RADIO CORPORATION OF AMERICA
RCA LABORATORIES
David Sarnoff Research Center
PRINCETON, N. J.

November 11, 1959

Harlan E. Anderson
EJCC Publication Committee
Digital Equipment Corporation
Maynard, Massachusetts

Dear Sir:

Enclosed are four copies of the final manuscript of my paper, "Negative Resistance Elements as Digital Computer Components". My biography is also included. The size of the slides I will use in my talk is 3 1/4" x 4".

Very truly yours,

Morton H. Lewin

Morton H. Lewin

MHL:at

Enclosures

File Copy

NEGATIVE-RESISTANCE ELEMENTS AS DIGITAL COMPUTER COMPONENTS

by

Morton H. Lewin
RCA Laboratories

ABSTRACT

The use of two-terminal negative-resistance devices as the basic switching elements in a digital system is discussed. Two fundamental problems analyzed are concerned with:

(1) Achieving logical gain at maximum possible repetition rate.

(2) Providing means to dictate the direction of flow of information in the system.

Circuits performing all the essential logical functions are presented, utilizing the "tunnel (Esaki) diode," a new high-speed negative-resistance semiconductor device, as the basic element. Single-ended and balanced circuit configurations are discussed. In addition, simple arrangements of a small number of tunnel diodes are derived to realize more complicated logical functions.

Part of the system described is a three-phase pulse power supply. Utilizing such a power source, all storage functions can be realized by dynamic storage techniques.

NEGATIVE-RESISTANCE ELEMENTS AS DIGITAL COMPUTER COMPONENTS

by

Morton H. Lewin

RCA Laboratories
Princeton, N. J.

INTRODUCTION

In determining the maximum repetition rate of a given switching circuit, the response of the switching device and the effect of other circuit parameters (including stray elements) must be taken into account. Although the switching speed is ultimately limited by the device, in many cases one never reaches this theoretical maximum because circuit limitations play the dominant role. To solve this problem, one is forced to devise extremely simple circuits with few components in order to minimize the effect of stray reactance. The use of two-terminal negative-resistance elements allows one to do this.

Shockley and Mason[1] have proposed that the ultimate high-speed semiconductor amplifying device is a two-terminal negative-resistance element. They reason that, since the speed of semiconductor components is basically limited by the transit time of carriers, the physical dimensions of devices operating in the highest frequency ranges must be extremely small. In the limit, fabrication problems dictate two-terminal active elements, where only one dimension need be small.

This paper is first concerned with the general problem of using two-terminal negative-resistance devices as the only active switching elements in a digital system. Specific circuits are then discussed, using a particular voltage-controlled negative-resistance device as an example. Much of this treatment can be adapted to other negative-resistance elements.

---

[1] Shockley, W. and Mason, W. P., "Dissected Amplifiers Using Negative Resistance," Journal of Applied Physics, Vol. 25, No. 5, May 1954, p. 677.

GAIN

A combinational switching circuit is defined as a circuit whose outputs depend only on the present inputs. This is to be distinguished from a sequential switching circuit in which the outputs depend not only on the present inputs but also on the past history of inputs. Thus, a combinational circuit, by definition, has no memory.

Consider a system of combinational circuits employing negative-resistance devices as the active switching elements. The requirement of no memory dictates either monostable operation of the negative-resistance elements or bistable operation with a built-in reset to eliminate storage. (The possibility of combinational circuits composed of sequential sub-circuits is ignored on the grounds that such complicated circuits will reduce the maximum speed of the system.)

For the case of monostable operation, if one removes any one of the negative-resistance elements from the circuit, measures the static V, -I characteristic seen looking into the rest of the circuit from its two terminals and then superimposes this on the negative-resistance characteristic, there is always only one stable intersection, for all input combinations. For an all-passive circuit, such as a conventional diode gate, one intersection (operating point) is assured. Assume that this measured characteristic can be approximated by a straight load-line, in the region of interest. (A design to insure monostability is feasible only if this characteristic is "well-behaved" (i.e.,

monotonic) in the region where it intersects the negative-resistance characteristic.) This leads to a simple situation which can be directly analyzed.

Typical voltage-controlled and current-controlled negative-resistance characteristics are shown in Fig. 1a. The two states for each device are most conveniently chosen as operation in the two positive-resistance regions on both sides of the negative-resistance region. Thus, for a voltage-controlled element, the state is defined by the voltage across the device, and for a current-controlled element by the current through it. Under the conditions described above, the circuits to be analyzed become those shown in Fig. 1b. The combination of R and the power source represents the Thevenin equivalent of the linearized measured characteristic.

Monostable operation can be achieved in two ways as indicated by the load-lines in Fig. 1a. In the first case, labeled "I", $R < R_n$ min. for the voltage-controlled element and $R > R_n$ max. for the current-controlled element, where $R_n = |dV/dI|$ in the negative resistance region. In the second case, labeled "II", R does not satisfy this incquality but the power supply values are chosen to result in one intersection. Thus, for case I monostable operation results regardless of the power supply parameters (assuming no reactance), while for case II the power supply values must be chosen to avoid bistable operation.

A DC coupled system with no reactive elements will be assumed. The output terminals and equivalent load resistance

R_L are shown in Fig. 1c. Input signal sources are also included. R_L represents the load furnished by other gate circuits in the net. R_1 represents the contribution to R of the internal parameters of the circuit under consideration. The series or parallel combination of R_1 and R_L, as appropriate, yields R. For the voltage controlled case, R_L can vary from ∞ to some minimum value and for the current-controlled case, from 0 to some maximum value. The fact that R_L varies as indicated is a direct result of the two-terminal nature of all components. For example, an examination of the possible configurations using voltage-controlled elements reveals that, in general, the output current from a stage in a given state depends on the states of the circuits being driven.

The values of R_1 and I_s or V_s must be chosen to assure monostability for all loads. Thus, they must be chosen such that only one intersection (of the type shown in Fig. 1a) occurs for $R_L = \infty$ in the voltage-controlled case and $R_L = 0$ in the current-controlled case. If these conditions are satisfied, monostable operation is assured for all R_L.

Recall that any reactance in the circuit is assumed negligibly small. For case I, looking into the circuit from the output terminals, the load resistor R_L sees a net positive resistance for all voltage-current conditions. Hence, there is no possiblity that an increment of energy delivered to the load will be greater than that supplied by the signal source, for any value of R_L. For case II, assuming a rectangular

signal pulse which raises the load-line sufficiently to cause the operating point to switch to the other positive-resistance region, a simple calculation[2] reveals that the input energy is at least as great as the output energy. Thus, the requirement of monostability, in the absence of adequate reactance, leads to a circuit which has no gain.

If one now allows the use of appropriate reactive elements (i.e., capacitance in parallel with the current-controlled device and inductance in series with the voltage-controlled device), as shown in Fig. 1d, gain can be achieved. Note that the added reactance cannot simply be greater than zero but must be greater than a certain minimum established by stray elements the properties of the negative-resistance device. (For an AC coupled system, the reactive coupling elements must also be taken into account.) In this case the gain arises from the fact that energy stored in the reactive element is delivered to the load when the negative-resistance device is triggered by a small signal. Such circuits have been treated in the literature[3,4,5]. It is shown there that the recovery time associated with the reactance is a factor which limits the maximum repetition rate of the circuit.

Bistable-with-reset operation allows one to achieve gain without the use of reactive elements. Since the furnishing of

---

[2] See Appendix.

[3] Farley, B. G., "Dynamics of Transistor Negative-Resistance Circuits," Proc. IRE, Vol. 40, Nov. 1952, 1497-1508.

[4] Anderson, A. E., "Transistors in Switching Circuits," Proc. IRE, Vol. 40 Nov. 1952, 1541-1558.

[5] Lo, A. W., et al, "Transistor Electronics" Prentice Hall 1955.

a reset signal may be considered to be an additional function of the power supply, effectively a time-varying power source is now being considered. One possible arrangement is to let the power supply (current or voltage) deliver a continuous train of rectangular pulses, such that during each pulse (excitation) the negative-resistance device can go to either one of its two states, depending on input conditions. The "reset" is then the termination of the excitation pulse. It can be seen that such a power supply also serves as a master clock. If one now calculates[6] the transition and recovery times for such a system and compares this to the system with DC power supplies and reactive elements, it is evident that the former scheme has the higher maximum repetition rate.

DIRECTIONALITY

Another fundamental problem is concerned with making the system unilateral. For example, since the negative-resistance element is a two-terminal device, when one terminal is grounded, the other must act as both the input terminal and the output terminal. One must therefore provide some means to dictate the direction of flow of information in the system (i.e., to make a circuit directional, so that a signal propagates from input to output). Some possible techniques for achieving directionality include use of passive elements such as Hall effect couplers or gyrators, use of non-linear interstage coupling elements such as conventional diodes, synthesis of

---

[6] See Appendix.

three-terminal circuit configurations with some unilateral properties, and separation of input and output functions in time using a time-varying power supply. Some of these techniques, as applied to circuits involving voltage-controlled elements, are discussed in more detail following the treatment of basic logic circuits.

POWER SUPPLY

Assuming the bistable-with-reset mode of operation, with the momentary removal of power supply excitation as the method of resetting, the waveform shown in Fig. 2a represents an acceptable source waveform. The sequence of operations performed by each stage is then as follows: After having been reset, a given circuit is energized to an initial state. If the combination of inputs presented to it is favorable, it will switch to its other state. The state of the circuit is then detected by the next stages. Finally, the circuit is reset, energized again and ready to receive a new combination of inputs.

If the entire system is powered from the same source, all circuits are reset simultaneously. The energizing pulses must then be wide enough to allow signals to propagate from the inputs of the system to its outputs, so that the repetition rate is limited by the longest signal propagation time expected. To increase the repetition rate, the system is broken up into small groups of gates such that each group is reset immediately after it has performed its function. A sequence of resets is

then required in order that information will continue to propagate and will not be erased. These requirements can be satisfied by a multiphase power supply such as, for example, the three-phase waveform shown in Fig. 2b. Using this method, a given gate or group of gates is powered by one phase, drives other circuits powered by the next phase and is driven by still other circuits powered by the previous phase. The excitation pulses overlap in time such that information propagates between two stages during the period when both are energized simultaneously. Considering the block B in Fig. 2b, one can see that the beginning of its supply pulse, T_1, corresponds to an "input" region and the end of the pulse, T_2, to an "output" region. The three-phase arrangement shown is characterized by the fact that there is always a group of circuits in the deenergized condition at any given moment. As a result, in many cases spurious signals are prevented from propagating. The similarity between this scheme and the multiphase clock systems used in conventional machines is only superficial. Here the clock source is also the power supply.

GENERALIZED ANALYSIS

I. Load-curves:

Consider a two-terminal "black-box" A whose static V,I characteristic is given by either $I_A = g_1 (V_A)$ or $V_A = f_1 (I_A)$. (No space) ~~(I_A). The box may simply hold a single negative-resistance~~ The box may simply hold a single negative-resistance element or may include a more complicated arrangement of

elements whose composite two-terminal V,I characteristic is given by the above equations. f_1 (or g_1) may be any continuous function and is <u>not</u> single-valued in both V and I if there are any negative-resistance regions. Now consider a second two-terminal "black-box" B whose static V,I characteristic is given by either $I_B = g_2(V_B)$ or $V_B = f_2(I_B)$. This will correspond to the device which determines the load-curve. If $g_2(V_B) = V_B/R$ (i.e., $f_2(I_B) = I_B R$), then the device is the resistor R, mentioned before, and the load-curve is a straight load-line. In general, however, both f_1 and f_2 are non-linear, negative-resistance characteristics. The two cases of interest are the following configurations:

(a) A constant-voltage source V_s across the series combination of elements A and B.

(b) A constant-current source I_s feeding the parallel combination of elements A and B.

The pertinent equations are:

CASE (a)		CASE (b)
$I_A = I_B$		$V_A = V_B$
$V_S = V_A + V_B$		$I_S = I_A + I_B$
$V_A = f_1(I_A)$	[1]	$I_A = g_1(V_A)$
$V_A = V_S - f_2(I_A)$	[2]	$I_A = I_S - g_2(V_A)$

The equilibrium points or quiescent operating points for a circuit are determined by the intersection points of the two curves [1] and [2]. In either case, curve [1] is the characteristic of A and curve [2] is the load-curve determined by the

power supply and the characteristic of B. For case (a), the load-curve is the image of B's V,I characteristic reflected through the current axis, translated in the positive voltage direction a distance V_S. For case (b), the load curve is the image of B's V,I characteristic reflected through the voltage axis, translated in the positive current direction a distance I_S. The only stable operating points are those determined by the intersection of two positive-resistance regions.

II. Composite Characteristics:

The determination of the composite V,I characteristic of two or more two-terminal elements, given their individual characteristics, is important in the analysis of negative-resistance circuits. To graphically obtain the composite characteristic from the individual curves, one follows these simple rules:

(1) For two elements in series, each point of the composite curve with coordinates V_1, I_1 is obtained by choosing any I_1 and letting $V_1 = V_{A_1} + V_{B_1}$, where V_{A_1} is the voltage across element A at the current I_1 and V_{B_1} is the voltage across element B at the current I_1. Thus, one adds the voltages across the individual elements at the same current.

(2) For two elements in parallel, each point of the composite curve with coordinates V_1, I_1 is obtained by choosing any V_1 and letting $I_1 = I_{A_1} + I_{B_1}$, where I_{A_1} is the current through element A at the voltage V_1 and I_{B_1} is the current through element B at the voltage V_1. Thus, one adds the current through the individual elements at the same voltage.

TUNNEL DIODE

The remainder of this discussion is concerned with one particular voltage-controlled negative-resistance element. Similar or "dual" treatment can be given to current-controlled devices.

The device to be considered was first reported by Esaki[7] and has since been investigated by others[8]. Since the phenomenon responsible for the unique characteristics of the device is the tunneling phenomenon predicted by Quantum Mechanics, the device has been called the tunnel diode. It holds promise of being an extremely fast element. Units with time constants of a fraction of a millimicrosecond have been fabricated. Preliminary tests verify that the device is capable of very high speed operation.

For the purposes of this analysis, the V,I characteristic of the tunnel diode will be assumed. Descriptions of the physical operation of the device are given by Esaki[7] and Sommers[8].

---

[7] Esaki, L., "New Phenomenon in Narrow Germanium p-n Junctions," Phys. Rev. 109, Jan. 1958, p. 603.

[8] Sommers, H. S. Jr., "Tunnel Diodes as High-Frequency Devices," IRE Proc., July 1959, p. 1201.

Chang, K. K. N., "Low-Noise Tunnel Diode Amplifier," IRE Proc. July 1959, p. 1268.

Chang, Nelson, et.al, "Tunnel Diodes for Low Noise Amplification," IRE WESCON, San Francisco, Aug. 1959.

Aarons, Holonyak, et al, "Germanium and Silicon Tunnel Diodes-Design, Operation and Application," IRE WESCON, San Francisco, Aug. 1959

The static voltage-current (V, I) characteristic for a typical Germanium unit is shown in Fig. 3. Typical values for the critical points are indicated. The inverse slope R_o of each positive resistance region is of the order of a few ohms.

Since the tunnel diode is such a low impedance element, it is not practical to assume that a constant voltage source is available to supply power to many units. In view of the fact that the source impedance of any realizable voltage source will be of the order of that of its load, it is more practical to assume that the power to individual units is supplied from current sources. In cases where a voltage source is desired, an individual auxiliary device for each circuit is necessary to simulate it. (This is demonstrated later.) Therefore, in line with previous discussions, a three-phase square-wave current source as shown in Fig. 2b will be assumed.

THRESHOLD GATE

Consider the circuit shown in Fig. 4. Assume the input terminals are connected to output terminals of other similar circuits. As long as the tunnel diode D is in the 0 (low voltage) state, the current into D, in addition to I_S, is approximately $M(V_1-V_0)/R$, where M is the number of driver units which are in the 1 (high voltage) state and I_s, V_1 and V_o are defined in Fig. 3. D will switch to the 1 state only if $I_S + M(V_1-V_0)/R > I_o$, the high threshold current at which the resistance becomes negative. Once it has switched to the 1 state, it will remain there even though the current into D is substantially less than I_S (corresponding to loading),

as is evident from an examination of the characteristic. Thus, the circuit is capable of logical gain, since it can now furnish a number of output current increments $(V_1-V_0)/R$ to the next stages. The output of the threshold gate, then, is 1 only if the total number of 1 inputs is greater than or equal to some integer T. I_S is adjusted to result in the correct logical function (i.e., the correct T). For an OR gate, T is one; for an AND gate, T equals the number of inputs; to generate the CARRY output in a full adder, for example, the number of inputs is three and T equals two, etc. The circuit must be reset back to the operating point below the threshold in order to be able to perform its function again.

It is evident that the merit of this circuit depends primarily on the uniformity of diode characteristics and the power supply tolerances involved. The maximum variations in I_0, I_S, V_1 and V_0 dictate the minimum current increments for reliable switching. Advances in fabrication techniques have already resulted in high yields of diodes matched well enough that a reliable logic system involving such circuits appears readily realizable.

The operation of the "single-ended" threshold gate, described above, relies on the accurate determination of the operating point on the negative-resistance characteristic. A balanced or symmetrical circuit offers advantages in many applications. Consider the series combination of two tunnel diodes. Their composite characteristic is shown by the solid curve in Fig. 5a. If a voltage V_1 is applied across the series combination, it

is possible for the circuit to exist in either of two states (i.e., one diode in the high voltage state and the other in the low-voltage state and vice versa). This is depicted in Fig. 5b where D_2 and V_1 determine the load-curve across the characteristic of D_1. If the voltage V_1 is applied as a pulse as is done in the proposed system, one can determine to which state the circuit goes by a small signal at the junction of D_1 and D_2[9]. This can be explained by noting that during the rise of the pulse, the current through D_1 and D_2 builds up to the point where both are very near the crest of the hill. The small current into the junction is sufficient to determine which diode breaks down. Thus if this current is positive, D_1 goes to the 1 state and if it is negative D_2 goes to the 1 state and D_1 is forced to the 0 state.

The difficulty in obtaining a constant-voltage pulse source to drive a large number of such low impedance circuits has already been mentioned. However, the tunnel diode has another important property in that it can simulate a low impedance voltage source, of magnitude V_1, if the current through it is greater than I_o, the high threshold current. This property is utilized to arrive at the final form of the balanced circuit, shown in Fig. 5c. As is shown in Fig. 5a, the dotted load-curve formed by D_3 and I_s intersects the characteristic of the series combination of D_1 and D_2 at the appropriate point, if I_s is large enough. The circuit is now powered by the more realizable current source.

[9] This scheme was suggested by A. Lo.

The logical functions OR, AND and THRESHOLD are achieved by requiring that the current into the junction be positive only when at least one, all or some of the inputs are 1's, depending on the function desired. This requires a reference current or bias as shown. The source of this reference current can be another tunnel diode again acting as a voltage reference.

From the above description one can see that the balanced circuit has several advantages over the single-ended scheme. First, the sensitivity of the circuit depends only on the matching of the two negative-resistance elements and not on the exact values of the critical points of the characteristic. Second, the sensitivity is virtually independent of reasonable power supply variations.

INVERTER

The composite characteristic of a tunnel diode D_1 in series with a resistance R_1 is shown by the solid curve in Fig 6a. R_1 is chosen to be approximately R_N, the magnitude of the linear approximation to the negative-resistance (see Fig. 3). Suppose points "a" and "b" were the only stable points for the circuit (corresponding approximately to a voltage V_1 applied across the series combination). Taking the voltage across the resistor as the output voltage, we have that point "a" yields a 1 output (high voltage; high current through the resistor) and "b" yields a 0 output (low voltage; low current through the resistor). Thus, if the circuit is always at "a" with an 0 input and at "b" with a 1 input, it would realize the inversion function.

To make "a" and "b" the only stable operating points, one can again use another tunnel diode D_2 to simulate a voltage source. The inverter circuit is then as shown in Fig. 6b, and the intersections of the dashed load-curve determined by D_2 and I_s with the composite characteristic of D_1 and R_1 in series are shown in Fig. 6a. To clarify the operation further, one can plot the composite V,I characteristic of the entire configuration of D_1, D_2 and R_1. It is shown in Fig. 6d. The horizontal (constant-current) load-line formed by I_s is indicated.

Since the voltage at point x (Fig. 6b) is high for both operating points, one must include some provision for adding a constant to the normal output voltage levels of the driver tunnel diode, in order that 1 driver output can furnish the current necessary to bring the inverter over the "a" hill to "b". This can be accomplished for a single-ended gate by the addition of a resistor R_2 to the circuit, as shown in Fig. 6c. The voltage of terminal T (during excitation) is greater than the normal output voltage by a constant amount $I_s' R_2$, assuming negligible loading at T. By adjusting R_2 so that the 0 output voltage is approximately V_1, the 1 output voltage is then approximately $2V_1$, and the required operation can be achieved.

Assuming the pulse excitation scheme described before, the operation of the inverter is now clear. Whenever the circuit is excited and the input is a 0, the inverter moves to point "a",

stays there, and the output is a 1. If the input is a 1, there is sufficient current input to bring the inverter over the hill in the characteristic to point "b" and the output is a 0. Note that for this latter case the output waveform will show a transient high voltage before reaching the low voltage 0 output. The next stage must therefore be powered by the next phase so that it is only interested in the voltage level at the end of the pulse (i.e., the "output function" region). For that case such operation is satisfactory.

Note that the height of the "a" hill in Fig. 6d depends on the value of R_1. By adjusting I_s to lie sufficiently below the crest of this hill and driving the circuit from a number of "elevated" outputs, one can obtain the logical function of a threshold gate whose output is inverted (i.e., NOT, OR-NOT, AND-NOT, etc.).

UNILATERALIZATION

Unilateral operation can be defined as operation in which signals can propagate in one direction only. This is required to insure that spurious signals are not generated in the system. The most obvious way to insure unidirectional operation is to use normal diode rectifiers as coupling elements. Current between stages can then flow only in one direction. Other methods are possible in which the coupling between stages is resistive. For example, considering the inverter circuit driven by an elevated output, one can see that when the input to the inverter is 0, there is essentially no current in the coupling resistor. When the input is 1, there is a relatively high current in the coupling resistor. Thus, again, the current in the coupling resistor flows only in one direction. By reversing

the positions of R_1 and D_1 (Fig. 6), so that the output is taken across the tunnel diode, one has a threshold gate with this unilateral property. Another unilateralization method is associated with the ability of the power supply to separate input and output functions in time. This scheme is effective for the balanced type of threshold gate. The circuit is receptive to an input signal only during a very short time (i.e., the rise time of the power supply pulse), after which it "locks" into one state or the other.

MULTILEVEL CIRCUITS

Consider the inverter configuration (Fig. 6b) in which I_s is reduced so that a third stable operating point "c" exists. This is indicated by the dotted curves in Fig. 6 (a and d). Note that point "c" yields a 0 output. Assume that R_1 has been reduced sufficiently to make the height of the "a" hill, in Fig. 6d, comparable with the height of the "c" hill. Let the circuit have two inputs, driven from the normal outputs of other tunnel diodes. The circuit operates in the following fashion: If the two inputs are both 0, "c" is a stable point and each time the circuit is excited the output is a 0. If one of the inputs is a 1 while the other is a 0, there is enough current input to make the circuit move over the first hill to point "a" where it is stable and the output is a 1. When both inputs are 1, there is sufficient current input to make the circuit move over both hills to point "b", and the output is again 0. Thus, the output is 1 only when the two inputs are different. This is the EXCLUSIVE-OR (modulo-2 sum) function.

One can also realize the SUM output for a full adder using a slightly different configuration. Consider the circuit shown in Fig. 7a. The operating points can be found by plotting the load-curve, determined by I_s and the characteristic of the series combination of D_2 and D_3 (Fig. 5a), across the characteristic of D_1 in series with R_1 (Fig. 6a). This situation is depicted in Fig. 7b. I_s is chosen so that there are four stable intersections, labeled 0_a, 1_a, 0_b and 1_b. These correspond to 0 and 1 outputs as explained before. The composite V,I characteristic of the whole configuration of three tunnel diodes and the resistor is shown in Fig. 7c. Note the four intersections with the constant-current (horizontal) load-line I_s. They are also labeled appropriately.

There are three inputs corresponding to two binary digits and the CARRY from the previous digit. The circuit operates in the following manner: When all three inputs are 0, each time the circuit is excited it moves to point 0_a and is stable there so that the output is a 0. When one of the inputs is a 1 while the others are 0, there is enough input current that the circuit moves over the first hill (Fig. 7c) to point 1a, where it is stable, and the output is a 1. When two inputs are 1, the circuit moves over the first <u>two</u> hills to point 0_b and the output is again 0. For three 1 inputs, the circuit moves over all <u>three</u> hills to point 1_b and the output is again 1. Thus we have

Number of 1 inputs	SUM output
0	0
1	1
2	0
3	1

This fulfills the SUM function of a full adder. To realize the CARRY function, one simply uses a threshold gate, of the type described before, which has the same three inputs and which gives a 1 output when the number of 1 inputs is two or greater.

STORAGE

Since any negative-resistance element can exist in two stable states with the proper DC load-line, it is possible to use such a device to store information. The term "static storage" can be applied to this situation (DC load-line), because the voltage or current level of the negative-resistance device is fixed when it is storing a particular bit. This type of storage might be used in the memory of a digital computer.

Storage is also necessary in the logic section of a computer. Here another means of storage, known as "dynamic storage," is directly compatible with the three-phase pulse-overlap system. Dynamic circuit techniques are used in the SEAC and DYSEAC computers[10]. The method involves the circulation of information around a closed loop, so that a circulating pulse represents a 1 and no pulse circulating represents a 0. In the original circuits using this technique, the pulse is introduced at one end of a delay line. At the other end it is amplified, reshaped and clocked and is then returned to the delay-line input. The delay-time is adjusted so that the pulse makes one trip around

---

[10] Elbourn, R. D. and Witt, R. P., "Dynamic Circuit Techniques Used in Seac and Dyseac," IRE Trans. on Electronic Computers, March 1953, pp. 2-9.

the loop in one clock period.

Consider the circuit shown in Fig. 8. All blocks under A are powered by phase A, all under B by phase B, etc. The block with the arrow represents a delay gate (one-input OR gate). This takes the place of the delay-line. Because of the phase relationship between the three power sources (see Fig. 2), it is possible to close the loop as shown. The circulation of a 1 or a 0 is thus made possible. The circuit has built-in amplification, reshaping and clocking. The particular circuit shown in Fig. 8 is a basic flip-flop, where S is the "set to 1" input and R is the "reset to 0" input.

The basic flip-flop can be included in more complicated storage circuits. Fig. 9a shows a binary counter and Fig. 10 shows one stage of a shift register. Other circuits involving dynamic storage techniques are possible.

Note that no time need be lost in obtaining an output from a dynamic circuit, even though the information stored is in the form of a circulating pulse. For example, the binary counter of Fig. 9a may be represented by a single block powered by the appropriate phase, as far as the input and output terminals are concerned. This is shown in Fig. 9b where the input comes from a circuit powered by phase B, the binary counter is considered powered by phase C and the output goes to a circuit powered by phase A.

EXPERIMENTAL VERIFICATION

In order to investigate the operation of tunnel diode logic circuits in a small sub-system, one cell of a simple experimental arithmetic unit was constructed and tested. A block diagram of

the cell is shown in Fig. 11. All of the fundamental logic circuits, including dynamic storage, are evident. The cell contains a storage loop, a full adder and auxiliary read-in and read-out gates for shifting right and left, complementing the input from memory, and reading out to memory.

The schematic diagram for the unit is given in Fig. 12. Fig. 13 contains photos of the complete experimental circuit. The unit contains 27 tunnel diodes. Resistive coupling is used throughout. It is powered from a transistorized power supply which delivers a three-phase, 1 mc, 10 volt square-wave. This repetition rate was chosen to most easily demonstrate the fundamental principles involved. The inputs to the system are DC levels simulating the output voltages of the tunnel diode (i.e., 0 = 50 mv, 1 = 450 mv), with the correct internal impedance.

Typical waveshapes, taken across one of the diodes in the storage loop, are shown in Fig. 14. (a) shows a circulating 1, after the loop has been set and (b) shows a circulating 0, after the loop has been reset. One can also make the bit stored in the loop alternate between 0 and 1 as shown in (c). This is accomplished by making $A_r = 1$, so that the storage loop is cleared and the SUM output of the full adder is gated into the loop, by letting $A_h = 1$, so that one of the inputs to the adder becomes the bit presently stored in the loop, and by allowing any one of the other inputs to the adder to equal ~~x~~ 1 (i.e., either C_i or A_c or $t_i = 1$), so that the sum output becomes the complement of the bit presently stored. Thus, each cycle the complement of the bit previously stored is read into the loop and the stored bit alternates are shown.

The peak currents (I_o) of the tunnel diodes used in the experimental cell range from 1.9 to 2.6 ma. The capacity of each diode is of the order of 100μμf. Peak-to-valley current ratios vary between 5 and 8. The currents from the power supply to each of the logic circuits were adjusted for proper operation. Observed switching times (see Fig. 14 d and e) are of the order of 50 mμs.

The experiment demonstrates a number of important facts concerning tunnel diode logic circuits. First, it demonstrates reliable operation of all fundamental logic circuits in a realistic system. These circuits includes OR, AND, THRESHOLD, NOT and EXCLUSIVE-OR. Second, it demonstrates that such circuits can supply logical gain. For example the OR gate in the dynamic storage loop has a fan-in of 3 and a fan-out of 5. The circuit contains two tunnel diode in cascade. Third, it demonstrates agreement between rough estimates of switching time, based on the time constant of the device (capacity times magnitude of average negative resistance), and the actual switching time.

ACKNOWLEDGEMENTS

The author wishes to thank A. W. Lo, G. B. Herzog, J. C. Miller and A. G. Samusenko of RCA Laboratories and Professors E. J. McCluskey, Jr. and W. H. Surber, Jr. of Princeton University for many interesting discussions and helpful suggestions.

This work was supported by the Bureau of Ships under contract with RCA.

APPENDIX

I. DC power supplies, monostable operation, zero reactance, case II:

The voltage controlled case will be considered. Dual treatment can be given to the current-controlled case. Referring to Fig. 1c, in order to insure monostable operation for all R_L, one must choose I_s and R_1 such that only one intersection occurs for $R_L = \infty$. Assuming R_L is very large, a particular limiting case is shown by the solid load-line in Fig. 15. For a given I_s, the negative reciprocal of the slope is $R_{max}^{R_{1\,max}}$ for monostable operation. Then, for any finite R_L, the load-line changes as shown by the dashed line "a". For a square pulse of input current of magnitude Δi and width Δt, we have that the energy input = $\Delta v\, \Delta i\, \Delta t$, where these values are depicted graphically in Fig. 15. The energy output is $\Delta v^2 \Delta t / R_L$. Thus, we must compare $\Delta v / R_L$ to Δi to determine which increment of energy is greater. These are also found graphically in Fig. 15. An examination of the geometrical constructions involved shows that if the conditions stated above are satisfied, $\Delta v / R_L \leqslant \Delta i$ for all R_L. Therefore, energy input $\geqslant$ energy output.

II. DC power supplies with reactance vs. pulse power supply:

Again, the voltage-controlled case will be considered. Comparison is being made between the two circuits shown in Fig. 16. C is the sum of the stray capacity plus the capacity inherent in the negative resistance device. The path of operation

looking into the parallel combination of C and NR is assumed to be approximately the dashed path shown in Fig. 16. This will be true, in the case of circuit (1), if L is sufficiently large. Under these conditions, one can assume that the transition times (paths "a" and "c") are comparable for the two circuits. If L is not large enough, the transition time of circuit (1) can be considerably larger than that of circuit (2), and in addition, the path followed is no longer horizontal[11].

We wish to compare the recovery times (paths "b" and "d") of both circuits. Specifically, let us calculate the time to go from point "1" to point "2" (path "d"). For circuit (2) the rise time is approximately 2.2 R_oC, since the voltage follows a simple exponential with an R_oC time constant. For circuit (1), using Laplace transform techniques, the transform of v is given by

$$V(s) = \frac{E_o R_o}{s\left[R_oLCs^2 + (L + R_oRC)s + R_o + R\right]} \qquad (1)$$

Assuming that

$$\frac{4}{LC\left(\frac{1}{R_oC} - \frac{R}{L}\right)^2} < \sim 0.4, \qquad (2)$$

which is equivalent to assuming that no oscillations occur, one finds that the roots of the characteristic equation are approximately

---

[11] Cunningham, W. J., "Introduction to Non-Linear Analysis," McGraw-Hill, 1958, pp. 106-114.

$$s_1 = -\frac{R}{L} - \frac{1}{L/R_o - RC}$$

$$s_2 = -\frac{1}{R_o C} + \frac{1}{L/R_o - RC} \,. \qquad (3)$$

The approximate solution is therefore

$$V(s) = \frac{A}{S} + \left[\frac{As_2}{s_1 - s_2}\right] \frac{1}{s - s_1} + \left[\frac{As_1}{s_2 - s_1}\right] \frac{1}{s - s_2} \qquad (4)$$

where

$$A = \frac{E_o R_o}{R_o + R} \,. \qquad (5)$$

An examination of this solution shows that the time constant of the dominant exponential is always greater than $R_o C$ for values of L and R consistant with the approximation (2).

Thus circuit (1) has the lower maximum repetition rate.

CAPTIONS FOR FIGURES

Fig. 1. (a) Load-lines for monostable operation
(b) Equivalent circuits
(c) Load and signal source included
(d) Addition of reactive element

Fig. 2. (a) Power supply waveform
(b) Three-phase power source

Fig. 3. Tunnel diode static characteristic
(Numbers indicated for typical Germanium unit.)

Fig. 4. Single-ended threshold gate

Fig. 5. (a) Characteristic of two tunnel diodes in series and load-curve formed by D_3 and I_S
(b) Bistable operation
(c) Balanced threshold gate

Fig. 6. (a) Characteristic of tunnel diode and resistor in series and load-curves formed by D_2 and I_S
(b) Inverter circuit
(c) Provision for obtaining elevated output
(d) Composite characteristic of D_1, D_2 and R_1

Fig. 7. (a) SUM output circuit for full adder
(b) Determination of operating points
(c) Composite characteristic

Fig. 8. Dynamic flip-flop

Fig. 9. Dynamic binary counter

Fig. 10. Shift register stage

Fig. 11. Block diagram of experimental unit

Fig. 12. Schematic diagram of experimental circuit

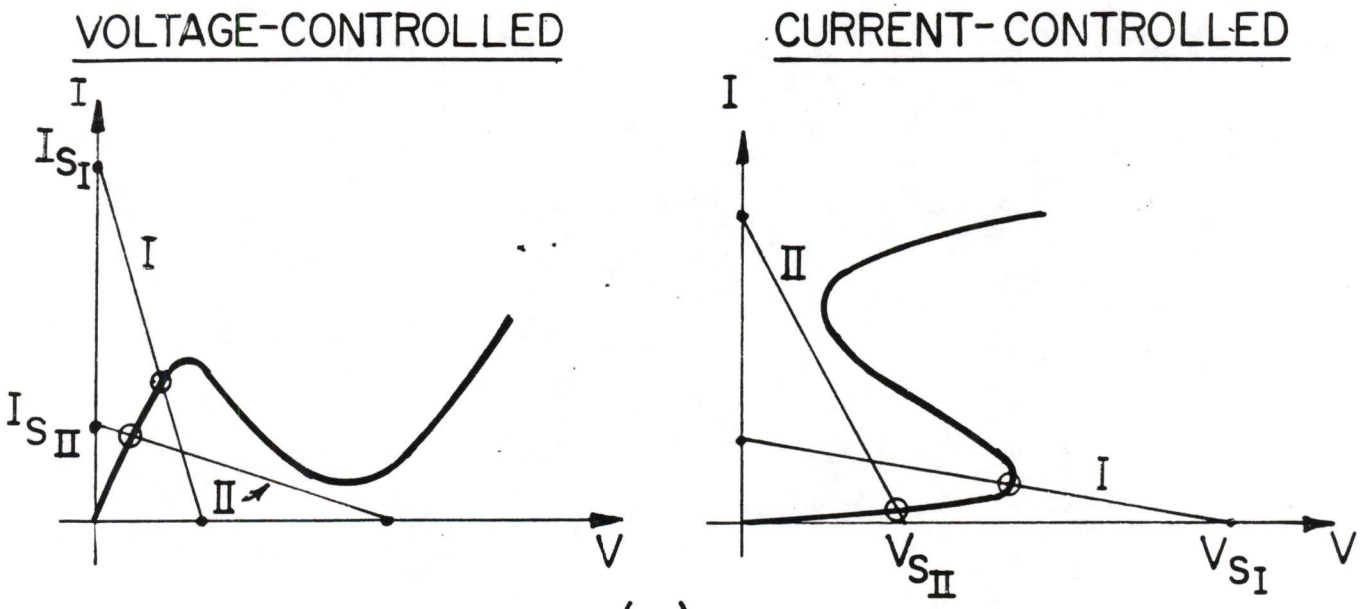
Fig. 13. Front and rear views of experimental cell
(Tunnel diodes are mounted under finger contacts.)

Fig. 14. Typical waveforms
(a) Circulating 1
(b) Circulating 0
(c) Alternating 1 and 0
(d) Rise time
(e) Fall time

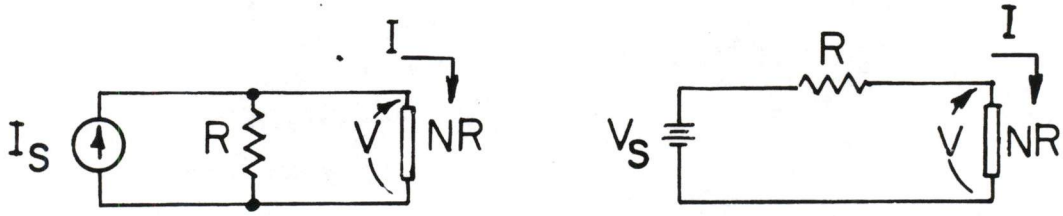
Time goes from left to right. Vertical scales are 0.13 v/div with base-line at bottom. Horizontal scales are 0.5 μs/div for (a), (b) and (c) and 20 mμs/div for (d) and (e).

Fig. 15. Graphical comparison of Δi and $\Delta v/R_L$

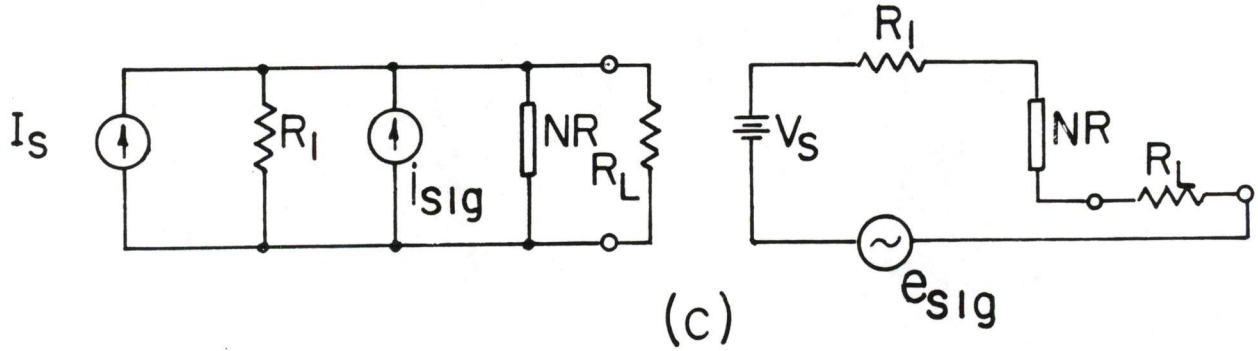
Fig. 16. Comparison of monostable and bistable-with-reset modes. Assumed path of operation.


VOLTAGE-CONTROLLED
CURRENT-CONTROLLED
I
I_{S_I}
$I_{S_{II}}$
I
II
V
I
II
I
$V_{S_{II}}$
V_{S_I}
V

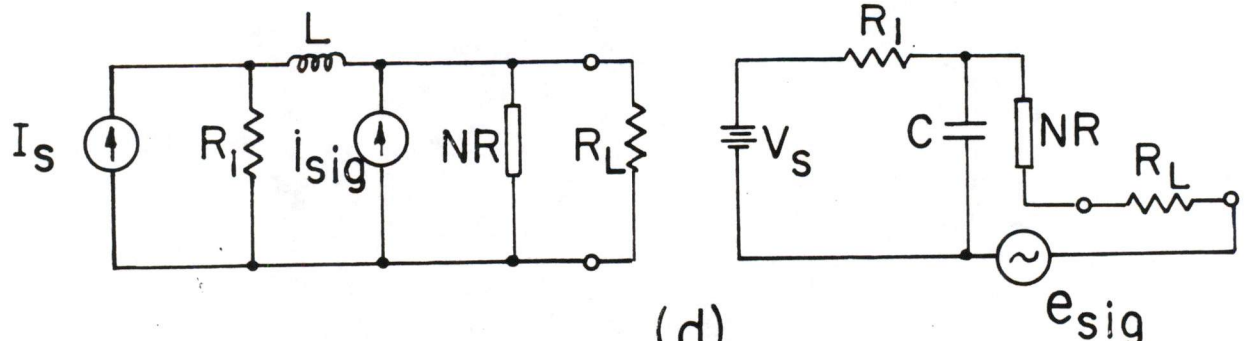
(a)


I_S
R
V
NR
I
V_S
R
V
NR
I

(b)

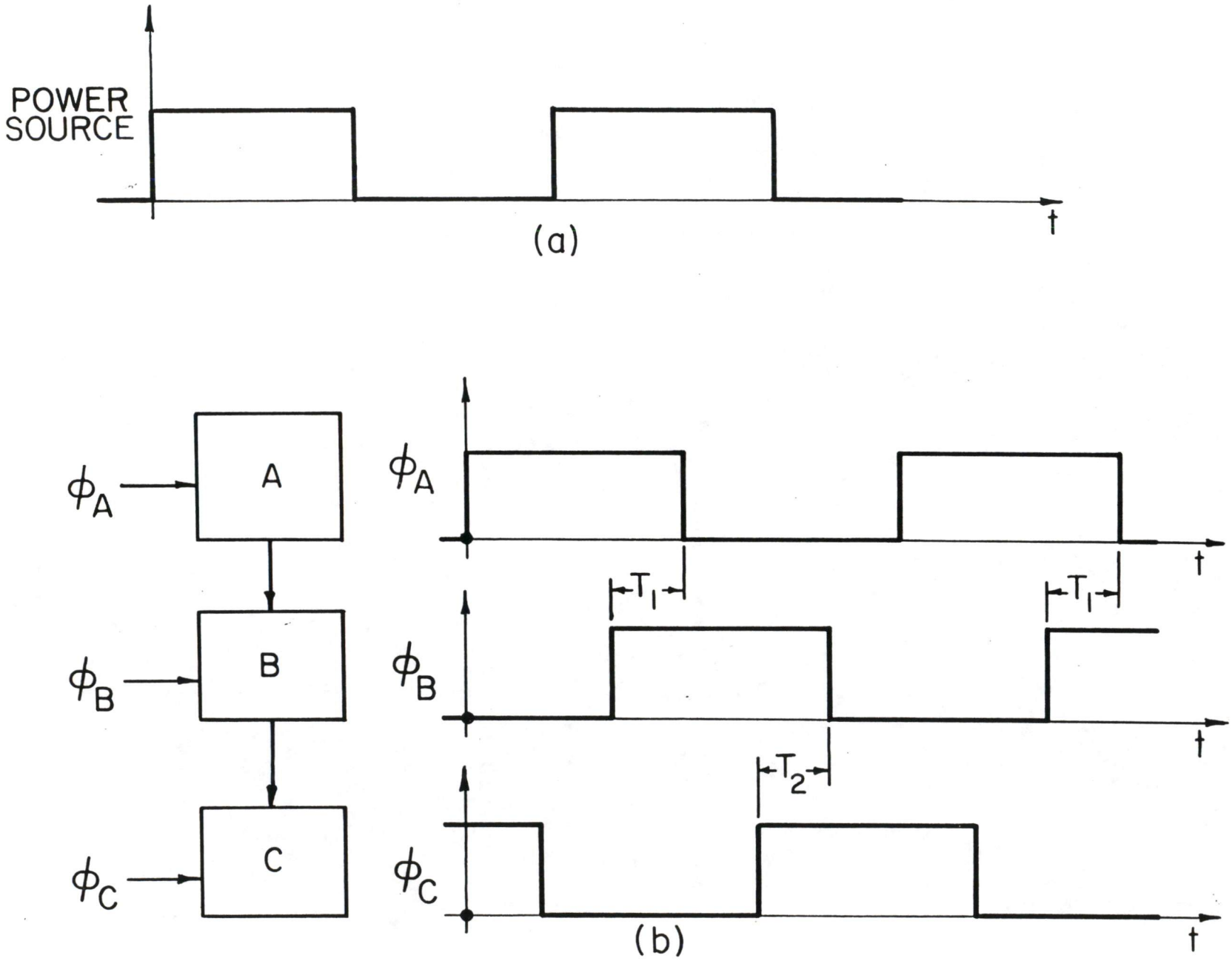

I_S
R_I
i_{sig}
NR
R_L
R_I
V_S
NR
R_L
e_{sig}

(c)


L
I_S
R_I
i_{sig}
NR
R_L
R_I
V_S
C
NR
R_L
e_{sig}

(d)

Fig. 1


POWER
SOURCE
t
(a)
A
B
C
ϕ_A
ϕ_B
ϕ_C
T_1
T_1
T_2
t
t
t
(b)

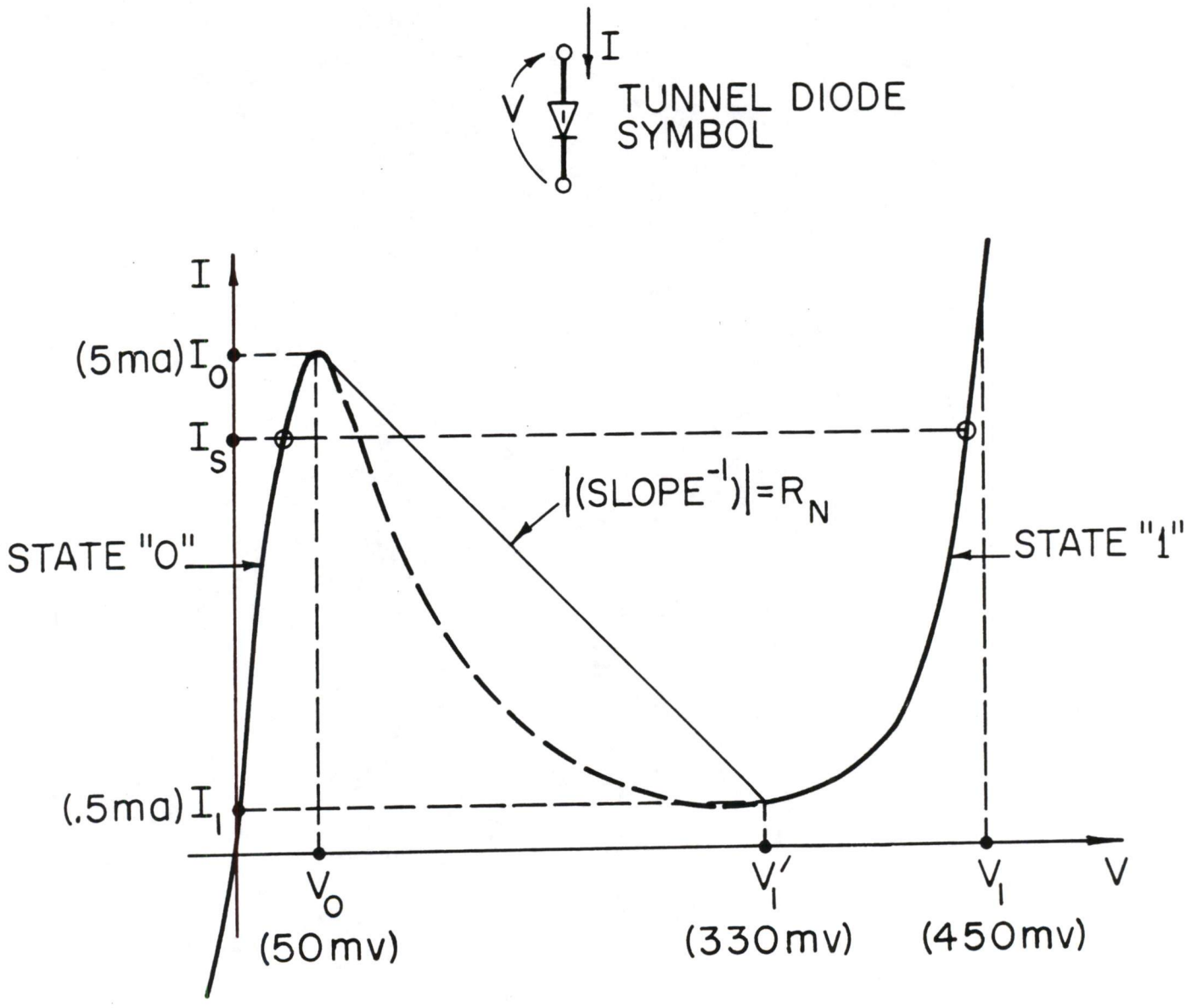

I
V
TUNNEL DIODE
SYMBOL
I
(5ma) I_0
I_s
STATE "0"
$|(\text{SLOPE}^{-1})| = R_N$
STATE "1"
(.5ma) I_1
V_0
(50mv)
V_1'
(330mv)
V_1
(450mv)
V

Fig. 2

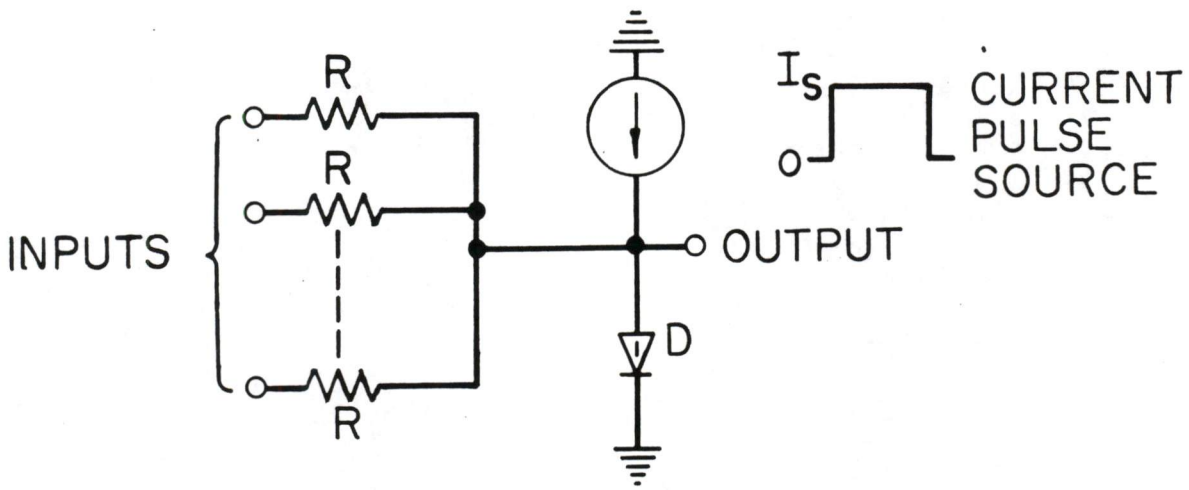
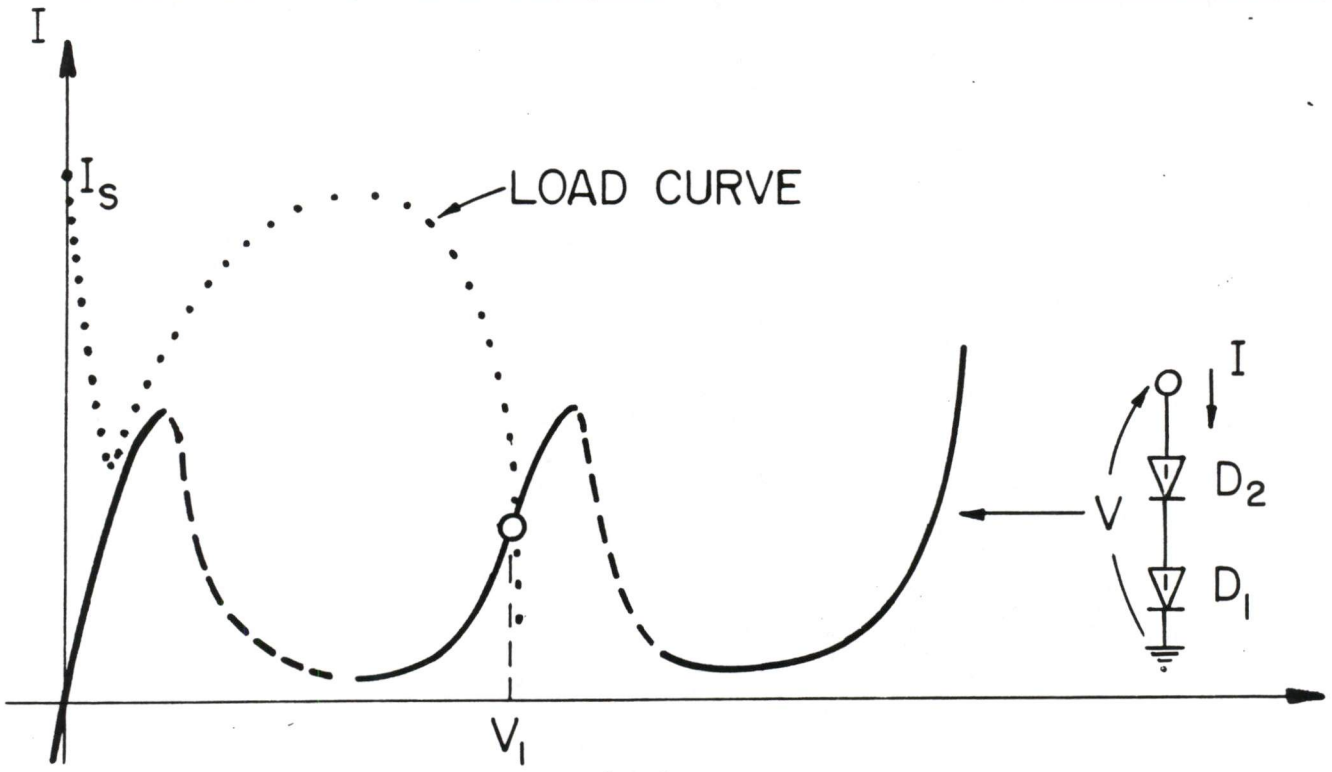
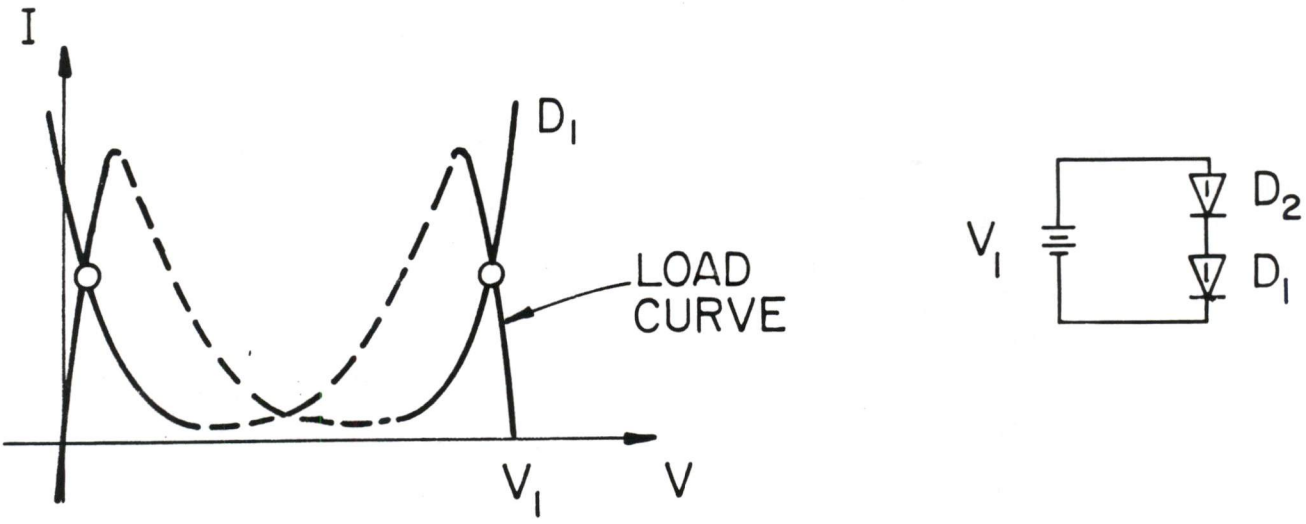


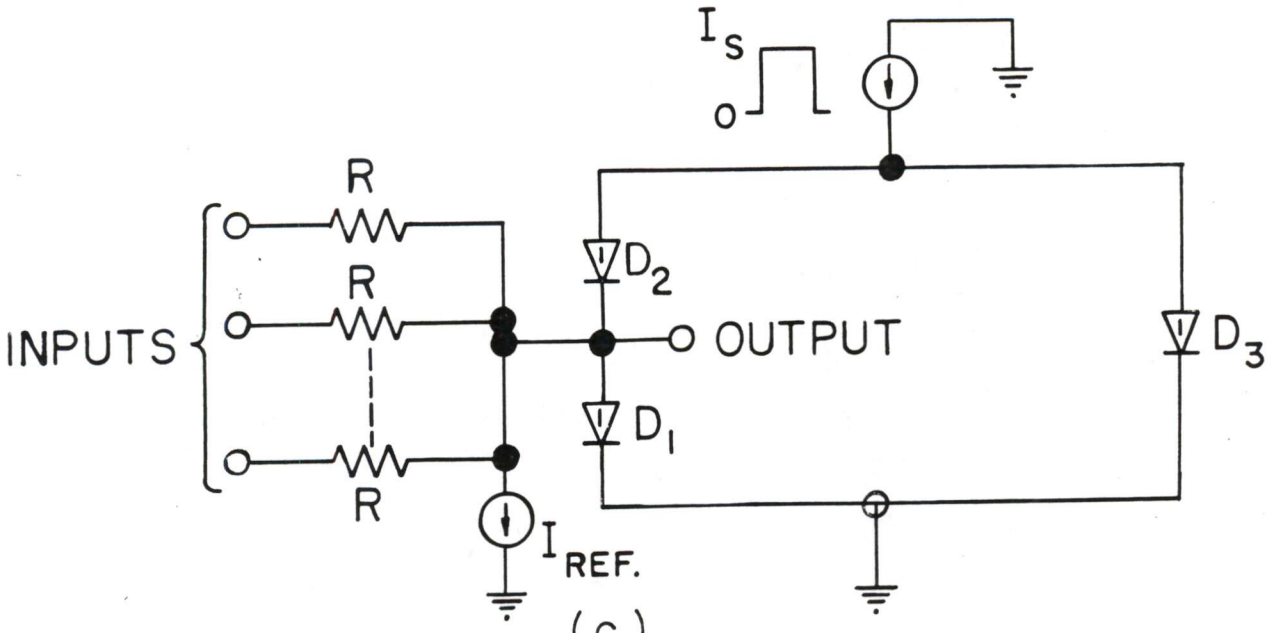
Fig. 4


I
I_S
LOAD CURVE
I
V
D_2
D_1
V_1

(a)

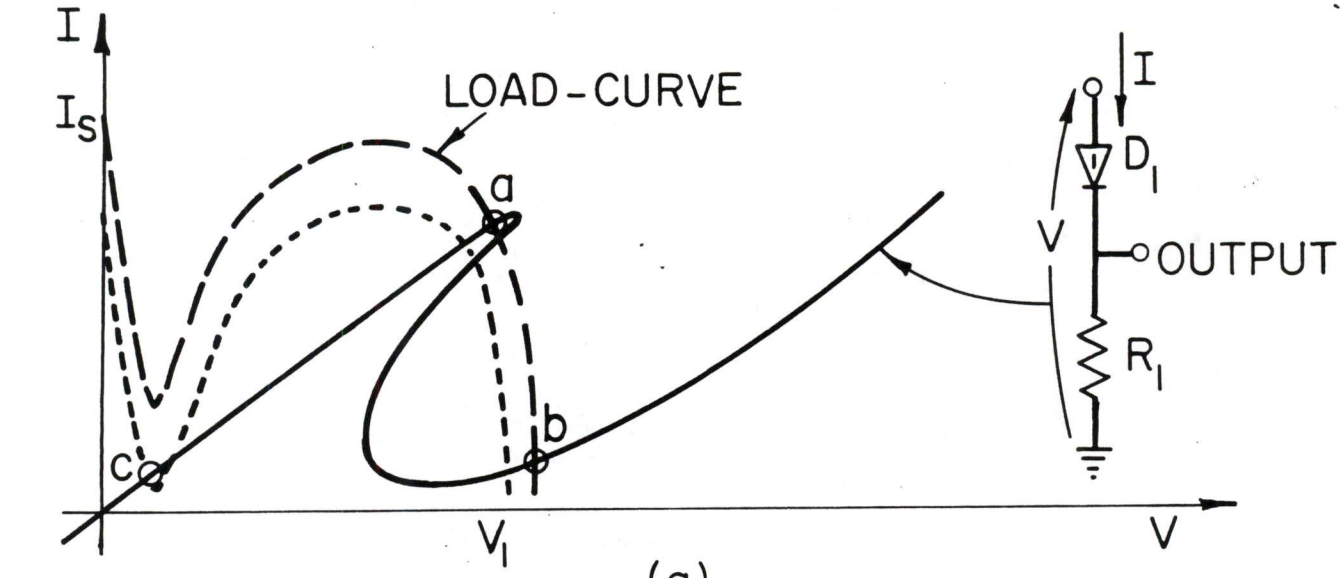

I
D_1
LOAD CURVE
V_1
V
V_1
D_2
D_1

(b)

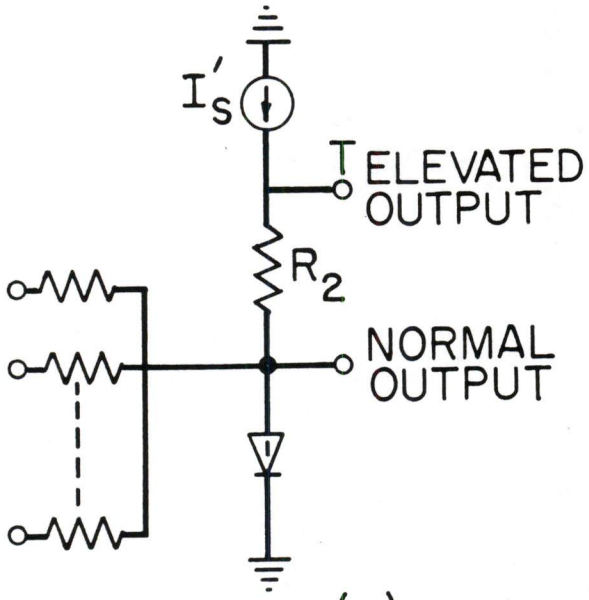

I_S
0
R
R
R
INPUTS
D_2
D_1
OUTPUT
D_3
$I_{REF.}$

(c)

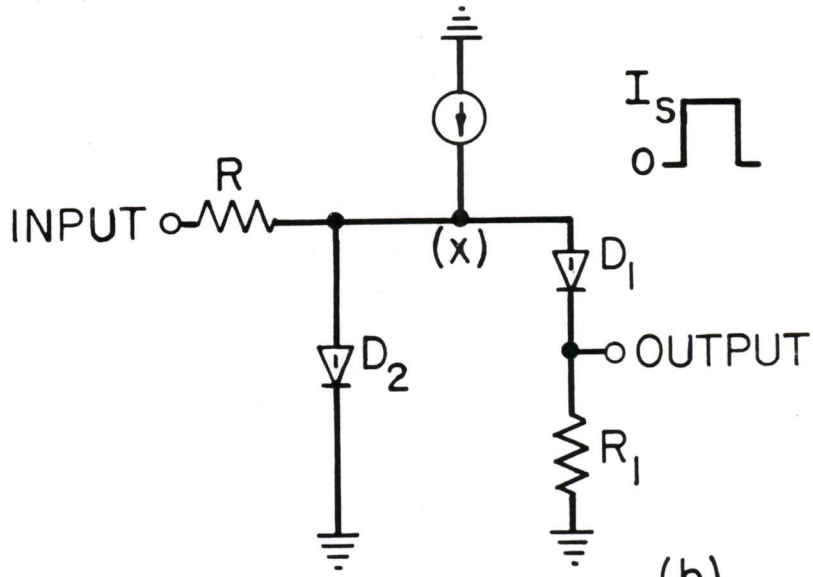
Fig.


I
I_S
LOAD-CURVE
a
b
c
V_1
V
I
D_1
V
OUTPUT
R_1

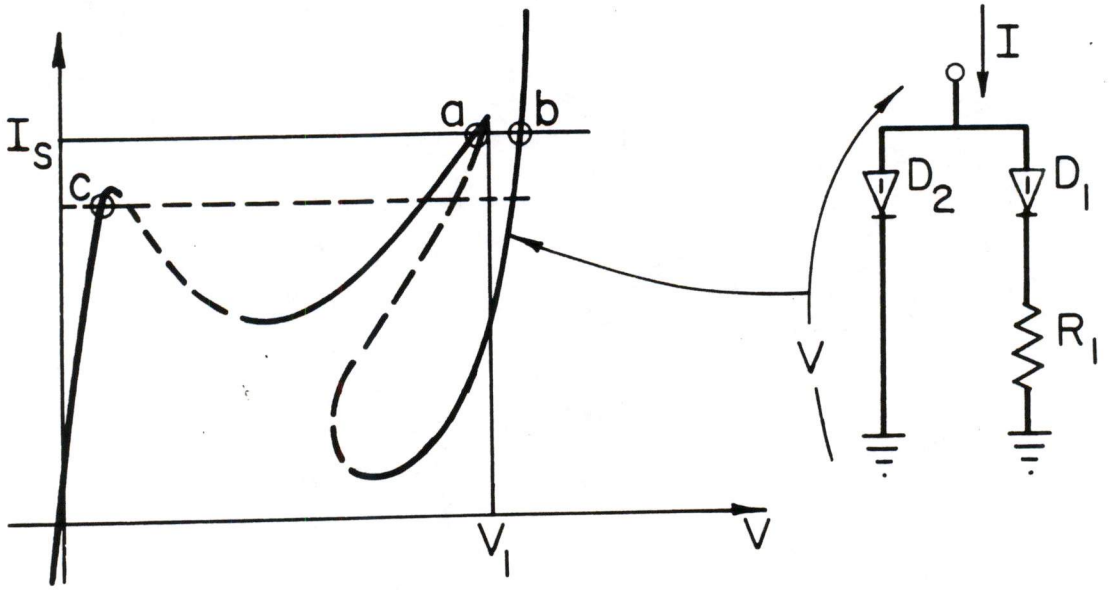
(a)


I_S'
ELEVATED
OUTPUT
R_2
NORMAL
OUTPUT

(c)

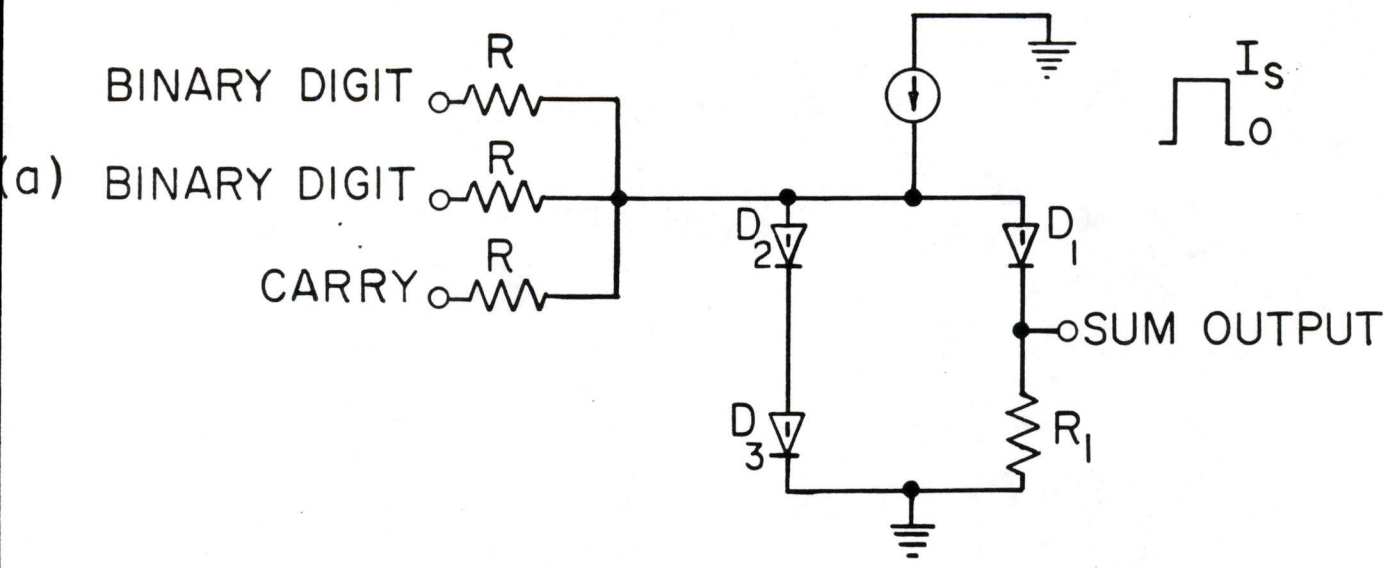

I_S
0
R
INPUT
(x)
D_1
D_2
OUTPUT
R_1

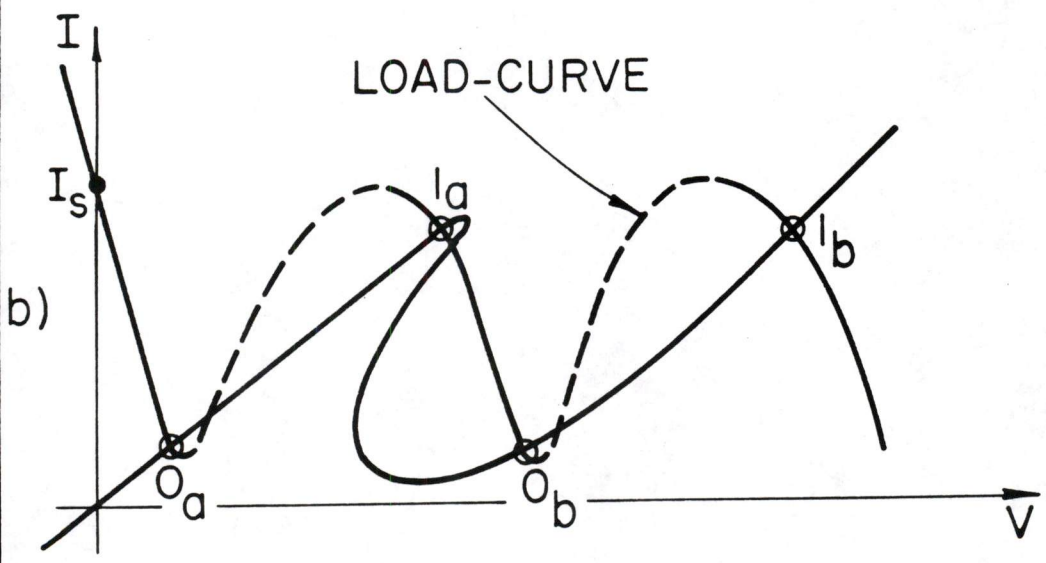
(b)

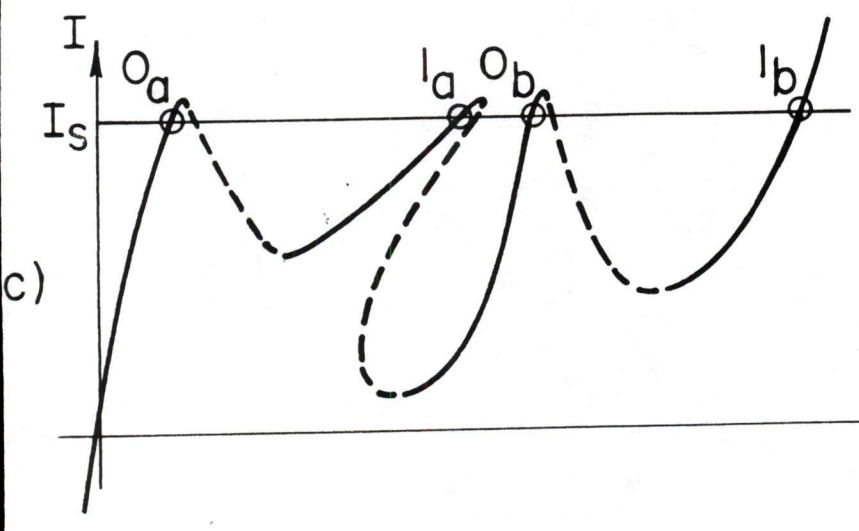

I_S
a
b
c
V_1
V
I
D_2
D_1
V
R_1

(d)

Fig. 1


(a)
BINARY DIGIT
BINARY DIGIT
CARRY
R
R
R
D2
D1
D3
R1
SUM OUTPUT
Is
0


(b)
I
Is
LOAD-CURVE
Ia
Ib
Oa
Ob
V


(c)
I
Is
Oa
Ia
Ob
Ib
V

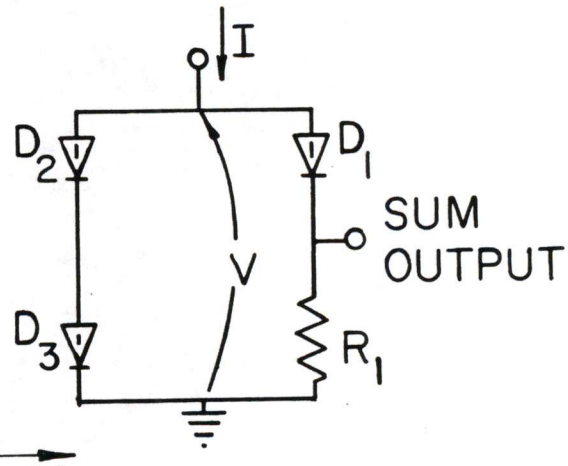

I
D2
D1
D3
R1
V
SUM
OUTPUT

Fig. 7

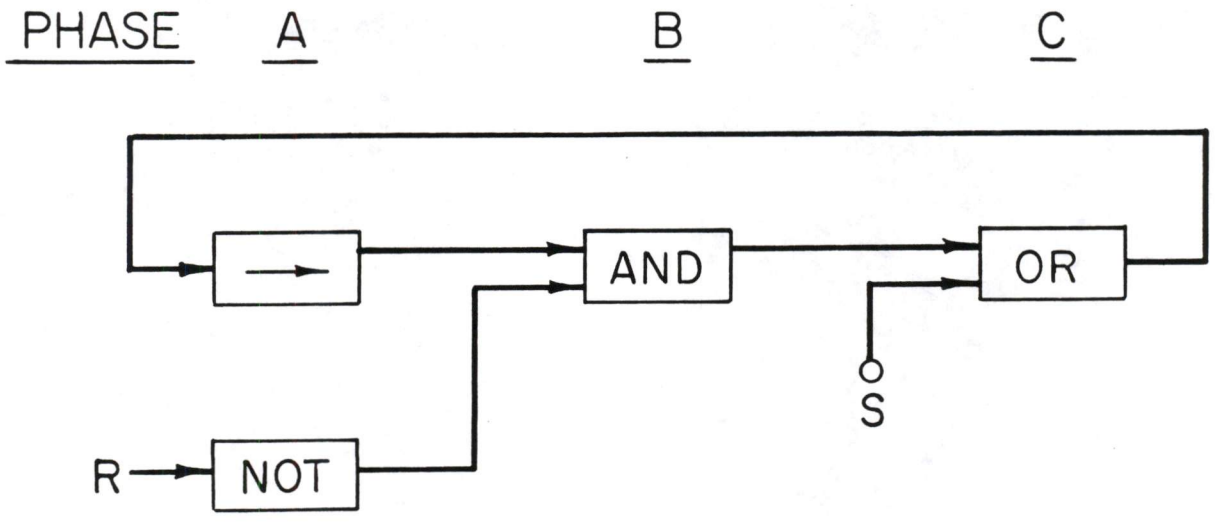
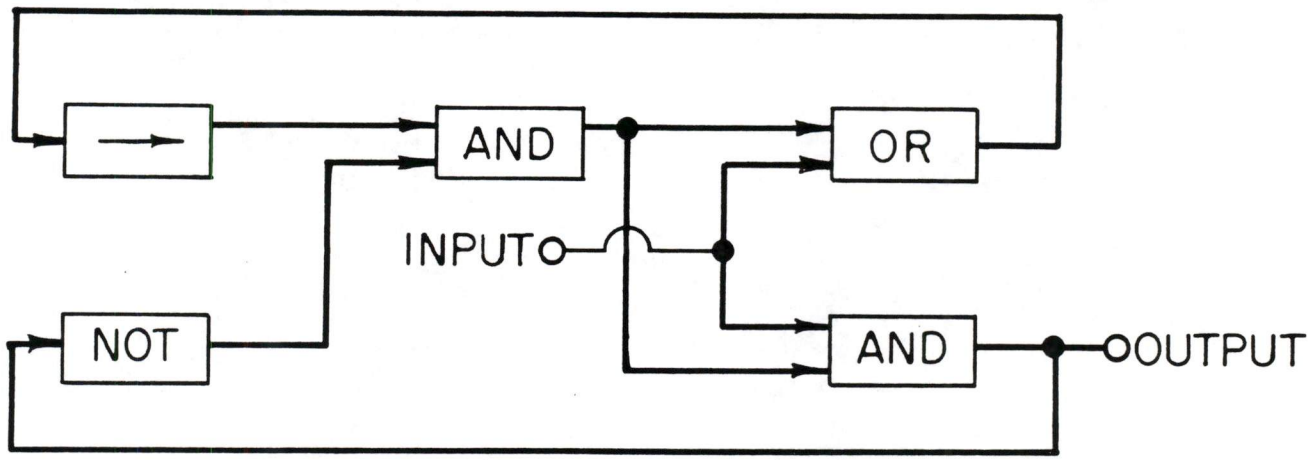


Fig. 8

PHASE: A B C

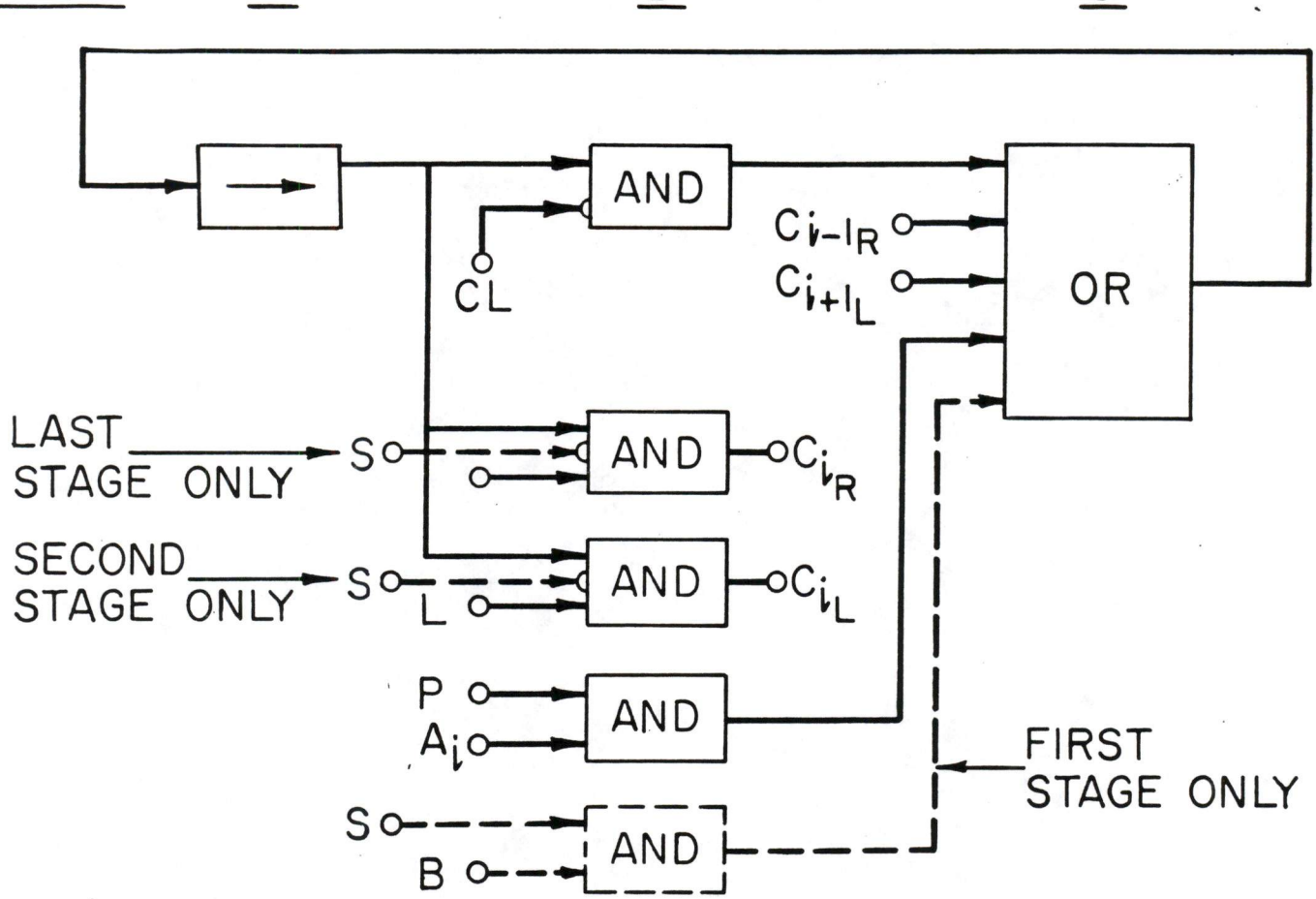


(a)

PHASE: C

(b)

Fig. 9



LEGEND

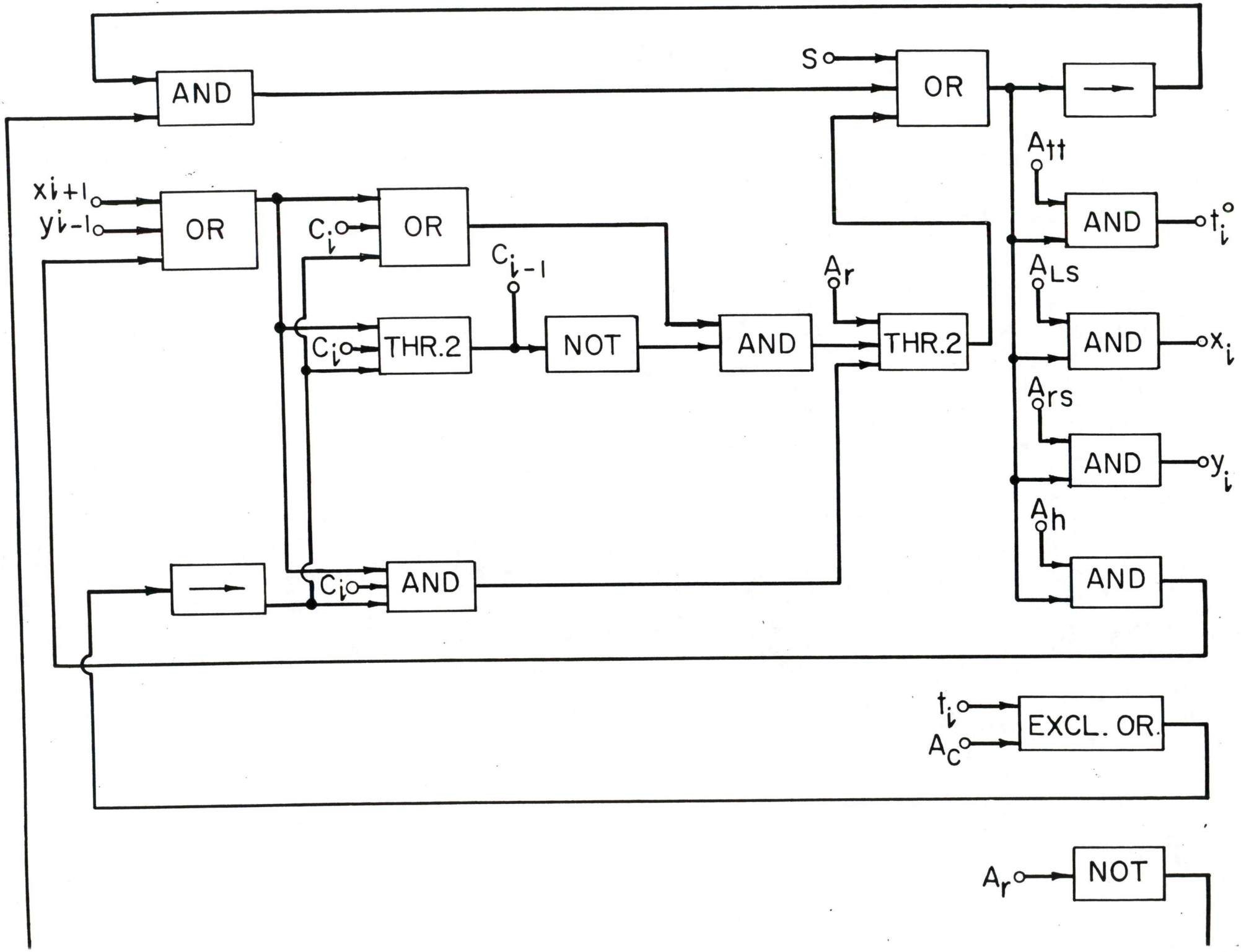
CL = CLEAR
R = SHIFT RIGHT - READ SERIAL (NORMAL ORDER)
L = SHIFT LEFT - READ SERIAL (REVERSE ORDER)
P = WRITE PARALLEL
S = WRITE SERIAL
A_i = PARALLEL INPUT
B = SERIAL INPUT
C_i = TRANSFER OUTPUT
C_{n_R} = SERIAL OUTPUT (NORMAL ORDER)
C_{1_L} = SERIAL OUTPUT (REVERSE ORDER)

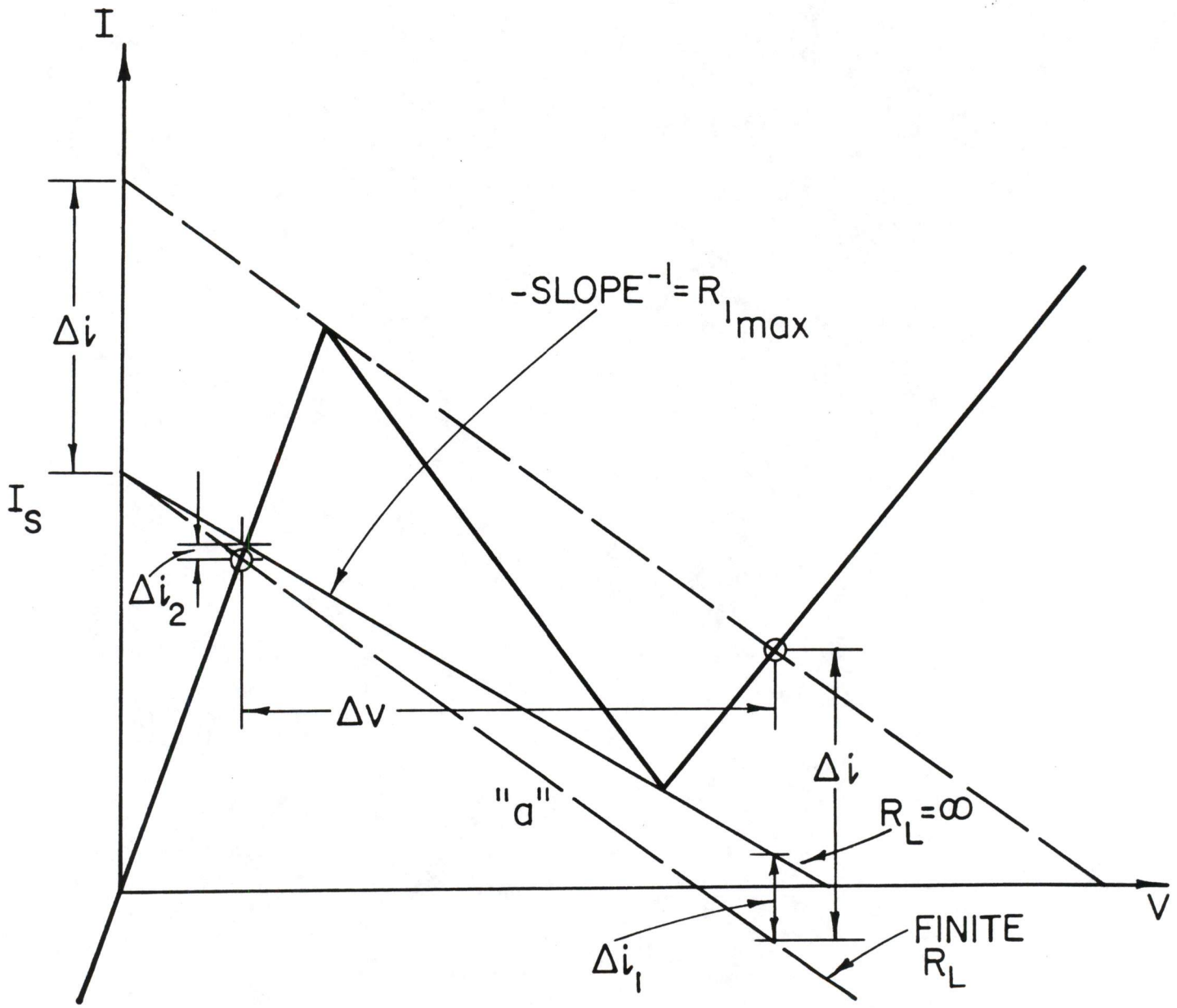
NOTE

$C_{-1} = C_n$
$C_{n+1} = C_1$
⟶| = INHIBIT INPUT

S ALWAYS ACCOMPANIED BY CL AND R OR L
P ALWAYS ACCOMPANIED BY CL

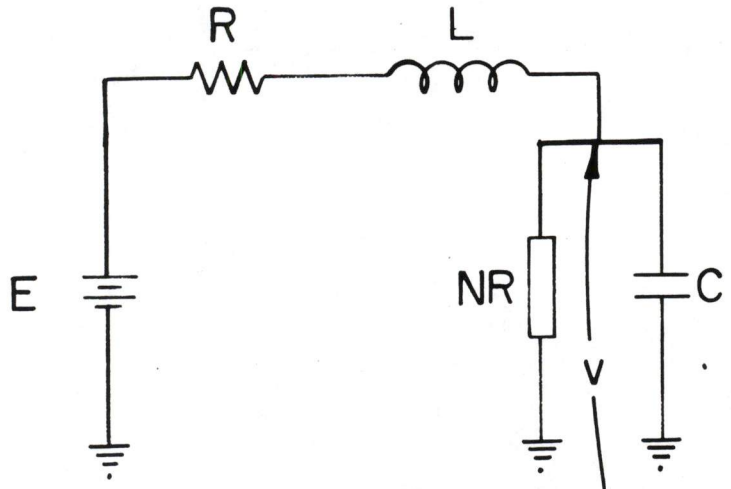
Fig. 1


AND
S
OR
A_{tt}
AND
t_i
x_{i+1}
y_{i-1}
OR
C_i
OR
C_{i-1}
A_r
A_{LS}
C_i
THR.2
NOT
AND
THR.2
AND
x_i
A_{rs}
AND
y_i
A_h
C_i
AND
AND
t_i
A_C
EXCL. OR
A_r
NOT

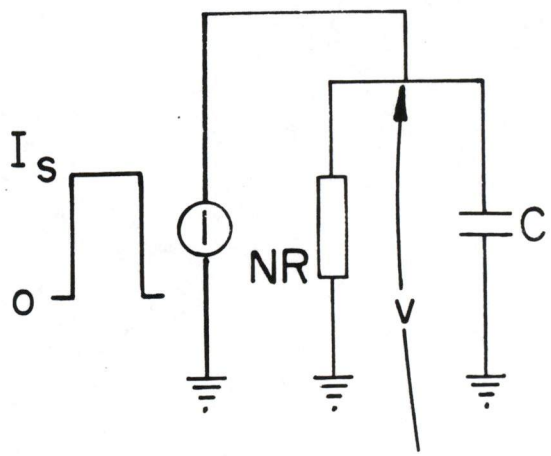


$$\frac{\Delta v}{R_L} = \Delta i_1 - \Delta i_2 \leq \Delta i$$

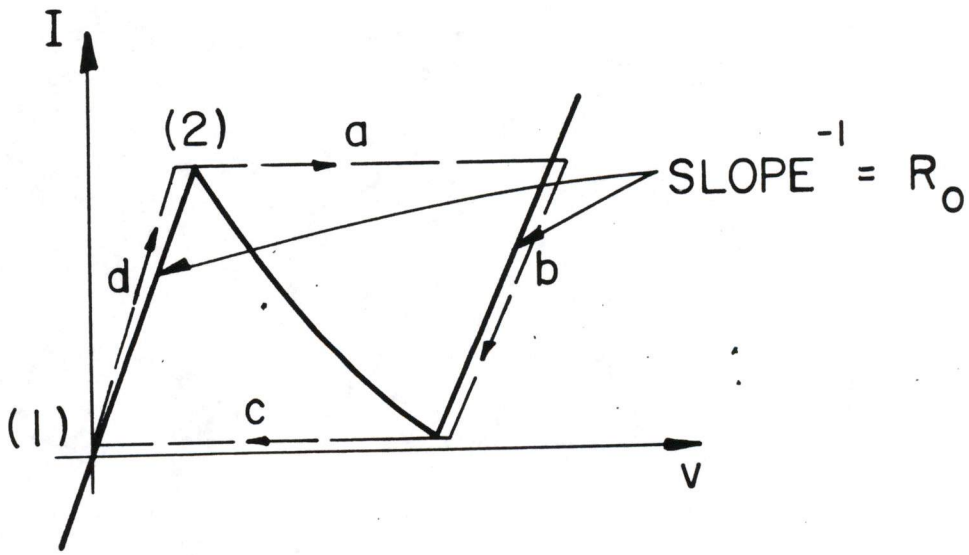
Fig..



(1.)



(2.)

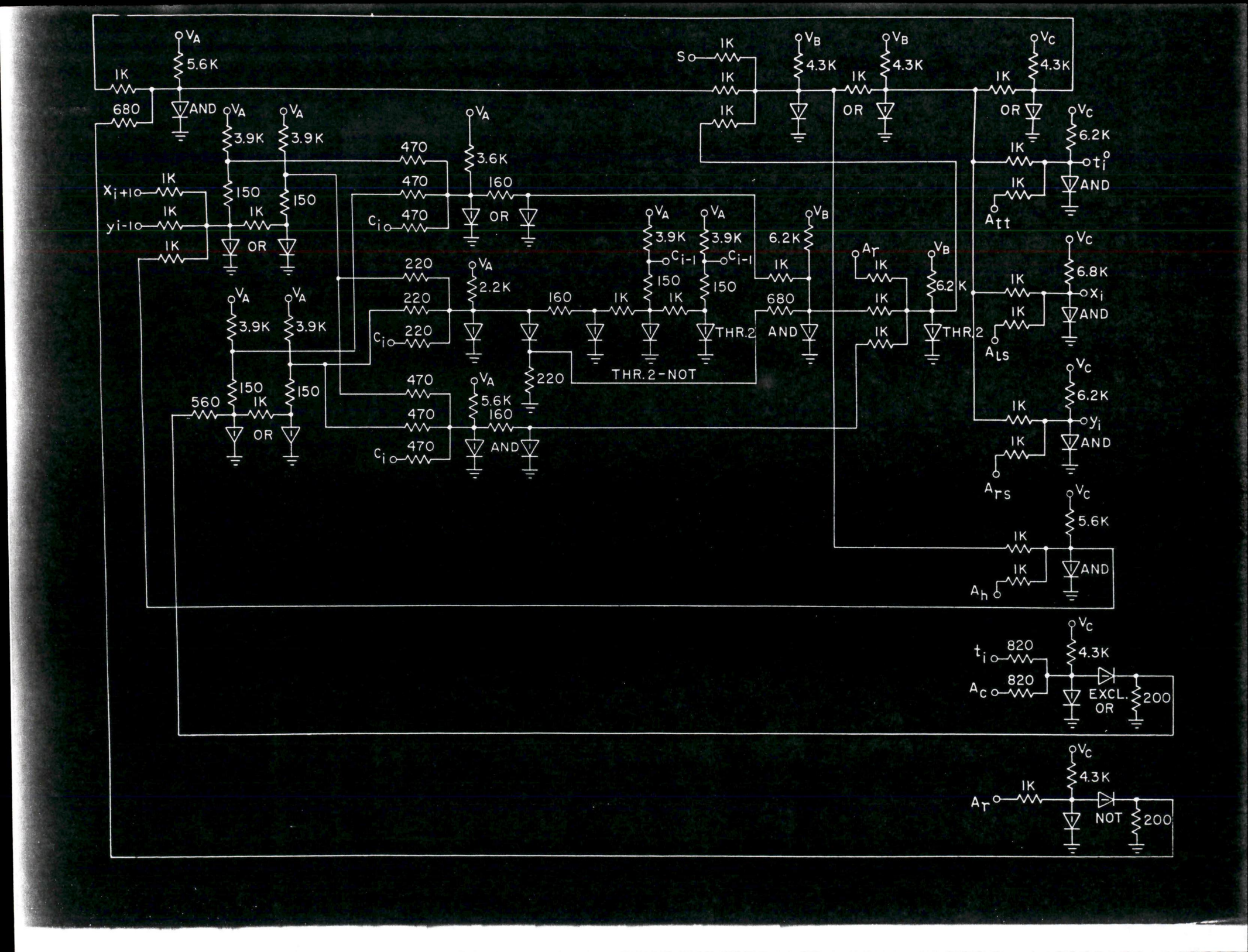


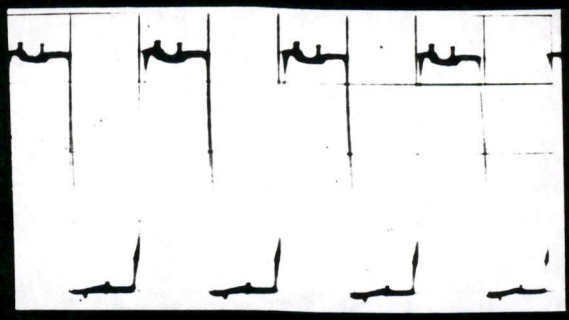
$\mathcal{L}(v(t)) = V(s)$

Fig. 1

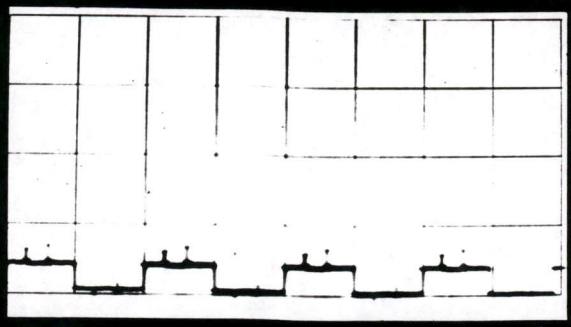
11332

11331

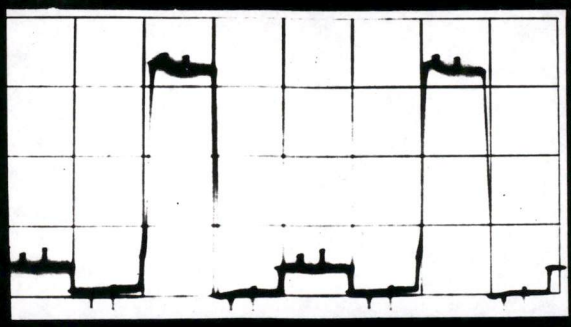

VA
5.6K
IK
680
AND
3.9K
150
OR
470
3.6K
160
220
2.2K
5.6K
560
Xi+1o
yi-1o
Ci o
VB
4.3K
S o
6.2K
Ci-1
THR.2
THR.2-NOT
VC
6.8K
Att
Als
Ars
Ah
Ar
ti
Ac
820
EXCL. OR
NOT
200



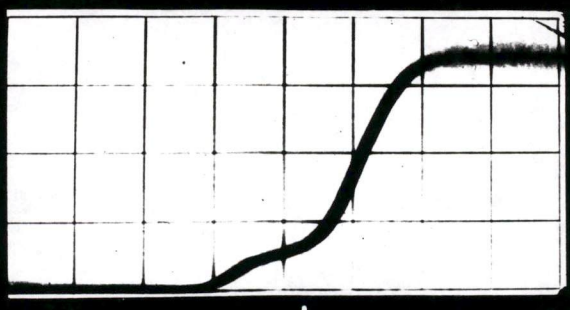
a



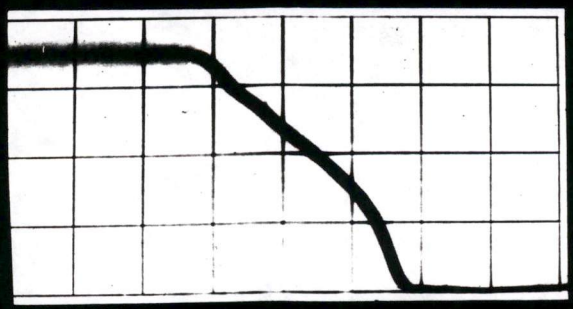
b



c



d



e

FIG 14

RAYTHEON

BIOGRAPHICAL SKETCH

Norman Lourie

BS in Electrical Engineering Worcester Polytechnic Institute 1948

MS in Electrical Engineering Northeastern University 1956

Circuit Design Engineer at Philco and Tracerlab 1948-1951

Since 1951, he has been with the Raytheon Computer Laboratory, later to be known as the DATAmatic Division of Minneapolis-Honeywell Regulator Company.

Presently, an Assistant Project Manager for the Honeywell 800 and a Department Manager in charge of Central Processor Equipment Design.

Member of Institute of Radio Engineers and Association for Computing Machinery.

ERRATA SHEET

"CONTROL AND ARITHMETIC TECHNIQUES

IN A MULTIPROGRAMMED COMPUTER"

Page 11, line 10

"B_i shifted into from B_{i+4} " should read:

"B_i shifted into B_{i+4}"

Page 13, line 7

"A block diagram of the type" should read:

"A block diagram of this type"

File Copy

ARITHMETIC AND CONTROL TECHNIQUES IN A MULTIPROGRAM COMPUTER

N. Lourie, H. Schrimpf, R. Reach and W. Kahn

In the design of a data processor for commercial applications, the designer is constantly striving for better machine performance for little or no increase in cost. In the system design of the Honeywell 800 Transistorized Data Processing System, several design concepts were utilized to help achieve this objective. One of these techniques involves the use of a small auxiliary memory to efficiently aid in the control of the high speed central processor. A second technique uses a new word organization that results in a faster and less costly arithmetic element. These design concepts which were incorporated into the Honeywell 800 System are discussed in this paper.

In a modern transistorized computer, the speeds that are economically achieveable in the Central Processor are very often much higher than necessary to keep up with peripheral devices. The concept of time sharing the Central Processor among several programs in order to utilize otherwise wasted time then becomes attractive. In order to achieve this time-sharing automatically without the use of cumbersome supervisory routines, at least one sequence counter per program is required. If a small coincident current memory running out of phase with the Main Memory were available, a relatively liberal number of programs could easily be run simultaneously by assigning these sequence counters to this control memory. Also, since additional memory locations become economical, it is now simple to assign each program two sequence counters for greater flexibility. These are known as the sequence and cosequence counters respectively. The Honeywell 800 has 8 pairs of

sequence counters, thus allowing the simultaneous operation of eight independent programs. To illustrate how this is performed, Table 1 shows a possible state of the various counters. If, upon starting, the first order is specified from the sequence counter, 00050 will be read from location 2 of the Control Memory, and the contents of 00050 in the Main Memory will be read and performed as an order. The sequence counter will then be incremented by unity so that 00051 will be immediately reinserted into address 2 of the Control Memory as the location of the next order to be performed under control of the sequence counter in program 1. If the previous order in program 2 specified that the cosequence counter was to be used to obtain the next order, the contents of address 35 will then be read out of the Control Memory and 03002 will then be used as a Main Memory address to select the next order performed. Similarly the computer will then cyclicly perform one order from each program. Some orders that leave useful information in the Central Processor do not relinquish control to another program, so that occasionally, several orders from one program will occur before any orders from another program are performed. The multiply order is an example of an order that requires such treatment since there is still a low order product that may be required after the completion of the order. Because the Control Memory is running simultaneously but out of phase with the Main Memory, this multiple operation not only is extremely flexible. but is performed without loss of speed.

Each of the sequence and cosequence counters in the Honeywell 800 has associated with it in the Control Memory another register known as a history register. Whenever a sequence or a cosequence counter is modified because of a sequence change, the associated history register is changed so

that it contains the address that the sequence or cosequence counter would have contained if there were no sequence change. With this feature available, the programmer can easily sequence change into a subroutine and then, at some later time, revert back to the main routine. Table 2 gives a numerical example of this use of a history register.

The same Control Memory can be extremely useful for control of information to and from peripheral devices and the Main Memory. Eight input registers and eight output registers have been reserved in the Control Memory for controlling the transfer of data between peripheral devices and Main Memory. Each of these Control Memory registers is uniquely associated with an input or output trunk. When an input trunk signals that it has a word available, the Central Processor is interrupted at the end of the next memory cycle. The buffer control register in the Control Memory associated with this input trunk is read, and the contents used to select a Main Memory address into which the word from the input trunk can be inserted. The contents of the buffer control register is then incremented by unity and immediately placed back into the same Control Memory location. Thus, the next word from the same input trunk will be inserted into the next highest memory location. In the case of a reverse tape read order, the contents of the Control Memory register is decremented by unity prior to insertion back into the Control Memory. so that the information from tape will be in correct order regardless of the direction of tape motion. Similarly, words are delivered from the Main Memory to a peripheral device in the case of a write order.

If more than one trunk is on demand at the same time, a simple buffer traffic control system establishes priority, and the trunks are processed one at a time. After all input and output trunks are processed, the machine control then reverts back to the main programs.

To increase the average data rate from tape and improve the utilization of tape space, it is desirable to place more than one item on a block of tape. When this is done, it would be desirable to be able to place each individual item in a different section of the memory. In order to accomplish this "distributed reading or writing", a set of address locations that will serve as the starting locations for each of the consecutive items after the first item, is inserted into the main memory. The starting location of this group of beginning item addresses is placed into a control memory address called a distributed item counter. There is one distributed item counter for each input trunk and each output trunk. As before, when a block of information is read into the memory, the initial item is placed into main memory locations as specified by the associated buffer control register. However, when a special bit configuration representing an end of item is sensed, the contents of the main memory location as specified by the distributed item counter is read into the buffer control register, thus creating a new starting address for the next item. The distributed item counter is incremented by unity prior to reinsertion into the control memory, to prepare for the next item. When this change of item location occurs, one extra memory cycle is required for all the associated housekeeping. Distributed writing is performed in a similar manner.

A numerical example of the handling of a four item block is shown below:

Read Buffer Control Register contains	01400
Distributed Read Item Counter contains	00100
Main Memory Address 00100 contains	01500
Main Memory Address 00101 contains	01600
Main Memory Address 00102 contains	01700

With these constants located as shown, the first item would be placed in consecutive memory address locations starting with 01400, the second item starting with location 01500, the third item starting with location 01600, and the fourth item starting with 01700.

Two locations in the control memory are reserved for each of the eight programs to serve as counters for such orders as multiply and multiple transfer orders between groups of memory locations.

The control memory also functions as an aid to indexing addresses. Referring to Figure 1, the base address "y" is read out of one of eight index registers that are available for each of the eight programs. An eight bit augmenter "m" specified by the order is either added to or subtracted from this base address to form the indexed address for the main memory. The base address "y" is reinserted into the control memory unmodified. One extra memory cycle time may be required to perform an order if any of the addresses in that order are indexed.

Indirect addressing is another feature that can easily be accomplished by use of a control memory. In this mode of operation, a number "x" is read from a specified control memory location and used as an address in the main memory. The number "x" is incremented by a constant "n" prior to reinsertion in the control memory, so that next time this control memory contents is used as a main memory address, a different main memory location will be addressed. Thus, with one order it is possible to operate on a whole series of main memory addresses without the necessity for order modification.

The control memory also serves to store a constant U that will give rise to an unprogrammed transfer of control if special situations such as end of tape, addition overflow, or read error occur. When one of these situations arises, the constant U is incremented by n, and an unprogrammed transfer of control to address U+n occurs. The constant "n" is a function of the type of situation that calls for the unprogrammed transfer of control. Since there is an Unprogrammed Transfer Register for each of the eight programs, eight independent U constants can be stored.

A mask index register is available for each program, such that any one of 64 mask constants can be called out of main memory by using one of the mask type orders and an incrementing constant.

A summary of the assignment of control memory locations is shown below:

Address	Description
0	AU-CU Control Counter No. 1
1	AU-CU Control Counter No. 2
2	Sequence Counter
3	Co-Sequence Counter
4	Sequence History Register
5	Co-Sequence History Register
6	Unprogrammed Transfer Register
7	Mask Index Register
8-15	Index Registers 0 through 7
16-27	General Purpose and Indirect Addressing Registers

These 28 locations are repeated 8 times so that each of the eight programs has a unique set of these registers.

In addition, there are eight each of the following registers which are associated with input-output. These registers are not uniquely associated with any program, but are available for convenient assignment to any program.

Address	Description
28	Read Address Counter
29	Distributed Read Address Counter
30	Write Address Counter
31	Distributed Write Address Counter

When the computer designer initially considers the specifications of the Arithmetic Unit of a digital computer, one of the prime considerations is the method performing addition. A good design will be capable of meeting the speed specifications with a minimum of hardware. The format of the bits in the Arithmetic Unit is an important factor in fulfilling this objective. Various formats for a 48 bit word are discussed below.

Serial

A pictorial representation is shown in Figure 2. In this arrangement, as well as all others to follow, for the sake of simplicity it is assumed that the A and B operand each reside in a 48 bit flip-flop register, each stage of which is capable of shifting. At the completion of the addition, the sum will be located in the B register. Since the addition is taking place only one bit at a time, a minimum of equipment is required. However, in order to achieve a reasonably fast add time, relatively high speed shifting flip-flops would be required. For instance, with 4 MC flip-flops, 24 microseconds would be required for a complete addition with end around carry.

Parallel-Serial

A pictorial representation of a 48 bit parallel-serial accumulator with 4 bits in parallel and 12 digits in serial is shown in Figure 3. Other geometries could be used here, but this one is a very important one, inasmuch as 4 bits in parallel can be used for a binary coded decimal digit. Since the adder is now a 4 bit adder instead of a 1 bit adder, more time will be required to propagate carries in the adder, thus resulting in a slower information shifting rate for a given circuit capability, than in the serial adder. Using a 125 millimicrosecond carry propagation per stage and a 1.33 MC shifting rate, the add time will be 18 microseconds, again including end around carry propagation. The addition time is not too much faster than the example given in the serial adder, but the speed requirement of the flip flops is reduced.

Parallel

When the ultimate in addition speed is required, a complete parallel accumulator as shown in Figure 4 is often used. To achieve the fullest speed advantages, the carry propagation time should be completely asynchronous. With no "carry hopping" or "end of carry" sensing, an equivalent add time with the same circuits and assumptions above would be 6 microseconds. If decimal add were added, another 1.5 microseconds would be required. Assuming the speed-up techniques suggested previously are used, average add times on the order of 1 to 2 microseconds are feasible, but the increase in the number of logical statements is substantial. The number of logical statements required without these speed-up techniques is about 12 times as many as the parallel serial adder since the full add logic is required for each of the 48 stages.

Upon examining the requirements of the response time of any adder stage in the parallel accumulator, it is noted that although any stage is required to propagate a carry in a short time, once that stage has responded, it rests for the remainder of the carry time, resulting in a very inefficient use of the inherent speed available. If good speed-up techniques are used, then this inefficiency is greatly reduced. This observation then suggests that a parallel accumulator without speed-up techniques is an extremely wasteful device. It was this observation that led to the invention of the parallel-serial-parallel accumulator. The parallel-serial-parallel accumulator is an efficient extension of the parallel-serial accumulator which results in speeds comparable with that of a parallel accumulator with no speed-up techniques, but with approximately one-fourth the number of logical inputs to the logical expressions for the adder.

Parallel-Serial-Parallel

The parallel-serial-parallel arrangement described here consists of three parallel 16 bit parallel-serial registers. Each 16 bit parallel-serial register has the bits of a 4 bit character in parallel, with 4 characters in serial. Each of the three 16 bit groupings is referred to as a major character. A major character is divided into four 4-bit groups, called minor characters, the bits of which appear in parallel. Major character 1 contains bits 0 - 15, major character 2 contains bits 16 - 31, and major character 3 contains bits 32 - 47. In 4 pulse times, the sum within each major character is computed. Carries generated as a result of these additions are then propagated and added into the next major character in the next 4 pulse times. At the end of these 8 pulse times, the probability* that the carries (including end-around) will be finished propagating and the answer will be correct is $1 - \frac{3}{2^{17}} = 0.999977$. Carry propagation completion can be sensed by means of a three leg buffer, and as much additional time as necessary (8 pulse periods maximum) allowed to complete the carry propagation.

Figure 5 shows a 48 bit binary PSP accumulator capable of either addition or subtraction. The equations for this are shown below.

S = Subtract; $\overline{S}$ = Add

A_n = Addend

B_n = Initial augend and final result

C_n = Carry functions

P_n = Final sum functions

CC_n = Character carry from each adder

T_{4n} = Timing function every 4th clock time such that $T_{4n} \cdot CC_n$ is the carry from the highest order minor character in each major character.

* Assuming linearly distributed random numbers for operands.

Equations for Major Character 1 Adder

$$C_0 = T_{4n}\ CC_3 + \bar{T}_{4n}\ CC_1$$

$$C_1 = C_0\ B_0 + \bar{S}\ A_0\ B_0 + \bar{S}\ A_0\ C_0 + S\ \bar{A}_0\ B_0 + S\ \bar{A}_0\ C_0$$

$$C_2 = C_1\ B_1 + \bar{S}\ A_1\ B_1 + \bar{S}\ A_1\ C_1 + S\ \bar{A}_1\ B_1 + S\ \bar{A}_1\ C_1$$

$$C_3 = C_2\ B_2 + \bar{S}\ A_2\ B_2 + \bar{S}\ A_2\ C_2 + S\ \bar{A}_2\ B_2 + S\ \bar{A}_2\ C_2$$

$$CC_1 = C_3\ B_3 + \bar{S}\ A_3\ B_3 + \bar{S}\ A_3\ C_3 + S\ \bar{A}_3\ B_3 + S\ \bar{A}_3\ C_3$$

$$B_{12} = P_0 = A_0\ B_0\ C_0 + A_0\ \bar{B}_0\ \bar{C}_0 + \bar{A}_0\ B_0\ \bar{C}_0 + \bar{A}_0\ \bar{B}_0\ C_0$$

$$B_{13} = P_1 = A_1\ B_1\ C_1 + A_1\ \bar{B}_1\ \bar{C}_1 + \bar{A}_1\ B_1\ \bar{C}_1 + \bar{A}_1\ \bar{B}_1\ C_1$$

$$B_{14} = P_2 = A_2\ B_2\ C_2 + A_2\ \bar{B}_2\ \bar{C}_2 + \bar{A}_2\ B_2\ \bar{C}_2 + \bar{A}_2\ \bar{B}_2\ C_2$$

$$B_{15} = P_3 = A_3\ B_3\ C_3 + A_3\ \bar{B}_3\ \bar{C}_3 + \bar{A}_3\ B_3\ \bar{C}_3 + \bar{A}_3\ \bar{B}_3\ C_3$$

B_i shifted into from B_{i+4} where $0 \leq i \leq 11$

Equations for the major character 2 adder are the same with subscripts on A_n, B_n, C_n increased by 16, CC_1 substituted for CC_3, and CC_2 substituted for CC_1.

Equations for the major character 3 adder are the same with subscripts on A_n, B_n, C_n increased by 32, CC_2 substituted for CC_3, and CC_3 substituted for CC_1.

The average add time with 125×10^{-9} carry propagate time and 1.33 MC flip-flops is approximately 6 microseconds.

This accumulator has been organized in such a manner that decimal arithmetic using binary coded decimal representation can be incorporated easily, since each minor character is a binary coded decimal digit. To include decimal arithmetic two areas need to be changed. The first is involved with rectification of the binary sum where either the binary coded decimal sum is greater than nine, or a major character carry was generated. This can easily be done by inserting the logic between B_{13} and B_9, B_{14} and B_{10}, and B_{15} and B_{11} below.

D = Decimal $\quad \bar{D}$ = Binary

$$B_8 = B_{12}$$

$$B_9 = \bar{D}\ B_{13} + \overline{CC_1}\ \bar{B}_{15}\ B_{13} + D\ B_{15}\ B_{14}\ \bar{B}_{13} + D\ CC_1\ \bar{B}_{13}$$

$$B_{10} = \bar{D}\ B_{14} + \overline{CC_1}\ \bar{B}_{15}\ B_{14} + \overline{CC_1}\ B_{14}\ B_{13} + D\ CC_1\ \bar{B}_{15}\ \bar{B}_{13} + D\ CC_1\ B_{15}\ \bar{B}_{14}\ B_{13} + CC_1\ B_{14}\ \bar{B}_{13}$$

$$B_{11} = \bar{D}\ B_{15} + \overline{CC_1}\ B_{15}\ \bar{B}_{14}\ \bar{B}_{13} + D\ CC_1\ \bar{B}_{15}\ \bar{B}_{14}\ B_{13} + CC_1\ B_{15}\ B_{14}\ B_{13}$$

The above can be verified with a simple truth chart.

The second area that needs change is the generation of inter-digit carries. This is accomplished by adding a few terms to C_0 to take care of those cases where the decimal sum or difference is between 10 and 15.

$$C_0 = T_{4n}\ CC_3 + \bar{T}_{4n}\ CC_1 + \bar{T}_{4n}\ D\ B_{15}\ B_{14} + \bar{T}_{4n}\ D\ B_{15}\ B_{13} + T_{4n}\ D\ B_{37}\ B_{36} + T_{4n}\ D\ B_{37}\ B_{35}$$

The two corrections required the addition of only 60 x 3 = 180 diodes to the accumulator. No amplifiers were added. The probability that the answer will be correct after 8 pulse times is $1 - \frac{3}{2000} = 0.9985$. For these cases, up to 9 more pulse times may be required for the correct answer.

In addition to allowing a very economical method of arithmetic, the PSP format allows other machine simplifications over a parallel format. In particular, it is possible to transmit a 48 bit word to remote portions of the machine by means of time sharing 12 lines, resulting in a decreased number of cable drivers. It also allows the use of one flip-flop and 3 pulses of delay line to store 4 bits of information in those cases where the other 3 flip-flops are not required for manipulation reasons. A block diagram of the type of storage is shown in Figure 7.

Parallel-Serial-Parallel Arithmetic Example

Figure 8 is an example illustrating two numbers being added together in a parallel-serial-parallel adder. The zeros immediately beneath the operands indicate the initial state of the carry circuits at the beginning of the addition. All of the numbers on a line with t1 indicate the computation being performed during the first pulse period. The 11 indicates that the "4" and "7" in the low order position of major character 3 is added to form a 1 sum and a 1 carry, and, simultaneously, the "5" and "4" in the low order portion of major character 2 is added to form a "9" with 0 carry indicated by the 09, and, similarly, 1 and 6 are being added to form 07.

During the next pulse period, t2, (and through the same addition circuits which produced the "t1" addition) "2" plus "9" in major character 1, together with the previous carry, "0", forms the 11 on the t2 line. Similarly "1" plus "8" plus "0" carry forms 09 and "7" plus "1" plus a "1" carry forms 09.

This process is repeated until the t1 line of word cycle 2. At this point, the carry from the high order position of major character 1 is added to the low order digit in the partial sum of major character 2. The carry is then propagated as shown until the 3699 is "corrected" to 3700. Similarly, the 9991 formed in major character 3 is corrected to 9992.

As seen in this example, the addition is completed after 8 pulse periods.

In summary, the Parallel Serial Parallel format provides a fast arithmetic speed at a relatively economic cost of logical circuitry. The addition of the Control Memory to the system provides a wide range of flexibility in order to achieve an efficient usage of the computer.

The authors of this paper wish to extend credit to all those at the Datamatic Division of Minneapolis Honeywell Regulator Company, whose ideas contributed to the creation of this data processing machine.

Use of Control Memory for Simultaneous Program Operation

Program	Sequence Counter	Location	Consequence Counter	Location
1	00050	2	00600	3
2	02090	34	03002	35
3	•	66	•	67
4	•	98	•	99
5	•	130	•	131
6	•	162	•	163
7	•	194	•	195
8	•	226	•	227

Table 1

Use of Sequence History Register to Relocate after Subroutine

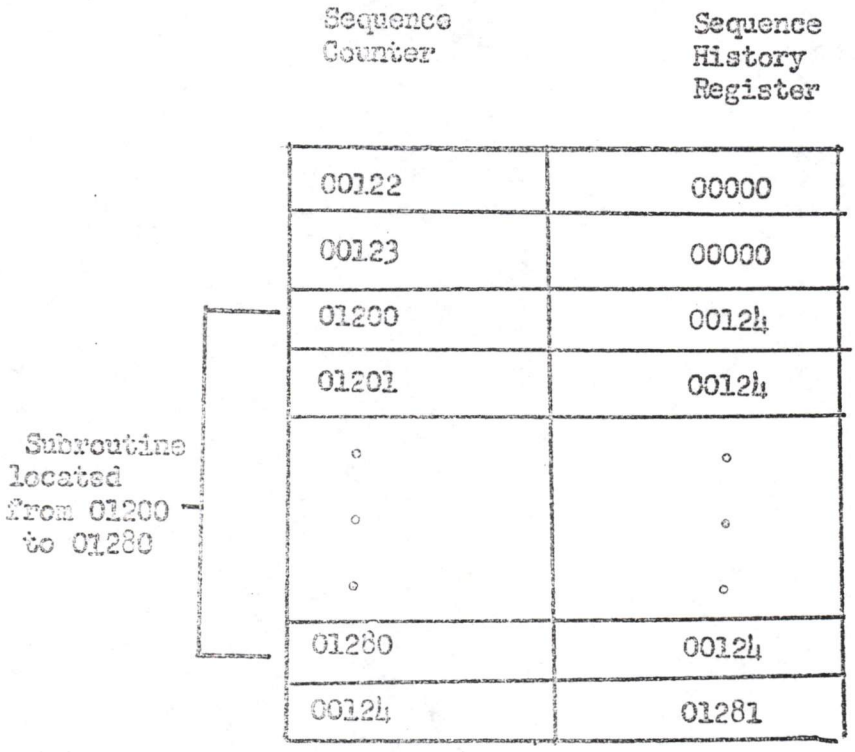


Table 2

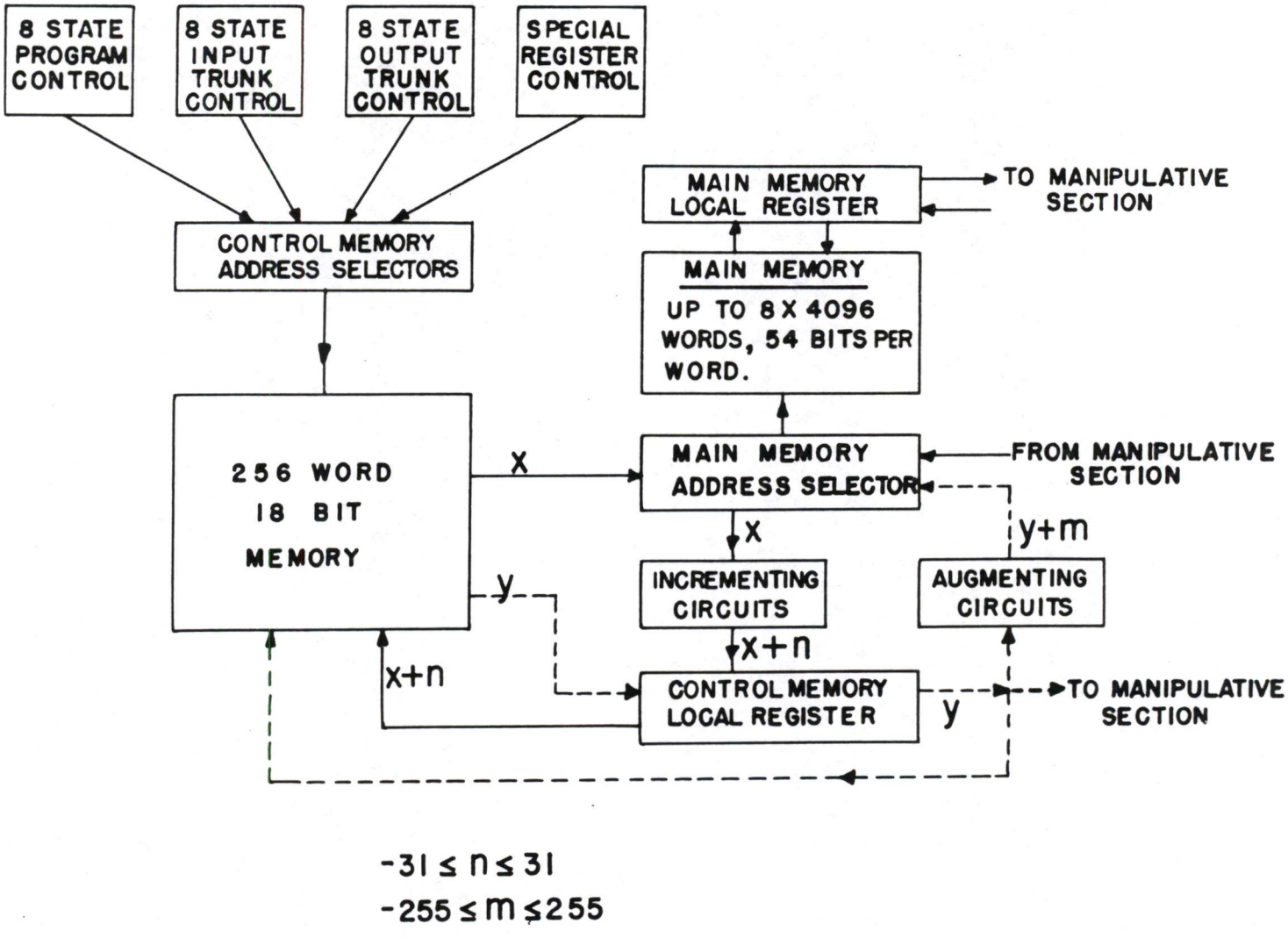


FIG. 1, CONTROL MEMORY IN A MULTIPROGRAM COMPUTER

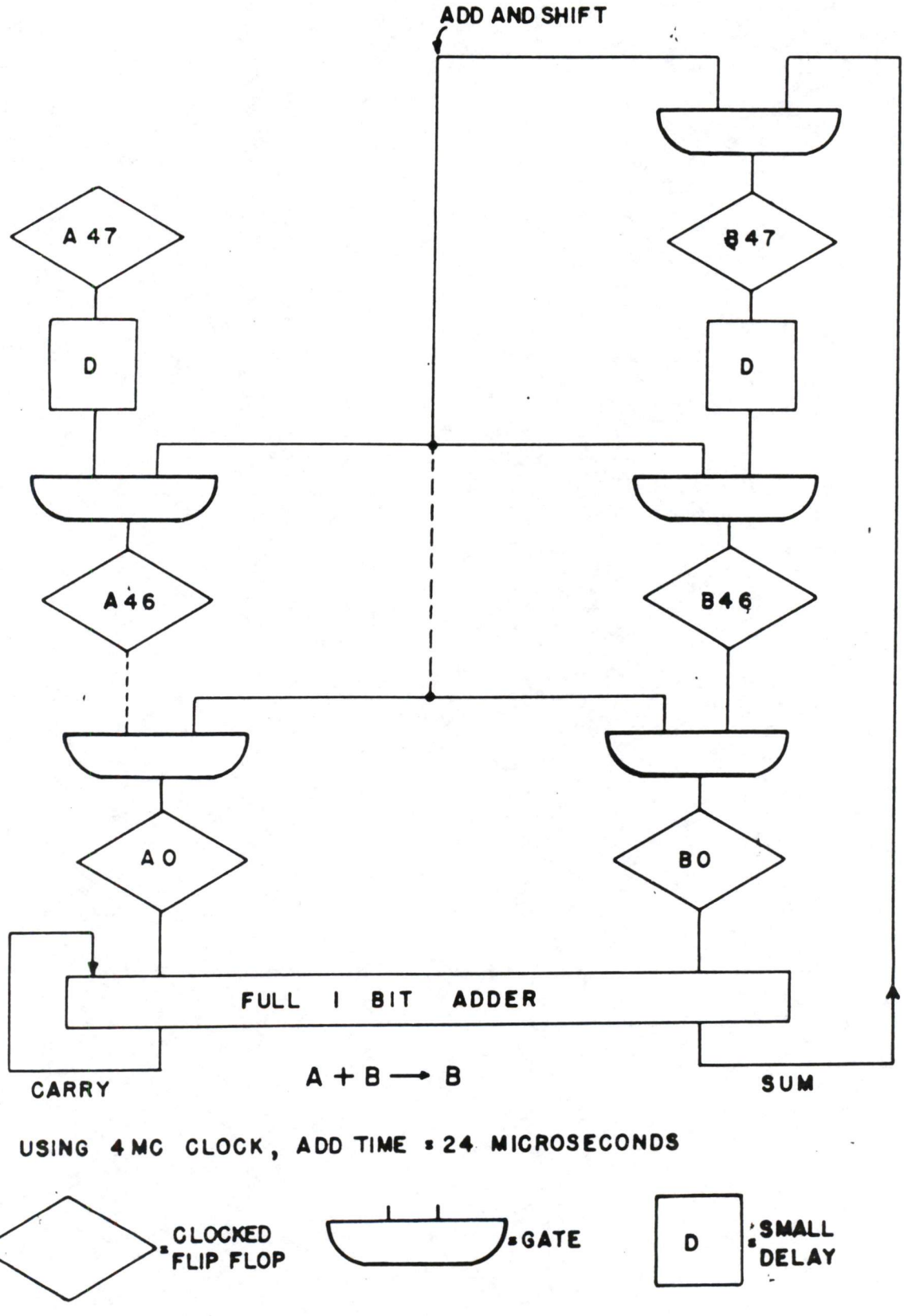
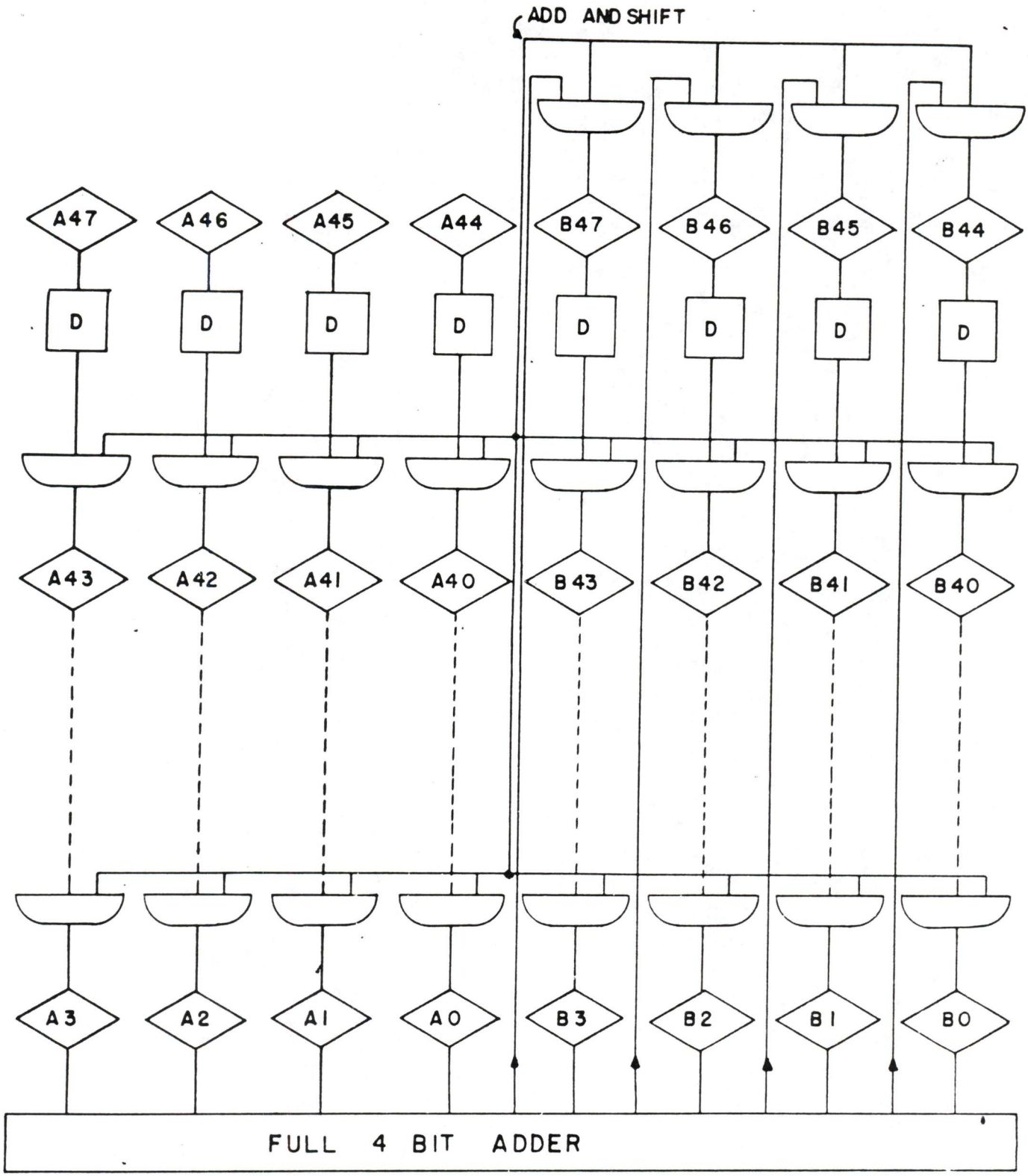
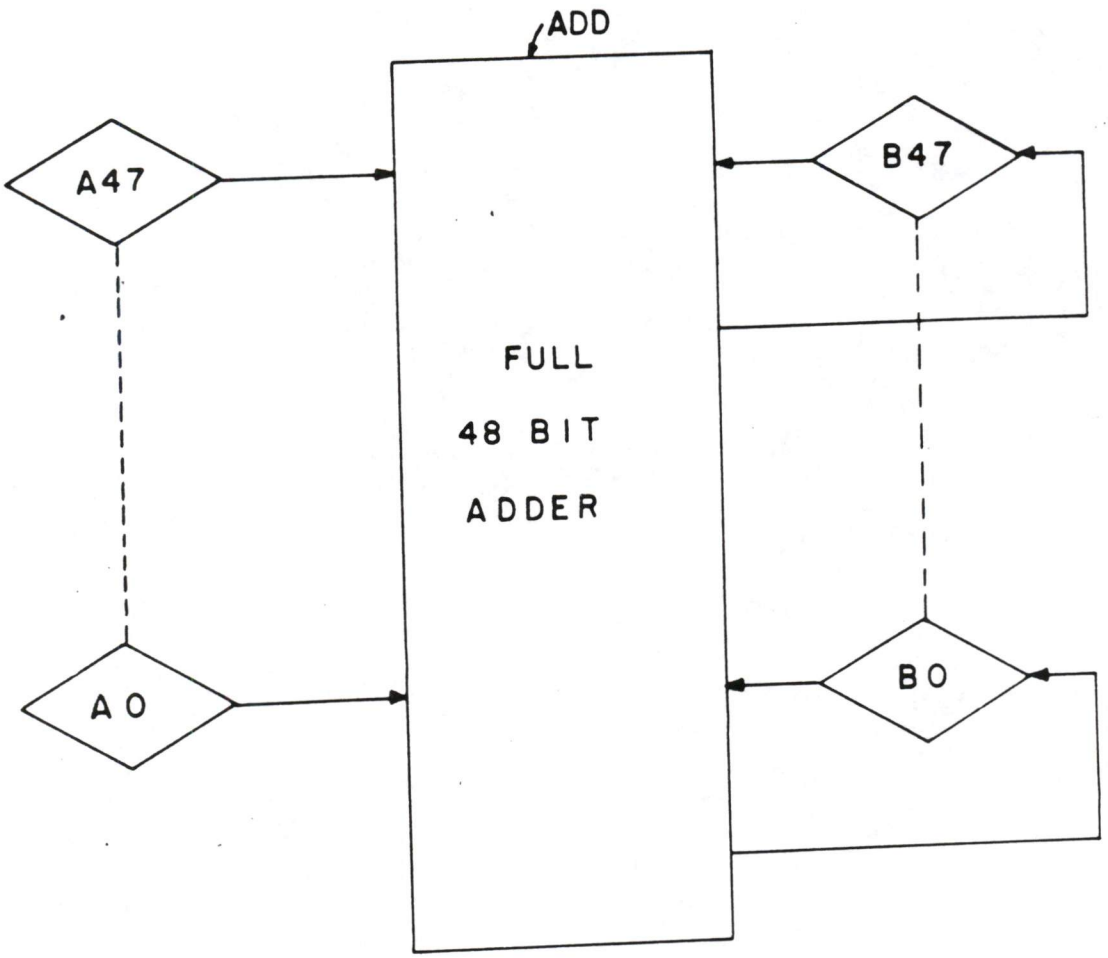


FIG. 2, 48 BIT SERIAL ACCUMULATOR



USING 1.33 MC CLOCK, ADD TIME = 18 MICROSECONDS $A + B \longrightarrow B$

FIG.3, 48 BIT PARALLEL SERIAL ACCUMULATOR. 4 BITS ≡ 1 CHARACTER IN PARALLEL, 12 CHARACTERS IN SERIAL.

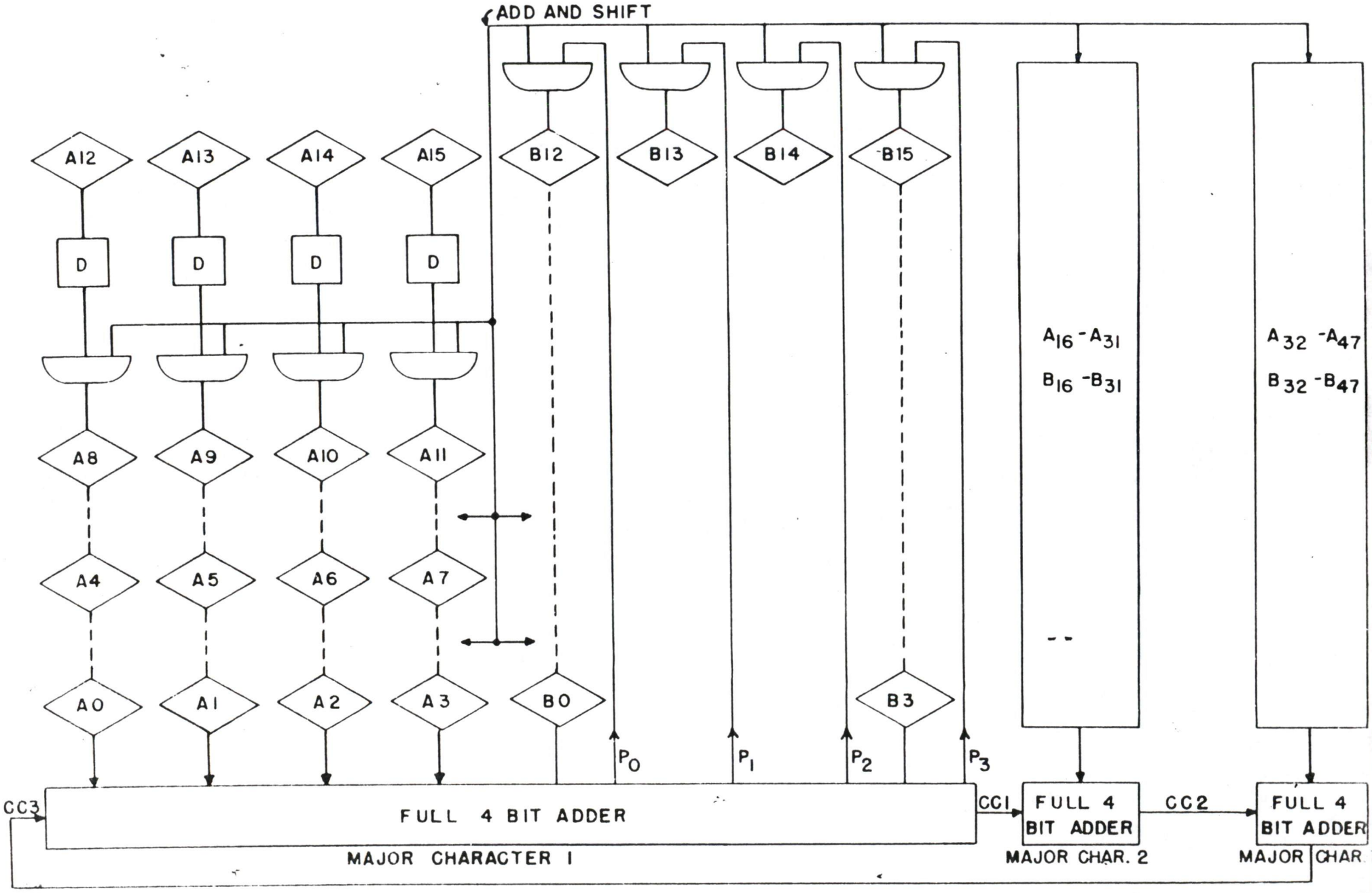


USING CARRY PROPOGATE TIME OF 125×10^{-9} SECONDS PER STAGE.

BINARY ADD TIME = 6 MICROSECONDS

DECIMAL ADD TIME = 7.5 MICROSECONDS

FIG.4, 48 BIT PARALLEL ACCUMULATOR


ADD AND SHIFT
A12
A13
A14
A15
B12
B13
B14
B15
D
D
D
D
A8
A9
A10
A11
A4
A5
A6
A7
A0
A1
A2
A3
B0
B3
A16 -A31
B16 -B31
A32 -A47
B32 -B47
P0
P1
P2
P3
CC3
FULL 4 BIT ADDER
CC1
FULL 4 BIT ADDER
CC2
FULL 4 BIT ADDER
MAJOR CHARACTER 1
MAJOR CHAR. 2
MAJOR CHAR

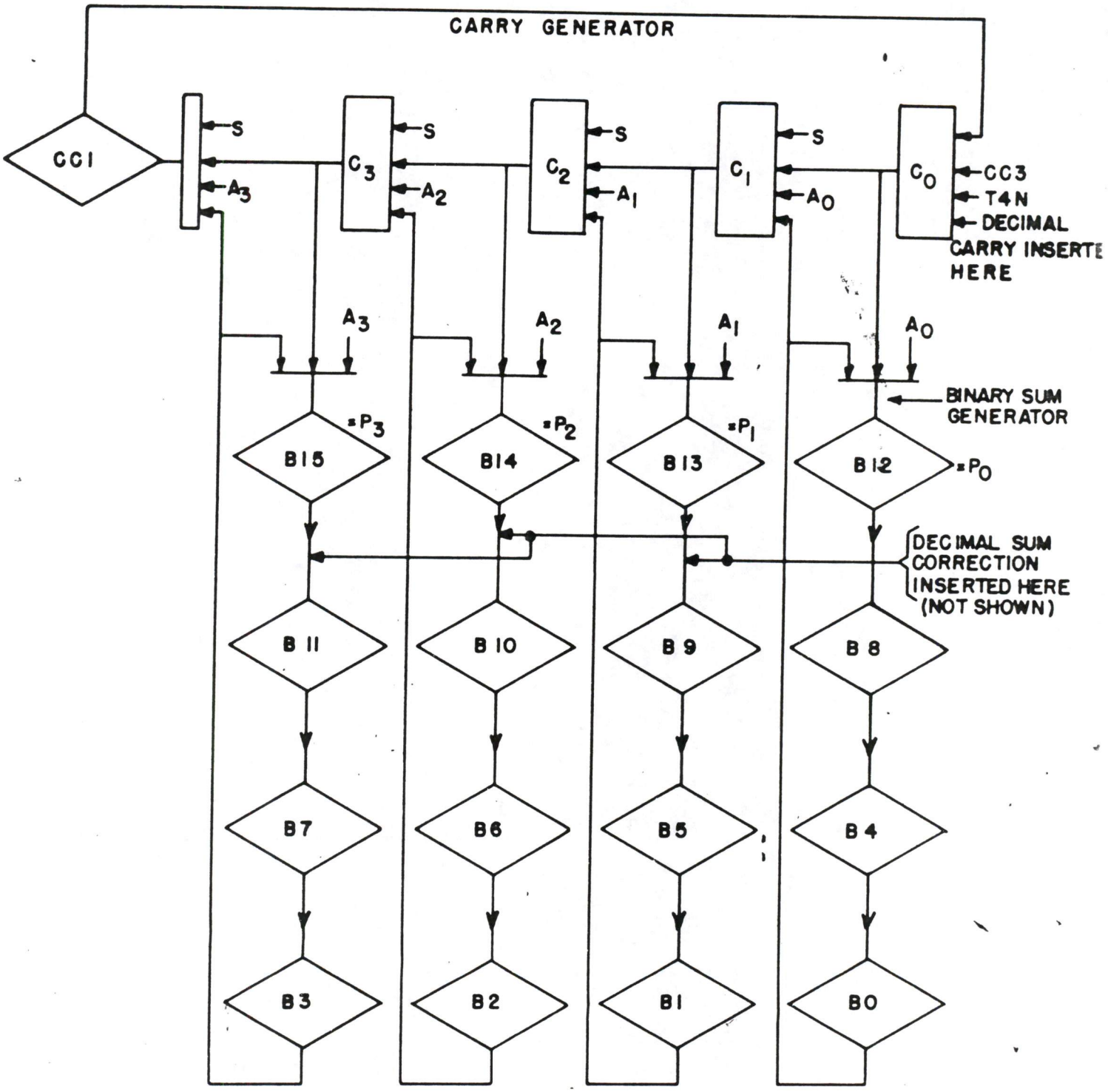


FIG. 6, LOGICAL ORGANIZATION FOR MAJOR CHARACTER I OF PARALLEL SERIAL PARALLEL ACCUMULATOR.

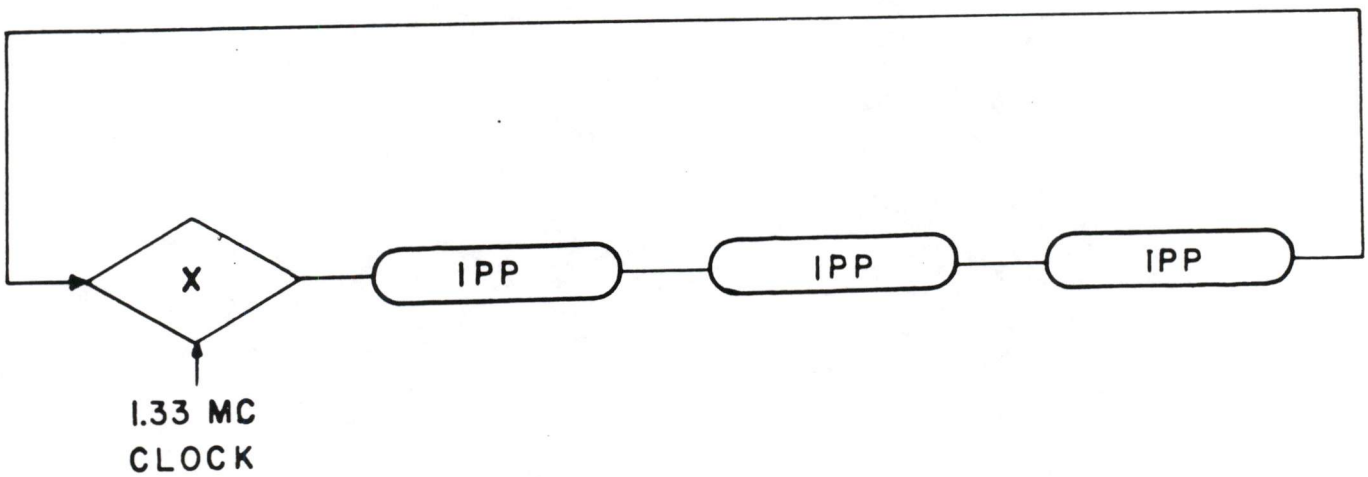


FIG.7, USE OF ONE FLIP FLOP AND 3 PULSES OF DELAY TO STORE 4 BITS.

PARALLEL— SERIAL—PARALLEL WITH VARIABLE CYCLE

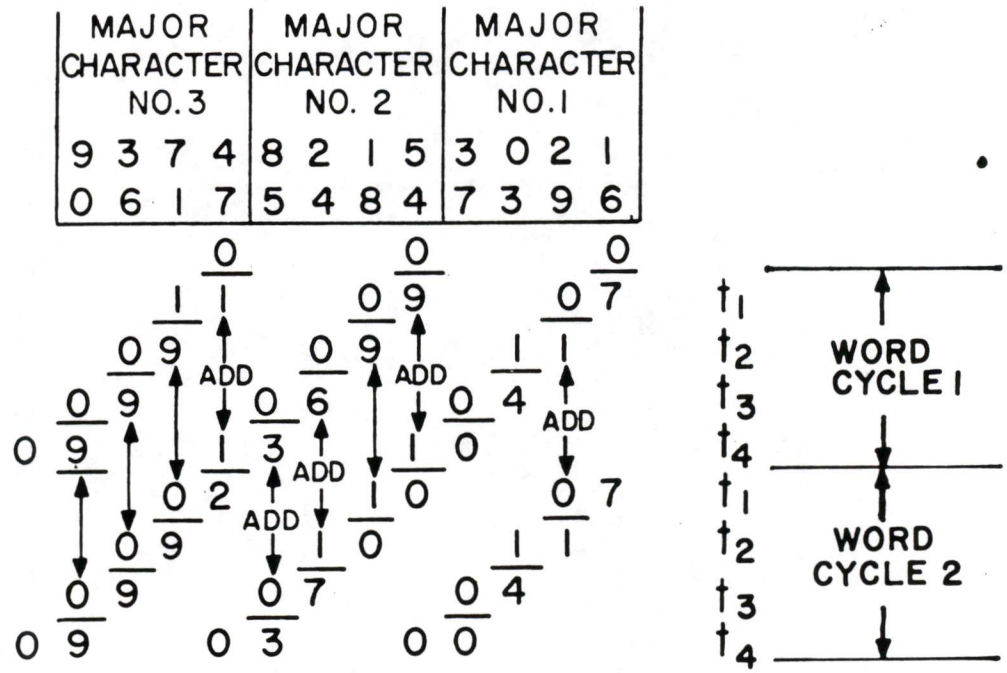


FIGURE 8

RADIO CORPORATION OF AMERICA
RCA LABORATORIES
David Sarnoff Research Center
PRINCETON, N. J.

November 12, 1959

Mr. Harlan E. Anderson
EJCC Publication Committee
Digital Equipment Corporation
Maynard, Massachusetts

Dear Mr. Anderson:

Please find enclosed four copies of the manuscript and original figures of the paper entitled

"SOLID STATE MICROWAVE HIGH SPEED COMPUTERS"

which I am to present on December 1 at the forthcoming 1959 EJCC.

Trusting it reaches you before the fateful Sunday, November 15, 1959.

Sincerely,

Jan Rajchman

Jan A. Rajchman

JAR:at

Enclosures

File Copy

SOLID STATE MICROWAVE HIGH SPEED COMPUTERS

by

Jan A. Rajchman

R.C.A. Laboratories, Princeton, N. J.

ABSTRACT

Two types of semi-conductor devices offer possibilities to speed computer rates up to thousand megacycles. (1) Variable capacity diodes in parametric subharmonic phase locked oscillators have permitted pumping up to 10 KMC and making of 4 KMC-pumped 100 mc logic circuits. Junction type diodes of order of magnitude higher speed capability were developed. (2) Tunnel diodes switching in less than 10^{-9} seconds were developed. Scaled-down frequency logic pulse type circuits were demonstrated. Arrays of tunnel diodes promise random access memories with access cycles of 10^{-8} to 10^{-7} seconds. - Both diodes were developed in microcapsules fitting within microwave transmission boards and dissipate sufficiently low power to permit necessary high packing density for ultra fast computers. Special circuit techniques for use of two terminal single port devices were developed.

SOLID STATE MICROWAVE HIGH SPEED COMPUTERS

by

Jan A. Rajchman

R.C.A. Laboratories, Princeton, N. J.

I. INTRODUCTION

This paper presents results of an effort aimed at developing the principles and technology required to speed the rate of computers up to the order of thousand megacycles. The approach is based on the use of two types of two-terminal semi-conductor devices: the variable capacity diode and the tunnel diode in combination with microwave techniques for the couplings within the computer.

Both devices provide amplification of binary signals by mechanisms depending on negative resistance. Their speed limitation is primarily due to the capacity of the junction and internal series resistance and can be two orders of magnitude higher than that of transistors which are limited by to drift time of minority carriers. The variable capacity diode can be used for computer logic in parametric phase locked oscillators according to concepts described by Goto[1] and Von Neumann[2]. The negative resistance of the tunnel diode can provide amplification and gain directly. Both devices have only two terminals, i.e. a single port for the input and output, so that special methods are required to give direction to information flow. These methods and the means to perform the other necessary functions of storing and gating signals are described in the following.

II. PARAMETRIC PHASE-LOCKED SUB-HARMONIC OSCILLATOR (PLO) COMPUTERS

1. Principle of Operation

Consider a tuned circuit composed of a fixed inductance and a capacity whose value depends on the voltage across it. (i.e. junction diode) Let the tuned circuit be excited by a frequency 2f which is approximately equal to twice the resonant frequency of the circuit. (Fig. 1) This excitation will tend to produce oscillations at frequency f in the circuit, and oscillations will actually be sustained if the excitation is sufficiently intense and the losses in the circuit are sufficiently small. This effect is a special "degenerate" case of a broad class of parametric excitation effects. The general theories of parametric oscillations, as well as the particular theory of this degenerate case have been reported by several authors[3,4,5,6,7].

The reason for the build up of oscillations can readily be understood by a simple physical reasoning. Let us assume that every time at which the capacity has maximum charge, the value of the capacity is reduced, as would be the case if the plates were pulled apart. The work necessary to reduce the capacity increases the energy stored in the condenser. The value of the capacity is restored to its initial value at the instant of which the charge in the condenser is zero. In this way a certain amount of energy is added to the circuits at every half cycle of the oscillation. If this increase of energy is greater than the loss of energy in the half cycle due to damping in the circuit, the amplitude of oscillations will grow. It is easy to see that this "pumping" of energy occurs at twice the frequency of oscillations and therefore can sustain an oscillation of either of two opposite phases. In the actual case of a variable capacity diode, the change of capacity is due to the voltage of the pump source applied to it instead of the mechanical work

necessary to pull the plates apart, but the effect is analogous.

The oscillations are sustained in either of two opposite phases which are locked to the phase of the pump and can be used to denote "zero" and "one" of a binary digit. The phase-locked-oscillator, PLO, constitutes thus a storage cell. The steady state phase depends on the conditions under which oscillations start. If a small locking signal at the frequency f is present in the tank, oscillations will build up in the phase closest to the phase of the locking signal. The input locking signal is thus "amplified". (Fig. 1).

Logic can be performed by arraying the PLO's in three or more groups, which are separately activated either by pump modulation or diode bias gating. Every PLO is loosely coupled to PLO's in other groups, the pattern of couplings determining the logic task to be performed. The groups are clocked in succession with some overlap, i.e. a given clock is turned off after the next one is turned on. This sequence causes information to flow in a given direction despite the bilateral character of the PLO. A PLO will start at the phase determined by the phase of the majority of oscillating PLO's to which it is coupled. The majority decision can be exploited directly in many circuits or can be reduced to "and" or "or" decisions by the use of a reference signal on one input. For example with two inputs, and a reference in phase zero, the output will be in phase π only when both inputs are in phase π. Negation is easily obtained by phase inversion. In a typical example of logic circuit, figure 2, each PLO may be connected to two inputs, two outputs and one reference, or to five other PLO's. Consequently, the input is at most one fifth of the output of preceding PLO's. Thus, a minimum "logic gain" of five is required. In a simple shift register, figure 3, minimum

logic gain is two.

There is a certain increase of amplitude of oscillation at each cycle which depends on the parametric pumping, i.e. specific variation of capacity and power, and on the losses of the circuit which are made up of the useful loading and unavoidable circuit dissipations. To build up the amplitude by a factor corresponding to practical logic gains, about 5 cycles of oscillations or 10 pump cycles are required in typical PLO's. Therefore to obtain 1000 mc information rates, i.e. phase switching in about 3×10^{-10} sec, pump frequecies of about 30 KMC or higher are required.

2. Experimental Results

An experimental program ultimately aimed at PLO computers pumped at frequencies of about 30 KMC resulted in the following:

A. Microwave Sub-harmonic Oscillators

Microwave circuits obtained by photographic engraving of copper clad insulating boards, known as strip transmission lines, and point contact diodes in conventional microwave cartridges, were used for PLO's pumped at 4 KMC. A typical early configuration (Fig. 4 and 5) included, a 2 KMC quarter-wave resonator with diode in shunt at one end, a 4 KMC resonator bar isolating pump and oscillating circuit, d-c return for optimum bias, and one or more loosely coupled inputs and outputs[8,9,10,11,12]. Output-vs-input power characteristics (Fig. 6) show broad operating range, efficiencies of a few percent, and required pump power levels of about 100 mw. Typical more recent configuration utilize a series connected gold bonded diode (Fig. 7) and multiple impedance matched antennas for couplings inputs and outputs. The characteristics (Fig. 8) show uniform couplings to various

antennas, and broad operating regions.

Methods of switching phase, first investigated in lumped parameter circuits pumped at 5 mc,[13] demonstrated great flexibility of "phase script" and yielded quantitative relations between logic gain and build-up cycles under a variety of conditions (Fig. 9). Qualitative confirmation of these results was obtained through more elaborate experiments with PLO's pumped at 4 MC using mercury wetted relay pulsers and travelling-wave oscilloscopes. Typical results: rise from noise level to saturation in 10 nanoseconds (nanosecond = 10^{-9} sec) and decay in 1.5 nanoseconds when the diode was pulsed slightly into conduction.

B. Microwave Computer Techniques and PLO Logic

Microwave transmission line techniques provide methods for linearly combining signals to exploit direction of transmission. For example a hybrid ring can be used to translate amplitude modulated to phase modulated signals and visa-versa through appropriate combination with a CW signal[9]. (Figure 10). Another example is the use [11] of a hybrid ring fed by two PLO's energized at pumps $\pi/2$ out of phase, so as to obtain cancellation at one (input) terminal and reinforcement (output) at another. (Figure 11). This provides undirectional information flow and thereby lowers the required logic gain and permits the use of a two rather than three clocks. Experimentally two PLO's have been balanced so that only 5% of the power appeared in the input.

A full stage of a binary adder with two PLO's and four hybrid rings was made.[10,11] It operates through linear combinations of phase-script signals and PLO's acting as a majority decision elements and amplifiers (Figure 12). Operation with pulses at repetition of 100 mc was obtained. Other types of adders

were made also.

A 500 mc binary scaler was made[8] using a PLO deliberately tuned at a frequency slightly different from half the pump frequency so that its phase changed for every momentary deactivation of the pump lasting for a time sufficient to allow the natural oscillations to drift more than $\pi/2$ in phase.

In general, making of computer subsystems with PLO's can take advantage of well developed microwave strip transmission techniques. Boards with 3 layers, with ground planes on both sides of transmission strips, permit compact subsystems without deleterious radiation pick-ups of 2 layer boards.

C. PLO Random Access Memory

The random access memory made of PLO's would be particularly suitable in a machine with PLO logic. This possibility was investigated[13]. In a two dimensional array of PLO's continuously activated by a pump, the access problem consists of (1) selectively establishing the desired phase in a selected PLO without disturbing the phase of any other and (2) of interrogating the phase of any selected PLO without ambiguity due to possible masking signals from all other PLO's.

The writing problem (1) is solved easily by forced switching. It is possible to choose the amplitude of locking signals such that each separately is too small but together the two signals are strong enough to change the phase of the PLO.

The reading problem (2) requires a more elaborate artifice. For example a standby PLO in addition to the storing PLO can be used for each bit. Each standby PLO is loosely coupled to its associated storing PLO and is also loosely

coupled to a read-out circuit. For read-out, the standby PLO, normally not activated, is selectively activated by the coincidence of two bursts of pump energization. It starts to oscillate at the phase of the associated storing PLO and thereby conveys the sought phase information to the read-out circuit. The signals of all other storing PLO's are effectively blocked from masking the read-out signals since their standby read-out PLO's are not activated.

Experimental memories have operated successfully at frequencies of less than 10 mc but complexities of technology with early designs of PLO's have made operation at microwave frequencies difficult. Recent improved designs would greatly facilitate the memory design.

D. Variable Capacity Diodes for PLO's

Microwave cartridge point contact diodes of commercial type and laboratory units made by specially developed techniques, as well as gold bonded diodes mounted within the boards in quarter wave series resonant circuits, permitted the experiments reported above but had serious drawbacks due to limitations in speed and wide variations from unit-to-unit. A program to develop junction types has resulted in a clear understanding of the limiting factors[14] and diodes of practical design with an order of magnitude better performance.

The simultaneous realization of low series resistance r_s within the diode and high variation of capacity with respect to voltage ($\frac{1}{C}\frac{dC}{dV}$, around a value C_o arbitrarily taken at -1 volt) requires that the impurity concentration be not uniform but have a specially designed profile. This was realized in solution-grown and out-diffused p-n germanium junctions. High cut-off frequency f_o for reasonable impedance and low power requires that the capacity be low and the area of the junction be of the order of 10^{-6} sq. inches.

In addition to the constants of the semi-conductor proper it is essential to minimize the capacity C_c, the series inductance L_s and the resistance of the contact of the diode encapsulation. A special micro encapsulation was developed (Figure 13, 14) in which the case capacitance is only .5 μμF and the lead inductance is only .3 μμH. It consists of a ceramic ring of .085" in diameter sealed hermetically between two metal plates with a metal finger entering from one side. A thin wire contacts the dot on the germanium wafer and is soldered to the finger. The diode is inserted in the printed winding boards and the upper and lower tabs can be directly soldered to the printed lines. The resulting circuits are not only superior in performance but simple to construct. Typical constants of a microencapsulated variable capacity diode are C_o = 1 μμF, r_s = 1 ohm, f_o = 150 KMC L = 300 μμH, and C_c = .6 μμF. Capacity voltage sensitivity, difficult to measure directly, is relatively high as judged by improved PLO performance.

These diodes have permitted the design of 10 KMC pumped PLO's with which gains of 20 db, rise time of 2 to 3 nanoseconds, and efficiencies of 10% were realized. Corresponding information switching rate would be about 300 mc as judged from the extrapolation of 100 mc rates of 4 KMC pumped PLO's.

III. TUNNEL DIODE COMPUTERS

1. Tunnel Diodes

Abrupt junction diodes made of very highly doped material exhibit a negative resistance at small forward bias. This effect was described by L. Esaki[15] who interpreted it as due to quantum tunneling. As the negative resistance makes amplification possible and the effect is inherently fast, it was realized that the tunnel diode is particularly suited for high speed computers. The device was

investigated in detail and a number of germanium units especially adapted for this use were made.[16]

The current voltage characteristic (Figure 15) exhibits in the forward direction, a maximum, a drop corresponding to negative resistance (-R), a minimum, and a subsequent rise. The negative resistance is well understood by semiconductor theory, and can be thought as due to a diminution in the number of electrons which can tunnel through a potential barrier as that barrier is lowered, a seemingly paradoxical fact resulting from the decrease in electronic states adjacent to the potential barrier.

The gain bandwidth product as well as the upper frequency of oscillation realizable were found[16] to be inversely proportional to RC, where C is the physical capacity of the diode junction. The time constant RC, independent of junction area, can be small despite the large value of C (typically 3 $\mu F/cm^2$). This is because the resistance R can be made very small, as it is an inverse exponential function of the impurity concentration. High concentrations are obtainable by suitable doping techniques. Time constants as low as 5×10^{-11} seconds have been determined by measuring R and C separately. In another experiment a tunnel diode was switchied by means of a mercury wetted relay and the resulting switching was observed on a sampling oscilloscope. Rise and decay times less than 10^{-9} seconds and a plateau of about 10^{-9} seconds were observed. Tunnel diodes oscillators of 1600 mc were made. Recently oscillations as high as 10,000 mc were reported[19].

It is necessary that the series resistance r and the inductance L of the diode and its mount be sufficiently small to make the time constants rC and L/R small compared to RC. Only very short lead-in can be tolerated. Microencapsu-

lations with wide, short, closely spaced terminals have been designed and permit direct incorporation of diodes in low impedance (typically 10 ohms) transmission lines. It turns out that the ceramic microencapsules originally designed for the variable capacity diodes is particularly suitable for the tunnel diode. (Figure 14). These units have only about 300 μμH of series inductance.

Hundreds of tunnel diodes were fabricated on a laboratory scale[16]. These were germanium types with impurity concentrations of about $2.5 \times 10^{19}/cm^3$. Various sizes were made with peak currents varying between 1 and 700 ma. The area of the junction in these diodes is about 5×10^{-5} cm^2 which entails a capacity of about 100 μμF.

2. Tunnel Diode Logic Circuits

Logic switching can be performed by tunnel diodes because their characteristics have sharp thresholds permitting gating and negative resistance permitting signal amplification. Signals are baseband pulses, in contrast to the carrier modulated signals of the PIO. Two distinct regions of the characteristic are used: a voltage range below the voltage of the current peak to denote "0" and a range above the voltage of the current valley to denote "1". (Figure 15). With germanium tunnel diodes the low "0" state is typically less than 50 mv and the high "1" state about 450 mv.

Gain is obtained through a triggering action. An operating point P, with current I_s and voltage V_s, is established near the maximum of the characteristic (I_o, V_o) through appropriate biasing of the power supply. The input signal adds a relatively small increment of current (or voltage) to go over the "hump". This causes the diode to switch to the high state "1". An output current as large as

the difference between the maximum I_o and the minimum I_1 can be obtained without loosing the state "1". The ratio of output-to-input currents, i.e., the gain, can thus be large if the operating point is very near the maximum. Practical nearness of biasing depends on the uniformity of diodes which was found sufficient in experimental diode batches to permit logic gains of 4 to 6. These were observed in circuits pulsed at a rate of one megacycle. The excess of the input signal over the value required to reach the maximum has an appreciable effect on the speed of triggering as it determines the rate at which the capacity of the diode is charged to reach the triggering point. Therefore some reduction of possible logic gain must be suffered to obtain high switching speeds.

Logic switching can be accomplished by properly interconnecting simple logic elements. Each such element, made of one or more tunnel diodes, serves as a bit-store, as an amplifier, and as a threshold gate. The gating is of the majority type with which simple majority, "and", and "or", decisions are obtained with proper choice of the threshold levels. Three main types of logic elements were investigated: (1) bistable, consisting of a tunnel diode in series with a resistance or resistance network (Fig. 16) (2) bistable, consisting of two tunnel diodes in series (Fig. 17) (3) monstable, using diodes in series with inductances. (Fig. 18).

(1) In the first bistable logic element the single port P is coupled resistively to a number of inputs and outputs as well as to a source of current I_s which can be pulsed (Fig. 16). In the absence of I_s the contributions of currents to the diode are so small that its voltage is small, on the "0" part of the characteristic. The value of I_s is so chosen that when the activating pulse is applied the total current will either be smaller or greater than the maximum I_o depending on the sum of the inputs and therefore the voltage will remain small or will be

switched abruptly to the "1" part of the characteristic. This change of voltage influences in turn other logic elements to which it is coupled when these are activated. The logic elements are in three or more groups which are clocked by overlapping pulses, in a manner similar to the PLO clocking. An experimental unit[18] has been made to demonstrate this type of logic. It contains a storage loop, a full adder stage, means to shift the right and left into registers and other functions including inversion. The unit contains 27 tunnel diodes with peak currents of 2 ma. Observed switching times with these early relatively slow diodes were about 50 nanoseconds. The unit was driven by a three-phase clock at 1 mega-cycle.

An alternative to the pulsed power supply for energizing the bistable elements is the use of a DC supply and resetting pulses. Any element of the network is set to state "1" directly according to the combined outputs of preceeding elements and is reset to state "0" by a clocking pulse. Setting of the elements of a group occurs immediately following resetting.

(2) In the second bistable type the logic element is made of two tunnel diodes in series (Fig. 17). To these diodes is applied a voltage from an a.c. source, - pulsed or sine wave, - of an amplitude sufficient for one diode to be in the high "1" state but insufficient for both to be in that state. The polarity of a relatively small input voltage applied to the mid-point is sufficient to determine which diode will trigger. This triggering will cause a greater voltage swing of the same polarity at that point and in effect will amplify the input signal. Logic networks can be obtained by arranging the logic elements in three or more groups and resistively coupling the midpoints of the elements of different groups.

The pattern of connections determines the desired logic operations. The groups are energized in succession by overlapping voltage waves in a manner similar to the clocking of the PLO's. A particularly simple system is the use of a 3 phase sine wave power supply. Negation is obtained by inverting transformers. Practical obtainable logic gains depend on matching of diodes in pairs rather than general uniformity of diodes as in the single-diode logic elements.

(3) Monostable logic elements are obtained with a tunnel diode in series with an inductance biased to a point P near the maximum of the characteristic (Fig. 18). A relatively small voltage can trigger the diode to the high state and thereby produce a relatively high voltage swing at the same point. Logic networks can be obtained by resistive couplings between logic elements in a manner similar to bistable elements. Resetting to the low state is produced automatically by the voltage induced in the inductance. Asynchronous operation can thus be obtained. In an experimental 7 stage delay chain, very uniform successive triggering was observed. Synchronous operation is possible also by superimposing clocking pulses on the inputs but it appears that higher speeds are realizable for a given logic gain in bistable circuits in that mode of operation.

All three types of circuits utilize logic elements in which the inputs and outputs are on the same single terminal. To insure separation of input and output functions, i.e., direction of information flow, several methods are possible. In the above described systems separation in time was used by multiple phase clocking, the inputs at any one logic elements being effective at a different time than the outputs. Directionality by electrical separation can be obtained through the use of unidirectional couplings. Unfortunately adequate rectifying diodes of speed comparable to that of the tunnel diodes are not available at the present time.

Directionality can be achieved also by making the level of energy of successive logic elements in a chain or the couplings between them progressively weaker so as to insure that the setting influence of the inputs dominates over the backflow influence of the outputs.

The multiple phase clocking is a simple solution to the directionality problem and the key to the successful use of single port two terminal devices. An increase of speed can be realized if several logic steps instead of a single one are made at each clock pulse. This is possible by using a cascaded arrangement of logic elements driving each other asynchronously. In these circuits means for directionality of information flow other than clocking must be provided.

From the above considerations it is evident that simple logic circuits of complete generality can be made from tunnel diodes. The signals are direct pulses and require no carrier. Energization can be either by multiphase sine wave or pulsed ac or by a combination of dc and clocking pulses. Experimental circuits have demonstrated gerneral system flexibility. The speed of the circuit can be very high. With early relatively slow experimental diodes having 2 ma peaks, one megacycle repetition rates were demonstrated but switching times were short enough to permit 10 megacycle rates. With newer faster 20 ma units logic elements were switched in times permitting 100 megacycle rates. The speed capabiltites of tunnel diodes are still being increased dramatically so that there is great promise for realizing logic circuits with 1000 megacycle rates.

3. Tunnel Diode Memory

Random access memories can be made using arrays of tunnel diode and give promise to be very fast. Each bit is stored by a current driven tunnel diode

having two stable voltages (Fig. 19). Row and column buses are resistively coupled to each diode. (Fig. 20). Any selected diode can be set to one or the other state by voltage pulses of appropriate polarity and amplitude on the corresponding buses. The memory can be organized for coincident bit addressing requiring two-to-one selection discrimination or for word addressing needing only three-to-one discrimination. The maxima and the minima of the characteristics provide the necessary thresholds and are sufficiently uniform to make possible either of these coincident write-in systems.

Read-out is obtained by driving the selected element or elements to the high state "1" and observing whether or not switching results. In a second following writing cycle the elements which have changed state are restored by appropriate control of the writing circuits in a manner analogous to that used in conventional core memories. The read-out signal can be obtained by direct pick-up from a common circuit resistively coupled to all elements. Read-out can be obtained also through inductive or radiative pick-up of high frequency which can be generated by the selected diode in several ways. A resonant circuit, which may be a simple stub at microwave frequencies, is associated with each element. In one method, to read, a write "0" is applied to the selected element and thereby switches it or not. If there is switching the relatively large voltage excursion through the negative region of the characteristic shock excites the tuned circuit and the resulting natural frequency oscillation is sensed on a circuit loosely coupled to all elements. (Fig. 21) In another method, selective readout addressing circuits impress signals at frequency f_1 on the selected row bus and f_2 on the selected column bus. In the high state the curvature of the voltage current characteristic is about 4 times greater than

in the low state, producing a corresponding ratio of amplitudes of the beat frequency $f_1 - f_2$ to which the elemental circuits are tuned.

Experiments with small arrays and array skeletons have demonstrated (1) two-to-one coincident write-in with reasonable tolerances of operation despite the use of experimental diodes with relatively wide variations of characteristics (2) adequate discrimination direct read out pulse signals with word addressing. (3) inductive read-out signals of high discrimination with ringing frequencies as high as 250 mc and beat frequency of about 1 kmc with bit coincident addressing.

Drivers of tunnel diode arrays must be able to supply pulses of relatively large power. The required current is large because all parallel connected half-selected elements load the selected lines, and the required voltage is large because the voltage of the series current regulating resistance must be several times greater than the voltage swing of the diode. Typically hundreds of milliamperes and several volts must be provided. It is unlikely that transistors will be adequate to drive large arrays at high speed (although a line of a arrays was driven in less than 10^{-8} sec.). The best promise for solving the driver problem lies in the tunnel diode itself. Sufficient voltage can be obtained by connecting a number of tunnel diodes in series, and adequate current may be obtained by using sufficiently high current units. Such arrangements have been operated and appear adequate.

Tunnel diode random access memories offer at the present time a good and possibly the only promise for achieving a cycle time of 10^{-8} seconds necessary of a memory associated with 1000 megacycle rate logic: the tunnel diodes themselves are or about to be fast enough and there does not seem to be any insurmountable

system problem. Experiments to date have demonstrated the essential write-in and read-out steps and have indicated a solution for the drivers. Furthermore it appears that propagation delays along addressing lines can be kept low enough so as to be insignificant. This results chiefly from the small size of the diode and of the resulting array.

IV. CONCLUSIONS

Computer logic by microwave carrier techniques and the junction diode PLO has been demonstrated in elementary subsystems operating at one hundred megacycles, and single elements switching in times corresponding to 300 megacycles. Microencapsulated improved diodes promise to provide thousand megacycle rates with power supplied at 30 KMC or higher.

Computer logic by tunnel diodes, already demonstrated at low rates, promises to be possible at thousand megacycles. Sufficiently uniform diodes and diodes capable of fractional nanosecond switching have been made. Random access memories with cycle times of the order of 10 nanoseconds appear possible using arrays of tunnel diodes. Tunnel diode computers operate with direct pulses and are powered by DC and/or AC at signal frequency.

Two-terminal semiconductor devices provide thus the manipulative elements required to gate, store, and amplify binary signals in nanoseconds. Furthermore, simplicity, small size, and power dissipation per element of tens of milliwatts permit packing of 10 or more elements per cubic inch. Therefore reasonably comprehensive computers with several thousand logic and several tens of thousands of memory elements can be made in a volume of less than two feet in diameter. Un-

avoidable delays due to signal propagation, of about 6 to 8 inches per nanosecond in normal transmission lines, are thus kept at about one nanosecond. This presents no serious difficulty in a "thousand megacycle" machine which can be assumed to have elementary logic functions executed in about one nanosecond and a memory cycle time of 10 nanoseconds.

We can, therefore, look forward to a new era of billion-bit-per-second information handling machines which are likely to produce as large, if not larger, an impact on the information processing art as was produced a decade ago by the introduction of present million-bit-per-second machines.

V. ACKNOWLEDGMENTS

The work reported in this paper is the result of the joint efforts of a number of scientists and engineers of the Radio Corporation of America in its Laboratories, Electronic Data Processing, Tubes and Semi-Conductor and Materials Divisions. The author has attempted to present a summary of their ideas and experimental results. Detail papers by various authors have been published already and others are following. The work was supported by government contract NObsr-77523.

REFERENCES

1. Eichi Goto, "On the Application of Parametrically Excited Nonlinear Resonator", Denki Tsushin Gakkai-shi, (Oct. 1955)

2. John von Neumann, "Nonlinear Capacitance or Inductance Switching, Amplifying, and Memory Organs", U. S. Patent 2,815,488, (Dec. 3, 1957), assigned to IBM.

3. J. J. Stocker, Nonlinear Vibrations in Mechanical and Electrical Systems, Interscience Publishers, Inc., New York (1950).

4. W. J. Cunningham, Nonlinear Analysis, McGraw-Hill, New York.

5. K. L. Kotzebue, "A Semiconductor-Diode Parametric Amplifier at Microwave Frequencies", Stanford Electronic Laboratories Technical Report, No. 49, (Nov. 1958).

6. A. Uhlir, Jr., "The Potential of Semiconductor Diodes in High Frequency Communications", Proc. of the IRE, Vol. 46 (June 1958), p. 1099-1115.

7. J. M. Manley and R. E. Rowe, "Some General Properties of Nonlinear Elements - Part 1 - General Energy Relations", Proc. of the IRE, 44 (July 1956), p. 904-913.

8. W. R. Beam, D. J. Blattner and F. Sterzer, "Microwave Carrier Techniques for High Speed Digital Computing", Symposium on Microwave Techniques for Computers, Washington, D. C., March 12, 1959, Transactions of the IRE on Electronic Computers, Sept. 1959.

9. F. Sterzer and D. Blattner, "Fast Microwave Logic Circuits", IRE National Convention, March 1959 also Eastern Joint Computer Conference, Dec. 3-5, 1958.

10. F. Sterzer, "Microwave Parametric Sub-Harmonic Oscillators for Digital Computing", Proc. of the IRE, July 1959.

11. F. Sterzer, "RF Circuits Using Sub-Harmonic Oscillators", PGMIT National Symposium Harvard University, Cambridge, Mass., June 1959.

12. F. Sterzer and W. R. Beam, "Parametric Sub-Harmonic Oscillators", Digest of Technical Papers, Solid State Circuits Conference, Philadelphis, Pa., Feb. 1959.

13. L. S. Onyshkevych, W. F. Kosonocky and A. W. Lo, "Parametric Phase-Locked Oscillator - Characteristics and Applications to Digital Systems", Symposium on Microwave Techniques for Computers, Washington D. C., March 12, 1959. Transactions of the IRE on Electronic Computers, Sept. 1959.

14. J. Hilibrand, C. W. Mueller, C. F. Stocker and R. D. Gold, "Semiconductor Parametric Diodes for Microwave Computers", Symposium on Microwaves Techniques for Computers, Washington, D. C., March 12, 1959, Transactions of the IRE on Electronic Computers, Sept. 1959.

15. L. Esaki, Physical Review, 109 (1958), 60-s.

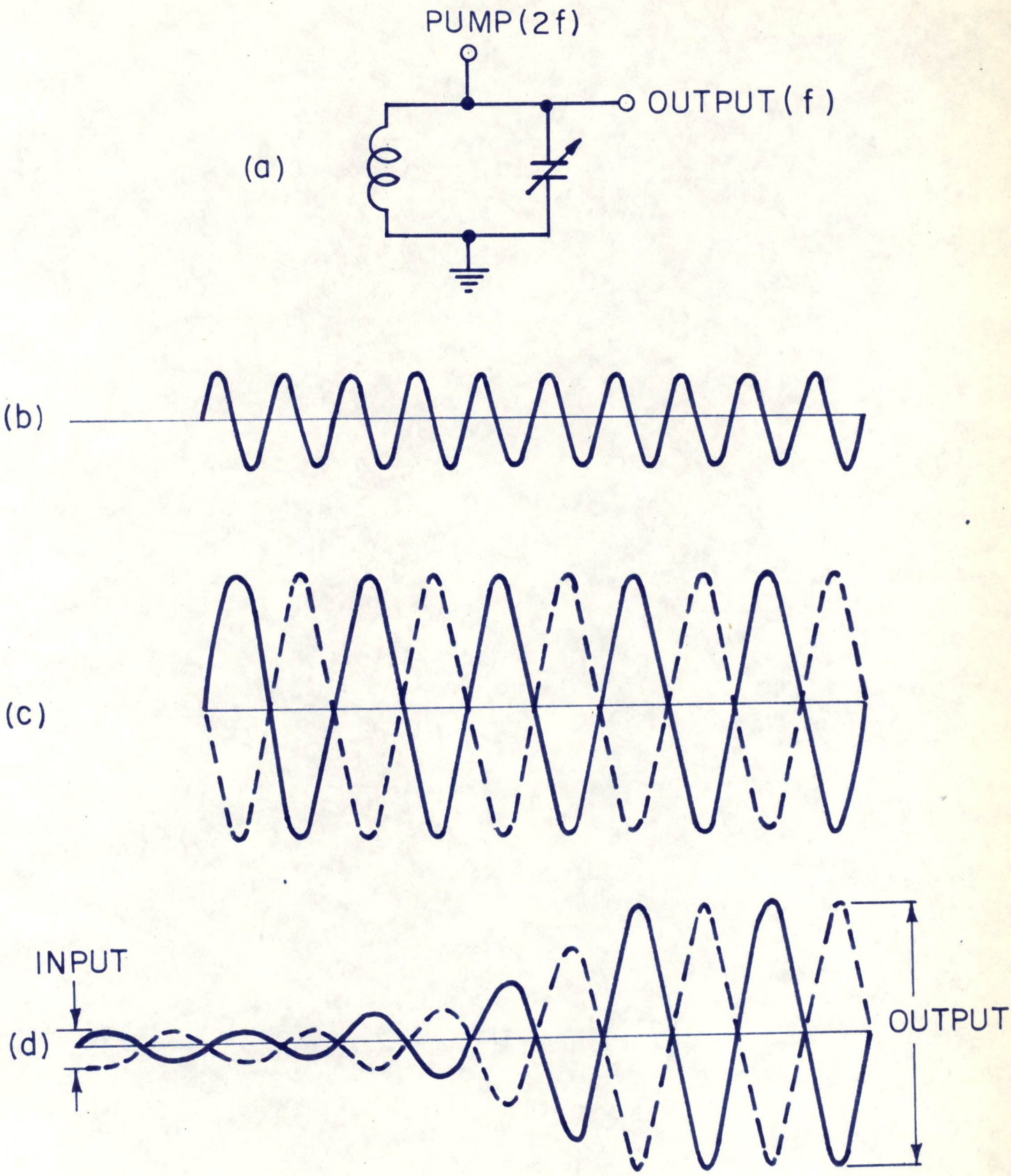
16. H. S. Sommers, Jr., "Tunnel Diodes as High Frequency Devices, "Proc. of the IRE, (July 1959), p. 1201.

17. R. L. Wigington, "A New Concept in Computing", Proc. of the IRE, Vol. 47 - No. 4, (April 1959), p. 516-523. (An account of J. von Neumann ideas - reference 2).

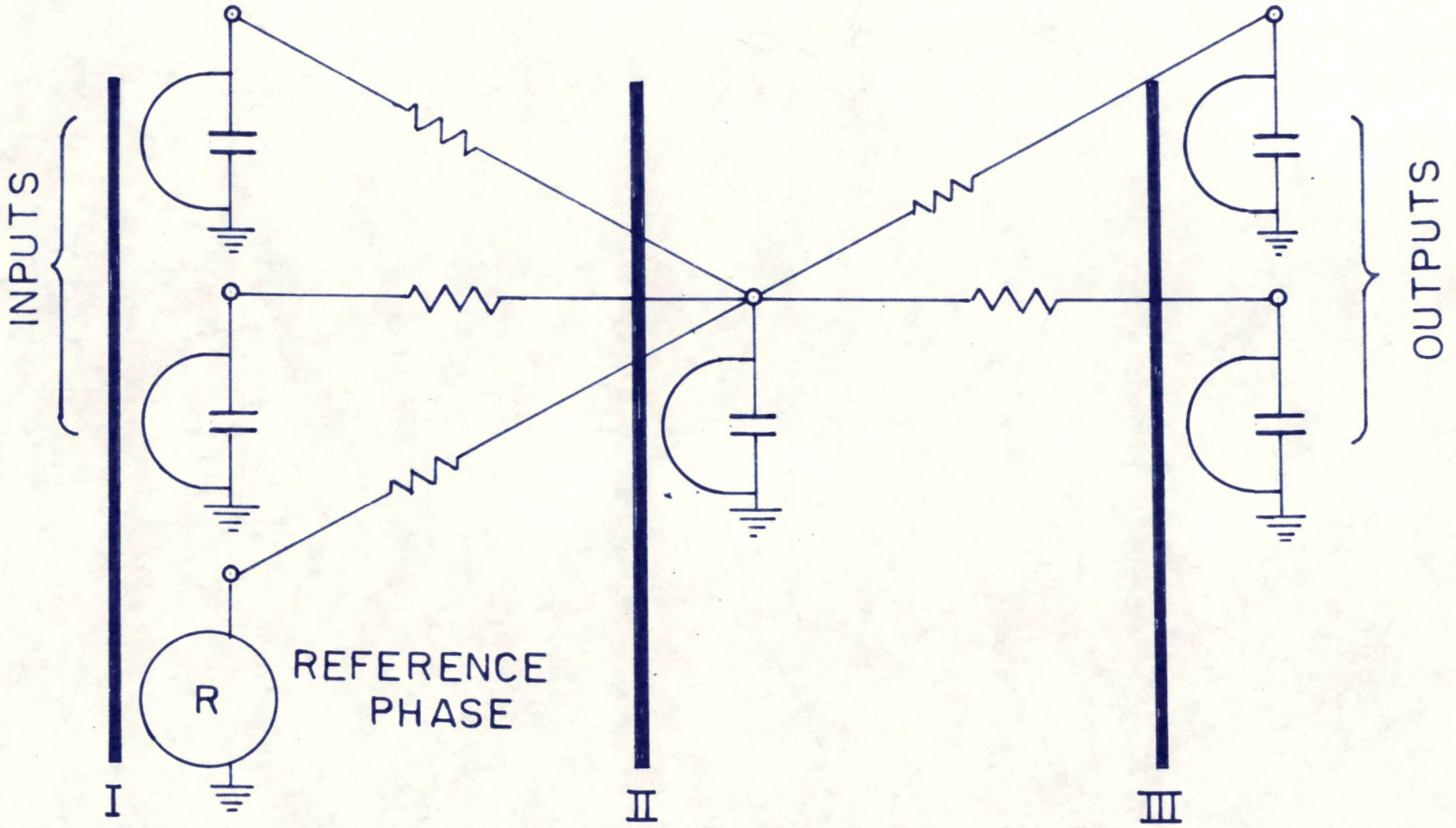
18. M. H. Lewin, "Negative Resistance Elements as Digital Computer Components", Paper of this 1959 EJCC.

19. R. N. Hall, "Tunnel Diodes", 1959 Electron Devices Meeting, October 29, 1959, Washington, D. C.

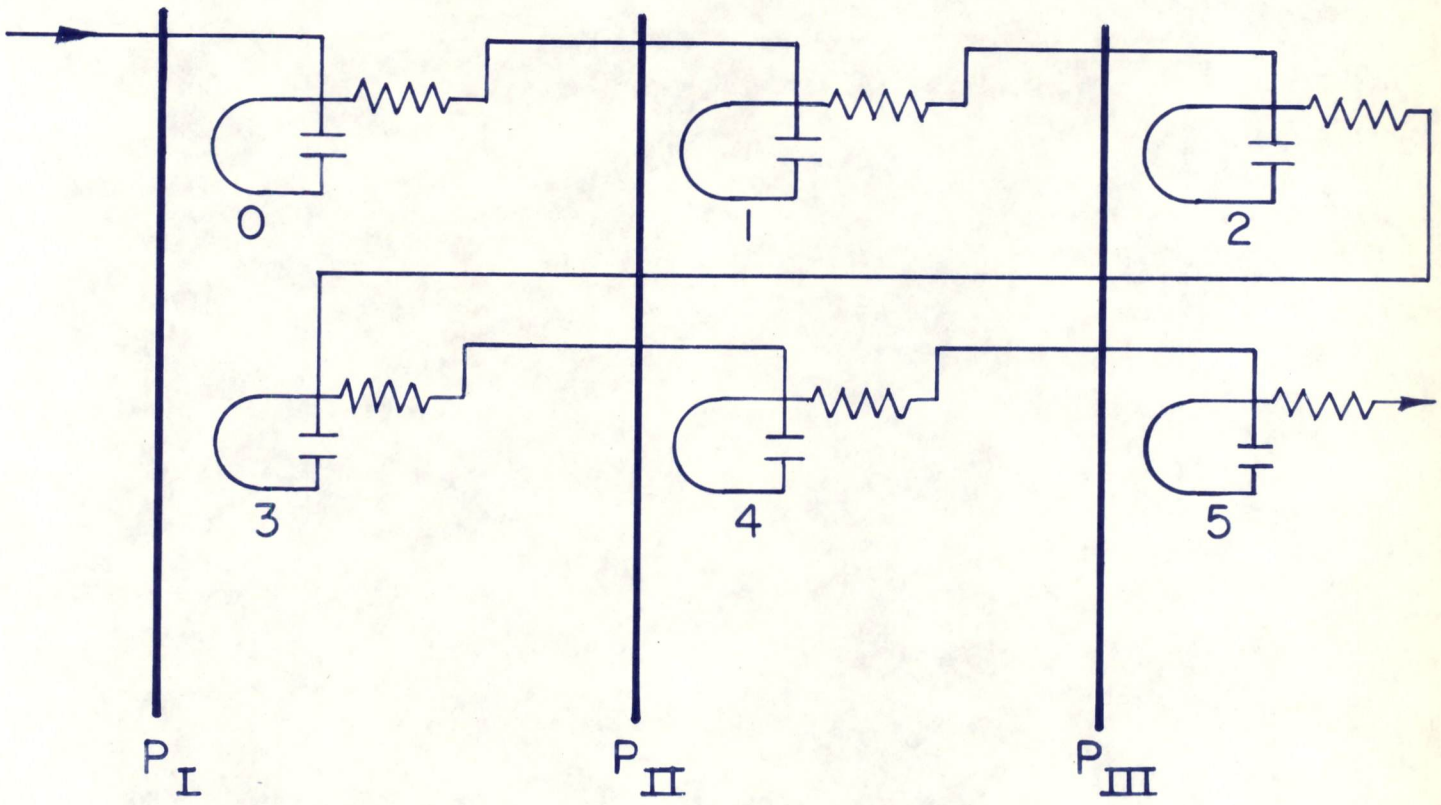
List of Figures and Captions

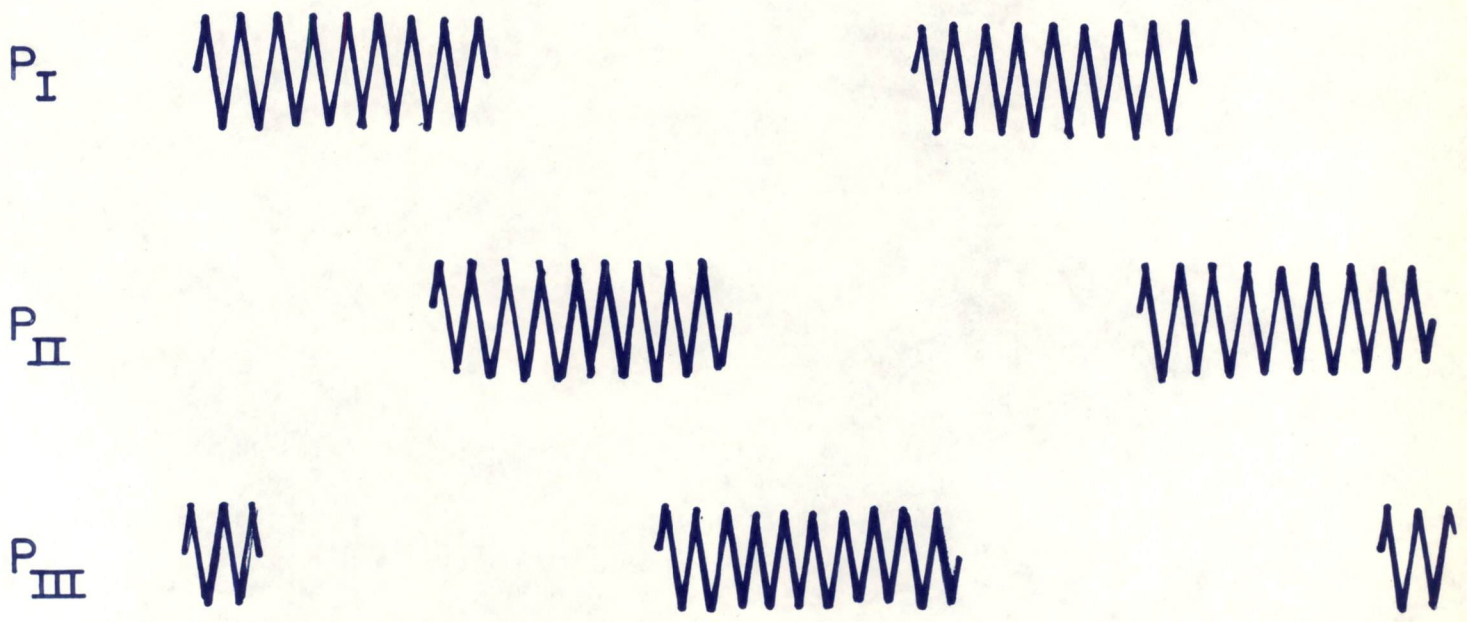


PRINCIPLE OF PARAMETRIC PHASE-LOCKED OSCILLATOR



PLO GENERAL LOGIC





SHIFT REGISTER WITH THE THREE PHASE PUMP SYSTEM.

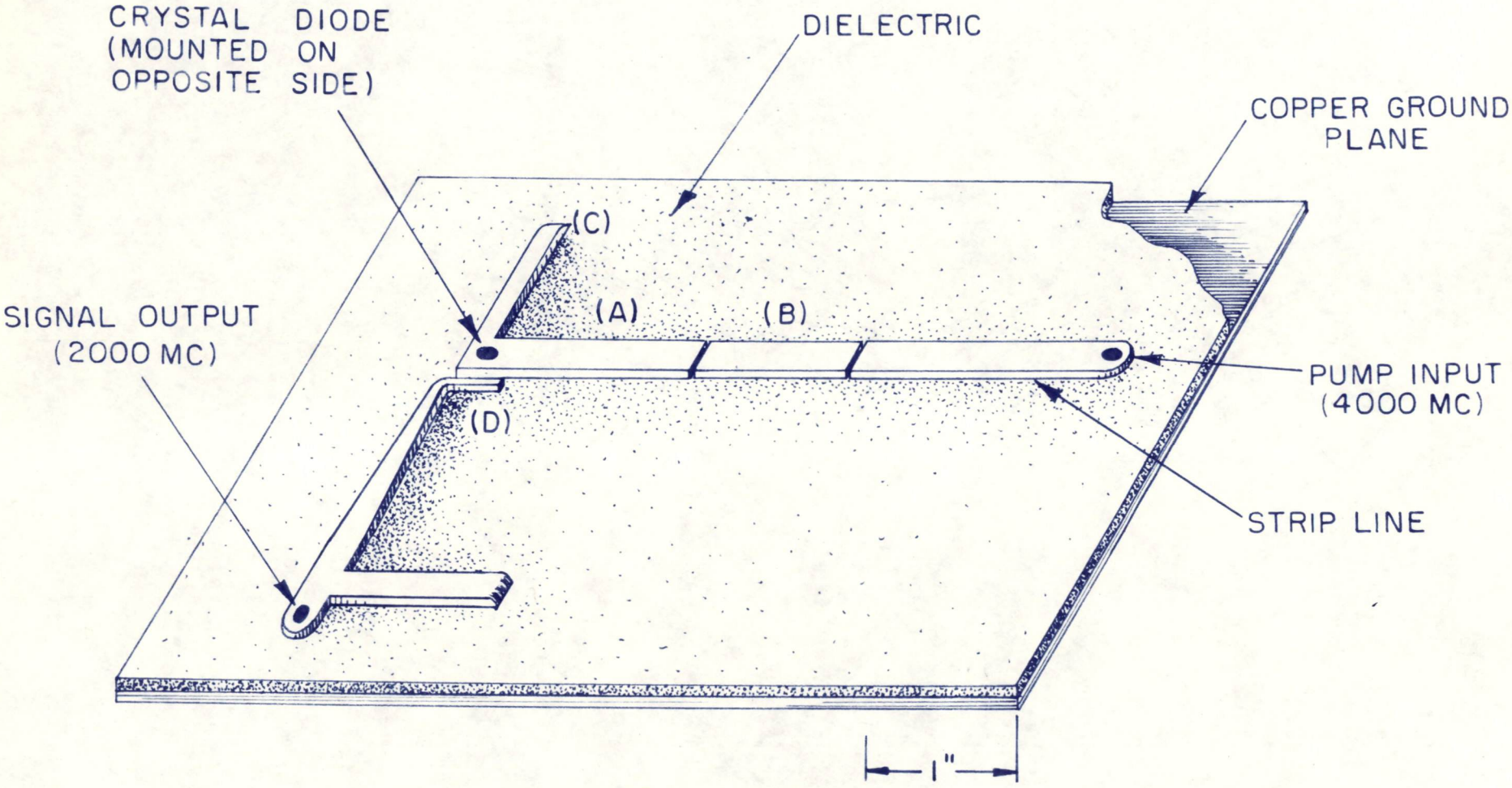
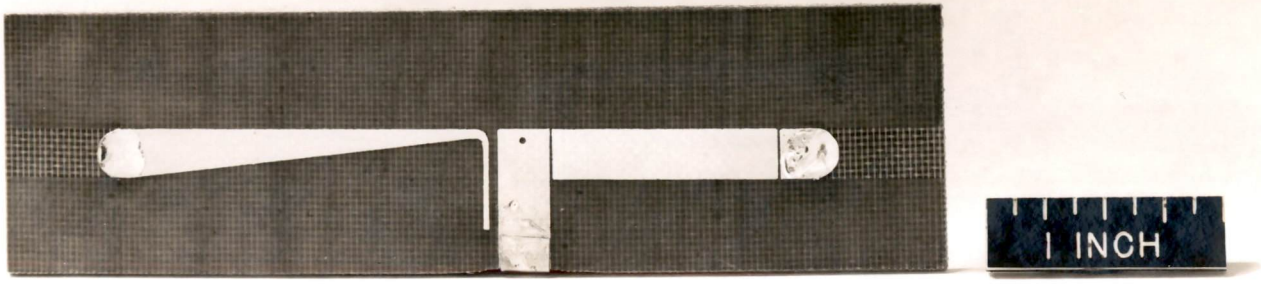
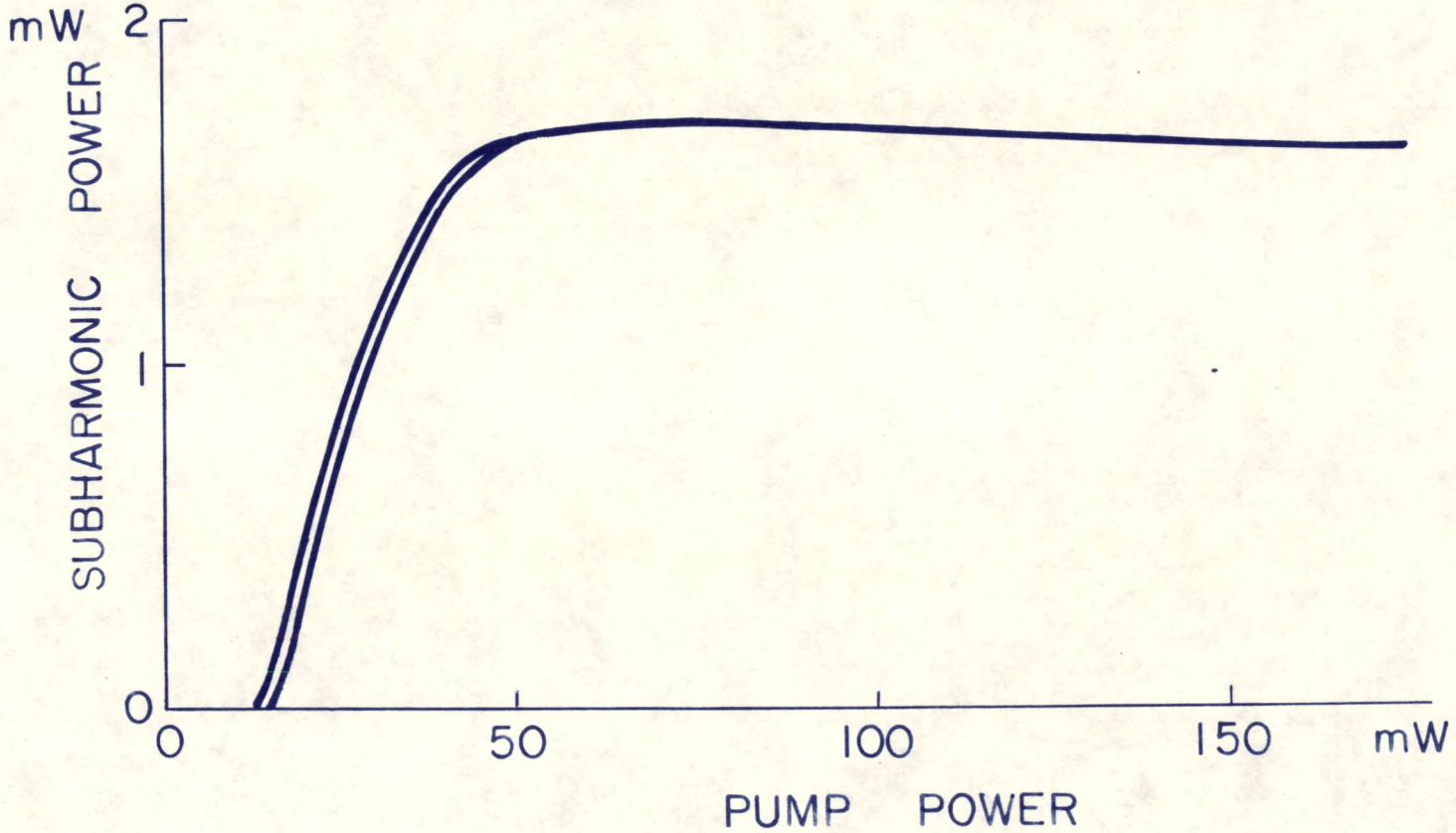


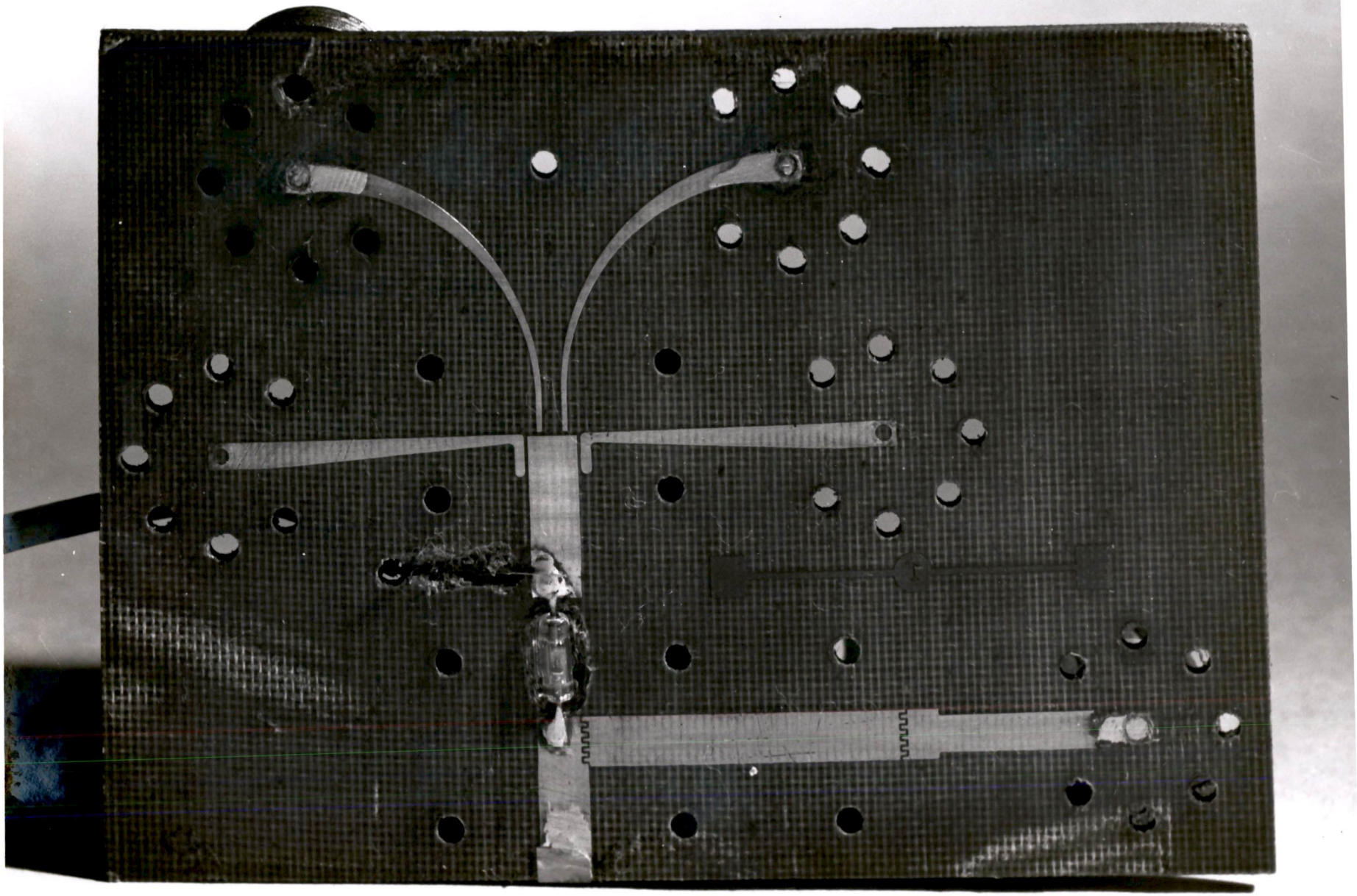
FIG. 2 MICROWAVE SUBHARMONIC
PHASE LOCKED OSCILLATOR



11234

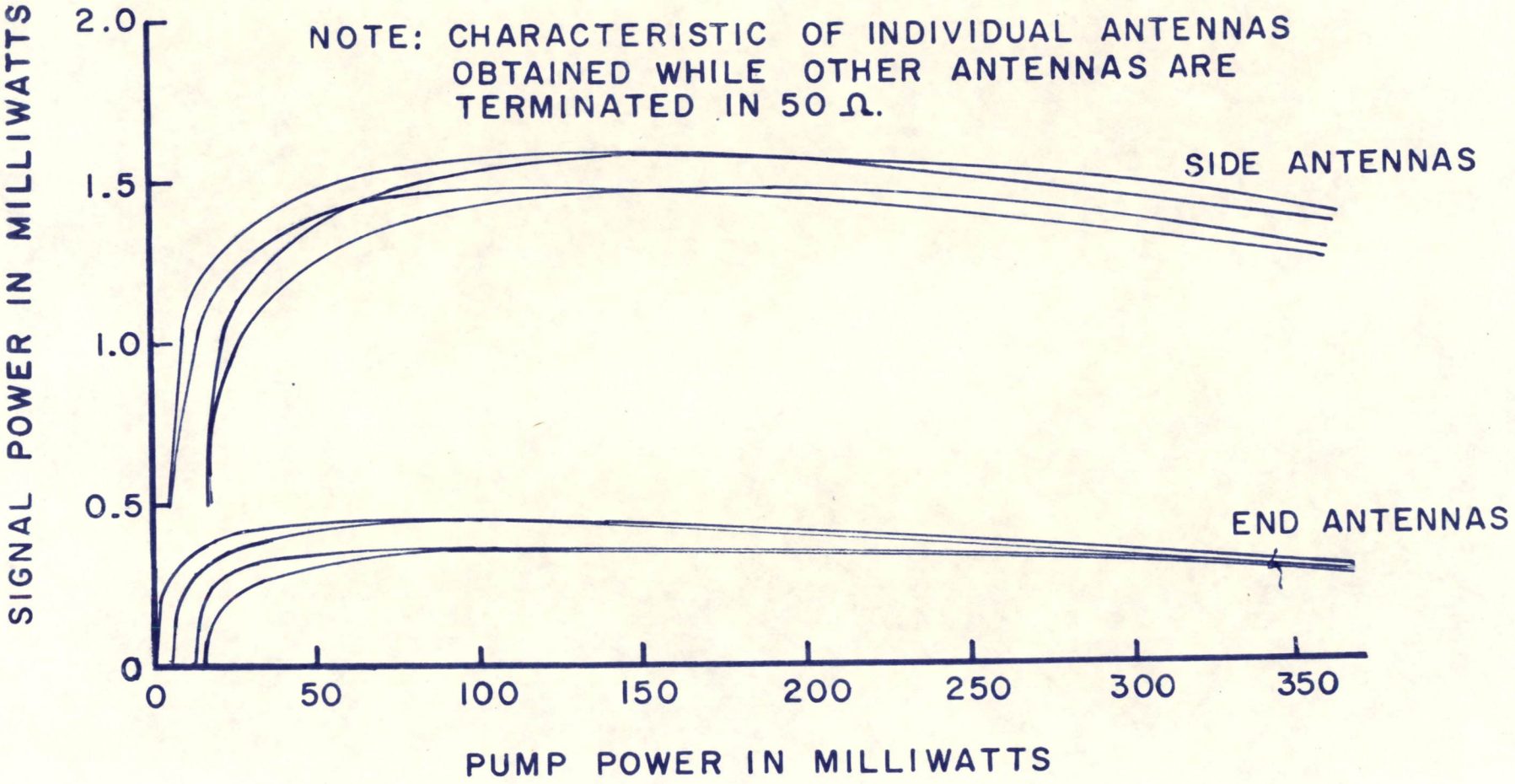


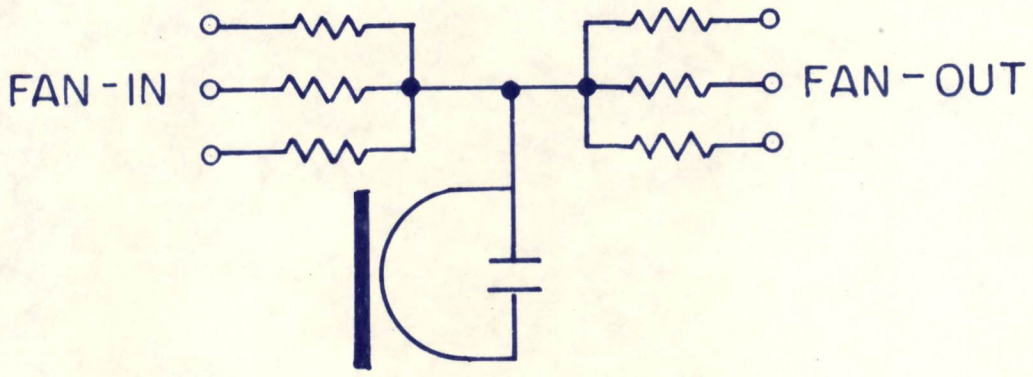
PLO CHARACTERISTIC

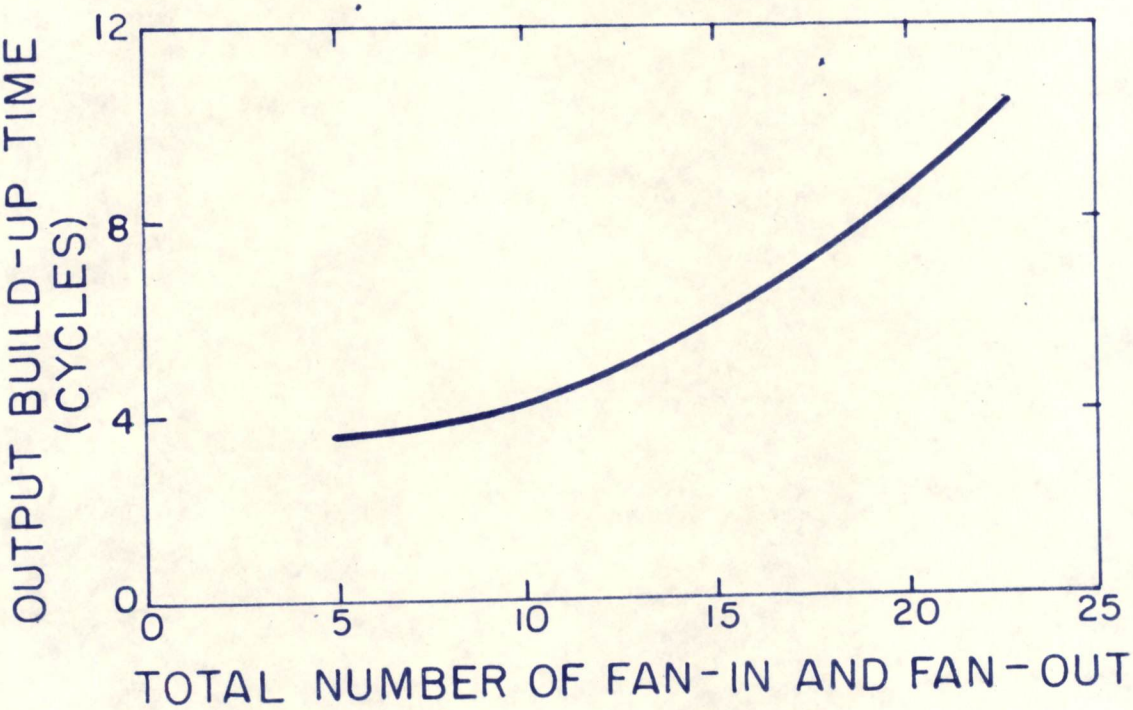


STATIC CHARACTERISTICS OF DESIGN #8586-48

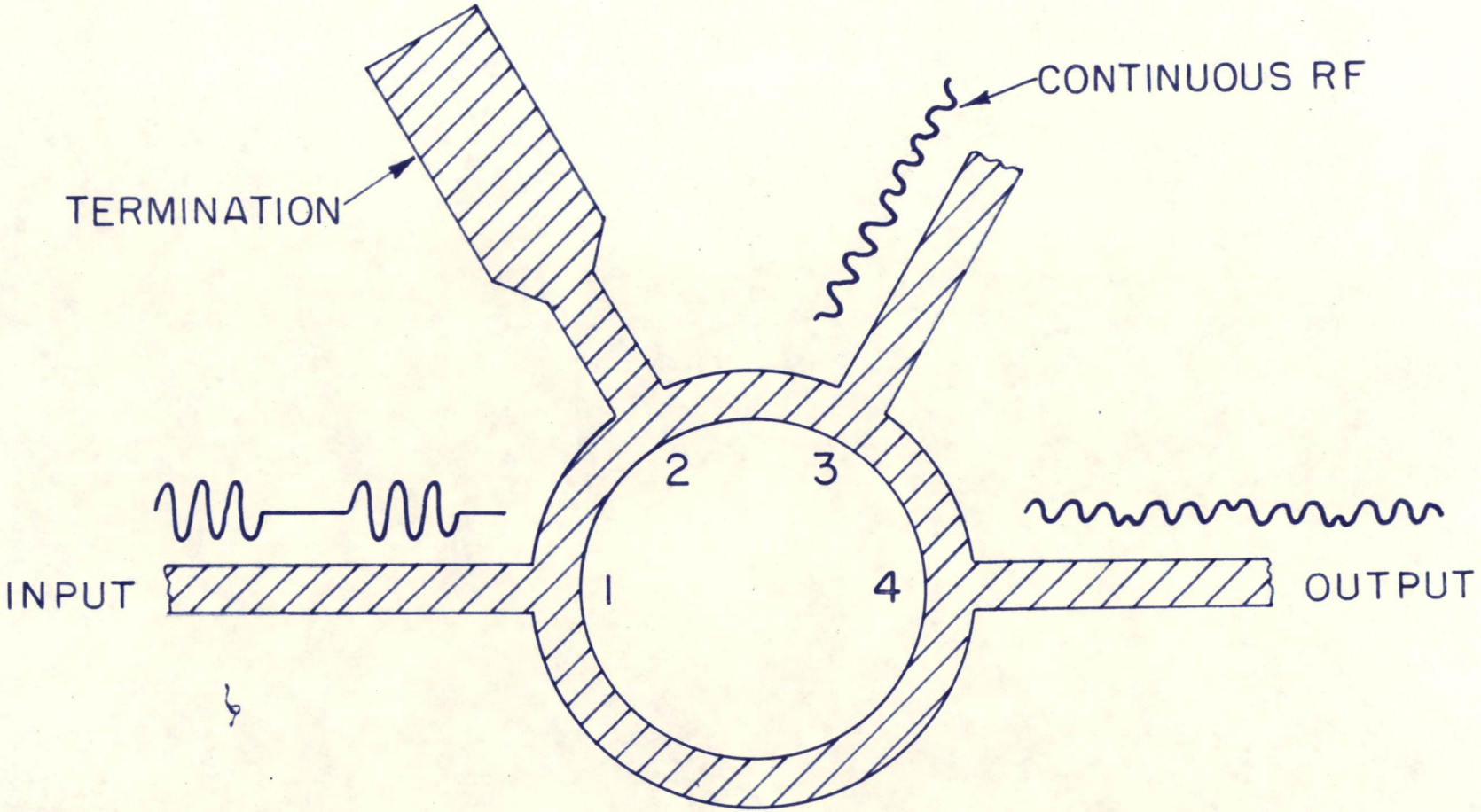
$R_b = 1K$, $V_b = 1.62$ V.



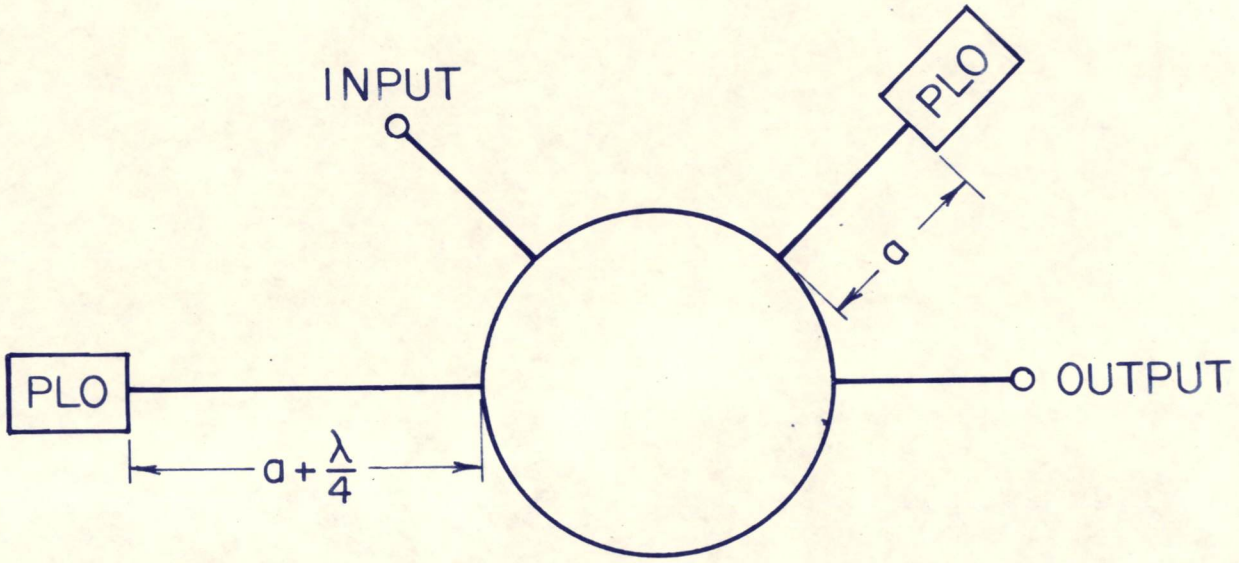




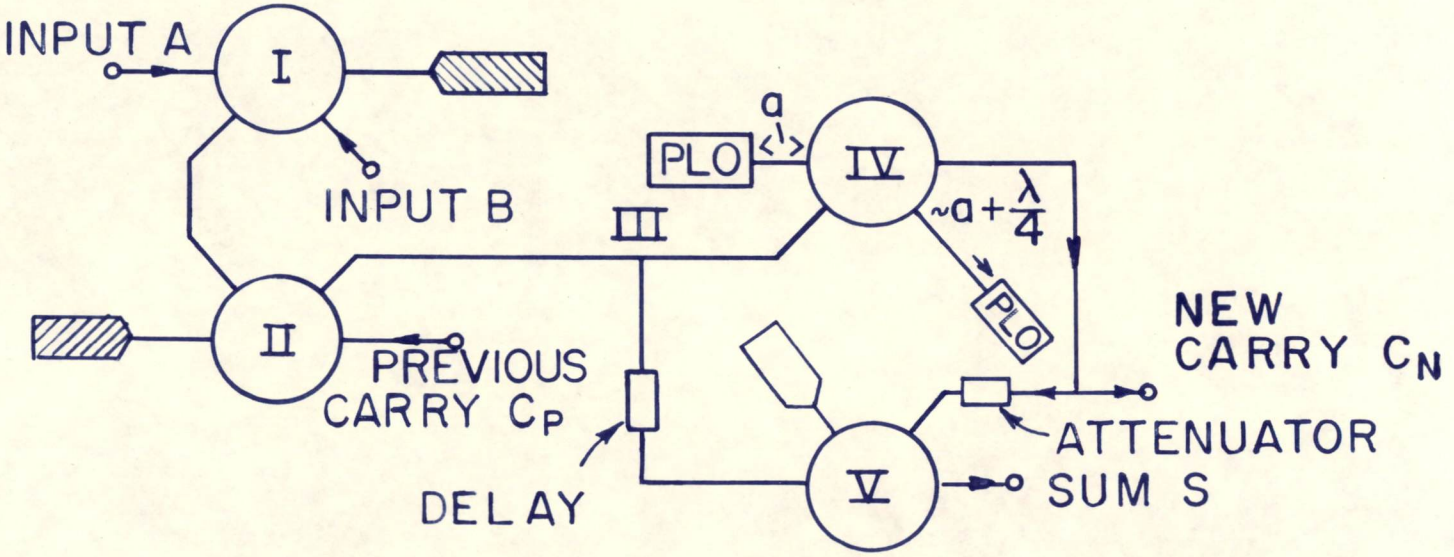
NUMBER OF CYCLES OF BUILD-UP AS A FUNCTION OF LOGIC GAIN



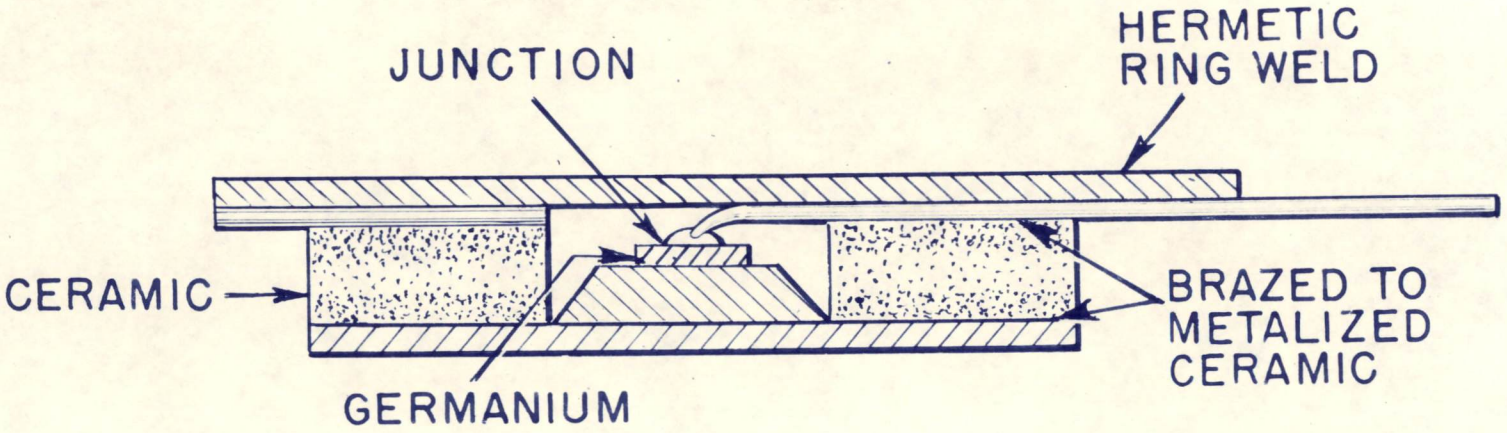
HYBRID CIRCUIT FOR TRANSLATING
AMPLITUDE-TO-PHASE SCRIPTS



CIRCUIT FOR SEPARATING INPUT AND OUTPUT OF PLO.

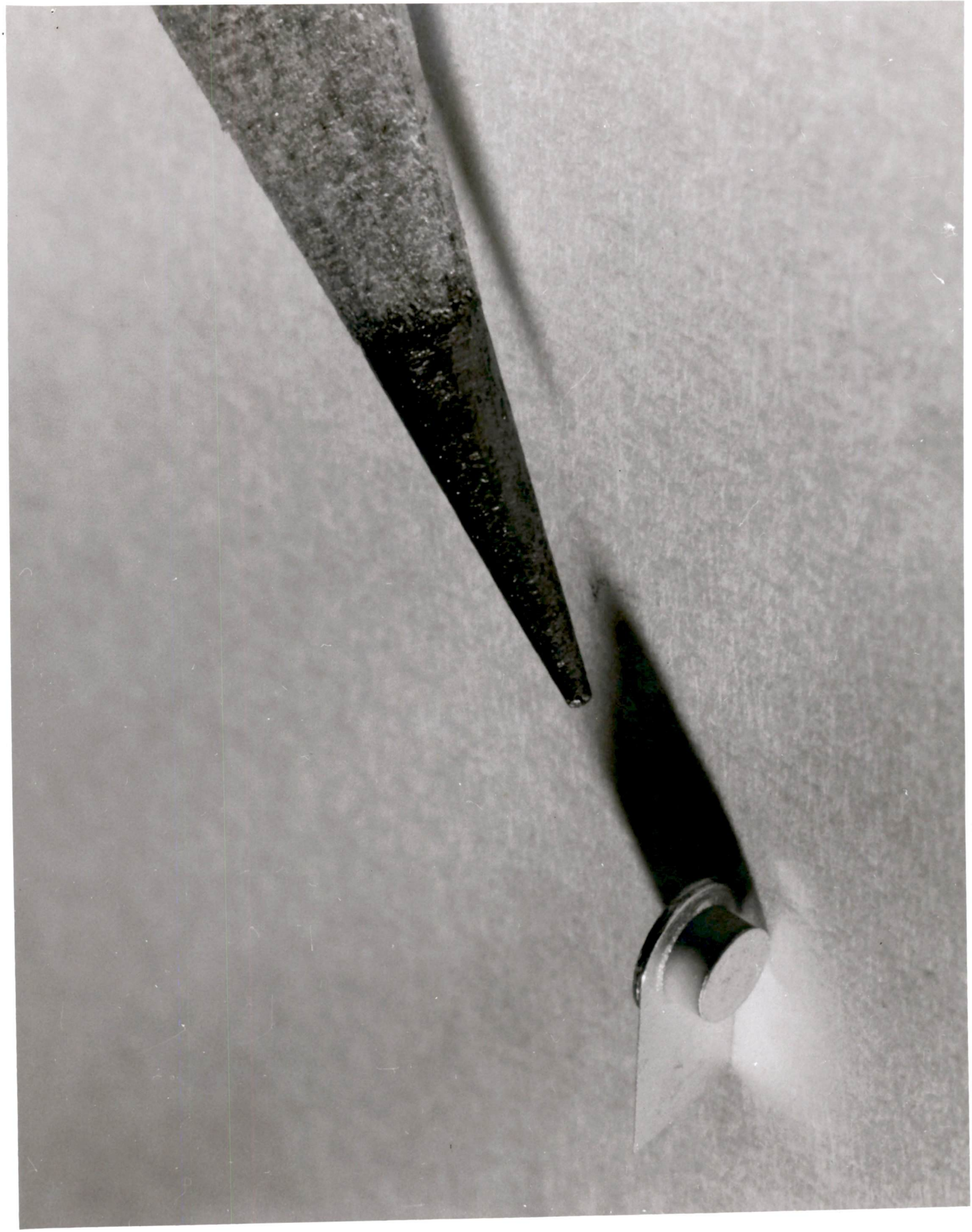


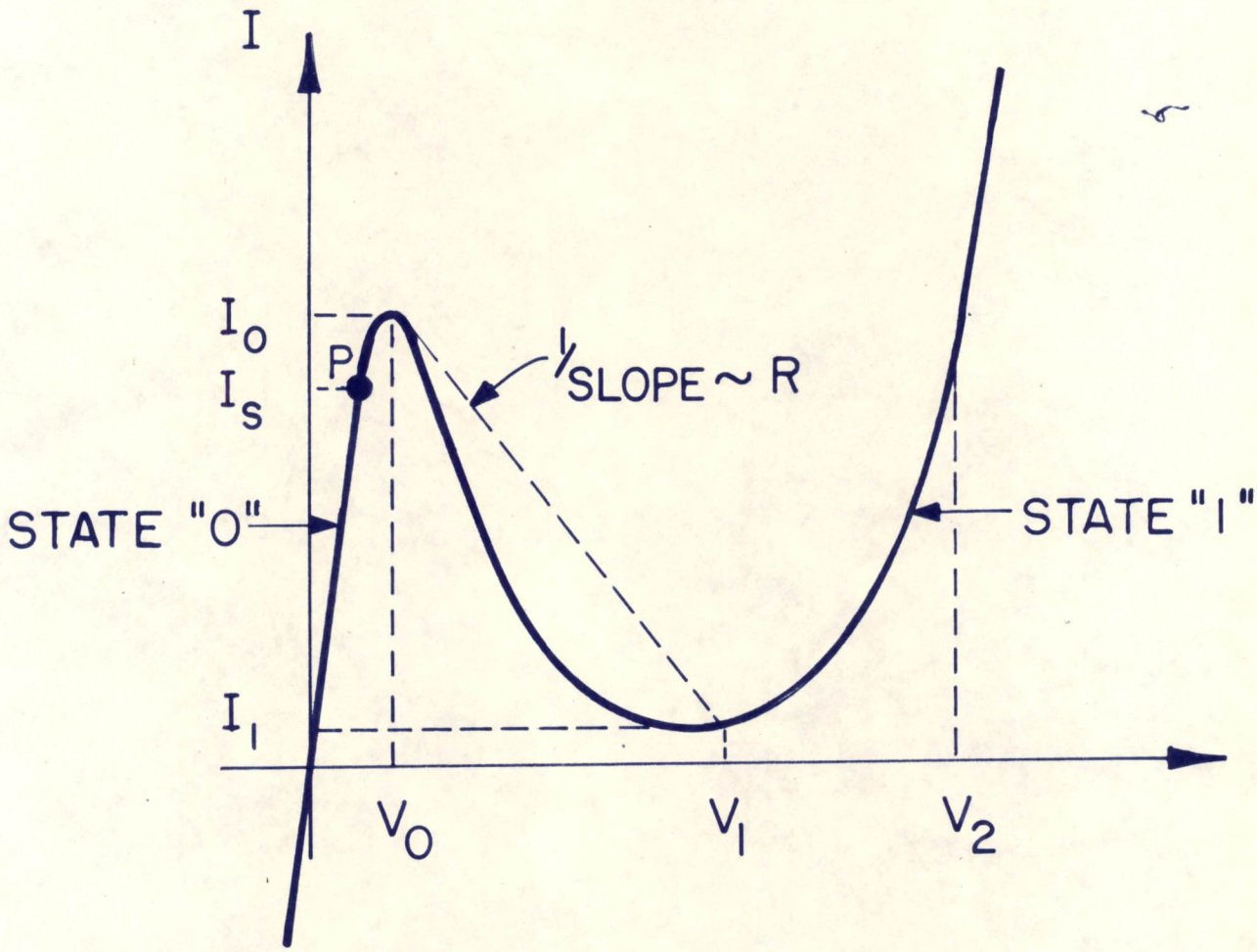
BINARY ADDER



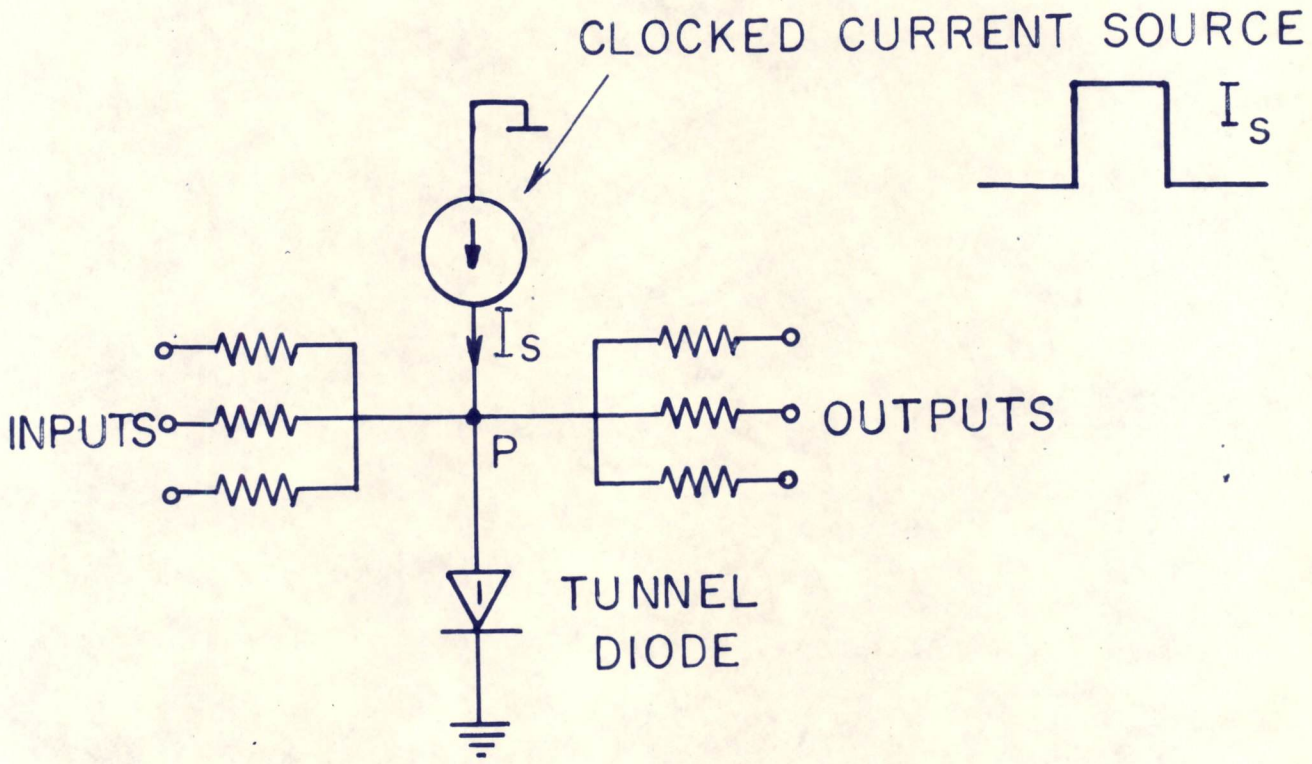
INDUCTANCE = .5 mμh.
SHUNT CAPACITANCE = .3 μμf.
RESISTANCE = .15 OHM @ 2 Kmc.

CROSS-SECTION OF DIODE

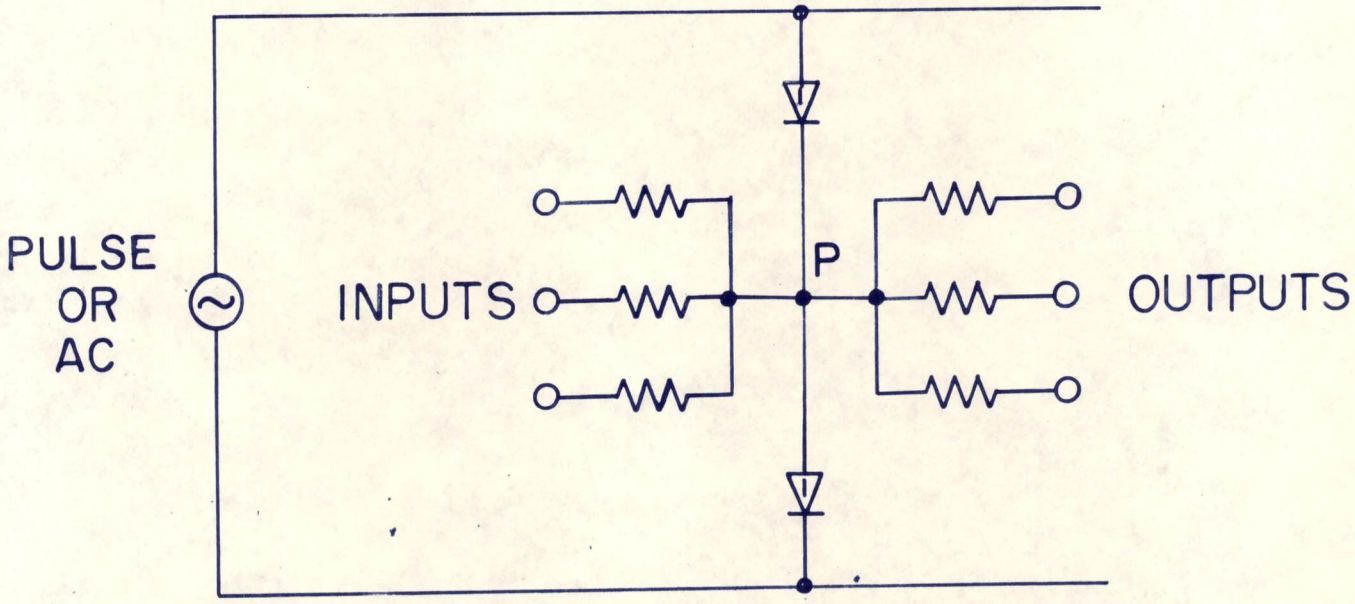




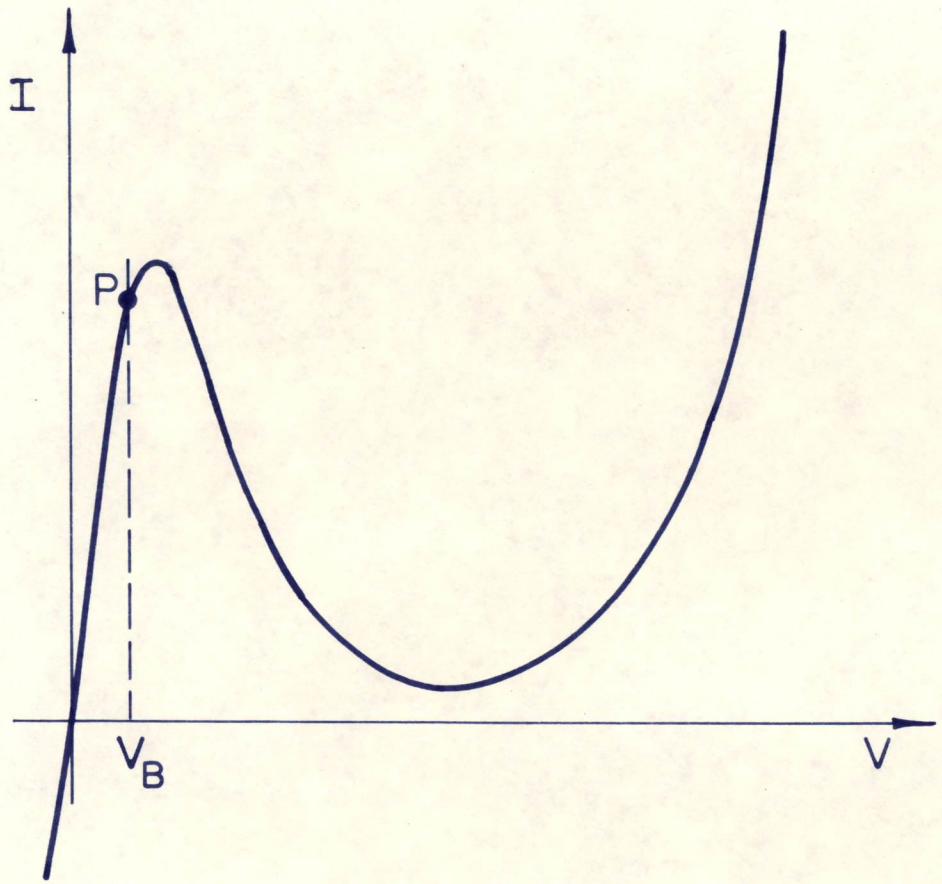
TUNNEL DIODE CHARACTERISTIC

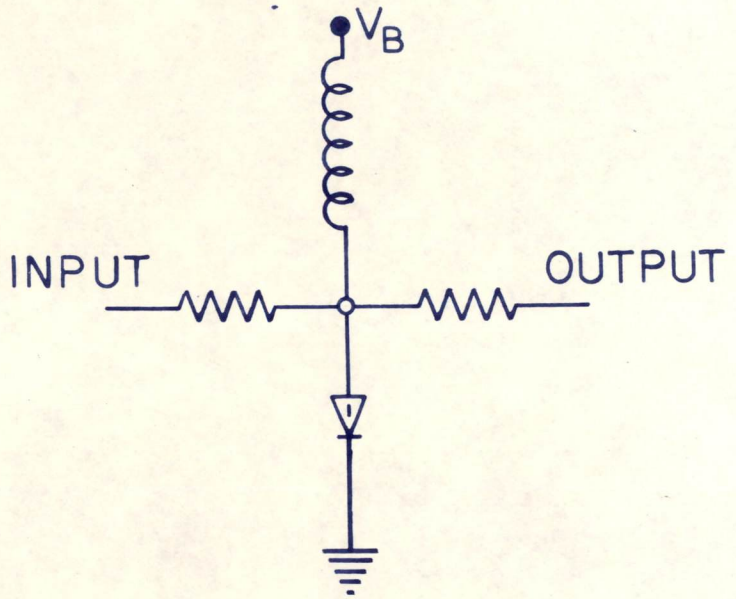


TUNNEL DIODE BISTABLE LOGIC ELEMENT
SINGLE DIODE TYPE

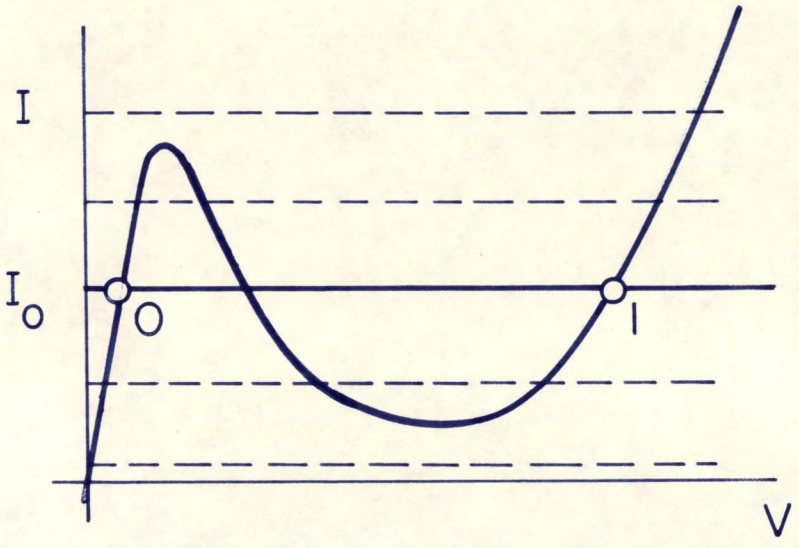


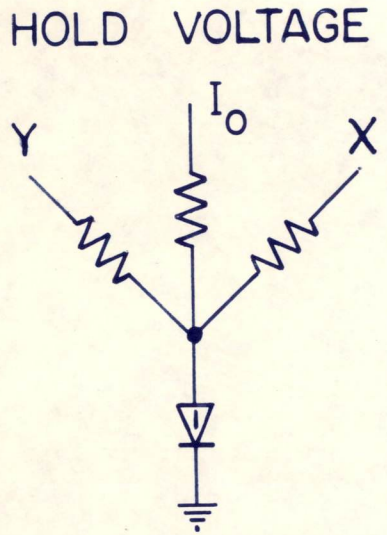
TUNNEL DIODE BISTABLE LOGIC ELEMENT
SYMMETRIC TYPE



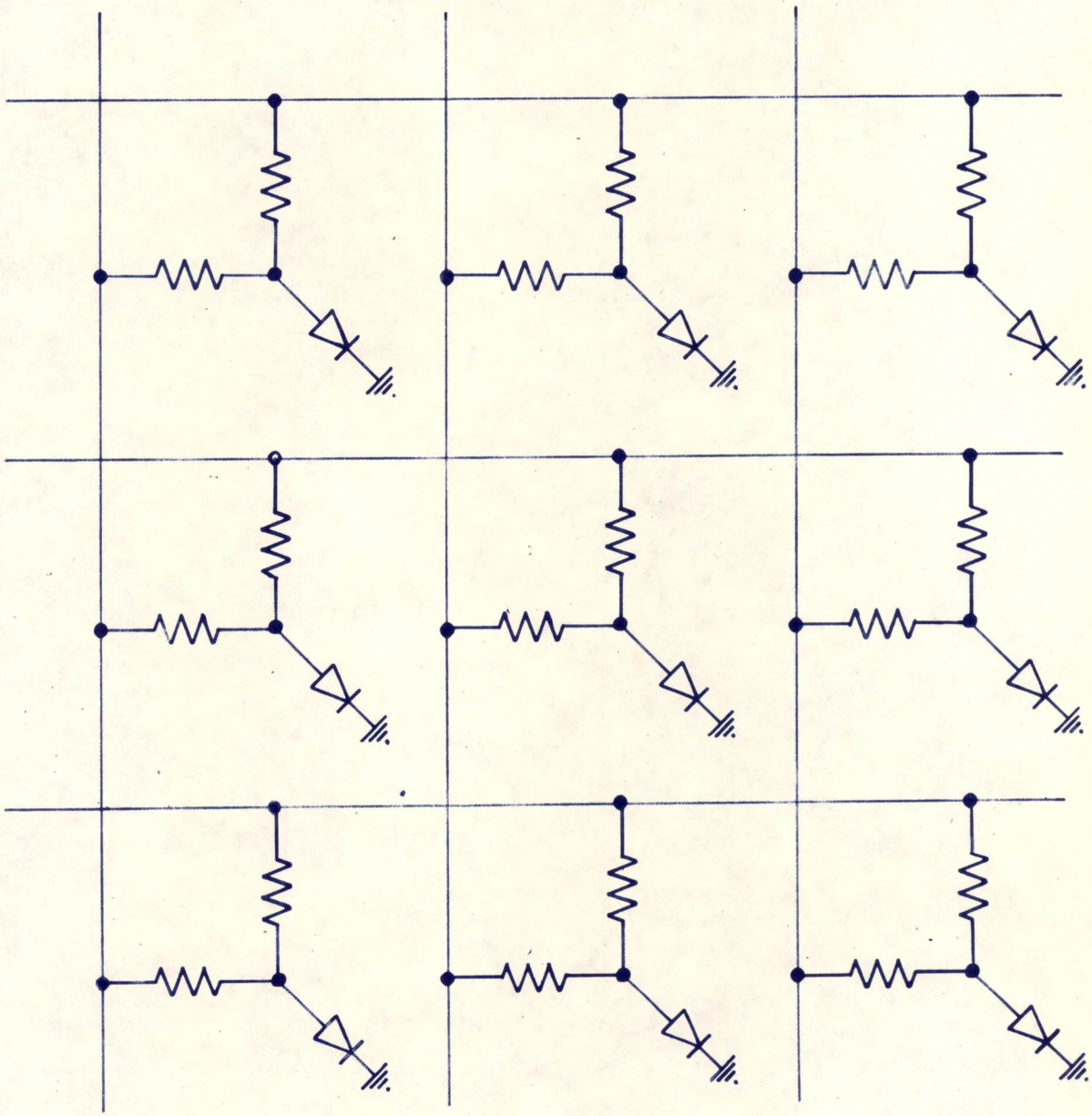


TUNNEL DIODE MONOSTABLE LOGIC ELEMENT

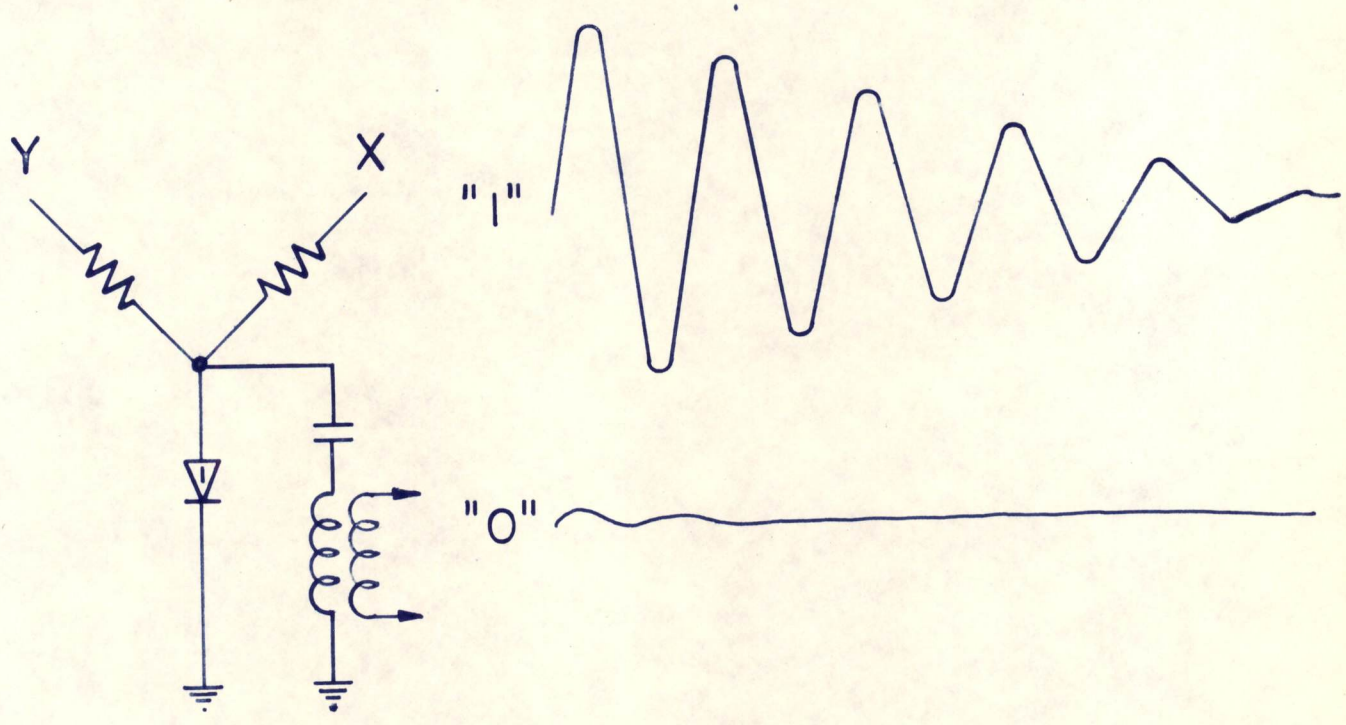




CURRENT COINCIDENCE WRITE - IN



TUNNEL DIODE MEMORY ARRAY



READ-OUT BY SENSING RINGING

RAMO-WOOLDRIDGE a division of Thompson Ramo Wooldridge Inc.

P.O. BOX 90534 AIRPORT STATION · LOS ANGELES 45, CALIFORNIA · OREGON 8-0511 · OSBORNE 5-4651

November 18, 1959

Mr. Harlan E. Anderson, Chairman
1959 EJCC Publication Committee
Digital Equipment Corporation
Maynaird, Massachusetts

Dear Mr. Anderson:

SUBJECT: Presentation of a paper at the EJCC.

We regret that we failed to send you a resume of the engineer reading our paper entitled "A High Speed, Small Size Magnetic Drum Memory Unit for Sub-miniature Digital Computers". Enclosed you will find a resume relating to Mr. R. A. Howard who will read the paper for us.

We are preparing two sets of slides, size 3 1/4" x 4 1/4" for use in the presentation.

Yours very truly,

M. May
Member of Senior Staff
Digital Control Department

Enclosures (1)

5500 WEST EL SEGUNDO BOULEVARD

SHORT BIOGRAPHY OF

Roy A. Howard

B.S. in Electrical Engineering at Oregon State College.
Post Graduate work in Mathematical Physics.
Member of Institute of Radio Engineers.
Ten years' experience in Digital Computer Development for Data Processing and Control purposes including Automatic Control of machine tools and semi-conductor test equipment.

While a Naval Officer, Mr. Howard was responsible to the Navy for overseeing a computer development at ERA. Later at the Hughes Aircraft Company he was responsible for system engineering and logical design of parts of a Data Processing System and a line of Automatic Machine Tool Controls.

RAMO-WOOLDRIDGE a division of Thompson Ramo Wooldridge Inc.

P.O. BOX 90534 AIRPORT STATION · LOS ANGELES 45, CALIFORNIA · OREGON 8-0511 · OSBORNE 5-4651

November 13, 1959

Mr. Harlan E. Anderson, Chairman
1959 EJCC Publication Committee
Digital Equipment Corporation
Maynard, Massachusetts

Dear Mr. Anderson:

Enclosed are four copies of the manuscript for our paper entitled "A High Speed, Small Size Magnetic Drum Memory Unit for Subminiature Digital Computers". An abstract is attached to each manuscript.

The illustrations made up as you requested (one complete set) are enclosed. Reproductions of the illustrations are also bound in the four copies of the manuscript. We hope this will make checking the manuscripts more simple.

Yours very truly,

Michael May

Michael May
Member of the Senior Staff
Digital Control Department

MM:jhb
Enc.

5500 WEST EL SEGUNDO BOULEVARD

ABSTRACT

A High Speed, Small Size Magnetic Drum Memory Unit for Sub-Miniature Digital Computers.

A magnetic drum memory unit has been designed and constructed for use in future sub-miniature digital computers. The unit has been designed to meet the requirements of MIL-E-5400 Class 2, withstand shock loads of 16g, and operate under continuous vibration of 10g peak, from 0-2000 cps. Its operating temperature range is - 54°C + 125°C. It will operate up to any altitude as normally packaged. The unit has a memory capacity of 300,000 bits. 122 tracks are provided on a 4 1/2" long 2 1/2" diameter recording drum. At the normal speed of 12,000 r.p.m., the clock frequency is 546 kc. The unit utilizes four 1.3" by 1.25" shoes to hold assemblies of 27 read-record heads, and smaller shoes to hold circulating register heads. Head to drum spacing is maintained by an air film between the shoe and the drum. Recording is achieved with silicon transistor circuits using 6 volts supply and 100 ma peak recording current. The signal level, when reading, is about 12 mv at 546 kc. The size of the unit is 3.7 X 3.7 X 7.4". Its weight is 11.3 lbs including a case for hermetic sealing.

This development was performed under contract for the Wright Air Development Center of the USAF.

A HIGH SPEED, SMALL SIZE MAGNETIC DRUM MEMORY UNIT FOR SUBMINIATURE DIGITAL COMPUTERS

M70-9U40

By

M. May

G. P. Miller

R. A. Howard

G. A. Shifrin

4 November 1959

RAMO-WOOLDRIDGE

A Division Of

THOMPSON RAMO WOOLDRIDGE INC.

LOS ANGELES 45, CALIFORNIA

CONTENTS

A HIGH SPEED, SMALL SIZE MAGNETIC DRUM MEMORY UNIT FOR SUBMINIATURE DIGITAL COMPUTERS

A memory with dimensions compatible with microminiature assemblies is required for future computers to be used in missiles and aircraft. A drum memory is described which can fulfill this need. The bit rate of 546 kc makes possible a 20-bit serial word time of the order of 40 microseconds. For a computer with add and multiply times of 40 microseconds, the drum memory described is adequate. Moreover, the technique described can be extended to provide a 20-microsecond word time by doubling the rotational speed of the drum, and to 10 microseconds or less by reading out two or more bits in parallel. A memory capacity of 15,000 twenty-bit words is available in the 7.4 x 3.7 x 3.7 inch total unit size, which is adequate for the type of computations usually made in an aircraft or missile. The advantage of such a drum memory as compared with a ferrite core memory, for example, is in cost, size, and ability to perform over wide temperature ranges. The disadvantage of the lack of immediate access to any address can for the most part be overcome by suitable programming precautions.

The magnetic drum development was performed under contract for Wright Air Development Center of the USAF to determine whether recording densities of 500 to 1000 bits to the inch and more than 30 tracks to the inch could be achieved in a small unit which would meet the requirements of MIL-E-5400, Class 2. The description of the development is broken down as follows:

1. General Description of Memory Unit
2. Selection of Magnetic Coating
3. Design of Suitable Read-Record Heads
4. Mechanical Design Details

GENERAL DESCRIPTION OF MEMORY UNIT

Size	3.7 x 3.7 x 7.4 inches over-all
Power	400 cps, 3 phase, about 30 watts
Weight	11.3 pounds
Motor	Mounted inside recording drum
Tracks	30 per inch, total of 122 tracks
Recording density	350 bits per inch using Manchester phase modulated recording
Clock frequency	546 kc
Total storage capacity	300,000 bits plus timing tracks and spare tracks
Speed	12,000 rpm approximately

Figure 2 shows a partly assembled unit. To achieve the high degree of stability required for high density recording over a wide temperature and vibration range, an especially rigid unit was constructed. The framework and most critical parts are made of stainless steel selected to have a coefficient of expansion to match that of the ball bearings. A cross section drawing of the rotating part of the unit is shown in Figure 1.

The recording drum is made up of an internally mounted, 400-cps, 3-phase induction motor whose stator (1) is attached to a fixed shaft (3). The squirrel cage type rotor (2) is fixed inside a steel cylinder (8) which provides magnetic shielding and forms a mounting for a nonmagnetic stainless steel cylinder (9). This cylinder is plated with nickel-cobalt by an electroless method to form the recording surface. A shoe holding 27 read-record heads can be seen resting on the recording surface in Figure 2. This shoe is loaded with a 6-10 pound force against the recording surface when the drum attains full speed. Since the shoe and 27 heads weigh less than 1.5 ounces, accelerations of 10 g's have little effect

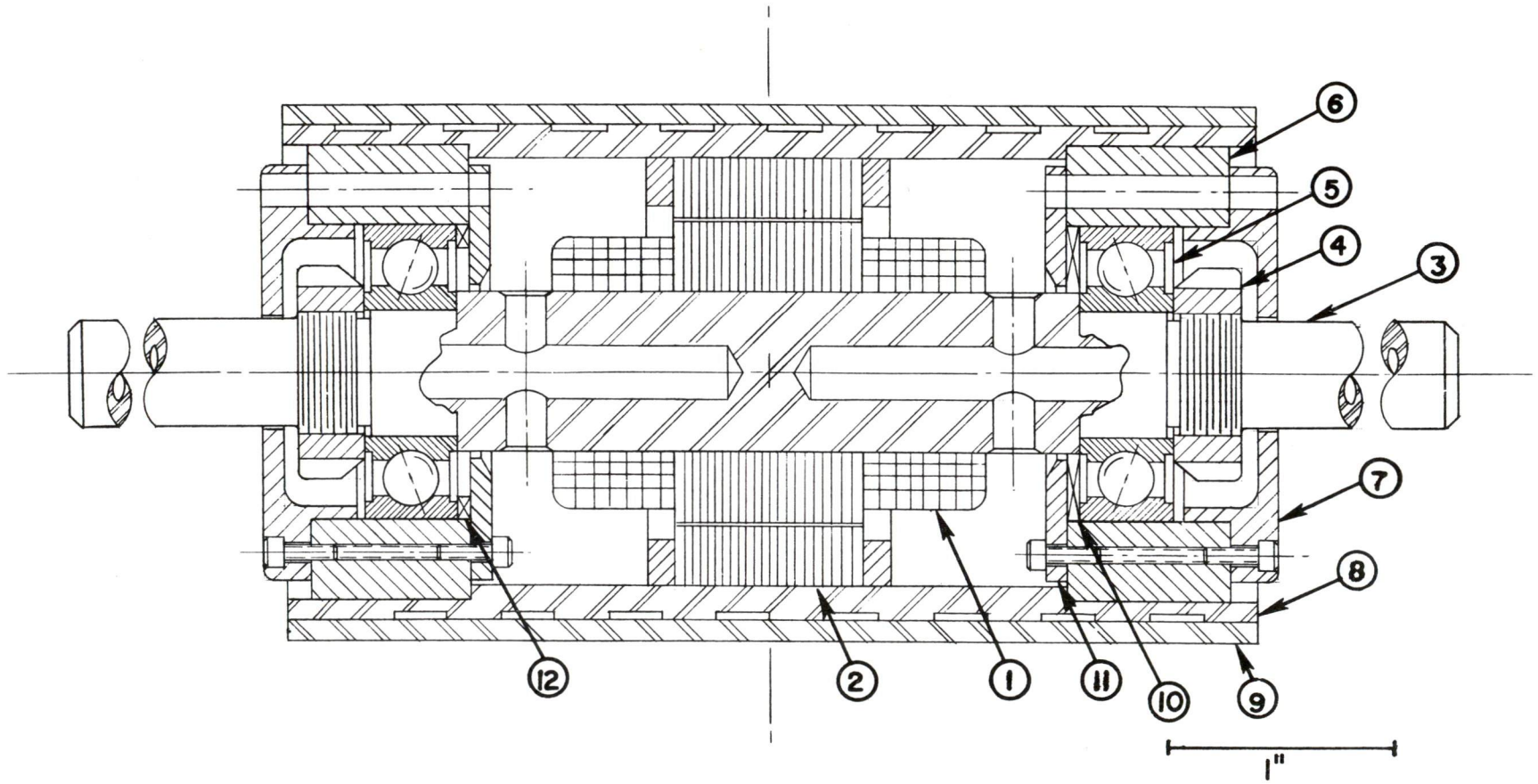


Figure 1. Cross Section of Rotor Assembly

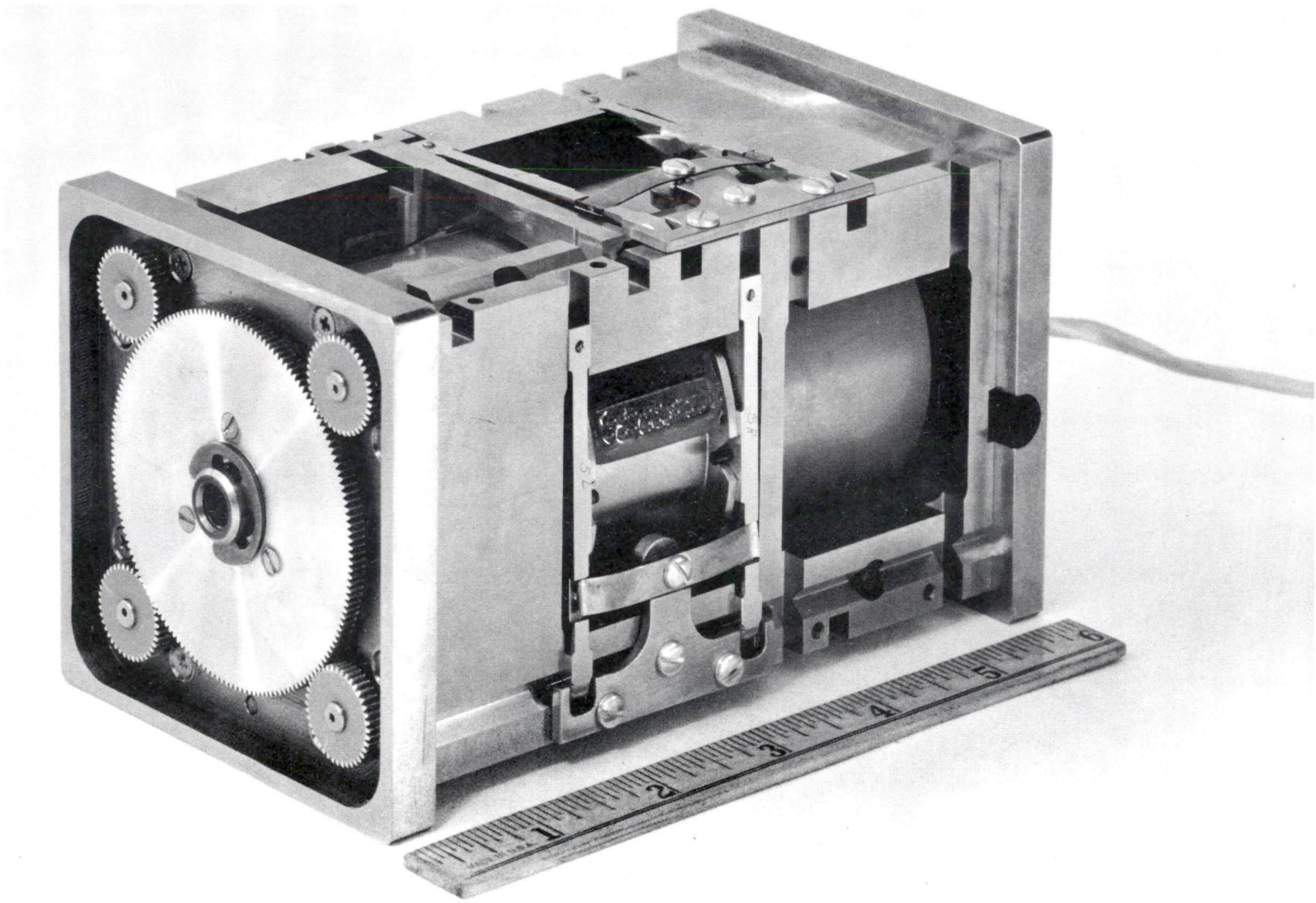


Figure 2. Photograph of Partly Assembled Magnetic Drum

on the head spacing (which is maintained by an air film between the shoe and the drum). The shoe is positioned radially by means of pivoted arms. The pivots are held in V-grooves to eliminate any possible play.

A gear wheel can be seen which turns cam shafts mounted down the length of the four corners of the framework. The cam shafts take the pressure off the shoes for starting or stopping. A very small motor (not shown) will be mounted to turn the gear wheel against a spring when the drum has attained full speed. Upon the removing of the driving power, the spring will turn the large gear wheel and take the load off the shoes.

The shoes are self-aligning and no adjustments other than spring pressure are required. The use of two independent arms loaded by a single cantilevered spring achieves this self-alignment.

Positions for four large sized shoes are visible in Figure 2. On the other faces of the frame similar mounting spaces for smaller shoes are provided. These shoes are intended to hold both read and record heads for circulating registers.

The electrical characteristics are summarized as follows.

Recording. Peak currents of 100 ma are required for recording. The current is built up linearly during half a bit time for the Manchester type recording. A silicon transistor push-pull circuit with 6 volts on the collectors is used for the recording amplifiers.

Reading. The read signal is about 10 mv peak-to-peak at 546 kc and about 30 mv at 273 kc. No noise is noticeable on the signal under test conditions. Using Manchester or variable phase type recording, no transients are apparent beyond one recorded bit before and after each word. Pattern sensitivity has been eliminated by the use of narrow pole piece heads described later. Typical read signals made up of single eight-bit words are shown in Figure 3.

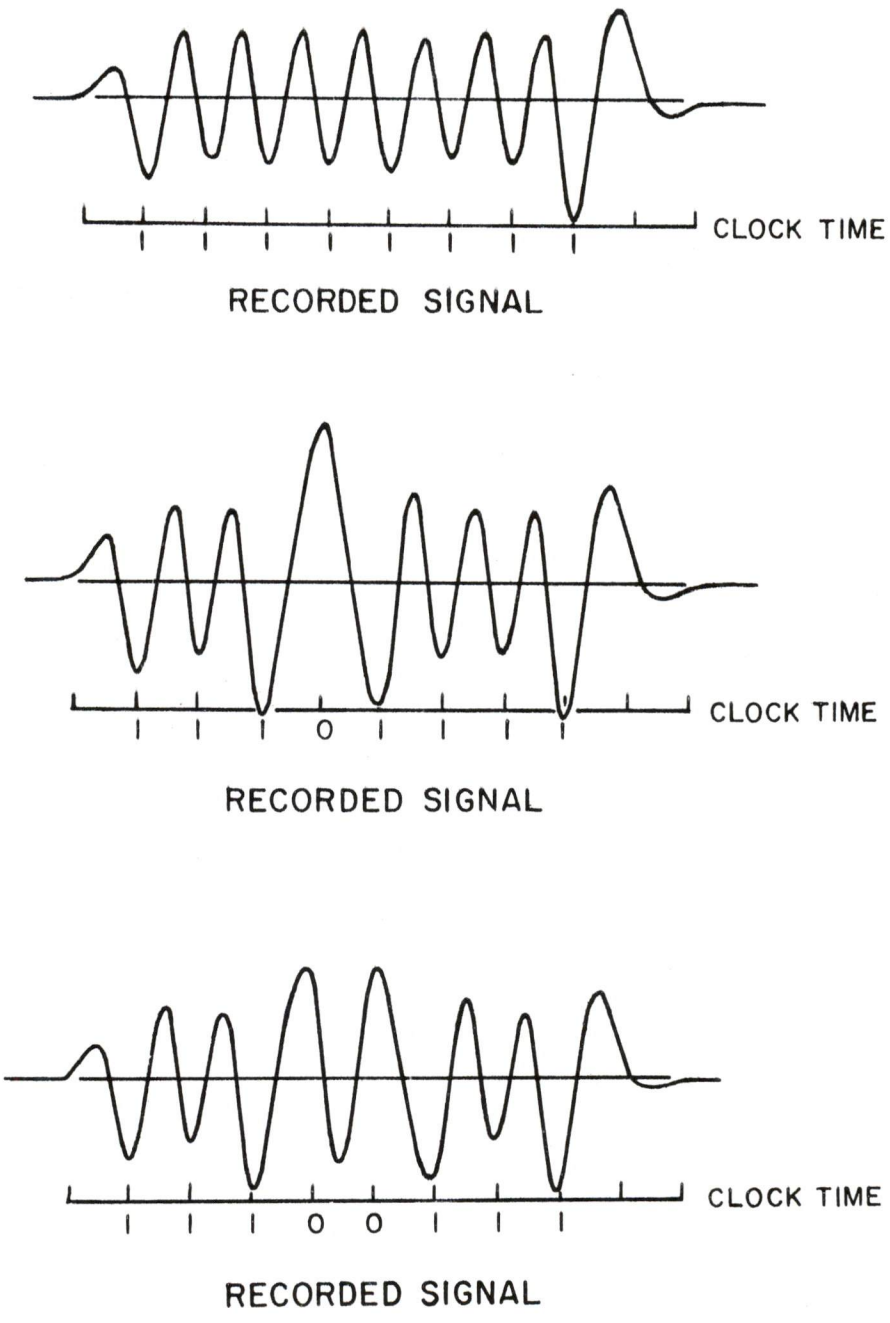


Figure 3. Read Signals with Clock Times Indicated

SELECTION OF MAGNETIC COATING

For digital recording the head-to-drum spacing should be of the order of one-tenth or less of the length of the recorded bits to achieve customary margins of operation. For 350 or more bits per inch, a head-to-drum spacing of less than 300 microinches is indicated. For both temperature ranges of -55° to $+125^{\circ}C$ and high shock and vibration, the small head-to-drum spacing required of 300 microinches cannot be maintained unless the head is made to bear on the recording surface. Gas lubrication is satisfactory for the maintenance of spacing in this range. For practical reasons it is desirable that a small particle of dirt (or accidental mistreatment during assembly or service) not do appreciable harm to the recording surface. This puts a requirement on the durability of the magnetic coating. For this reason magnetic plating is preferred to oxide films. There is an optimum plating thickness (which in practice turns out to be of the order of 100 microinches) for 350 bits per inch as is shown later. Since oxide coatings are usually ground after application, there would be an especially difficult problem in grinding them down to a uniform thickness of 100 microinches. Thus it became necessary to develop a suitable plating. Electroplated nickel-cobalt alloys have been tried and work perfectly well magnetically. They can be plated as thinly as desired and have been tested at thicknesses of 60 microinches and less. Mechanically this plating is not the best that can be obtained since it is not especially hard and has not been made to have both a high coercive force and adhesion strength comparable to the bulk material strength. This type of coating is magnetically satisfactory, but slight damage may put several 0.03 inch wide tracks out of operation due to local peeling of the coating. Nickel deposited by the Brenner electroless process forms a very hard coating which has excellent adhesion and hardness after suitable heat treatment. This coating markedly improves the wear resistance of almost any material that might

be used to make the drum. A modification of the Brenner process to include cobalt produces an alloy which has good recording characteristics. This alloy is satisfactory magnetically without heat treatment but can be made harder with heat treatment.

DETERMINATION OF MAGNETIC PROPERTIES AND THICKNESS

The signal read from the recording surface will be

$$E \text{ peak-to-peak} = 2\phi\omega \cdot n \times 10^{-8} \text{ volts} \qquad (1)$$

ϕ = maximum number of flux lines in the head

n = number of turns on head

ω = frequency in radians per second

This assumes that the readback signal is essentially sinusoidal. The parameter ϕ will be less than the flux lines remaining in the recorded dipoles after magnetization since not all the lines can be made to link the head. It will be proportioned to track width. It will be dependent on B_r and H_c for the magnetic coating.

For a thin magnet which is very wide, it can be shown that

$$H = H_o - (2t / \pi L)(B-H) \qquad (2)$$

where

H_o = applied field

H = effective magnetizing field

t = plating thickness

L = length of the recorded dipole

B = magnetic induction

A nickel-cobalt plating having a coercive force of 320 oersteds and a saturation induction of about 6000 gauss was selected. The ratio t/L can be varied so that a demagnetizing H just intersects the corner of the B-H loop for the material. Since L is fixed by

the recording density, t is selected so that the residual induction is near the maximum induction, thus taking advantage of the squareness of the hysteresis loop of the nickel-cobalt alloy. A greater thickness would provide no greater residual flux because of demagnetization, but would require a greater recording magnetomotive force and would magnetize more slowly due to eddy current effects. Thus both magnetic plating material and its thickness can be optimized for the drum memory.

Figure 4a shows an actual B-H loop of a nickel-cobalt plated film to show the effect of thickness on the residual induction due to demagnetization. A line is drawn of slope determined by t/L which intersects the B-H loop at the point of residual induction.

Figure 4b shows a similar B-H loop for a heat-treated, nickel-cobalt alloy chemically deposited by the Brenner process. The squareness is not as good as that obtained by electroplating, but it is expected that this could be improved.

The B-H loops were taken on actual plated 2-1/2 inch diameter by 4-1/2 inch long stainless steel cylinders before they were mounted on the drum assembly. (See Figure 5 for photograph of the B-H tester.)

The B-H loops were taken by magnetizing the plating axially in a solenoid whereas recording takes place around the periphery of the drum. There was some doubt as to whether or not anisotropic effects would invalidate this measurement, and so several disks were plated and tested along various axes in the B-H tester. Very little change in B-H characteristics was noted as the direction of magnetization was changed. The disks were purposely ground so that the effect of grinding marks would be observed if they set up an easy direction of magnetization.

DESIGN OF SUITABLE READ-RECORD HEADS

The design goals called for 350 bits per inch recording density and at least 30 tracks per inch. Reading resolution of 350 Manchester cells per inch requires coupling as much flux as is

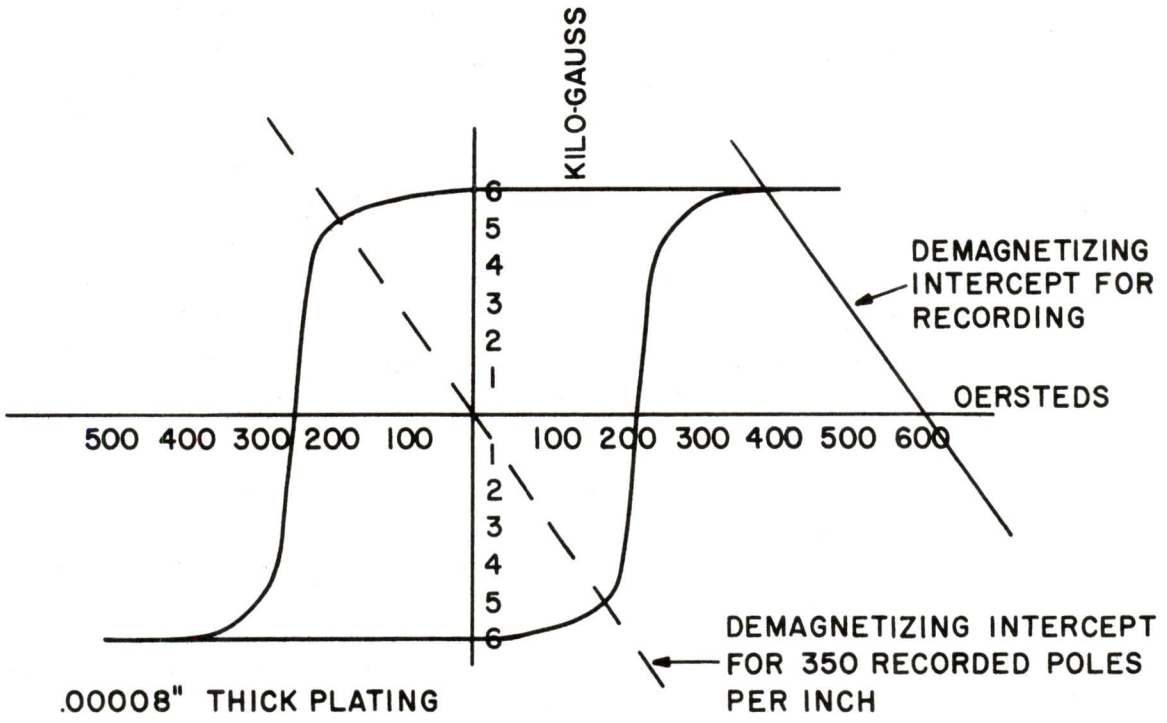


Figure 4a. Actual B-H Loop for Electroplated Nickel Cobalt

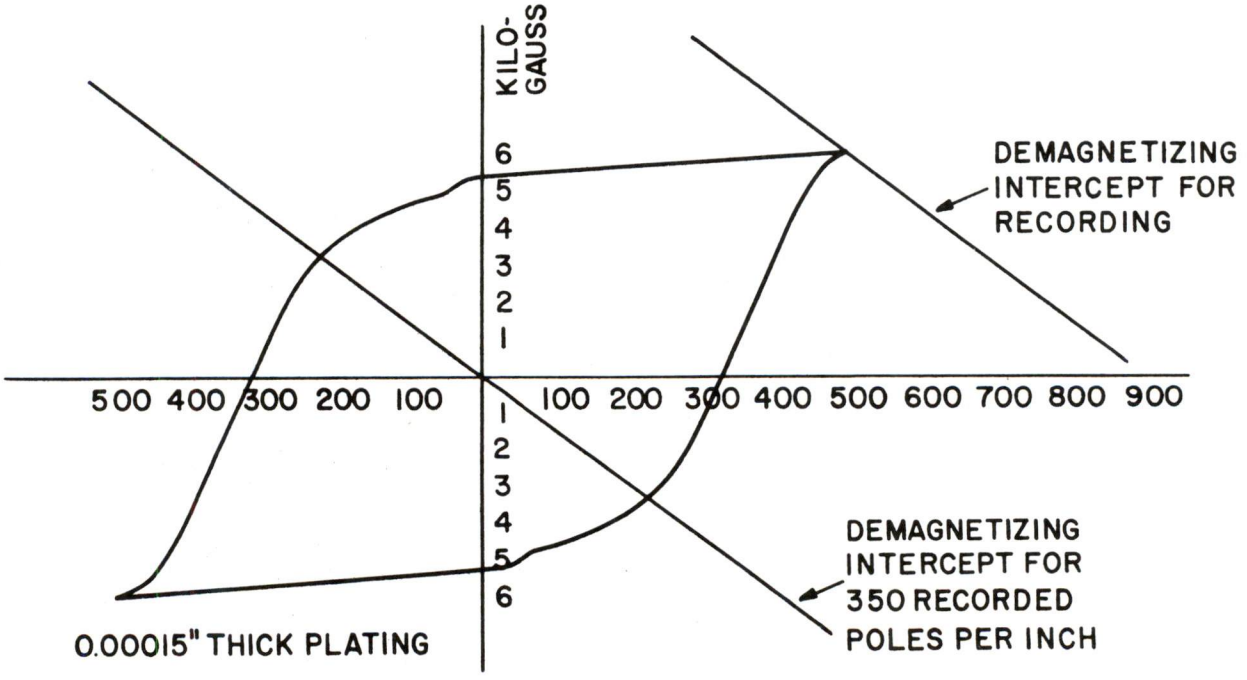


Figure 4b. Actual B-H Loop for Heat Treated Electroless Nickel-Cobalt

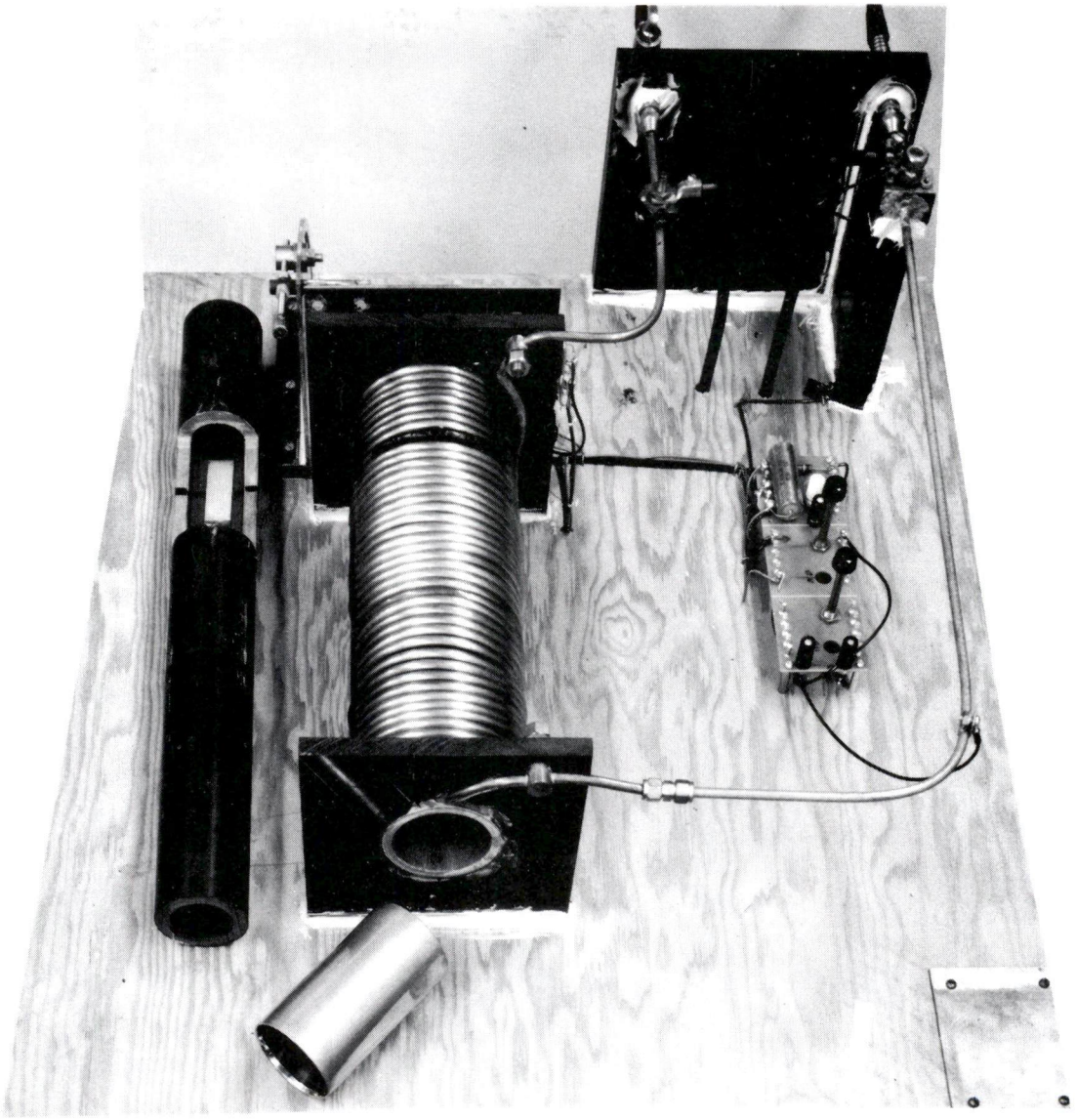


Figure 5. B-H Loop Tester

possible from a 0.0014 inch long magnetic dipole into a magnetic structure around which are wound a number of turns of wire. Coupling much of the flux requires a head gap of the order of 0.0004 inch and head to recording surface spacing smaller than 0.0001 inch. However, a compromise can be made which will cause a loss of signal but not necessarily loss of operational margins. Recording densities of more than 1000 bits per inch have been obtained in systems using a single floating head assembly. However, this usually is accomplished with very closely spaced heads and wider tracks than 0.025 inch. In the interest of economy and development time a compromise which utilized many heads mounted in a single air floated pad was adopted. To make the construction problem easier, a head-to-drum spacing of 200-300 microinches was adopted. This limits a practical digital recording system to the region of 500 recorded bits to the inch. In the present system a recording density of 350 bits per inch is used, but this does not represent the practical system limit. The floating pads holding about 27 heads are of the order of 1.3 inches by 1.25 inches. Economies in space and cost are achieved by this mass mounting method which at present requires the use of recording densities of 500 bits per inch and less. The problems of recording and reading will be discussed separately although it is highly desirable that a compromise head be used which can both record and read. Apart from economy it greatly relaxes mechanical tolerance problems.

Recording. Figure 6 shows an idealized read-record head at its pole face. If the resistivity of the pole pieces were high so that eddy currents could be neglected, the amp turns required for recording and the read signal obtained per turn of the head winding could be quite closely calculated. Such a head is most difficult to make and the desirable spacing to the recording surface of 50 microinches or less is also most difficult to obtain in multiple head assemblies. The performance of the idealized head is of interest, however, for comparison with the compromise design which has been presently adopted but which clearly could be improved. To determine the recording amp turns required, let the B-H loop (Figure 4b) be assumed to be the B-H loop for the recording surface. For the 0.00015 inch thick plating whose B-H loop is shown on Figure 4b assuming

$$L = 0.0014'' \text{ (350 bits per inch Manchester recording)}$$

It can be seen from Figure 4b that 500 oersteds are required to saturate the magnetic plating at 6000 gauss. From Equation (2) we find that

$$H_o = 900 \text{ oersteds approximately for 500 oersteds effective magnetizing force}$$

Two parallel lines are shown on Figure 4b, whose intersections with the B-H curve and H axis give the residual flux density and the recording force required. This gives a flux density after magnetization of 3200 gauss. If the curve of Figure 4a were used, magnetizing force of 600 oersteds would give a remnant density of 5500 gauss. However, because the electroplated coating is thinner (80 microinches versus 150 microinches), the remnant flux would be only 90 percent of that obtained for the electroless plating.

The overriding consideration for selecting the electroless plating was its hardness and resistance to wear.

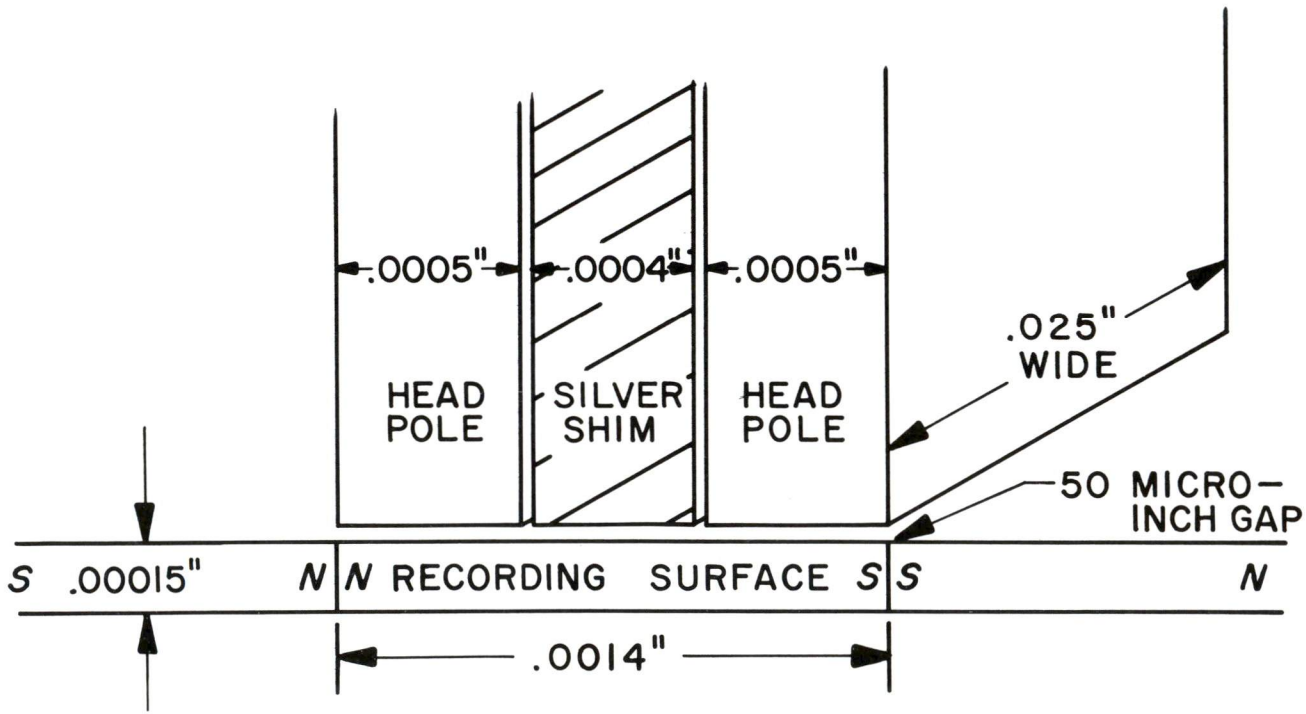


Figure 6. Ideal Geometry for Recording 350 Bits/Inch

The remnant flux for a recorded dipole 0.0014 inch long, 0.025 inch wide and 0.00015 inch thick would be about 7.7×10^{-2} lines for a flux density of 3,200 gauss.

About 2.5-amp turns must be provided for magnetizing the plating. In the ideal head (Figure 6) 14.5×10^{-2} lines must be maintained across two gaps in series to saturate the coating at 6,000 gauss. The gap dimensions are 50 microinches in extent, 0.025 inch long and 0.0005 inch wide. This infers an average flux density in the air gap of 1,800 gauss, the maintenance of which will take about 0.36-amp turns. The maintenance of flux in a very small continuous permalloy or ferrite circuit will take a negligible extra number of amp turns.

In practice, sufficient amp turns must be provided to generate a large number of fringing lines which form closed circuits around the side of the head and under and over the recording surface. If the ideal head as drawn in Figure 6 were made, 3-4 amp turns would be sufficient for recording on the magnetic coating specified.

In practice, allowance has been made for the fact that the air gap may be 300 microinches instead of 50 microinches since this is much more readily achieved in a multiple assembly holding 27 heads. The best compromise for recording also includes making the silver shim gap larger than would appear ideal for small head-to-drum spacings since the flux density drops off rapidly in terms of the head gap dimension. A practical though not very efficient head would utilize 0.001 inch wide pole pieces with a 0.001 inch wide silver shim. (See Figure 7.) Such a head records with 15-amp turns but gives a slightly greater read signal using 30-amp turns. Since these figures are large compared with the calculated 3-amp turns, it is clear that recording efficiency was sacrificed in order to make the head easier to fabricate and less sensitive to spacing than the ideal head. This inefficiency becomes important only if the recording circuitry becomes large or impractical. A head made to the dimensions shown on Figure 7

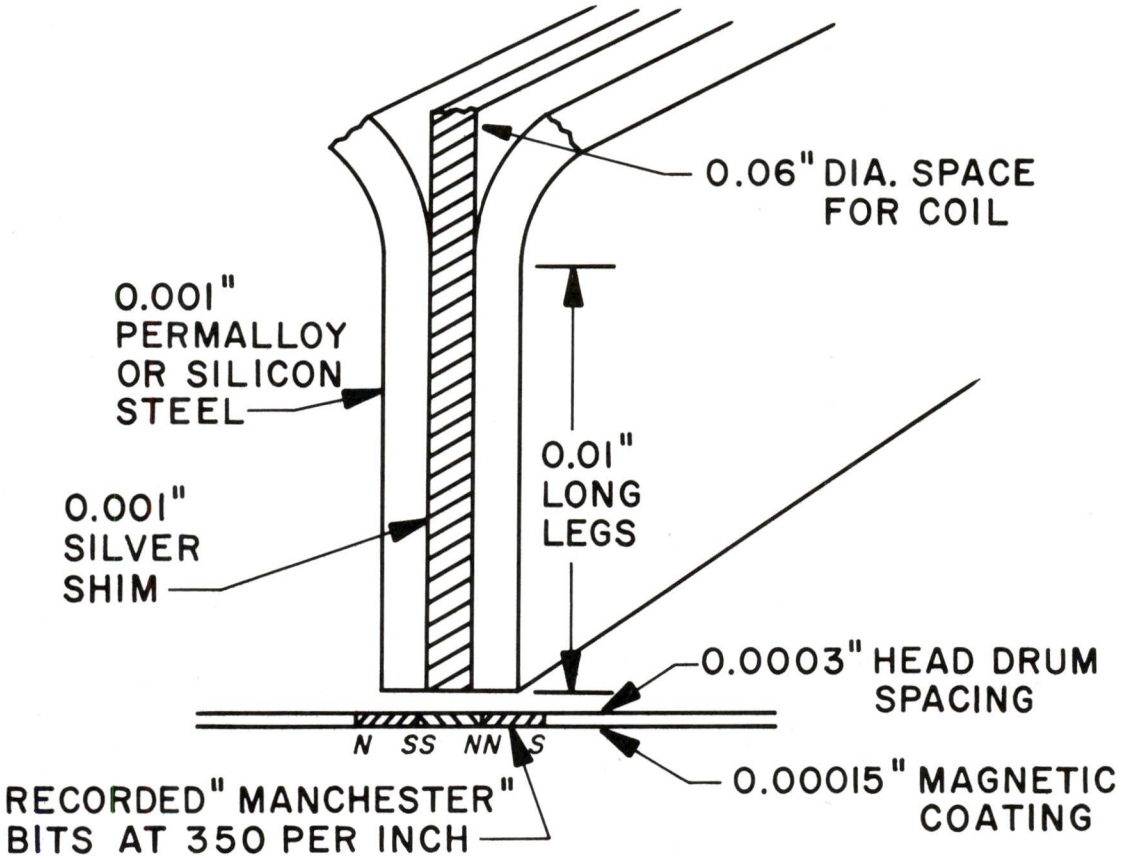


Figure 7. Practical Geometry for Recording

has been driven at 546 kc with a silicon transistor circuit using 6 volts on the collectors and 100 to 200 ma peak current. Since this circuit is quite acceptable for a microminiature computer, recording efficiency can be sacrificed if this results in a net savings in manufacturing cost. The practical geometry of Figure 7 clearly looks inefficient magnetically, but economy and ease of manufacture are in its favor. The 0.01 inch long legs are highly desirable for mechanical structure since a clamp holds the permalloy against the silver shim. The silver shim is wide for the size of the recorded dipole, but head spacing is far less critical than if the silver shim were closer to a more reasonable appearing dimension. Laminating the legs of the magnetic structure will improve the performance since penetration of the magnetic field at 546 kc is about 10 percent into either side of the material (assuming nonsaturation) for the half amplitude point. In practice, excess drive is used which causes the penetration to be greater than the 10 percent mentioned above. The penetration is greater because the permeability of the material is lowered as it becomes saturated, resulting in an increased speed of propagation in the saturated region. The final choice of magnetic head is likely to be a compromise between the schemes shown in Figures 6 and 7. For practical reasons, the dimensions shown in Figure 7 make a good starting point for the development of a useful system.

The magnetic head structure is made of 0.001-inch permalloy rather than ferrite which would be too hard to handle in sufficiently small sizes. Under less than ideal conditions for recording, there are very marked transients where recording starts and stops, since some recording on a minor hysteresis loop takes place under the full region of the magnetic head. As recording density is increased without scaling down the head gap and head-to-recording surface spacing, this problem becomes more marked. For a chosen minimum head-to-drum spacing, the useful recording density can be greatly increased if the magnetic structure of the

recording head is reduced to the smallest dimensions possible so that its influence does not appreciably extend beyond the recorded dipole. Care must be taken in using legs of small cross sectional area because there is not a large excess of flux over the amount required to saturate the coating. Flux leakage may prevent recording altogether unless the over-all head structure is kept very small.

Design of Read Head. It was shown earlier that a 150 micro-inch thick recording surface with the B-H loop characteristic of Figure 4b would have 7.7×10^{-2} flux lines at the center of a 0.025 inch wide recorded dipole. An ideal head would intercept these lines (and even increase the available flux by reducing the demagnetization). If the flux change were sinusoidal (any other wave form would give greater peak-to-peak volts) the read signal would be

$$E = \phi\omega \cos \omega t \times 10^{-8}$$

where ϕ is the total flux in the recorded magnetic dipole and $E = 2\phi\omega \times 10^{-8}$ peak-to-peak volts per turn of the reading head. At 546 kc which is the maximum frequency used, a signal approaching 5.2 mv per turn could be expected from an idealized structure. With this ideal structure, it would be easy to determine that resolving signals at a much higher density would be possible and thus it would most likely be used at a density where it would give much less than the theoretical maximum signal. The magnetic head tested with the memory system described falls far short in obtaining the maximum obtainable signal at maximum density. In fact the presently used heads develop a signal in the range of 12 mv peak-to-peak at 546 kc as against a possible 780 mv calculated for a 150-turn head. Reference to Figure 7 indicates that unlike the situation in Figure 6 where more than half the flux would couple the head windings, only a small part of the flux will

be useful in generating a read signal. Calculation of the exact magnetic flux coupling in this situation is most difficult, but a glance at a scale drawing makes the finding of 1/66 of the possible signal quite plausible. The fact that the low output is tolerated is a compromise between signal level, and economy and ease of manufacturing the heads. Since an excellent signal-to-noise ratio and margins in clock pulse timing are obtained in this situation, the compromise is quite tolerable.

At 273 kc which represents the pattern 0 1 0 1 in Manchester recording, the read signal obtained is about 30 millivolts in comparison with a possible 390 millivolts if all the flux in the recorded poles interlinked the head winding. The loss of signal by a 13 to 1 ratio is explained by the presence of an air gap, which provides a substantial reluctance in series with the head structure, and also by the fact that the head structure itself does not have a zero reluctance. Figure 8 shows the response of the read head versus recording density and indicates that the head shown in Figure 7 is being used beyond its optimum density.

Figure 3 shows signals read by the head with clock times indicated. As can be seen, the signals can be interpreted with adequate reliability since there is no noise or mistiming in evidence.

There is, of course, much room for improvement of the magnetic head; however, each improvement increases the difficulty of making the head and the increased cost must be balanced against the economic benefits of the improvement.

Construction of the Magnetic Head. Figure 9 shows the essential detail of the magnetic head. In assembling these heads the lower part is insulated and slipped into an aluminum tube. The tube is compressed forming a subassembly which can be tested. The subassembly heads are clamped into a holder (Figure 10)

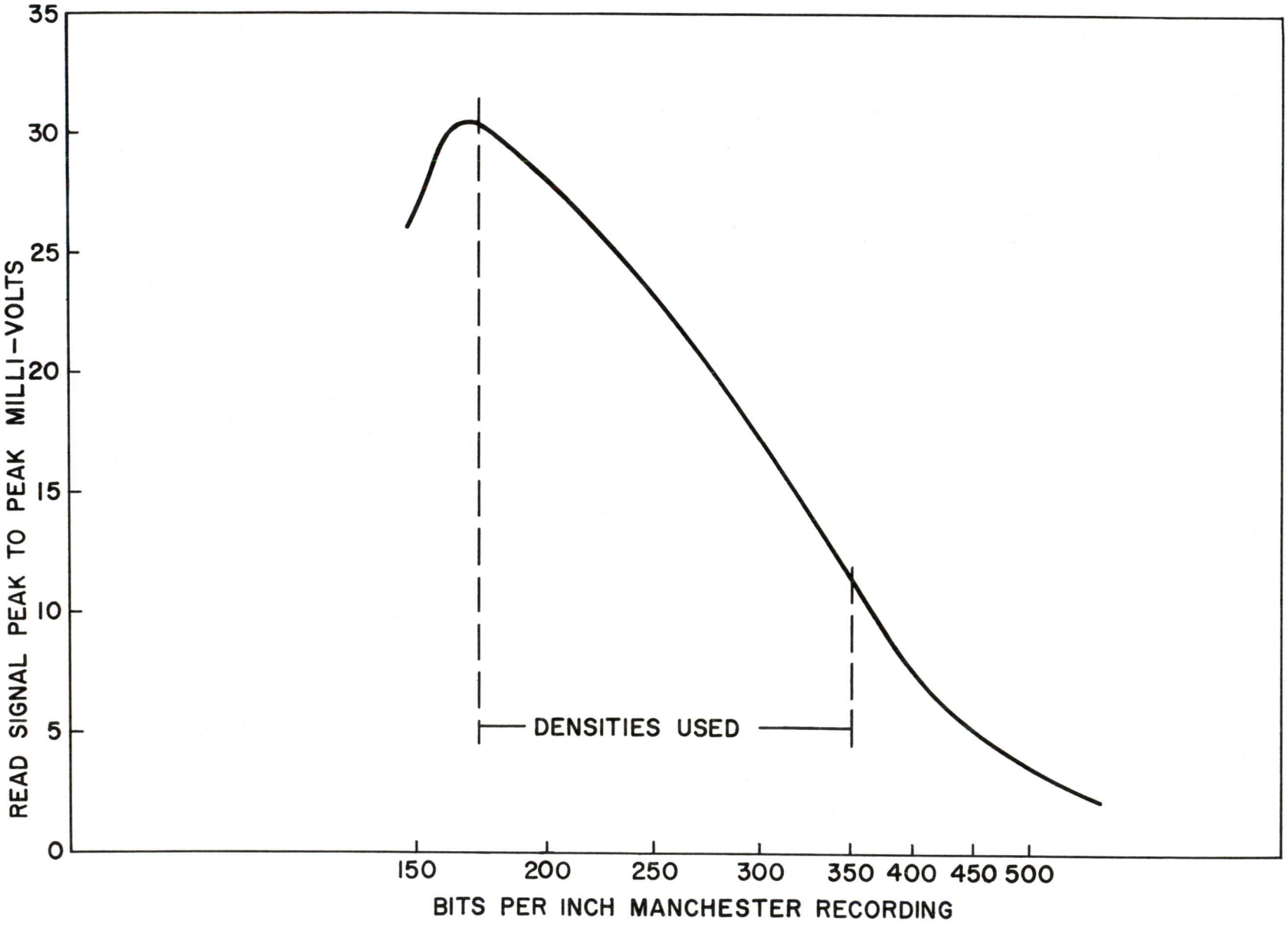


Figure 8. Typical Response for Magnetic Read-Record Head

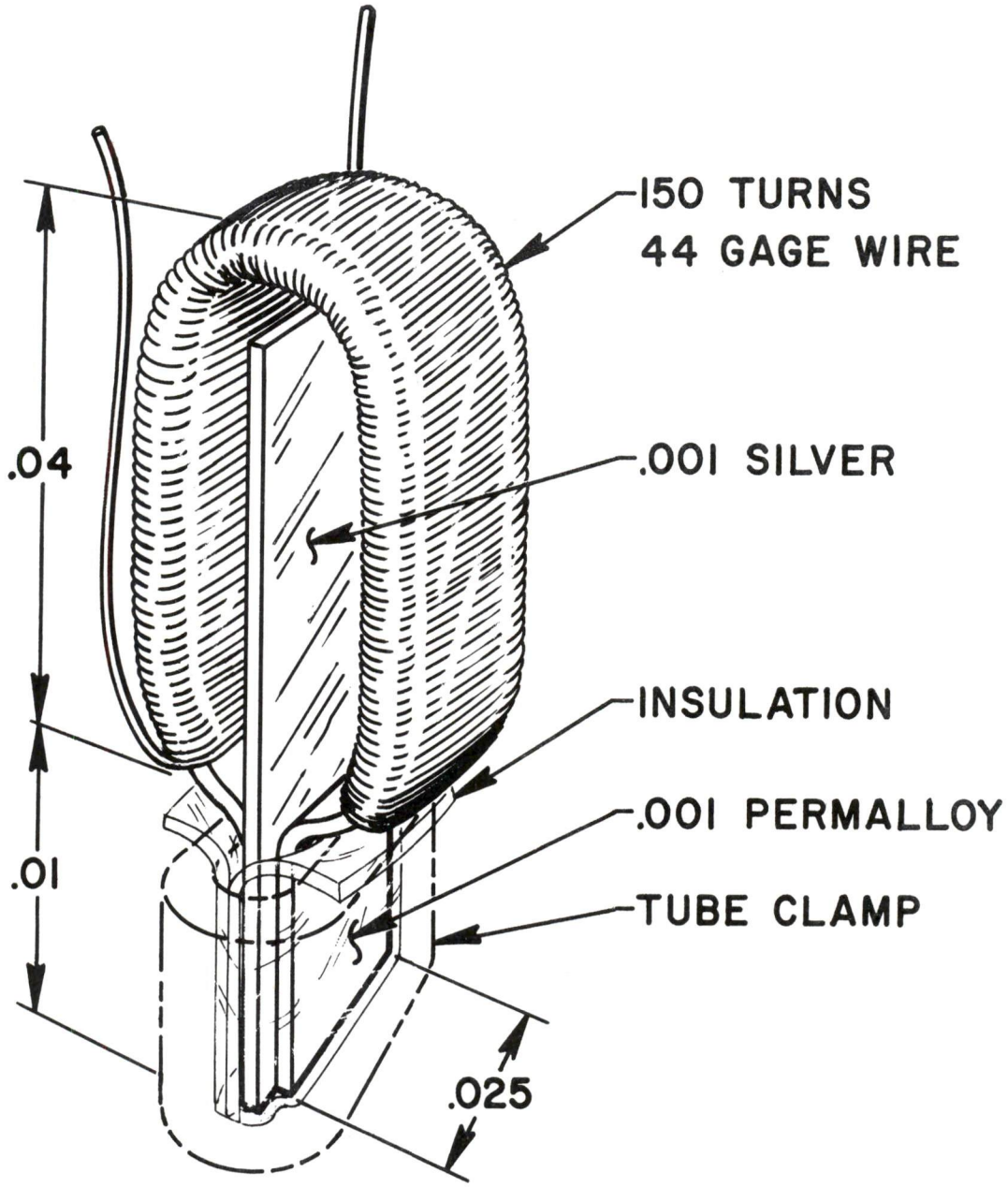


Figure 9. Essential Details of the Magnetic Head

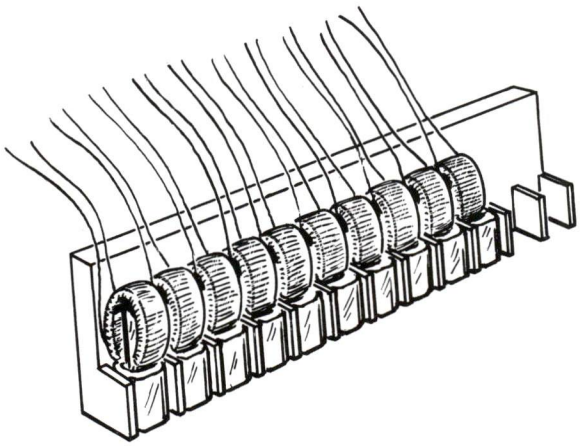


Figure 10. Partly Assembled Magnetic Heads

and fixed in place with a suitable high temperature epoxy resin compound. Two such assemblies are made with the heads staggered so that with the assemblies mounted 15 to the inch a track density of 30 per inch is achieved. The assemblies are then mounted in the shoes.

MECHANICAL DESIGN DETAILS

Principle of Operation of Head Support Mechanism. A rotating drum moves a considerable volume of air in its close vicinity even though the drum surface is quite smooth by normal standards. This phenomenon is due to a boundary layer effect. That is, air molecules which are immediately in contact with the drum tend to adhere to that surface. Due to the viscosity of the air, the air molecules immediately about this initial layer are dragged along and as the distance from the drum surface increases, the velocity of the air molecules which are dragged along decreases. With this concept in mind, it is seen that if a stationary surface which is curved to match that of the drum is held near the rotating drum surface, the air will be dragged between the two surfaces. Since the air will also tend to adhere to the second surface, there will be a drag or friction force as shown in Figure 11. If this stationary surface is inclined to the drum surface so that the space decreases in the direction of rotation, the air which is dragged in is squeezed into a progressively smaller space as is shown schematically in Figure 12. This squeezing effect is of course a compression process, and pressure forces normal to the two surfaces develop. If this second surface is held in place by a spring force of proper magnitude, it will be held off the drum to a distance where the fluid pressure force equals the spring load force. When such a condition exists, the layer of fluid which separates the two surfaces is referred to as a hydrodynamic lubricating film and such surfaces which react in this manner are referred to as a self-acting bearing. In the example given, air is used as the lubricating fluid; however, any fluid, liquid, or gas which will adhere to the bearing surfaces without causing damage will perform in this manner.

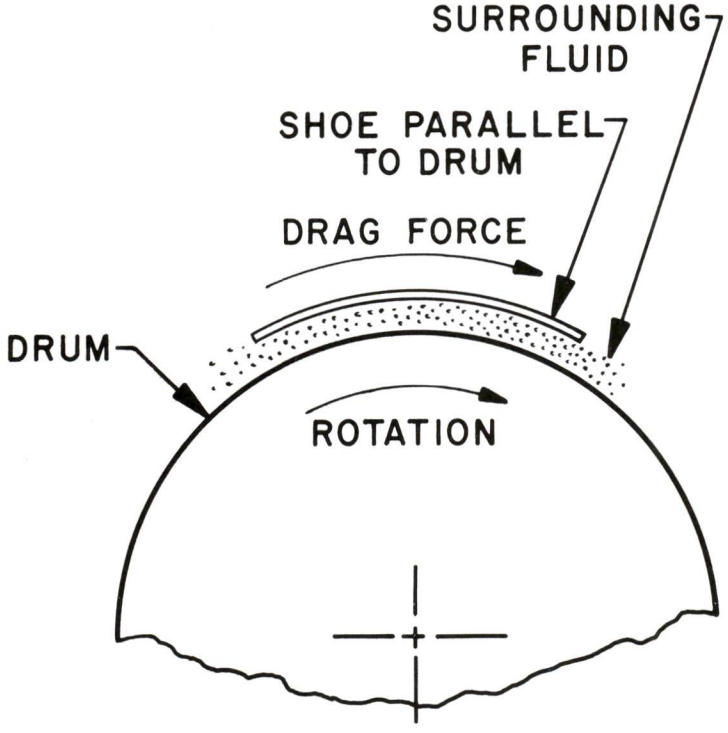


Figure 11. Schematic View of Drum With Shoe in Parallel Position

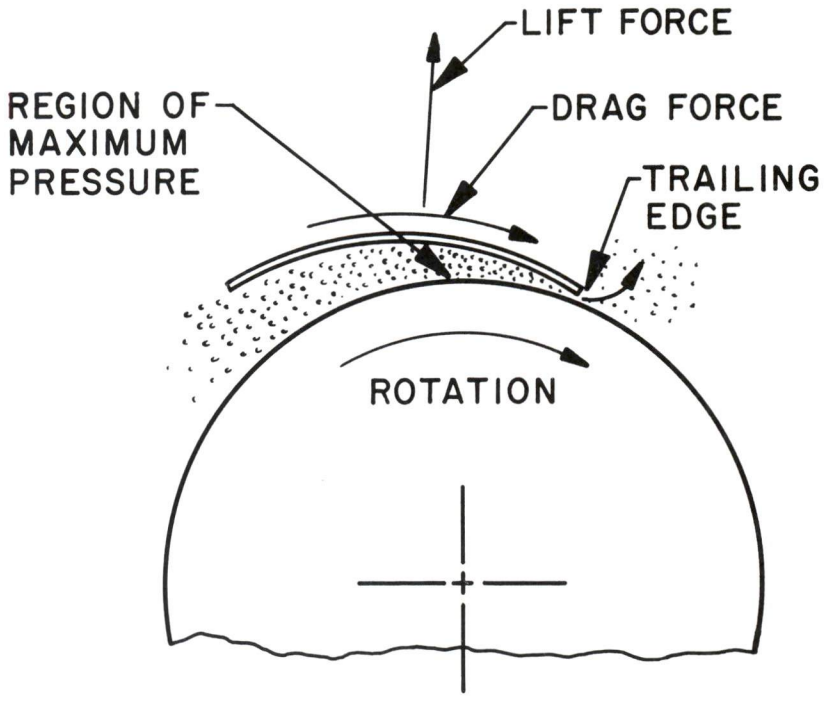


Figure 12. Schematic View of Drum with Shoe at Angle to Develop Wedge of Lubricant

The theoretical aspects of this phenomenon were first proposed by O. Reynolds about 75 years ago, and solutions of his equation for the incompressible lubricating films have been well accepted in the literature on bearing lubrication. In recent years considerable attention has been directed toward the case of the compressible or gas lubricating film for many promising advantages such as chemical stability, extremely low friction, the maintenance of close clearance between moving parts, and the use of the ambient gas as a lubricant. The technology of the analysis of the bearing using a compressible fluid as a lubricant as in the example (Figure 12) is quite involved and beyond the scope or purpose of this paper. Work on this phase for use in the design of such bearings for use in memory drums is in progress; and for the technology the reader's attention is directed to the list of references.

Properties of Lubricating Film Supported Shoe. The lubricating film supported shoe possesses certain unique properties which make it a useful device for the support of a recording head. The most important properties will be described below. For the purpose here, let us denote the angle between the drum and shoe surfaces as the attack angle α, the edge farthest from the drum surface as leading edge, and the edge nearest the drum as the trailing edge. Drum rotation is in the direction from the leading edge towards the trailing edge.

Figure 13 shows the typical relationship of the pressure force which can be developed under typical operating conditions. It will be noted that at the operating conditions shown in the figure, a mean pressure of 1.5 psi gage at a trailing edge spacing of 400 microinches is obtained. As the trailing edge spacing is decreased, the mean pressure increases at a rapid rate so that at a spacing of 200 microinches the mean pressure has increased virtually four fold or inversely with the square of the spacing. This characteristic is most desirable from electrical and mechanical points of view for

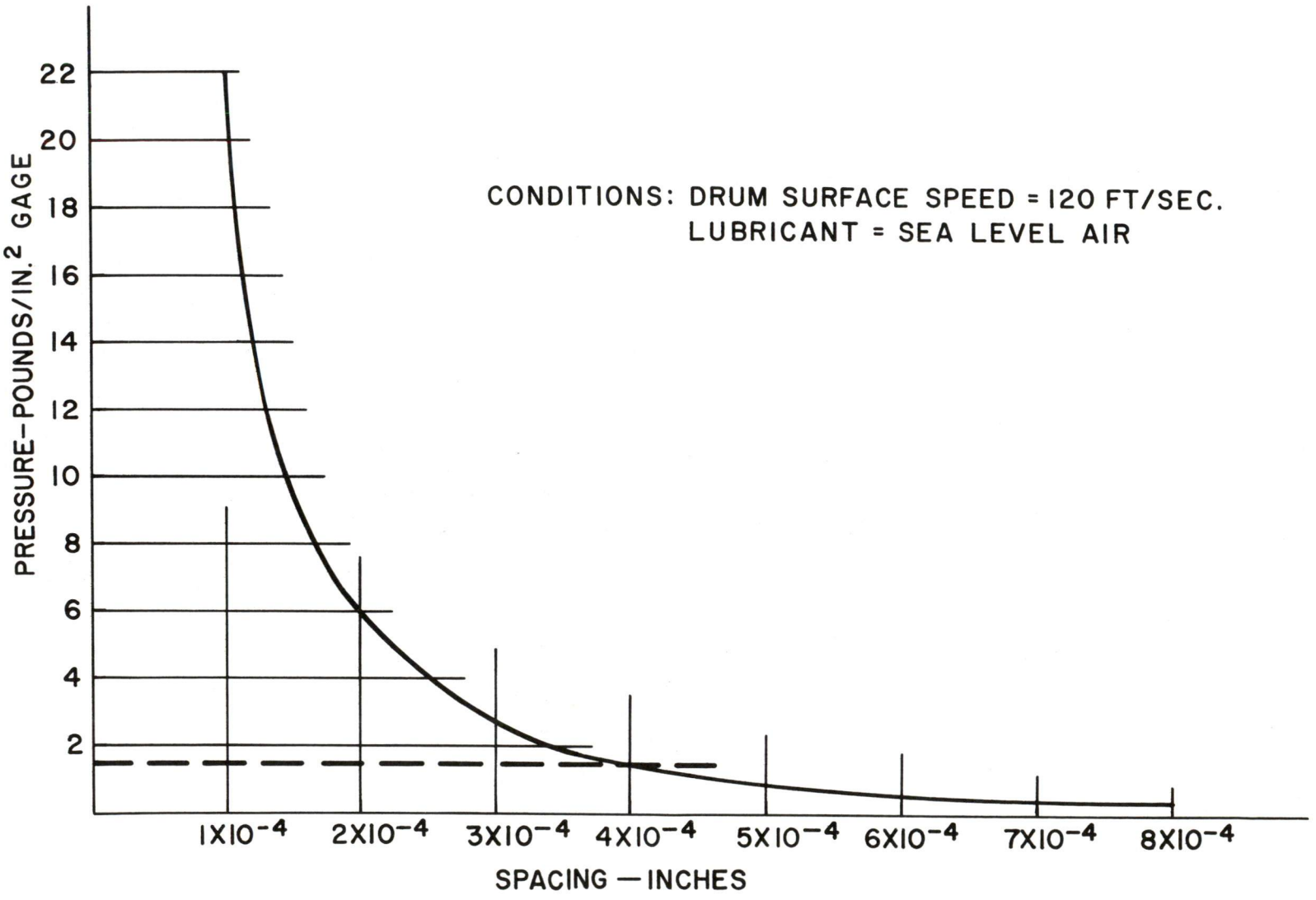


Figure 13. Mean Pressure vs. Trailing Edge to Drum Spare

recording drum applications. For any fixed design as the load is increased the shoe and hence recording head to drum spacing is decreased. This is, of course, helpful to the electrical performance as far as output signal is concerned. As a greater load is applied to the shoe, the ratio of the applied load to the weight or inertia of the shoe and its associated mechanism is increased. When this ratio is increased, the ability of the shoe to withstand accelerations and run out irregularities of the drum is also increased. In the case of the drum which is the subject of this paper, the effective area of the shoe is 1.6 square inches and its normal operating load is 10 pounds, which gives a mean pressure of about 6 psi. The shoe with recording heads in place has an effective weight of about 1.5 ounces, and so the load-to-weight ratio is slightly over 100. Since at this operating condition, slight changes in the spacing result in a considerable change in the lift force, there is available a large force to restore the proper head to drum spacing. Let us consider an example at the conditions cited above. In a broad sense, since the curve shown in Figure 13 is one of a force vs displacement, the lubricating film may be regarded as a spring of variable rate. If the displacements are left small, the lubricating film may be approximated by a linear spring and the slope of the curve may be taken as the spring constant. For the conditions cited above, this slope or linear spring constant is about 100,000 pounds per inch for a shoe of the given effective area. The spring rate of the spring used to produce the load force would have to be added to this rate; however, since this spring would have a rate of about 100 pounds per inch, it is virtually insignificant in its effect on the natural frequency of the system of forces acting on the shoe. The spring rates of 100,000 pounds per inch acting on the effective mass of the shoe give a resonant frequency of more than 3000 cycles per second. Thus it follows that such a mechanism is quite capable of withstanding accelerations of 10 g's up to 2000 cps without seriously affecting the output electrical signal.

Another unique property of the floating shoe is its inherent stability. Figure 14 shows a typical pressure distribution between the trailing and leading edge of the shoe. As the attack angle α is increased, the center of pressure shifts towards the trailing edge, and similarly as the angle α is decreased the center of pressure shifts to the leading edge. Let us fix a certain shoe geometry and allow the shoe to pivot about an axis at the center of pressure and parallel to the drum surfaces. Now if the shoe is tipped so that the angle α is increased, the center of pressure moves toward the trailing edge. This action develops a turning moment on the shoe. The turning moment is in the direction required to return the shoe to the original position. Similarly, when the shoe is tipped to an angle less than the stable angle, a turning moment of opposite sign develops to return the shoe to the original stable position. From experience it has been found that the system has sufficient damping to make it stable. Thus it follows that the location of the pivot axis is not critical, for the shoe will tend to seek a value of the angle α so that the action line of the center of pressure will pass through the pivot axis.

Design Requirements. The design of a mechanism to make use of the lubricating film supported shoe or for keeping a recording head in proper location with respect to the drum recording surface requires careful attention to the precision requirements of the mechanism. The development of a design framework which requires a minimum of very precise parts which are amenable to precision manufacturing techniques is necessary to the successful execution of the task. It is not only necessary to have surfaces which are geometrically true, but it is also required that the proper geometric relationship between the various parts be accurately maintained. The most important of these relationships is the alignment between the shoe and the drum. It is essential that the center of curvature of the shoe be maintained parallel to the axis of rotation of the drum. The limits of accuracy

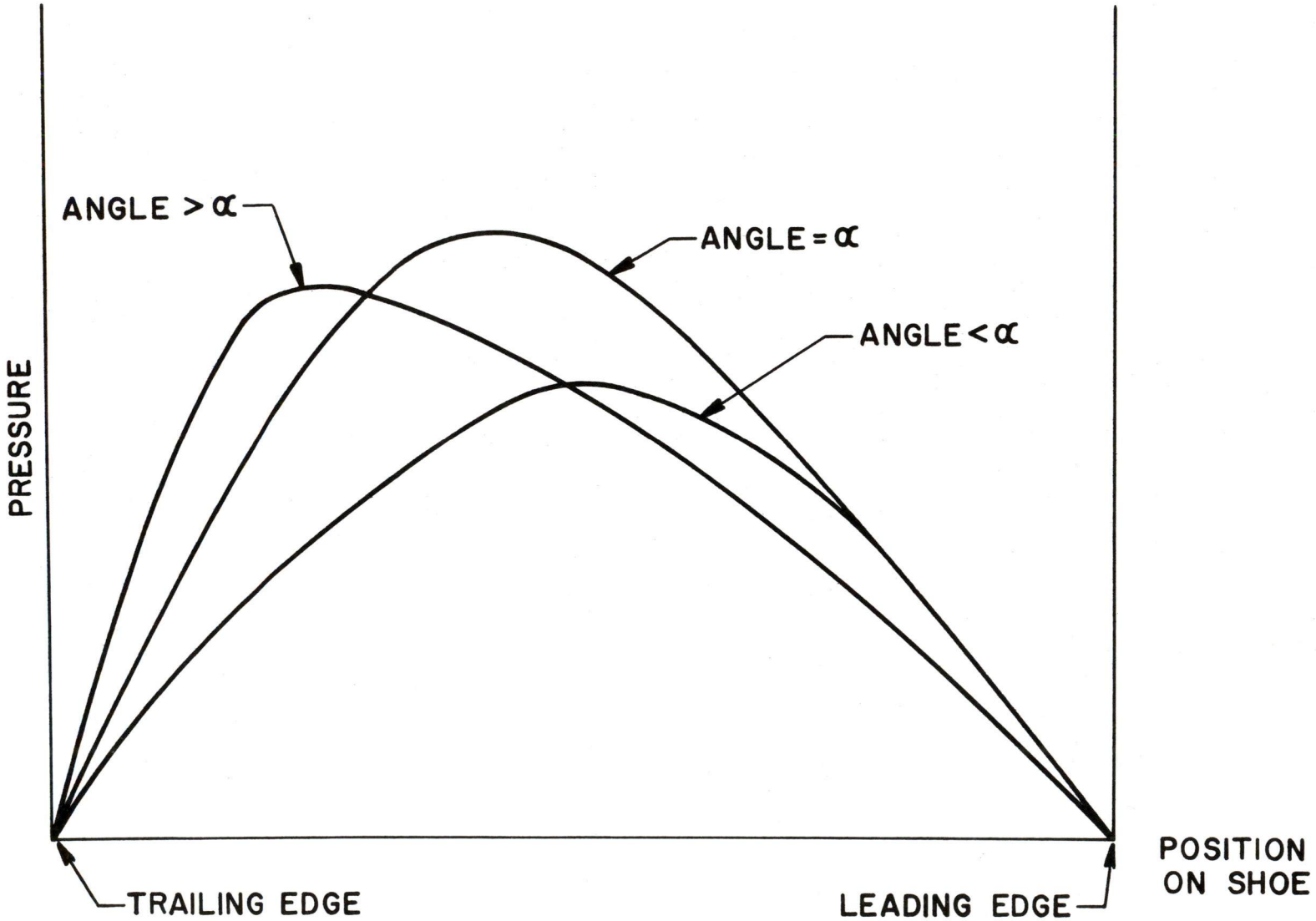


Figure 14. Typical Pressure Distribution Box Change in Attach Angle α

required are dependent upon the particular design and the performance required. For the design the out-of-parallelism is kept to less than 3 parts in 10,000. The other important requirement is that the load on the shoe be uniformly distributed so that tipping does not occur. As will be shown later, the load on the shoe of the subject drum is applied at two points. The difference between these forces is kept to a value less than 7 percent. The tolerances given above are those used in the design of the drum with due allowance for possible manufacturing tolerance and also the expected deflections of the mechanical parts.

Figures 2, 15, and 16 show the drum in various stages of assembly. It will be noted that the rotating portion of the drum is set into a very rigid frame, and access to the drum recording surface is through appropriately located cutouts in this frame. A V-groove is machined into the sides of this frame so that it is accurately parallel to the axis of the drum. Guide slots for radius arms are machined at precise right angles to the V-groove. Each of the radius arms are provided with polished sapphire pivot pins which are cemented in place in an assembly fixture. The centerline distance between the pins is accurately maintained so that it is virtually the same for a given pair of arms associated with a given shoe. One pin of each arm operates in the V-groove of the frame, while the other pin operates in a V-groove in the shoe. The V-groove in the shoe is located in the line of action of the center of pressure, and it is made accurately parallel to the axis of the cylindrical surface of the shoe. To prevent smearing of the pole pieces of the recording heads, the curvature of the shoe is ground by means of a contoured abrasive wheel so that the lay of the grinder marks is parallel to the head gaps. Final finishing is done on a cylindrical lapping tool which has a diameter 0.1 percent greater than the drum. The load for the shoe is supplied by the spring which is adjusted by a single centrally located screw. By this means, equal forces are applied to each side of the shoe. The load forces the pins to seat

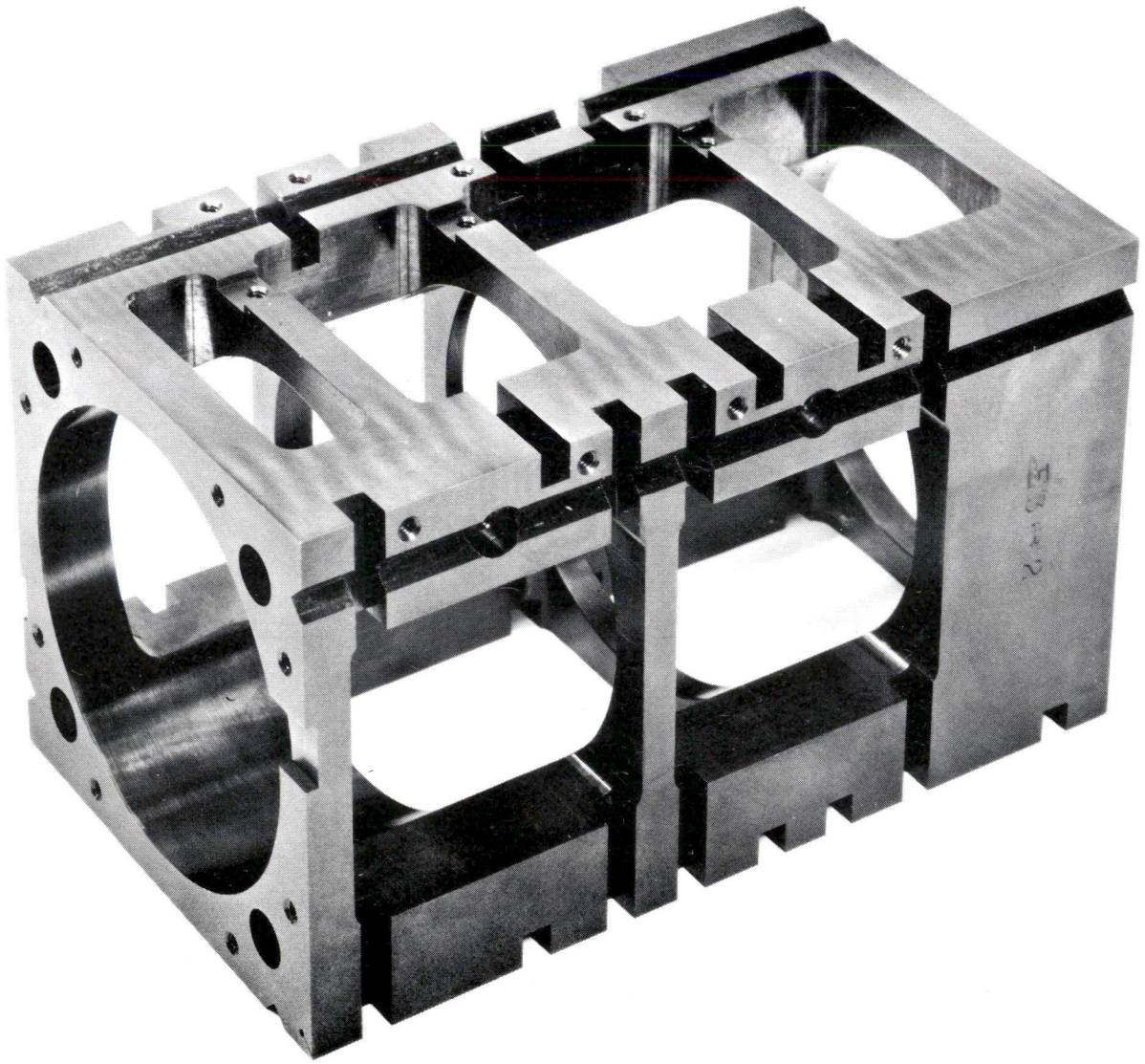


Figure 15. Main Frame

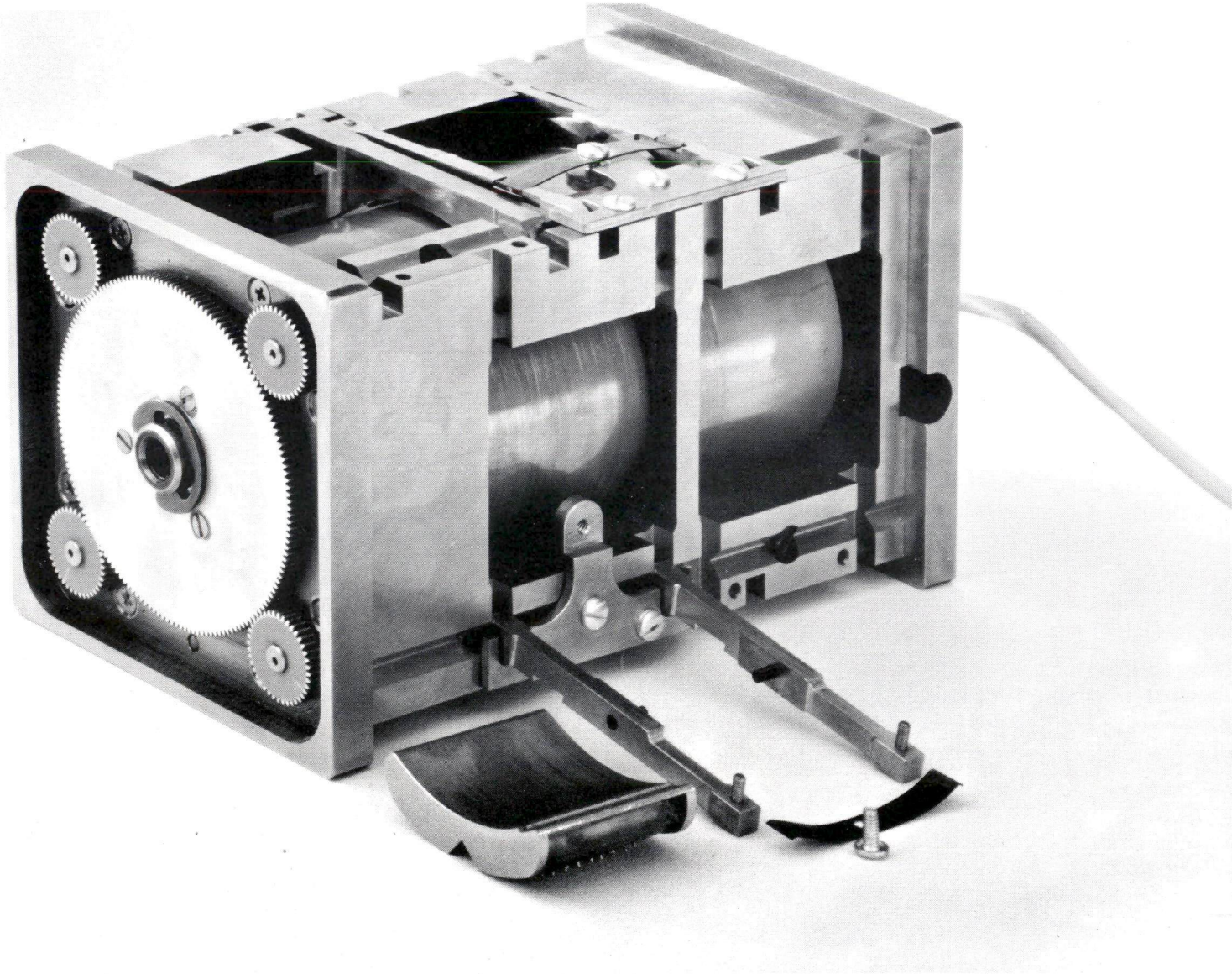


Figure 16. Shoe and Radius Arms

in the V-grooves of the shoe and frame and precisely locate the shoe with respect to the drum so that the axis of the drum and shoe are parallel within the extremely close limits previously cited.

Special consideration must be given to start and stop conditions, for without sufficient drum speed the lubricating wedge or film cannot develop and a high-friction condition will exist. To prevent this, it is necessary to unload the shoe and lift it slightly off the drum surface until sufficient speed for normal operation is attained. For stopping the drum, the procedure is reversed. There are basically two methods by which this may be accomplished. One method involves removing the spring load until operating speed is reached. The second method involves introducing lubricant under pressure through a very small hole in the shoe into the space between the shoe and drum. If sufficient lubricant (in the subject drum it is air) is supplied, the shoe will be lifted off the drum surface. After operating speed is reached, this supply of air may be shut off and normal operation resumed. This latter method requires the use of an air compressor, a fact which makes it somewhat unattractive for airborne use. The first method is used in this drum design. It will be noted that in Figure 16 the radius arms extend from the side of the frame which has the V-groove to the opposite side. At this side of the frame, the ends of the arm can ride on a simple eccentric cam which is operated by the small gears. During normal operation, these ends of the arms are free of the cam. For offspeed operation the cam is rotated to a position where the ends of the arms are lifted. Since the mechanism is extremely rigid, a movement of less than one mil of the end of the arm is sufficient to transfer the spring load from the shoe to the cam. In this condition, the lubricating film between the drum and shoe must support the weight of the shoe. Since the weight of the shoe is very much less than the operating load, the resulting friction is negligible. If the magnetic coating is very durable, the slight contact between the shoe and drum under these conditions is not serious and may be eliminated completely by operating

the drum with the axis in a vertical position. When the shoe is in this free condition, a state of instability may develop if the cam is inadvertently set to lift the ends of the arm too high. Should this condition develop, serious damage to the drum and shoe surface will occur. To eliminate this possibility, one arm of each pair for a given shoe is provided with a spring-loaded pin as shown in Figure 17. This pin is allowed to act upon the side of a shoe to cause a small amount of friction damping. Since the load at which the shoe is operated is much higher than the weight of the shoe, this damping friction does not affect the operation any noticeable amount.

The main frame, as almost all other parts of the drum assembly, is made of a precipitation-hardening stainless steel. For the sake of rigidity and precision, it is fabricated from one piece of stock and provided with generous ribs.

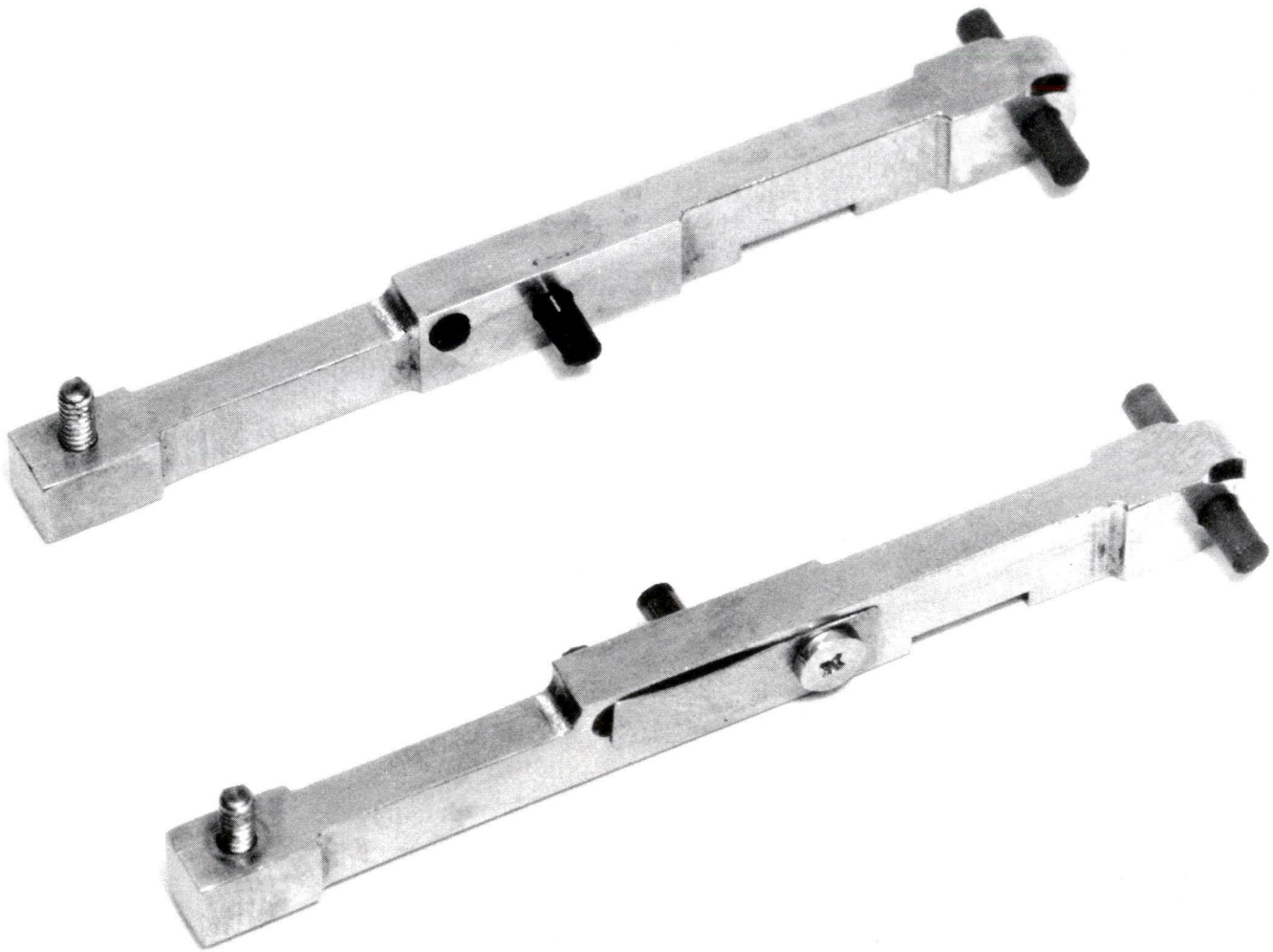


Figure 17. Typical Pair of Radius Arms

REFERENCES

(1) W. A. Gross. "A Gas Film Lubrication Study" Part I, "Some Theoretical Analyses of Slider Bearings," IBM Journal of Research and Development, vol. 3, No. 3 (July 1959), pp. 237-255.

(2) W. A. Michael. "A Gas Film Lubrication Study" Part II, "Numerical Solution of the Reynolds Equation for Finite Slider Bearings," IBM Journal of Research and Development, vol. 3, No. 3 (July 1959), pp. 256-259.

(3) R. K. Brunner, J. M. Harker, K. E. Haughton, and A. G. Osterlund. "A Gas Film Lubrication Study" Part III, "Experimental Investigation of Pivoted Slider Bearings," IBM Journal of Research and Development, vol. 3, No. 3 (July 1959), pp. 260-274.

(4) D. D. Fuller. Theory and Practice of Lubrication for Engineers, John Wiley and Sons, Inc. New York, N. Y. (1956).

(5) A. Kingsbury. "Experiments with an Air-Lubricated Journal" Journal of the American Society of Naval Engineers, vol. 9 (1897), pp. 267-292.

(6) J. S. Ausman and M. Wildman. "How to Design Hydrodynamic Gas Bearings" Product Engineering, (25 November 1957) pp. 21-28; 103-106.

(7) A. Brenner, Grace E. Riddell. U. S. Patent No. 2532283, (5 December 1950).

(8) A. Brenner, Grace E. Riddell. "Deposition of Nickel and Cobalt by Chemical Reduction," Bureau of Standards Research Paper No. RP 1835, vol. 39 (November 1947).