10^4	0.004-inch diameter, 0.010-inch bauds, 10 feet per second, 0.05 cents per foot
Magnetic tape.....	10^7	3×10^9	10^5	10^5	¼-inch tape, five parallel tracks, 0.020-inch digit spaces, 10 feet per second, 0.1 cent per foot
Mercury delay line.	....	2×10^4	10^6	1×10^6	
RG-64U cable.....	....	5	10^6	5×10^6	
Selectron*........	....	10^3	10^4	5×10^3	
Deflection cathode-ray tube*.......	....	10^2	10^5	2.5×10^5	

* Parallel storage; access time equal to reading time plus writing time.

An approximation of practical reading and writing speeds is also given. The figures shown in this table are by no means precise, but may be of value in making order-of-magnitude estimates.

REFERENCES

Reliability of Components

1. Brotherton, M., *Capacitors* (D. Van Nostrand Company, Inc., New York, 1946), pp. 45–55.

2. Eckert, J., Jr., Reliability of Parts *in Theory and Techniques for the Design of Digital Computers*, Vol. II (Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, 1948), pp. 20-1–20-13.

3. Michael, F. R., Tube Failures in ENIAC, *Electronics*, Vol. 20, No. 10, pp. 116–119 (October, 1947).

4. Radio Corporation of America, *Special Red Tubes* (RCA Tube Department, Harrison, N.J., 1948). Pamphlet.

Possibilities for Checking

5. Alexander, S. N., Input and Output Devices for Electronic Digital Calculating Machinery, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 248–253 (1948).

6. Bloch, R. M., R. V. D. Campbell, and M. Ellis, The Logical Design of the Raytheon Computer, *Mathematical Tables and Other Aids to Computation*, Vol. III, No. 24, pp. 286–295 (October, 1948).

7. Eckert, J. P., Jr., Reliability and Checking *in Theory and Techniques for the Design of Digital Computers*, Vol. IV (Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, 1948), pp. 35-1–35-16.

8. Williams, S. B., Reliability and Checking in Digital Computing Systems *in Theory and Techniques for the Design of Digital Computers*, Vol. IV (Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, 1948), pp. 34-1–34-16.

9. Williams, S. B., Bell Telephone Laboratories Relay Computing System, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 41–53 (1948).

Size, Cost, and Speed

10. Sheppard, C. B., Transfer between External and Internal Memory, *Annals of the Computation Laboratory of Harvard University*, Vol. XVI, pp. 267–271 (1948).

INDEX

441