

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

Subject: EMAR: AN EXPERIMENTAL MEMORY ADDRESS REGISTER

To: D. R. Brown

From: W. A. Clark

Date: 10 August 1955

Approved:

N. L. Daggett  
N. L. Daggett

Abstract: It has been proposed by the Systems Design and Memory Sections that an experimental memory address register for the 256 x 256 memory selection system be constructed as a precursor to the TX-0 MAR. This note describes the register and its control and is intended to serve as a basis for further discussion of the proposal.

### Introduction

The 256 x 256 coincident current memory array requires a 16-bit register to hold the address of the selected word during the read-write cycle. It has been proposed that an experimental memory-address register (EMAR) and associated control be built up in advance of the final design for the TX-0 computer. Figure 1 shows the proposed setup. The register proper would use transistor circuits approximately like those of the TX-0 MAR but control would consist of Burroughs' test-equipment units. The entire setup would provide valuable data on matrix decoding, line driving, and timing problems and would serve as a testing device for the memory selection system and individual memory planes as they become available. Target date for the completion of EMAR would be 1 October 1955.

### Description

The output of EMAR would be signals large enough to control the core-switch drivers over distances of about 40 feet. Inasmuch as there are 4 distinct sets of drivers (one set for each coordinate (u or v) of a core-switch; one core-switch for each coordinate (x or y) of the memory plane), it will be convenient to consider EMAR as divided into 4 groups of digits corresponding to the sets of drivers. These will be called the xu, xv, yu, and yv groups (see Figure 2).

Except for minor differences in the end digits of EMAR, all groups are identical to one another, and it will be sufficient to describe only one group in detail. The xv group is shown in Figure 3. Notice that an additional level of selection takes place at the drivers themselves: one of 2 rows of 8 drivers each is selected (using the driver cathode circuits) by digit 10 during the gating period and one driver in each row is selected (using the driver grid circuits) by the active line from the decoder. Thus, only one driver of the 16 becomes active during the gating period.

The principal difference between the logic of EMAR and that of the TX-0 MAR is that EMAR would be able to count. This ability permits cycling through the entire range of addresses. (In TX-0 the equivalent counting operation takes place in the memory buffer register.) It may be desirable to include display decoders and scope for visual checking of the selection system.

### Control

A simple circulating pulse delay line control is adequate for EMAR. Figure 4 shows one possible form which includes one flip-flop (IFF) for generating the inhibit gating level, one (RFF) for the read gating level, and one (MB FF) to buffer information being read into and out of one memory plane.

Push-button pulse-generators would be required for the following functions:

- 1) start (inject pulse into the loop),
- 2) clear MB FF,
- 3) set MB FF,
- 4) clear EMAR,
- 5) preset EMAR (to value held in toggle switch register).

Signed: \_\_\_\_\_

W. A. Clark

WAC/dg

Distribution:	Group 63 Section Leaders	Attachments:	Figure 1
	Staff, Systems Design Section		Figure 2
	Staff, Memory Section		Figure 3
	Staff, Logical Design Section		Figure 4
	Staff, New-Components-and-Circuits Section	Dwg. Nos.	A-63554
			A-63555
			A-63556
			A-63557

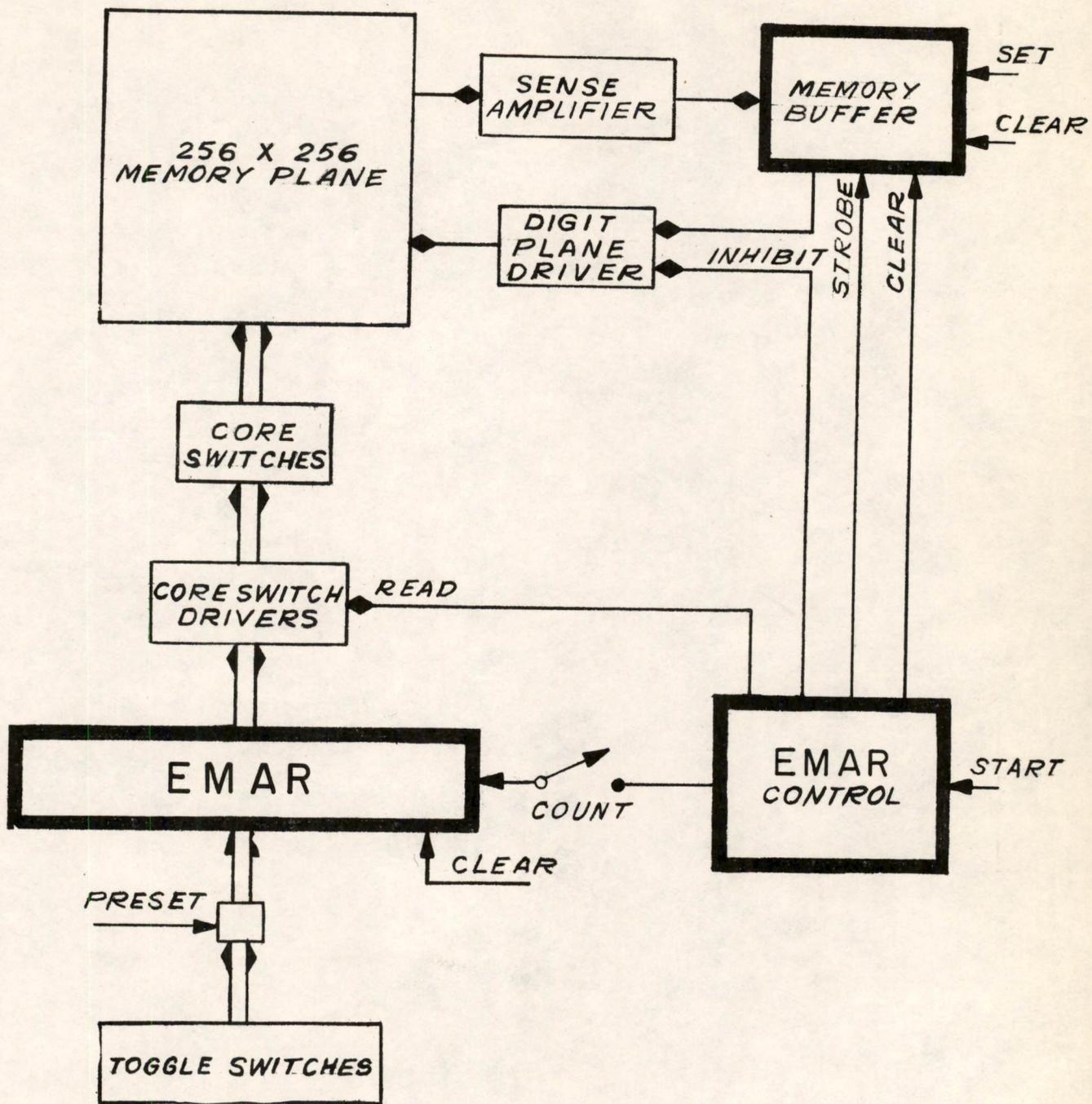


FIG. 1  
THE EMAR SETUP

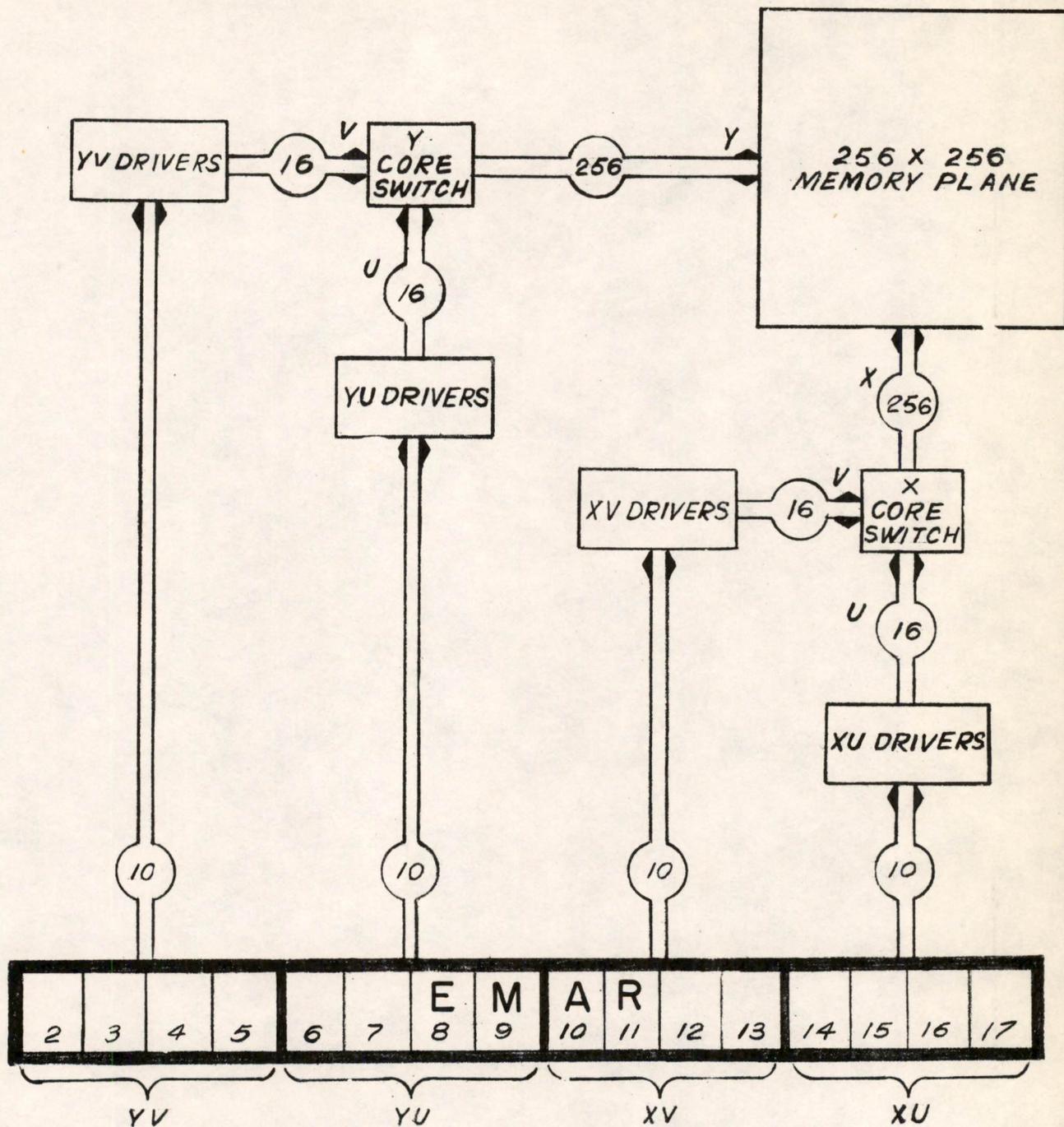


FIG. 2

EMAR AND THE MEMORY SELECTION SYSTEM:  
 (DIGITS ARE NUMBERED AS THEY WOULD BE IN TX-0)

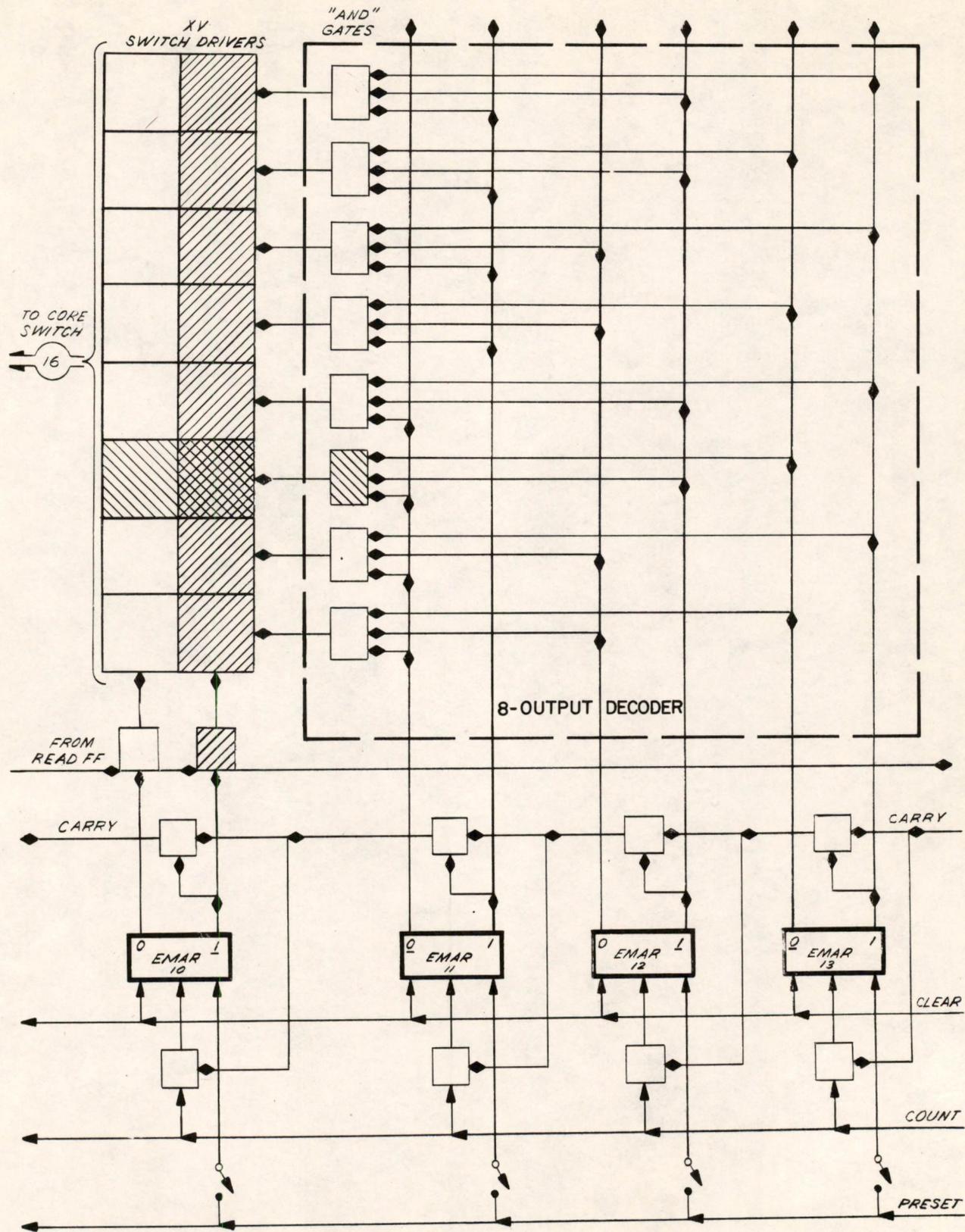


FIG. 3

THE XV GROUP

EMAR FF'S HOLD 1010. ACTIVE GATES ARE SHADED, AND THE ACTIVE SWITCH DRIVER IS DOUBLY-SHADED.

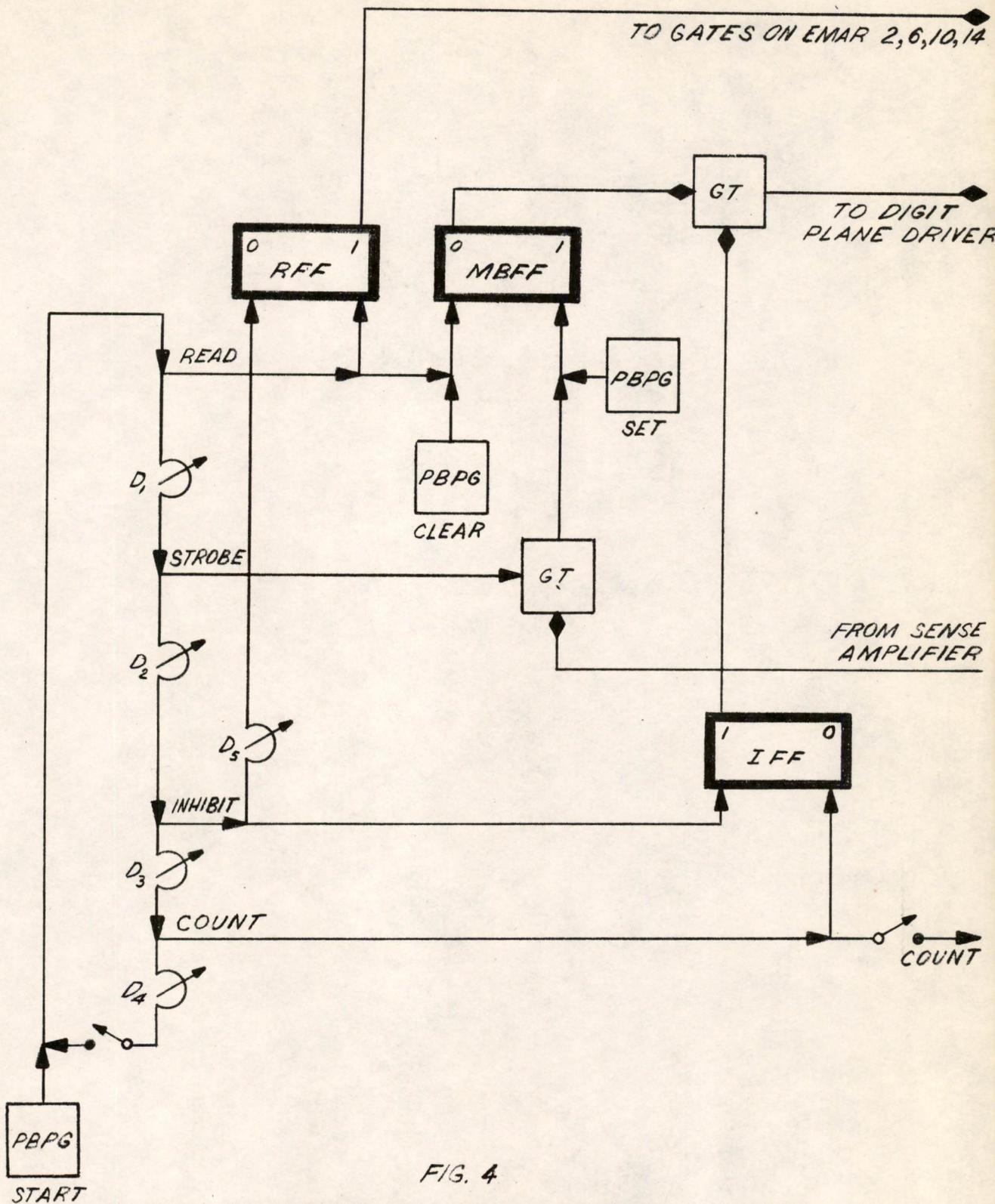


FIG. 4  
EMAR CONTROL AND MEMORY BUFFER

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: Organization and Responsibilities of MTC Technicians

TO: All MTC Technicians

FROM: H. L. Ziegler

DATE: 11 August 1955

APPROVED: W. A. Hosier  
W. A. Hosier

ABSTRACT: In recent months there has been considerable confusion concerning supervision of MTC technician work. This memo is to clarify this situation as well as outline the general duties of the technicians.

To avoid unnecessary confusion in work assignments of MTC technicians all future assignments will be made by Stan Olsen. Technicians receiving direct requests to do special jobs should refer such requests to Stan for his decision.

Watch or duty assignments will be made by Joe Salvato as in the past. Duty technicians are expected to be in the vicinity of the Control Console at all times unless they have been temporarily relieved at their request. "Swaps" of duty hours are not to be made unless approved by both Joe and Stan.

Starting immediately the duty technician will assume certain computer operation duties. These consist of making all temporary computer changes requested by programmers and listing these on the "Live Messages Only" blackboard; operating the PETR, the Flexo, (punch and printer) the IBM card machine, and, in general, seeing that paper tapes and punched cards are handled properly and are kept in good condition. In this way there should be considerably fewer mistakes due to new users of the computer not being familiar with correct operation procedures. Also, any faulty operation of equipment will be known to MTC maintenance personnel much sooner than with the present system. These new duties do not alter the present ones which include responsibility for seeing that the Log is kept up and that the console area is kept clean.

Efficient troubleshooting of the computer requires a thorough knowledge of the design logic as well as a familiarity with electronics and circuitry. The emergency conditions under which most troubleshooting is done prevent the troubleshooter from giving a detailed and descriptive

"running account" of his method and his thought sequence. It is up to the technician to have sufficient knowledge of the computer system to follow and assist the troubleshooter in his step-by-step analysis of the trouble. All this implies that technicians when not on duty or assigned to special work such as construction should spend a reasonable portion of their time studying the MTC system.

To make this task somewhat less discouraging the system will be divided into sections with each technician having primary responsibility for the equipment in his section. When he has had time to learn that particular section he will be "rotated" to the next one to learn it. A technician moving to a new section of equipment will then be available as a "consultant" to his replacement in the old section he has left.

Computer time is expensive and if we are to justify shutting down for four hours on Monday morning we must produce as much improvement as possible to show for it. Therefore, work on the computer during this period has top priority over everything else.

Signed: \_\_\_\_\_

*H L Ziegler*  
H. L. Ziegler

HLZ/eta

Distribution List

E. Albanese (15 copies)

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: POSITIVE BIAS AS APPLIED TO SURFACE-BARRIER TRANSISTOR SWITCHING CIRCUITS

To: Group 63 Staff

From: Kenneth H. Konkle and Edmund U. Cohler

Date: August 12, 1955

Approved: Torben Meisling  
Torben Meisling

Abstract: Successful operation of switching circuits using surface-barrier transistors requires that the output of an "on" stage be less than the input required to initiate switching of the following "off" stage. Positive bias allows this requirement to be met with transistors having low beta, or with noise in the system. An equivalent circuit, which replaces the positive bias current with a battery in series with the base and collector resistor, allows these circuits to be easily analysed. Positive bias reduces the collector current in the driver and the base current to the loads. This reduction of base current can be minimized by redesigning the driver so as to maintain the collector current at its former value. By the addition of positive bias, flip-flops are made d-c stable with low beta transistor. Positive bias can also be applied to pulse transformer circuits.

Introduction:

Successful operation of switching circuits (Fig. 2A) using surface-barrier transistors depends on the shape of their input and transfer characteristics in the grounded-emitter connection (Fig. 1). With the average transistor,  $V_C$  in the saturated region is less negative than -0.1 volt; and fortunately, this voltage, when applied to a succeeding stage, is not sufficient to bring it out of the "off" region. If, in Fig. 2A,  $V_{B1}$  is less negative than -0.1 volt, it follows that Q1 will be off,  $V_{B2}$  will be more negative than -0.35 volt, Q2 will be on, and  $V_{C2}$  and  $V_{B3}$  will be less negative than -0.1 volt. Conversely, if  $V_{B1}$  is more negative than -0.35 volt, it follows that Q1 will be on,  $V_{C1}$  and  $V_{B2}$  will be less negative than -0.1 volt, Q2 will be off, and  $V_{B3}$  will be more negative than -0.35 volt. This is shown graphically in Fig. 3, the per stage transfer characteristic for the circuit of Fig. 2A. This S-shaped characteristic, produced by the shape of the

transfer characteristic of Fig. 1, makes the switching circuit a two state device which is tolerant to changes in transistor characteristics, component values, input voltage, and noise. The object of positive bias is to shift these characteristics so that a greater difference exists between  $V_C$  in the saturated region and the  $V_B$  required to bring the transistor out of the off region. This allows switching circuits to be designed which have even greater tolerance to noise and to transistors having low initial beta, low beta at high current, high saturation collector voltage, and end resistance.

#### Circuit and Equivalent Circuit:

Positive bias is applied to the circuit of Fig. 2A as shown in Fig. 2B. Either of these circuits can be represented by the block diagram of Fig. 2C. The coupling network is shown in detail in Fig. 4A. This network is linear and can be handled by conventional circuit analysis techniques. Removing the positive bias current source and replacing it with batteries gives the equivalent coupling network of Fig. 4B, where:

$$E_{CB} = R_C \cdot I_P \quad (1)$$

$$E_{PB} = R_B \cdot I_P \quad (2)$$

Application of this equivalent circuit to a switching circuit with several loads and positive bias gives Fig. 4C, where:

$$E_{CB} = R_{C0} \sum_{j=1}^N I_{Pj} \quad (3)$$

$$E_{PBj} = R_{Bj} \cdot I_{Pj} \quad (4)$$

#### Applications:

##### Inverters:

Computer logic building blocks, such as inverters, level gates, and pulse gates, using surface barrier transistors are fundamentally the switching circuit of the type shown in Fig. 2. The design requirements for these circuits are that, with the available drive, the transistor must be made to saturate and, with the transistor off, its collector circuit must be able to supply sufficient drive to saturate succeeding stages. These requirements become difficult only when it is decided that for system simplicity all inverters must be identical and that an

inverter must be capable of driving a given number of loads.

Although positive bias does give greater tolerance to noise and saturation voltage, it reduces the drive to the loads. This reduction in drive is caused by the lower effective drive voltage. The effective drive voltage is  $E_{CC}$  in the circuit without positive bias and is  $E_{CC} + E_{CB} + E_{PB}$  for the circuit with positive bias. As can be seen from equation (3),  $E_{CB}$  is increased with each additional load. This loss of drive can be somewhat compensated for by several circuit changes which do not change the operating point of the drive transistor from that without bias.

The saturation collector current in the drive transistor without bias is:

$$I_{Con} = (E_{CC} - V_{Con})/R_C \approx E_{CC}/R_C \quad (5)$$

where  $I_{Con}$  is collector saturation current,  $V_{Con}$  is saturation collector-emitter voltage. The circuit values when bias is used are indicated by primes:

$$I'_{Con} = \frac{E'_{CC} + E_{CB} - V'_{Con}}{R'_C} = \frac{E'_{CC} + R'_C \sum^N I_P - V'_{Con}}{R'_C} \quad (6)$$

If the transistor operating point in the saturated region is to be the same in both cases then:

$$I_{Con} = I'_{Con} \text{ and } V_{Con} = V'_{Con}$$

hence:

$$\frac{E_{CC} - V_{Con}}{R_C} = \frac{E'_{CC} + E_{CB} - V_{Con}}{R'_C}$$

The circuit parameters  $R'_C$  and  $E'_{CC}$  can be varied in any way to satisfy this equation. Two apparent methods are to hold either  $R_C = R'_C$  or  $E_{CC} = E'_{CC}$ . If  $R_C$  is held constant then:

$$E'_{CC} = E_{CC} - E_{CB} = E_{CC} - R_C \sum^N I_{P_j} \quad (7)$$

If  $E_{CC}$  is held constant it can be shown that:

$$\frac{1}{R'_C} = \frac{1}{R_C} + \sum^N \frac{-I_{P_j}}{(E_{CC} - V_{Con})} \quad (8)$$

The circuit for the constant  $R_C$  compensation will be that of

Fig. 4C with  $E_{CC}$  for the driver replaced by  $E'_{CC}$  of equation (7). This does not appear to be an attractive method of compensation except in the case where a fixed number of loads are used, such as in a flip-flop. A variation of this circuit, which might be called the constant  $I_P$  compensation, could be used where a definite maximum number of loads is specified. Positive bias current would be fed into the collector circuit for each load not connected and  $E'_{CC}$  would be set at a value:

$$E'_{CC} = E_{CC} - E_{CB_{max}}$$

The constant  $E_{CC}$  compensation would be achieved by having a compensating resistor,  $R_k$ , connected from each load to  $E_{CC}$  where:

$$R_k = (E_{CC} - V_{Con}) / -I_P \quad (9)$$

This would automatically adjust the value of  $R'_C$  to match the amount of bias current in the collector circuit of the driver. These circuits are shown in Figs. 5A and 5B with positive bias current supplied from a large positive voltage  $E_{pp}$  through a high resistance,  $R_p$ .

The above methods of compensation will not maintain the drive to the loads at the value present without bias; they are only effective in raising the effective drive voltage or in reducing the output impedance of the driver. The drive to a particular number of loads can be maintained by reducing the base resistance. The table of Fig. 6 indicates the base current of the load inverters vs. their number for the various circuits described above using transistors having the input characteristic of Fig. 1. The base resistance in the last three circuits tabulated were reduced so as to maintain the drive to two loads equal to that obtained without bias.

Returning now to the transfer characteristic of inverters, Fig. 7 shows the new per-stage transfer characteristic obtained when 0.1 volt of positive bias is added and  $E_{CC}$  and  $R_B$  are adjusted to give drive equal to that obtained for the circuit, whose transfer characteristic is given in Fig. 3.

#### Flip-Flops:

Positive bias may be applied to the flip-flop circuit designed by this group. Fig. 8A shows the basic flip-flop circuit, and Fig. 8B, the equivalent circuit obtained when positive bias is used. In designing the flip-flop, positive bias is first neglected, and circuit parameters are specified so as to supply the amount of drive required for stability and speed. Positive bias is then included and the previously determined

circuit parameters are changed to maintain the same transistor operating points. The addition of positive bias increases the margin with respect to internal and external noise and to variation in transistors. In addition the dynamic operation of the flip-flop is somewhat improved because of decreased hole-storage recovery time and improved waveshapes.

Consider a flip-flop which has previously been designed without positive bias according to the above considerations; e.g.,

$$E_{CC} = -1.5 \text{ volts, } I_{Con} = -4.33 \text{ ma., } I_B = -0.493 \text{ ma.,}$$

$$R_C = 300 \text{ ohm, and } R_B = 2.32K \text{ ohm.}$$

The circuit may now be redesigned to include positive bias, with new parameter values indicated by primes. The circuit parameters  $R_C$ ,  $I_C$ , and  $I_B$  will be kept constant. From equation (7)

$$E_{CC}' = E_{CC} - 2R_C I_P \quad (10)$$

Since  $I_B$ ,  $R_C$ , and  $E_{CCeff}$  are the same, the equations giving the new value of  $R_B$  can be found from:

$$R_B' = R_B I_P / (I_B - I_P). \quad (11)$$

Finally the base voltage of the off transistor will be:

$$V_{Boff}' = V_{Con} + R_B I_P I_B / (I_B - I_P) = V_{Con} + E_{PB} \quad (12)$$

where without positive bias:

$$V_{Boff} = V_{Con} \quad (13)$$

This particular voltage is the critical one in determining tolerance of the flip-flop to changes in transistor beta. It has been found that  $V_{Boff}$  must be less negative than -0.15 volt to maintain d-c stability. It has also been found that for  $V_{Con}$  more negative than -0.2 volt, the transistor is in the active region and hence:

$$V_{Con} = E_{CCeff} - \beta R_C I_B \quad (14)$$

It has been arbitrarily decided that a beta of 6.6 will be the "end of life" value for a collector current of 5 ma. Substituting the previously mentioned circuit parameters in equation (14) gives  $V_{Con} = -0.525$  volt. Using equation (12) with this value of  $V_{Con}$  and  $V_{Boff} = -0.15$  gives  $I_P = 0.24$  ma. This value of  $I_P$  when used in equation (11) gives  $R_B' = 1.56K$  ohm and in equation (10),  $E_{CC}' = 3.14$  volts. The above design

used equal bias currents for the flip-flop and amplifier; however, by slight modification of the equations unequal currents could have been used.

Pulse Transformer Circuits:

The circuit of Fig. 8 shows how positive bias could be applied to a pulse transformer circuit. Due to the fact that a very low output impedance is desirable in this circuit, the internal impedance of  $E_{PB}$  should be kept small.

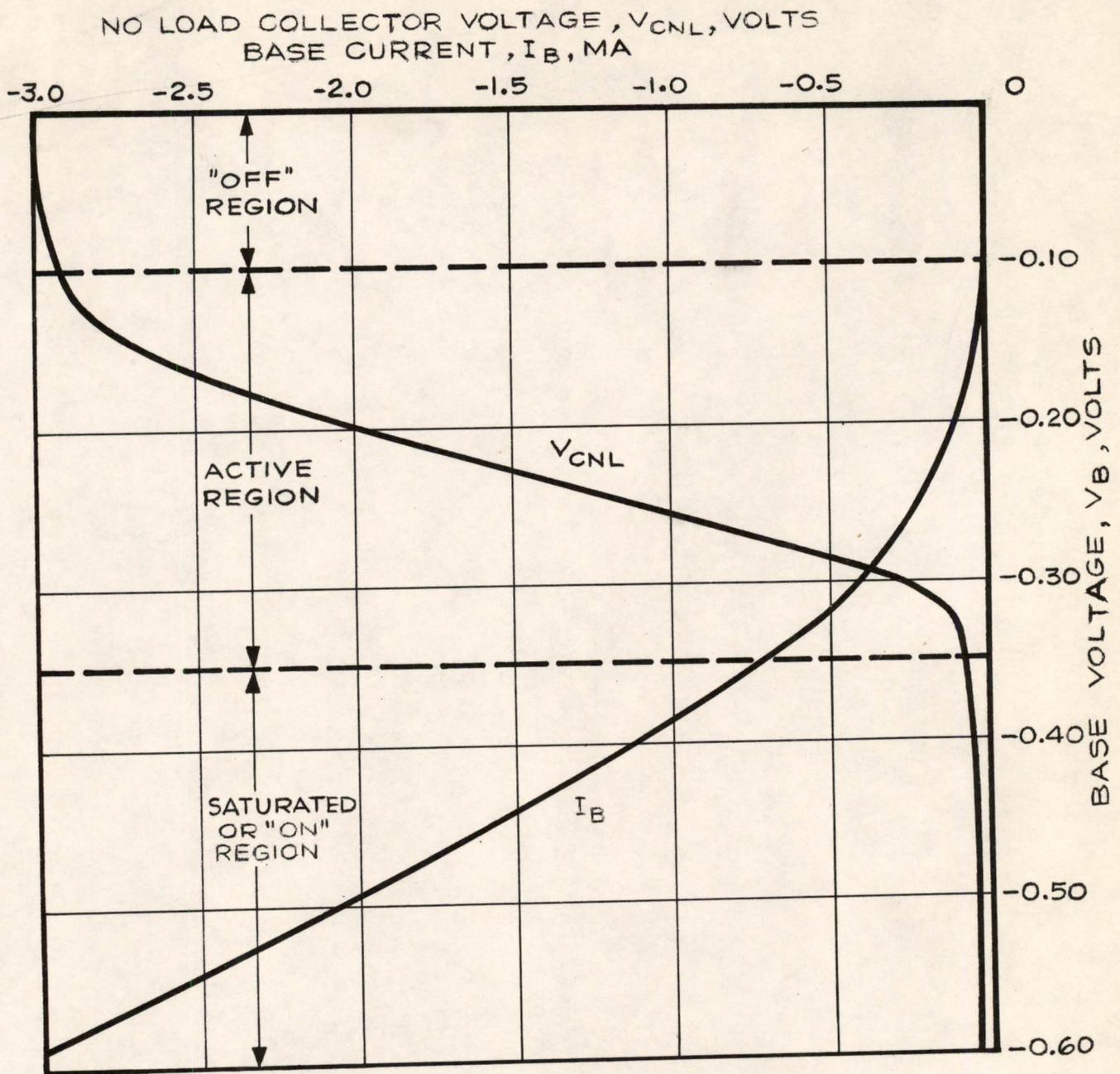
Kenneth H. Konkle  
Kenneth H. Konkle

Edmund U. Cohler  
Edmund U. Cohler

KHK/dg

Distribution: Group 63 Staff

Attachments:	Fig. 1	Drawing Numbers	A-63589
	Fig. 2		A-63590
	Fig. 3		A-63591
	Fig. 4		B-63592
	Fig. 5		B-63593
	Fig. 6		
	Fig. 7		A-63594
	Fig. 8		B-63595



SBT=NO. 121  
 $E_{CC} = -3.0$  VOLTS  
 $R_C = 0.50$  K $\Omega$

FIG. 1

INPUT AND TRANSFER CHARACTERISTICS OF  
 A SURFACE-BARRIER TRANSISTOR IN THE  
 GROUNDED EMITTER CONNECTION

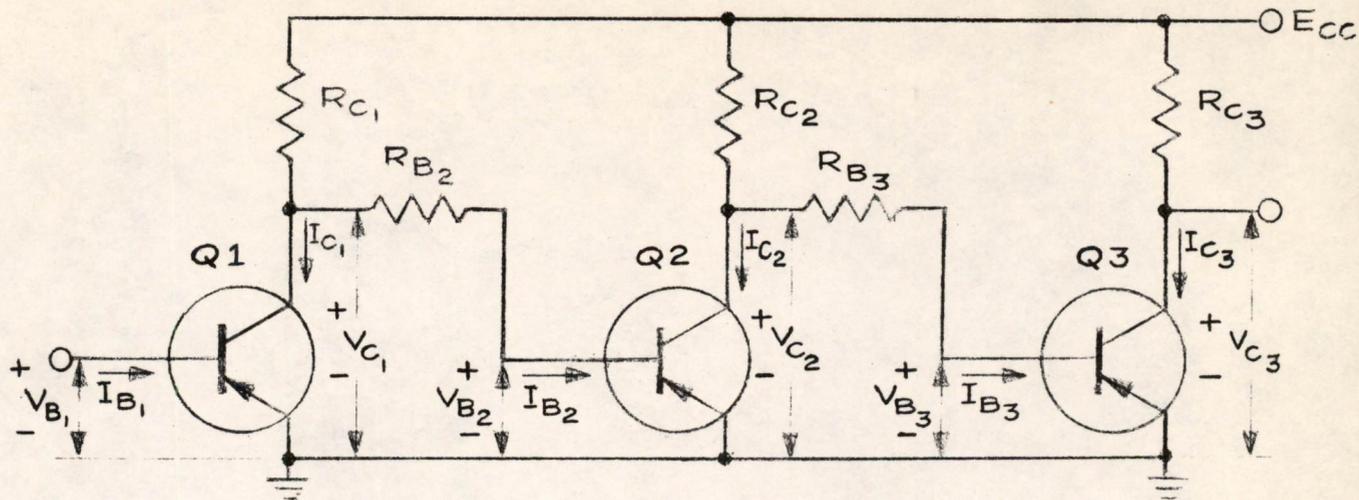


FIG. 2A  
BASIC SWITCHING CIRCUIT

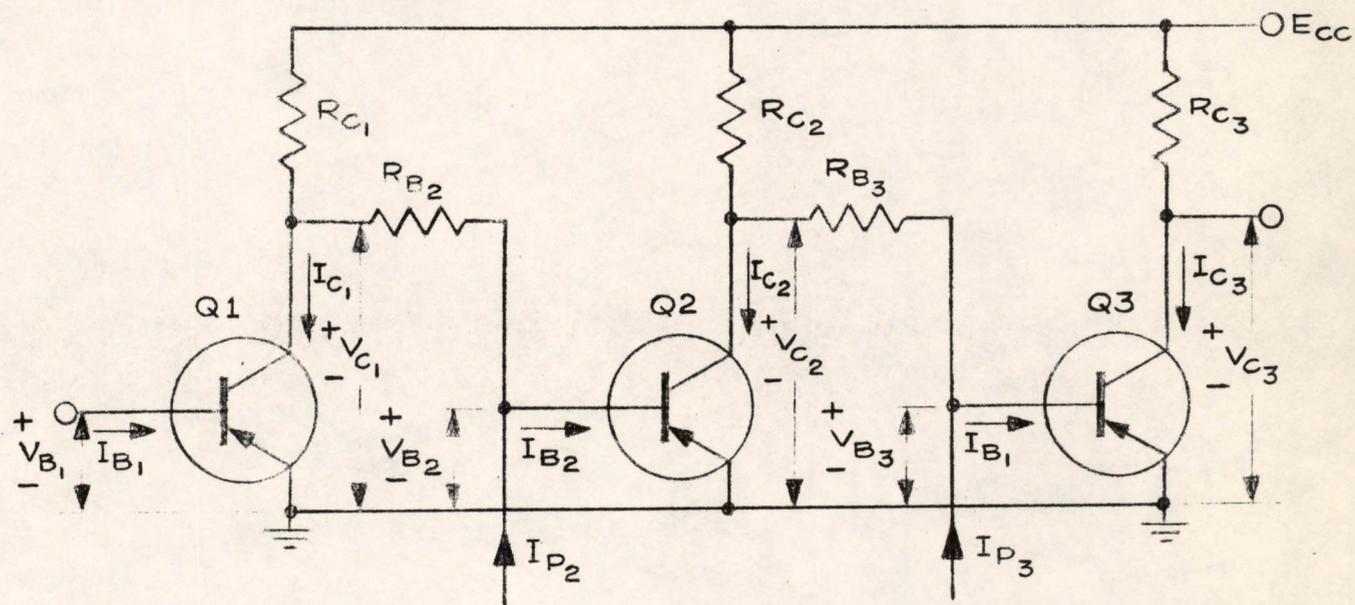


FIG. 2B  
SWITCHING CIRCUIT INCORPORATING POSITIVE BIAS

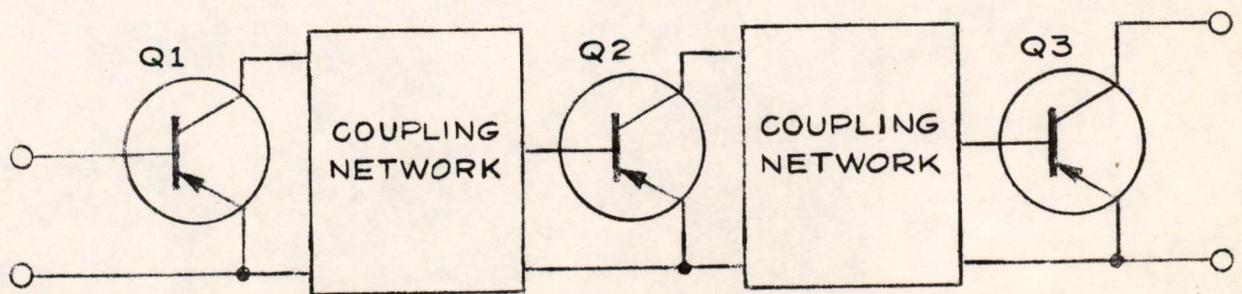
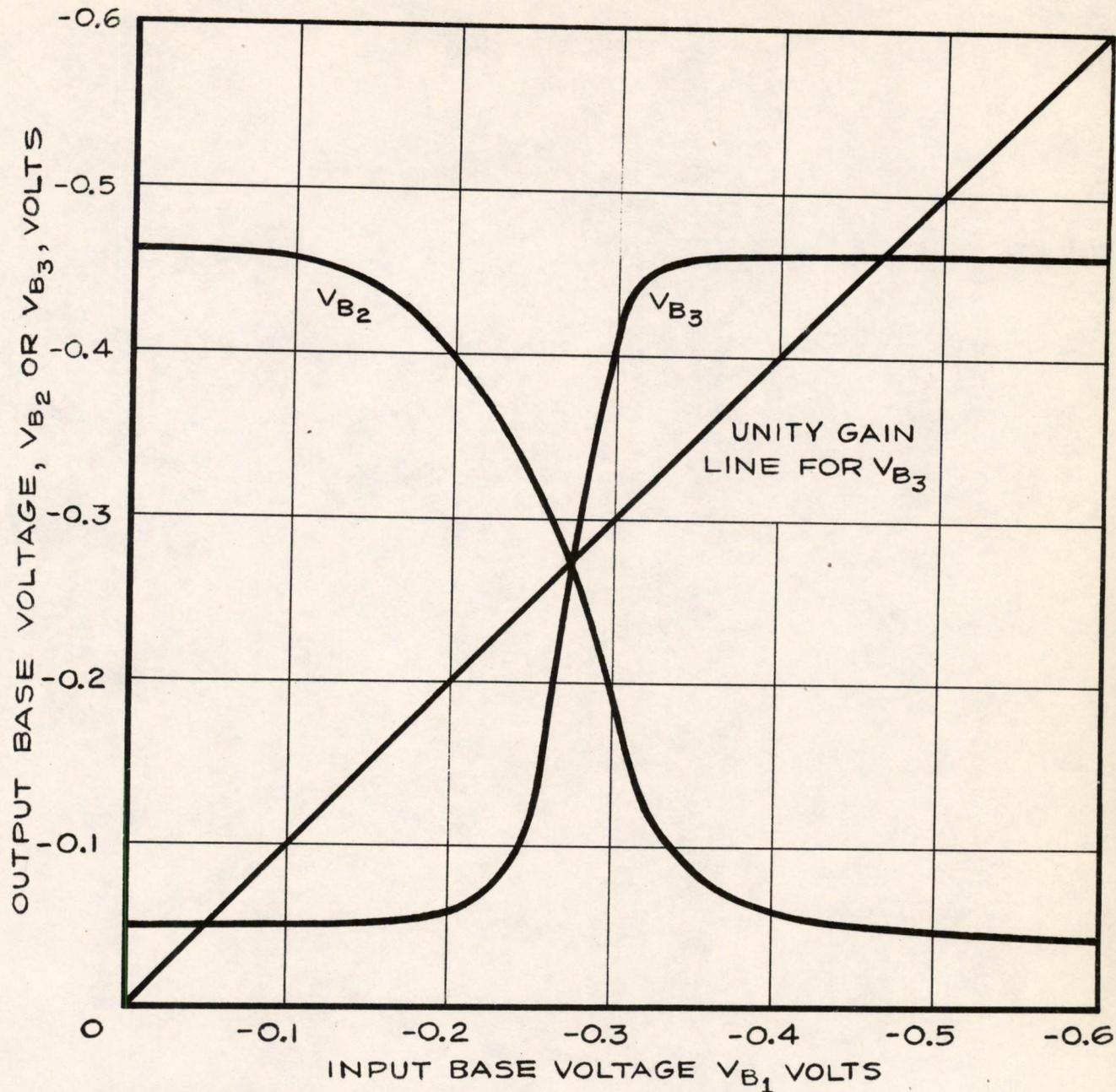


FIG. 2C  
SWITCHING CIRCUIT BLOCK DIAGRAM



$E_{CC} = -30V$   
 $R_C = 500\Omega$   
 $R_B = 1000\Omega$

FIG. 3

PER STAGE TRANSFER CHARACTERISTIC OF SWITCHING CIRCUIT OF FIG. 2A

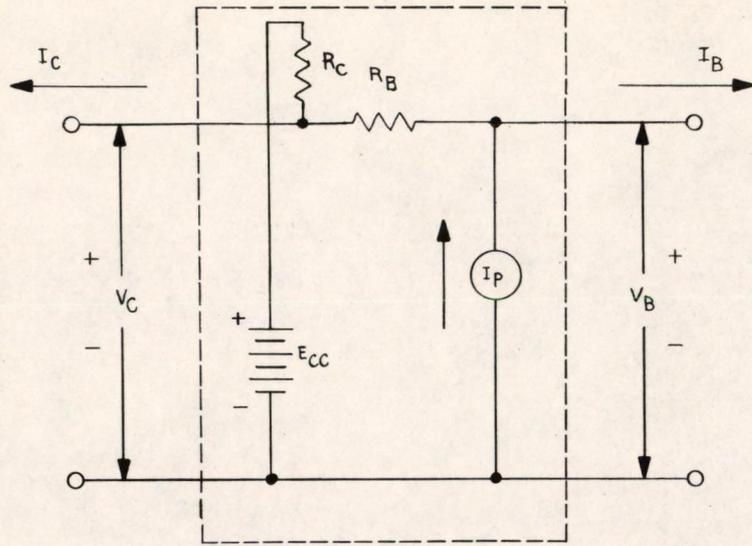


FIG. 4A  
COUPLING NETWORK

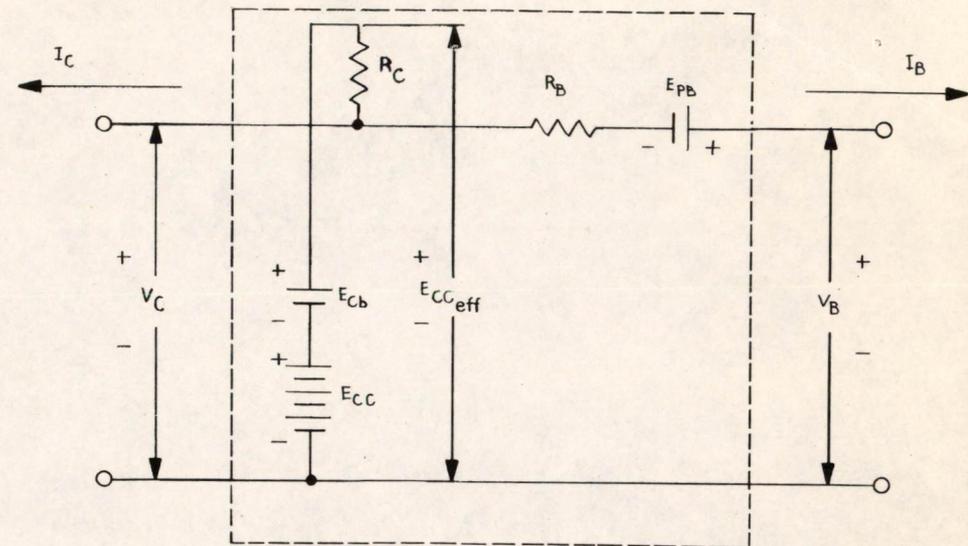


FIG. 4B  
EQUIVALENT COUPLING NETWORK

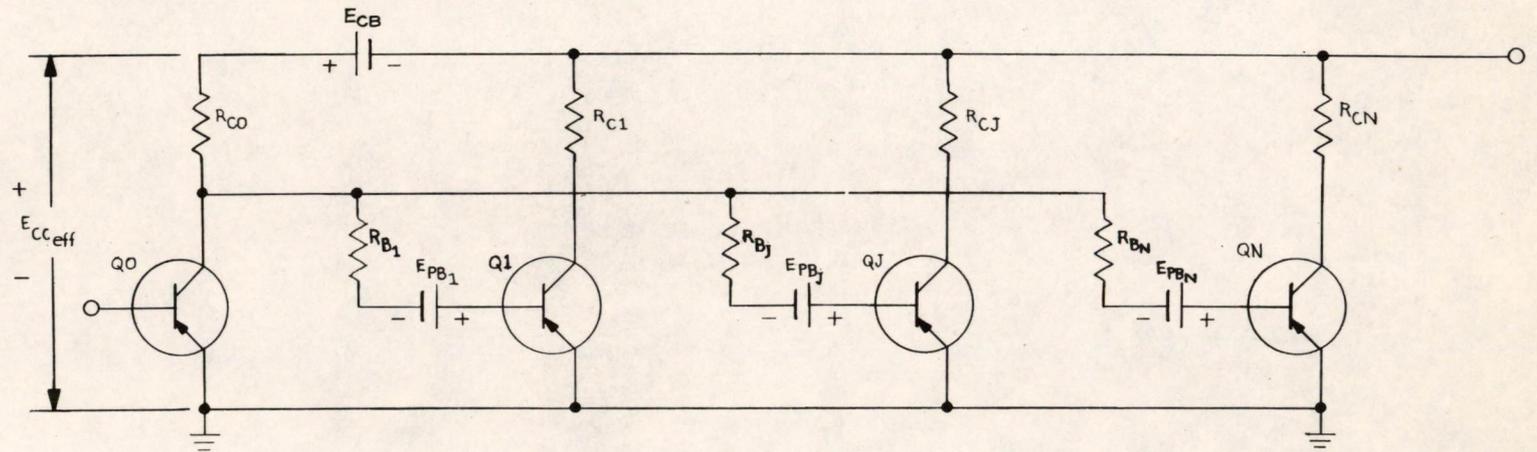


FIG. 4C  
SWITCHING CIRCUIT WITH SEVERAL LOADS AND POSITIVE BIAS

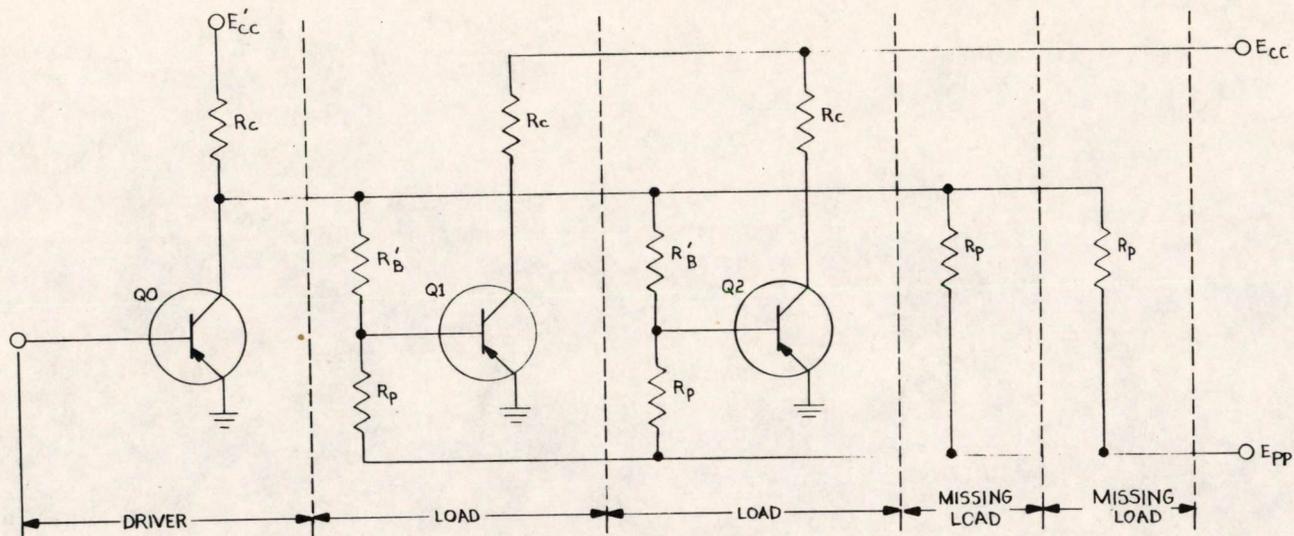


FIG. 5A

CONSTANT IP COMPENSATION

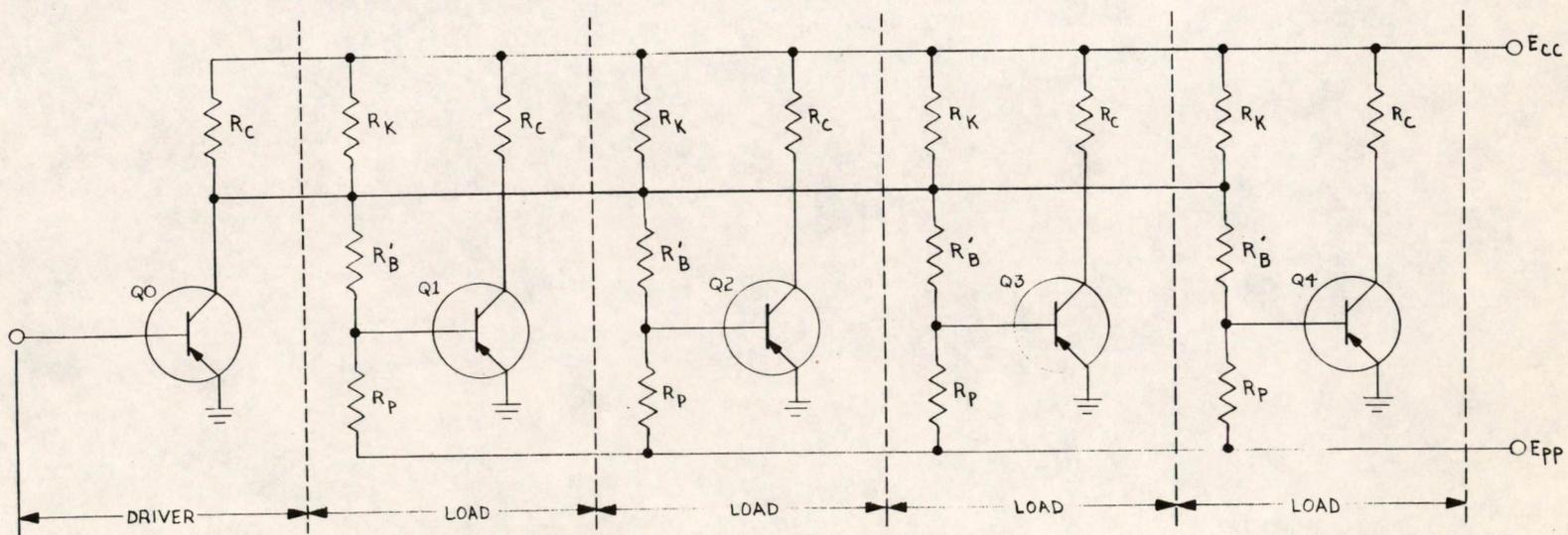


FIG. 5B

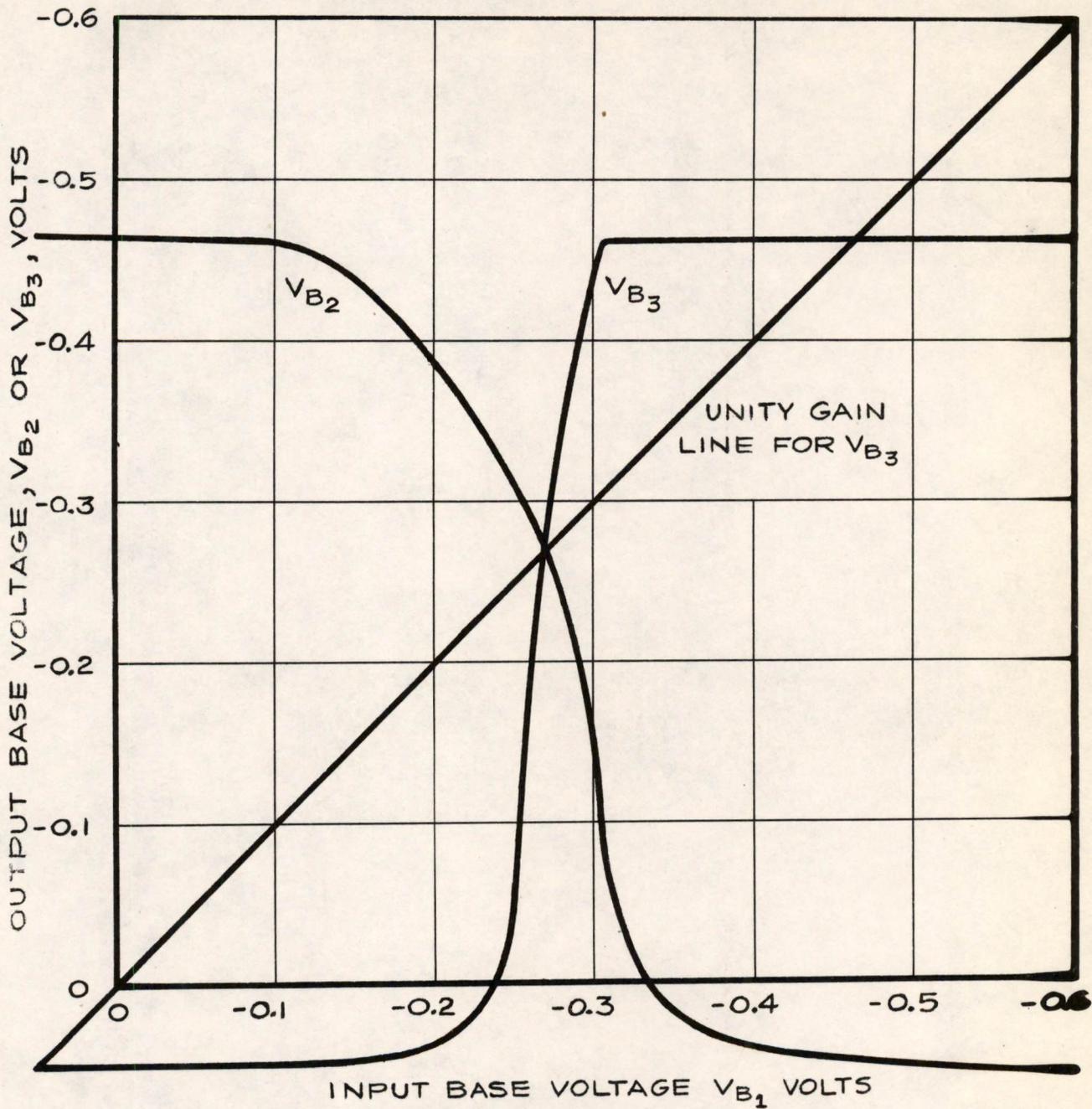
CONSTANT ECC COMPENSATION

Drive Current,  $I_B$ , to Loads in ma.

Number of Loads	1	2	3	4	5	6
No positive bias $E_{CC} = -3.0$ volts $R_C = 0.5K$ ohm $R_B = 1.0K$ ohm $I_P = 0$ ma $E_{PB} = 0$ volts	1.68	1.29	1.04	0.87	0.75	0.66
Positive bias but no circuit change $E_{CC} = -3.0$ volts $R_C = 0.5K$ ohm $R_B = 1.0K$ ohm $I_P = 0.1$ ma. $E_{PB} = 0.1$ volt	1.56	1.15	0.89	0.71	0.59	0.49
Positive bias and lowered $R_B$ only $E_{CC} = -3.0$ volts $R_C = 0.5K$ ohm $R_B = .88K$ ohm $I_P = .115$ ma $E_{PB} = 0.10$ volt	1.76	1.29	0.99	0.79	0.65	0.54
Constant $R_C$ or $I_P$ compensation $E_{Ceff} = -3.0$ volts $R_C = 0.5K$ ohm $R_B = .93K$ ohm $I_P = .107$ ma. $E_{PB} = 0.1$ volt	1.70	1.28	1.03	0.85	0.73	0.63
Constant $E_{CC}$ compensation $E_{CC} = -3.0$ volts $R_C = 0.5K$ ohm $R_B = .85K$ ohm $R_K = 25K$ ohm $I_P = .120$ ma. $E_{PB} = 0.1$ volt	1.77	1.29	1.03	0.85	0.73	0.64

Fig. 6

Table of Drive Current vs. Number of Loads for  
 Different Types of Compensation.



$E_{c\text{eff}} = -3.0\text{V}$      $I_p = 0.107\text{ MA}$   
 $R_c = 500\ \Omega$          $E_{PB} = 0.10\text{ VOLTS}$   
 $R'_B = 930\ \Omega$

FIG. 7

PER STAGE TRANSFER CHARACTERISTIC OF SWITCHING CIRCUIT OF FIG. 2B

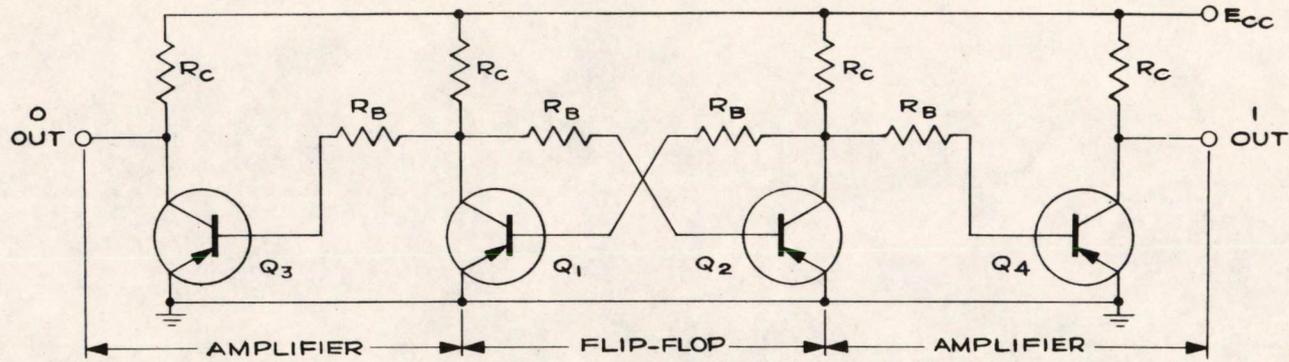


FIG. 8A  
BASIC FLIP-FLOP CIRCUIT

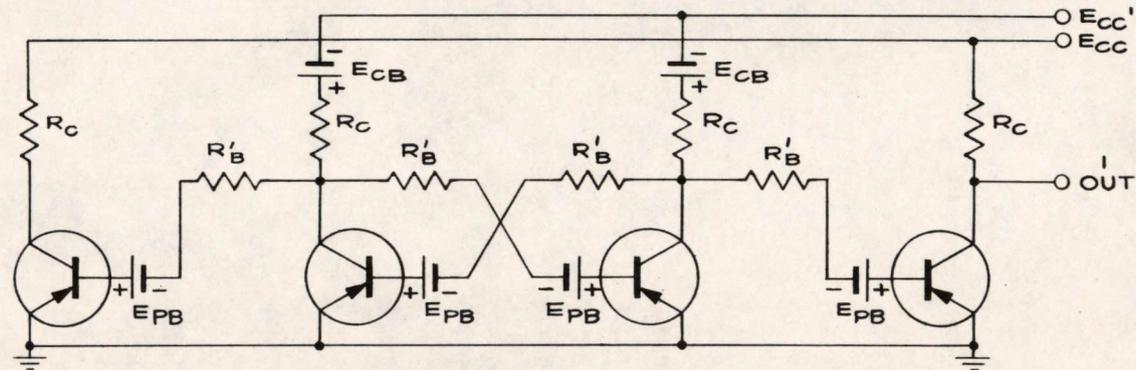


FIG. 8B  
EQUIVALENT FLIP-FLOP CIRCUIT WITH POSITIVE BIAS

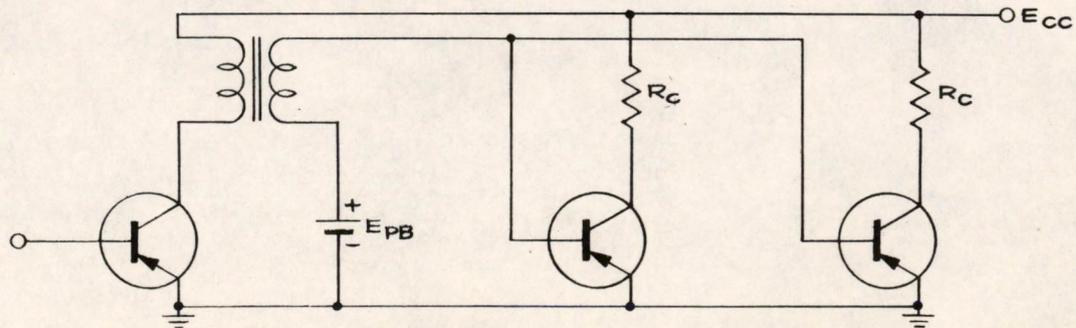


FIG. 8C  
PULSE TRANSFORMER CIRCUIT INCORPORATING POSITIVE BIAS

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: THE CRYOTRON -- A SUPERCONDUCTIVE COMPUTER COMPONENT

To: David R. Brown

From: Dudley A. Buck

Date: August 22, 1955

Approval:

*Torben Meisling*  
Torben H. Meisling

Abstract: The study of nonlinearities in nature suitable for computer use has led to the cryotron, a device based on the destruction of superconductivity by a magnetic field. The cryotron, in its simplest form, consists of a straight piece of wire about one inch long with a single-layer control winding wound over it. Current in the control winding creates a magnetic field which causes the central wire to change from its superconducting state to its normal state. The device has current gain, that is, a small current can control a larger current and it has power gain so that cryotrons can be interconnected in logical networks as active elements. The device is also small, light, easily fabricated, and dissipates very little power.

1. The Cryotron Principle

Before describing the cryotron as a circuit element and potential computer component, the basic physical phenomena underlying its operation will be described.

1.1 Superconductivity

Superconductivity was discovered in 1911 by H. Kammerlingh Onnes at Leiden, three years after he succeeded in liquifying helium. While extending electrical resistance measurements to this new low-temperature region he found that the resistance of mercury drops suddenly to zero at 4.12 K. Soon many other materials were shown to display this same unusual behavior. Niobium becomes a superconductor at 8 K, lead at 7.2 K, vanadium at 5.1 K, tantalum at 4.4 K, tin at 3.7 K, aluminum at 1.2 K, and titanium at 0.5 K. In addition to 21 elements, many alloys and compounds are superconductors with transition temperatures ranging between 0 and 17 K.<sup>1,2</sup>

1. Superconductivity, D. Schoenberg, (Book) Cambridge University Press, 1952.
2. Superfluids, Vol. 1, F. London, (Book) Wiley, 1950.

This document is issued for internal distribution and use only by and for Lincoln Laboratory personnel. It should not be given or shown to any other individuals or groups without express authorization. It may not be reproduced in whole or in part without permission in writing from Lincoln Laboratory.

The research reported in this document was supported jointly by the Department of the Army, the Department of the Navy, and the Department of the Air Force under Air Force Contract No. AF 19(122)-458.

The resistivity of many superconductive materials is relatively high at room temperature, especially those which have high transition temperatures such as niobium, lead, tantalum, etc. It is interesting that relatively poor conductors become superconductors at low temperatures whereas good conductors such as gold, silver, and copper do not. The resistivity of superconductive materials drops as they are cooled. Just above their superconductive transition, the resistivity is between  $10^{-1}$  and  $10^{-3}$  of their room temperature resistivity, depending on the purity and mechanical strain in a particular sample.

Below the superconductive transition the resistivity is exactly zero. That it is truly zero is vividly demonstrated by an experiment now in progress by Professor S. C. Collins at M.I.T. wherein a lead ring has been carrying an induced current of several hundred amperes since March 16, 1954, without any observable change in the magnitude of the current.

## 1.2 Destruction of Superconductivity by a Magnetic Field

The foregoing discussion of the superconductive transition is valid only in zero magnetic field. With a magnetic field applied, the onset of superconductivity occurs at a lower temperature. If the intensity of the magnetic field is increased, the transition temperature is still lower. A plot of the transition temperature as a function of the applied magnetic field is more or less parabolic in shape, levelling out as absolute zero is approached. Such a plot for several common elements is given in Figure 1.

If the temperature is held below the transition temperature for one of these materials, the resistance of that material is zero. Its resistance will remain zero as a magnetic field is applied until that magnetic field reaches a critical value. Above this value the normal resistance returns. If the field is lowered, the resistance disappears.

Raising and lowering the magnetic field thus controls the resistance of the material in the magnetic field by causing it to shift from its superconducting state to its normal state and back without changing the temperature. In Figure 2, this operation corresponds to moving up and down on a vertical (constant-temperature) line which has its lower end in the superconducting region and its upper end in the normal region. If the operating line is moved to a lower temperature, the magnetic field required to reach the normal region is greater. For each of the materials which becomes superconducting, there is a temperature about 0.2 K below the zero-field transition which allows operation with rather small magnetic fields--between 50 and 100 oersteds. For lead, this temperature is about 7.0 K, for tantalum about 4.2 K, for tin about 3.5 K, for aluminum about 1.0 K. Tantalum has been used in many of the early experiments at M.I.T. because 4.2 K is the boiling point of helium at a pressure of 1 atmosphere and therefore the temperature of most storage tanks for liquid helium. Higher temperatures (up to 5.2 K) involve raising the pressure on the liquid helium bath; lower temperatures (down to about 1.0 K) involve lowering the pressure. At 4.2 K, then, experiments do not involve sealing of the lead-in wires.

In a typical cryotron, the resistance being controlled is in the form of a straight piece of wire about 1 inch in length. The magnetic control field is generated by current in a single-layer winding which is wound over the central wire (Figure 3, top). The central wire is analogous to the plate circuit of a vacuum tube and the control winding is analogous to the control grid. In this case, the plate resistance is zero in the cutoff region and rises rapidly as grid-current cutoff is reached.

### 1.3 Superconducting Control Winding

The control winding is made of a superconducting wire which has a relatively high transition temperature. Niobium (formerly called columbium) is used because it has a very high transition temperature and can be drawn into fine wire which is strong. Lead or lead-plated wire is a second possible control-winding material.

At the temperature used, the control winding remains a superconductor at all times, and would remain so even in magnetic fields much higher than those being used to control the central wire. Therefore, there is no resistance in the control winding. A magnetic field, once established, needs no further energy for its support; the control current is maintained against zero back voltage. Similarly, all interconnecting wire is also superconducting.

## 2. The Cryotron as a Device

### 2.1 Static Characteristics

The resistance of the central wire of a typical cryotron is plotted as a function of current in the control winding in Figure 4. The central wire is called the gate circuit. This particular cryotron is made by winding a single layer of 0.003-inch insulated niobium wire over 0.009-inch bare tantalum wire. The insulation on the niobium is heavy Formvar. The finished winding has 250 turns per inch. In the midportion of the winding, the magnetic field due to a current is 124 oersteds per ampere. When the control current of Figure 4 is translated into magnetic field intensity, the highest transition field is seen to be about 40 oersteds.

As current in the tantalum gate circuit is increased, the transition control current becomes lower. This effect is due to the additional magnetic field at the surface of the tantalum wire created by the gate current. This field, commonly called the self-field of the wire, limits the amount of current which can be carried by a superconductor. The effect was, in fact, discovered shortly after the discovery of superconductivity when Onnes and his coworkers (1913) tried to make a powerful electromagnet out of their newly discovered zero-resistance materials. When the current in their superconducting solenoid reached a certain critical value, its resistance suddenly reappeared. When the discovery was made that magnetic fields cause restoration of resistance, it was quickly seen by Silsbee (1916) that the limit on the current that can be carried by a superconductor is due to the magnetic field created by that

current. The magnetic field, H, at the surface is given by:

$$H = \frac{I}{\pi d}$$

where H is in ampere-turns per meter  
I is in amperes  
d is in meters.

If H is given in oersteds, I in amperes, and d in mils (thousandths of an inch) this becomes

$$H \text{ oersteds} = \frac{157.5}{d \text{ mils}} I$$

It will be noted that the transition characteristics are very sharp for high gate currents. The additional sharpness is a peculiarity of the measuring technique wherein a current is passed through the gate circuit and the voltage across the gate circuit is measured. When resistance suddenly appears,  $I^2R$  loss in the gate circuit causes heating which lowers the critical field and speeds switching.

The magnetic field due to the control winding is along the axis of the central wire while the self-field of the wire due to its own current is tangential to the wire. The two fields thus add in quadrature and the resulting net field is the vector sum of two fields. Results indicate that the superconducting central wire reaches its critical field when the net field reaches a certain value, regardless of which way the net field points in relation to the center line of the wire. In one experiment, the curves of Figure 4 were reproduced exactly for all four combinations of positive and negative control and gate current. There is no reason to suspect that there would be anisotropy in the critical field for different orientations of the net field with respect to the wire axis as long as the control field is longitudinal, especially since the wire is polycrystalline. Thus the cryotron has an interesting property as a circuit element. Control is independent of the sign of the control current--- it depends only on the magnitude. Furthermore, when the gate circuit is ON, that is, in its superconducting state, current can flow in either direction, unlike a vacuum tube which can pass current only in one direction.

## 2.2 Current Gain

Because the two fields add in quadrature, the self-field of the wire has less effect on the threshold control current at low gate current than it does at high gate current. The locus of threshold control current points as a function of gate current is an ellipse. The ratio of major axis to minor axis of the ellipse is the ratio of the magnetic field produced by a current in the control winding to that produced by the same current in the gate circuit. This ratio is also called the current gain of the cryotron. If the current gain were less than unity, it would not be possible to control one cryotron with an identical cryotron because more current would be required to bring the second cryotron to its control-

current threshold than the first cryotron could handle through its superconducting gate circuit.

The control field is related to the control current by the number of turns per inch in the control winding, and the self field is related to the gate current by the diameter of wire used in the gate circuit. Current gain,  $K$ , is simply given by:

$$K = \pi d \frac{N}{l}$$

For a given pitch control winding and a given gate wire diameter, the current gain is specified. Figure 5 is a plot of lines of constant  $K$  as a function of winding pitch and gate wire diameter. For the cryotron whose characteristics are plotted in Figure 4,  $K = 7$ . The current gain actually observed for a given cryotron is often less than calculated, presumably due to the constriction of supercurrents by small normal regions which nucleate about flaws in the wire surface. Control-current threshold points thus form a locus in the gate current-control current plane which lies on an ellipse of smaller major-to-minor axis ratio.

### 2.3 Power Gain and L/R Time Constant

The input power to a cryotron, exclusive of eddy current and relaxation losses, is the product of the energy stored in the magnetic field of the control winding and the frequency at which the control winding is energized:

$$P_{in} = \frac{f L_c I_c^2}{2}$$

The input power is reactive. In an oscillator circuit the input inductance can be resonated with a linear capacitor to minimize input losses. In computer pulse circuitry, however, the control windings are untuned. The entire amount of power is therefore dissipated.

The output power of a cryotron can be approximated as follows: Consider a cryotron amplifier delivering square waves of equal on and off periods to a resistive load. The gate circuit shunts the current when superconducting and allows part of it to flow through the load when normal. Maximum power transfer occurs when the load resistance,  $R_L$ , is made equal to the normal resistance of the gate circuit  $R_g$ . Under this condition, average load power is given by:

$$P_{ave} = \frac{I_g^2 R_L}{2}$$

Power gain, G, can be approximated by:

$$G = \frac{\text{power out}}{\text{power in}} = \frac{I_g^2 R_L}{f L_c I_c^2} = \frac{1}{f} \left( \frac{I_g}{I_c} \right)^2 \frac{R_g}{L_c}$$

In the pulse circuits of section 3, the gate current of one cryotron becomes the control current of another;  $I_g = I_c$ . For this condition, the frequency at which the power gain becomes unity is:

$$f_{\text{max}} = \frac{R}{L_c}$$

which is the reciprocal of the L/R time constant of the circuit. The L and R are on different cryotrons, but since large numbers of identical cryotrons are involved, one can speak of the L/R time constant of a given cryotron as being the fundamental time constant of the circuitry.

If a given cryotron is made longer while holding the pitch of the control winding constant, the resistance and inductance increase together such that the L/R time constant is not affected. The L/R time constant is thus independent of cryotron length.

If the diameter of a given cryotron is made smaller while holding the pitch of the control winding constant, the resistance increases inversely as the diameter squared, while the inductance decreases directly as the diameter squared. The L/R time constant thus decreases as the fourth power of the diameter.

The current gain of the cryotron drops if the diameter is made smaller while holding the pitch of the control winding constant because the current-carrying capacity of the gate circuit decreases directly with the diameter. If the current gain is to be held constant by increasing the pitch of the control winding proportionately as the diameter is made smaller, the inductance remains constant and the L/R time constant decreases as the square of the diameter. One thus pays rather dearly for current gain. The circuits of section 3, below, are operated with a minimum of excess current gain.

The resistivity of the normal state varies over several powers of ten among the various superconductors. The L/R time constant varies inversely as the resistivity. An increase in speed of circuit operation can therefore be achieved by alloying superconductors to increase resistivity.<sup>3</sup>

---

3. B. Serin, "The Magnetic Threshold Curve of Superconductors," Chapter VII in Progress in Low-Temperature Physics, edited by C. J. Gorter, Interscience Publishers, 1955.

Circuit speed can also be increased by using a hollow central wire. Superconductivity is a skin effect, penetrating but a few hundred atom layers, and therefore the core of a wire can be removed and the wire will still have zero resistance in its superconducting state. The resistance in the normal state, however, will be higher by the ratio of the original cross-sectional area to the new cross-sectional area. The core need not actually be removed, provided it is made to have a relatively high resistivity. Wire with a high-resistivity core and a superconducting shell can be fabricated by vapor plating.

#### 2.4 Eddy Currents

It has been shown in Faber<sup>4</sup> that the delay,  $\tau$ , due to eddy currents in the destruction of superconductivity of a wire by a longitudinal magnetic field is:

$$\tau_e = \text{const.} \frac{\mu}{\rho} \frac{d^2 H}{(h-H_c)}$$

where  $H$  is the external magnetic field,  $H_c$  is the threshold magnetic field and  $\rho$  is the resistivity. The switching time varies directly as the square of the diameter and inversely as the resistivity, and is a function of the amount by which the threshold magnetic field is exceeded.

As the circuits of section 3 are speeded up by making cryotron diameters smaller, there will be a speed range where eddy currents become important. Lowering the diameter still further and increasing the pitch proportionally should then increase the speed as the inverse square of the diameter, since both circuit  $L/R$  time constants and eddy current time constants decrease proportionally.

The observed time constants of the free-running multivibrator of section 3 are of the same order of magnitude as the calculated  $L/R$  circuit time constants. Eddy current effects should become important during the next order of magnitude increase in speed.

The transition from normal to superconductor also involves delays and a somewhat different switching mechanism.<sup>5,6</sup> A supercooling effect is important. A nucleus of superconducting material forms at one spot on the wire surface, sweeps around the wire, and then grows along the wire. Extrapolation of slow velocity data on tin rods in fields just barely below the threshold field indicate that in cryotron operation, velocities of the order of tens of centimeters per microsecond ought to be encountered

- 
4. T. E. Faber, "The Phase Transition in Superconductors II. Phase Propagation above the Critical Field," Proceedings of the Royal Society, A, 219, pp. 75-88 (1953).
  5. A. B. Pippard, Kinetics of the Phase Transition in Superconductors, Philosophical Magazine, 7, 41, p. 243 (1950).
  6. T. E. Faber, "The Phase Transition in Superconductors III. Phase Propagation below the Critical Field," Proceedings of the Royal Society, A, 223, pp. 174-194 (1954).

with a current gain of two. As soon as a superconducting path is established over the surface of the wire, the cryotron is in its superconducting state—even if the center of the wire requires additional time to become superconducting. While it is not anticipated that this transition will be a major source of delay, it is interesting to note that this delay is one which depends on the length of the cryotron.

As circuit speeds are increased by increasing the resistance of the central wire, thereby shortening  $L/R$  circuit time constants and minimizing eddy current effects, a fundamental limit to the ultimate speed exists in the form of relaxation losses. The exact frequency region in which these losses will become predominant is not known, but from experiments with superconducting coaxial cable and wave guide resonators, an estimate is available which places the limit between 100 megacycles and 1000 megacycles.

### 3. Cryotron Computer Circuitry

The low impedance level of cryotron circuitry dictates a high-impedance power supply (current source) with circuit elements connected in series. Each element allows the current a choice among two or more paths only one of which is superconducting; all of the current flows through the superconducting path. The current encounters zero back voltage except when the paths are changing. The standby power is therefore zero. Several circuits, representative of those found in digital computers, are described below.

#### 3.1 Flip-Flop

A bistable element, one of the most common in a digital computer, can be made using two cryotrons. The two gate circuits are each in series with the control winding on the other and the two paths are in parallel (Fig. 3, bottom). If the current is established in one of the two paths, that current makes the alternate path resistive. Current in one path, once established, will therefore continue to flow indefinitely in that path.

Two additional cryotrons can be added to the circuit, one in series with each branch, in order to place the flip-flop in the desired state. A pulse on one of the two input cryotrons momentarily places a resistance in that side. Both sides are then resistive, and the current divides between them. If the power supply current is not larger than twice the critical current of the cryotrons, both sides of the flip-flop will become superconducting. One side of the flip-flop has a resistance inserted by the input cryotron, however, and the current thus chooses the other side. Once the current builds up in the other side, it makes the side on which the input cryotron is being pulsed resistive, and therefore the pulse in the control winding of the input cryotron can be removed; the current will continue in the new path.

Two more cryotrons can be added to the circuit for sensing the state of the flip-flop. Placed with their control windings each in series with one of the two sides of the flip-flop, one of the read-out

cryotrons is resistive and the other is superconductive. The gate circuits are joined and a read-out current pulse is applied at the junction. The read-out pulse will choose one path or the other, depending on the state of the flip-flop. The flip-flop with read-in and read-out cryotrons is shown in Figure 6.

Any number of input cryotrons can be added in series with those already described (Figure 7) to set the flip-flop to one state or the other. Connected as such, they are OR gates; any one of them acting alone can set the flip-flop. Similarly, additional cryotrons can be added with their gate circuits in parallel with the control winding of the input cryotron already described, behaving as AND gates (Figure 9). The flip-flop set current is bypassed through one or more of these parallel gates unless all of them are resistive. This latter connection involves superconductors in parallel, in which case the current divides inversely as the inductance of the parallel paths.

Additional read-out cryotrons can be added in series with those already described. Since their control windings are superconducting, the additional cryotrons do not add any resistance to the flip-flop. The additional inductance increases the L/R time constant of the circuit, however, lengthening the transition time between states.

### 3.2 Multivibrator

Three flip-flops made of one-inch pieces of the cryotron stock whose characteristics are given by Figure 4 have been studied in a multivibrating circuit (Figure 9). The read-out cryotrons of flip-flop A are connected in such a way as to set flip-flop B to the state opposite that of A. A similar connection between B and C causes C to assume the state opposite to that of B, and a similar connection between C and A causes A to assume a state opposite to that of C. Since there are an odd number of stages, the ensemble free-runs through the sequence given in Table I.

Flip-flop	Time Period								
	0	1	2	3	4	5	0	1	2
A	0	0	0	1	1	1	0	0	etc.
B	0	1	1	1	0	0	0	1	etc.
C	1	1	0	0	0	1	1	1	etc.

Table I. Sequence of Multivibrator Flip-Flop States

ZERO is defined as conduction through the upper cryotron of the flip-flop pair and ONE is defined as conduction through the lower.

The time taken for transition from one time period to the next is a function of the transfer current. If transition occurs at a fixed threshold current value, the final value of the rising current in a given control winding determines the fraction of the  $L/R$  time constant required to reach that threshold value. If the final value is (a) times the threshold value, the time required to reach the threshold is given by:  $t = L/R \ln(a/a-1)$ . The particular multivibrator circuit described completes the round-trip through its six time periods at the rate of 100 to 1,000 times per second depending on transfer current. The higher frequency gives individual time periods of 167 microseconds duration.

To monitor the transitions of one of the flip-flops, an additional cryotron gate is added with its control winding in series with one side of the flip-flop. A current source is connected to its gate circuit. When the control current is zero, the gate circuit is a superconductor and the voltage is zero. When the control current reaches the threshold value, the gate circuit becomes resistive and develops a voltage which is amplified and displayed. Typical values are:  $R = 0.01$  ohm,  $I = 100$  ma;  $V = 1$  millivolt. The true current waveform is not preserved by the monitoring gate due to its inherent nonlinearity plus the sharpening of its transition due to  $I^2R$  heating as it becomes resistive.

### 3.3 Multiterminal Switch

Distributing a pulse among several wires can be accomplished by a cryotron switch (Figure 10). Information is fed into the switch from cryotron flip-flops, here represented by toggle switches. One flip-flop causes the odd or even rows of the switch to be resistive, a second flip-flop causes odd or even pairs to be resistive, a third flip-flop causes odd or even fours to be resistive, and so on. A single path survives as a superconductor, and all of the read current follows that path and thence to the load. With the flip-flops set as shown with binary input 101, row 5 is selected. This particular switch can thus be used as a binary-to-octal converter.

### 3.4 Binary Adder

The principles embodied in the flip-flop and switch can be used to design the stages of a binary adder. The task to be done by each stage is represented by Table II. The (n)th digits of the two numbers to be added are combined with the carry from the (n-1)th stage to form the (n)th digit of the sum and the carry to the (n+1)th stage. Since there are eight possible combinations of the three inputs, an accumulator design can center about the eight-position switch already described. The three inputs operate the control windings and the eight output leads, one of which carries current, can actuate cryotron gate circuits which set up paths to determine sum and carry digits. Figure 11 shows such a stage. The carry input actuates either the upper four rows or lower four rows of the switch, thus eliminating one of the control-winding pairs. The eight gates which operate the sum flip-flop are connected with four in series in each of two parallel paths. The element which is caused by the switch to be resistive diverts the current to the path opposite itself setting

INPUT A	INPUT B	CARRY IN	SUM	CARRY OUT
1	1	1	1	1
0	1	1	0	1
1	0	1	0	1
0	0	1	1	0
1	1	0	0	1
0	1	0	1	0
1	0	0	1	0
0	0	0	0	0

Table II. Binary Addition Table

the sum flip-flop to its proper state. A similar group of gates develops the carry for the following stage. Note that all circuits are in series from a current source power supply.

The foregoing binary adder design is described to illustrate the way in which switches and gates can be interconnected. A design having fewer cryotrons per stage is available wherein the carry is handled by a lattice network shown in Figure 12. The label beside each of the six control windings indicates when it is to be energized. The  $A = B = 0$  and  $A = B = 1$  windings can each be made of two cryotrons in a parallel AND circuit and then directly excited from the A and B flip-flops, or the current necessary to excite them can be derived from a four-position cryotron switch. The latter method has an advantage since  $A = B$  and  $A \neq B$  currents are useful in forming the sum digit. After the current has passed through the  $A = B = 1$  and  $A = B = 0$  coils in the carry network, the two coil ends can be combined to provide a current  $A = B$ . This involves the outputs of two of the four output terminals of the 4-position switch. The other two output terminals can be tied together directly to provide a current if  $A \neq B$ . The sum digit can then be simply formed in the following way:

Note that the sum is ONE if  $A = B$  and the carry in is ONE; ZERO if  $A = B$  and the carry in is ZERO; ONE if  $A \neq B$  and the carry in is ZERO; ZERO if  $A \neq B$  and the carry in is ONE. The  $A = B$  and  $A \neq B$  currents can therefore be used to route the carry input currents to the proper side of the sum flip-flop. Figure 13 is a schematic drawing of such a stage of an accumulator, abbreviated in that the A and B flip-flops are not shown, nor is the transfer link from the sum flip-flop back to the A flip-flop shown (used during accumulative addition and subtraction). In this design one notices the convenience of interconnecting cryotrons without regard to d-c levels, very much as relay contacts are placed in relay computer circuitry.

### 3.5 Stepping Register

Stepping registers are commonly used for receiving digital information in serial form at one pulse repetition frequency and after a predetermined number of binary bits have been stored, shifting the information out at a different frequency. A second common use for shifting registers is to accomplish the conversion between digital information in serial and parallel form. The stepping registers in common use are made of vacuum tubes, transistors, or magnetic cores. Cryotrons can also be used in the same service. Each stage of the shift register consists of two cryotron flip-flops with read-in and read-out cryotrons. One transfer circuit sets the second of the two flip-flops of each stage to correspond to the state opposite that of the first. The coupling link to accomplish this is similar to the one described in section 3.2 which interconnects stages of the multivibrator. A second transfer circuit sets the first flip-flop of each stage to correspond to the state opposite that of the second flip-flop of the stage to its left. A line of such stages serves as a shifting register, capable of shifting digital information to the right. Information (ONE's or ZERO's) fed into the first flip-flop in synchronism with the second of the two transfer pulses (called ADVANCE B pulse), will advance through the stepping register one stage for each pair of transfer pulses, ADVANCE A and ADVANCE B, which are displaced in time. Figure 14 shows two stages of a cryotron stepping register. Parallel output gates are not shown.

### 3.6 Coincident-Current Circuits

Many interesting circuits can be made of cryotrons with two or more control windings wound over each other in such a way that the net magnetic field affecting the central wire is due to the sum of the magnetic fields of the individual windings. The d-c cryotron characteristics of Figure 4 are sufficiently sharp in their transition between superconducting and normal states to allow the transition to result from the sum of two half-amplitude fields or even three one-third-amplitude fields. A coincident-current circuit of this type is useful for the selection of cryotron flip-flops placed at the intersection of the rows and columns of a matrix. A one-half-amplitude pulse is applied to the flip-flop control windings along a row, and a similar pulse to the flip-flop control windings along a column. The flip-flop at the intersection of that row and column can thus be placed in one of its two states; all other flip-flops in the matrix are unaffected.

If two such control windings are operated in opposition in such a way that the magnetic field of one subtracts from that of the other, a gate circuit of the "exclusive OR" type is available, wherein a flip-flop is set if either pulse A or pulse B occurs, but not if they both occur. Operation in this manner takes advantage of the fact that control depends on the magnitude of the controlling field and not on its polarity.

#### 4. Engineering a Cryotron System

##### 4.1 Low-Temperature Environment

The most unusual requirement of a cryotron system is that it operate at a temperature near the absolute zero. Ten years ago this requirement would have precluded serious thought of such a system. Today, however, such an operating temperature is relatively easy to achieve.<sup>7</sup> This change is mainly due to the work of Samuel C. Collins whose helium liquifiers revolutionized the field of low-temperature physics. Arthur D. Little, Inc. of Cambridge, Massachusetts, has built seventy Collins helium liquifiers of a 4-litre-per-hour capacity. The liquifier at M.I.T. liquifies 27 litres per hour. Storage of liquid helium has also improved. Commercially available double Dewars which use liquid nitrogen in the outer Dewar lose less than one percent of their liquid helium per day.

The heat dissipated by a cryotron system causes evaporation of the helium. If the average power dissipated per cryotron is  $10^{-4}$  watt, an estimate based on present experimental units, a 5,000-cryotron computer would dissipate one-half watt. The latent heat of vaporization of liquid helium at 4.2 K is 5 calories per gram,<sup>8</sup> its density is 0.1257, and therefore one-half watt corresponds to an evaporation rate of 0.93 litre per hour. A continuous system which recycles helium would be most economical for a stationary installation; a ten- or twenty-litre charge at the time of launching would suffice for portable systems.

The temperature of a liquid helium bath can be controlled by controlling the pressure of the bath. Table III gives the boiling point of helium at various pressures. Below 2.19 K, the so-called lambda-point, liquid helium exhibits unusual properties which may prove useful in a cryotron system. A second phase of liquid helium appears which acts as a second fluid free to move through the first fluid with no friction. This zero-viscosity component is able to conduct heat with zero temperature gradient. It thus flows intimately in and around any structure immersed in it and allows rapid conduction of heat away from the structure. If heating is a problem in a cryotron system, operation in this temperature region should provide a solution. It is interesting, incidentally, to watch a liquid helium bath being pumped down. It may boil rather vigorously until the temperature drops below 2.19 K at which point the surface becomes perfectly still; heat is conducted through the liquid and liberated at the surface rather than on the container walls which causes boiling.

---

7. C. A. Swenson and A. G. Emslie, "Low-Temperature Electronics," Proc. IRE, vol. 42, No. 2, pp. 408-413; February 1954.

8. W. H. Keeson, Helium, (Book), Elsevier 1942.

Pressure m.m. Hg.	Temperature degrees K	Pressure m.m. Hg.	Temperature degrees K
0.001	0.657	720.	4.156
0.01	0.791	730.	4.170
0.1	0.982	740.	4.184
1.	1.269	750.	4.198
10.	1.743	760.	4.211
100.	2.638	770.	4.225
200.	3.067	780.	4.239
300.	3.368	790.	4.252
400.	3.605	800.	4.266
500.	3.803	900.	4.40
600.	3.975	1000.	4.52
700.	4.127	1500.	5.03
710.	4.141	1720.	5.20

Table III. Boiling Point of Helium

#### 4.2 Physical Construction

Figure 15 shows some experimental cryotron circuits. They are mounted at the ends of three-foot cupro-nickel tubes for immersion in a liquid helium storage vessel. Power supply and signal wires come up through the center of the tube. The experiments read chronologically from the large probes on the right which were used for d-c characteristic measurements to the three-flip-flop multivibrator circuit on the left which contains nineteen active elements. A closeup of the latter experiment is shown in Figure 16. The individual elements are those whose d-c cryotron characteristics are given by Figure 4. Spotwelding has been used to interconnect niobium and tantalum wires. Nickel lugs, while not superconducting, have proven useful for mounting. They both spotweld and solder nicely and careful design minimizes the resistance they introduce. The feasibility of using superconductive etched-wiring boards is under study. In these, lead would form the superconductive paths.

Many materials are used in the construction of circuits to operate in liquid helium. Ordinary wire insulation (enamel, silk, glass, Formex, Formvar, etc.) shows no sign of failure after repeated immersion. One experiment using wooden coil forms glued together with Duco Cement was successful. Scotch Electrical Tape, while it freezes, seems to hold well. Commercially available feedthrough and standoff insulators have been used without any sign of cracking. Metals in general are much stronger at

extreme low temperatures. Some are relatively good thermal insulators (stainless steel and cupro-nickel) and may be used for mechanical support. There is no basis for the common impression that everything falls apart just below JAN specifications (-85 C).

#### 4.3 Input, Output, and Power Supply

Input pulses to cryotron circuits involve current amplitudes which are easily achieved in the terminal equipments commonly associated with digital computers. Since the voltage level is low, input of information to a cryotron system involves no unusual problems.

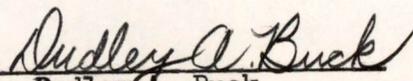
Connecting the output pulses of a cryotron system to terminal equipment, on the other hand, is difficult due to the low power level of the cryotron circuitry. Power cryotrons can be designed to increase the power level, but it appears that vacuum-tube or transistor amplifiers are necessary to bring the level up to that of most output equipments. Magnetic amplifiers with superconductive control windings are an interesting possibility for power amplification.

Power supplies for cryotron systems are easy to achieve. The low impedance of the circuitry dictates a current-source power supply. A battery with a series resistance is adequate.

#### 5. Conclusion

The cryotron in its present state of development is a new circuit component having power gain and current gain so that it can be used as an active element in logical circuits. It is easily and inexpensively fabricated from commercially available materials and its size is small. Extrapolating the volume occupied by the present experimental circuits to larger numbers of components indicates that a large-scale digital computer can be made to occupy one cubic foot, exclusive of refrigeration and terminal equipment. The power required by such a machine extrapolates to about one-half watt, once again excluding refrigeration and terminal equipment. The reliability of cryotron circuitry is not known, but it is anticipated that operation in an inert helium atmosphere at a temperature near to absolute zero where chemical activity and diffusion processes are essentially stopped promises a high degree of reliability. The circuit noise level is similarly not known, but due to the low temperature, very little thermal fluctuation noise is anticipated. The device is at present somewhat faster than electromechanical relays, but far slower than vacuum tubes and transistors. A program is under way to increase the speed.

Signed

  
Dudley A. Buck

DAB/jg

Drawings Attached:

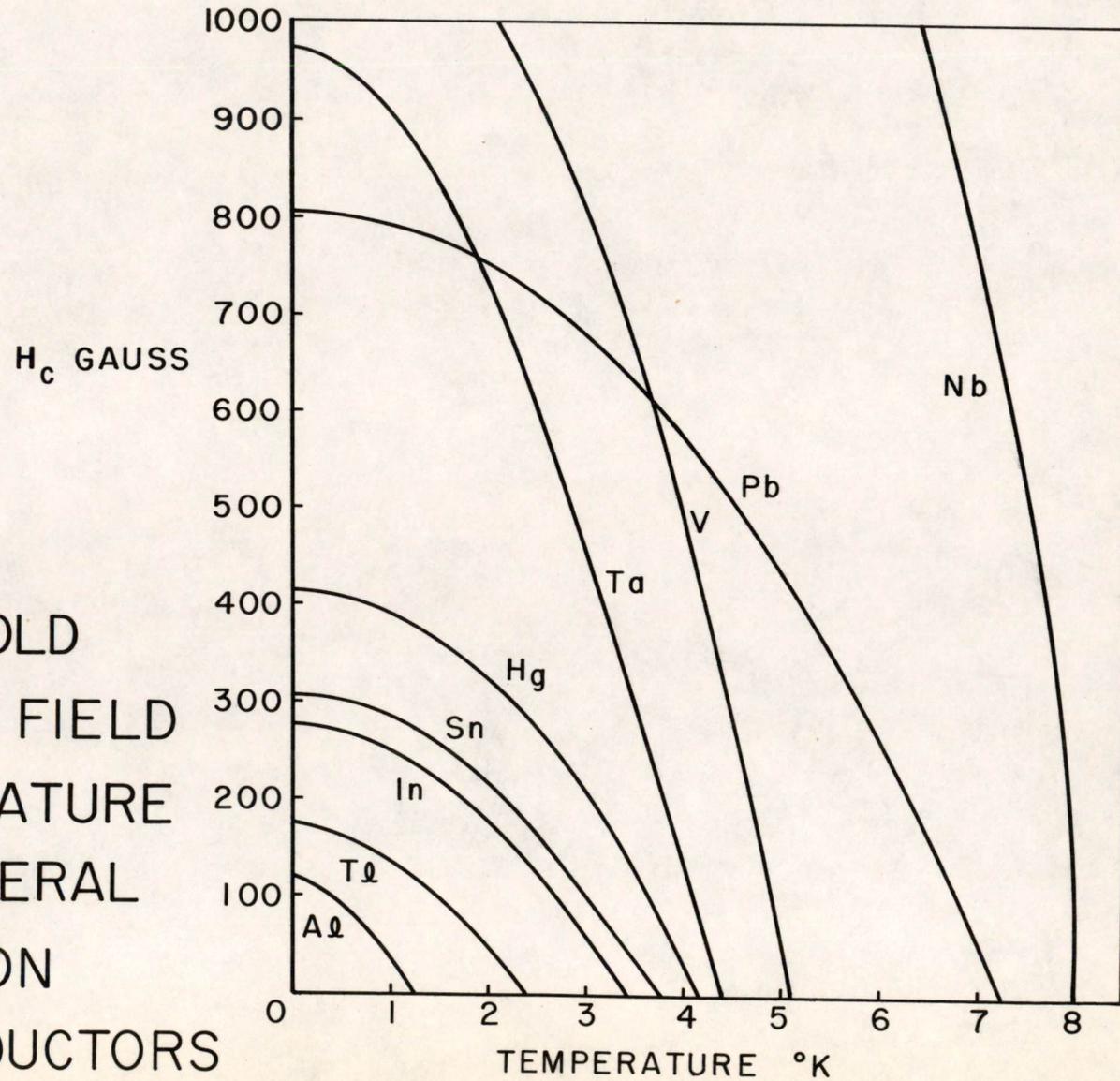
Figure 1 - C-63334  
Figure 2 - A-63087  
Figure 3 - A-63088  
Figure 4 - B-63091  
Figure 5 - B-63491  
Figure 6 - A-63090  
Figure 7 - C-63294  
Figure 8 - C-63291  
Figure 9 - C-63293  
Figure 10 - C-63292  
Figure 11 - C-63338  
Figure 12 - B-63335  
Figure 13 - C-63336  
Figure 14 - C-63337  
Figure 15 - A-63679  
Figure 16 - A-63678

Distribution List:

Group 63 Staff  
Group 35 Staff  
Group 65 Staff  
M. A. Herlin - D-321  
B. G. Farley  
R. P. Mayer  
L. L. Sutro  
R. F. Jenney  
R. W. Sittler  
A. Shortell, Barta  
Prof. W. K. Linvill  
Prof. S. C. Collins - 41-204  
Prof. J. B. Wiesner, 20A-122  
Prof. R. M. Fano, 20A-118  
Prof. S. J. Mason, 20A-118  
Prof. A. R. von Hippel, 4-244  
Prof. S. H. Caldwell, 10-303A  
Prof. R. H. Frazier, 10-403  
Prof. T. S. Gray, 4-212  
Prof. A. Kusco, 10-159  
Prof. R. M. Saunders, 10-198  
Prof. R. B. Adler, 20A-118  
Prof. A. B. Van Remnes, 3-401  
Prof. D. J. Epstein, 20C-105  
Prof. D. A. Huffman, 10-338  
R. D. Thornton, 10-441  
N. DeClaris, 4-232  
A. Smakula, 20B-136  
N. Alperin, Barta

C-63334  
F-2969  
SN-1220

FIG. 1  
THRESHOLD  
MAGNETIC FIELD  
vs TEMPERATURE  
FOR SEVERAL  
COMMON  
SUPERCONDUCTORS



A-63087  
F-2862  
SN-1169

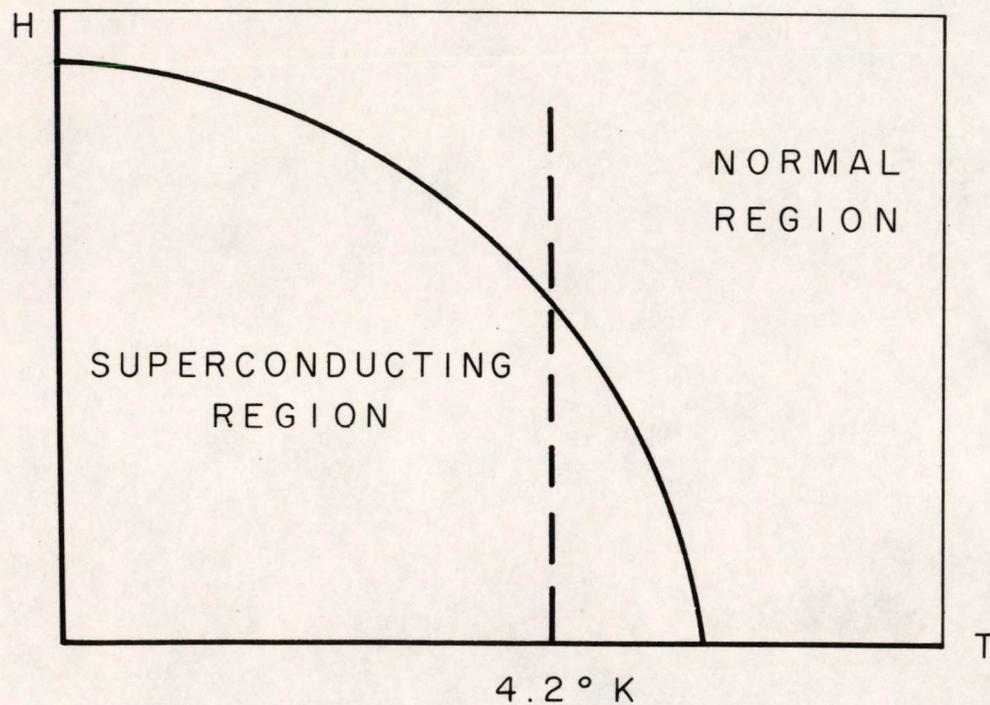
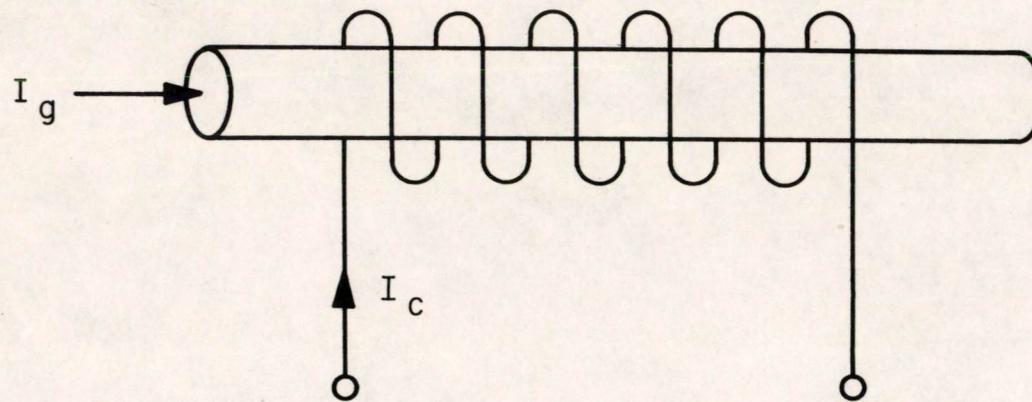


FIG. 2  
THRESHOLD MAGNETIC FIELD  
AS A FUNCTION OF TEMPERATURE  
FOR A SUPERCONDUCTOR



SINGLE CRYOTRON

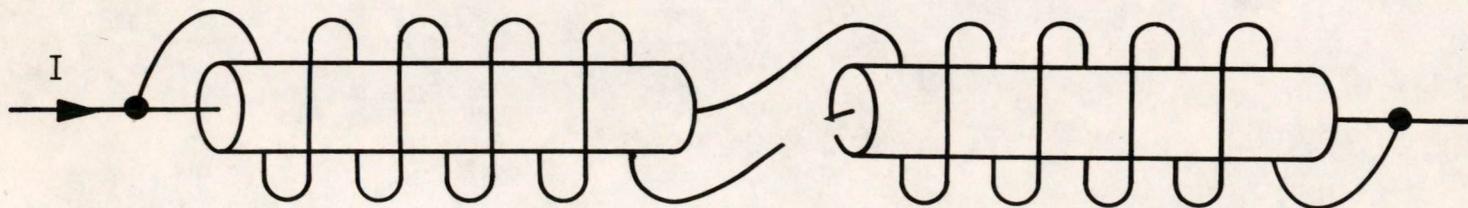


FIG. 3  
CRYOTRON BISTABLE ELEMENT  
(FLIP-FLOP)

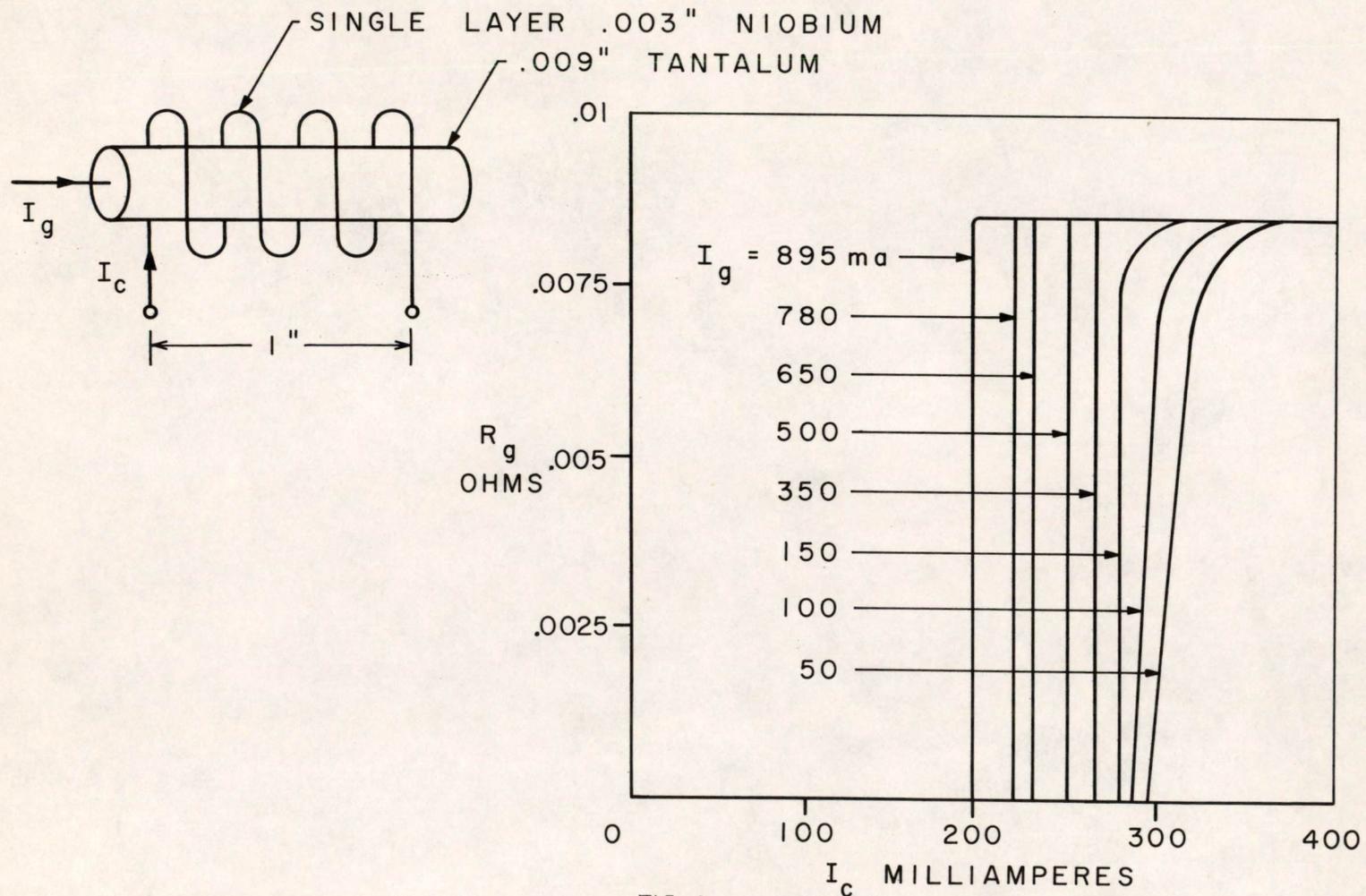


FIG. 4  
CRYOTRON GATE RESISTANCE  
vs. CONTROL CURRENT

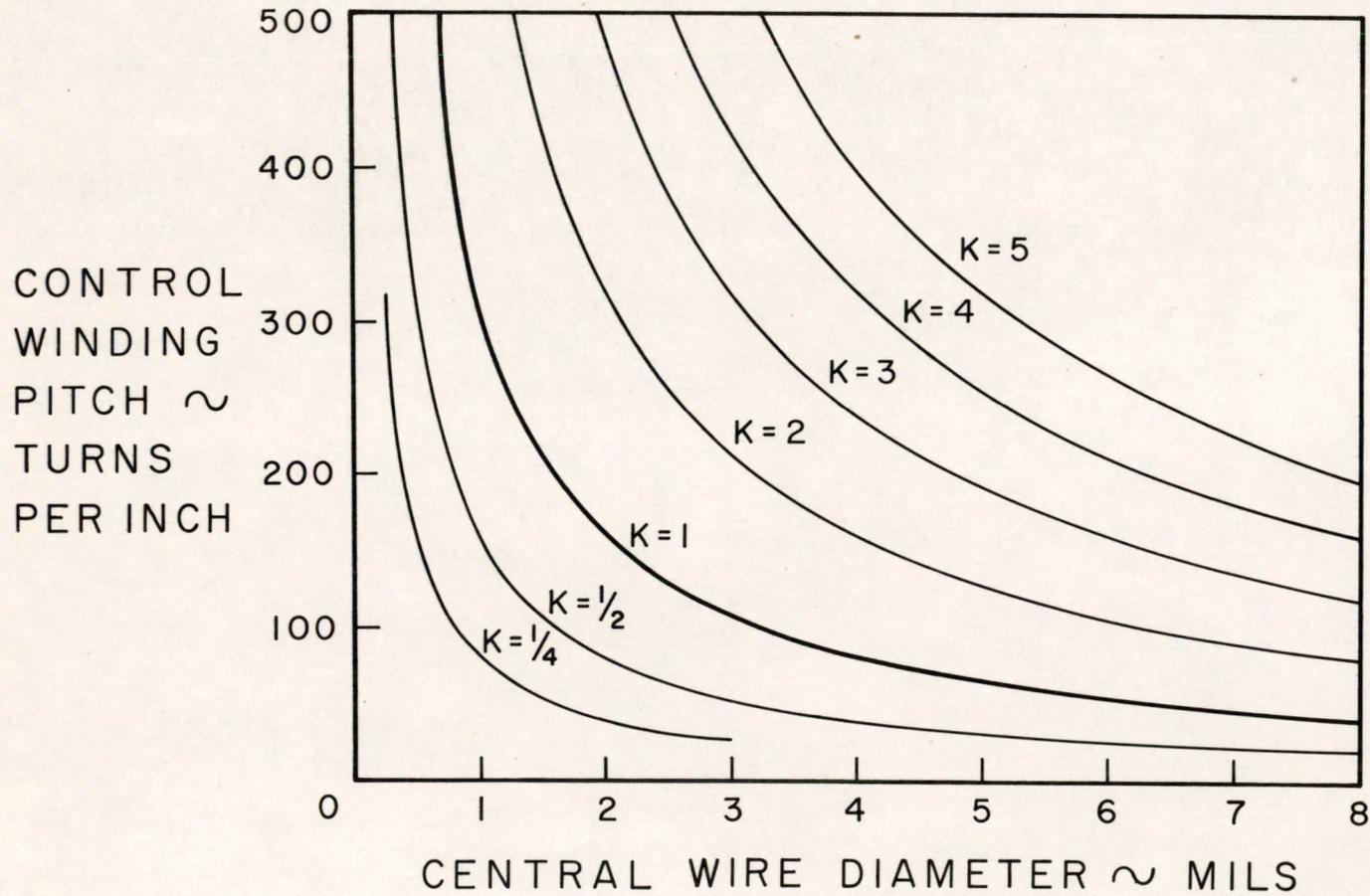


FIG. 5

CURRENT GAIN vs CONTROL WINDING  
PITCH AND CENTRAL WIRE DIAMETER

A-63090  
F-2864  
SN-1171

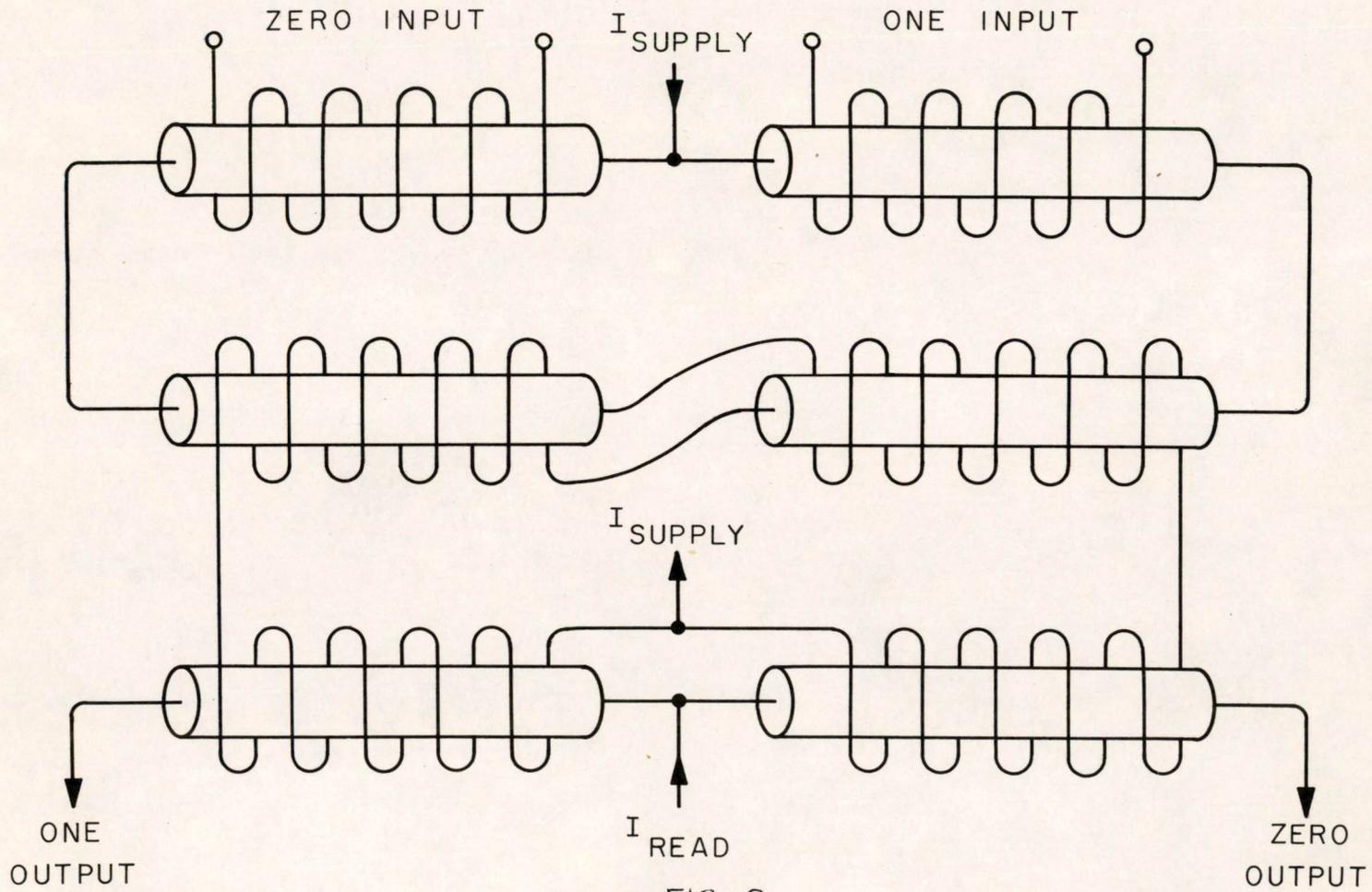
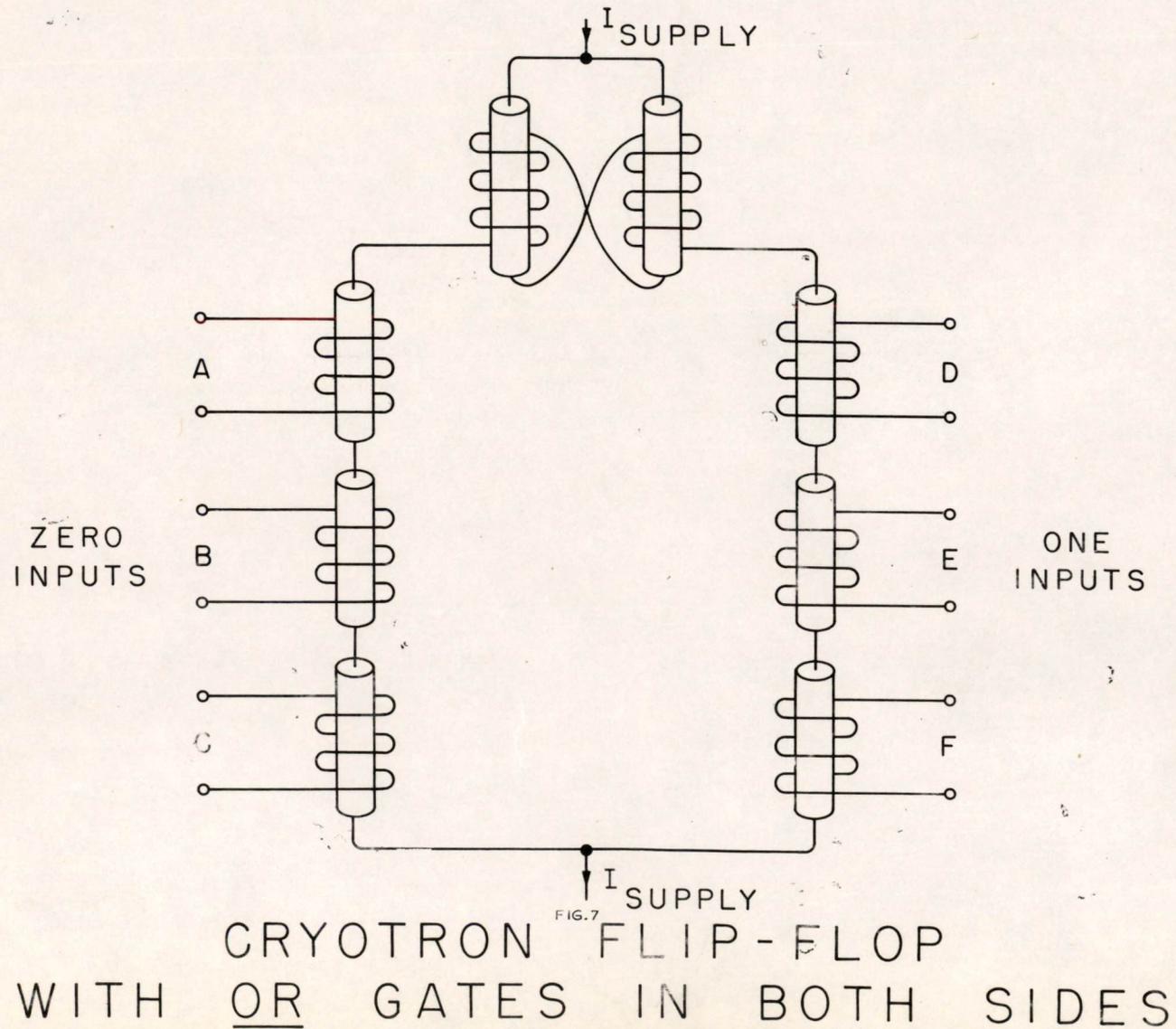


FIG. 6

CRYOTRON FLIP-FLOP WITH READ-IN  
CRYOTRONS AND READ-OUT CRYOTRONS

C-63294  
F-2968  
SN-1219



C-63291  
F-2965  
SN-1216

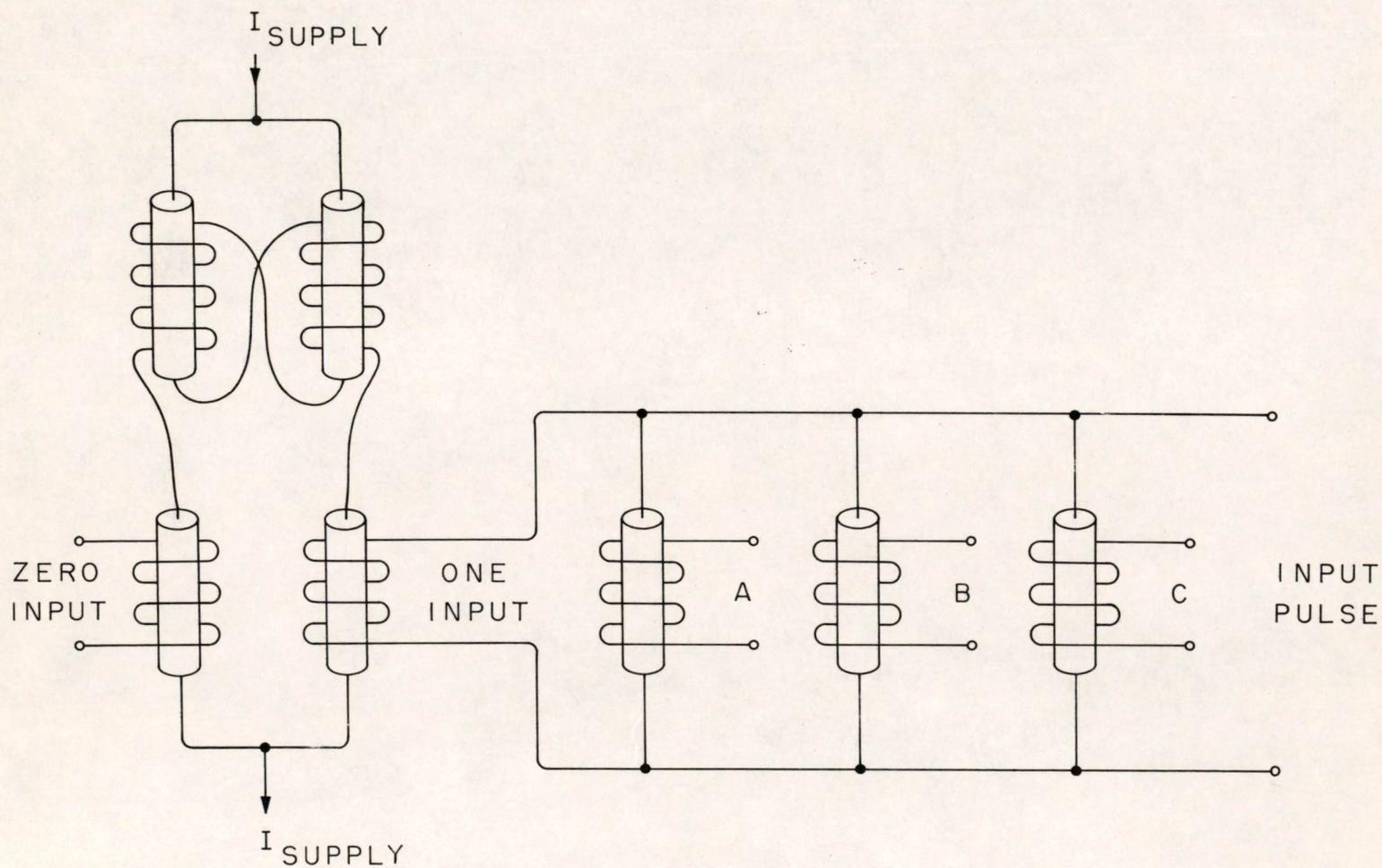


FIG. 8

CRYOTRON FLIP-FLOP  
WITH AND GATES IN ONE SIDE

C-63293  
F-2967  
SN-1218

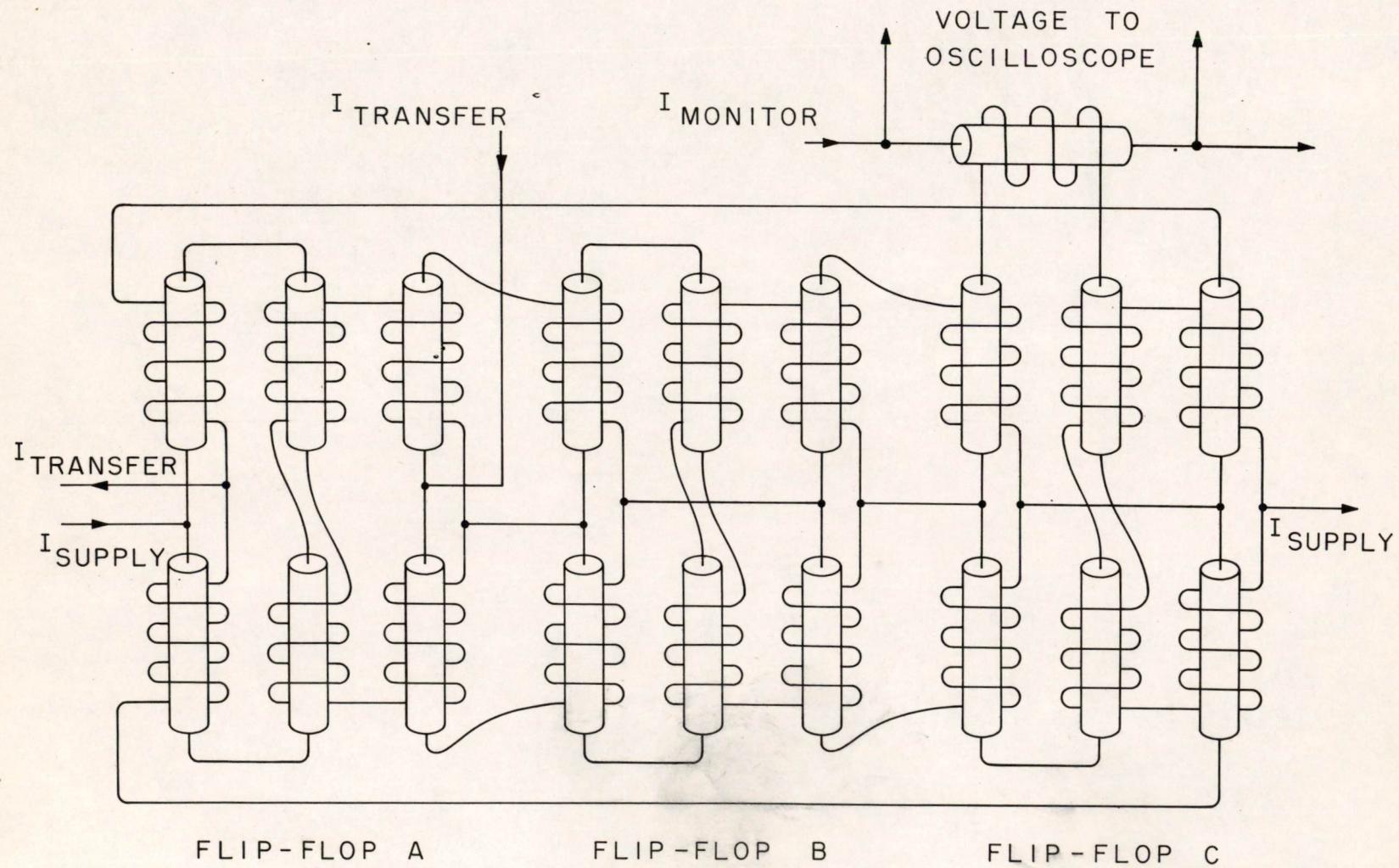


FIG. 9

# 3-FLIP-FLOP CRYOTRON MULTIVIBRATOR

C-63292  
F-2966  
SN-1217

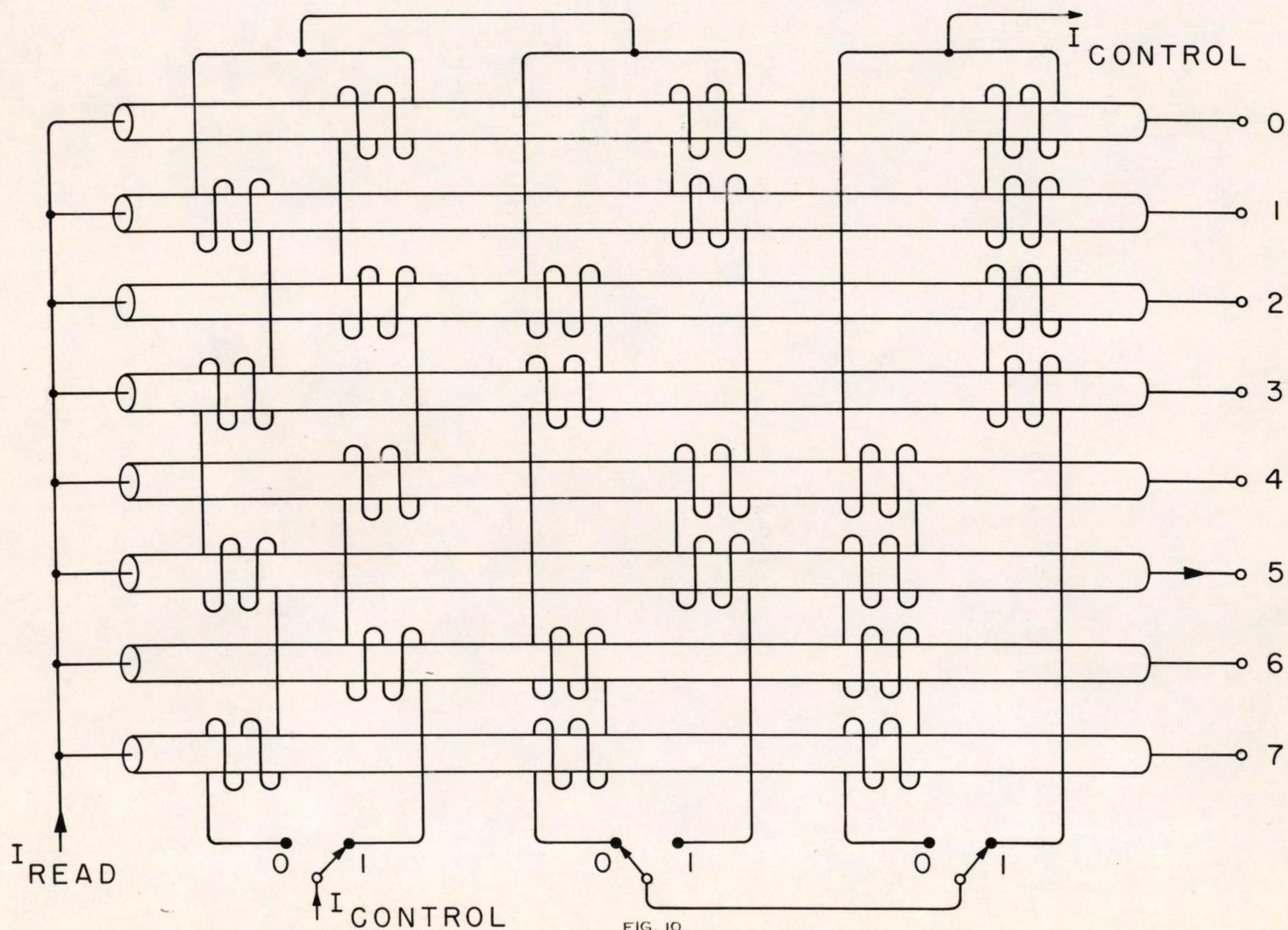


FIG. 10

# 8-POSITION CRYOTRON SWITCH

C-63338  
F-2973  
SN-1224

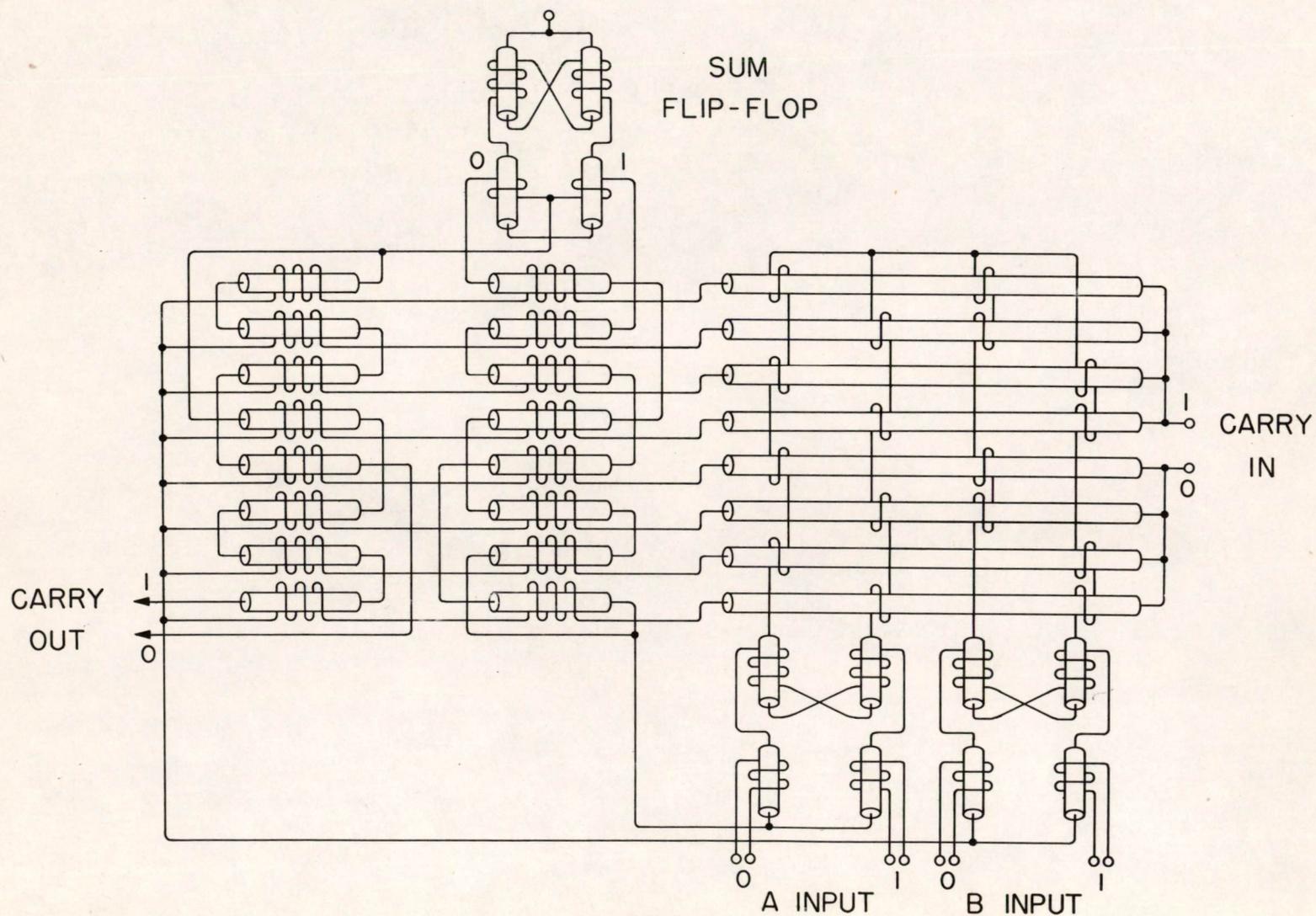


FIG. II

ONE STAGE OF A BINARY ADDER

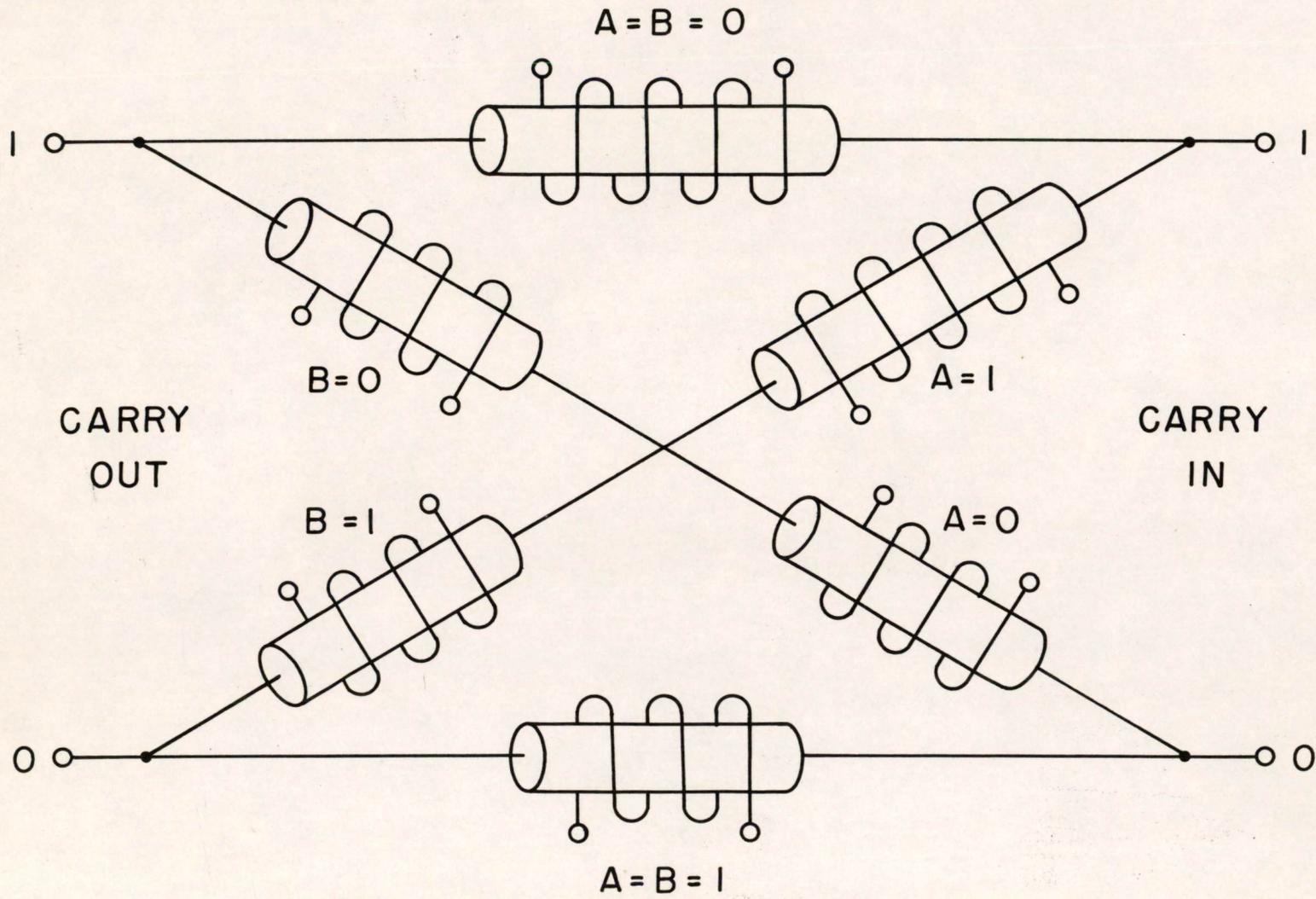


FIG. 12

# CARRY NETWORK

C-63336  
 F-2971  
 SN-1222

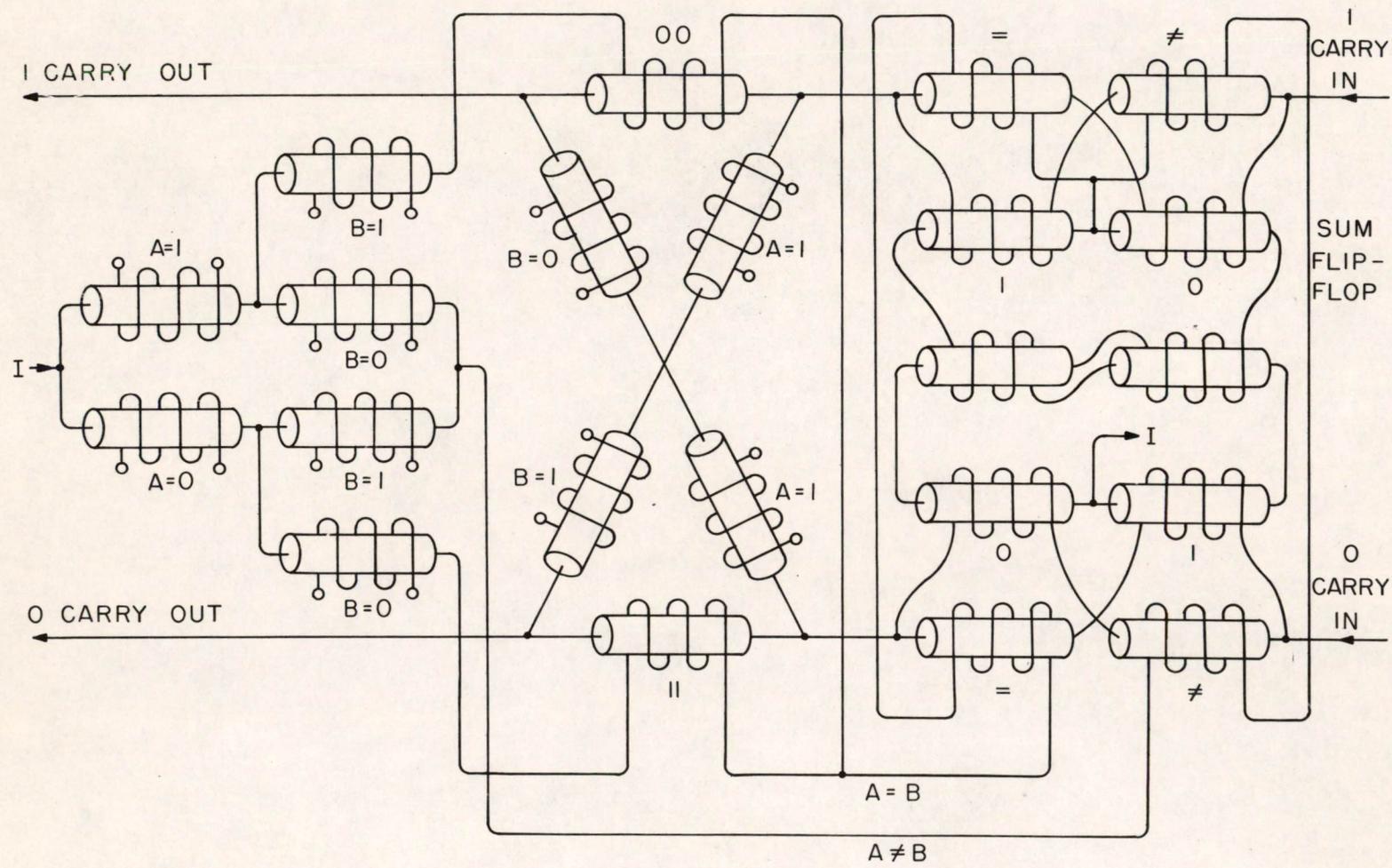


FIG. 13  
 BINARY ACCUMULATOR STAGE

C-63337  
F-2972  
SN-1223

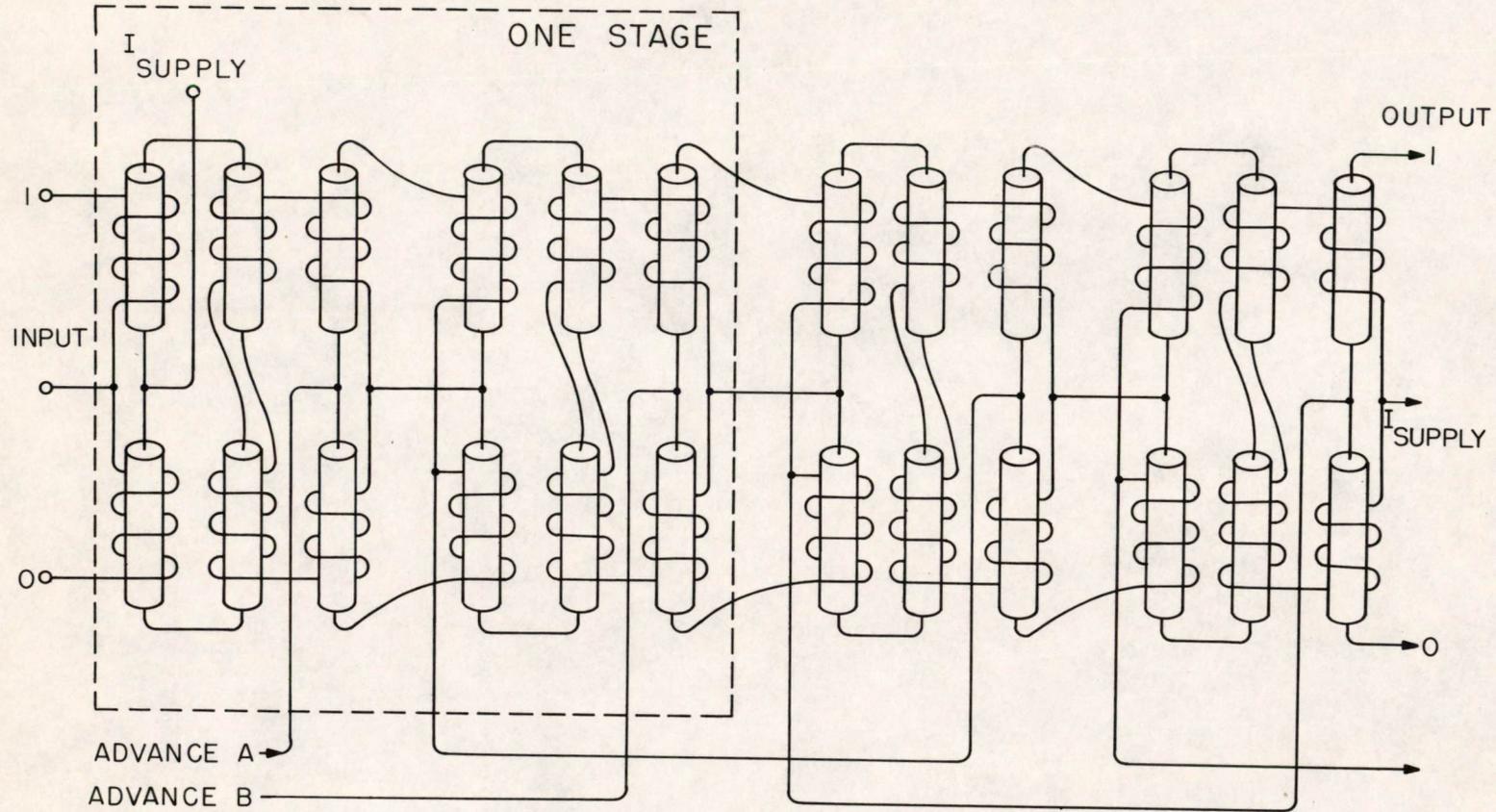


FIG. 14

TWO STAGES OF A  
CRYOTRON STEPPING REGISTER

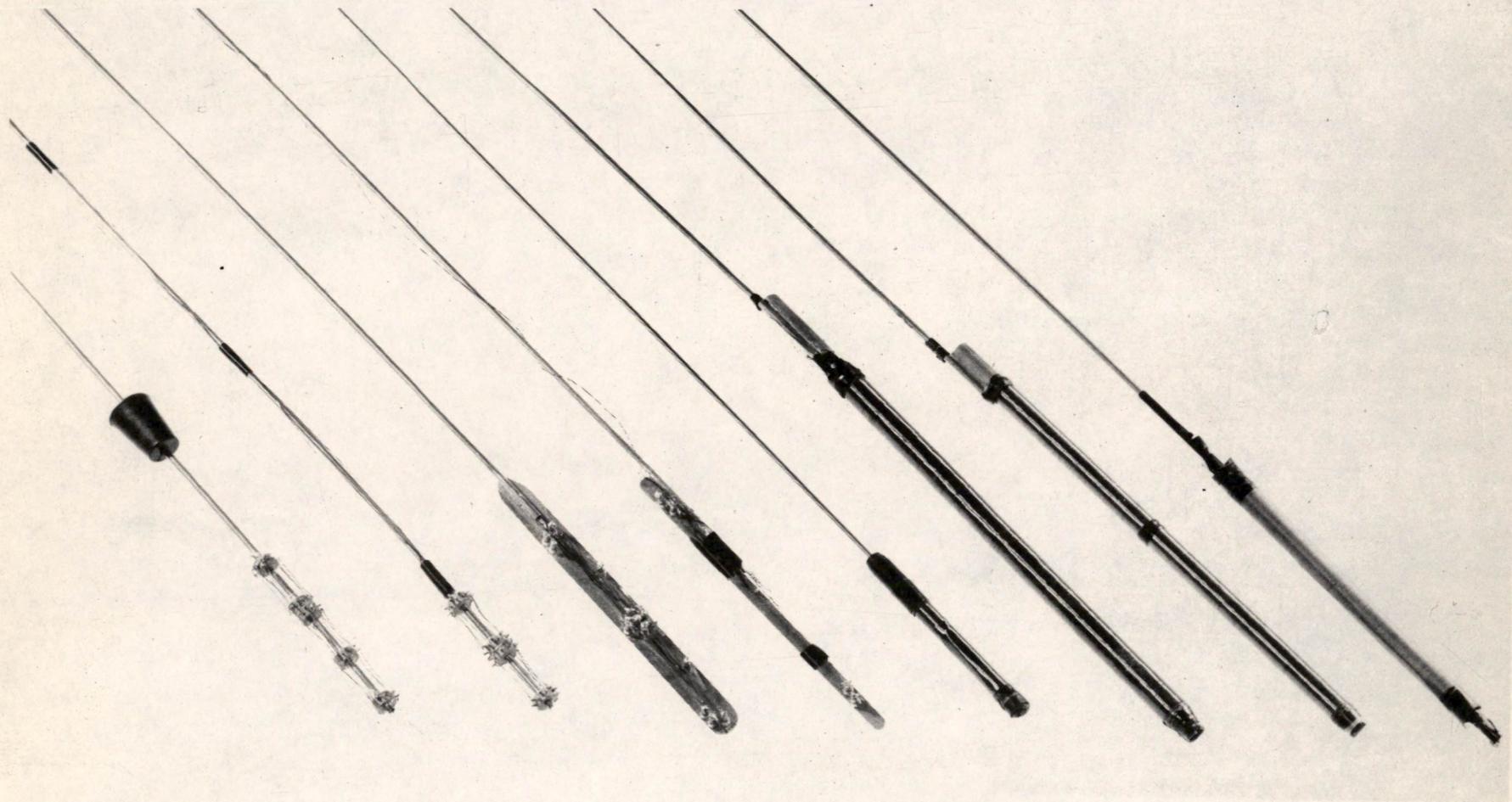


FIG. 15  
EXPERIMENTAL CRYOTRON CIRCUITS

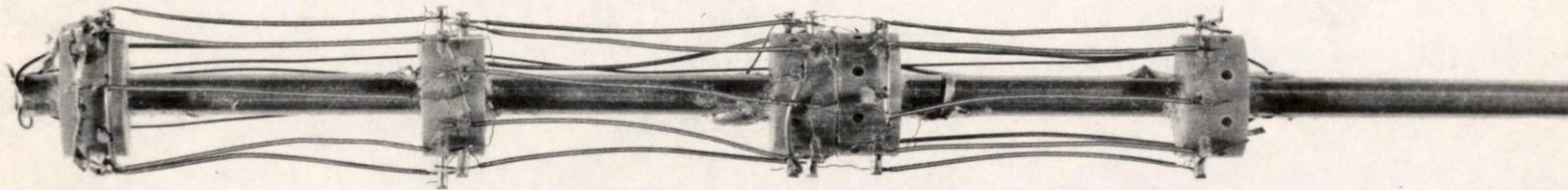


FIG. 16  
3-CRYOTRON-FLIP-FLOP MULTIVIBRATOR

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: PULSE TRANSFORMER AMPLIFIERS

To: David R. Brown

From: Melvin M. Cerier

Date: September 7, 1955

Approved: Torben Meisling  
Torben Meisling

Abstract: A comparison between the inverter amplifier and the transformer-transistor amplifier shown in Figs. 1 and 3 reveals that the transformer-transistor amplifier is more satisfactory in several respects. The transformer type of amplifier can usually drive a larger load than the inverter amplifier. An experimental transformer type of amplifier was built and is described. The approximate power dissipated in a grounded emitter transistor under pulse conditions is calculated, and it is shown that a surface-barrier transistor under these conditions can carry large currents before power dissipation ratings are exceeded.

### Introduction

Situations arise, in parts of a system, where one pulse amplifier is called upon to drive many bases. Simple inverter amplifiers are often not satisfactory for this purpose. An inverter amplifier usually cannot drive a lot of bases into saturation, and the fact that this type of amplifier inverts the pulse often makes the logical circuit design more complicated than it might otherwise be. Both of these difficulties might be avoided by using transformer coupling. The properties of transformers that make this possible are; (a) isolation between primary and secondary, and (b) matching of the primary circuit impedance to the load. The matching property allows the choice of a load impedance for the amplifier transistor without much regard for the driven impedance. The isolation property allows the output to have the same polarity as the input.

This note is an attempt to explain the disadvantages of the inverter amplifier and to show what might be expected from a transformer-transistor amplifier.

In the appendix, a region of operation is described for a grounded emitter transistor. The limits of operation are the maximum allowable supply voltage and the maximum allowable average power dissipated.

The Simple Inverter

Figure 1 shows a typical circuit of a simple inverter driving many bases. The assumption is made that either all transistors are "off" and look like open circuits, or are saturated and look like short circuits between emitter and collector. When the voltage at the base of the amplifier transistor is negative, this transistor is assumed to be saturated and the voltage at its collector is zero. This zero voltage appears at the base of the driven transistors and keeps them turned "off." When the voltage at the base of the amplifier transistor is zero it is assumed that this transistor is "off" and the voltage at its collector is negative. This negative voltage saturates the driven transistors. All of the load current must flow through  $R_{c1}$ . When the amplifier transistor is changed from "off" to "on," the driven transistors are changed from "on" to "off." The load is turned off because the input is shorted to ground through the collector to emitter short of the amplifier transistor. The output of this amplifier is inverse of the input, i.e. the output is negative when the input is zero. The amplifier transistor must carry maximum current whenever it is "on."

If the current gain of the transistors at saturation is  $\beta$ , the base current  $I_b$ , necessary to saturate a driven transistor is:

$$I_b = \frac{V_{c2}}{\beta R_{c2}} \quad (1)$$

The condition for operation is that the actual base current,  $i_b$ , be greater than the amount necessary to saturate. This condition can be written

$$i_b \geq \frac{V_{c2}}{\beta R_{c2}} \quad (2)$$

If each base is assumed to look like a battery  $e_b$  in series with a resistance  $r_b$ , as in Figure 2, the current available from the inverter,  $ki_b$ , is given by

$$ki_b = \frac{V_{c1} - e_b}{R_{c1} + \frac{R_b + r_b}{k}} \quad (3)$$

where  $k$  is the number of driven stages. Eliminating  $i_b$  from (2) and (3) and solving for  $k$ , results in an expression which shows the maximum number of stages that can be driven into saturation from a simple inverter.

$$k \leq \beta \left( \frac{R_{c2}}{R_{c1}} \right) \left( \frac{V_{c1} - e_b}{V_{c2}} \right) - \frac{R_b + r_b}{R_{c1}} \quad (4)$$

1. E. U. Cohler, "Typical SBT Static Characteristics," Lincoln Laboratory Memorandum 6M-3649, May 31, 1955.

The number of stages that can be driven can be made large by making  $R_{c2}$  larger than  $R_{c1}$ ,  $(V_{c1} - e_b)$  larger than  $V_{c2}$ , and  $(R_b + r_b)$  as small as possible compared to  $R_{c1}$ . In most cases, however, the driving stage has the same circuit values as the driven stages. If  $e_b$  is small compared to  $V_{c1}$  and  $(R_b + r_b)$  is small compared to  $R_{c1}$  the maximum number of stages that can be driven is,

$$k_{\max} = \beta. \quad (5)$$

If the stages are not all alike, it would be desirable to make  $R_{c1}$  small, and  $V_{c1}$  large, in order to increase  $k$ . This would mean the saturation current for the driving stage would be comparatively large.

For a typical case assume that  $R_{c1} = R_{c2} = 600$  ohms,  $V_{c1} = V_{c2} = 4$  volts,  $e_b = 0.3$  volts,  $R_b + r_b = 300$  ohms, and  $\beta = 7$ . Then,

$$k \leq 6,$$

or the maximum number of stages than can be driven is 6.\*

#### Transformer-Transistor Amplifier

Figure 3 shows a typical circuit for a pulse transformer amplifier driving many bases. A resistance,  $R_t$ , might be necessary to terminate properly the pulse transformer, or a transmission line driven by the transformer.

The amplifier transistor should never be turned on for long periods of time since the full supply voltage would appear at the collector when the transformer saturated, and a great deal of power would be dissipated in the transistor. Normal operation would be as follows. A pulse of negative current would be applied to the base of the amplifier transistor. This transistor would saturate and a pulse of current would flow in the primary. Secondary current would flow which would turn the driven transistors "on." Since the time integral of the secondary voltage must be zero, there will be an overshoot of the secondary. After the input pulse is over, the secondary voltage will tend to go positive. This positive voltage could be used to advantage to unsaturate the driven transistors. The amplifier transistor in this case will only carry maximum current when the load is a maximum.

Again the base current necessary to saturate the driven transistors is given by (1). Also, it is assumed that all transistors are either "off" or "on." There are two restrictions that must be placed on the relations among all of the circuit parameters to insure proper operation of the circuit. The first condition is that the secondary voltage must be large enough to saturate all of the driven transistors. Making the same assumptions as before, this condition can be written,

\* In practice, the charge and discharge time constants of  $C_b$  are small compared to a pulse length. The energy stored by this capacitor is usually small compared to the energy delivered to the transistor. Therefore,  $C_b$  is neglected in both the preceding and following calculations.

$$\frac{\frac{V_{cl}}{n} - e_b}{R_b + r_b} \geq I_b \quad (6)$$

Solving for n gives an expression which places an upper limit upon the turns ratio.

$$n \leq \frac{V_{cl}}{e_b + (R_b + r_b) I_b} \quad (7)$$

The other restriction is that there be enough current available to saturate all of the driven transistors. The current available will be limited either by the maximum current allowed to flow in the collector of the amplifier transistor, or by the current available to drive the base of this transistor. In either case there is some maximum collector current,  $I_c$ , available. This condition can then be written,

$$n I_c \geq k I_b + \frac{V_{cl}}{n R_t} \quad (8)$$

If it is assumed that the base current available for the amplifier transistor is the same as that necessary to saturate the driven transistors,  $I_c$  is given by,

$$I_c = \beta I_b \quad (9)$$

Combining equation (9) with (8) and solving for k,

$$k \leq \beta n - \frac{V_{cl}}{n I_b R_t} \quad (10)$$

For a typical example choose the following conditions:  $V_{cl} = 4$  volts,  $V_{c2} = 3$  volts,  $e_b = 0.3$  volts,  $R_b + r_b = 300$  ohms,  $R_{c2} = 600$  ohms,  $\beta = 7$ , and  $R_t =$  . Then from (8)  $n = 2.22$ , and  $n = 2$  is the maximum turns ratio. From (10)  $k \leq 14$ , and  $k = 14$  is the maximum number of stages that can be driven. Note that this is better than a 2 to 1 improvement over the simple inverter case.

### Experimental Work

A transformer-transistor amplifier was built using a Lincoln Laboratory stock transformer No. 724-001. This is a Ferramic H core with 27 turns on the primary and 9 turns on the secondary. This amplifier was designed to put out 0.1 microsecond pulses at a pulse repetition frequency of 5 megacycles. The design procedure was strictly experimental.

The circuit diagram of this amplifier is shown in Figure 4. The resistors in the primary and secondary provide proper damping; the diode puts the primary resistor in the circuit only for the part of the wave that overshoots. The inductance in the secondary is to make the overshoot last only 0.1 microsecond.

Figure 5 shows the measured output voltage plotted as a function of the input voltage. This curve has a sharply rising part and a flat part as would be desirable in a system. The sharply rising part would be useful to discriminate against noise in a system, and the flat part would be useful for pulse height standardization. The curve does not show a gain of over unity because of the voltage drop across the large resistance in the base circuit. This large resistance provides protection against large input pulses that might be obtained from Burroughs equipment.

The output pulse width is shown plotted against input pulse amplitude in Figure 6. The output pulse width was measured from points where the voltage was 10 percent of the maximum voltage. The pulse width never gets larger than 0.1 microsecond. This curve has a sharply rising part and a flat part, also. From Figures 5 and 6 it is easy to see that for any input voltage above 2 volts the outputs have similar shapes.

Figures 7, 8, and 9 show some input and output waveforms. The pictures show that the overshoot lasts only 0.1 microsecond. That the amplifier output level does not change with the pulse repetition rate is shown by the pulse burst input-output pictures. The output waveform is shown for a continuous 5 megacycle input.

This amplifier was used to drive a pulse distributor. The load consisted of two flip-flops and another amplifier similar to the one shown. The system worked well and was quite stable and insensitive to changes in the input amplitude, and frequency. The system was also tried with 70  $\mu$ sec pulses at about 7 mc and it worked well under these conditions.

It was found that this amplifier could drive into saturation as many as ten grounded emitter stages with a saturation current of 4.5 milliamperes each.

### Conclusions

The output of a transformer-transistor amplifier can just as easily be inverted as not, which might make logical circuit design easy.

Because of the added flexibility of the turns ratio, the transformer-transistor amplifier can drive more bases than a simple inverter.

The transistor in the simple inverter is called upon to carry maximum current whenever it carries current. Thus this transistor must switch maximum power whenever it switches. Since the transistor in the transformer-transistor amplifier only carries maximum current when the load is maximum it must switch maximum power only when the load is maximum.

The biggest disadvantage of the transformer type of amplifier is the difficulty of transformer design. Relating the performance of the pulse transformer to the design parameters is extremely difficult. The best design procedure seems to be an experimental one. On the other hand, the design of a simple inverter is fairly straightforward.

Signed Melvin M. Cerier  
Melvin M. Cerier

MMC/jg

Drawings attached:

Figure 1 and 2 - A-63154  
Figure 3 - A-63155  
Figure 4 - A-63156  
Figure 5 - A-63157  
Figure 6 - A-63158  
Figure 7, 8, 9 - A-63159  
Figure 10 - A-63160  
Figure 11 - A-63161

cc: Group 63 Staff

## APPENDIX

Calculations of Power Dissipation

Figure 10 shows the circuit and waveforms used for the calculation of power dissipation for a grounded emitter with a resistance load. A linear approximation is made and it is assumed that the output voltage has the form shown in part (b). Part (c) shows the assumed collector current waveform. Obviously, the average power for the first half is the same as for the second half of the pulse. For the first half of the pulse the collector voltage is,

$$e_c = V_c \left( \frac{t}{T} - 1 \right), \text{ for } 0 < t < T. \quad (12)$$

and the collector current is,

$$i_c = \frac{V_c}{R_c T} t; \text{ for } 0 < t < T. \quad (13)$$

The instantaneous power dissipated in the collector is,

$$P_c(t) = \frac{V_c^2}{R_c T} \left( \frac{t^2}{T} - t \right). \quad (14)$$

If the current gain of the transistor is  $\beta$  and the equivalent circuit of the base is that shown in Fig. 2, the power dissipated in the base is,

$$P_b(t) = \frac{V_c}{\beta R_c T} t e_b + \frac{V_c^2}{\beta^2 R_c^2 T^2} t^2 r_b \quad (15)$$

The total power dissipated in the transistor is approximately the sum of the collector and the base power. The average power over the first half of the pulse is,

$$P_{av} = \frac{1}{T} \int_0^T P_c(t) dt + \frac{1}{T} \int_0^T P_b(t) dt$$

$$P_{av} = \frac{1}{6} \frac{V_c^2}{R_c} + \frac{1}{2} \frac{V_c e_b}{\beta R_c} + \frac{1}{3} \frac{V_c^2 r_b}{\beta^2 R_c^2} .$$

If the duty cycle is limited to 50 percent, the maximum average power is,

$$P_{av} = \frac{1}{12} \frac{V_c^2}{R_c} + \frac{1}{4} \frac{V_c e_b}{\beta R_c} + \frac{1}{6} \frac{V_c^2 r_b}{\beta^2 R_c^2} \quad (16)$$

Expressing this in terms of the saturation current and supply voltage, this becomes,

$$P_{av} = \frac{1}{12} I_s V_c + \frac{1}{4} \frac{I_s e_b}{\beta} + \frac{1}{6} \frac{I_s^2}{\beta^2} r_b \quad (16a)$$

For a surface -barrier transistor, the maximum average allowable power is 9 milliwatts. If the average power described by equation (16a) is set equal to 9 milliwatts and  $V_c$  plotted versus  $I_s$  the curve shown in Fig. 11 results. This curve is plotted assuming that  $e_b = 0.3$  v,  $r_b = 125$  ohms, and  $\beta = 2$ . The choice of  $\beta = 2$  is made because the current gain of a transistor goes down as the current gets larger. Thus choosing  $\beta = 2$  is a bad condition that allows for large currents.

Any operating condition described by a point between the curve and the axes would be allowable from a power dissipation point of view. Due to the possibility of avalanche, the supply voltage is usually not allowed to be greater than -4.5 v. Thus operation at any point within the shaded region would be allowable within the approximations described.

The assumption of a linear approximation to the actual waveforms is not a very accurate one. However, since the waveforms have this general shape in actual practice, the calculations should be good enough to get better than an order of magnitude idea of what the power dissipation is.

At the present time, the effects on a surface-barrier transistor of a large inverse voltage on the collector are not known. The shaded area of Fig. 11 assumes that the voltage at the collector never gets larger than the avalanche voltage. When the load on the transistor is a transformer, the inverse voltage at the collector will be greater than the supply voltage by the amount of the overshoot at the transformer primary. The effect of this peak inverse voltage is not now known.



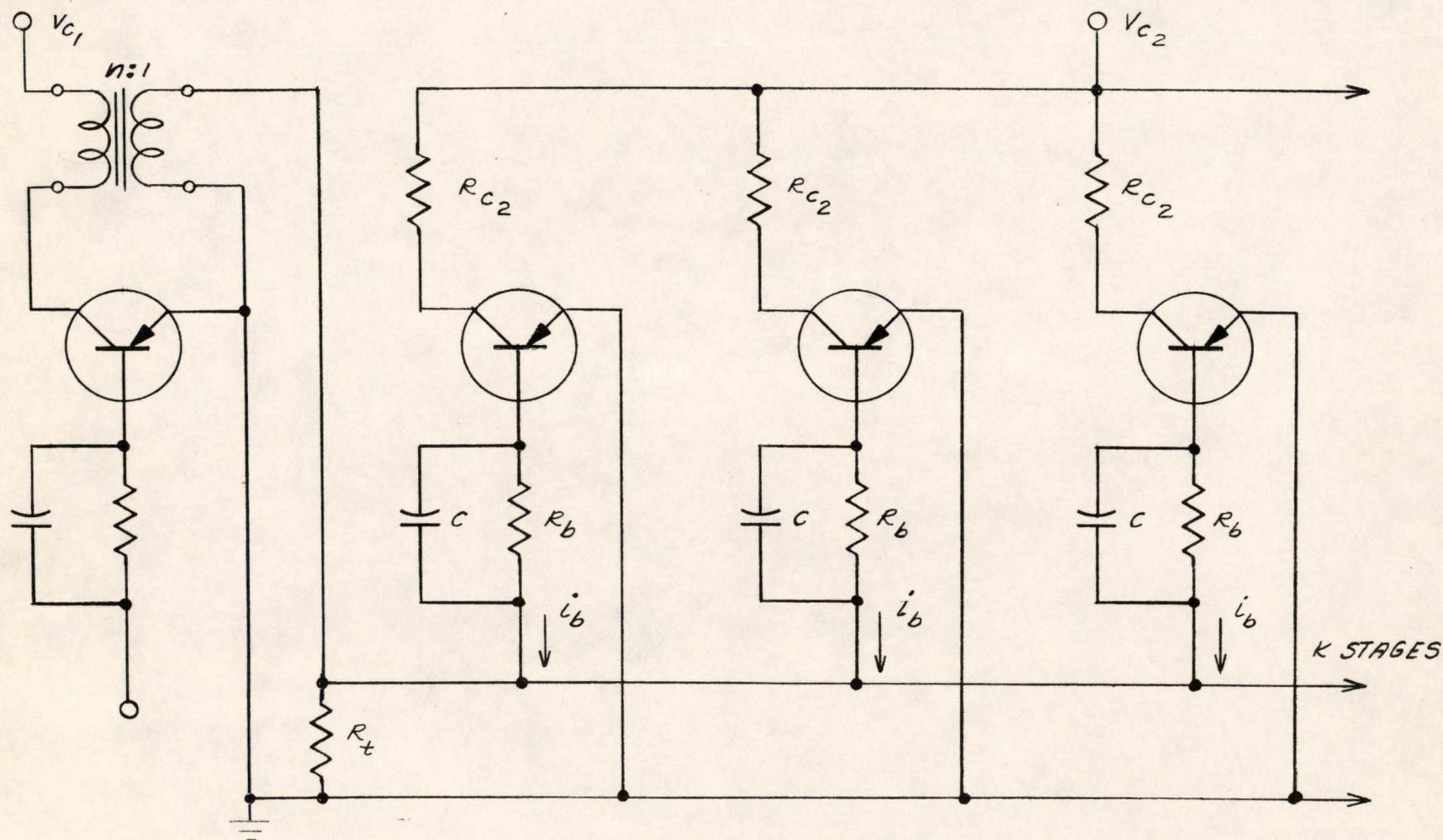


FIG. 3

CIRCUIT DIAGRAM OF TRANSFORMER - TRANSISTOR AMPLIFIER  
DRIVING K GROUNDED EMITTER STAGES

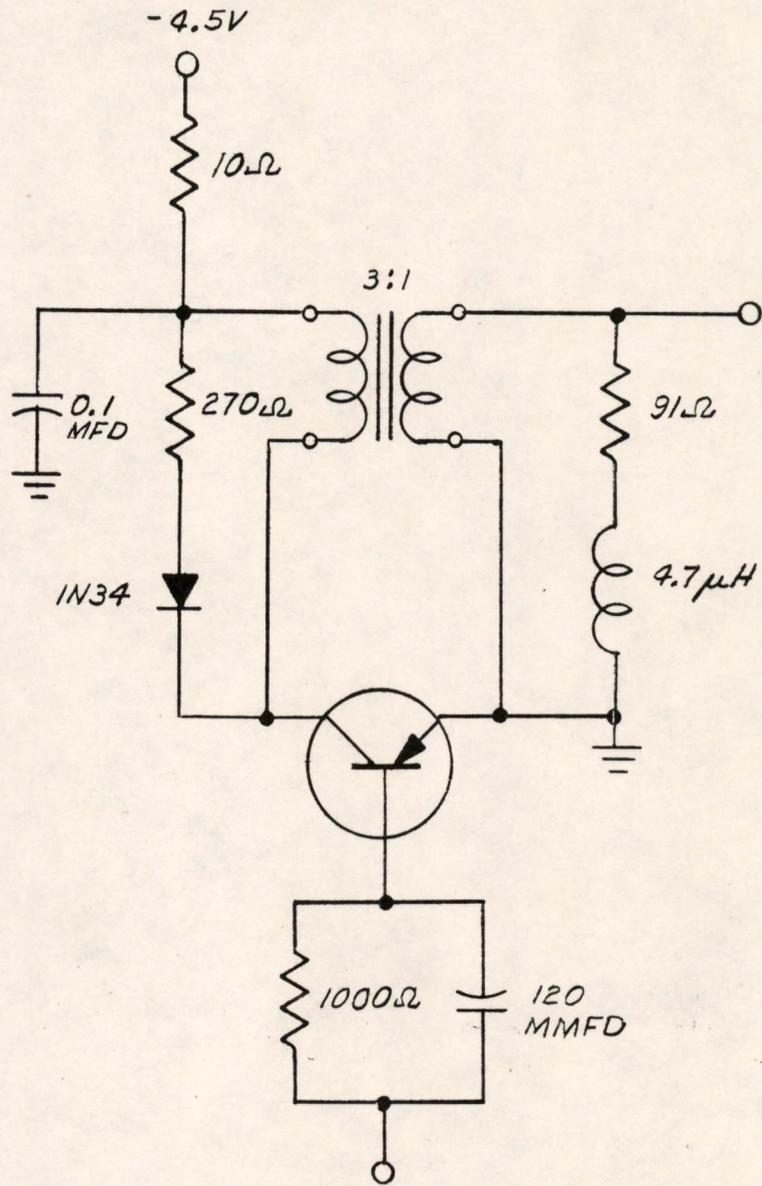


FIG. 4  
 CIRCUIT DIAGRAM OF  
 EXPERIMENTAL PULSE AMPLIFIER

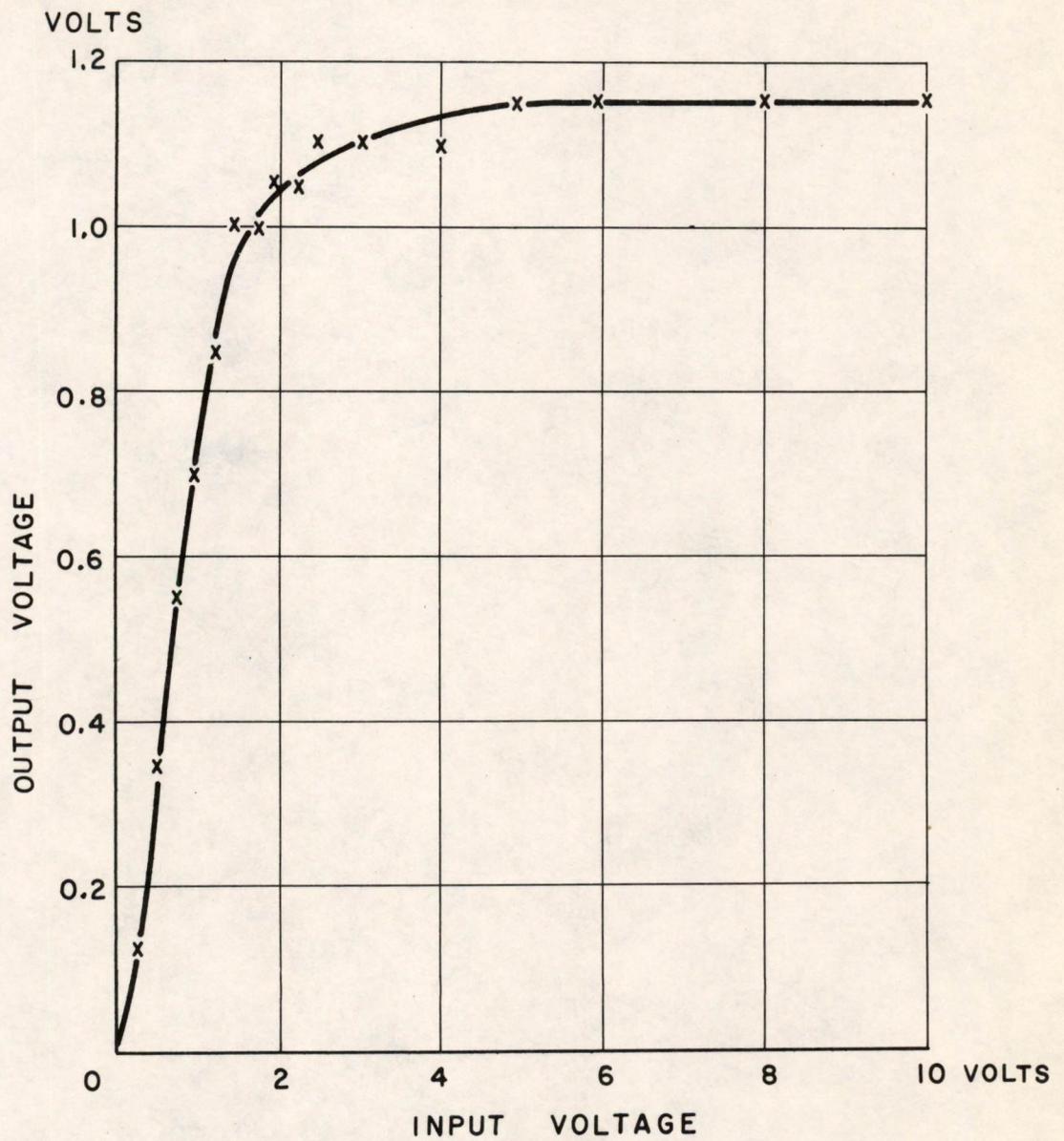


FIG. 5

OUTPUT MAGNITUDE vs INPUT MAGNITUDE FOR A 0.1  
MICROSECOND PULSE THROUGH THE TRANSFORMER  
TRANSISTOR PULSE AMPLIFIER

MICROSECONDS

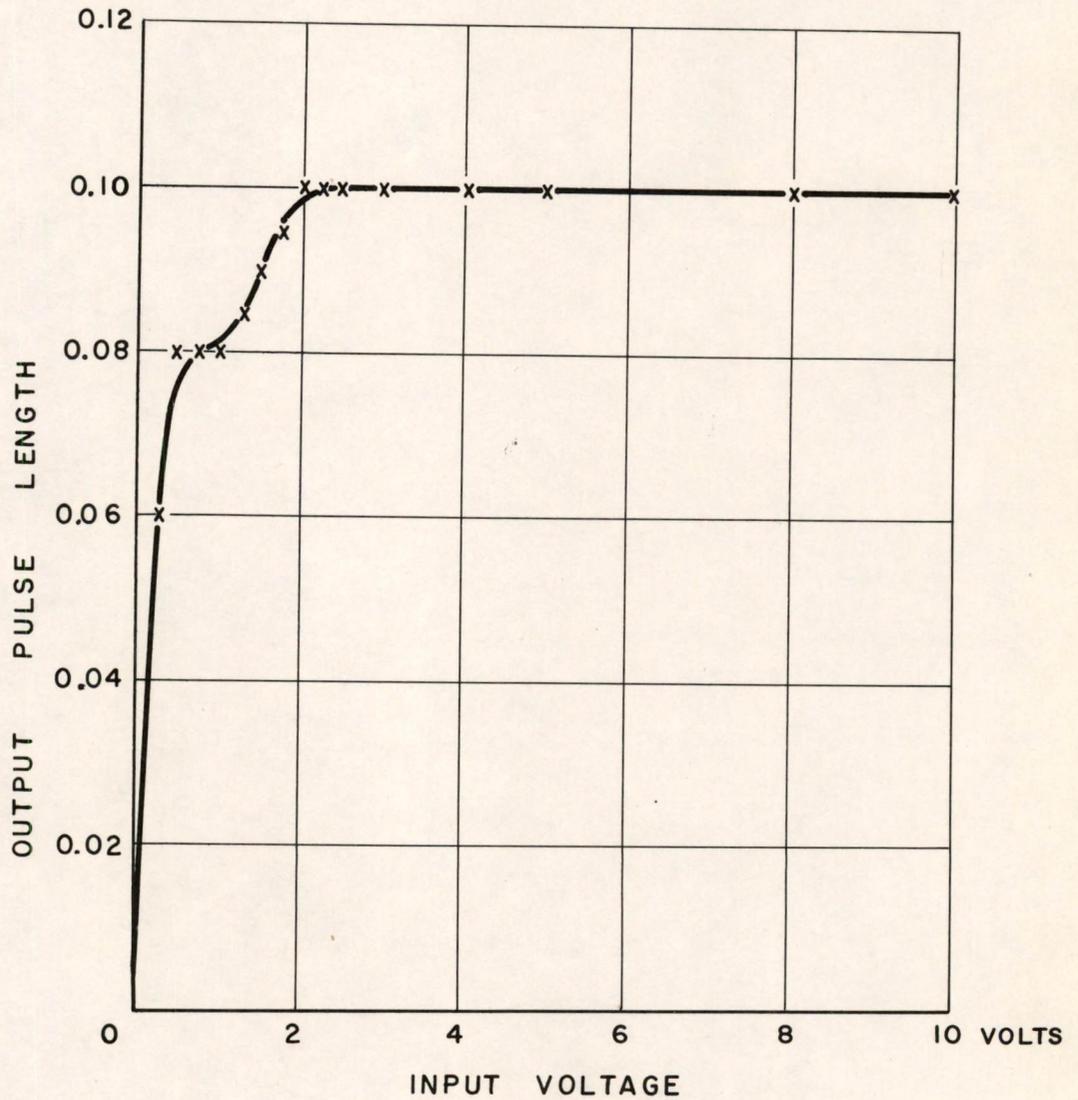
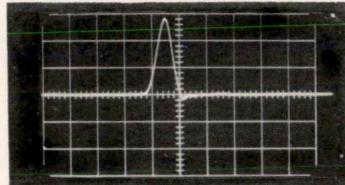


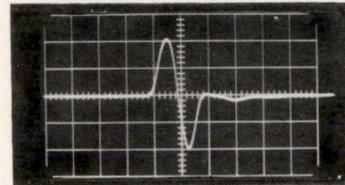
FIG. 6

OUTPUT PULSE LENGTH vs INPUT MAGNITUDE FOR A 0.1  
MICROSECOND PULSE THROUGH THE TRANSFORMER  
TRANSISTOR PULSE AMPLIFIER

NOTE:  
ALL PICTURES ARE  
INVERTED.

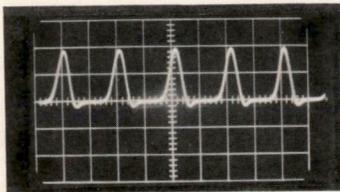


INPUT  
0.5 VOLTS/CM  
0.1  $\mu$  SEC/CM

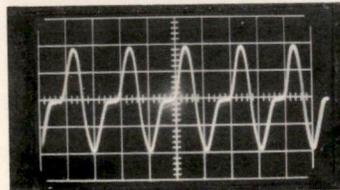


OUTPUT  
0.5 VOLTS/CM  
0.1  $\mu$  SEC/CM

FIG. 7  
INPUT AND OUTPUT FOR  
A SINGLE PULSE

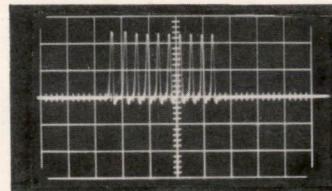


INPUT  
1 VOLT/CM  
0.1  $\mu$  SEC/CM

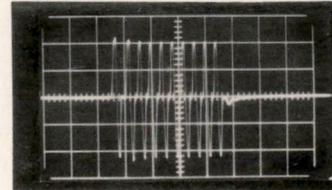


OUTPUT  
0.5 VOLTS/CM  
0.1  $\mu$  SEC/CM

FIG. 8  
INPUT AND OUTPUT AT  
5 MEGACYCLES

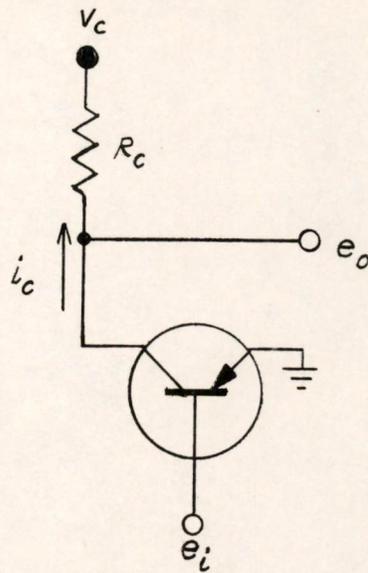


INPUT  
1 VOLT/CM  
0.5  $\mu$  SEC/CM

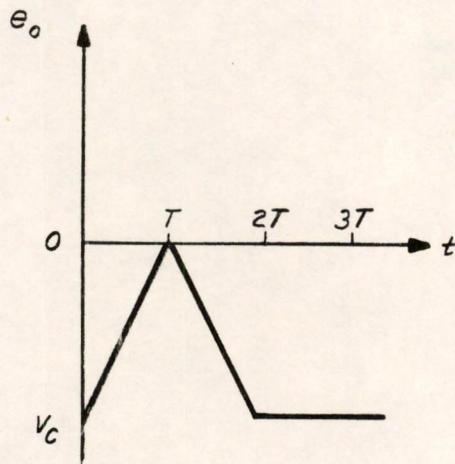


OUTPUT  
0.5 VOLTS/CM  
0.5  $\mu$  SEC/CM

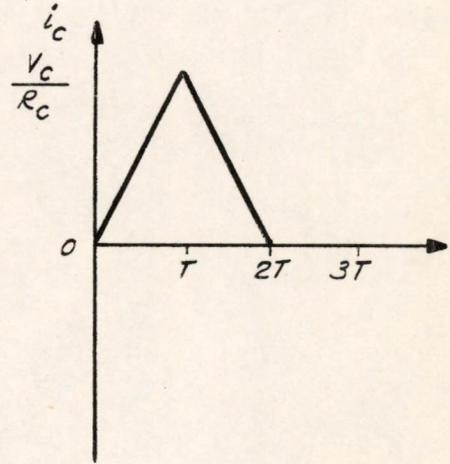
FIG. 9  
INPUT AND OUTPUT FOR  
A BURST OF PULSES



a) CIRCUIT DIAGRAM



b) OUTPUT VOLTAGE



c) COLLECTOR CURRENT

FIG. 10

CIRCUIT DIAGRAM AND WAVEFORMS USED TO  
 CALCULATE POWER DISSIPATION WHEN THE  
 LOAD IS RESISTIVE

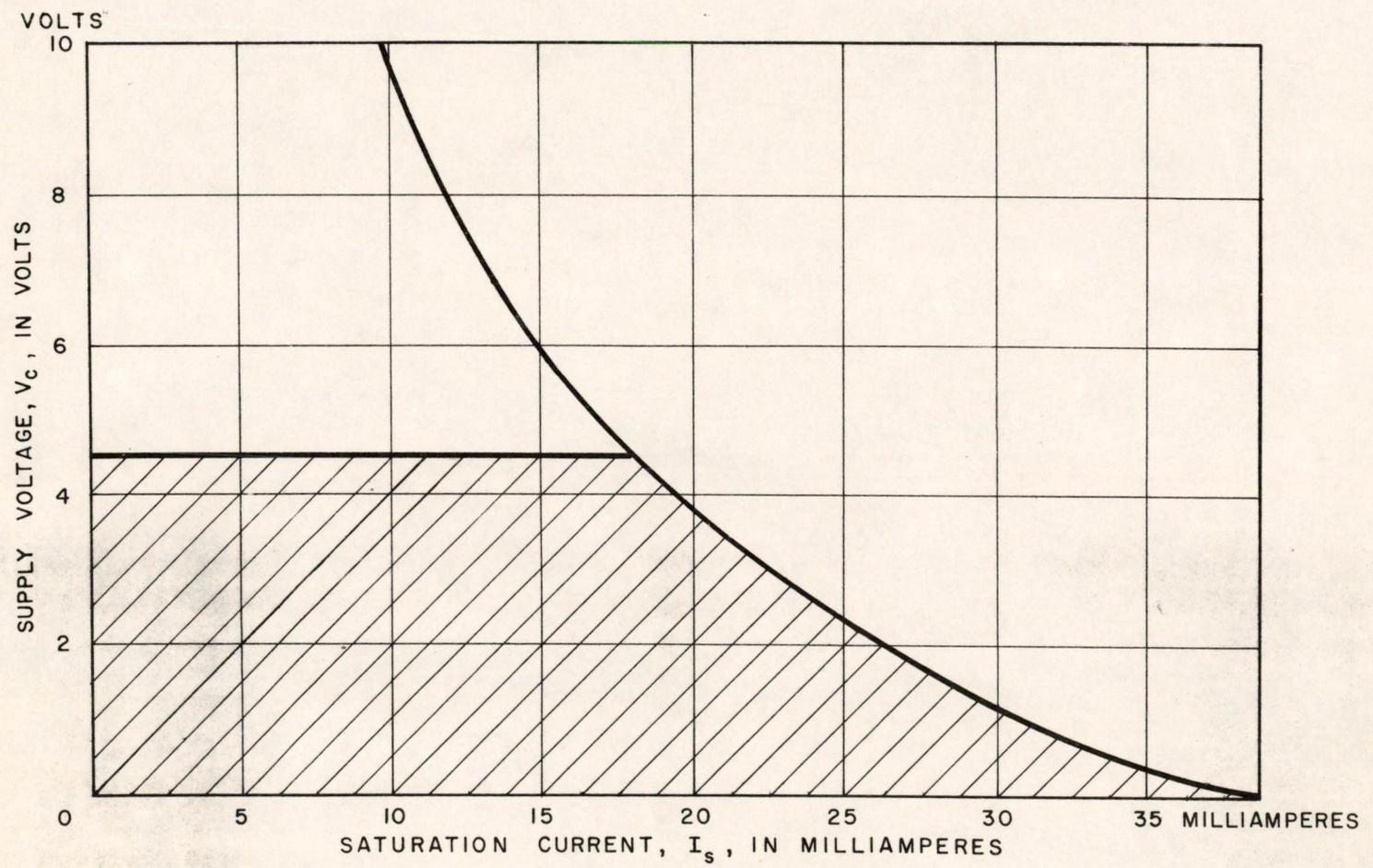


FIG. II

LIMITS OF PULSE OPERATION FOR A GROUNDED EMITTER TRANSISTOR WHEN THE AVERAGE POWER DISSIPATED IS LIMITED TO 9 MILLIWATTS

K. H. Olsen

3191

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: SBT Hole Storage - 1

To: Distribution List

From: Charles T. Kirk, Jr.

Date: September 27, 1955

Approved: Torben H. Meisling  
Torben H. Meisling

Abstract: This is the first of a series of reports covering an investigation of hole storage phenomena in SBT's. This report is concerned with storage time in SBT's employed as common-emitter switches under open base conditions. A circuit for measuring storage time is described in the first section of the report. The second section is devoted to a first-order, theoretical analysis of the hole storage mechanism. The analysis results in the following equation for storage time:

$$\text{Storage time} = t_1 = \frac{Q}{-I_c(\text{sat})} \left\{ \left[ \frac{1}{\beta(\text{sat})} - \frac{1}{\beta(\text{act})} \right] + \ln \left| \frac{\beta(\text{act})}{\beta(\text{sat})} \right| \right\}$$

where the terms of this equation are defined in the text. A comparison is made between this expression for storage time and storage times measured experimentally in the final section of this report.

Introduction

One of the most important limitations on the speed of saturating transistor-switching-circuits is due to minority-carrier storage (hereafter referred to simply as hole storage). The principal effect of hole storage is to introduce a delay in the turn-off time of a saturated transistor. During this delay time, the transistor behaves as though it is still on even though a turn-off signal has been applied. The most unfortunate part of this phenomenon is that the delay is dependent upon the circuit environment in which the transistor is placed, and the transistor parameters. Consequently, the delay due to hole storage varies from circuit to circuit, and from transistor to transistor.

The purpose of this investigation is to find out how circuit conditions and transistor parameters affect storage time. Once storage

This document is issued for internal distribution and use only by and for Lincoln Laboratory personnel. It should not be given or shown to any other individuals or groups without express authorization. It may not be reproduced in whole or in part without permission in writing from Lincoln Laboratory.

The research reported in this document was supported jointly by the Department of the Army, the Department of the Navy, and the Department of the Air Force under Air Force Contract No. AF 19(122)-458.

time can be related to the transistor parameters and its circuit environment, we can design the transistors and/or circuits to minimize storage time consistent with good switching performance.

A Circuit for Measurement of Storage Time

Figure 1 shows a simple switching-circuit in which storage time of a transistor can be observed under controlled circuit conditions. The circuit employs an SBT transistor in the common-emitter configuration as a simple switch. A mercury relay in series with the base current supply is used to turn the transistor on and off. The on circuit conditions can be measured by holding the relay closed and measuring the value of  $I_c(\text{sat})$  and  $I_b(\text{sat})$  by means of the milliammeters in the collector and base circuits.  $I_c(\text{sat})$  and  $I_b(\text{sat})$  can be controlled independently over a range of 0-10 ma by adjusting the resistances in the collector and base, respectively. Storage time is independent of the value of  $R_b$  and  $R_c$  since the base circuit is open and  $I_c = I_c(\text{sat})$  during this time.

Typical waveforms showing the behavior of the collector current,  $I_c$ , and the base current,  $I_b$ , upon opening the relay are shown in Figure 2. Storage time (designated as  $t_1$  in Figure 2) is defined as the time during which  $I_c = I_c(\text{sat})$  for  $I_b = 0$ .

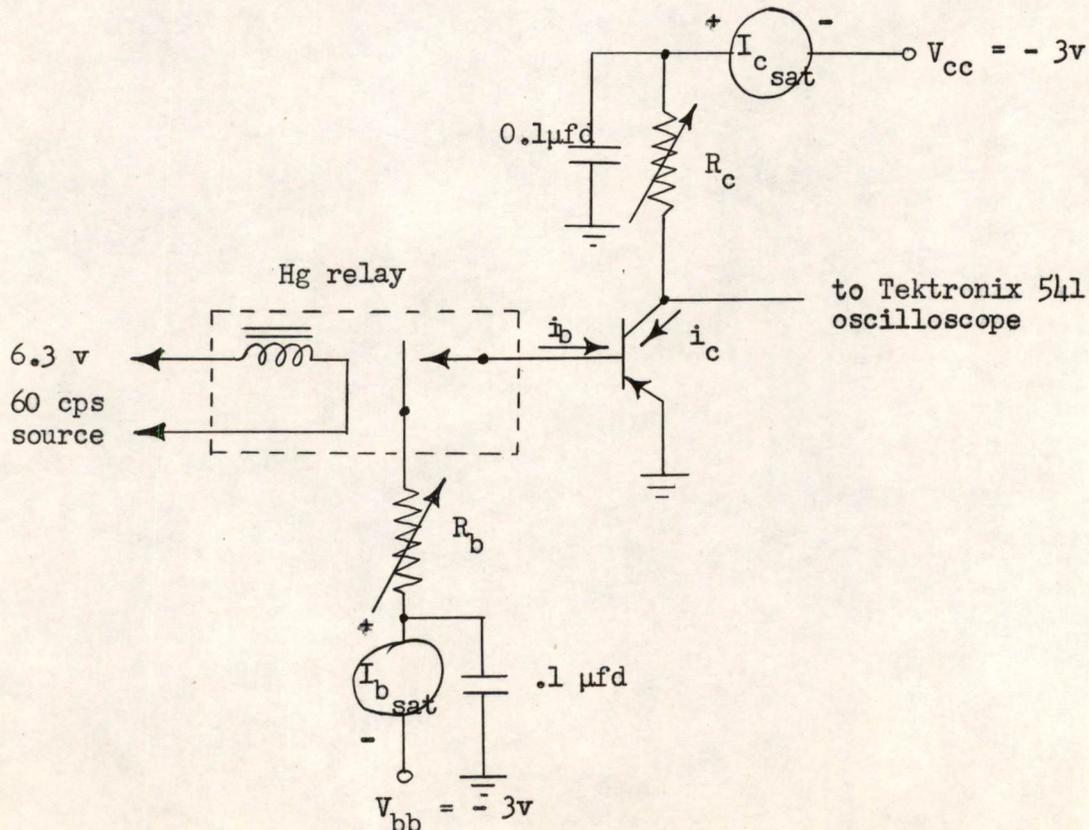


Figure 1. Circuit for Measuring Storage Time

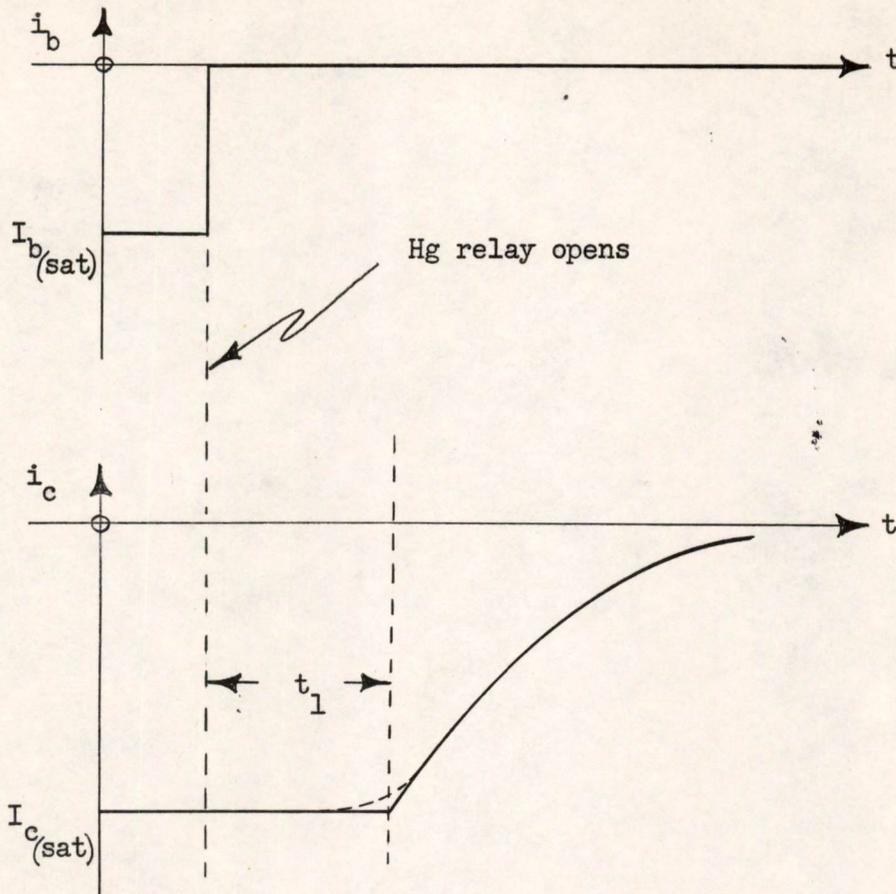


Figure 2. Typical Waveforms Obtained from Circuit of Figure 1  
Showing Definition of Storage Time

#### Analysis of the Transistor Hole Storage Mechanism

The proceeding analysis of the hole storage mechanism under the circuit conditions of Figure 1 is based on the idealized one-dimensional transistor model of Figure 3. The resultant agreement of the theoretical case and the actual case therefore will depend on how closely the one-dimensional model approximates the SBT. It was for this reason that the grounded-emitter configuration of the transistor was chosen to study the hole storage mechanism. In this configuration, the decay of holes stored in the base region of the transistor for  $I_b = 0$  is relatively slow. Thus, the volume effects of hole storage due to sideways diffusion are masked by the slow decay of holes in the central region of the base between the emitter and collector.

In Figure 3, the abscissa or x-dimension represents distance through the base region in going from the emitter to collector junctions. The ordinate or p-dimension represents the excess hole or minority carrier density at any point in the base region. Ordinarily,  $p$  is a

function of distance,  $x$ , and time. The excess hole densities at the boundaries of the emitter and collector junctions are defined to be  $p_e$  and  $p_c$ , respectively.

The following conditions are assumed to hold in the base region of the transistor model:

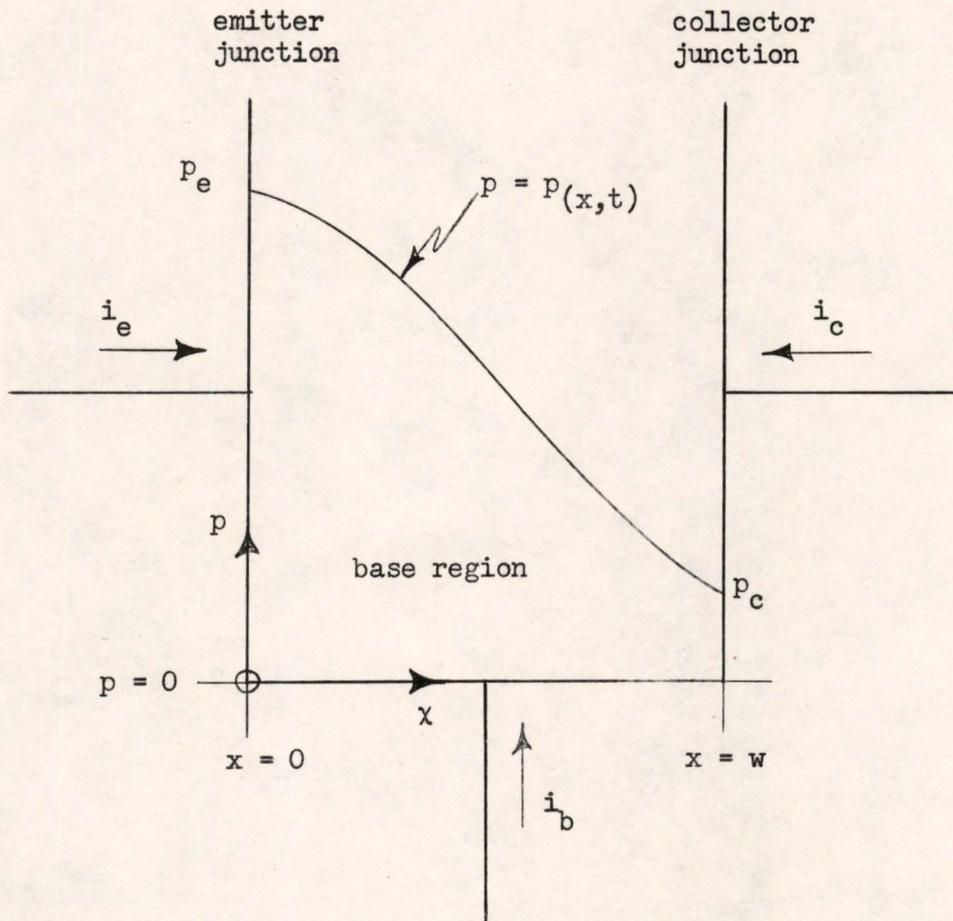


Figure 3

One-Dimension Model of SBT Base Region

- (a) The base region is field free.
- (b) The surface and volume recombination is negligible.

Under these conditions hole flow in the base will be by diffusion only, and the hole density distribution at any time  $t$  will be governed by a diffusion equation having the form

$$\frac{d^2 p}{dx^2} = \frac{1}{D} \frac{dp}{dt} \quad (1)$$

where  $D$  is the diffusion constant for holes. Furthermore we shall assume that  $dp/dt$ , the rate of decay of excess holes in the base region, is sufficiently small such that

$$\left. \begin{array}{l} \frac{dp}{dt} \\ \text{for all } x \\ \text{and all } t \end{array} \right\} \doteq 0 \quad (2)$$

Thus we can write that

$$\frac{d^2 p}{dx^2} \doteq 0 \quad \text{for all } t \quad (3)$$

Equation (3) states that the solution of  $p$  at any time  $t$  is a straight line which may be obtained from the proper boundary conditions at time,  $t$ , provided Equation (2) is true. Finally we shall make the additional assumption that the total collector current,  $I_c$ , is given by the hole current at the collector.

The assumptions we have made in the above discussion can be listed in an equivalent but more concise form as follows:

- (1) The hole storage phenomena of the SBT in the grounded emitter configuration can be represented by the one dimensional transistor model shown in Figure 3.
- (2)  $\alpha \doteq \gamma(p_e)$  where  $\gamma$  is the injection efficiency of the emitter.
- (3)  $\frac{d^2 p}{dx^2} \doteq 0$  for all  $t$
- (4)  $I_c \doteq I_c(p)$

In order to facilitate the analysis, it is assumed that the relay in the circuit of Figure 1 has been closed long enough, prior to  $t = 0$ , for the transistor to reach its equilibrium saturation state. The steady-state hole distribution in the base is given by the  $t = 0$  curve in Figure 4.  $P_e(\text{sat})$  and  $P_c(\text{sat})$  are the equilibrium hole densities for the saturation state at the emitter and collector junction boundaries of the base region, respectively.

At  $t = 0$  the relay opens. Initially, the excess hole density distribution remains as it was prior to the opening of the relay since all currents remain finite. The emitter hole current,  $I_e(p)$ , at any time represents the rate at which holes are injected into the base by the emitter junction. Similarly, the collector hole current,  $I_c(p)$  represents the rate at which holes are injected into the base by the collector junction. The time rate-of-change of the total excess hole charge in the base at any time,  $t$ , can then be written as

$$q \frac{dP_T}{dt} = I_e(p) + I_c(p) \tag{4}$$

where

$$P_T = \int_0^w P(x,t) dx \tag{4a}$$

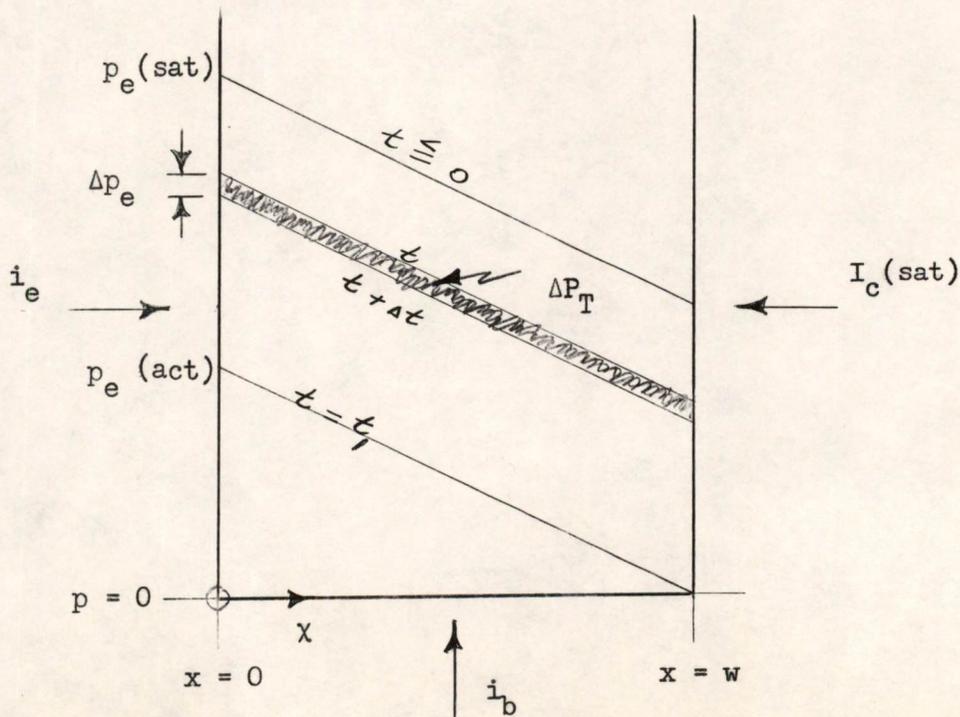


Figure 4  
Model of SBT Showing Idealized Behavior of Hole Density Distribution During Storage Time

Now Kirchoff's law states that for all  $t$

$$0 = I_e + I_b + I_c = I_e(p) + I_e(n) + I_b + I_c(p) \quad (5)$$

For  $t > 0$  equation (5) reduces to

$$0 = I_e + I_c = I_e(p) + I_e(n) + I_c(p) \quad (5')$$

since  $I_b = 0$  for  $t > 0$ .

It is obvious from equation (5') that

$$-I_c(p) \doteq -I_c > I_e(p)$$

and, since  $I_c < 0$ , according to the circuit in Figure 1,

$$\frac{dP_T}{dt} < 0$$

The total amount of excess holes in the base must then decrease with time for  $t > 0$ . This process can be described analytically by a differential equation which relates the excess hole density at the emitter,  $p_e$ , to time. Assumption (3) allows us to write that

$$I_e(p) = \alpha(p_e)I_e \quad (6)$$

Eliminating  $I_e$  from equations (6) and (5') and substituting the resulting expression for  $I_e(p)$  into equation (4) we find that

$$q \frac{dP_T}{dt} = I_c \left[ 1 - \alpha(p_e) \right] \quad (7)$$

As we discussed previously (see equation 3) the solution of  $p$  for any time,  $t$ , is linear in  $x$ . This solution is subject to the boundary condition that

$$I_c \Big|_{t_1 \gg t \gg 0} \equiv I_c(\text{sat}) = -qD_p \frac{dP}{dx} \Big|_{x=w} \quad (8)$$

$$\text{where } I_c \Big|_{t_1 \gg t \gg 0} = \text{a constant}$$

by definition.

Now since the solution is a straight line and the collector current is constant during the storage time

$$\left. \frac{dp}{dx} \right|_{x=w} = \left. \frac{dp}{dx} \right|_{\text{for all } x} = \text{a constant} \quad (9)$$

Therefore the solutions of  $p$  for any two different values of  $t$  in the range  $t_1 \geq t \geq 0$  will be parallel to each other. Two successive solutions of  $p$  are shown in Figure 4 for times  $t$  and  $t + \Delta t$  where  $t_1 > t \geq 0$ . The total change of  $P_T$  over the interval  $\Delta t$  can then be written as

$$\Delta P_T = w \Delta p_e \quad (10)$$

Dividing both sides of equation (10) by the interval  $\Delta t$  and taking the limit as  $\Delta t$  goes to zero, we have

$$\frac{dP_T}{dt} = w \frac{dp_e}{dt} \quad (10')$$

Now substituting the expression for  $dP_T/dt$  given in equation (10') for the left-hand side of equation (7) we have that

$$qw \frac{dp_e}{dt} = I_c(\text{sat}) \left[ 1 - \alpha(p_e) \right] \quad (11)$$

If  $\alpha$  (actually the injection efficiency  $\gamma$ ) can be expressed as a function of  $p_e$ , then equation (11) can be solved for  $p_e$  as a function of time by the method of separating variables.

Webster<sup>†</sup> has derived an approximate expression relating injection efficiency to  $p_e$  for a transistor operating in the active region. For the present we shall assume that this relation also holds in the saturation region as well. Thus we have from Webster that

$$\alpha(p_e) = \frac{1}{1 + \frac{\sigma_b w}{\sigma_e L_e} \left( 1 + \frac{p_e}{N_d} \right)} \quad (12)$$

<sup>†</sup> Webster, W.M., "On the Variation of Junction-Transistor Current Amplification Factor with Emitter Current," Proc. I.R.E., Vol. 42, pp. 914-920; June 1954.

Eliminating  $\alpha(p_e)$  from equations (11) and (12) and solving the resulting equation by the method of separating variables, we obtain that

$$q \frac{\sigma_e}{\sigma_b} L_e N_d \left[ \frac{\sigma_b w}{\sigma_e L_e} \frac{p_e}{N_d} + \ln \left| 1 + \frac{p_e}{N_d} \right| + C \right] = I_c(\text{sat})t \quad (13)$$

At time  $t = 0$ ,  $p_e = p_e(\text{sat})$  (see Figure 4) and the constant,  $C$ , must be given by the expression

$$- \left[ \frac{\sigma_b w}{\sigma_e L_e} \frac{p_e(\text{sat})}{N_d} + \ln \left| 1 + \frac{p_e(\text{sat})}{N_d} \right| \right] = C$$

Substituting this expression for  $C$  in equation (13), we find that the time required for the hole density at the emitter to decrease from  $p_e(\text{sat})$  to  $p_e$  is given by the equation

$$t = \frac{q \frac{\sigma_e}{\sigma_b} L_e N_d}{-I_c(\text{sat})} \left\{ \frac{\sigma_b w}{\sigma_e L_e} \left[ \frac{p_e(\text{sat})}{N_d} - \frac{p_e}{N_d} \right] + \ln \left| \frac{1 + \frac{p_e(\text{sat})}{N_d}}{1 + \frac{p_e}{N_d}} \right| \right\} \quad (14)$$

Figure 4 shows that the storage time,  $t_1$ , may be defined as the time required for  $p_e$  to decay from  $p_e(\text{sat})$  to  $p_e(\text{act})$ . Thus,  $t_1$  can be obtained from equation (14) by taking the value for  $p_e$  in equation (14) as  $p_e(\text{act})$ .

Equation (12) can be rearranged to have the form

$$\frac{\sigma_b w}{\sigma_e L_e} \left( 1 + \frac{p_e}{N_d} \right) = \frac{1}{\alpha(p_e)} - 1 = \frac{1}{\beta(p_e)} \quad (12')$$

If we now define that for  $p_e = p_e(\text{sat})$ ,

$\beta(p_e) = \beta(\text{sat})$  and for  $p_e = p_e(\text{act})$ ,  $\beta(p_e) = \beta(\text{act})$ , then, by making the appropriate substitutions of Equation (12') into equation (14) for the case where  $p_e = p_e(\text{act})$ , we can eliminate  $p_e(\text{sat})$  and  $p_e(\text{act})$  from the equation to obtain

$$t_1 = \frac{Q}{-I_c(\text{sat})} \left\{ \left[ \frac{1}{\beta(\text{sat})} - \frac{1}{\beta(\text{act})} \right] + \ln \left| \frac{\beta(\text{act})}{\beta(\text{sat})} \right| \right\} \quad (15)$$

where  $Q = q \frac{\sigma_e}{\sigma_b} L_e N_d$

The storage time is seen to be a function of the circuit conditions,  $I_c(\text{sat})$  and  $\beta_{(\text{sat})}$ , the transistor parameter,  $\beta_{(\text{act})}$  and the physical parameters of the transistor as given by Q.  $I_c(\text{sat})$  is the collector current for  $t < 0$ .  $\beta_{(\text{sat})}$  may be taken as the ratio of  $I_c(\text{sat})$  to  $I_b(\text{sat})$  for  $t < 0$ .  $\beta_{(\text{act})}$  is defined as the active common emitter current gain at  $I_c = I_c(\text{act})$  for  $V_c = 0$  †. Q is a constant (for a given transistor) which is related to the physical parameters of the transistor (see equation 15) and has the dimensions of charge.

Comparison of Experimental Data and Theory

The circuit of Figure 1 was used to obtain measurements of storage times for various circuit conditions. Several sets of storage times were taken for four different values of the saturation collector current,  $I_c(\text{sat})$ . Each set was obtained by fixing  $I_c(\text{sat})$  and recording the storage times for different values of the saturation base current,  $I_b(\text{sat})$ . Figure 5 shows a plot of storage time versus  $\beta_{(\text{act})}/\beta_{(\text{sat})}$  for a sample transistor with  $I_c(\text{sat})$  as a parameter. The experimental data is plotted as points on the graph while the curves represent storage time vs  $\beta_{(\text{act})}/\beta_{(\text{sat})}$  as obtained from equation (15). The values of  $\beta_{(\text{act})}$  used in obtaining the theoretical curves were obtained by measuring the  $\alpha$  of the transistor at zero collector-to-base voltage and  $I_c = I_c(\text{sat})$  for each of the four values of  $I_c(\text{sat})$  shown on the graph. The value for Q can be obtained by substituting the measured value of  $t_1$  into equation (15) for a particular setting of  $I_c(\text{sat})$  and  $I_b(\text{sat})$  and solving equation (15) numerically for Q.

The actual values of Q that are used in the theoretical expression were found by averaging the individual values of Q obtained from the experimental data for each value of the parameter  $I_c(\text{sat})$ . The average value of the Q's obtained are shown in Table I.

TABLE I

Average Values of Q for the Various Values of the Parameter  $I_c(\text{sat})$

SBT No. 1796

$I_c(\text{sat})$ ma.	$Q_{\text{avg}}$ $\mu\text{ coulombs}$
10	700
7.5	600
5	500
2.5	325

†  $\beta_{(\text{act})}$  is defined for  $I_c = I_c(\text{sat})$  and  $V_c = 0$  since these are the approximate circuit conditions that exist at the time when the transistor comes out of saturation and becomes active, i.e. at the time  $t = t_1$ .

SBT  
No. 1796

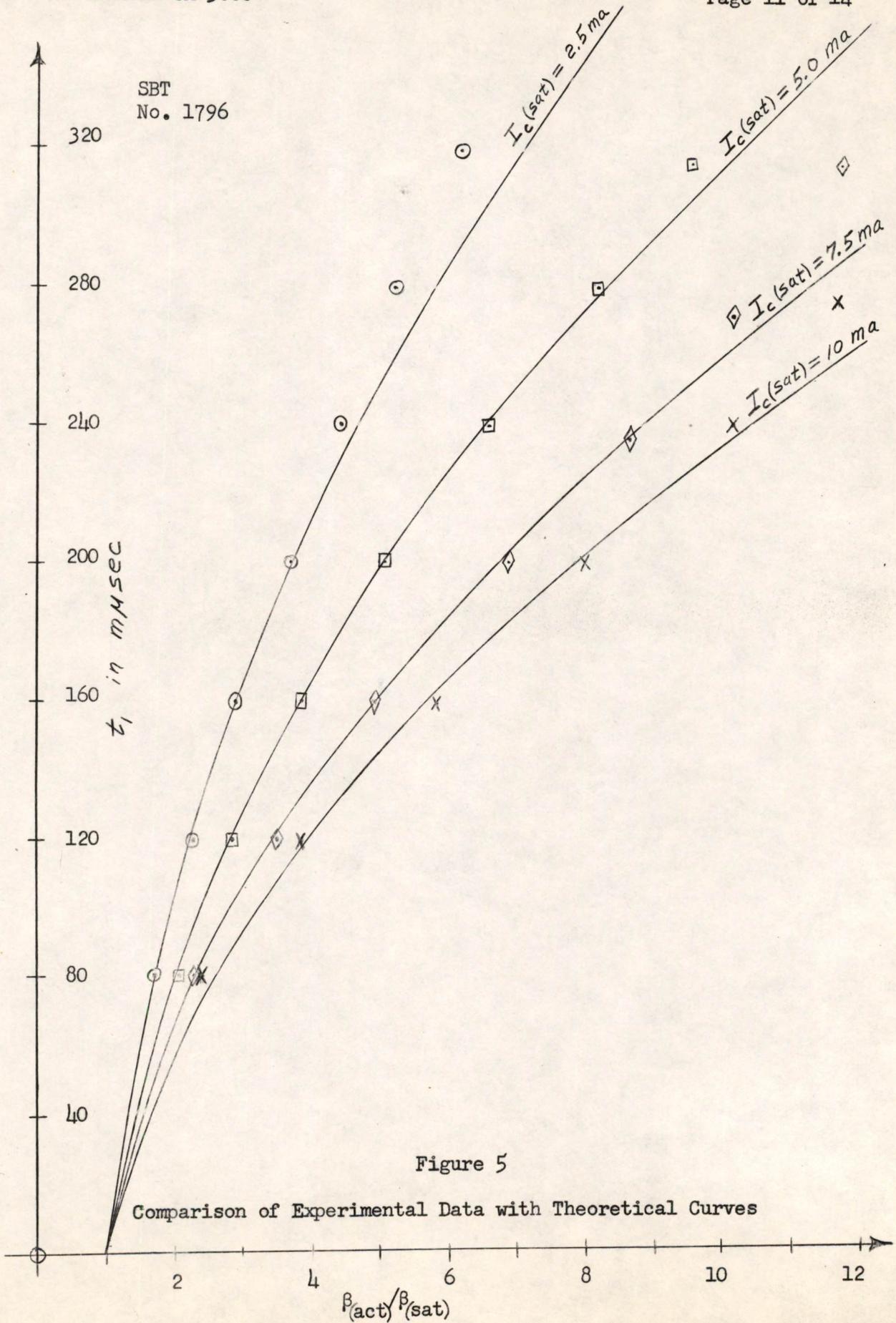


Figure 5

Comparison of Experimental Data with Theoretical Curves

According to the theory,  $Q$  is a constant for a given transistor. This fact, however, is at a variance with the calculated values of  $Q$  given in Table I. Another discrepancy between the theory and the experimental data can be seen in Figure 5 where for large values of  $\beta_{(act)}/\beta_{(sat)}$  (the range in which the transistor is heavily saturated) the measured values of storage time begin to deviate from the theoretical curves. These two difficulties probably arise because of two limiting assumptions we made in deriving the theoretical expression for storage time. These are that (1) the SBT is assumed to behave like the one-dimensional model of Figure 4 and (2) the common base  $\alpha$  of the transistor is due entirely to the emitter injection efficiency.

It has been observed experimentally that some sideways diffusion takes place in a saturated SBT<sup>†</sup> which results in storage of holes in the outer regions of the base (that is outside the base region between the emitter and collector). Consequently, when the transistor is heavily saturated our one-dimensional model no longer holds. One effect of storing holes in the outer extremities of the base region would be to lengthen storage time since these holes would then have to diffuse back from the outer regions of the base to the collector. Such a situation would account for the discrepancy that occurs in Figure 5 in which the measured storage times for large values of  $\beta_{(act)}/\beta_{(sat)}$  are longer than the theory predicts.

The more fundamental difficulty with the theoretical results, that of  $Q$  varying proportionately with  $I_c(sat)$  cannot be explained on the basis of the sideways diffusion phenomena. In fact according to the sideways diffusion effect,  $Q$  should vary inversely as  $I_c(sat)$ . Some insight into the problem of  $Q$  varying with  $I_c(sat)$  can be gained by reconsidering the idealized hole storage mechanism, developed in the previous section, in the light of practical experience. According to our model of the SBT, the rate of change of stored holes in the base is given by the algebraic sum of collector hole current and the emitter hole current. For the common emitter open base situation, this rate is given by equation (7) as

$$\frac{d P_T}{dt} = I_c(sat) [1 - \alpha]$$

Thus, according to equation (7), as

$$I_c(sat) \rightarrow 0$$

$$t_1(\text{storage time}) \rightarrow 0$$

<sup>†</sup> Sideways diffusion is the diffusion of holes in a direction perpendicular to the emitter-collector axis. This effect has been observed experimentally, and will be described in more detail in a subsequent report.

since,  $dP_T/dt$ , the rate of change of stored holes will approach zero with  $I_C(\text{sat})$ . It can be shown experimentally, however, that  $t_1$  approaches a finite constant as  $I_C(\text{sat})$  approaches zero. The only other mechanisms by which we can account for the reduction of holes in the base region when  $I_C(\text{sat}) = 0$  are volume and surface recombination. The most important of the two is probably surface recombination since the volume lifetime of a hole in the base is in excess of 100  $\mu\text{sec}$ . A more accurate expression for the total rate of change of stored holes in the base then requires that equation (7) be modified by an additional term that takes the removal of stored holes by surface recombination into account. Thus,

$$\frac{dP_T}{dt} = I_C(\text{sat}) (1 - \alpha) - I_{SR}$$

where  $I_{SR}$  is the surface recombination hole current.

Webster<sup>†</sup> has shown that for high emitter currents (in the order of 10 ma) the loss of holes due to surface recombination becomes quite small compared to the loss due to injection efficiency.<sup>††</sup> Consequently, the expression for storage time derived in the previous section should be approximately correct for values of  $I_C(\text{sat}) \geq 10$  ma. The correct value of  $Q$  in the case of our particular SBT is then approximately 700<sup>†††</sup> as obtained from the experimental data taken for  $I_C(\text{sat}) = 10$  ma. If this value of  $Q$  is now used to obtain the storage time curve for  $I_C(\text{sat}) = 2.5$  ma., then the values of storage time predicted by the theory will exceed the measured values of storage time. The necessary reduction in  $Q$  needed to get the theoretical curve to fit the data is thus a correction for the error introduced by neglecting surface recombination.

### Conclusion

A relatively simple one-dimension transistor model can be used to describe the hole storage mechanism in the SBT. The expression for storage time derived from this model agrees reasonably well with measured values of storage time.

A comparison of the theoretical expression of storage time and the experimental data revealed the following difficulties with the present transistor model:

---

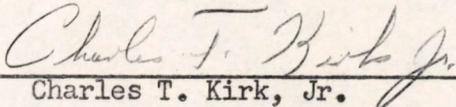
<sup>†</sup> Webster, W.M., Op. Cit.

<sup>††</sup> This is shown by Webster for the active region of operation; however, this situation should still hold approximately true for the saturation region where  $\beta_{\text{act}}/\beta_{\text{sat}}$  is small.

<sup>†††</sup> An extrapolated plot of  $Q$  vs  $I_C(\text{sat})$  for the values given in Table I shows that  $Q$  approaches a limiting value of 800  $\mu\text{coulombs}$  for  $I_C(\text{sat}) \geq 15$  ma.

- (1) At high saturation levels ( $\beta_{act}/\beta_{sat} > 10$ ), the measured values of storage time were found to be larger than the values predicted by the theory.
- (2) For low value of collector currents,  $I_c(sat) \sim 1.0$  ma, the measured values of storage time were found to be lower than the values predicted by the theory.

The first of these two difficulties is believed to be caused by the sideway diffusion of holes into the outer extremities of the base. The second difficulty is thought to be caused by the loss of holes through surface recombination.

  
Charles T. Kirk, Jr.

CTK/md

Distribution:

Group 63 Staff

B. Lax, Group 35  
D. T. Stevenson, Group 35  
R. H. Rediker, Group 35  
R. H. Kingston, Group 35  
R. H. Baker, Group 24

K.H. Olsen  
B-191

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: A NOVEL METHOD FOR MEASURING HOLE-STORAGE CHARGE IN DIODES  
To: See Distribution List  
From: Edmund U. Cohler  
Date: October 25, 1955  
Approved: Torben H. Meisling  
Torben H. Meisling

Abstract: Hole-storage charge in most diodes can be measured by a new method. A constant current is applied to a condenser through the diode, and the condenser is then allowed to discharge through the diode in the reverse direction. The reverse current through the diode (while the condenser is discharging) occurs in two stages. In the first stage the current is quite high, and the stored holes are discharged. In the second stage the diode reverse current discharges the condenser at a very slow rate. The change in charge on the condenser during the first stage provides a measure of the hole-storage charge. This method is described in more detail in the following paper.

Introduction:

Reverse-recovery of crystal diodes has long been known to be a problem in magnetic-core shift registers of the single-core-per-bit type. In such circuits the core acts as a current source that charges either a delay line or a capacitor through a crystal diode. The diode prevents the stored charge from leaking off of the capacitor after the core completes its switching. An equivalent circuit for this type of operation is shown in Fig. 1.

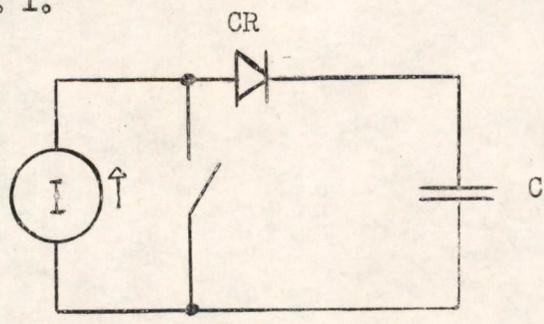


Figure 1

This document is issued for internal distribution and use only by and for Lincoln Laboratory personnel. It should not be given or shown to any other individuals or groups without express authorization. It may not be reproduced in whole or in part without permission in writing from Lincoln Laboratory.

The research reported in this document was supported jointly by the Department of the Army, the Department of the Navy, and the Department of the Air Force under Air Force Contract No. AF 19(122)-458.

In early shift registers attempts were made to use alloy junction diodes. These diodes were found to have such poor recovery characteristics that nearly half of the charge on the capacitor was lost through reverse-recovery current in the diode. This difficulty suggested the recovery test to be described.

### Test Method

In order to determine the value of a diode, for applications in which the voltage across the unit was reversed immediately subsequent to passage of forward current, it was decided to test the diode in a standard circuit which would put it under roughly equivalent conditions. The circuit of Fig. 2 was devised to allow ease of measurement and maximum flexibility. This device applies a constant current through the diode to a capacitor of any desired capacitance, for any desired amount of time. The current generator is then shorted out and the capacitor allowed to discharge through the back resistance of the diode. The amount of charge lost from the condenser can be measured by simply noting the change in voltage across the condenser. Moreover, the observation does not require fast response oscilloscopes as in many other tests.

### Theoretical Basis

The amount of "hole storage" in a forward biased diode is a function of the amount of time in saturation and the magnitude of the saturation current.<sup>1</sup> By supplying a variable current for a variable amount of time this device allows one to make tests for a large range of saturation conditions. Moreover, a correct choice of time and capacity can provide simulation of many actual situations.

When the diode is saturated, the distribution of holes in the base is roughly as shown in Fig. 3a. Upon reversal of the voltage across the diode, the current becomes limited by the end resistance of the diode only. (Fig. 3b.) This high current continues for a very short time. The current then decreases rapidly, and is a function of the initial forward current only.<sup>2</sup> In a very short time the current becomes roughly equal to the normal reverse current of the particular diode (Fig. 3c.) When this condition prevails, the charge that was stored in the diode is just about completely dissipated. This means that the change in the voltage on the capacitor can be used to measure the total charge that was stored in the diode. This discharge may take place within a very short span of time, say 0.2 microseconds or less. However, the response of the oscilloscope need not be that fast, because one needs to measure only the eventual charge on the capacitor which changes very slowly.

1. Jones, Nolan T., "Minority Carrier Storage in Diodes and Transistors", Thesis, August, 1954.
2. Kingston, Robert H., "Switching Time in Junction Diodes and Junction Transistors", Proc. I.R.E. Vol. 42, No. 5, May, 1954.

$p$  = hole density  
 $x$  = distance from junction

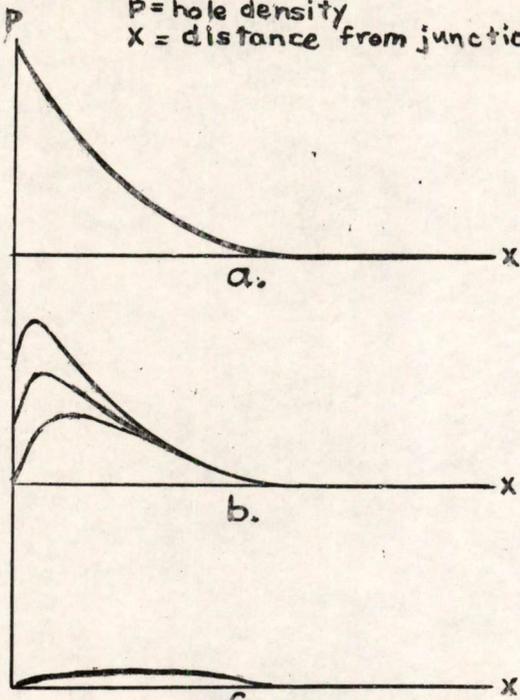


Figure 3

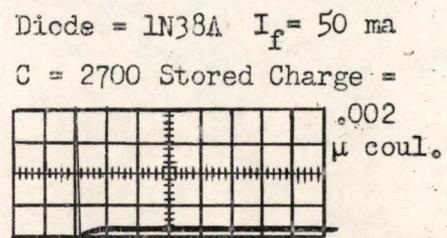
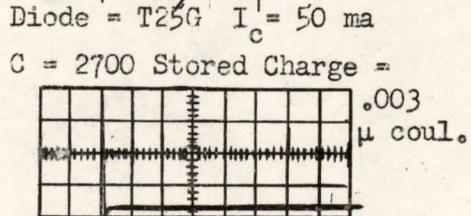
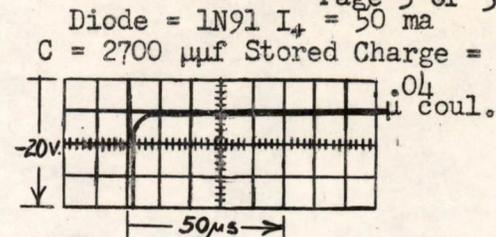


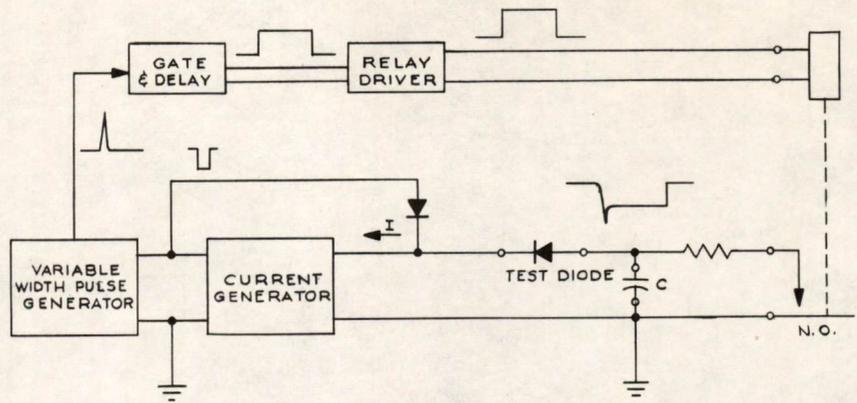
Figure 4

It might be noted that the lifetime of holes in the material has not been mentioned. Usually, the lifetime of holes in the diode is long compared to the time required to discharge the diode in this manner. However, certain types of diodes may achieve very good recovery characteristics by dint of an artificially reduced lifetime. In such diodes, the results of this test would be relatively meaningless as a measure of the total stored charge, but would still reveal the operation to be expected under certain circuit conditions. Moreover, the lifetime must be taken into account when any diode is used in a circuit which causes the diode to recover very slowly.

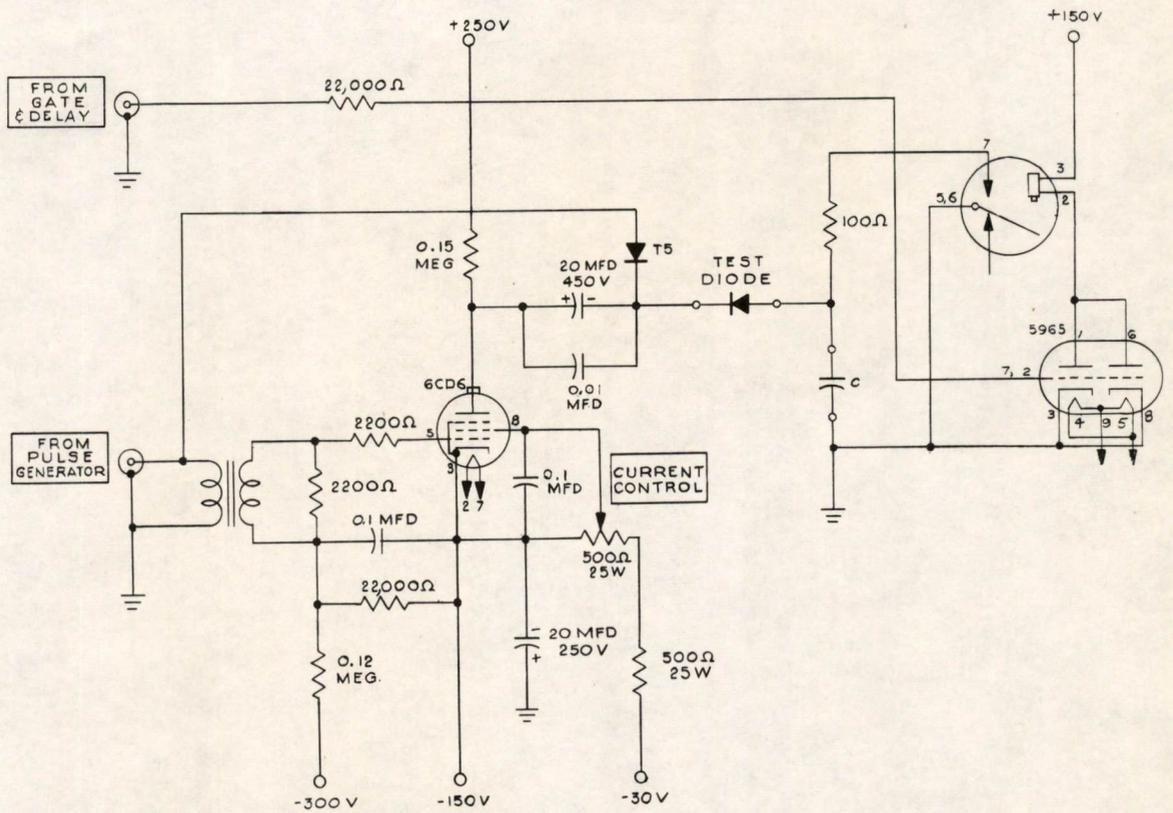
Circuit Operation

The circuit, as shown in Fig. 2, is quite conventional in operation. The variable width pulse generator supplies a voltage pulse which serves the dual purpose of turning on the current generator (6CD6) and turning off the switching diode (T5). When the voltage of the pulse generator returns to ground (from some negative value) the condenser, C, has been charged to some negative voltage and now discharges through the test diode, the switching diode, and the internal impedance of the pulse generator. The relay and relay driver (5965) are operated some time after the charging and discharging of the condenser through the test diode, in order to bring the initial charge on the condenser back to zero before the next test cycle. The pulse generator is a Rutherford. Typical waveforms are shown in Fig. 4 and show the relative recovery characteristics of various types of diodes. Finally, it might be noted that a certain amount of the recovery is due to capacity, and this cannot be differentiated from hole storage by the above method.

Signed Edmund U. Cohler  
 Edmund U. Cohler



BLOCK DIAGRAM



CIRCUIT FOR DETERMINING REVERSE RECOVERY CHARGE OF DIODES

C-59295

ELECTRICAL ENGINEERING DEPARTMENT  
MASTER'S THESIS PROPOSAL

TITLE: THE APPLICATION OF TRANSISTORS TO MULTIPOSITION SELECTION SWITCHES

BRIEF STATEMENT OF THE PROBLEM:

A transistor multiposition selection switch is to be designed, constructed, and analyzed which will receive a coded binary input from  $n$ -pairs of terminals controlled by flip-flops, and which has a selected output on one of  $2^n$  output lines. The selected output may in turn control the equipment necessary to obtain random access to a magnetic memory<sup>1, 2, 3</sup> With the coming of larger memories (e.g. 256 x 256 x 37) and transistorized computers, the present diode matrix switches<sup>4, 5</sup> are not practical. A possible solution to this problem is a dynamic switch designed around surface barrier transistors to provide operation at faster rise and fall times and with lower power requirements.

HISTORY OF THE PROBLEM:

The first multiposition switch where  $n$ -pairs of inputs selects one of  $2^n$  outputs was conceived independently by Jan Rajchman and Perry Crawford during 1941. This first switch<sup>6, 7</sup> has been generalized to many different forms as will be shown in the following discussion.

Resistance Matrix Switch

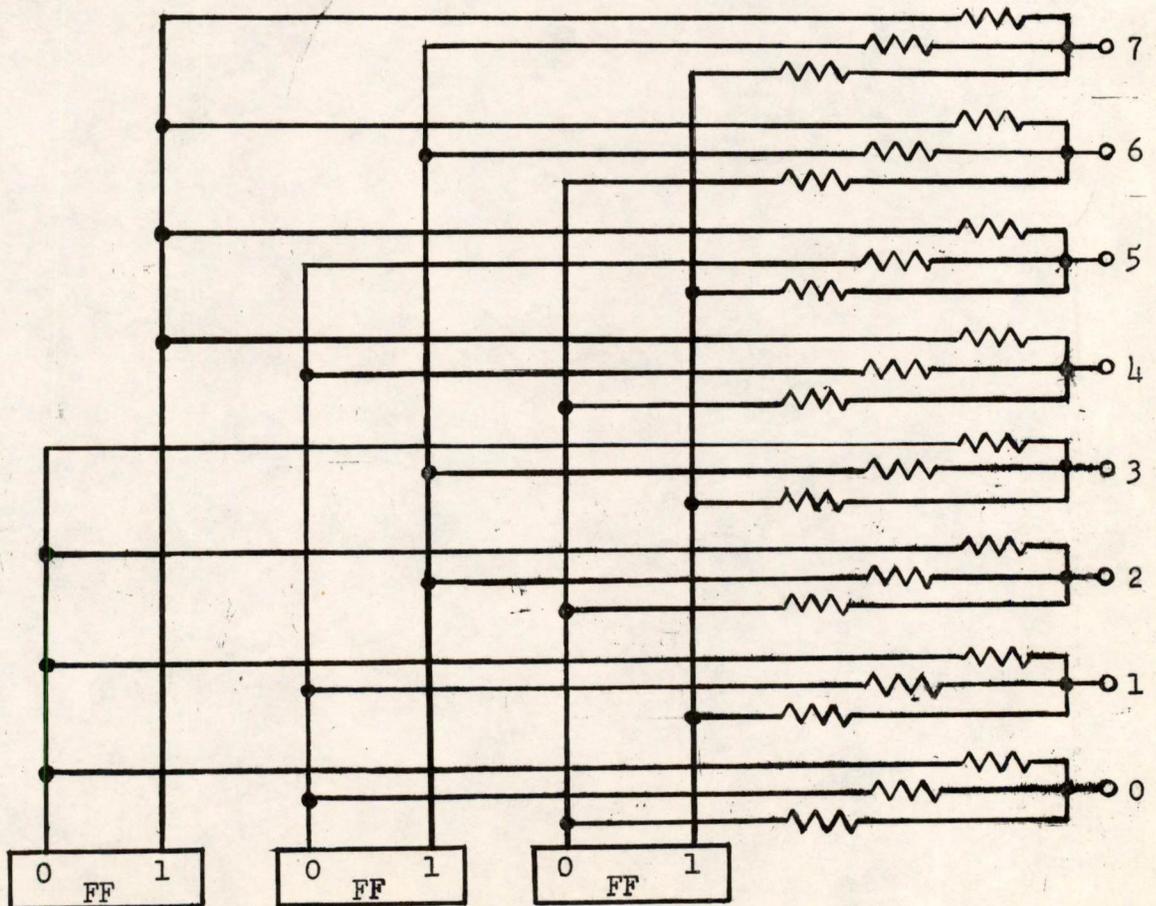
A typical resistance matrix switch is shown in Figure 1. This switch has three input pairs and eight output lines (the number of things taken three at a time is eight and in general  $n$ -pairs of inputs have  $2^n$  outputs).

---

1. Superscripts refer to the bibliography at the end of the proposal.  
This document is issued for internal distribution and use only by and for Lincoln Laboratory personnel. It should not be given or shown to any other individuals or groups without express authorization. It may not be reproduced in whole or in part without permission in writing from Lincoln Laboratory.

The research reported in this document was supported jointly by the Department of the Army, the Department of the Navy, and the Department of the Air Force under Air Force Contract No. AF 19(122)-458.

The operation can be explained as follows: Assume that the voltage at each input pair is either zero or E volts and the selected output terminal is number 7. Then the selection switches must be set to 111 and the selected terminal will be at E volts. Terminals 3, 5, and 6, will be at  $2E/3$  volts; terminals 1, 2, and 4, will be at  $E/3$  volts; and terminal 0 will be at zero volts. In general, there are  $(n+1)$  possible output voltages,  $E, \frac{n-1}{n}E, \frac{n-2}{n}E, \dots, 0$ . The distribution of output terminals among these voltages is given by the  $(n+1)$  row of the Pascal triangle. The output voltage, which is at E volts, must have some type of discriminator to reject the other voltages on the nonselected terminals. The effective switching voltage,  $E - \left(\frac{n-1}{n}\right)E$ , equals  $E/n$ . As n becomes large, it is apparent that selection becomes more difficult and thus the practical size of switches of this type is limited.



RESISTANCE MATRIX SWITCH

FIGURE 1.

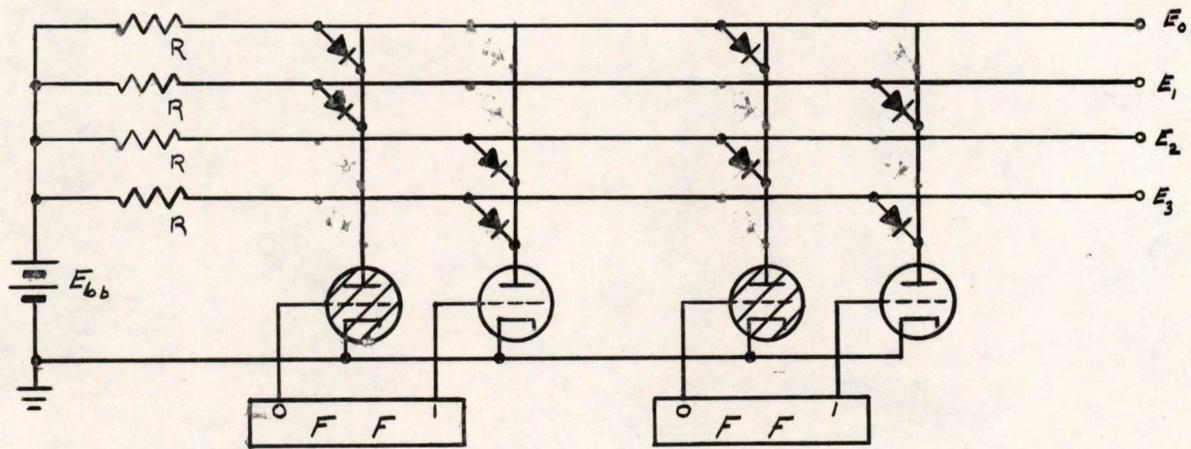
Diode Matrix Switch

By replacing the resistors in the matrix of Figure 1 with crystal diodes, as shown in Figure 2a, a circuit is obtained which can have a larger effective switching voltage. Plate current through the buffer amplifiers must flow through the resistors, R. The crystal matrix is connected so that current always flows through all the resistors but one. The terminal associated with that resistor, being at the higher voltage than the other terminals, is the selected terminal. In order to make the explanation easily understood, assume that the crystal-rectifier has a forward resistance,  $R_f$ , and a back resistance,  $R_b$ , which are independent of the magnitude of the current through the rectifier. To determine whether a crystal rectifier may be represented by  $R_f$  or  $R_b$ , the polarity of the voltage across the rectifier can be established by inspection of the circuit. An "on" buffer amplifier may be represented by the static plate resistance of the tube,  $r_p$ , and an "off" buffer amplifier by an infinite resistance. The equivalent circuit can be represented as shown in Figure 2b; and if  $R_f$  is assumed to be much smaller than R, the equivalent circuit can be simplified to that of Figure 2c (this is a very good approximation provided that R is at least ten times greater than  $R_f$ ). The general equivalent circuit for the above assumptions is shown in Figure 2d. The effective switching voltage from this equivalent circuit becomes

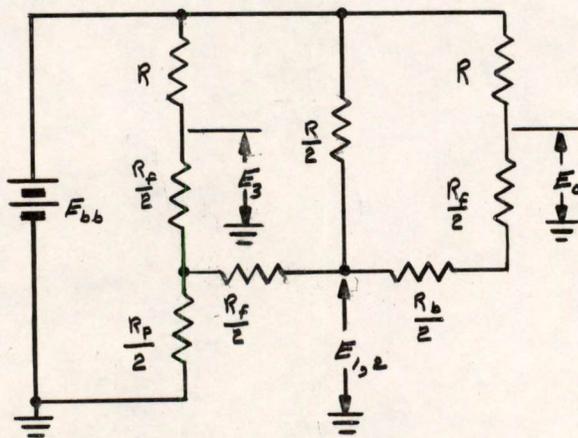
$$\Delta E = \frac{R R_b E_{bb}}{(2^{n-1}) (2^{n-1-1})} \left[ R_p \left( R + \frac{R}{2^{n-1}} + \frac{R_b}{(2^{n-1-1})^n} \right) + \frac{R}{2^{n-1}} \left( \frac{R_b}{2^{n-1-1}} + R_n \right) \right]$$

The simplified equivalent circuit shows that  $\Delta E$  will decrease as n increases or  $R_b$  decreases; however, this type of switch allows one to have

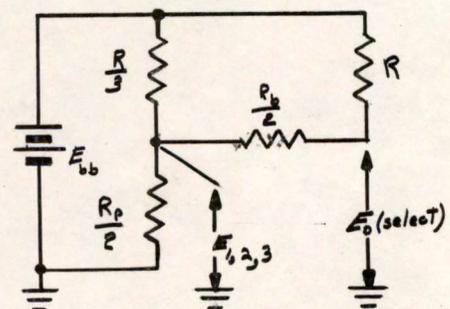
much greater switching voltages compared to a resistance switch of the same size. 64-position switches of this type are presently being used in the MTC digital computer at Lincoln Laboratory. Switches of this magnitude, however, require large amounts of power to drive them, i.e., transmitting tubes, in order that the voltage at the output terminals will rise rapidly (e.g. one-half microsecond). Thus, in larger sizes this type of switch becomes uneconomical, not due to the cost of the diodes, which are relatively inexpensive, but due to the required complexity of driving equipment.



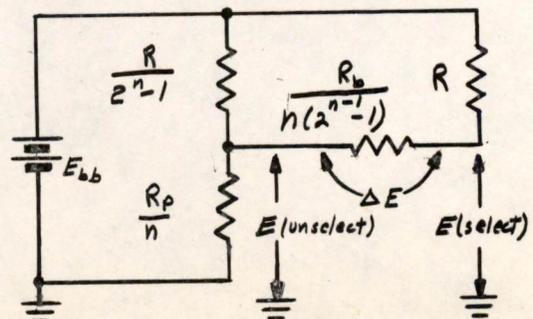
(a)



(b)



(c)

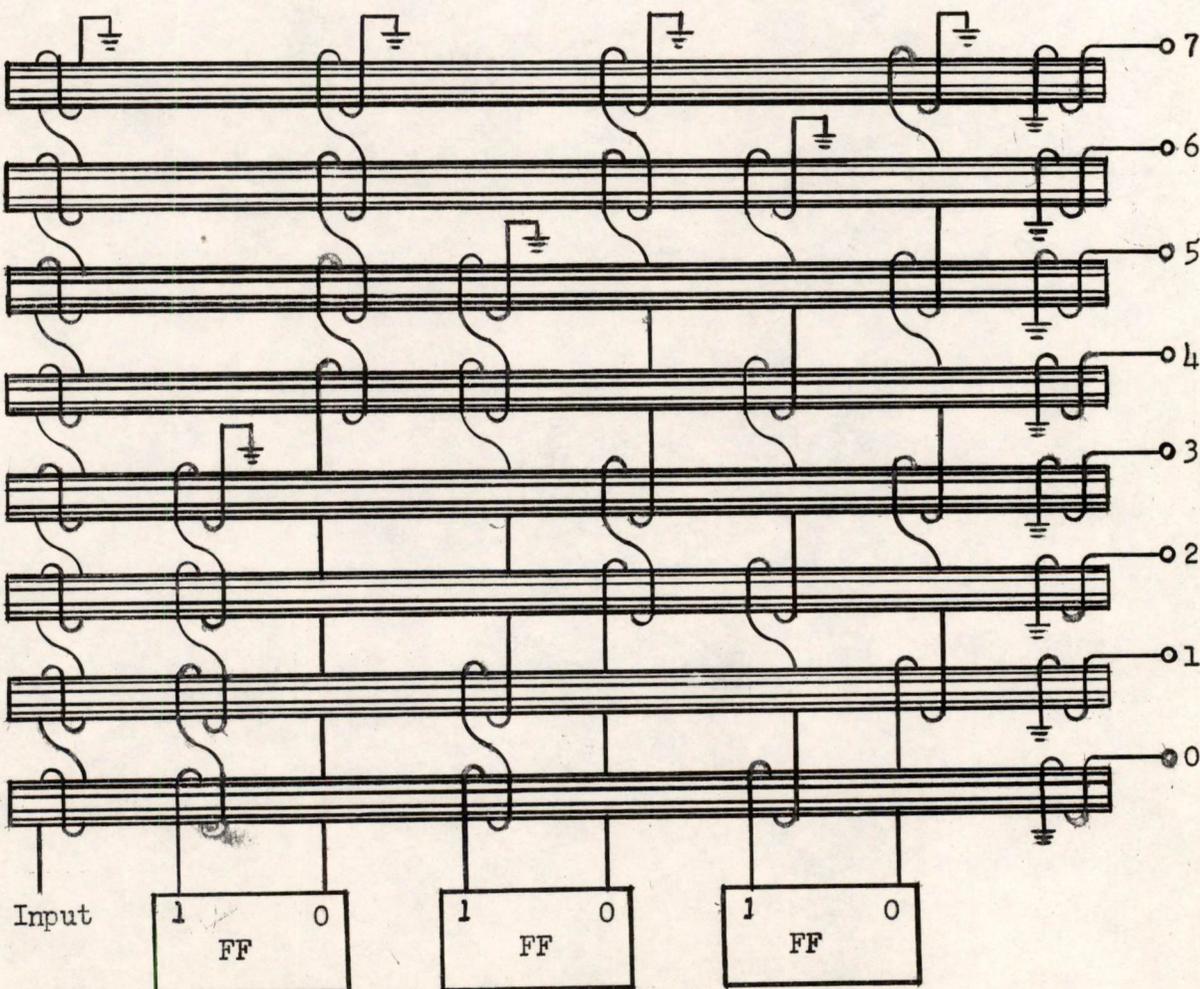


(d)

FIGURE 2.

Magnetic Matrix Switch<sup>8,9,10</sup>

The structure of the magnetic matrix switch (Figure 3) is similar to the diode matrix switch just discussed. Output selection is made possible by the use of saturable cores which have rectangular hysteresis loops. It is assumed that the output current of each flip-flop is great enough that the presence of an input pulse will cause a net flux change only in the selected core. Each core has an output winding and an input winding as in a transformer. The primary disadvantages of this type of switch are slow operating speeds and undesirable noise levels.

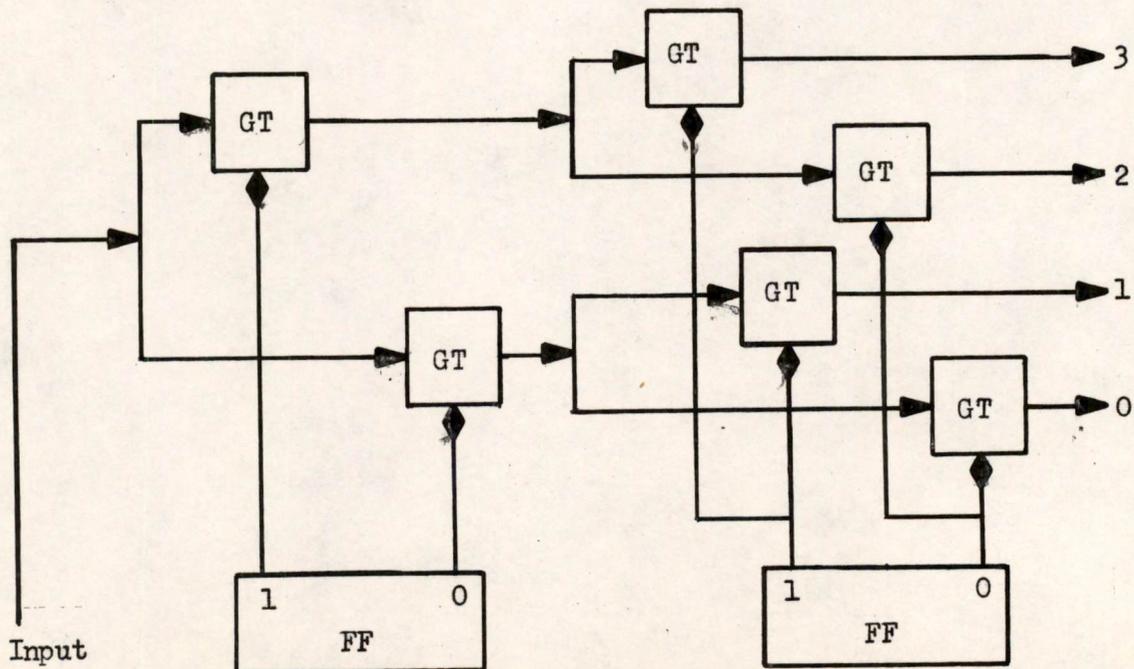


MAGNETIC MATRIX SWITCH  
FIGURE 3.

Gate Tube Whiffle-Tree Switch

Gate tubes can be connected to form multiposition switches. Figure 4 is a 4-position gate tube whiffle-tree switch. A two digit order is applied to the two input flip-flops. The output will be determined by the series of gate tubes that are selected. For example, to select output number 1, the coded input to the flip-flops would be 01. Input pulses pass through the series of gate tubes that complete a path to the selected output. The disadvantages of this type of switch are the delay inherent in each gate unit and the large number  $(2^{n+1} - 2)$  of gate tubes required to make a moderately large switch.

A gate tube with multiple control grids<sup>11</sup> to eliminate these difficulties has been suggested. However, gate tubes with n control grids have not been developed with an n much larger than 3 which restricts the maximum practical size of this type switch.



GATE TUBE WHIFFLE-TREE SWITCH

FIGURE 4.

PROPOSED PROCEDURE

Transistor Switching

In the past five years transistor switching circuitry has been investigated for computer work and is now beginning to play an important role in the computer field; however, multiposition selection switches have not been investigated. It is anticipated that an all-transistor multiposition selection switch can eliminate many of the difficulties found in the other types of switches. One promising transistor for this particular application is the Philco surface barrier transistor currently under mass production. This study will center about several switching configurations of surface barrier transistors as outlined in the Probable Procedure.

Probable Procedure

1. Transistor "and" and "or" circuits will be analyzed with regards to speed, rise and fall times, power considerations and limitations, maximum size, etc. Shown in Figure 5 are typical "and" and "or" circuits with their respective transmissions, according to the Boolean Algebra notation (the output is at zero volts).

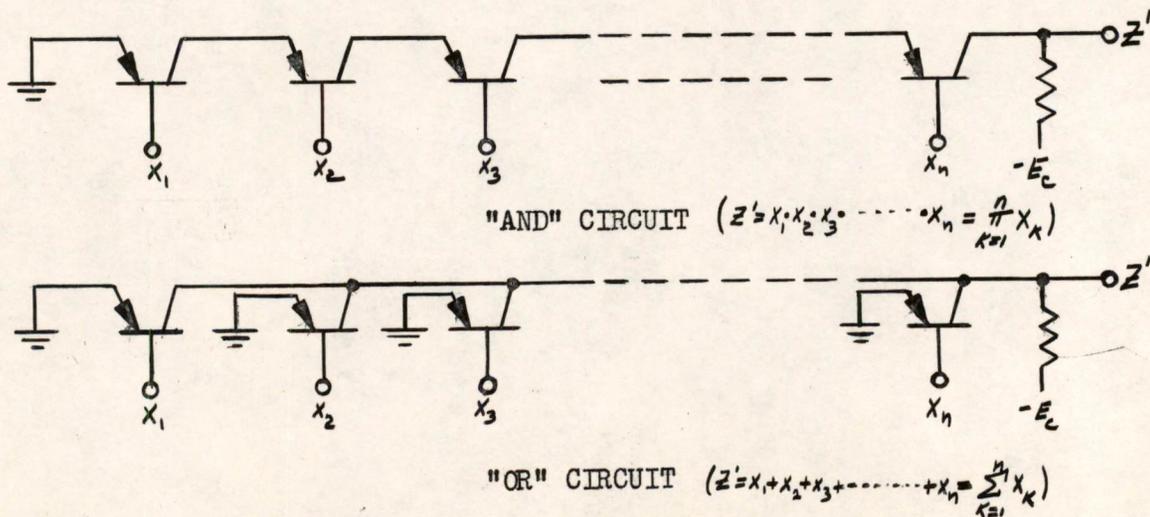
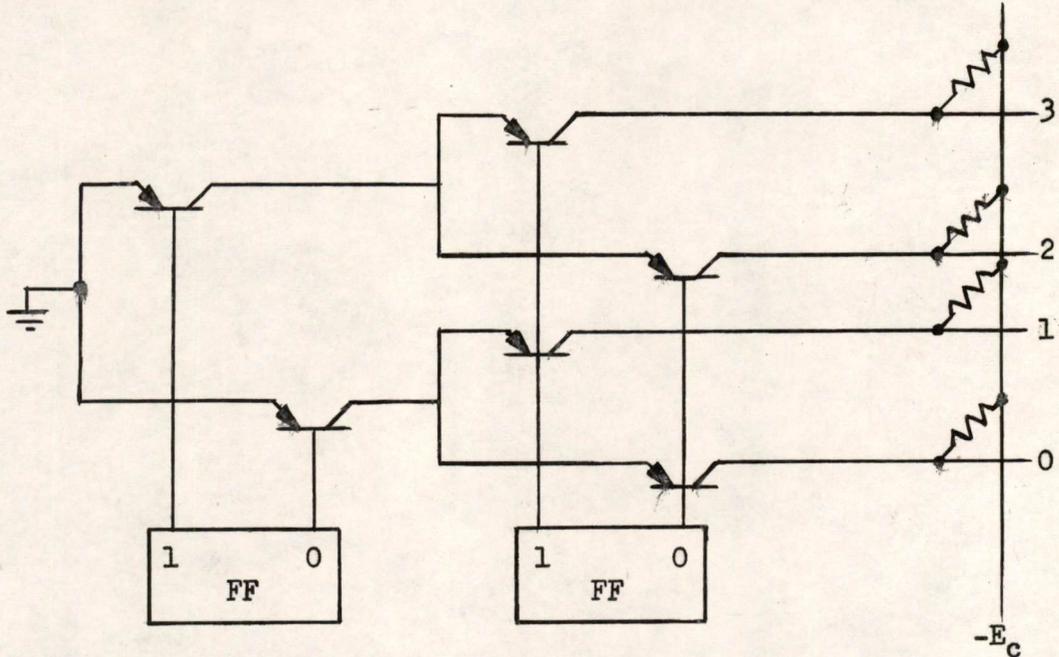
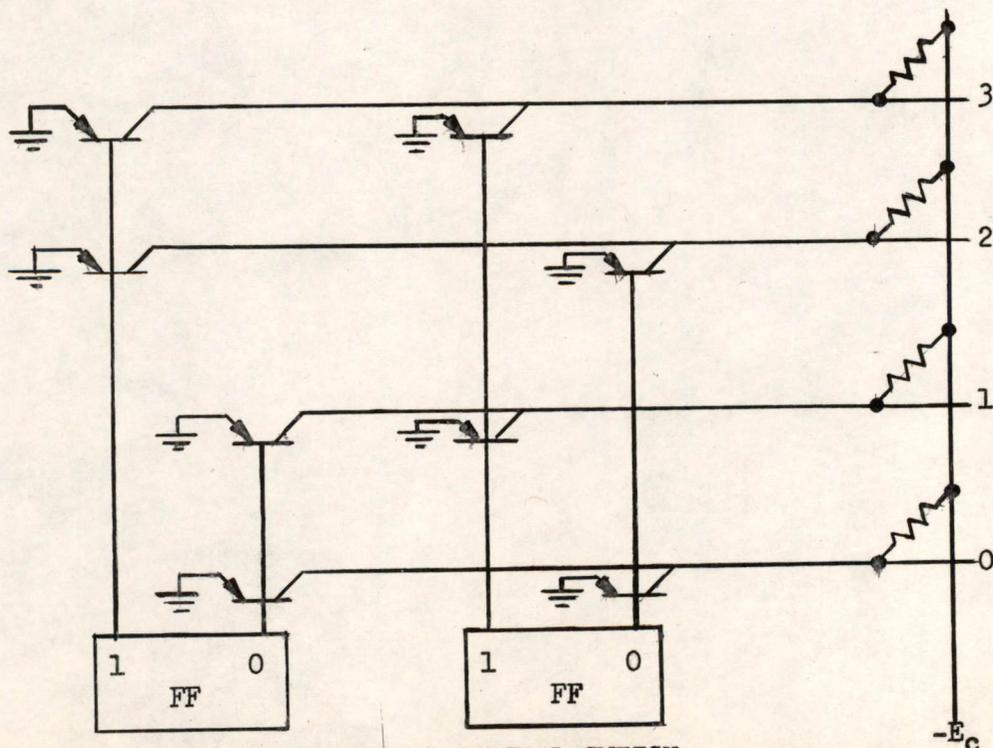


FIGURE 5.

2. Transistor switches employing the use of "and" and "or" circuits will be investigated. These switches will also be referred to as control switches for reasons which will become apparent later. Shown in Figure 6 are typical "and" and "or" types of control switches.



FOUR POSITION "AND" CONTROL SWITCH

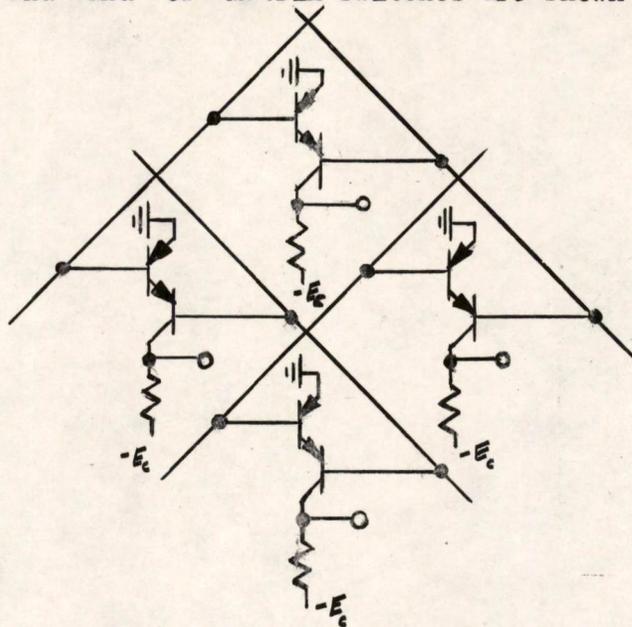


FOUR POSITION "OR" CONTROL SWITCH

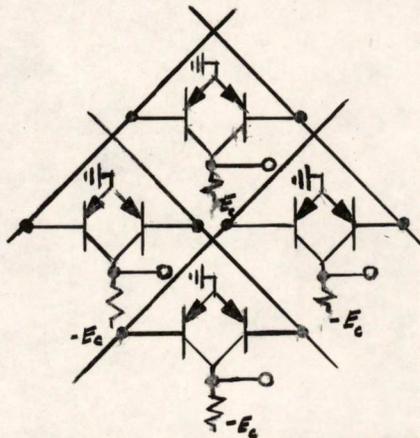
FIGURE 6.

3. Transistor "and" and "or" types of matrix switches will be analyzed.

Two typical "and" and "or" matrix switches are shown in Figure 7.



FOUR OUTPUT "AND" MATRIX SWITCH

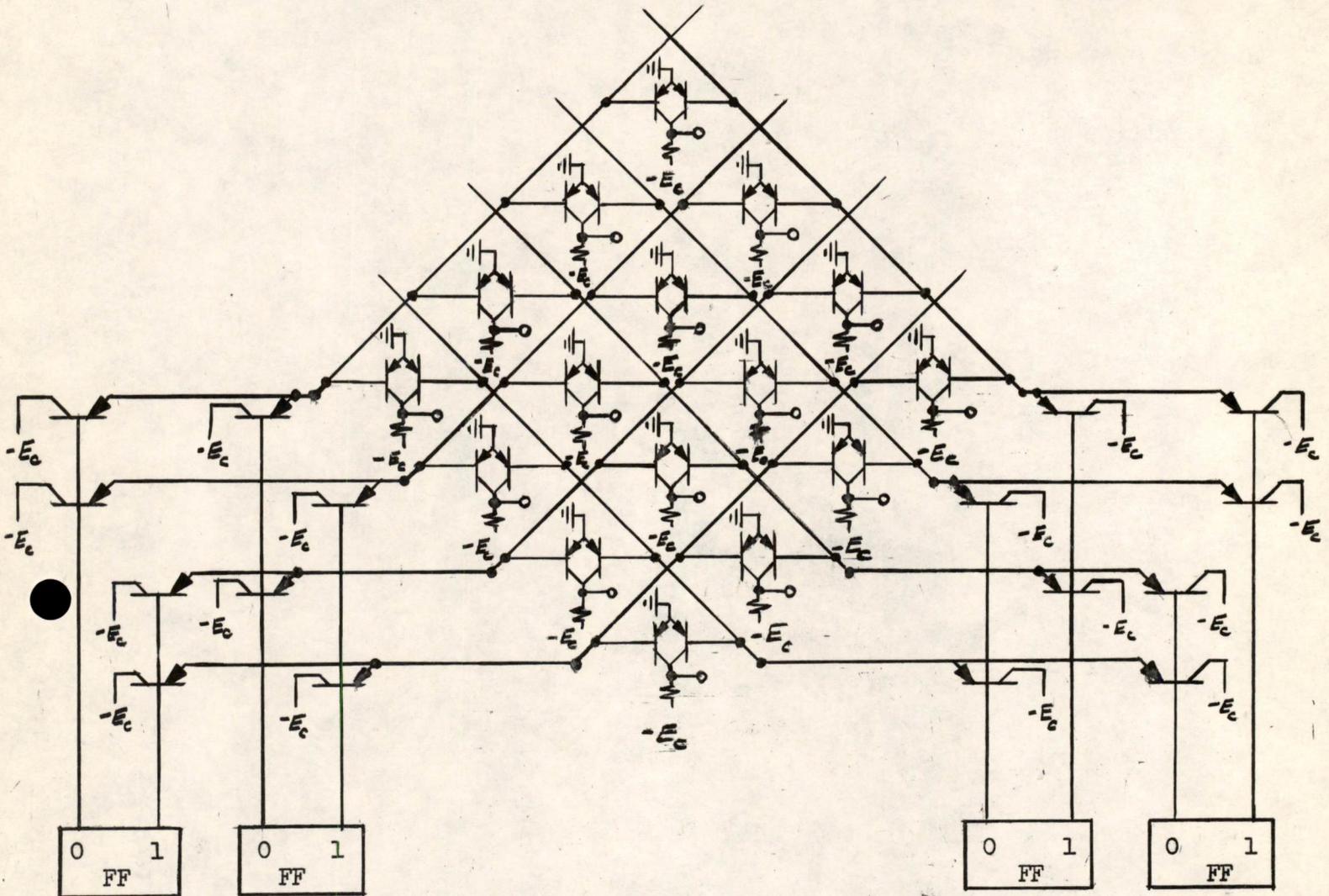


FOUR OUTPUT "OR" MATRIX SWITCH

FIGURE 7.

4. A transistor multiposition selection switch consisting of a matrix type switch controlled by two control switches will be designed, constructed, and analyzed. Shown in Figures 8 and 9 are typical multiposition switches. Static and dynamic analysis with consideration of economy will be studied for optimum design procedures.





SIXTEEN OUTPUT MULTIPosition SWITCH ( $-E_c$  SELECT)

FIGURE 9.

EQUIPMENT NEEDED:

Lincoln Laboratory will supply the test equipment and components necessary for this thesis. This equipment will primarily be composed of the Burrough's Standard Test Equipment.

ESTIMATED DIVISION OF TIME:

1. Preparation of the proposal.	50 hours
2. Further study of the literature.	25 hours
3. Experimental work and analysis.	150 hours
4. Correlation of results and formulation of deductions and conclusions.	75 hours
5. Preparation of the thesis report.	<u>100 hours</u>
Total	400 hours

SIGNED: Paul G. Griffith  
Paul G. Griffith

Date: September 29, 1955

SUPERVISION AGREEMENT:

In my opinion this problem is adequate for a Master's thesis;  
therefore, I am willing to jointly supervise the research and  
evaluate the thesis.

APPROVED: William K. Linvill  
W. K. Linvill  
Associate Professor of  
Electrical Engineering

APPROVED: Dudley A. Buck  
D. A. Buck  
Instructor in  
Electrical Engineering

PGG/dg

Distribution: Staff, Group 63  
R. H. Baker

BIBLIOGRAPHY

1. Forrester, Jay W., "Digital Information Storage in Three Dimensions Using Magnetic Cores", Servomechanisms Laboratory, Report R-187, May 1950.
2. Papian, William N., "A Coincident-Current Magnetic Memory Unit," Servomechanisms Laboratory, Report R-192, Sept. 1950.
3. Papian, William N., "New Ferrite-Core Memory Uses Pulse Transformers," ELECTRONICS, pp. 194-197, March 1955.
4. Brown, David R., "A High-Speed Multi-Position Electronic Switch," Servomechanisms Laboratory, Report R-157, March 1949.
5. Brown, David R., and Rochester, N., "Rectifier Networks for Multi-position Switching," Proceedings of the IRE, Feb. 1949.
6. Rajchman, Jan A., "Electronic Computers," RCA Laboratories, Contract # NOS-93260, 1942.
7. Crawford, Perry Orson, Jr., "Automatic Control by Arithmetic Operations," Physics Department, Master's Thesis, 1942.
8. Olson, Kenneth H., "A Multi-Position Magnetic Switch and its Incorporation into A Coincident-Current Memory," Lincoln Laboratory, Memorandum 1282, Sept. 1951.
9. Olsen, Kenneth H., "A Magnetic Matrix Switch and Its Incorporation Into A Coincident-Current Memory," Lincoln Laboratory, Report R-211, May 1952.
10. Rajchman, Jan A., "Static Magnetic Matrix Memory and Switching Circuits," RCA Laboratories, RCA Review, June 1952.
11. Rubinoff, Morris, "An Input Device Using Multiple Gates," Harvard University, Proceedings of a Symposium on Large-Scale Digital Calculating Machinery, 1948.

12. Davidson, George A., "A Transistor Selection System for a Magnetic-Core Memory," Lincoln Laboratory, Memorandum 6M-3717, June 1955.

6M-3938

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: THE LOGICAL STRUCTURE OF DIGITAL COMPUTERS  
The Turing Machine

To: Class Registrants

Abstracts to: All Lincoln Division Heads and Group Leaders

From: W. A. Clark

Approved: W. A. Clark  
W. A. Clark

Date: 5 October 1955

ABSTRACT: The logic of the Turing machine as a symbol manipulator is described and examples of counting and sorting are explained. A set of problems is included.

## INTRODUCTION

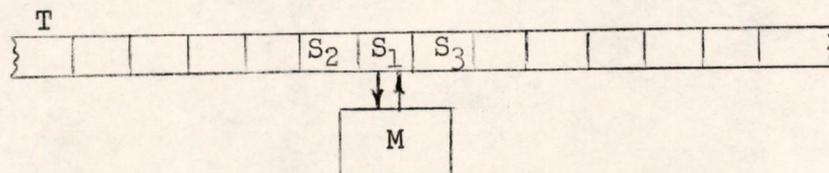
The subject of this course is the Logical Structure of Digital Computers. By "computer logic" one means the set of rules which the computer follows in carrying out its operations. Logical structure is to be distinguished from physical structure. The electronic components, wires, motors, and other hardware, comprising the physical structure of the computer, do no more than mechanize the operating rules defining the logical structure of the computer. It is completely irrelevant to the logic that the computer is built of vacuum tubes, or relays, or paper, as long as the rules are properly represented and followed.

The digital computer is essentially a symbol-manipulating machine. It accepts a set of symbols defining a problem to be solved and the data on which to operate. It then performs various operations on these symbols according to the rules defining its logical structure, and thereby produces a new set of symbols which comprise the solution to the problem. The rules thus take the form of a set of statements describing the manner in which certain symbols are to be replaced with new symbols. (Consider, for example, a particular sequence of five symbols. One useful rule in a computer dealing with this sequence would be: "If the first symbol is a '1', the second '+', the third '1', the fourth '=', and the fifth 'x', then replace the symbol 'x' with the symbol '2'.")

## THE TURING MACHINE

A simple abstract model of the general symbol-manipulation process (and, therefore, of digital computer logical structure) was formulated by the British mathematician, Turing<sup>1</sup> as a conceptual aid in proving certain results in mathematical logic. He defined a class of symbol processing mechanisms which he called simply "automatic machines," but which are now generally known as "Turing machines."

The elements of the Turing machine are illustrated below:



<sup>1</sup>A. M. Turing: On Computable Numbers, with an Application to the Entscheidungs Problem, Proc. Lond. Math. Soc., series 2, V24, pp. 230-265, 1936.

A machine,  $M$ , having a finite number of internal configurations or states operates on an infinitely long tape,  $T$ , which is divided into cells. Each cell is capable of bearing one symbol from a specified, finite set of symbols,  $S_0, S_1, S_2 \dots S_n$ , e.g., the alphabet, the digits, etc. The machine deals with one cell at a time (called the scanned cell) and can read the symbol in this cell and write a new symbol in its place, or move to the next cell to the right or to the left.

An operation is carried out concurrently with a jump from one machine state to another. This action is completely determined by the current state of the machine and the currently scanned symbol. Each move results in a new configuration of machine and tape in which the scanned symbol and machine state determine the next move, and so on.

A notation will now be described and some examples of Turing machines presented. This material will differ from that in Turing's original presentation, but the essential features are retained.

The operations to be discussed are:

- 1) Replace the scanned symbol,  $S_i$ , with the symbol  $S_j$ , abbreviated:

$$S_i : S_j$$

where  $i$  and  $j$  may have any particular values  $0, 1, 2, \dots n$ . The symbol  $S_0$  will represent blank tape to complete the description.

- 2) If the scanned symbol is  $S_i$ , move to the next cell on the right:

$$S_i : R$$

- 3) If the scanned symbol is  $S_i$ , move to the next cell on the left:

$$S_i : L$$

These operations can be abbreviated:

$$S_i : T_k \quad \left\{ \begin{array}{ll} T_k = \text{print } S_k & k = 0, 1, \dots n \\ T_k = \text{move "R"} & k = n + 1 \\ T_k = \text{move "L"} & k = n + 2 \end{array} \right.$$

The rules by which the machine operates are then formulated in terms of these operations and the internal states of the machine. Each rule will be of the form:

If	the machine is in state $p$	$(p = 1, 2, \dots, M)$
and	the scanned symbol is $S_i$	$(i = 0, 1, \dots, N)$
then	carry out operation $T_k$	$(k = 0, 1, \dots, N + 2)$
and	jump to state $q$	$(q = 1, 2, \dots, M)$

which will be abbreviated to the quadruple:

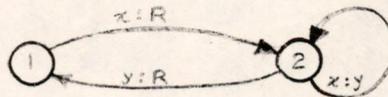
$$(p, S_i : T_k, q)$$

For example, the quadruple  $(3, x : R, 14)$  means "If the machine is in state 3 and the scanned symbol is  $x$ , move to the next cell on the right and jump to state 14."

The logical structure of the machine is thus specified by a finite set of quadruples of the above form.

The ordered pair of symbols  $p, S_i$  will be called a determinant, since it determines the subsequent move of the machine according to the remaining terms in the quadruple. To be consistent, the requirement is imposed that no two quadruples describing a given machine can have the same determinant.

The logical structure of a Turing machine may be represented conveniently as a network in which each node corresponds to a state of the machine and each directed branch between nodes corresponds to a jump between states. The branch is labeled with the operation which occurs during the jump. For example, the network



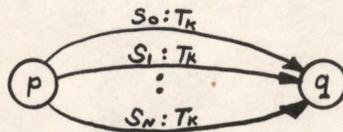
corresponds to the set of three quadruples

$$\begin{aligned} &(1, x : R, 2) \\ &(2, x : y, 2) \\ &(2, y : R, 1) \end{aligned}$$

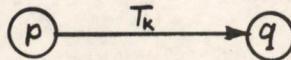
A drawing of the network for a given machine is variously called a state diagram or transition diagram. Two conventions which simplify

the drawing of transition diagrams for processes involving a large number of symbols  $S_0, S_1, S_2, \dots, S_n$  are the following:

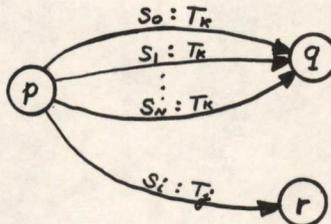
- a) If all branches from a given state,  $p$ , lead to state  $q$ , and involve the same operation,  $T_k$



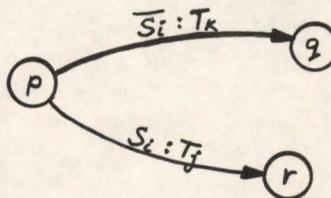
then the diagram may be abbreviated to



- b) If all branches except the one for a particular scanned symbol,  $S_i$ , lead from state  $p$  to state  $q$ , and involve the same operation,  $T_k$



then the diagram may be abbreviated to

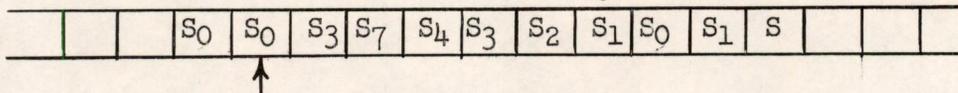


where  $\overline{S_i}$  is read "not  $S_i$ ."

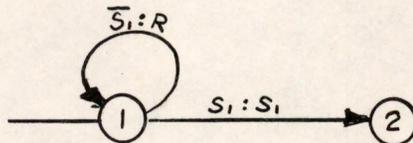
A few examples of Turing machines will now be given to illustrate the preceding definitions and concepts.

Example 1

Given a tape on which the symbols  $S_0, S_1, S_2, \dots, S_n$  appear in any order and number. The machine starts in state 1 scanning any cell to the left of a cell holding the symbol,  $S_1$



The machine is to "hunt" for the first cell to the right which holds  $S_1$  and jump to state 2 when it is scanning this cell. The transition diagram is



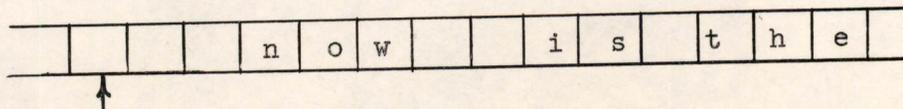
The machine remains in state 1 and scans the next cell on the right until  $S_1$  is found, whereupon it jumps to state 2 with no change of symbol on the scanned cell.

Example 2

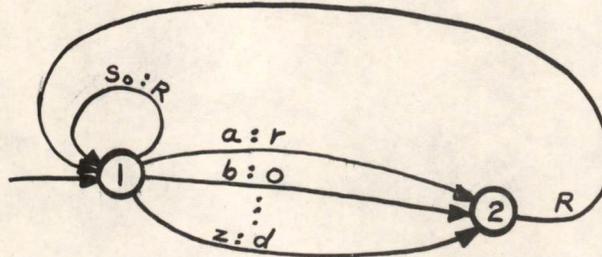
Consider a process of simple cryptographic encoding. Letters of the alphabet are to be scrambled according to the code

a → o  
 b → r  
 c → p  
 .  
 .  
 .  
 z → d

e.g., the word "cab" becomes "por", etc. The message to be encoded is printed on the tape with an arbitrary number of spaces between words.



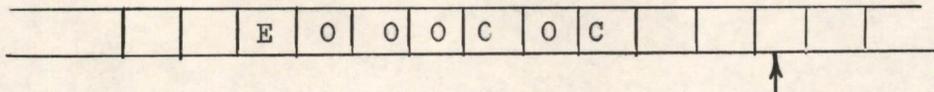
The machine starts in state 1, scanning the indicated cell. Its diagram is



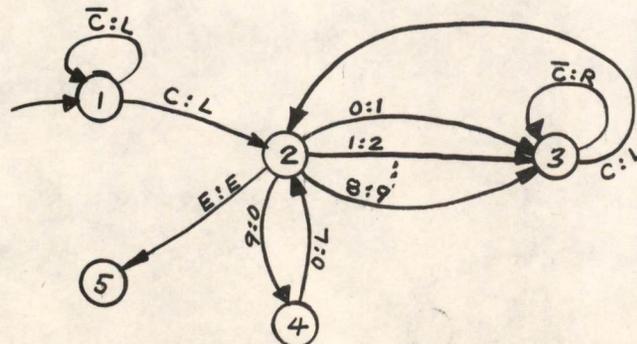
The symbol  $S_0$  represents blank tape. The machine continues indefinitely, changing letters of the message to their equivalent code value.

### Example 3

A block of five cells, each holding the digit 0, is separated from the rest of the tape by the symbol E on the left and C on the right.



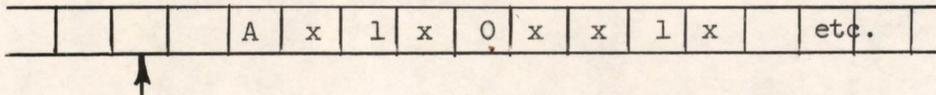
The machine starts in state 1, scanning a cell to the right of C. It is to print in succession the 5-digit decimal numbers from 0 to 99,999 on the marked block of cells, reset them to 0, and then jump to state 5. Its diagram is



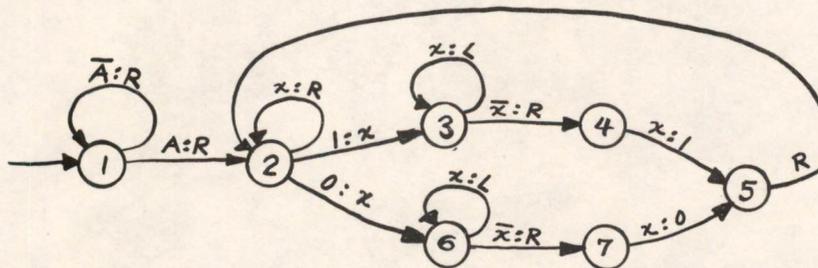
The machine finds the first digit following symbol C and jumps to state 2. If this digit is not 9, it is replaced with the next larger digit and the machine jumps to state 3. In state 3, the machine moves right until it finds C and backs up one cell, jumping back to state 2. When the digit is 9, it is replaced with 0 and the machine jumps to state 4 and moves left one cell (corresponding to a "carry" from one digit position to the next), returning to state 2 again. Clearly, this process results in the printing of the required sequence of numbers up to 99,999. At this point, the repeated sequence of transitions (2, 9 : 0, 4) (4, 0 : L, 2) resets the five cells to 0 and the process terminates with (2, E : E, 5).

#### Example 4

Consider a tape marked with A, 1, 0, and x in the following manner:

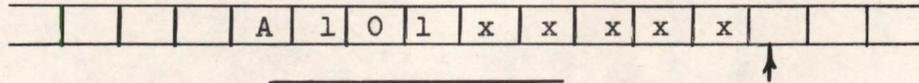


The 1's and 0's are intermixed with x's to the right of A in an arbitrary way. The machine starts in state 1, scanning a cell to the left of A, and is to compact the sequence of 1's and 0's into a block following A. The order of the 1's and 0's is to be retained. The diagram is



In state 1, the machine finds the first symbol to the right of A and jumps to 2. In state 2, the machine skips over cells holding x's and finds the nearest 0 or 1, replaces it with an x and jumps to state 6 or 3, respectively. In the

case illustrated, the nearest non-x symbol is a 1, and the machine will go to state 3. In state 3, the machine finds the leftmost x, jumps to 4, prints a 1, and moves to the right, returning to state 2 via 5. Note that in taking either the upper or lower branch,  $(2, 1 : x, 3)$  or  $(2, 0 : x, 6)$ , the machine in effect "remembers" that the last symbol scanned was a 1 or 0, respectively. The reader should verify that the tape illustrated becomes:



It is possible to restrict the Turing machine to two symbols, 0 and 1, without loss of generality.

Consider a problem which is expressed in terms of four symbols:  $S_0, S_1, S_2, S_3$ . These can be encoded into groups of 1's and 0's in many ways. For example:

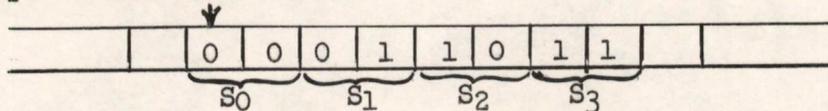
$S_0 = 000$   
 $S_1 = 001$   
 $S_2 = 010$   
 $S_3 = 100$

(1)

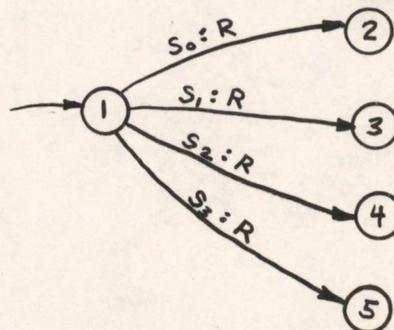
$S_0 = 00$   
 $S_1 = 01$   
 $S_2 = 10$   
 $S_3 = 11$

(2)

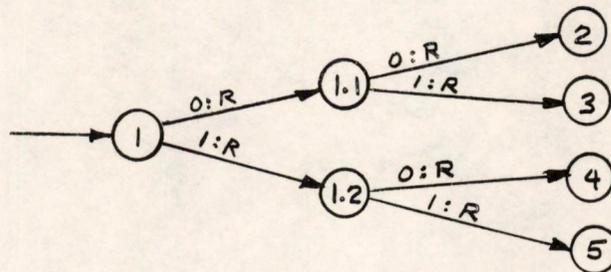
Case (2) will be discussed in more detail. The tape is divided into groups of two cells each:



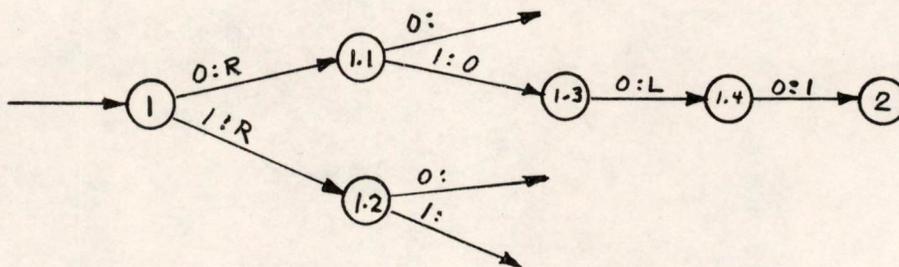
Then consider a section of the transition diagram of a machine dealing with  $S_0, S_1, S_2, S_3$  which is of the form



This is equivalent to the following in which only 0 and 1 are used:



Similarly, a change of symbol, say  $(1, S_1 : S_2, 2)$  can be represented as:



Note that the interpretation of the sequence of 1's and 0's depends on the direction of travel. This dependence could be eliminated by using a symmetrical code:

$$\begin{cases} S_0 = 000 \\ S_1 = 010 \\ S_2 = 101 \\ S_3 = 111 \end{cases}$$

We have now shown that the operations of any Turing machine can be reduced to the set

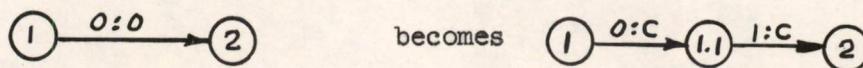
$$\begin{cases} \text{print 0} \\ \text{print 1} \\ \text{move right one cell} \\ \text{move left one cell} \end{cases}$$

This can be reduced still further. The only possible symbol-printing situations are:

$$\begin{matrix} 0 : 1 \\ 1 : 0 \end{matrix} \Bigg\}$$

$$\begin{matrix} 0 : 0 \\ 1 : 1 \end{matrix} \Bigg\}$$

The first two cases involve a change of information on the tape; the second two do not. We can define an operation "complement," abbreviated "C", to replace the two print operations. The cases involving no change of symbol are replaced with two complement operations done in sequence:



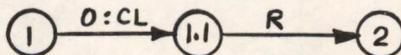
Thus, the set of Turing machine operations reduces to:

$$\left. \begin{matrix} C \\ L \\ R \end{matrix} \right\}$$

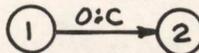
This can be reduced again to any of the three sets:

- (1)  $\left\{ \begin{matrix} CL & \text{complement and move left} \\ R & \text{move right} \end{matrix} \right.$
- (2)  $\left\{ \begin{matrix} L & \text{move left} \\ CR & \text{complement and move right} \end{matrix} \right.$
- (3)  $\left\{ \begin{matrix} CL & \text{complement and move left} \\ CR & \text{complement and move right} \end{matrix} \right.$

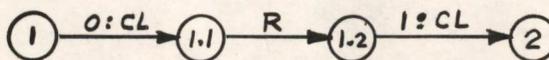
The proofs for (1) and (2) reduce to showing that "CL" can be broken into "C" and "L":



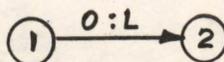
is the equivalent of



and

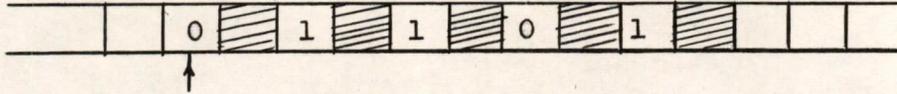


is the equivalent of

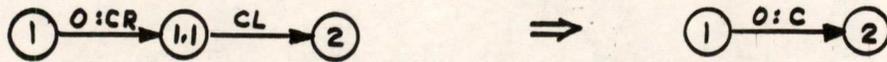


The other possible actions on 0 and 1 for cases (1) and (2) are proved in a similar manner.

Case (3) requires a different arrangement of symbols on the tape. Alternate cells are used to hold the symbols of the problem. The cells in between aid in "phasing" the complementing, but hold no significant information:



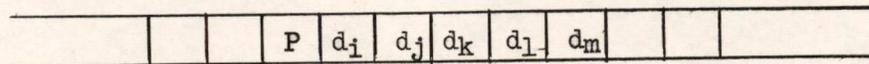
The equivalent forms are:



The other forms can be obtained directly from these.

PROBLEMS

- 1.1) A Turing machine tape is marked in the following manner:



Symbol "P" marks the beginning of a block of 5 different decimal digits,  $d_i, d_j, d_k, d_l, d_m$ . (e.g.,  $\overline{70934}$ ).

Describe (draw a state-and-transition diagram of) a machine which will rearrange the digits in descending order on the 5 cells following "P" and move on to the right when finished. Any additional symbols and cells may be used in the process providing they are erased upon completion. The machine is to start on any cell to the left of "P".

- 1.2) Restate problem 1.1 in terms of a tape on which only the symbols " $S_0$ " and " $S_1$ " appear. (Invent a suitable code for the symbols "P", blank, 0, 1, . . . 9, etc., and describe the initial tape configuration). Redraw the state-and-transition diagram accordingly, using only the operations "complement" (change  $S_0$  to  $S_1$  or  $S_1$  to  $S_0$ ), "move right", and "move left", abbreviated "C", "R", and "L", respectively.
- 1.3) Non-erasable tape can be defined as tape on which it is possible to write a symbol in a given cell only if the cell is blank. Show that it is possible for any Turing machine to use non-erasable tape.

6M-3938, Supplement 1

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: THE LOGICAL STRUCTURE OF DIGITAL COMPUTERS  
The Universal Turing Machine

To: Class Registrants

Abstract to: J. C. Proctor, C. W. Farr

From: W. A. Clark

Date: 20 October 1955

Abstract: The complexity of a Turing machine can be measured by the number of quadruples defining its logical structure. At a certain level of complexity, it becomes possible to design a Turing machine which is universal in the sense that it can perform any calculation which any other Turing machine can perform. The Universal Turing Machine achieves this generality by having the ability to simulate other Turing machines, even those which are more complex. This simulation process is described in terms of the manipulation of the quadruples themselves. An example of a Universal Turing Machine is presented in detail, and a set of related problems is included.

*W. A. Clark*

W. A. Clark

WAC/jhk

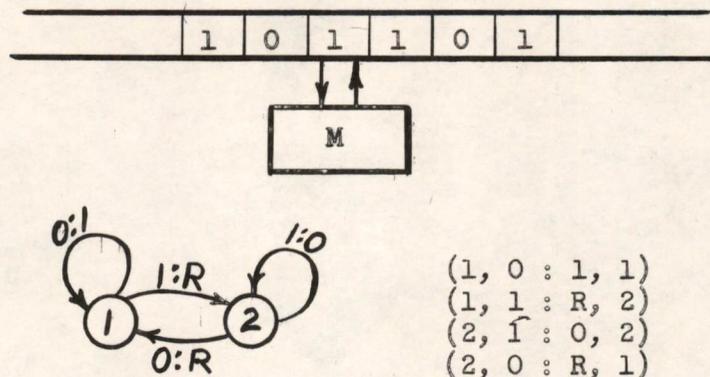
THE UNIVERSAL TURING MACHINEPreliminary Remarks

We have seen that a very general class of Turing machines can be defined which use a single tape on which only the symbols "0" or "1" appear in any cell ("0" corresponding to blank tape) and which perform only the operations "print 0," "print 1," "move right one cell," and "move left one cell." For convenience in subsequent reference, we will call these machines "Class A Turing machines" to distinguish them from other machines which use more symbols, different operations, more tapes, etc. Within a given class one might measure the complexity of a particular machine by the number of its internal states or, perhaps better, by the number of quadruples required to describe its logical structure. Thus, a 100-quadruple Class A Turing machine would be more complicated than a 10-quadruple Class A Turing machine.

It seems reasonable to attempt to relate a machine's complexity to its capability. For example, it is possible to combine a machine,  $M_1$ , which is capable of counting, with a machine,  $M_2$ , which is capable of ordering a set of numbers, and thereby obtain a more complicated machine,  $M_3$ , which is capable of both ordering and counting. It might be supposed that it is always possible to increase the generality of a machine by increasing its complexity in this way. The fact is, however, that there is a critical complexity beyond which no further increase in generality can be guaranteed! That is, at a certain level of complexity it becomes possible to design a Turing machine which is universal in the sense that it can perform any calculation which any other Turing machine can perform, even if the other machine is more complicated than the universal machine. The universal Turing machine achieves this generality by having the ability to simulate any machine whose calculation it is required to duplicate. The tape of the simulated machine appears as a designated sequence of cells on the tape of the universal machine. We will consider these points in more detail later.

Quadruple Manipulation

The simulation is itself a symbol manipulation process in which the symbols represent the set of quadruples describing the simulated machine. As an example of the manipulation of quadruples, consider the following simple Class A machine and the set of quadruples describing its logical structure:



This machine, starting in state 1, will print the sequence 10101010..  
 ..... Its operation will now be described in terms of the quadruples  
 and scanned symbols:

Define active quadruple to be that quadruple which describes the action of the machine at any given point in the process. The active determinant is then the determinant found in the active quadruple (see page 3). The first term of the active quadruple will be called the initial state, and the last term, the final state. The second term of the active quadruple is, of course, the scanned symbol, and the third term is the specified operation.

In the illustration, if the machine is in state 1, the active quadruple is the second one in the list, namely:

$$(1, 1 : R, 2)$$

The machine will move one cell to the right and jump to the final state 2. State 2 thus becomes the initial state of the next machine action and the new scanned symbol is a "1". Therefore, the next active determinant will be:

$$\underline{2, 1}$$

The next active quadruple can be found by again examining the list of quadruples and finding the quadruple which starts with the determinant 2, 1. In this case, it is the third determinant:

$$(2, 1 : 0, 2)$$

The machine will print a "0" and jump to the final state 2. The scanned symbol is now "0" and the next initial state is 2;

thus, the next active determinant will be

2, 0

and the next active quadruple is found to be

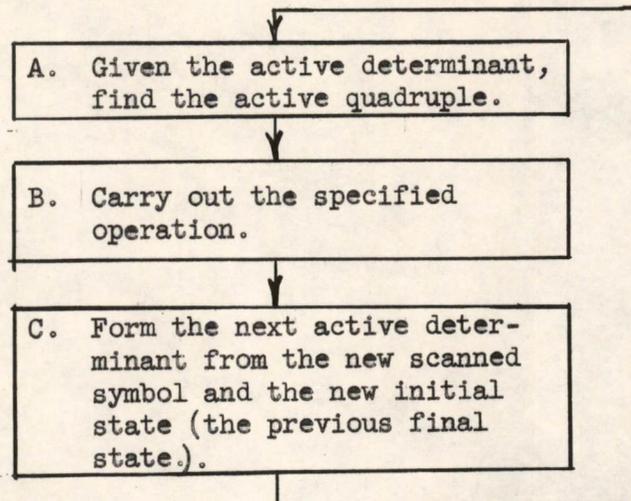
(2, 0 : R, 1)

etc.

This example illustrates the use of the set of quadruples in describing a sequence of machine actions.

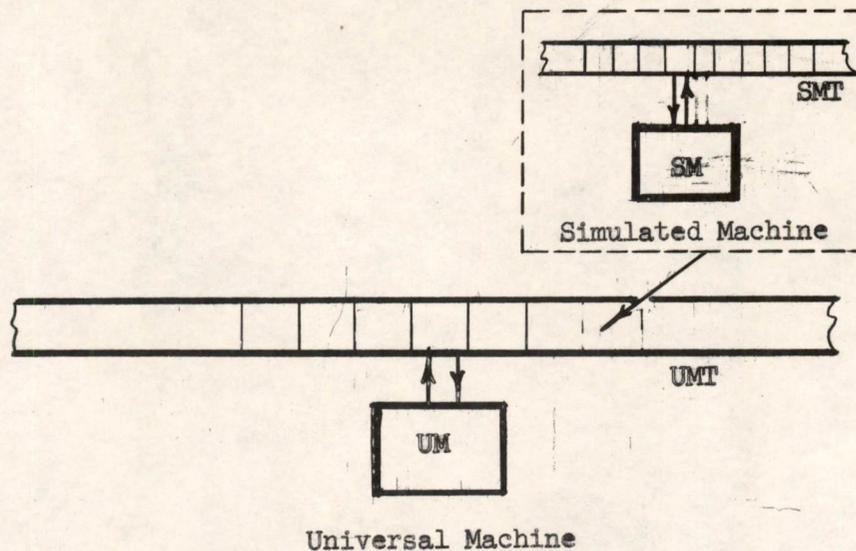
#### General Description of the Universal Turing Machine

The basic simulation process can be represented in the following way:



It is, of course, necessary to start this sequence with the first active determinant at A.

The Universal Turing Machine, UM, will be designed to carry out the above steps A, B, and C for any list of quadruples describing a given simulated machine, SM, which will be encoded in a suitable manner and printed on the universal machine tape, UMT. As we mentioned earlier, this tape will also hold a sequence of cells which correspond to the cells of the simulated machine tape, SMT. It will also require cells on which to print the active determinant symbols, and cells to mark significant points on the tape, e.g., the SMT scanned cell.



Before UM begins its calculation, the quadruples defining SM are printed on UMT, the cells on UMT corresponding to SMT are marked to match the initial configuration of SMT, and the first active determinant is printed on UMT. UM is started in a specified initial state scanning a specified cell on UMT. It then carries out steps A, B, and C without end and prints the results of SM's calculation on the designated UMT cells corresponding to SMT. It is assumed that the set of symbols used by SM is included within the set of symbols used by UM.

#### Detailed Description of a Class A Universal Turing Machine

The general description of the previous section will now be related to a particular Class A Universal Turing Machine. It is seen that the first problem is that of finding a suitable code using the symbols 0 and 1 to represent quadruples, determinants, etc. The coding scheme presented here is the work of E. F. Moore who employed it in a description of a 3-tape universal machine<sup>2</sup>. The second problem is that of finding a suitable arrangement of the symbols on the universal machine tape. Finally, a description of the universal machine itself must be developed.

Consider first the coding of a Class A machine quadruple  $(r, S_i : T_k, s)$ . Each determinant must be one of the two forms:

---

<sup>2</sup>E. F. Moore: A Simplified Universal Turing Machine, Bell Telephone System Monograph 2098, presented at the Meeting of the Association for Computing Machinery, Toronto, Canada, Sept. 8, 1952.

or

$r, 0$

$r, 1$

where  $r$  takes on any integral value  $1, 2, \dots, M$  (for an  $M$ -state machine). The specified operation,  $T_k$ , is any one of the four forms:

$O$   
 $1$   
 $R$   
 $L$

and finally, the final state,  $s$ , is an integer from 1 to  $M$ .

The scheme proposed by Moore is the following:

Determinant:

Code  $r, 0$  as a block of  $3r + 1$  successive 1's  
       "  $r, 1$  " " " "  $3r + 2$  " 1's

Operation:

Code	$O$	as	$O$	immediately	following	determinant.
"	$1$	as	$OO$	"	"	"
"	$R$	as	$OOO$	"	"	"
"	$L$	as	$OOOO$	"	"	"

Final State:

Code  $s$  as a block of  $3s$  successive 1's immediately following the operation.

For example, the quadruple  $(1, 0 : 1, 1)$  would become:

111100111

and the quadruple  $(1, 1 : R, 2)$  would become:

11111000111111

A list of quadruples is coded by stringing together in any order the codes of member quadruples, separating one quadruple from the next by at least one 0. For example, the machine described on page 14

$(1, 0 : 1, 1)$   
 $(1, 1 : R, 2)$   
 $(2, 0 : R, 1)$   
 $(2, 1 : O, 2)$

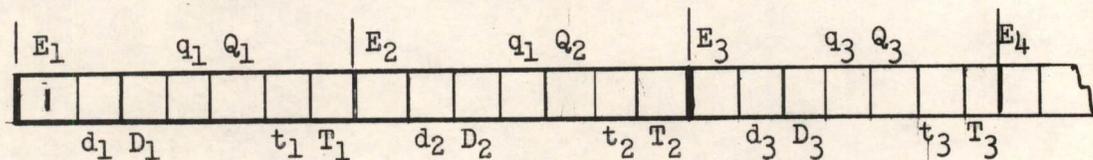
is completely (although not uniquely) coded by the sequence:

...00001111001110111111110111111000111111000111001111100011111100...

Notice that any block of ones in this sequence has a unique interpretation, e.g., a block of  $N$  ones represents a final state only if  $N$  is divisible by 3, etc. A block of zeros following a determinant represents an operation; a block of zeros following a final state is simply a separation.

We now proceed to describe the arrangement of symbols on the universal machine tape, UMT. UMT will need to be endless only on the right. The class of machines which the universal machine will simulate will also use tapes which are endless only on the right. This causes no restriction in the generality of these machines over that of doubly endless tape machines (see problem 2.1).

UMT will be divided into groups of 7 cells each. The quadruple list, active determinant, SMT cells, and various marking cells will be interleaved on UMT in the following manner:



UMT

The E cells are used to mark the end of the tape (only  $E_1$  holds a "1"; all other E cells hold "0". These are not changed.)

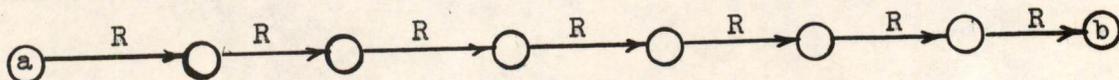
The D cells hold the active determinant.

The Q cells hold the list of quadruples.

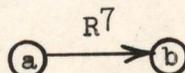
The T cells correspond to the cells of the simulated machine tape.

The cells labeled  $d$ ,  $q$ , and  $t$  are used to mark the following D, Q, and T cells, respectively. Only one cell of each will hold a "1" at any given point in the calculation. For example, a "1" in  $t_3$  indicates that  $T_3$  would be the scanned cell on SMT. The use of these marker cells will become clear later.

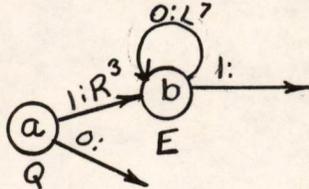
To go from cell  $Q_i$  to  $Q_{i+1}$  for example, it is necessary to slip over the intervening 6 cells. This process is diagrammed:



and will be abbreviated:



The process of locating the end of the tape will now be described. Suppose that UM is scanning some Q cell and is to find the cell E<sub>1</sub> if the scanned symbol is "1". The diagram is:



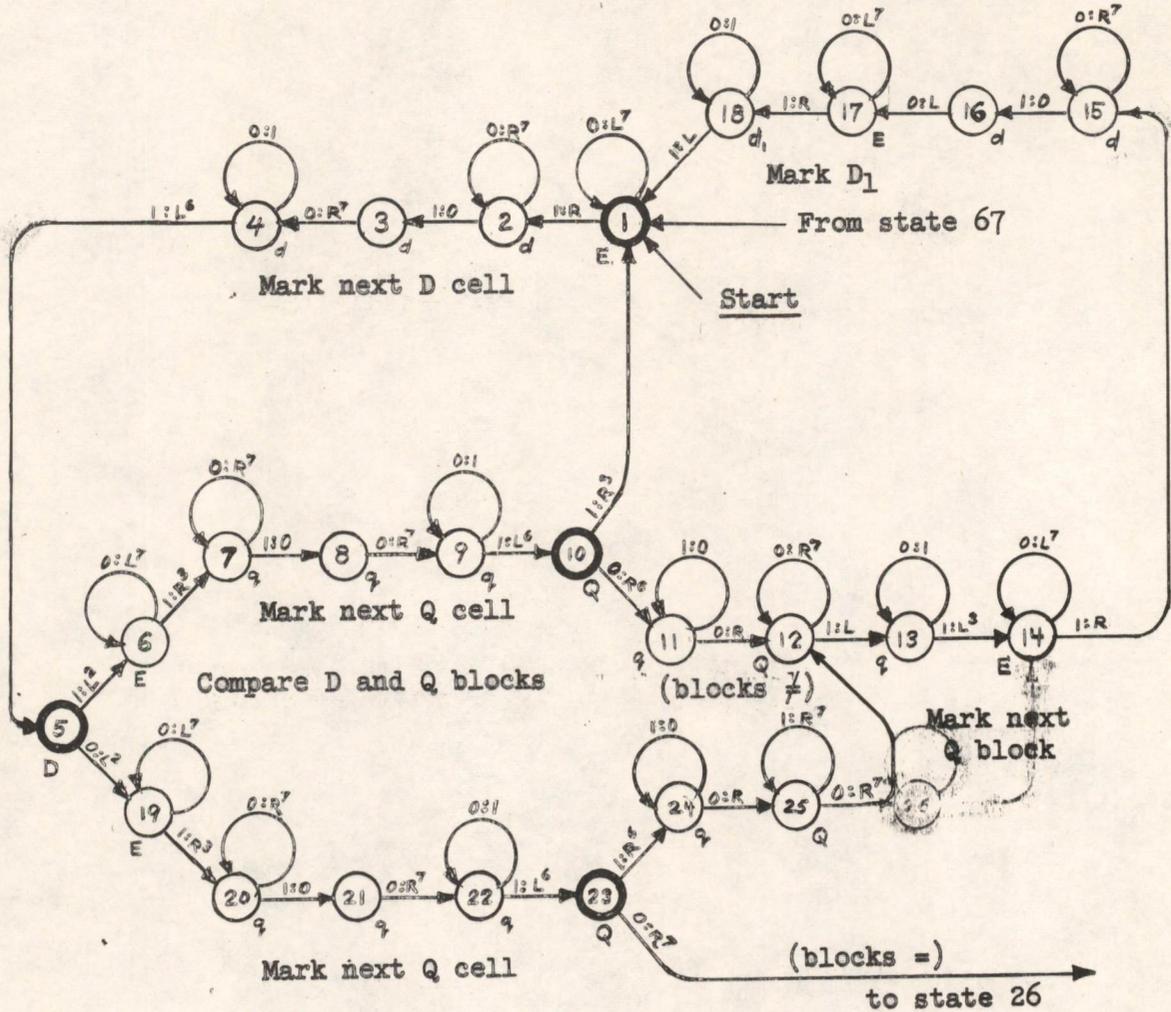
The letters Q and E below the state node in this diagram indicate which "phase" of the 7 tape phases the machine will end in after the transition to that state. The machine starts in state "a" scanning a Q cell, (i.e., in Q-phase). If the scanned symbol is a "1", the machine moves 3 cells to the right (to the nearest E cell) and jumps to state "b". Now in E phase, the machine scans successive E cells to the left until a "1" is found, which occurs on E<sub>1</sub> at the end of the tape.

UM will start in state 1 scanning E<sub>1</sub>. The quadruple list for SM is printed on the Q cells beginning with Q<sub>1</sub> and q<sub>1</sub> holds a "1" (q<sub>1</sub> marks Q<sub>1</sub>). The first active determinant is printed on D<sub>1</sub>, D<sub>2</sub>, D<sub>3</sub>, etc., and d<sub>1</sub> holds a "1" (d<sub>1</sub> marks D<sub>1</sub>). Finally, the T cells are marked according to SMT and the initial simulated scanned cell T<sub>s</sub> will be marked (t<sub>s</sub> holds a "1").

It will be noted that UM must move to the end of UMT before searching for a "1" on any marker cell in order to guarantee that the marked cell will not be missed.

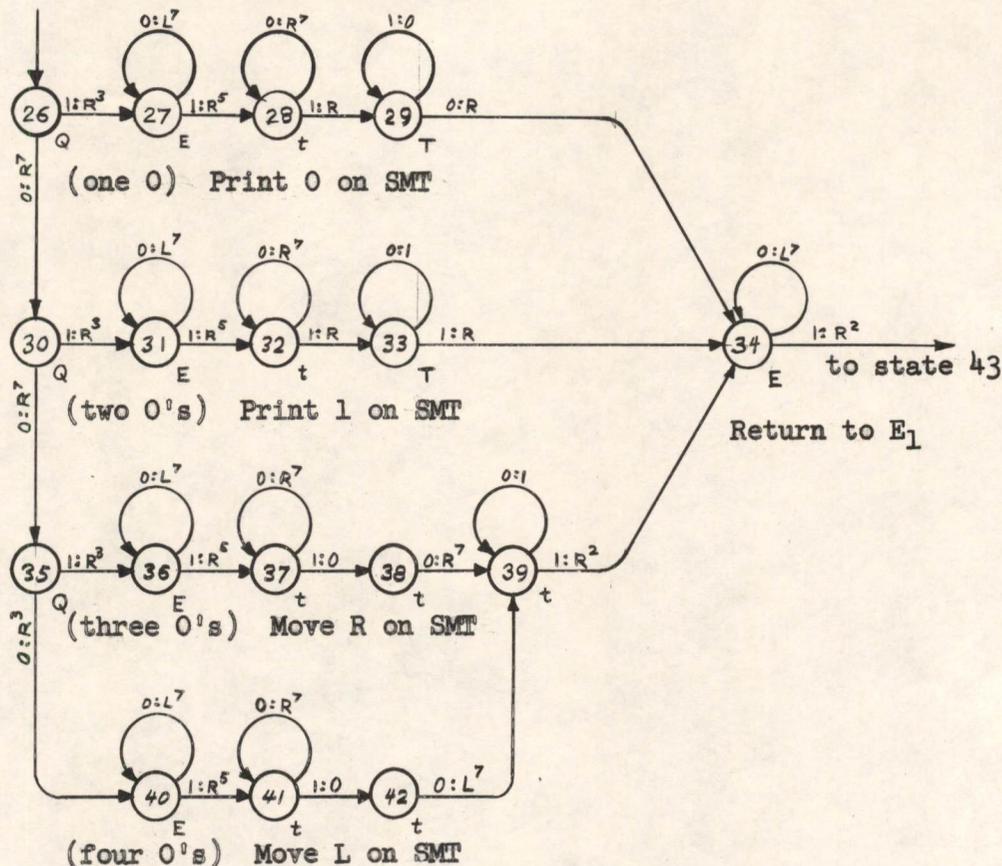
Parts A, B, and C of the operation of one particular Class A Universal Turing Machine will be described separately on the following pages.

A. Given the active determinant, find the active quadruple.



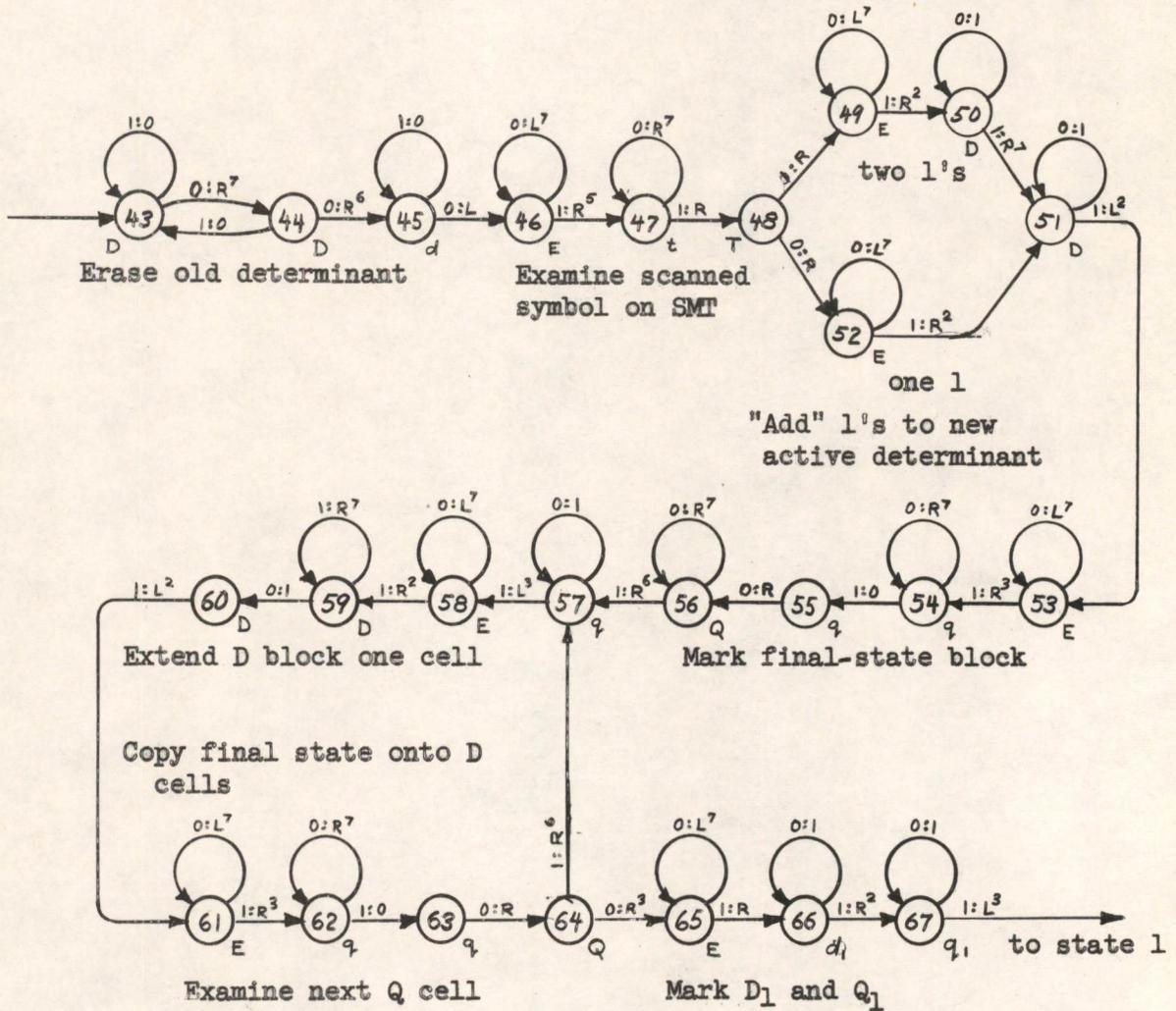
The next D cell is marked (a "1" is printed in the associated d cell) in states 2 to 4 and the previously marked D cell is examined (state 5). If it holds a "1", UM checks the currently marked Q cell (and marks the next Q cell) in state 10 and if it holds a "1", prepares to compare the next D and Q cells by returning to state 1. If the D and Q blocks are of unequal length, UM marks the next Q block of 1's, marks the beginning of the active determinant again, and starts the comparison with the new Q block by returning to state 1. When both D and Q cells hold 0's concurrently, the active quadruple has been found and UM jumps to state 26 to begin part B.

B. Carry out the specified operation.



The active quadruple has been located and UM determines which SM operation is specified by counting the number of successive 0's in the operation code. If one 0, UM finds the scanned cell on SMT and prints a "0"; if two 0's, a "1"; if three 0's, UM marks the T cell on the right of the currently marked T cell; if four 0's, UM marks the T cell on the left of the currently marked T cell. UM then returns to the end of the tape and moves on to part C.

C. Form the next active determinant from the new scanned symbol and the new initial state (the previous final state).



The old active determinant is erased (states 43 to 46) and the current scanned symbol on SMT is examined. If it is a "0", a "1" is printed on D<sub>1</sub> to start the next active determinant; if the scanned cell on SMT holds a "1", then a "1" is printed on both D<sub>1</sub> and D<sub>2</sub>. UM then combines the block of "1's" which code the final state of the active quadruple with the one or two 1's now on the initial D cells, thus forming the new active determinant. D<sub>1</sub> and Q<sub>1</sub> are marked (1's printed on d<sub>1</sub> and q<sub>1</sub>) and UM returns to state 1 to repeat A, B, and C for the next simulated transition.

PROBLEMS

- 2.1 Show that any problem which can be solved by a Turing machine using doubly infinite tape (infinitely long in both directions) can also be solved by a Turing machine using singly infinite tape.
- 2.2 Consider the class of Turing machines which use only the symbols "0" and "1" and which perform only the operations "print 0", "print 1", "move right one", "move left one", abbreviated O, 1, R, L, respectively.

How many one-state machines of this class are there? Two-state? N-state?

- 2.3 A Turing machine calculation which never uses more than a finite length of tape might be called limited; otherwise non-limited. (For example, the machine  $(1,0:0,1) (1,1:R,1)$  performs a limited calculation if the tape holds a "0" to the right of the scanned cell.)

Write the set of quadruples for each one-state machine of the class defined in 2.2 which is non-limited for all possible arrangements of symbols on the tape.

- 2.4 Design a Universal Turing Machine using as few internal states as possible. The design may use any finite number of symbols to aid in reducing the number of states required. Make a count of the number of states and the number of quadruples.

6M-3938, Supplement 2

Division 6 - Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington 73, Massachusetts

SUBJECT: THE LOGICAL STRUCTURE OF DIGITAL COMPUTERS  
Boolean Algebra

To: Class Registrants

Abstracts to: J. C. Proctor, C. W. Farr

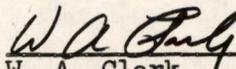
From: W. A. Clark

Date: 2 November 1955

Abstract: The symbol-printing operations in a Class A Turing machine can be described in terms of the tape cells themselves. For example, a machine which performs the sequence "If cell A holds "1" or if cell B holds "0", print "1" on cell C" is described by the statement:

$$(A^1 \text{ or } B^0):C^1$$

The manipulative aspects of this notation can be exploited in demonstrating that the rules for printing symbols define a Boolean Algebra.

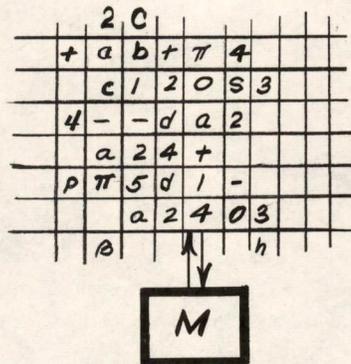
  
W. A. Clark

WAC/jhk

The Class A Universal Turing Machine described on the preceding pages provides us with one value for the critical complexity discussed earlier. A count of transitions shows that about 800 quadruples are required to describe this machine. Other universal machines using more symbols and fewer internal states have been designed (see also problem 2.4); Shannon has shown<sup>3</sup> that it is possible to design a two-state universal machine (and impossible to design a one-state universal machine) by using a large enough number of symbols.

#### Variations on the Turing Machine Theme

There are many possible variations to the Turing machine concept. In the example presented earlier, the machine operated on a linear array of symbols printed on an infinitely long tape. Another class of machine might be defined which operates on a two-dimensional array of symbols printed on an infinite plane.



The operations of this machine would include "move right," "move left," "move up," and "move down." Extension of these ideas to n-dimensional arrays readily follows.

A three-tape Turing machine was mentioned earlier<sup>4</sup>. In this case, the machine deals with three separate linear arrays scanning three cells simultaneously but operating on only one tape at a time. The transitions are described as sets of sextuples rather than quadruples, each determinant consisting of the initial state and the three scanned symbols. Again, extension to machines using more tapes or several n-dimensional arrays is possible.

Von Neumann has suggested a parts-manipulating machine analogous to

<sup>3</sup>C. E. Shannon and others: "Automata Studies," Princeton University Press (to be published shortly.)

<sup>4</sup>E. F. Moore: Op. cit.

to the Turing symbol-manipulating machine. This machine would operate in an environment containing hardware of various kinds such as nuts, bolts, wire, vacuum tubes, etc., and would construct another machine from these parts. Again, it is possible to design a universal constructing machine capable of constructing anything that any other constructing machine can construct. A case of particular interest is that of a machine which constructs a copy of itself. With the current trend toward automation, some of these ideas are being applied in practical situations.

### Summary Remarks

Before moving on to further discussion of the Turing machine and its relation to other topics, let us list some of the items which have been introduced in the preceding pages:

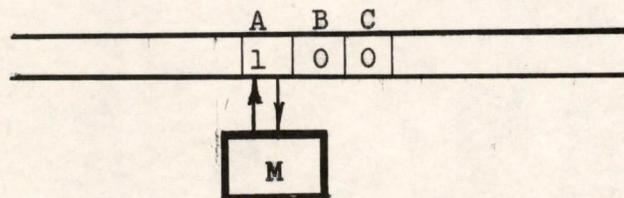
- |                        |                       |
|------------------------|-----------------------|
| 1. Logical structure   | 6. Coding             |
| 2. Operating rules     | 7. Machine complexity |
| 3. Stable states       | 8. Simulation         |
| 4. Transitions         | 9. Universality       |
| 5. Symbol manipulation |                       |

These are all items which will be discussed in more detail during the remainder of the course. It is interesting that the conceptually simple Turing machine serves to introduce so many of the basic ideas in the subject of digital computers.

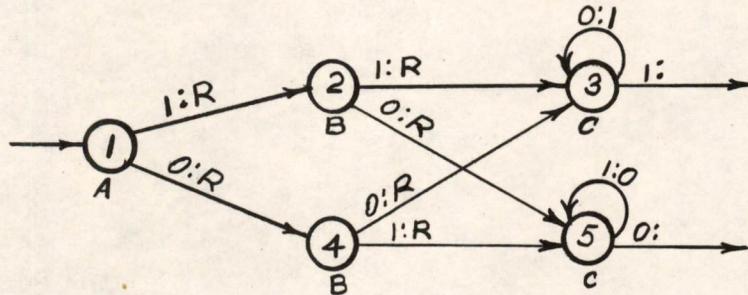
### BOOLEAN ALGEBRA

We now proceed to develop a manipulative notation which enables us to describe the action of two-symbol machines in terms of the cells themselves. This will lead to a kind of symbol-printing algebra which will be shown to have the properties of Boolean algebra.

Consider the following simple Class A Turing machine which operates on cells labeled A, B, and C.



for which the transition diagram is



The machine prints "1" on C if A and B hold the same symbol and prints a "0" on C if A and B hold different symbols. The symbol finally appearing in cell C depends on the symbols in A and B. There are, of course, several kinds of dependence relations possible, and the machine illustrated mechanizes only one of these relations. A word description of the illustrated process in terms of basic Turing machine operations would consist of the following pair of statements:

1. If cell A holds "1" and cell B holds "1",  
or if cell A holds "0" and cell B holds "0",  
then print "1" on cell C.
2. If cell A holds "1" and cell B holds "0",  
or if cell A holds "0" and cell B holds "1",  
then print "0" on cell C.

A notation which simplifies this description is one which employs superscripts to denote the symbol held in a given cell:

$A^0$  will mean "there is a '0' in cell A"  
and  $A^1$  will mean "there is a '1' in cell A"

Then we might agree that

$A^0$ : will mean "if there is a '0' in cell A"  
and  $:C^1$  will mean "print a '1' in cell C"

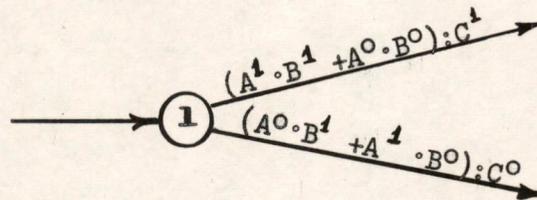
With these abbreviated forms, the statements describing the action of the illustrated machine become:

$$\left. \begin{array}{l} (A^1 \text{ and } B^1 \text{ or } A^0 \text{ and } B^0):C^1 \\ (A^0 \text{ and } B^1 \text{ or } A^1 \text{ and } B^0):C^0 \end{array} \right\}$$

The remaining simplification involves replacing "and" and "or" with shorter symbols. We will choose "." to replace "and" and "+" to replace "or." With these changes the statements become:

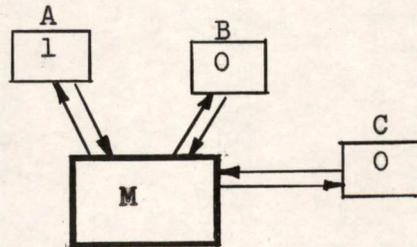
$$\left. \begin{array}{l} (A^1 \cdot B^1 + A^0 \cdot B^0):C^1 \\ (A^0 \cdot B^1 + A^1 \cdot B^0):C^0 \end{array} \right\}$$

and the transition diagram can be redrawn in the form:



which is considerably simpler and more compact than the original diagram.

It is interesting to note that in this new description the operations "R" and "L" do not appear. It is no longer a requirement that the cells A, B, and C be adjacent cells in a linear array. In fact, the new transition diagram equally well describes the action of a discrete-state machine which deals with several independent cells simultaneously:



We will return to this idea later.

#### Table of Combinations

The possible outcomes and the corresponding conditions of A and B can be represented conveniently in a table. On the left are listed all combinations of symbols found in A and B, and on the right the resulting symbol in C:

A	B	C
0	0	1
0	1	0
1	0	0
1	1	1

Thus, the first line of the table corresponds to  $A^0 \cdot B^0 : C^1$ , etc. The number of lines,  $k$ , in such a table is given by  $k=2^n$ , where  $n$  is the number of cells determining the symbol to be printed. The number of different tables is  $2^k$ , i.e.,

$2^{2^n}$  = number of functionally different machines  
which print "1" or "0" depending on the  
symbols held in  $n$  cells.

It should be noted that directions for printing only the 1's are sufficient to determine the complete table. Thus, it is sufficient to describe the illustrated machine by:

$$(A^1 \cdot B^1 + A^0 \cdot B^0) : C^1$$

from which the table is written:

A	B	C		A	B	C
0	0	1	→	0	0	1
0	1			0	1	0
1	0			1	0	0
1	1	1		1	1	1

0's appearing in all unfilled positions. Similarly, directions for the printing of 0's are also sufficient and may result in simpler descriptions in some cases. For example,

$$A^0 \cdot B^0 : C^0 \quad \text{and} \quad (A^1 \cdot B^0 + A^0 \cdot B^1 + A^1 \cdot B^1) : C^1$$

both describe the same machine.

We will say that two cells, C and D, are equivalent if C holds a "1" whenever D holds a "1" and C holds "0" whenever D holds "0". Thus, from a table, e.g.,

A	B	C	D	E
0	0	1	1	0
0	1	0	0	0
1	0	0	0	1
1	1	1	1	1

it is seen that C is equivalent to D and E is equivalent to A. These will be written  $C=D$  and  $E=A$ , respectively.

Thus, given

$$x^1:f^1 \quad \begin{array}{c|c} x & f \\ \hline 0 & 0 \\ \hline 1 & 1 \end{array}$$

it is seen from the table that  $x=f$ . This could have been obtained from the statement  $x^1:f^1$  by dropping superscripts and replacing ":" with "=", and the converse is of course also true. Using this rule, we would also obtain  $x^0=f^0$  from  $x^0:f^0$ .

From  $x^0:f^1$  we would obtain  $f=x^0$ . The table is:

$$\begin{array}{c|c} x & f \\ \hline 0 & 1 \\ \hline 1 & 0 \end{array}$$

Inspection of the table shows that the symbol in cell  $f$  is the complement of the symbol in cell  $x$ , i.e.,  $f=\text{complement of } x$ . Thus,  $x^0$  will be read "x complement" or "complement of x", the superscript "0" indicating the complementation.

From

$$\begin{array}{c} x^0:f^1 \\ f^0:g^1 \end{array} \quad \begin{array}{c|c|c} x & f & g \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}$$

we conclude that  $x = (x^0)^0$ , the double-complement rule.

Now we define two cells 1 and 0 in the following way: Cell 1 always holds the symbol "1" (see, for example, cell  $E_1$  of the UM described earlier) and cell 0 always holds the symbol "0". Evidently:

$$1^0 = 0 \quad \text{and} \quad 0^0 = 1$$

Consider the following printing statements for cells  $f_1$  through  $f_6$ :

$$\begin{array}{l} (1^1+1^1):f_1^1 \\ (1^1+0^1):f_2^1 \\ (0^1+0^1):f_3^1 \\ 1^1 \cdot 1^1:f_4^1 \\ 1^1 \cdot 0^1:f_5^1 \\ 0^1 \cdot 0^1:f_6^1 \end{array}$$

The table of combinations is then

1	0	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$
1	0	1	1	0	1	0	0
1	0	1	1	0	1	0	0
1	0	1	1	0	1	0	0
1	0	1	1	0	1	0	0

from which we conclude that

$$\begin{array}{ll} 1+1=1 & 1 \cdot 1=1 \\ 1+0=1 & 1 \cdot 0=0 \\ 0+0=0 & 0 \cdot 0=0 \end{array}$$

These results illustrate properties of an arithmetic which is like ordinary arithmetic for the "dot" operation (multiplication) but unlike ordinary arithmetic for the "plus" operation (addition). We have described the fundamental operations of Boolean Algebra.

Evidently these operations are commutative, i.e., "x or y" is the same as "y or x" and "x and y" is the same as "y and x". Symbolically

$$x + y = y + x \quad \text{and} \quad x \cdot y = y \cdot x$$

The "and" and "or" operations are also associative. Using the parenthesis to denote grouping, the following statements hold:

$$x + (y + z) = (x + y) + z \quad \text{and} \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

Consequently, the order and grouping of symbols in any Boolean expression is arbitrary. The distributive properties of the "and" and "or" operations can be established by constructing the table of combinations for the forms:

$$\begin{array}{ll} (x^1 \cdot y^1 + x^1 \cdot z^1):f_1^1 & (x^1 + y^1) \cdot (x^1 + z^1):f_3^1 \\ x^1 \cdot (y^1 + z^1):f_2^1 & (x^1 + y^1) \cdot z^1:f_4^1 \end{array}$$

x	y	z	$f_1$	$f_2$	$f_3$	$f_4$
0	0	0	0	0	0	0
0	0	1	0	0	0	0
0	1	0	0	0	0	0
0	1	1	0	0	1	1
1	0	0	0	0	1	1
1	0	1	1	1	1	1
1	1	0	1	1	1	1
1	1	1	1	1	1	1

Comparison of columns in the table shows that  $f_1=f_2$  and  $f_3=f_4$ , i.e.,

$$x \cdot y + x \cdot z = x \cdot (y + z) \quad \text{and} \quad (x + y) \cdot (x + z) = x + y \cdot z$$

As in ordinary algebra, then, one can "multiply" through a "sum" (recall the process of "factoring"). Unlike ordinary algebra, Boolean algebra permits one to "add" through a "product". These forms will occur quite often in subsequent work.