RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

Published quarterly by
RCA LABORATORIES
in cooperation with all subsidiaries and divisions of
RADIO CORPORATION OF AMERICA

VOLUME XXIII | JUNE 1962 | NUMBER 2

CONTENTS

RCA REVIEW is regularly abstracted and indexed by *Applied Science and Technology Index, Bulletin Signalétique des Télécommunications, Chemical Abstracts, Electronic and Radio Engineer, Mathematical Reviews, and Science Abstracts* (I.E.E.-Brit.).

RCA REVIEW

BOARD OF EDITORS

COMPUTER MEMORIES — POSSIBLE FUTURE DEVELOPMENTS*

By

Jan A. Rajchman

RCA Laboratories
Princeton N. J.

Summary—*Limits in attainable speed and storage capacity for magnetic memories of all types are estimated on basic, technological, and economic grounds. Cycle times of 100 nanoseconds for 4096 words are limiting for micromagnetic ferrite and thin metal films. Capacities of several million bits are the maximum attainable with reasonable economy by any magnetic technique. Tunnel-diode memories can obtain cycle times of 15 nanoseconds and superconductive memories may have storage capacities of billions of bits, thus extending beyond the limits of magnetic memories.*

Introduction

THE STANDARD INTERNAL MEMORY of present computers is a random-access ferrite-core memory.[1] Storage capacity and speed of access determine its usefulness. These two parameters are plotted for a number of typical commercial or commercially announced memories in Figure 1. The largest memory has about two million bits, the fastest has a cycle time of about 0.8 microsecond. Considerable work is in progress throughout the world to increase speed, and some work is being done to increase capacity.

This paper considers the reasonable limits to speed and capacity which can be attained with ferrites, thin magnetic films, or any other magnetic technology. Superconductive and tunnel-diode memories which show promise to exceed these limits are also considered.

Limits of Speed

Let us consider the maximum speed attainable in magnetic memories. We shall consider the cycle time T required to address, to read, and to write or rewrite the information and leave the memory ready for another such access. This cycle time is the reciprocal of the maximum possible repetition rate. In a memory of N words of m bits

* Manuscript received 11 April 1962.

[1] J. A. Rajchman, "Computer Memories — A Survey of the State of the Art," *Proc. I.R.E.*, Vol. 49, p. 104, Jan. 1961.

each, the speed is independent of m (since $N \gg m$ in all cases of practical importance) and we may confine ourselves to the consideration of a single-bit N-word memory. The cycle time T can be expressed as the sum of time required for each of various functions.

$$T = 2t_s + t_t + t_g + t_a + t_n$$

where t_s is the switching time, t_t the transmission time, t_g the amplification time, t_a the addressing time, and t_n the preread delay time. These times can be derived as follows.

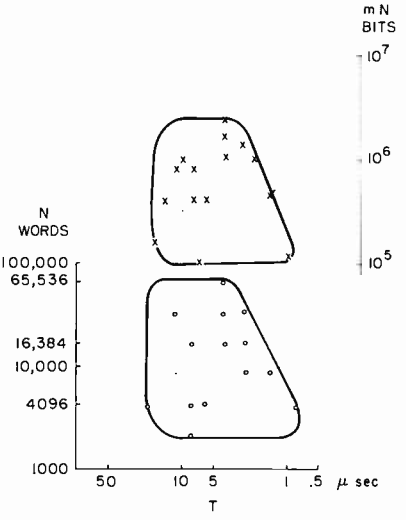


Fig. 1—Speed and storage capacity of present memories (1961).

Switching Time, t_s

The switching time, t_s, of a core or other magnetic element is found from the well-known relation

$$t_s\,(H - H_c) = S_w.$$

In a current-coincident mode, the magnetomotive force (mmf), H, can be no more than twice the coercive force, H_c, for otherwise the half-selected cores will switch; optimally, $H = 1.5H_c$. In general, this limits the speed with respect to a word-organized mode in which H can be many times H_c, so that $t_s \approx S_w/H$. The mmf, H, depends on the drive

current, I, and the length of the magnetic path. When the elements are packed as tightly as possible, the length of the magnetic path, in turn, is related to the element spacing, d. If we approximate geometrical factors which are slightly different for flat and circular elements, the following single relation can be assumed:

$$t_s = \frac{S_w d}{I},$$

where S_w is in oersted-seconds, d in centimeters, I in amperes, and t_s in seconds. A memory cycle necessarily involves two switchings of the core. This explains the factor of 2 in Equation (1).

Let us consider possible minimum values of t_s for ferrites and thin magnetic films. In either case a drive current of $I = 1$ ampere is about the maximum to be expected from the best transistor and about the maximum tolerable in small wires. In the case of ferrites, a value of $S_w = 0.5 \times 10^{-6}$ oersted-second is the lowest reported, so that with $d = 1.0$ mm, $t_s = 100$ ns (ns = nanosecond $= 10^{-9}$ second), and with $d = 0.1$ mm, $t_s = 10$ ns. In the case of films, values as low as 0.5×10^{-7} oersted-second have been consistently observed. But here the spacing cannot be arbitrarily small because a ratio of length to thickness of 10^3 to 10^4 is necessary to prevent deleterious demagnetizing effects. Spacings in the range of $d = 1$ cm, used in an early experiment, yield $t = 100$ ns, and spacings in the range of $d = 1$ mm, found possible in more-recent experiments, yield $t = 10$ ns.

Transmission Time, t_t

The transmission time, t_t, along the digit write and read windings which link all N elements is N times the delay at each element. If the spacing between elements is d centimeters and the electrical disturbance travels at K times the speed of light, c, then $t_t = Nd/Kc$. We will assume $K \approx 1/3$, not an unreasonable value for the transmission-line type of windings encountered in ferrite cores, ferrite plates, or thin-film memories. There is delay along both write and sense windings, but if these are grounded on opposite ends, the variations in write and read delay will compensate and have a sum of t_t. Typically $t_t = N \times 10^{-10}$, $N \times 10^{-11}$ or $N \times 10^{-12}$ for a spacing, d, of 1 cm, 1 mm, or 0.1 mm, respectively.

Amplification time, t_g

The voltage V obtained from the sense winding must be amplified sufficiently to set an output register, and then further to provide

voltage for the digit winding which inhibits (or partially drives) the whole plane. This voltage is made up of N times the voltage produced at each unselected element which, in turn, is the sum of the elastic voltages (unavoidably present in real nonideal rectangular-hysteresis-loop material) and the voltage due to the inductance of the windings between elements. Let us denote by r the ratio of these elastic voltages to inelastic voltage due to the switching of the selected core. In general $rN \gg 1$, so that the useful voltage, V, can be neglected in comparison with the sum of elastic voltages. The required voltage on the digit winding is thus

$$rVN + V \approx rVN.$$

The total voltage gain, G, required to amplify the sense voltage is therefore rN. Amplification necessarily entails time; a certain frequency bandwidth is necessary and a certain delay is encountered in the amplifier. We will make the simplifying assumption that in the best amplifier, with cascaded stages if necessary, the ratio of the gain, G, to the time, t_g, required to obtain it is a constant, $B = G/t$. This constant B is characteristic of the amplifying element used—tube, transistor or tunnel diode. Also we assume, optimistically, that the current gain which is required in addition to voltage gain is obtained without further loss of time. Therefore the minimum time required for amplification is $t_g = rN/B$.

A typical value of r is 0.1 for good hysteresis squareness and reasonably tight spacing of the elements. Values of B vary a great deal with transistors. Not uncommon are amplifiers for which $B = 10^9$ sec^{-1}. In advanced experimental types, B may be as great as 10^{10} sec^{-1}. Tunnel-diode amplifiers with an equivalent B of 10^{11} sec^{-1} have been made. The corresponding delays t_g are thus $N \times 10^{-10}$, $N \times 10^{-11}$, and $N \times 10^{-12}$.

In general, the input write signal must be amplified, and this takes time. Usually the input signals are at reasonably high level, however, so that this time can be neglected in comparison to that required to amplify the sense signals.

Address Time, t_a

The time, t_a, required to "address" is defined as the time required for decoding the address. In general, the inputs to the memory are at a sufficiently high level and need not be amplified before decoding. Consequently, this operation can proceed in a time t_a which is short with respect to t_g and will be neglected for simplicity.

"Preread Delay" Time, t_n

The time t_n is in practice a dominating factor. In general, the digit write (inhibit) is coupled to the sense winding so that a large signal appears in the read circuits during writing. This signal can be orders of magnitude greater than the desired read signal. Cancellation systems, balanced circuits, and other methods are generally used to keep this pick-up at a minimum, but it still produces a signal considerably greater than the sense signal. Cancellation cannot, in general, be perfect due to delta effects, i.e., due to the fact that elastic couplings are dependent on the remanent state of the core. Clamps can be used on the sense amplifier to prevent it from being paralyzed by this large signal. In practice most of these remedies do not avoid the necessity of allowing a certain time for the decay of the transients due to writing at one address before it is possible to sense at this or another address. This is the time t_n. In typical 5-μsec memories, t_n is about 1 μsec.

To gain an idea of the highest attainable speeds, let us consider only the switching time of the element, the delay through the array, and the time required for amplification. These times are dependent on basic properties of the magnetic material, the size of the memory, and the limits of the associated electronics. We neglect the very important "preread" delay time on the optimistic assumption that further ingenuity in balancing and clamping will render this delay negligible. Similarly we neglect the decoding time, t_a.

Figure 2 shows a number of curves relating the cycle time, T, to the number of words, N. The curves are for four switching times, 10^{-6}, 10^{-7}, 10^{-8}, and 10^{-9} second. For a one-bit word the cycle time is simply twice the switching time. As N increases, a term proportional to N is added, since both t_t and t_g are proportional to N. The curves are drawn for $t_t = t_g$. This is a convenience, and is justified since the efforts in the last few years in microminiaturization to reduce t_t and in semiconductor amplifier to reduce t_g seem to have roughly resulted in a balance between the importance of these two factors. In any case it is easy to see on the curves the effect of having one of the times dominating the other.

In the case of ferrite cores, outside diameters of 18 mils and inside diameters of 13 mils are found in advanced memories. These cores switch in about 10^{-7} second (100 nanoseconds) for a current of roughly one ampere. Spacings of $d = 1$ cm are commonly used. Present transistor circuitry is characterized by roughly $B = 10^9$ sec^{-1}. Therefore

the curve corresponding to $t_t + t_g = 2N \times 10^{-10}$ represents about the best state of the art. A memory of 1024 words has a minimum cycle time of .4 and .5 μsec, and one of 4096 words has a cycle of 1 μsec. Assuming an easily achievable spacing of $d = 1$ mm and somewhat more difficult, but possible circuitry, $B = 10^{10}$ sec^{-1}, the curve for $t_t + t_g =$

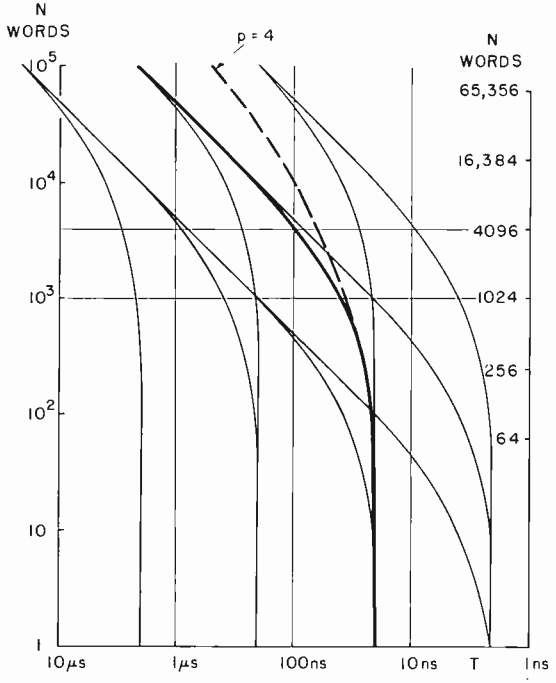


Fig. 2—Relation between speed and storage capacity of magnetic memories. Curves are for $t_s = 10^{-6}$, 10^{-7}, 10^{-8} and 10^{-9} second, and for $t_t + t_g = 2N \times 10^{-10}$, $2N \times 10^{-11}$, and $2N \times 10^{-12}$ second.

$2N \times 10^{-11}$ shows that cycle times of 200 ns for 1024 words and 300 ns for 4096 words are limiting. The next substantial improvement must result from a decrease in switching time. This can be obtained simply by reducing the size of the core. The degree of reduction depends on the skill of technology. It is unreasonable to assume a reduction below 2 or 3 mils. A slight improvement in the switching coefficient S_w, perhaps by 50% can be assumed. This would then lead to a switching time of 10 nanoseconds which, with $t_t + t_g = 2N \times 10^{-11}$, corresponds to the heavy curve in Figure 2. Typically, a memory with 4096 words could then be cycled at the limit in 100 nanoseconds.

The outstanding virtue of thin magnetic films is fast switching.[2,3] However, because elements smaller than about 1 mm would entail deleterious demagnetizing effects, and practical limits of drive current are about 1 ampere, the fastest usable elements still take about 10 nanoseconds to switch. Optimistically, with some reduction in size and a considerable improvement of S_w, a limiting switching time of 1 nanosecond could be assumed. Because a drastic reduction of size is not possible, the transmission time is at best $t_t = N \times 10^{-11}$ sec. The best transistor circuitry can perhaps yield a $B = 10^{10}$ sec^{-1}, as was assumed for ferrites, but here there are practical difficulties with dealing with a much smaller signal, so that it may be more difficult to achieve it. In any case the heavy line ($t_s = 10$ ns, $t_t + t_g = 2N \times 10^{-11}$) of the figure seem to be about the ultimate limit for both thin films and microferrites.

In the foregoing analysis we have assumed "pure" magnetic memories in which all the N magnetic elements of a given bit of the words are coupled through a single sense winding and a single digit winding. Considerable speed-up is possible by splitting these windings into p parts each. The delay time, t_t, and the required gain, and consequently the amplification time, t_g, are both proportional to N/p in such a "compound" memory. Furthermore, in the case of a current-coincident memory, the disturbs are reduced by the factor p. Of course, such partitioning requires more electronics. The number of required digit practice, the largest "pure" planes have 4096 cores. Planes with 16,384 drivers and sense amplifiers is increased by p. In present commercial cores are usually partitioned at least four times.

One could believe that with sufficient partitioning the effects of t_t and t_g could be reduced at will, but this is not possible since, because of wiring requirements, t_t cannot be zero and there is definite minimum time for amplification required for current gain and losses. A reasonable reduction of $t_t + t_g$ of about 4:1 can be expected. The result of such a reduction is plotted on the figure for the case $t_s = 10$ ns and $t_t + t_g = 1/4$ ($2N \times 10^{-11}$) showing this improvement with respect to the heavy limiting curve. With such partitioning, a memory with 16,384 bits could, in principle, be cycled in 100 ns.

TUNNEL-DIODE MEMORIES

What are the prospects of an increase of speed by an order of magnitude with respect to these magnetic memories?

---

[2] A. V. Pohm and E. N. Mitchell, "Magnetic Film Memories — A Survey," *Trans. I.R.E. PGEC*, Vol. EC-9, p. 308, Sept. 1960.

[3] J. I. Raffel, "Operating Characteristics of a Thin Film Memory," *Jour. Appl. Phys.*, Vol. 30, p. 605, April 1959.

A 10-ns cycle time could be obtained with an element switching in 1 ns. The capacity would be 4096 bits for elements spaced 0.1 mm apart, provided, of course, that the amplifiers would work at sufficiently high speed (See Figure 2). These severe requirements can be met by using tunnel diodes for the storage element as well as for the amplifiers and drivers. The bistable character of the device and its sharp nonlinearities permit operation of an array in voltage or current coincidence. Faster speed is obtained in a word-organized mode where coincidence is used only for writing, as in magnetic memories. Special "backward" tunnel diodes having a nearly flat region in place of the usual peak–valley region, can be used to couple the storing diodes to the sense amplifiers. This arangement reduces considerably the otherwise severe attenuation of the signal and thereby permits obtaining the output read signal in only 1 to 3 nanoseconds when tunnel-diode circuitry is used.[4]

Tunnel-diode memories with cycle times of 10 to 25 ns and capacities of about 1024 words are in an experimental state of development. Their probable region of utility is shown in Figure 4.

Limits of Capacity

Let us consider the largest storage capacity attainable in magnetic memories. One could think of the question simply in economic terms. For example, one could obtain a billion-bit memory by buying one thousand conventional memories of one million bits each. Such a memory would cost approximately one billion dollars. Furthermore, the delivery time would likely be very long. This can be appreciated by considering that automatic machines making and wiring elements at the rate of one per second would require more than 30 years to complete a billion-bit memory! Even with higher fabrication speeds, parallel operations, etc. a bit-by-bit construction would still be too lengthy to be acceptable. Clearly then, means must be found to speed up construction by a batch fabrication technique. Also, the cost must be reduced.

There is probably no fundamental limit to attainable storage capacity if the consequent loss of speed is accepted. A meaningful question is therefore to ask what price must one pay for an increase in capacity, let us say from todays million bits to a billion bits.

Considerable progress has been made with certain batch fabrication techniques. For example, as long as five years ago an apertured ferrite

[4] M. M. Kaufman, "A Tunnel-Diode–Tunnel-Rectifier 15-Nanosecond Memory," *Solid/State/Design*, Vol. 3, p. 23, Feb. 1962.

plate with 256 bits and with printed windings was developed.[5] More recently, some of the work with thin magnetic films which has been conducted in various laboratories is aimed primarily at batch fabrication[6] rather than, or in addition to, speed. Of significance are recently announced[7] etched permalloy sheets in which many elements and their necessary windings are fabricated by photoetching techniques. Also notable are efforts to make arrays by electroplating wires[8] or meshes. Various degrees of success have been achieved, but thus far no method has speeded up fabrication or reduced cost with respect to bit-by-bit fabrication by the sought factor of a thousand. It is likely, however, that further vigorous development along the lines already proposed or others will yield this factor within the next few years.

Unfortunately, this is not sufficient. Magnetic materials, even with the best hysteresis squareness, simply do not have sufficient nonlinearity and other properties to provide all the required switching without the aid of considerable electronics. To appreciate the amount of electronics necessary, let us estimate the number C of single semiconductor switches, transistors, or diodes required for a memory of N words of m bits each.

Addressing requires $2\sqrt{N}$ driving channels in current-coincident mode and N channels in a word-organized mode. Each channel must be dual to provide the two polarities. Furthermore, address decoding requires multiplying these numbers by at least 2. Therefore, about $8\sqrt{N}$ elements for current-coincident mode and $4N$ for word-organized mode are fair estimates.

Magnetic switches can be used for addressing. Most practical is an x–y d-c biased switch, in which the x drive overcomes the bias and the y drive provides the switching and useful output of the selected core. A large switch of this sort would of course have to be made by integration techniques similar to those used for the memory. The number of address drivers would be $2\sqrt{N}$, but here the number used need not be doubled since both polarities are obtained due to the restoring effect of the d-c bias. On the other hand, the inefficiency of the switch requires more power in the drivers, so that the estimate of $8\sqrt{N}$ semi-

---

[5] J. A. Rajchman, "Ferrite Apertured Plate for Random Access Memory," *Proc. I.R.E.*, Vol. 45, p. 325, March 1957.

[6] E. M. Bradley, "Making Reproducible Magnetic Film Memories," *Electronics*, Vol. 33, p. 78, Sept. 1960.

[7] G. R. Briggs and J. W. Tuska, "Permalloy-Sheet Transfluxor-Array Memory," *Proc. Magnetism and Magnetic Materials*, 7th Conf., Phoenix, Arizona, Nov. 13-16, 1961; supplement to *Jour. Appl. Phys.*, Vol. 33, p. 1065, March 1962.

[8] T. R. Long, "Electrodeposited Memory Elements for a Nondestructive Memory," *Jour. Appl. Phys.*, Vol. 31, p. 123, May 1960.

conductor switches for current-coincident addressing will still be approximately valid in this case.

For writing and reading, a single channel for each of the m bits would be sufficient if the magnetic elements were ideal, but because they are not for large N's, the writing and sensing windings must be partitioned. For writing, the maximum number N_w of elements per windings is $V_d/r\ V$, where V_d is the maximum voltage of the driving transistor. Typically if $V_d = 10$ volts, $V = 10$ mv, and $r = 0.1$, then $N_w = 10^4$.

For reading in a word-organized mode, the maximum number N_R of elements per sense winding is determined by the attenuation produced by the large series impedance of the winding. An estimate of $N_R = 10^4$ is not unreasonable. In a current-coincident mode the $2\sqrt{N}$ half-select disturbs each add a fraction, f, of the sense voltage, V. The magnitude of this fraction depends on the squareness of the hysteresis loop of cores, on their uniformity (since cancellation effects are used), and on the delta effect at strobing time. For extremely good cores, f is barely 400, so that if a signal-to-disturbs ratio of 2:1 is tolerated, $N = 10^4$ at the limit.

It is thus evident that when N exceeds $N_R = N_w = 10^4$, the sense and digit write windings must be partitioned. The partitioning factor p is N/N_w. We will assume that the sense amplifier, and the write amplifier with its logic rewrite, require 10 semiconductor elements each. Consequently the total number of elements for sensing and writing is $20\ m\ N/N_w$, assuming $N_R = N_w$.

The total number of semiconductor devices for addressing, writing into, and sensing from a memory of N words of m bits each is thus, for $N > N_w$, approximately

$$C = 8\sqrt{N} + 20\ m\frac{N}{N_w}, \qquad \text{(current-coincident mode)}$$

$$C = 4N + 20\ m\frac{N}{N_n}. \qquad \text{(word-organized mode)}$$

(For $N < N_w$ the additive terms are $20m$ when partitioning is not necessary.) These relations are plotted in Figure 3 for $m = 100$ and $N_w = N_R = 10^4$.

It is apparent that a billion-bit memory requires about twenty million semiconductor devices if operated in current coincidence, and 400 million if operated word organized. In any case it is clear that

batch fabrication of the semiconductor devices is indispensable. A great deal has been done, and even more has been said, about batch fabrication of semiconductor devices together with the necessary circuit elements, i.e., integrated circuits. To date, batch fabrication of about ten elements is a modest reality, but a number of experimental approaches for much greater integration show good promise.

What is then the overall prospect for a billion-bit magnetic memory? Great strides in magnetic and semiconductor batch fabrication are

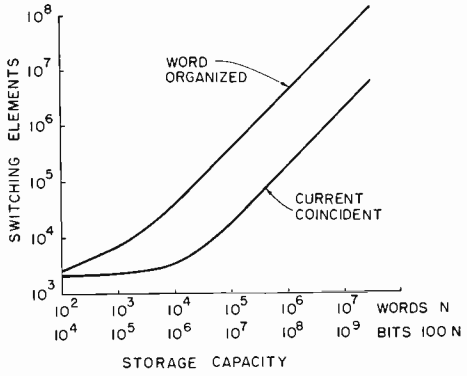


Fig. 3—Number of semiconductor switching elements required for given memory storage capacity.

necessary. At the moment, there is more progress in this respect in magnetics (where it is perhaps easier) than in semiconductors. It seems therefore that a current-coincident mode which taxes magnetics more severely for squarer and more uniform elements but requires less electronics will come into being before a word-organized mode which, while it is more lenient in magnetics, requires an order of magnitude more electronics. In any case, the development is likely to be gradual unless spectacular success in material improvement or a brilliant invention, or more likely both, should permit a sudden jump. More likely, this evolution may be stopped through the success of superconductive memories.

Superconductive Memories

Superconductive phenomena are essentially ideal for memory applications. Persistent supercurrents are a natural form of storage, and sharply defined thresholds between the superconductive and normal states permit switching. Moreover, thin superconductive film tech-

nology offers the unique possibility of simultaneous miniature batch fabrication of storage elements, addressing switches and all connections.

A short description of a superconductive continuous sheet memory may be in order since it has been reported only recently.[9] Two perpendicular sets of parallel suitably insulated lead strips are evaporated onto a continuous film of tin. When an x and a y strip carry a current, I, the magnetic field pattern at their intersection is at 45° with respect to the strips. The intensity of the field is maximum at the intersection and diminishes gradually with distance. Consequently, there is a region limited by a definite boundary within which the field in the tin sheet exceeds a critical value and renders the sheet normal and outside of which it does not. Within that region, it is possible to induce persistent currents and change the polarity of existing persistent currents. The final polarity obtained depends on the polarity of the primary currents and determines whether a one or a zero is stored. An element is switched only if the primary excitation is opposite to that which previously established its state, and only if it exceeds a certain definite threshold. The situation is thus quite analogous to a hysteretic magnetic element with a perfect square loop. A voltage is induced in a sense winding on the opposite side of the memory plane due to the local magnetic field leaking through the plane at the selected position. No field leaks elsewhere through the superconductive sheet because it is a perfect magnetic shield; thus, there are no disturb signals due to half excitations. Addressing is done by a decoding tree of cryotrons feeding the selected lines. These cryotrons are fabricated at the same time as the memory and are made of the same materials: lead and tin.

The superconductive memory can be used in a conventional semiconductor computer. The drivers for the cryotrons, the sense circuits, and the write and rewrite circuits are of the conventional transistor type. Here, however, in contrast to magnetic memories, the number of circuits increases only very moderately with capacity. For addressing, drivers are necessary only for the binary bits of the address. The writing and sensing circuits need not be partitioned.

The technique thus seems to offer all the prerequisites for making a large-capacity memory — integral miniature batch fabrication techniques and very small requirements of auxiliary circuits. Capacities of billions of bits may be achieved and thus furnish the computer art, for the first time, the large capacities available today only in electro-

---

[9] L. L. Burns, Jr., G. W. Alphonse, and G. W. Leck, "Coincident-Current Superconductive Memory," *Trans. I.R.E. PGEC*, Vol. EC-10, No. 3, p. 438, Sept. 1961.

mechanical memories. The accesses at electronic speeds will be measured in microseconds rather than in seconds or minutes.

The large capacity possible with superconductive memories justifies the use of liquid helium. Today, such use is a laboratory inconvenience, but tomorrow closed-cycle refrigeration units will be as commonplace as air-cooling or air-conditioning systems.

While the superconductive memory offers a unique solution to large capacity memories, it promises also to be fast. Its speed is chiefly limited by the time required to address it through the cryotron tree. Cycle times of about a microsecond may be expected with conventional cryotron designs for capacities of millions of words.

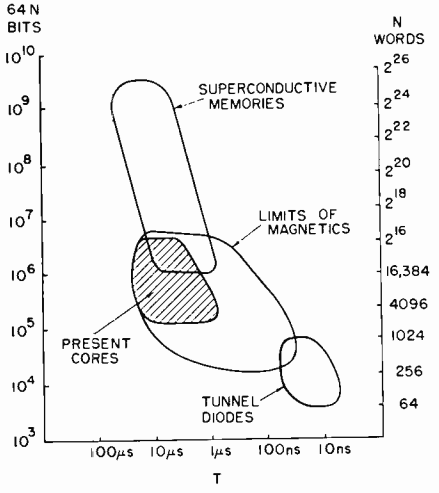


Fig. 4—Limits of speed and storage capacity of magnetic, superconductive and tunnel-diode memories.

Conclusions

The estimates and speculations described above can be summarized by the diagram of Figure 4 with the following remarks.

1. Present ferrite-core technology has achieved capacities of about two million bits and cycle times as short as 0.8 microsecond.

2. Extension of magnetic technology to faster memories is likely to be limited. The limit depends on the number of words and is about 100 ns for 4096 words or possibly for 16,384 words. Ferrite and thin-film technologies may be pushed to this limit. Both present serious engineering difficulties, which are perhaps more severe with thin films.

3. Tunnel-diode memories may extend speed beyond the limit of magnetic memories by a factor of 5 to 10, for capacities of about 1024 words.

4. Extension of magnetic memories to large capacities depends on economic factors. Substantial extension requires batch-fabrication techniques for both the magnetic elements and the semiconductor circuitry. Progress in semiconductor batch fabrication must be far more advanced to make this a real possibility.

5. Superconductive thin-film technology — and particularly the continuous sheet superconductive memory — offers the possibility of a very large capacity memory because of its ideal storage and switching characteristics combined with batch fabrication and high packing density.

Foretelling limits for any part of the explosively advancing art of electronics is perhaps foolhardy. Brilliant inventions or breakthroughs in materials, or both, may render reasonable limits of today completely meaningless tomorrow. This may well be the case for the estimates given here. Nevertheless, a reasonable attempt at circumscribing possible gains with given approaches helps to set the field in perspective.

The backbones of computers are memories and a possible course of events may be the development of other types of memories. The advent of random-access memories was essential to the coming into being of the modern stored-program computer as we know it. Memories with a more general type of access may well dictate new forms of computers. In these, there may be less emphasis on speed as more-direct solution methods may become possible. For example, it may be possible to address a memory through the content of part of the stored information itself and retrieve the remaining part associated with it. Such content-addressable or associative memories could make searching for information very simple and could either simplify or make unnecessary such tasks as ordering, merging, sorting, and collating of information which require so much time in today's serial processing. Content addressability entails mixing logic functions with storage functions. This is precisely what is necessary to achieve really large capacities in random-access memories, so that all the considerations on integrated batch fabrication are valid here. Superconductive techniques are thus a prime candidate for content-addressable memories and have already been used in small-scale experiments for that purpose.[10] Magnetic techniques with judicious use of semiconductors may well provide the first proto-

[10] A. E. Slade and C. R. Smallman, "Thin Film Cryotron Catalog Memory," *Symposium on Superconductive Techniques for Computing Systems*, May 1960.

types.[11] The real solution could be through the use of integrated semiconductor techniques such as thin-film transistors.[12] At this time it is too early to formulate a sensible judgment on the technique best suited to content-addressable memories, but it is already clear that memories of this type will be very important in the future.

---

[11] W. L. McDermid and H. E. Petersen, "Magnetic Associative Memory System," *IBM Jour.*, Vol. 5, No. 1, p. 59, Jan. 1961.

[12] P. P. Weimer, "Evaporated Circuits Incorporating a Thin Film Transistor," *1962 International Solid-State Circuits Conf.*, Phila., Feb. 1962.

HIGH-SPEED LOGIC CIRCUITS USING TUNNEL DIODES*

By

R. H. Bergman, M. Cooperman and H. Ur

RCA Electronic Data Processing
Camden, N. J.

Summary—The work reported here has been aimed at a set of logically complete, general purpose tunnel-diode–tunnel-rectifier logic circuits. Emphasis has been placed on high speed, reliability, and a design for use in large-scale computers. The circuits are d-c powered and do not require a high-powered clock source. Worst-case design philosophy has been utilized to obtain reliability. Monostable circuits are used to perform logic functions and obtain gain, while bistable circuits are used for storage in registers. Solutions to several important problems were necessary to achieve the required performance: the tunnel rectifier is used to provide directionality; monostable circuit recovery time is reduced by nonlinear biasing; the timing problem is reduced by gating pulses against levels; high-frequency signal distribution is achieved by the use of terminated transmission lines; and finally, inversion is achieved through use of a pulse-inverter circuit.

A complete design of the circuits is based on both static and dynamic analysis for obtaining optimum circuit parameters. The static analysis is based on worst-case variations of circuit parameters, including aging effects and safety factors for noise and cross-talk immunity. The dynamic analysis employing measured characteristics, graphical calculations, and analog and digital simulation is used to obtain waveforms, delays, and signal amplitudes.

Laboratory results confirm the calculated ones. Coupling between circuits including the effects of transmission lines is also analyzed. A detailed design is given for OR, AND and bistable circuits having one-nanosecond logic delays and 200-mc repetition rates.

Introduction

IN THE PAST SEVERAL YEARS widespread interest has been shown in the use of tunnel diodes in digital computers.[1,2] This interest has been due mainly to the extremely fast switching capabilities of the device. However, difficulties in obtaining directionality, gain, and inversion with this two-terminal device have led to circuit schemes requiring high-frequency a-c power supplies. The work de-

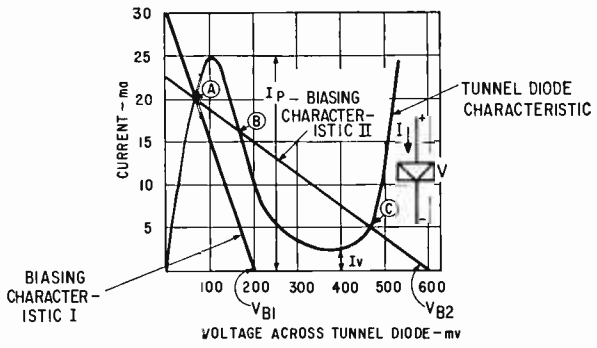
* Manuscript received 8 March 1962.

[1] R. H. Bergman, "Tunnel Diode Logic Circuits," *Trans. I.R.E. PGEC*, Vol. EC-9, p. 430, Dec. 1960.

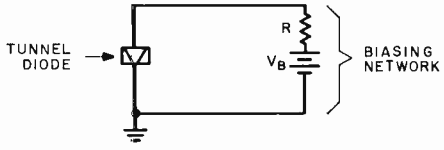
[2] M. H. Lewin, "Negative-Resistance Elements as Digital Computer Components." *Proc. Eastern Joint Computer Conf.*, No. 16, p. 15, Dec. 1959.

scribed in this paper was directed toward providing circuits which would operate from d-c supplies and would be capable of handling timing problems caused by the increased effect of wiring delays at these speeds.

A set of circuits has been designed on this basis with one nanosecond (10^{-9} second) logic delays and 200-mc repetition rates. Directionality is obtained by the use of low-peak-current tunnel diodes called tunnel



(a)



(b)

Fig. 1—(a) Two modes of operation for tunnel diode and (b) tunnel-diode circuit.

rectifiers. Gain is provided through the use of regenerative pulse amplifiers. For register and flip-flop a level-producing bistable circuit is employed. Timing problems are reduced by gating levels against pulses wherever necessary.

Monostable Circuits

The tunnel diode has properties which enable it to be used as a fast switching device. A typical characteristic of a tunnel diode, shown in Figure 1a, exhibits a voltage-controlled negative resistance. With ap-

propriate biasing, either a monostable or a bistable mode may be established.[3]

A circuit for obtaining these modes of operation is shown in Figure 1b. Two possible biasing characteristics that might result from this circuit are shown superimposed on the tunnel-diode characteristic of Figure 1a. Biasing characteristic I provides one stable operating point, designated by A; while biasing characteristic II provides two stable operating points, designated as A and C. The intersection at point B is unstable.

Basic Monostable Circuit

The monostable mode provides a practical way of obtaining amplification in tunnel-diode logic circuits. A simple monostable circuit is shown in Figure 2a, and operation of the circuit is illustrated graphically in Figure 2b. R and V_B form the biasing characteristic which intersects the tunnel-diode characteristic at point 1 of Figure 2b. To simplify the analysis, it will be assumed that the inductance L of Figure 2a is relatively large, so that there is negligible change of current in R as the diode switches from point 1 to point 3. Under these conditions, when a current step input is applied to node 1 of Figure 2a, it all goes into the tunnel diode. When the total current into the tunnel diode exceeds the tunnel-diode peak current, the operating point switches to point 3 along the trajectory indicated by the dotted lines in Figure 2b. The trajectory is a plot of the total current available to the diode versus the voltage across the diode. Thus, during the time that the diode is switching from point 1 to point 3, a current of $I_{in} + I_B$ is available, resulting in a constant-current dynamic load line of this magnitude.

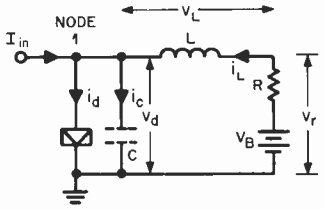
The speed of switching from point 1 to point 3 depends on how quickly the diode capacitance C can be charged to the new voltage. In general, the instantaneous rate of change in voltage across a capacitor C may be expressed as

$$\frac{dv}{dt} = \frac{i_c}{C}, \tag{1}$$

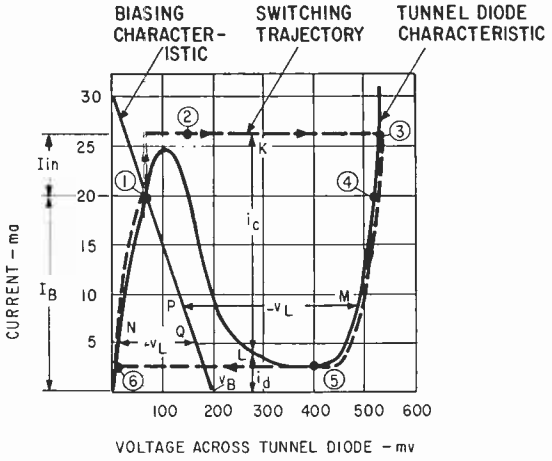
where v is capacitor voltage, i_c is the charging current, and t is time. In this case, i_c is the difference between the total available current, $I_B + I_{in}$, and the current i_d. The instantaneous charging current, i_c, is shown graphically in Figure 2b as the distance KL. Thus, for a given diode, the larger the distance KL, the faster the switching takes place.

[3] H. S. Sommers, "Tunnel Diodes as High Frequency Devices," *Proc. I.R.E.*, Vol. 47, p. 1201, July 1959.

(a)



(b)



(c)

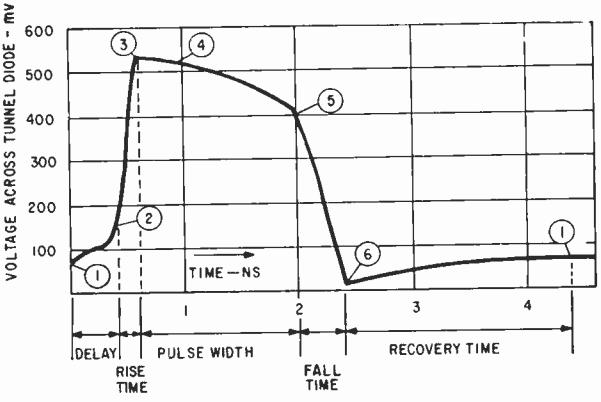


Fig. 2—Monostable circuit: (a) basic circuit, (b) graphical analysis of switching, and (c) output voltage waveform.

If the input is removed immediately after the diode reaches point 3, the operating point of the tunnel diode moves to point 4. This is not a stable state, however, and the current in L decays causing the dynamic operating point to proceed toward point 5. (If biasing characteristic II of Figure 1a had been used, there would be a stable intersection somewhere between points 4 and 5. This intersection would then become a static operating point and the diode would remain in that state indefinitely. This state represents a storage of information and forms the basic property upon which the bistable circuit to be described later is based.)

The speed of moving from point 4 to point 5 depends upon how quickly the current in the inductance can change. In general, the instantaneous rate of change of current in an inductance, L, may be expressed as

$$\frac{di_L}{dt} = \frac{v_L}{L}, \tag{2}$$

where i_L is the current in the inductance, v_L is the voltage across the inductance, and t is time. In this case, v_L is the difference between the biasing network voltage, v_r, and the tunnel-diode voltage, v_d (Figure 2a). The voltage v_L is shown graphically in Figure 2b as the distance PM. Thus, for a given inductance, the larger the distance PM, the faster the dynamic operating point moves (relaxes) from point 4 to 5. When point 5 is reached, the current in L continues to decrease since v_L is not zero. This action takes the diode once again into the negative-resistance region, causing it to switch to point 6 by a mechanism similar to that which caused it to switch from points 1 to 3.

Point 6 is still not stable since v_L is not zero. Note, however, that v_L has changed sign, which causes the current in L to increase until point 1 is reached. Here again, for a given inductance the larger the distance QN, the faster the operating point moves (relaxes) from point 6 to point 1. Point 1, the initial operating point, is stable since both v_L and i_c are zero.

This transient analysis is general, and also applies for nonlinear biasing characteristics which, as shown later, are more desirable. This method of analysis has been employed successfully to study the transient behavior of nonlinear circuits of much greater complexity.

The voltage v_d obtained from the switching trajectory of Figure 2b, is shown as a function of time in Figure 2c. The waveform of the complete cycle is subdivided into time intervals as indicated. The delay time is a function of i_c/C.

The magnitude of i_c at the tunnel-diode peak is called the overdrive, and is expressed as a percentage of the tunnel-diode peak current, I_p. The relationship between total switching time (delay plus rise time) and overdrive is shown in Figure 3. This curve is valid for the circuit of Figure 2a.

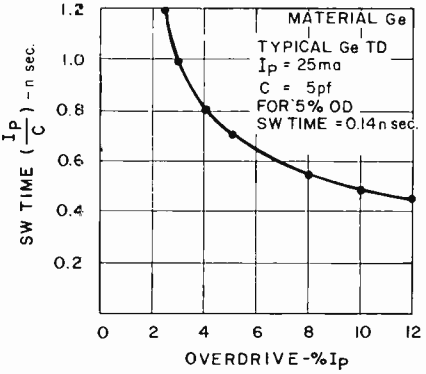


Fig. 3—Normalized switching time versus overdrive.

The rise time and fall time are also functions of i_c/C in their respective regions. The pulse width and recovery time are functions of L/v_L in their respective regions.

Thus, the monostable circuit provides a reshaped voltage pulse and can supply an output current approximately equal to I_B (Figure 2b) which is several times the magnitude of I_{in}. Thus, current amplification is obtained.

For high-speed operation, it is necessary to reduce the total cycle time of Figure 2c to a minimum. Most of this time is made up of the pulse width and recovery period; these can be reduced by reducing the inductance as explained previously. However, to obtain the operating speed of interest (5 nanoseconds total cycle time for tunnel diodes having $I_p/C = 5$ milliamperes/picofarad), the inductance has to be made so small (5 nanohenries, or 5×10^{-9} henries) that it no longer blocks I_{in} from flowing into R, as asumed in Figure 2a. Therefore, I_{in} has to be increased to I_{in1} as shown in Figure 4a. This reduces the current amplification, thus reducing the usefulness of this circuit.

This situation is remedied with the nonlinear biasing characteristic of Figure 4a. Such a characteristic is obtained with the current source, I_B, and tunnel rectifier (TR) of Figure 4b. The idealized characteristic of the tunnel rectifier is shown in Figure 5. In the circuit of Figure

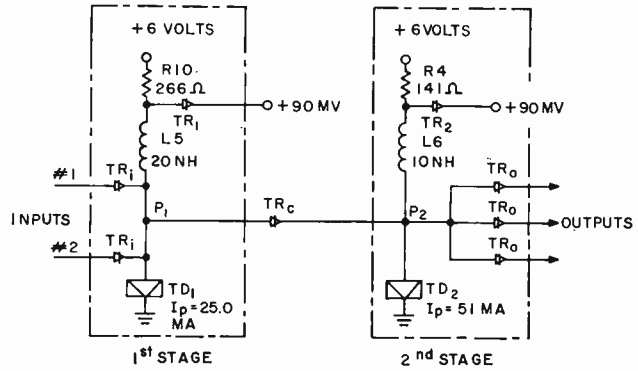


Fig. 6—Tunnel-diode OR gate.

2. TR_c provides directionality so that after TD_2 fires, TR_c blocks any reverse current flow, thus preventing TD_1 from acting as a load on TD_2.

3. The initial high resistance of TR_c minimizes static interaction between the two stages.

The purpose served by the input and output tunnel rectifiers as coupling elements between gates is the same as that served by TR_c in coupling the stages. There are, however, some additional requirements which must be met when interconnecting gates.

1. A higher degree of directionality. This is illustrated in Figure 7, where OR gate A and OR gate B drive OR gate C. Firing OR gate A causes OR gate C to fire. Unless good directionality is provided in the path between gate C and B, gate C will fire gate B. The same reasoning applies when gate B drives gate C.

2. The need for transmission lines to interconnect gates makes it

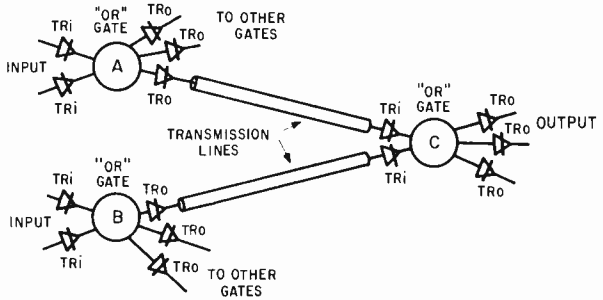


Fig. 7—Coupling between gates to prevent back triggering.

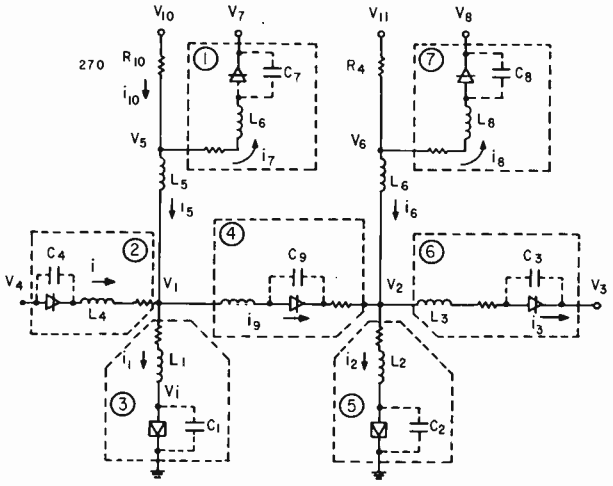


Fig. 8(a)—Equivalent circuit of OR gate.

desirable to use coupling elements at both ends of the transmission line (i.e., at input and output of gate) in order to minimize reflections.

In order to arrive at an optimum design of the OR gate shown in Figure 6, the dynamic behavior of the circuit was simulated on a computer. To obtain sufficiently accurate information, the stray circuit elements were included. This resulted in the equivalent circuit shown in Figure 8a. The switching trajectories obtained with an analog computer are shown in Figures 8b and 8c.

The loop in the trajectory of the first stage (Figure 8b) is the result of the second stage firing which causes an increase of current through the tunnel diode of the first stage.

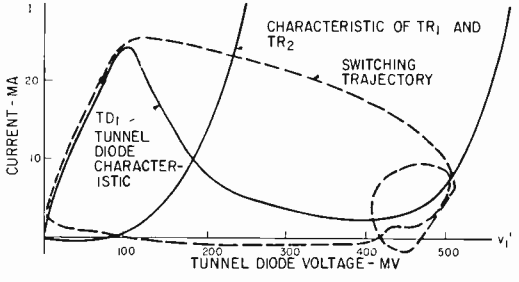


Fig. 8(b)—Switching trajectory of first stage of OR gate.

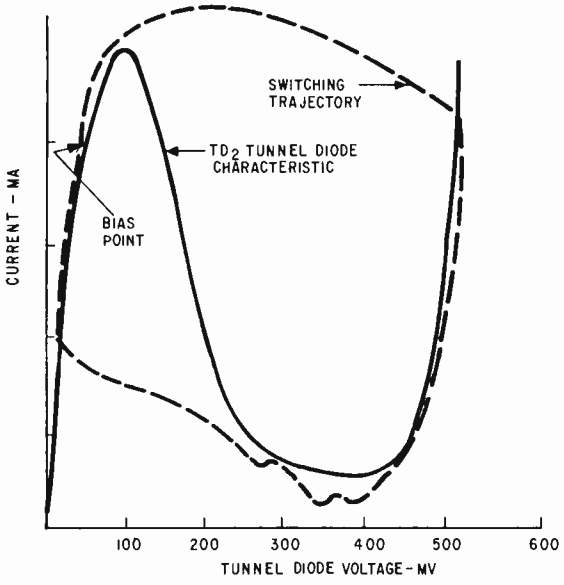


Fig. 8(c)—Switching trajectory of second stage of OR gate.

Tunnel Diode AND Gate

The OR gate produces an output if any one of the inputs is high. If, however, the bias current of the first stage of Figure 6 is reduced so that one input is insufficient to switch it, but the sum of the currents of two inputs are sufficiently large to produce switching, the basic property of an AND gate is obtained. Such an AND gate is shown in Figure 9.

One of the AND gate inputs is a level input obtained from a bistable circuit, the other input is a pulse from an OR gate or another AND

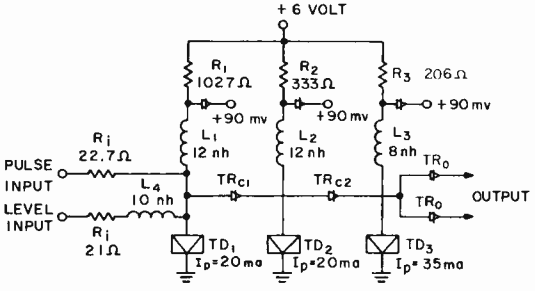


Fig. 9—Tunnel-diode AND gate.

gate. The use of a pulse gated against a level avoids the timing problems encountered when a pulse is gated against a pulse, particularly at high speeds.

Experimental Circuits

The OR gate of Figure 6 and the AND gate of Figure 9 exhibit the following performance:

OR Gate

2 inputs of 7.4 ma (minimum) each
3 outputs of 10 ma (nominal) each
Gate Delay: 0.7 ns (nominal value)
Repetition Rate: 200 mc (minimum value)

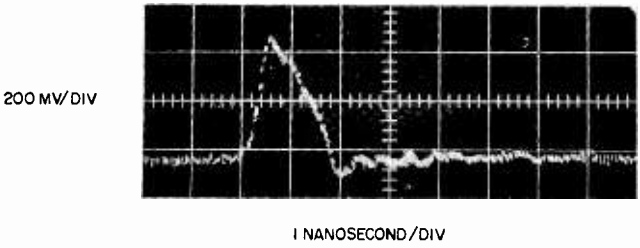


Fig. 10—Photograph of an OR gate output waveform.

AND Gate

2 inputs of 10 ma (nominal) each
2 outputs of 10 ma (nominal) each
Gate Delay: 1.2 ns (nominal value)
Repetition Rate: 200 mc (minimum value)

A typical output waveform of an OR gate is shown in Figure 10.

D-C Tolerance Analysis

To determine the gain obtainable from the circuits and specify components and their variation, a worst-case tolerance analysis was performed. In the analysis, variations of power supplies, bias resistors, tunnel diodes, and rectifiers were assumed. From these specified variations and other imposed conditions, the number of stages needed to provide required gain or fan-out was determined. Some of the additional conditions were: a guaranteed minimum overdrive when switching of a tunnel diode was desired; a safety factor when switching was not desired; and an allowance for undesired transients when the circuits were interconnected.

OR Gate

In the design of the OR gate, a minimum fan-in of two and a fan-out of three was required. The three relations shown below were used in the worst-case analysis of the OR gate.

$$I_{B\,\max} \leq I_{p\,\min} - I_S - I_{L1} \tag{3}$$

$$i_{\text{in}\,\min} \geq I_{p\,\max} - I_{B\,\min} + I_O + I_{L2} \tag{4}$$

$$i_{\text{out}\,\max} \leq I_{B\,\min} - I_{L3} - I_{V\,\max} \tag{5}$$

where I_B is the nominal bias current

I_p is the nominal peak current of the tunnel diode

I_S is the safety current when switching is not desired

I_L is the total leakage current from clamp and coupling diodes

I_O is the minimum overdrive current when switching is desired

I_V is the diode current at the desired output voltage

i_{in} is the input current

$i_{\text{out}\,\max}$ is the output current at the desired output voltage.

Equation (3) determines the maximum allowable bias current for the stage when safety and leakage currents are considered. The magnitude of the leakage current varies in different stages depending on the number and use of coupling diodes. In addition, the output-stage leakage factor must include feed-through signals from other gates to which the circuit is connected. This interaction is minimized by the use of rectifier diodes as coupling elements; however, the capacitance of the rectifiers permits some transient current feed-through.

Equation (4) determines the minimum requirement for input current. The input current is the difference between the maximum peak current and minimum bias current plus overdrive current and leakage current. The output current, Equation (5), is the difference between the minimum bias current and the maximum tunnel-diode current at the critical output voltage of the tunnel diode. The critical output voltage of the tunnel diode is found by superimposing a worst-case composite load line on the tunnel-diode characteristic. In the case of an output stage driving three loads, the worst-case load consists of two

maximum-current loads and one minimum-current load. For coupling between cascaded stages in a gate, the requirements are less stringent and the critical voltage is selected to be near the valley voltage of the tunnel diode.

The factors involved in determining the gain of a stage can be normalized with respect to the peak current of the tunnel diode. As an example of the design of the OR gate of Figure 6, the first stage will be considered. The following list specifies the component variations required in the design:

$I_p = 25.0$ ma $\pm$ 2% (initial tolerance plus effect of aging);

the tolerance of I_B is $\pm$ 4% ($\pm$ 2% due to bias-voltage variation, and $\pm$ 2% due to bias-resistor variation);

$I_S = 5\%\ I_p$

$I_O = 7\%\ I_p$

$I_{V\,\max} = \dfrac{I_p}{9}$

$I_{L1} = 0.9$ ma

$I_{L2} = 0.7$ ma

$I_{L3} = 0.2$ ma

From Equation (3), $I_{B\,\max} = 0.894\ I_p = 22.3$ milliamperes and (assuming a nominal bias voltage of +6 volts) the bias resistor is 266 ohms. From Equation (4), the minimum input required is $0.296\ I_p = 7.4$ milliamperes. From Equation (5), the output current available to drive the second stage of the gate is $0.704\ I_p = 17.6$ milliamperes.

Since the minimum input current required has been determined, the maximum input current for loading the driving stage can be calculated graphically once the tolerances of the coupling network are defined.

In a similar manner the input and output requirements of the second stage can be calculated and the over-all gain of the gate determined.

AND Gate

In the AND gate, the bias current of the first stage is reduced so that both inputs must be supplying current in order for the stage to switch. Also, the presence of only one current input must not cause the stage to switch. These two requirements place a limit on the maxi-

mum and minimum allowable individual input currents. These requirements are

$$i_{L\,\max} + I_{B\,\max} \leq I_{P\,\min} - I_S \tag{6}$$

$$i_{P\,\max} + I_{B\,\max} \leq I_{P\,\min} - I_S \tag{7}$$

$$i_{P\,\min} + i_{L\,\min} + I_{B\,\min} \geq I_{P\,\max} + I_O \tag{8}$$

where i_L is the level input current and i_P is the pulse input current.

The maximum and minimum level current and the pulse input current were determined by consideration of worst-case driving voltages and coupling impedances. The determination of these voltages is shown in a later section. For example, the maximum pulse input current occurs when the AND gate is driven from an OR gate with a single load, through a minimum impedance coupling network. The minimum pulse input occurs from a maximally loaded OR gate where the coupling impedance is maximum. The extreme pulse and level inputs were found to be

$$7.7 \leq i_P \leq 12.1 \text{ milliamperes}$$

$$8.1 \leq i_L \leq 12.1 \text{ milliamperes.}$$

Using these values and the component variations used in the design of the OR gate, Equations (6), (7), and (8) were solved to determine the peak current of the tunnel diode and the bias required. These were found to be 20 and 5.77 milliamperes, respectively. It was necessary to add two amplifier stages to provide sufficient gain for a fan-out of two. The resulting circuit is shown in Figure 9. It can be seen that an inductance was added in series with the level input. The purpose of this inductor is to keep the input current flowing for a short time after the first stage of the gate has fired; thus additional current is available during this time to drive the second stage of the gate. The penalty paid is additional time required for the level input to build up.

Bistable Circuit

As indicated in Figure 1, a tunnel diode with a suitable load line can be operated in a bistable mode to obtain storage.

The bistable circuit may be divided into a number of blocks performing various functions, as shown by Figure 11. Only one block on this diagram is bistable, the others are monostable and their only func-

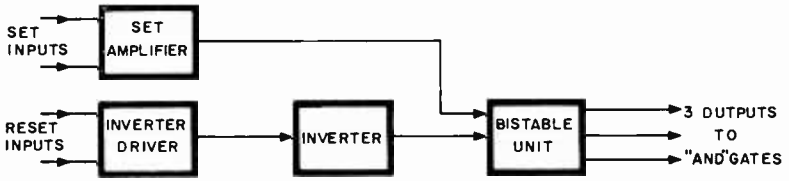


Fig. 11—Block diagram of bistable circuit.

tion is to switch the bistable unit to its "high" or "low" state via the available set and reset inputs. The set and reset inputs are positive pulses. Any one of the set inputs is amplified by the set amplifier which switches the bistable unit to its high-voltage state. This represents a logic "1". Any one of the reset inputs switches the inverter driver which activates the inverter to produce a negative pulse. This negative pulse switches the bistable unit to its low-voltage state, representing a logic "0".

In order to illustrate the operating principle of the bistable circuit, a simplified circuit diagram will be considered. This circuit is shown in Figure 12. Here, some of the stray reactive elements have been omitted. The circuit is divided into parts corresponding to those on the block diagram of Figure 11.

The set amplifier consists of a tunnel-diode monostable circuit described earlier. This circuit is designed so that an 8-milliampere set input is required to initiate switching. The output current of the set amplifier flows through TR_6 into TD_6 of the bistable unit.

The idealized switching characteristic of the bistable unit is shown

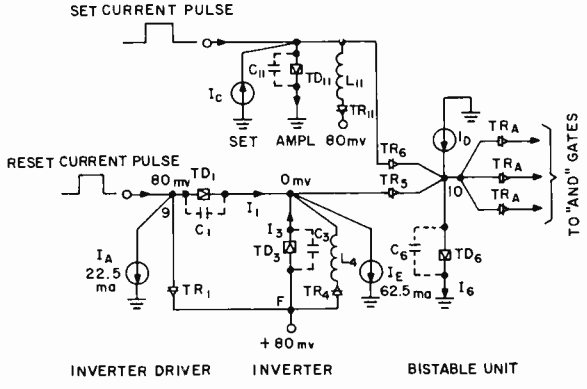


Fig. 12—Simplified diagram of bistable circuit.

in Figure 13. It consists of a tunnel diode *I–V* characteristic intersected by a load line at two stable points, 0 and 1c. This load line is formed by the combined input characteristics of the three AND gates which TD_6 is driving. Assuming that TD_6 is in the "0" state, the current supplied by the set amplifier causes TD_6 to switch over the peak along the indicated trajectory and stop at point 1c. The bistable circuit is, thus, set to its high state.

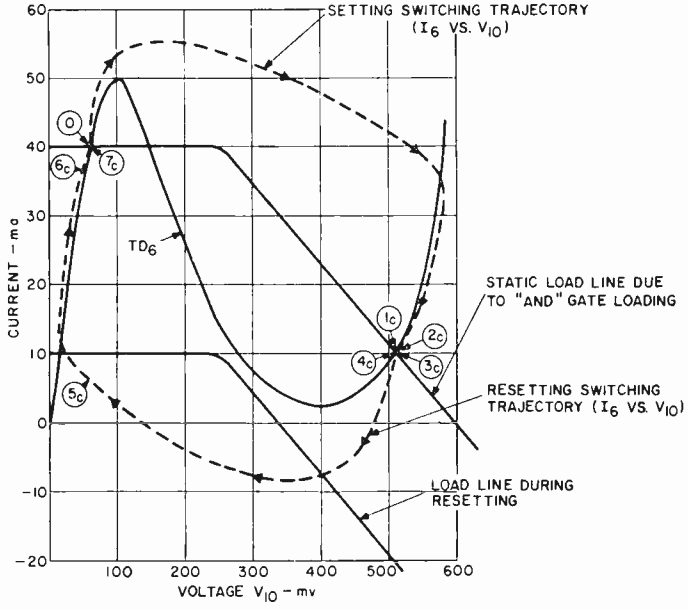


Fig. 13—Idealized switching characteristic of bistable unit.

The operation of the reset circuitry, which consists of the inverter driver and inverter is more complex and is, therefore, described in greater detail. The reset action may be divided into several steps listed in sequence of occurrence. After the application of a reset pulse (Figure 12),

1. the inverter driver switches
2. the inverter switches
3. the bistable unit switches to its low state
4. all circuits settle down to their original states (except the bistable unit which remains in the low state). This is referred to as the recovery period.

The idealized switching characteristic of the inverter driver and in-

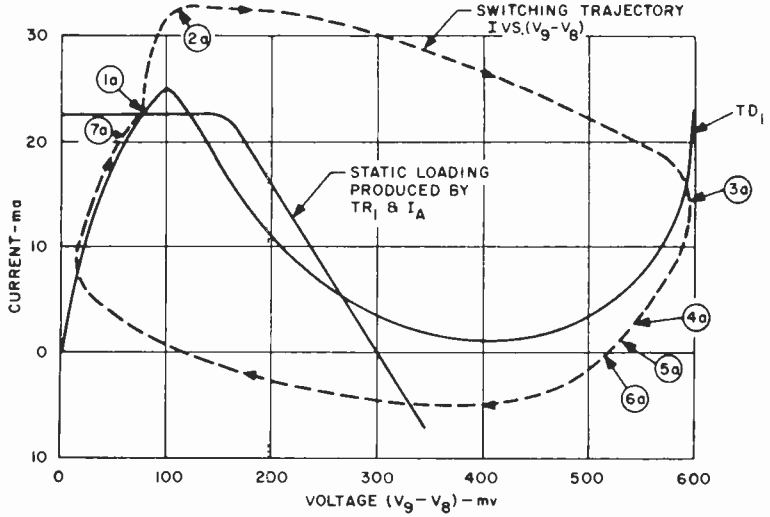


Fig. 14—Idealized switching characteristic of inverter driver.

verter are shown in Figures 14 and 15, respectively. Note that the characteristics and the trajectories have been idealized wherever necessary in order to simplify the analysis. Thus, at the steady state in Figure 12, the currents in TR_1, TR_5, TR_6, and TR_{11} are zero. Initially then, $I_1 = I_A = 22.5$ milliamperes, $I_3 = I_E - I_1 = 62.5 - 22.5 = 40$ milliamperes.

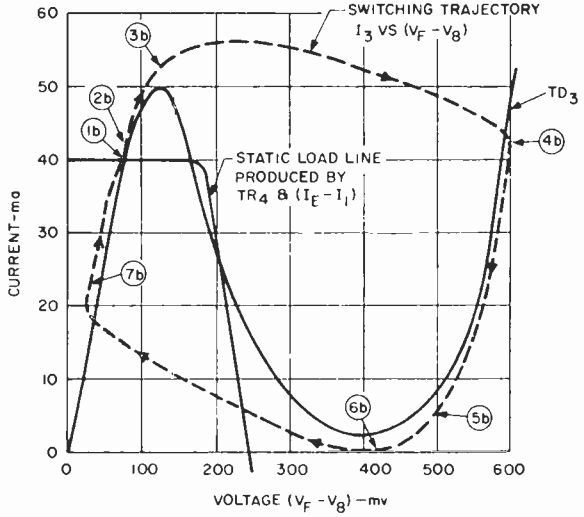


Fig. 15—Idealized switching characteristic of inverter.

When a 10-milliampere positive current pulse is forced into node 9, the current I_1 increases from 22.5 to 32.5 milliamperes. This current exceeds the peak current of TD_1, causing it to switch from point 1a to point 3a along the indicated trajectory of Figure 14. Up to this time, the voltage across TD_3 has undergone very little change. This is indicated by point 3b of Figure 15. The corresponding points reached by the bistable unit are shown in Figure 13 by the numbers with the "c" subscripts. The input has thus caused the voltage at node 9 of Figure 12 to increase from 80 to about 500 millivolts. This causes TR_1 to conduct. Since the current into node 9 is constant, conduction of TR_1 causes current I_1 in TD_1 to decrease.

As soon as current I_1 starts decreasing, I_3 starts increasing since $I_3 = I_E - I_1$. (Note TR_4 and TR_5 are not conducting at this time.) I_3 eventually exceeds the peak of TD_3 (point 3b Figure 15) causing TD_3 to switch along the indicated trajectory to point 4b. The voltage across TD_3 ($V_8 - V_F$) is now about —550 millivolts and the votlage, V_8, at node 8 is —470 millivolts. When TD_6 is high, the voltage at node 10 is approximately 500 millivolts. The voltage across TR_5 ($V_8 - V_{10}$) now becomes —470 —500 = —970 millivolts, which is sufficient to cause reverse conduction in TR_5. The idealized characteristic of TR_5 is shown in Figure 5.

Most of the current from the current source, I_D, flows in TR_5, causing the static loadline of Figure 13 to move down. When the loadline moves down such that the intersection at 1c becomes unstable or disappears, TD_6 becomes temporarily unstable and starts switching towards point 5c along the indicated trajectory. (It should be remembered that these are simplified trajectories merely showing the general path of switching. Any deviations not considered essential in describing the basic mode of operation have been omitted. Accurate trajectories are presented later.)

The resetting of the bistable unit is completed when it reaches the "0" state. As indicated by their static load lines, the only stable points for the inverter driver and inverter are 1a and 1b, respectively (Figures 14 and 15). Consequently, the inverter driver and inverter continue switching along their indicated trajectories until they reach these stable points. This completes the reset action.

Transient Analysis Using a Digital Computer

The circuit of Figure 12 was simplified for the purpose of explaining the basic mode of operation. In the practical case, however, actual characteristics and stray circuit elements must be considered. A schematic diagram representing the reset action of the actual circuit to a

fair degree of accuracy is shown in Figure 16. Here each tunneling device is shown to be shunted by its junction capacitance and in series with a stray inductance. Resistors R_6, R_8, and R_9 are the shunting resistances of the current supply sources. R_0 represents the output resistance of the preceding stage. The combination of TR_7, C_7, and L_7 represents the static and dynamic load due to the three AND gates.

L_4 and L_7 are externally added inductances. Their main function is to prevent a rapid change of current in TR_4 and TR_7. Thus, when TD_3 switches, most of the current flows into TD_6, providing efficient resetting.

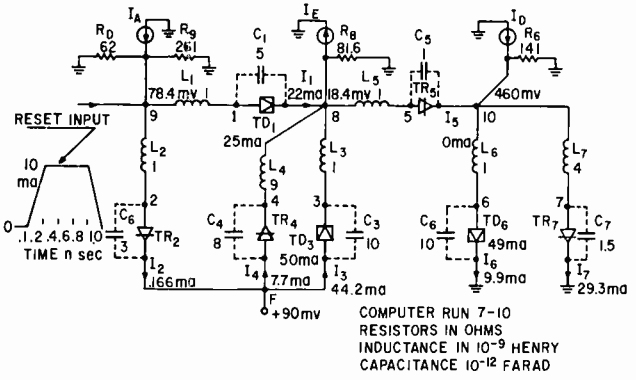


Fig. 16—Equivalent circuit used for analyzing the reset action of the bistable circuit on the RCA 501 computer.

TR_2, TR_4, and TD_3 are biased at +90 millivolts (a) to make the voltage at node 9 approximately the same as the voltage at the input of the other logic stages and (b) to prevent excessive conduction of TR_2 and TR_5 during steady-state conditions. The steady-state currents and voltages indicated on the diagram represent nominal operating conditions.

The differential equations governing the reset behavior of this circuit were solved with the RCA 501 digital computer. These results are presented in Figures 17 thru 20. Figures 17 thru 19 show the device characteristics and switching trajectories of the various stages. The nonlinear characteristics are represented by straight line segments for use in the digital computer. The trajectories are provided with time markers which indicate the speed with which switching progresses. The general trends of the trajectories are the same as those shown for the idealized cases. There are, however, some differences evident in the

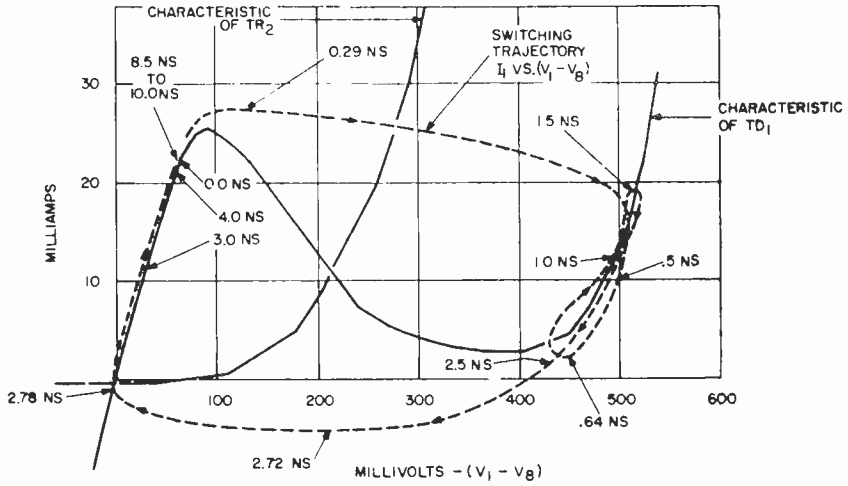


Fig. 17—Inverter driver switching characteristic calculated with the RCA 501.

loops of Figures 17 and 18. The loops are the result of a stage switching which causes a sudden voltage or current change. For example, the trajectory loop in Figure 18 in the time interval between 2.7 and

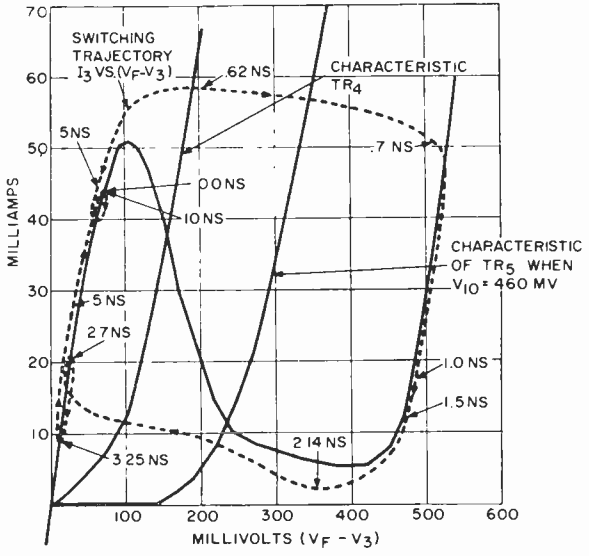


Fig. 18—Inverter switching characteristic calculated with RCA 501.

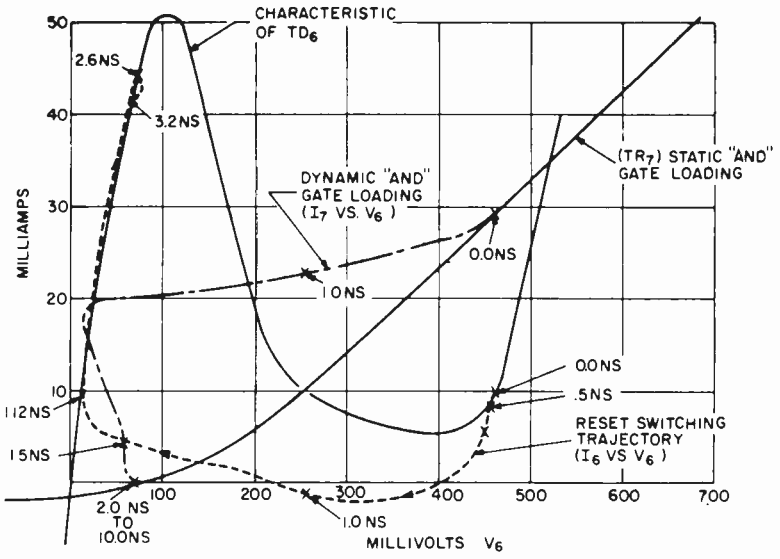


Fig. 19—Bistable reset switching characteristic calculated with RCA 501.

3.25 ns is caused by the increase in current in the inverter driver as shown by the time interval 2.78 to 3.0 ns of Figure 17.

Some of the significant voltage waveforms calculated with the computer are shown in Figure 20. The double peak in voltage V_9 is due to the fact that V_9 is the sum of the voltages across TD_1 and TD_3. Thus, during the time when the inverter output is negative, V_9 appears low even though TD_1 is in the high state. This may easily be seen by comparing the timing of V_9 and V_8 in Figure 20.

The computer results were compared with laboratory results and showed good agreement.

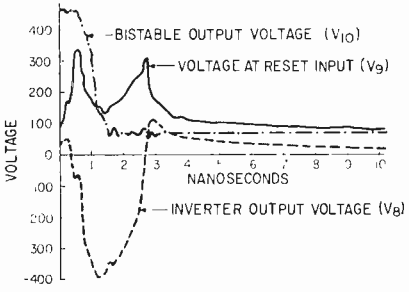


Fig. 20—Reset-voltage waveforms of bistable circuit.

D-C Tolerance Analysis

The choice of circuit components and voltages was based on consideration of the following four worst-case conditions:

1. setting of bistable unit,
2. resetting of bistable unit,
3. switching of inverter,
4. switching of inverter driver.

The component tolerances considered are given in Table I.

Table I

Parameter	Symbol	% Variation
Tunnel diode peak current	I_p	±2
Tunnel diode peak voltage	V_p	±5
Tunnel rectifier forward current		±5
6-volt supply voltage	V_a, V_c, V_d, V_e	±2
90-mv supply voltage	V_f, V_g	±10
All resistors		±2

Bistable Unit

As pointed out previously, the two stable states of TD_6 (Figure 12) are determined by AND gate loading. The current required for setting or resetting TD_6 is sharply affected by the impedance presented to it by the AND gates. Consequently, in making a tolerance analysis of the bistable unit, the variations in the AND gate input characteristic must be considered. This is illustrated in Figure 21, which shows the characteristic of TD_6 with its variations. Superimposed on this characteristic are some of the load lines the AND gates might present. The load line designated by A represents a worst-case condition where one AND gate receives a minimum input current level, which may not be sufficient to permit firing. This condition occurs when the input impedance to this AND gate is on the high extreme, and the input impedances to the other two AND gates are on the low extremes. This worst case is defined more accurately with the help of Figure 22. For example, AND gate #1 receives minimum current when R_{t1} and V_1 are maximum and R_{t2}, R_{t3}, V_2, and V_3 are minimum, where R_{t1}, R_{t2}, and R_{t3} represent the total resistance in branches 1, 2, and 3, respectively,

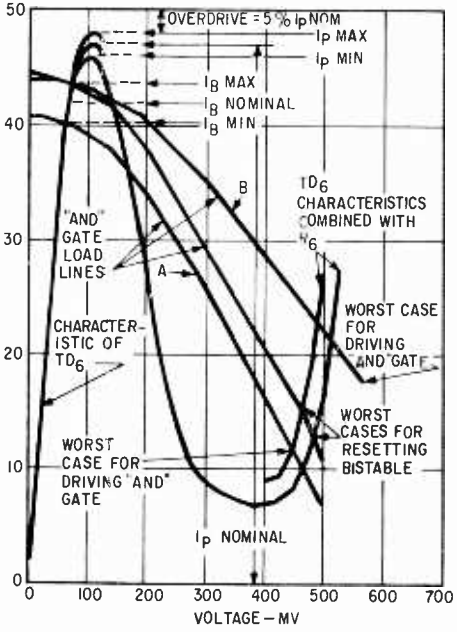


Fig. 21—Tolerance analysis of bistable diode.

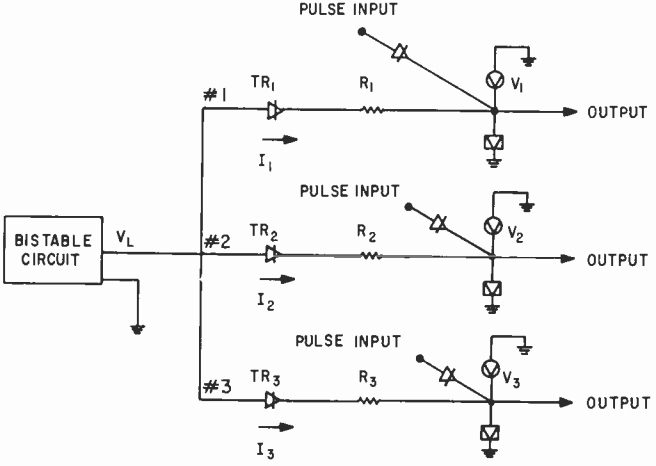


Fig. 22—Three AND gates driven by one bistable.

and where V_1, V_2, and V_3 are the diode voltages which may vary between 70 and 105 millivolts.

Referring to Figure 21, load line A intersects the characteristic of TD_6 at 450 millivolts. The bistable circuit is designed to make 450 millivolts the minimum high-state voltage under this worst-case condition. The choice of 450 millivolts is the result of a compromise between ease of resetting the bistable and minimum variation in bistable output voltage.

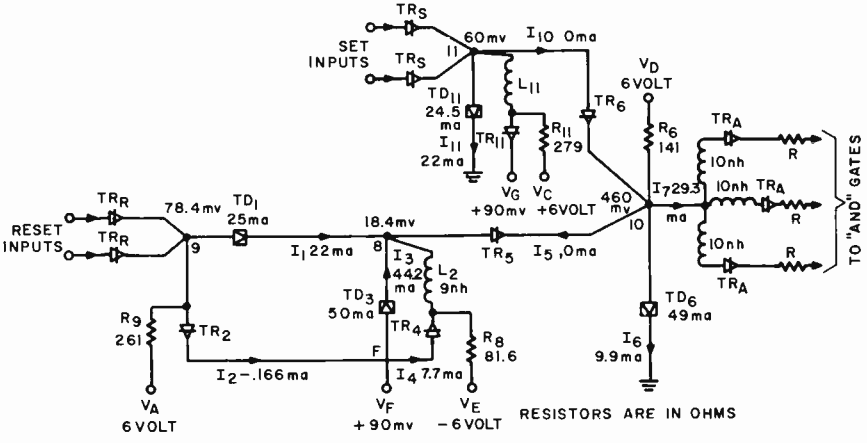


Fig. 23—Schematic diagram of bistable circuit with nominal values.

The point at which curve A intersects TD_6 in the low-voltage region gives $I_{B\,\min} = 40.2$ milliamperes (this is the minimum current that may flow in TD_6 when it is in the low-voltage state). If I_B is assumed to vary ± 4 per cent due to variations in supply voltage, V_D, and source resistance, R_G, of Figure 23, then

$$I_{B\,\min} = I_B - 0.04I_B \tag{9}$$

or

$$I_B = 42.0 \text{ milliamperes}, \tag{10}$$

and

$$R_G = \frac{V_D - V_{10}}{I_B} = \frac{6000 - 70}{42} \text{ ohms}, \tag{11}$$

where V_{10} is the voltage at node 10 when TD_6 is in the low-voltage state (Figure 23).

The choice of TD_6 is based on two considerations: (a) the peak

current of TD_6 should be as small as possible to allow the use of a smaller set input and (b) the peak current must be large enough to prevent TD_6 from switching without an input. In order to satisfy (b), the following relationship must be satisfied:

$$I_{p\ \min} \geqq I_{B\ \max} + 0.5\, I_p. \tag{12}$$

The last term in Equation (12) is a safety margin.

The variation in I_p is assumed to be $\pm$ 2 per cent; consequently, Equation (12) becomes

$$0.98\, I_p = 1.04\, I_B + 0.05\, I_p \tag{13}$$

or $I_p = 47$ milliamperes. The minimum set input is made large enough to provide an overdrive of $0.05\, I_p$ under worst-case conditions. This may be expressed as

$$\begin{aligned} I_{\text{set min}} &\geqq I_{p\ \max} + 0.05\, I_p - I_{B\ \min} \\ &= 1.02\, I_p + 0.05\, I_p - 0.96\, I_B \\ I_{\text{set min}} &= 10 \text{ milliamperes.} \end{aligned} \tag{14}$$

These values are indicated in Figure 21. The minimum output available from the set amplifier was calculated to be 16 milliamperes. Since the set requirements were less than that available from the set amplifier, the peak current of TD_6 was raised to 49 milliamperes. This permits the use of a more readily available tunnel diode for the bistable circuit.

TD_6 is most difficult to reset when maximum current is flowing through it. The two points which represent worst cases for resetting are indicated in Figure 21. The diode currents corresponding to these points are lower by about 3.5 millamperes because a combined characteristic of TD_6 and R_6 was used in the construction of Figure 21.

Having determined the nominal values of the circuit elements in the bistable unit, another worst-case condition must be evaluated. This condition is designated by load line B which results when the input level to a particular AND gate is maximum and presents the danger of firing with the level only. Again, this worst-case condition is defined more accurately with the help of Figure 22. For example, AND gate #1 receives maximum current when R_{t1} and V_1 are minimum, and R_{t2} and V_2 are maximum, and Gate #3 fires, causing its level current input

to go to zero.* The diode voltages may vary from 70 to 105 millivolts. When I_B of Figure 21 is maximum, the intersection of curve B with TD_6 gives the maximum high-state bistable voltage under worst-case conditions. This voltage is 510 millivolts.

Thus under worst-case conditions, the bistable output voltage may vary from 450 to 510 millivolts. This voltage variation must be taken into account in the design of the AND gate.

Inverter and Inverter Driver

Experiments and computations have indicated that in order to reset the bistable circuit, a 50-milliampere peak current tunnel diode is required for the inverter. This dictated the choice of a 25-milliampere peak current tunnel diode for the inverter driver. Tolerance analyses, similar to those shown for TD_6 in Figure 21, were carried out for the inverter and inverter driver.

Experimental Circuit

The complete bistable circuit is shown in Figure 23. The choice of circuit components and design was based on the d-c tolerance calculations and the digital computer results.

The set and reset inputs are applied through tunnel rectifiers to isolate the circuit from preceding logic stages.

The current sources are connected directly to the rectifiers TR_4 and TR_{11}, instead of to nodes 8 and 11 as shown in Figure 12. The difference between the two is very slight. However, the connection in Figure 23 is more desirable since it prevents R_8 and R_{11} from loading the tunnel diodes.

The inductances in series with the output rectifiers connected to TD_6 are used to obtain more efficient reset action. During the initial period of resetting, the series inductances prevent the currents in the AND gates from changing, thus permitting more of the reset current to go through TD_6. The addition of inductance, however, increases the time required for establishing a new level in the AND gate, reducing somewhat the over-all repetition rate.

The bistable circuit of Figure 23 exhibits the following performance:

2 set inputs of 7.4 milliamperes (minimum) each,

2 reset inputs of 8 milliamperes (minimum) each,

3 outputs of 10 milliamperes (nominal) each.

* It is assumed that no more than one AND gate is fired at one time.

*Gate Delay**

Set: 3 ns (nominal value)

Reset: 3 ns (nominal value)

*Delay between***

Set — Reset: 3 ns (nominal value)

Reset — Set: 3 ns (nominal value)

Reset — Set — Reset: 8 ns (nominal value)

Set — Reset — Set: 6 ns (nominal value)

An unassembled and an assembled bistable wafer are shown in Figure 24a. Care was taken in the wafer layout to keep stray reactances

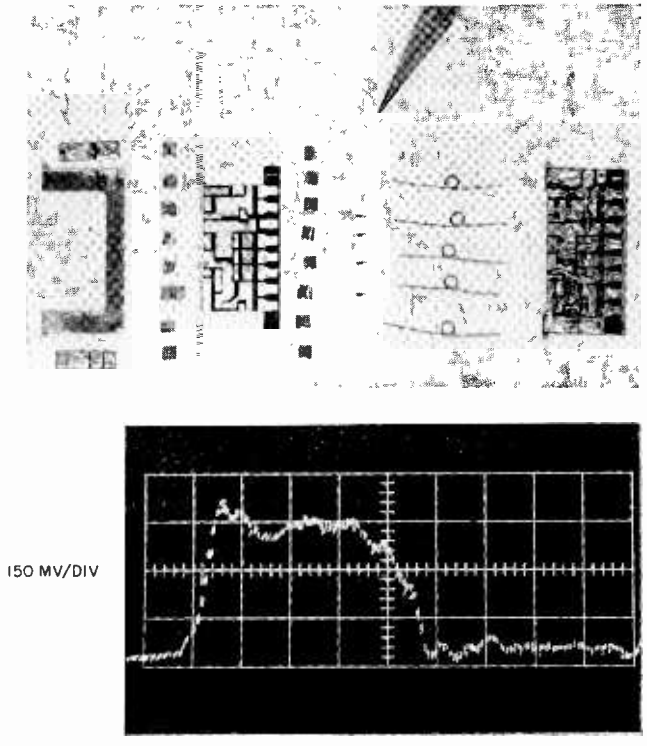


Fig. 24—Bistable wafer and output waveform of bistable circuit.

* Delay from the time the input has risen to the time the new level in the AND gate is established (exclusive of transmission-line delay).

** Delay from the time the input has risen to the time the new input may be applied.

to a minimum. A photograph of the output waveform resulting from a set input followed by a reset input 5 ns later is shown in Figure 24b. This waveform was taken at a point corresponding to node 10 of Figure 23.

Logic-Circuit Interconnection

As the rise times of logical circuits are shortened into the nanosecond region, increased attention must be paid to the methods by which these circuits are interconnected. Some of the problems which must be considered are crosstalk, the effect of common ground paths, and line termination. When the propagation times are long compared to the rise time of the interconnection system, the distributive effect of the interconnection should not be neglected. Also, as the speed of circuits is increased without an equivalent reduction in their volume, the interconnection delay becomes a greater and greater percentage of the total. At some point, most interconnections will have to be treated as transmission lines. For this reason, it is important to design the interconnection system with greatest efficiency, i.e., using low-dielectric-constant materials for maximum propagation velocity and terminating transmission lines so that efficient use is made of the transmitted power.

In the case of the circuits described here, typical logic delays are of the order of one nanosecond with rise times of 0.2 nanosecond. Fabrication requires typical interconnection lengths of 2 to 6 inches.

It was decided to use a miniature coaxial cable for all logic interconnections as shown in Figure 25. One of the requirements of the connection system was that its design be compatible with the circuits with respect to maintaining directionality and required level of input, output, and feedback signals.

Figure 26 shows the three configurations used to interconnect logic circuits. The first (a) is used to couple pulses from monostable gates to the OR gate. Rectifier diodes are used at both ends of the line to maintain a unilateral path. Also, due to the unloading effect of the OR gate at the load end of the cable when it switches, it is desirable to terminate the line at both ends to minimize ringing.

The second interconnection (b) is used to couple pulses to the AND gate. Once again a rectifier is used at the input to the line to isolate the load during initial switching of the driving gate. At the output end of the line, a resistor is used to provide better termination for the line. This is possible for the AND gate since all inputs to the gate must be high for the gate to switch. Under these conditions directionality is not a problem. The use of a resistor to terminate the line also

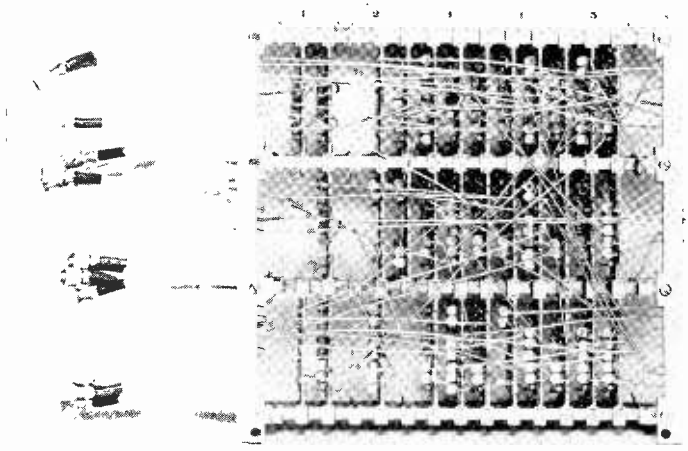


Fig. 25—Photograph of logic circuit interconnections.

has the advantage of providing a better-controlled input pulse. This is important because of the threshold property of the AND gate which requires tight control of maximum and minimum input signals. This is not the case in the OR gate, where only a minimum input current limit is required.

The level input to the AND gate (c) is coupled in a manner similar to the pulse except that inductors are added at each end of the line to increase the efficiency of the circuits at both ends of the line as described earlier.

Characteristic impedance of the line was selected to provide the required current to the loads. An impedance of 25.7 ohms was chosen on the basis of matching the requirements for the inputs to the AND gate, since this was the most critical circuit.

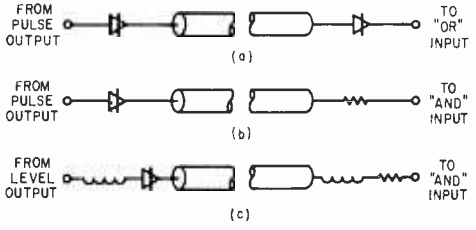


Fig. 26—Interconnection configurations.

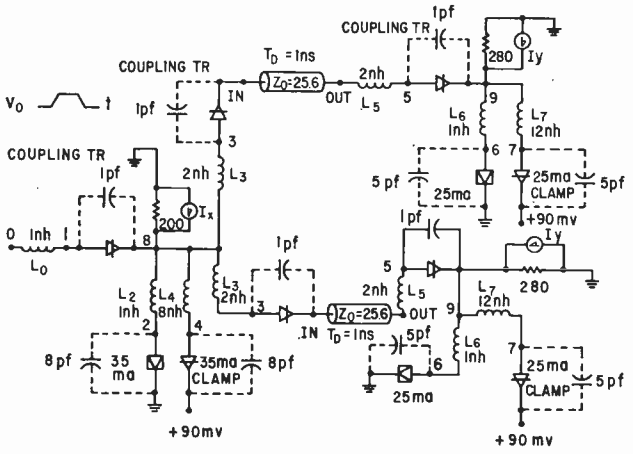


Fig. 27(a)—Model for studying transmission-line effects. Waveforms taken at various points on the model are shown in Figures 27(b) through 27(g).

The use of nonlinear line terminations makes it difficult to calculate the response of the interconnection to an input signal. To solve this problem, the circuit was simulated on the RCA 501 computer. In the simulation, the line was assumed to be ideal, consisting of only a delay. Figure 27a shows the general circuit model which was used in the simulation. A total of twelve runs were made with the computer to investigate the properties of the three coupling networks. Figures 27b to 27g show waveforms obtained in various points of the circuit of

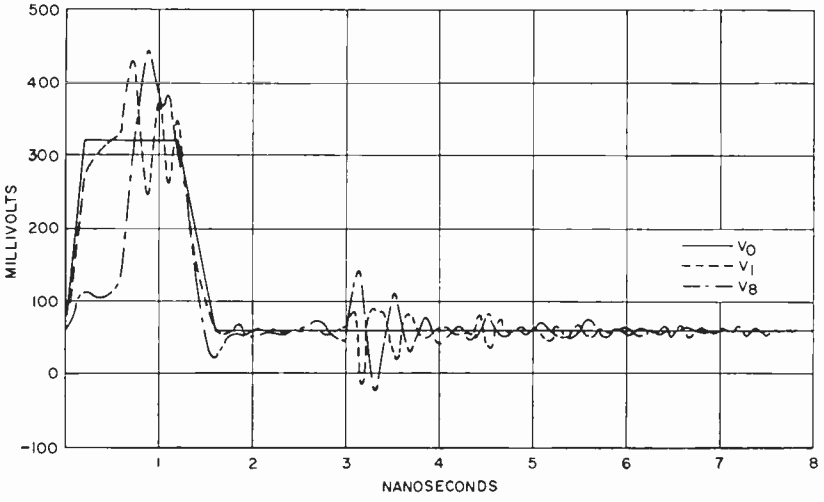


Fig. 27(b)

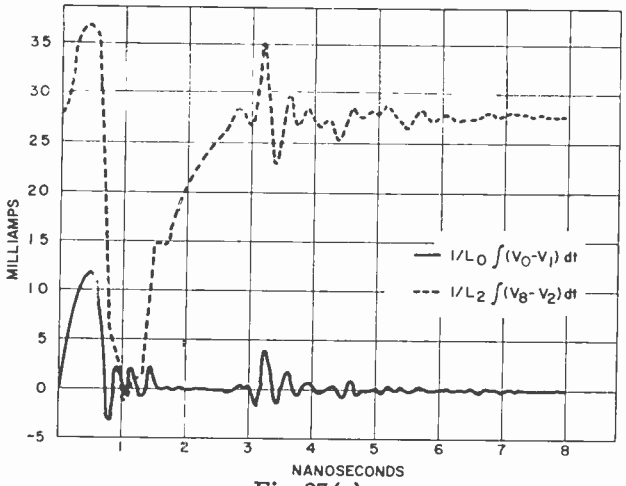


Fig. 27(c)

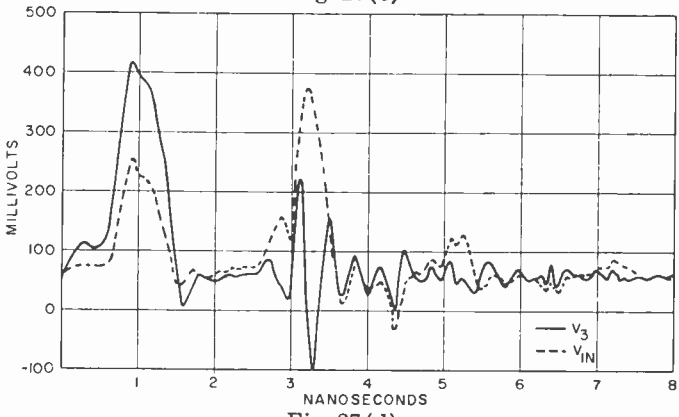


Fig. 27(d)

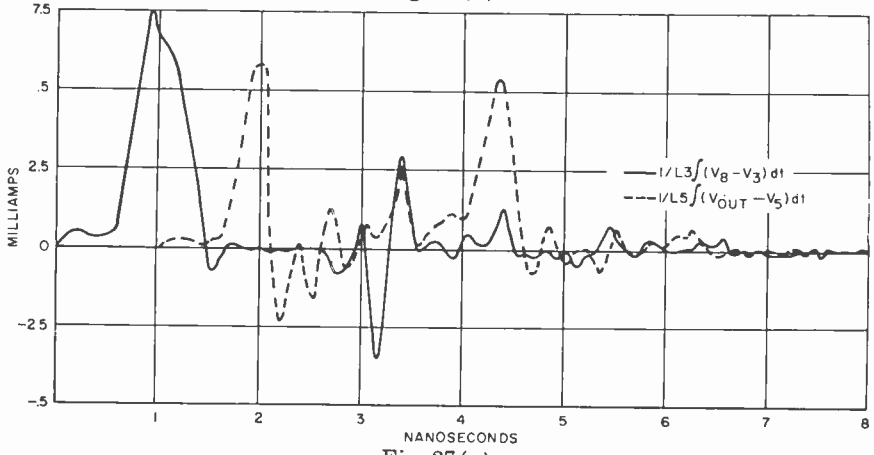


Fig. 27(e)

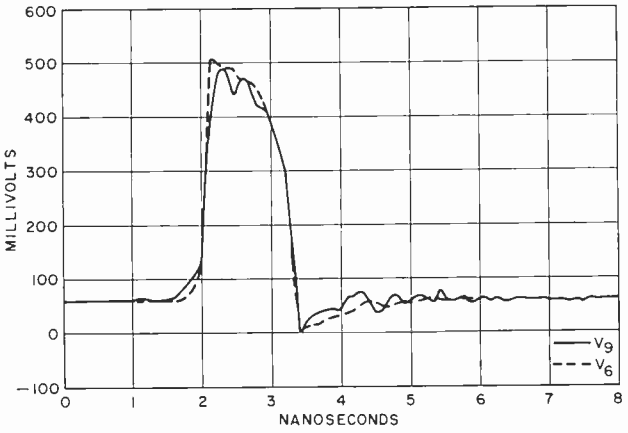


Fig. 27(f)

Figure 27a for a typical run which simulates a 35-milliampere tunnel-diode stage driving two OR gates. The results of the simulation indicated that the OR interconnection was approximately 90 per cent efficient in the transfer of current to the load in the nominal case, while the AND interconnections were better than 95 per cent efficient. The efficiency of a signal transfer effects the magnitude of input signal to a gate. At low efficiencies, less overdrive is available to switch the gate. The results are summarized in Table II. Included in this table

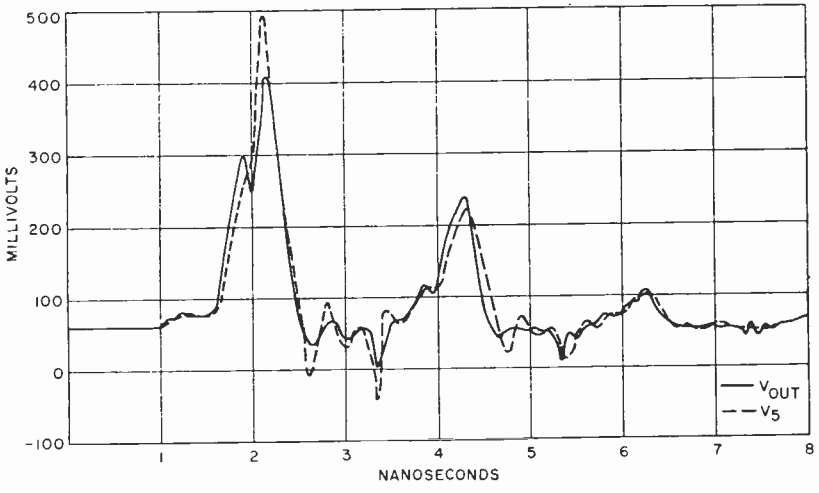


Fig. 27(g)

is the effect of variation of transmission-line characteristic impedance. The transfer of current was found to be quite insensitive to this variation in the line.

For pulse-carrying lines a maximum length of 6 inches is imposed due to the recovery time of the circuits involved. This limit is required due to the unloading action of the circuit at the output end of the line.

Unloading occurs when a monostable circuit at the terminal end of a line switches to the high-voltage state for a period of 1.5 nanoseconds. Since this high-voltage state is equal to the driving voltage,

Table II

Connection Type	Current Transmission Efficiency I_{out}/I_{in}	% change in Z_0	% change in I_{out}
OR	.897	0	—
OR	.948	+10	+1
OR	.840	—10	—1.7
AND (pulse)	.998	+10	+.75
AND	.897	—10	—.75
AND (level)	1.05	+10	—.45
AND	.969	—10	+.45

the load current drops. This unloading transient propagates back up the line to the driving circuit which has switched back to the low state by this time. The rectifier diode blocks most current feedback into the driving stage; however, a reflection occurs which travels back to the load circuit and reaches it during its recovery. If the tunnel diode is still sufficiently far from its peak, this signal can be absorbed without retriggering. For longer lines, special driving circuits have to be used.

Length of level lines are associated with the rate at which the AND gate at the end of the line is pulsed at its pulse input. Long lines are allowed if care is taken not to pulse the AND gate when a reflected unloading pulse returns to the gate.

Conclusion

Experiments and theoretical analysis have established the feasibility and practicality of all-tunnel-diode logic for digital computers. The major advantage of such a computer would be its extremely high speed

of operation. The three basic building blocks, OR gate, AND gate, and bistable circuit, have operated reliably under worst-case conditions with practical component and power-supply variations.

A group of circuits using a total of 300 gates mounted on frames, as shown in Figure 25, have been built to perform storage, shifting, counting, parity checking, and other operations. Preliminary results indicate reliable operation for these circuits over a period of several weeks with low maintenance required.

The major shortcoming of the circuits presented is the relatively low fan in and fan out of the gates. This limitation is not felt to be basic, however, and work is in progress to overcome it. Also, the operating speed is presently limited by the stray capacitance and inductance of the devices and circuit construction; improvements in these areas will increase the performance of tunnel-diode logic circuits. The results obtained and the foreseen improvements make the tunnel diode an increasingly important device for high-speed digital computers.

Acknowledgment

The work described here is the product of a joint effort by the Logic Circuit Group of the EDP Advanced Development section. Credit is due the following individuals for their part in this work: L. J. Bazin, for preliminary tolerance analysis; R. Cheuh, for tolerance analysis of the monostable circuits; E. C. Cornish, for tolerance analysis and laboratory development of the bistable circuit; and C. Pendred for simulation of the circuits on the RCA 501 computer. Some of the basic ideas which lead to the development of the resetting circuit were contributed by M. M. Kaufman.

The authors also wish to express their gratitude to their colleagues at Pennsauken, Somerville, and Princeton for their support and cooperation. Particular thanks are due R. K. Lockhart and J. N. Marshall for their guidance and support.

TUNNEL-DIODE BALANCED-PAIR SWITCHING ANALYSIS*

By

G. B. Herzog

RCA Laboratories
Princeton N. J.

Summary—Tunnel diodes, because of their extremely fast switching characteristic, have been extensively studied for ultra-high-speed computer use. Reported diode switching times of less than 0.1 nanosecond led to expectations that tunnel-diode logic circuits could operate at information rates of 1 kmc. The use of one tunnel diode as the biasing impedance for a second in a balanced circuit driven by opposite phases of an a-c clock appeared to give increased logic gain with reduced dissipation over that of single-ended bistable circuits with resistor constant-current bias sources. Experimental balanced-pair circuits were operated at 1.2 kmc verifying the fast switching characteristics of tunnel diodes. An analysis of the switching behavior of the balanced-pair circuit by computer solution of the circuit equations has led to the belief that this circuit can be operated reliably up to 250 mc by the proper choice of tunnel diodes.

Oscillatory instabilities due to the irreducible lead inductance set a logic-gain–switching-speed relationship which severely reduces the usefulness of the circuit above approximately 250 mc, although forced switching can be obtained at frequencies well beyond this limit. Unbalances further reduce the effective gain of the circuit so that the useful gain of the balanced pair at ultra-high speed is far less than previously predicted on the basis of static analysis. At modest speeds (100 mc and below), however, the balanced pair performs well, and a clear understanding of the switching action leads to a choice of slower tunnel diodes which actually perform better than faster high-current diodes.

Introduction

During the early history of the tunnel diode, the series combination of two diodes energized by opposite-polarity unipolar square waves (Figure 1) appeared to be one of the better tunnel-diode circuits suggested[1] for use in ultra-high-speed computers. By simultaneously applying positive bias to one diode and negative bias to the other, an instant is reached when both diodes are biased near their peak voltage. Intuitively, it appeared that a very small input current at the junction of the two diodes would be sufficient to insure that only one diode would switch across its negative-resistance

* Manuscript received April 17, 1962.

[1] First suggested at RCA Laboratories by A. W. Lo.

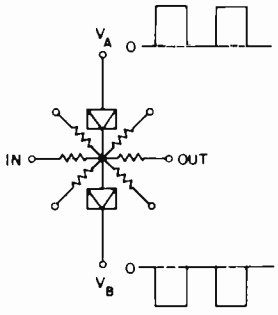


Fig. 1—Balanced-pair circuit.

region, absorbing most of the bias voltage. The junction of the two diodes would then assume the output potential of point A or point B in Figure 2 in response to the majority of the input signals, and this potential would be equal in magnitude to one of the unipolar bias voltages less the small voltage drop across the off diode. At low frequencies, the gain appeared to be limited only by the match of the diode peak current and the balance of the bias voltages. With reasonable tolerances on the interstage coupling resistors, large fan-outs were predicted for closely matched diodes,[2] assuming the load current to be 80 per cent of diode peak current. It was obvious that the diode shunt capacitance would play an important part in the operation of the circuit at high frequencies, but extrapolations from work with clock frequencies of from 1 to 30 mc suggested that improved diodes would allow operation at 1000 mc. Diodes with peak currents of 50 milliamperes and valley capacitance of 10 picofarads were soon avail-

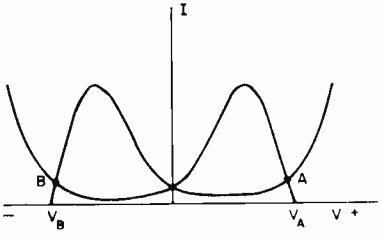


Fig. 2—Superimposed tunnel-diode characteristics.

[2] Chow, W. F., "Tunnel Diode Digital Circuits," *I.R.E. Transactions on Electronic Computers*, Vol. E.C.9, September, 1960, p. 295.

able for experimental use, and early in 1961 balanced circuits were operated at frequencies up to 1200 mc.[3]

Before high-speed digital systems could be constructed from balanced-pair gates of the type described, it was necessary to have a better understanding of the locking process, i.e., that part of the bias cycle during which one diode switches and the other stays in the low-voltage state. The fraction of the cycle during which the input must be of a given magnitude to insure proper locking determines the necessary circuit tolerances for a given fan power. Because of the impossibility of making accurate physical measurements of the circuit currents and voltages at high frequencies, and the difficulty of a pure analytical approach,[4] a computer solution of the circuit equations was used to give data on circuit operation. This data has led to a realization that the maximum operating frequency of this circuit is not limited by the diode speed but rather by the case inductance, which causes oscillatory instability. For 50-ma 10-pf diodes with a case inductance of 400 ph, the maximum practical operating frequency appears to be approximately 250 mc, although switching can be achieved at 1000 mc and above by sufficiently large input signals (pf = picofarad = 10^{-12} farad, ph = picohenry = 10^{-12} henry). The length of the locking period is determined by the circuit stability and may be an appreciable fraction of the bias period.

Equivalent Circuit

The balanced pair was represented by the equivalent circuit of Figure 3. Case plus lead inductance is represented by L, diode capacitance plus case capacitance by C, source resistance by r, load resistance by R_L, and the diode characteristic by $f(V)$ or $|-R|$. Although the junction capacitance is a function of diode voltage, comparison of circuit behavior with C held constant rather than variable indicated only trivial differences. Consequently, most of the computation was done with C assumed constant. The inclusion of the mutual inductance, M, between the two halves of the circuit represents a practical attempt

---

[3] Private communication—H. S. Miiller, RCA Laboratories.

[4] Because the author was at Cambridge University, England at the time he started this work, he was unaware that J. J. Gibson was simultaneously applying scattering techniques to solve this same problem. Similar results were obtained by both methods, and the waveforms obtained from the computer solution complement the analytical forms derived by Gibson; J. J. Gibson, "An Analysis of the Effects of Reactances on the Performance of the Tunnel-Diode Balanced-Pair Logic Circuit," (to be published).

to reduce the effect of lead inductance. The value of M was generally assumed to be zero, although a significant value of M was shown to be beneficial. The voltage sources, V_a and V_b, are basically sinusoidal varying potentials with a d-c pedestal:

$$V = E_{\text{d-c}} - A \sin \omega t. \tag{1}$$

The magnitude of the d-c component was chosen to assure resetting of the tunnel diodes on the negative swing and generally $E_{\text{d-c}} = A$, although values of $E_{\text{d-c}} < A$ were also used. To study the decay of the

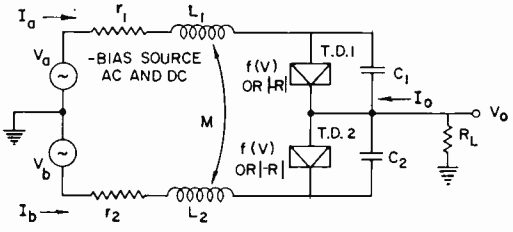


Fig. 3—Balanced-pair equivalent circuit.

switching transient, the bias voltage was often held at its maximum value after the sinusoidal buildup; that is,

$$V = E_{dc} - A \sin \omega t, \qquad \omega t \leqq \frac{3\pi}{2} \tag{2}$$

$$V = E_{dc} + A, \qquad \omega t \geqq \frac{3\pi}{2} \tag{3}$$

To simplify discussion, the frequency of the sine function will be referred to as the frequency of operation. Considerations in the choice of circuit parameters were that the load should represent at least 50 per cent of the diode peak current; the bias source resistance should be small enough so as not to make the output impedance high compared to the load resistance; and the diode peak current, capacitance, and lead inductances should be values readily achievable in production. These factors lead to a choice of a 50-ma, 10-pf, 10-ph diode with a bias source resistance of 1 to 3 ohms and a load resistance of 5 to 8 ohms.

To be more nearly correct, the load impedance should include series inductance L and shunt capacitance C. The purely resistive approximation was used on the grounds that the shunt capacitance would be small compared to the diode capacitance, and the series inductance in each coupling resistor would have negligible effect compared to the resistance in a properly designed system. This simplification made the switching action easier to analyze and reduced computing time. The qualitative effects of shunt capacitance and series inductance are easy to predict once the switching behavior of the basic circuit has been determined.

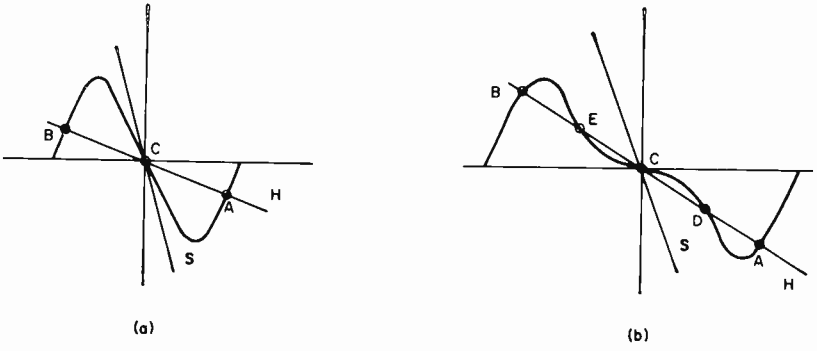


Fig. 4—Balanced-pair characteristic curves.

STATIC CHARACTERISTIC

To study the switching of the balanced-pair circuit, it is convenient to draw the static circuit characteristic as shown in Figure 4. This curve is obtained by plotting the difference of the two tunnel-diode currents against center-node voltage for a given bias voltage and source resistance, and represents the input characteristic at the junction of the two diodes. A load line on this characteristic intersects the possible output states. Although this characteristic holds for only one instantaneous value of the periodic bias, it is sufficient to consider some limiting cases of the characteristic.

The curve of Figure 4a represents the characteristic obtained when the total voltage across the circuit is approximately equal to the valley voltage of a single diode. The load line, S, represents a small resistance while H represents a larger load resistance. A useful output, indicated by points A or B, is obtained only with the larger load resistance, H.

Increased bias voltage causes reduced slope near the center of the characteristic and the possibility of five load-line intercepts, as shown in Figure 4b. Three intercepts, A, B, and C, represent possible stable outputs. The presence of intercept C indicates the possibility of zero output. This possibility will be discussed further after circuit stability is discussed.

Dynamic *V–I* Characteristics

Solution of the circuit equations as described in the Appendix provides data on the instantaneous circuit voltages and currents. A plot

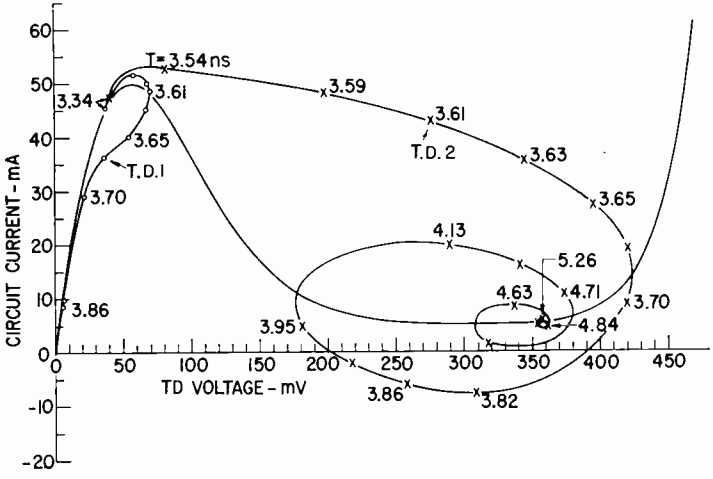


Fig. 5—*V–I* characteristic of switched unstable circuit.

of the circuit current against tunnel-diode voltage for each half of the circuit gives insight into the switching mechanism. A typical circuit *V–I* characteristic (Figure 5) for a 50-ma diode operated with a bias-voltage growth rate of 150 mc, as given by Equations (2) and (3), illustrates the case where the voltage on only one diode exceeds the peak value and that diode eventually stabilizes at a high voltage; the other diode stabilizes at a low voltage. The same type of curves for different circuit parameters show that although one diode may initially relax to the low-voltage state, it can eventually switch and get in step with the other diode in an oscillatory limit cycle (Figure 6). The behavior of a particular circuit depends on the amount of input signal, load resistance, and circuit stability.

Circuit Stability

The inductance and capacitance associated with the tunnel-diode negative resistance makes the balanced pair a potential oscillator circuit. The stability criterion that $L/rC > |-R|$ for oscillations provides a guide to the type of switching behavior to be expected from a given balanced-pair circuit. Two possible oscillatory modes exist. One mode

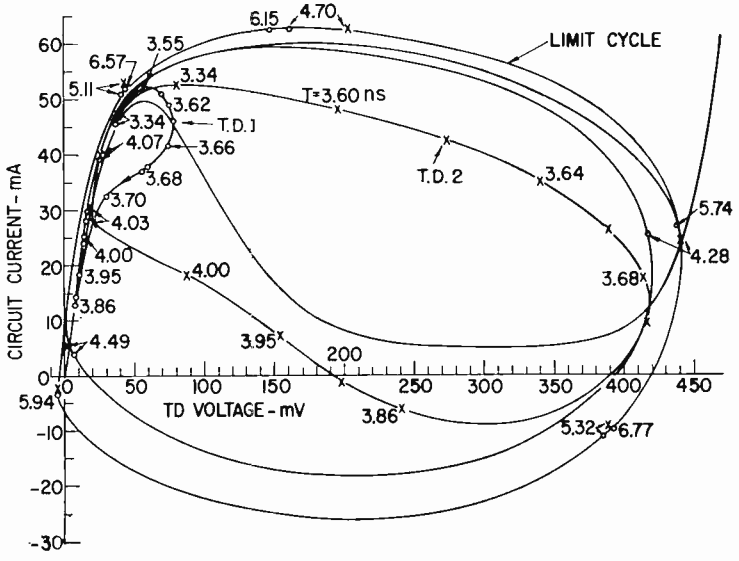


Fig. 6—*V–I* characteristic of oscillating unstable circuit.

involves an oscillatory current around the loop formed by the balanced-pair circuit; the second mode involves an oscillatory current into the load with the two halves of the balanced pair operating in parallel, though not necessarily at identical bias points on the tunnel-diode characteristic. In nearly all practical circuits having oscillatory tendencies, oscillations starting in the second mode degenerate into oscillations in the first mode. The stability of the balanced pair has therefore been defined in terms of the first type of possible oscillation. Because the circuit consists of a series combination of $2L$, $2r$, $2C$ and $2\ |-R|$, the stability of the circuit may be expressed in terms of a single L, r, C, and $|-R|$.

The following terms used to describe the balanced pair are arbitrary and should not be confused with similar terms used in nonlinear analysis:

Unconditionally Stable

If $L/rC < |-R|$ for the greatest negative conductance, i.e., smallest negative resistance encountered in traversing the tunnel-diode characteristic, the circuit is defined as unconditionally stable. With the assumption that the load is predominantly resistive, such a circuit must be stable in the second mode as well as the first.

Conditionally Stable

It is possible that at the maximum value of applied bias voltage, the intersection of the bias source resistance, r, with the tunnel diode characteristic defines a point where $L/rC < |-R|$, while for lower values of bias voltage $L/rC > |-R|$. Under these conditions, the energy gained during the switching process can exceed the energy loss, and continuous oscillations can be induced. A circuit of this type is generally referred to as a "hard" oscillator and requires a disturbance of appreciable amplitude to start the oscillations. A balanced circuit satisfying these conditions is defined as a conditionally stable circuit.

Unstable Circuit

If $L/rC > |-R|$ for each point on the tunnel-diode characteristic intercepted by the source resistance, the circuit is a "soft" oscillator and will be very prone to oscillation. Actually, the conditionally stable circuit and the unstable circuit act very similarly when switched in the standard fashion where the bias is increased from a small value to a large value. Greater differences are observed where the bias is decreased from an overly large value to the final large value for useful output. For this type of special operation, the unstable circuit will certainly oscillate while the conditionally stable circuit may not. Furthermore, circuits with such a small value of source resistance as to be classified as unstable tend to have oscillations initiated in the second mode which in turn excite oscillations in the first mode.

Switching Behavior

Unconditionally Stable Circuits

At low frequencies, with finite input signals and modest loads, the intuitive explanation of balanced-pair switching applies; i.e., the voltage on only one diode exceeds the peak value, and when this diode switches to the high-voltage state, the output assumes the voltage state of point A or B in Figure 4a. At high frequencies, where the capacitive current becomes comparable to the input-signal current, the

voltage on both diodes initially exceeds the peak value. This leaves the output voltage near point C in Figure 4a as the bias voltage continues to increase. If, for the largest bias voltage applied, the characteristic always has a shape similar to that of Figure 4a with three load intercepts, then the output must eventually reach one of the two stable output states A or B, providing the bias is maintained for a sufficiently long time. The greater the input signal, the faster the final output state is achieved. Figure 7 compares the output of a circuit utilizing

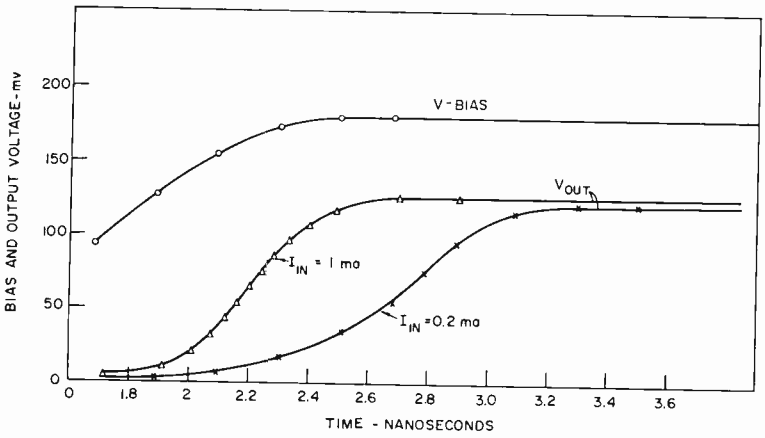


Fig. 7—Output waveform of stable circuit.

10-ma, 10-pf, 400-ph diodes at 300 mc for 0.2-ma and 1.0-ma input signals. Note that only with the 1.0-ma input signal is there appreciable output by the time the bias reaches its maximum value. Phase plane representation of the output, i.e., the first derivative of output voltage plotted against output voltage, provides a convenient means of comparing the switching behavior of different types of circuits. The phase plane representation corresponding to the output waves of Figure 7 is shown in Figure 8.

A stable circuit having a characteristic such as that shown in Figure 4b with 5 load-line intercepts may have a zero output, point C, for small input signals. If the rate of bias-voltage buildup is such that the output is still in a region near C when intercepts D and E appear, the desired output, points A or B, will never be achieved since both diodes will remain locked in their valley region. To insure an output, either the input must be quite large, or the maximum bias voltage must be kept small. In either case, the switching behavior of the

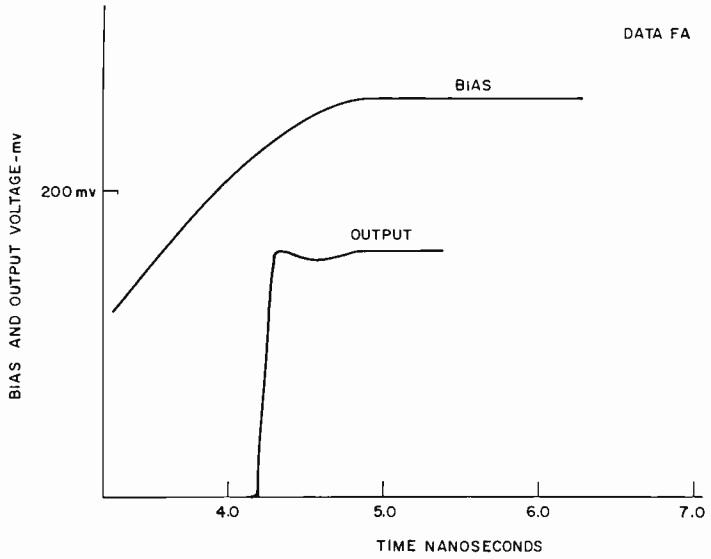


Fig. 9c—Output wave shape (circuit as in Figure 9a).

150 mc in this desired mode. The circuit has an inductance of 400 ph, a capacitance of 10 pf, a source resistance of 3 ohms, and a load resistance of 8 ohms, which represents 20 ma or a 40 per cent load for the 50-ma tunnel diodes used. The *V–I* characteristics of the same

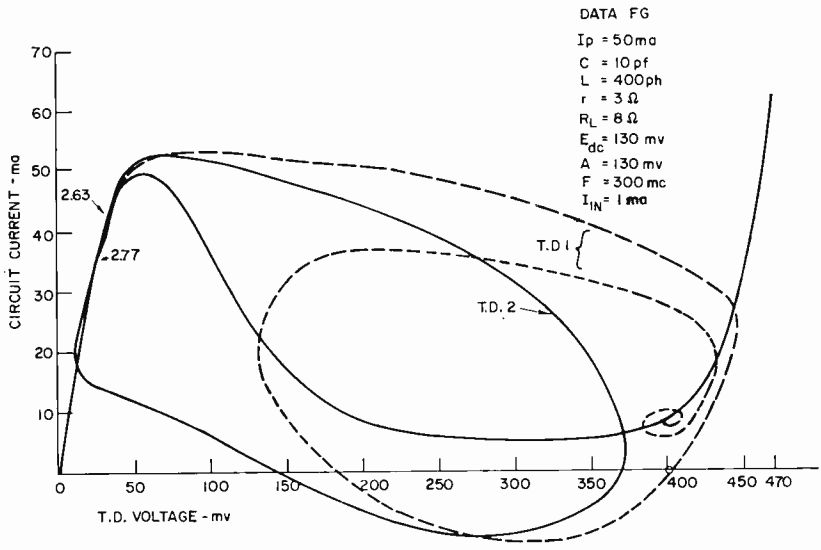


Fig. 10a—*V–I* characteristic of switched conditionally stable circuit operating at 450 mc.

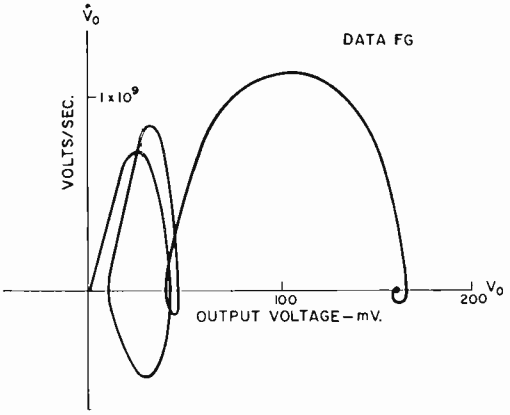


Fig. 10b—Phase plane of output voltage (circuit as in Figure 10a).

circuit driven by a 300-mc bias voltage are shown in Figure 10. The capacitive current is 2 ma at 300 mc compared to 0.7 ma at 150 mc, and consequently the voltage on both diodes far exceeds the peak value. The 1-ma input current is sufficiently large, however, to cause the desired diode to stabilize at a useful output, but only after a period of oscillation (Figures 10 and 11). At 450 mc, the 1-ma input is not sufficient to break the hard oscillations started by the switching transient and no useful output is obtained (Figure 12). An increased input of 2 ma insures the desired output after one cycle of oscillation (Figure 13), but at least 3 ma would be necessary to avoid the initial oscillatory

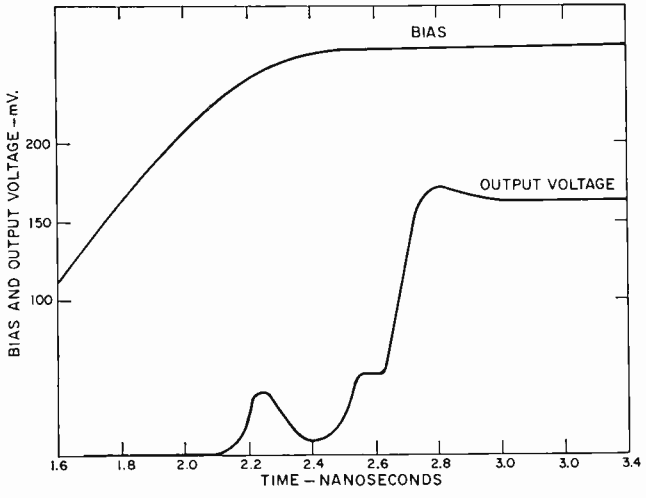
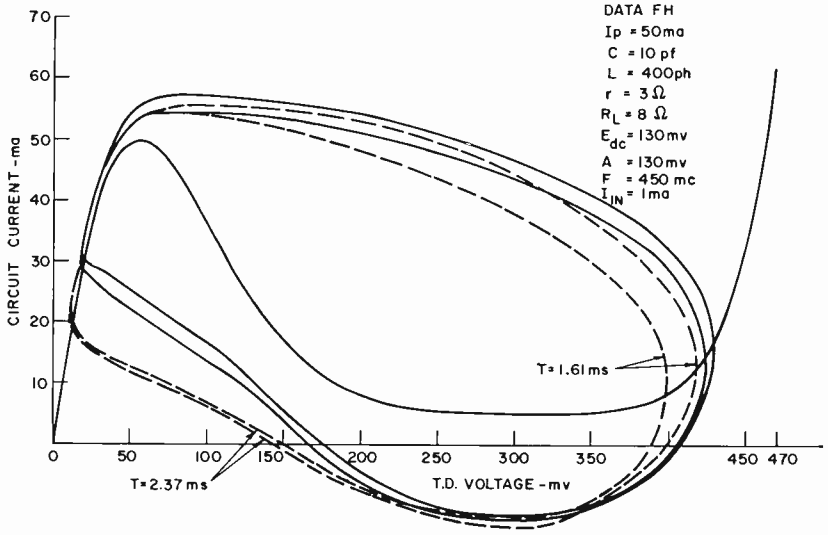
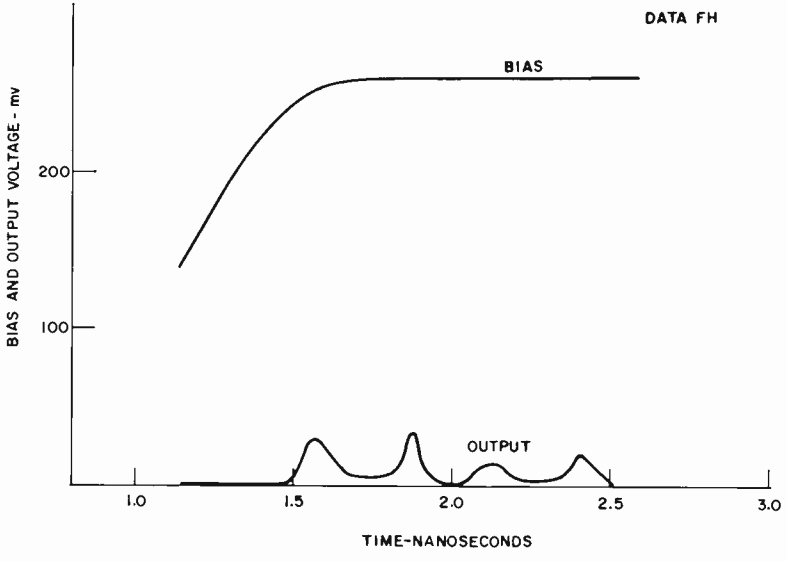


Fig. 11—Output wave shape (circuit as in Figure 10a).



(a)



(b)

Fig. 12—*V-I* characteristic (a) and output wave shape (b) of oscillatory conditionally stable circuit operating at 450 mc.

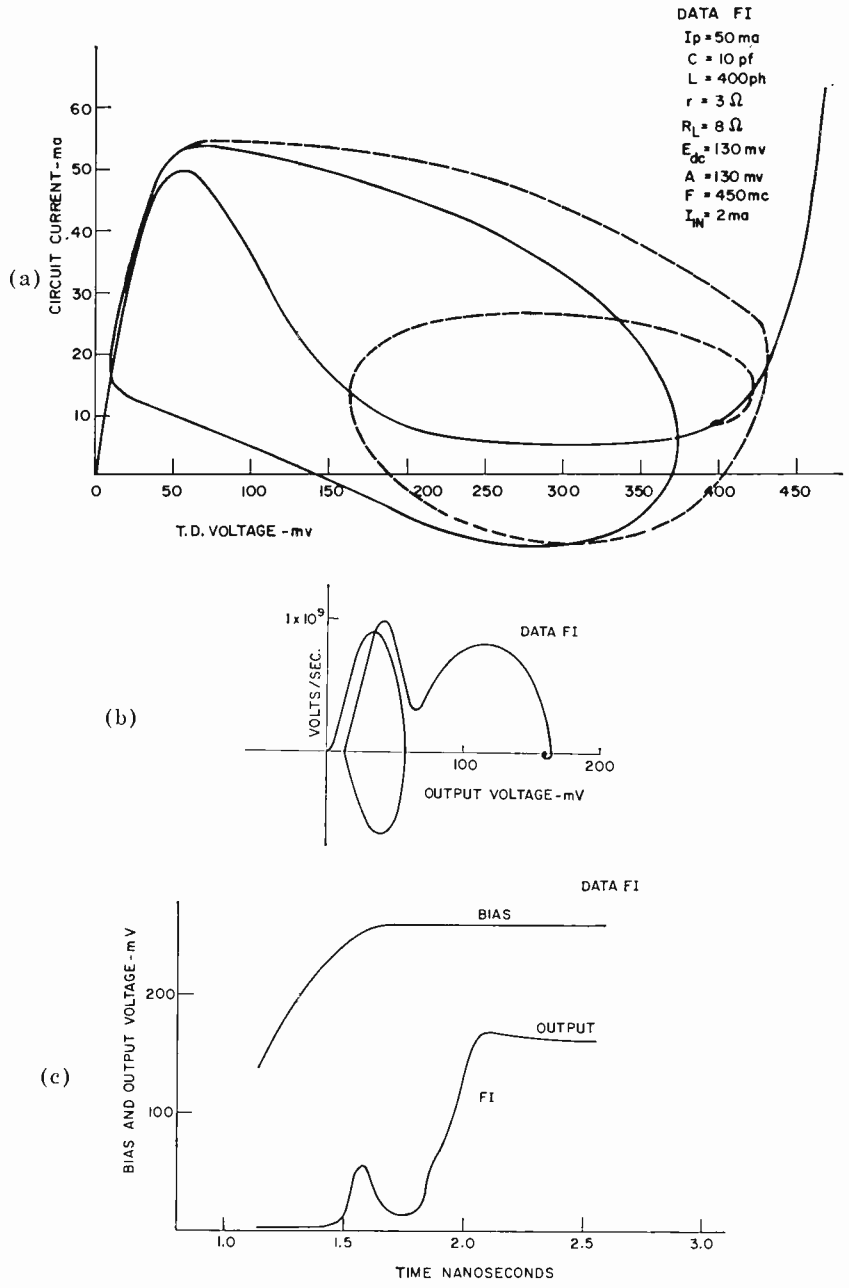


Fig.13—*V-I* characteristic (a), phase plane of output voltage (b), and output wave shape (c) of switched conditionally stable circuit operating at 450 mc.

condition that delays the desired output until the bias cycle is nearly over. This type of switching with the preliminary oscillatory transient is an undesirable mode of operation, because it requires the presence of the input current for a long period of time and produces an output for only a short part of the bias cycle (Figure 14).

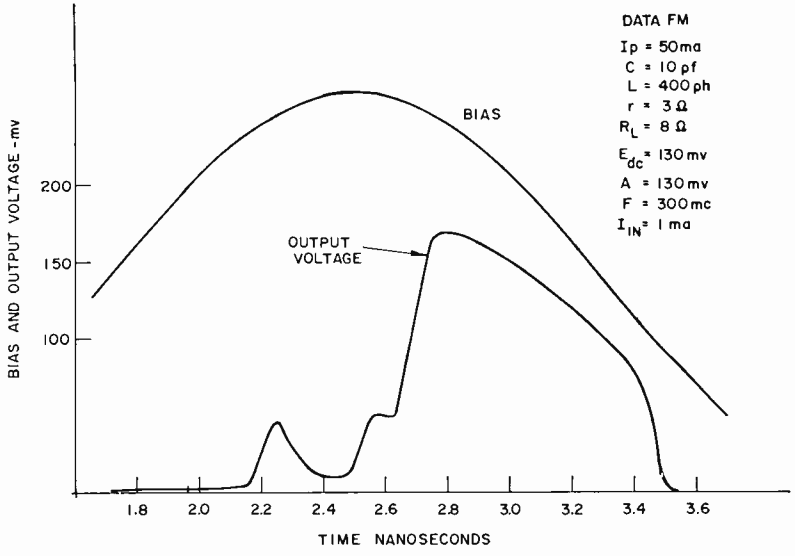


Fig. 14—Wave shape of switched conditionally stable circuit operating at 300 mc.

b. Very Small Source Resistance

If the source resistance is small, so that the circuit is classified as unstable, correct switching of the circuit is strongly dependent on the load resistance as well as the input current. Figures 5 and 6 illustrate the operation of a circuit at 150 mc with a 1-ohm source resistance, 400-ph lead inductance, 10-pf capacitance, and load resistances of 8 and 6 ohms, respectively. For the 8 ohm or 40 per cent peak current load, the input current of 2 ma is sufficient to prohibit the voltage on one diode from exceeding by any appreciable amount the diode's peak voltage. The selected diode switches to the high state and the circuit losses are sufficient to cause the transient to die out, resulting in the output waveshape of Figure 15.

If the load resistance is decreased to 6 ohms in an attempt to in-

crease the output, the initial switching action is much the same, as shown in Figure 6 (time period 3.34 to 3.86 ns), but the ultimate circuit conditions are much different (ns = nanosecond = 10^{-9} second). The diode that switches delivers energy to the circuit reactances as well as to the load. Because the load resistance is smaller than for the first case, the circuit losses are smaller, i.e., the Q is higher, and the oscillations fail to die out as with the 8-ohm load. The voltage on the switched diode is decreased to a small negative value due to the voltage

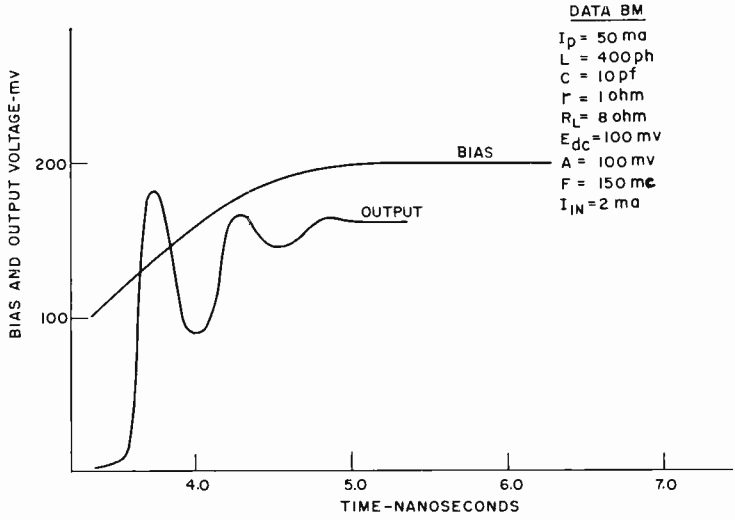


Fig. 15—Wave shape of unstable circuit operating at 150 mc.

drop across the lead inductance. The energy stored in the inductance then contributes an appreciable voltage that adds to the bias voltage on the next half cycle of the oscillation. Because the rate of change of bias voltage is now largely due to the very-high-frequency oscillatory voltage, the capacitive currents are much greater than during the initial switching action. This causes the voltage on both diodes to greatly exceed the peak voltage. Because the circuit is basically unstable, mode 1 oscillations continue for the duration of the bias period. Since the two diodes are oscillating in unison, the output voltage at the center node is essentially zero. The small output pulses that do appear are due to the imbalance of the circuit caused by the input-signal current. This type of circuit instability where a mode-2 oscillation induces a mode-1 oscillation may occur even at very-low-bias fre-

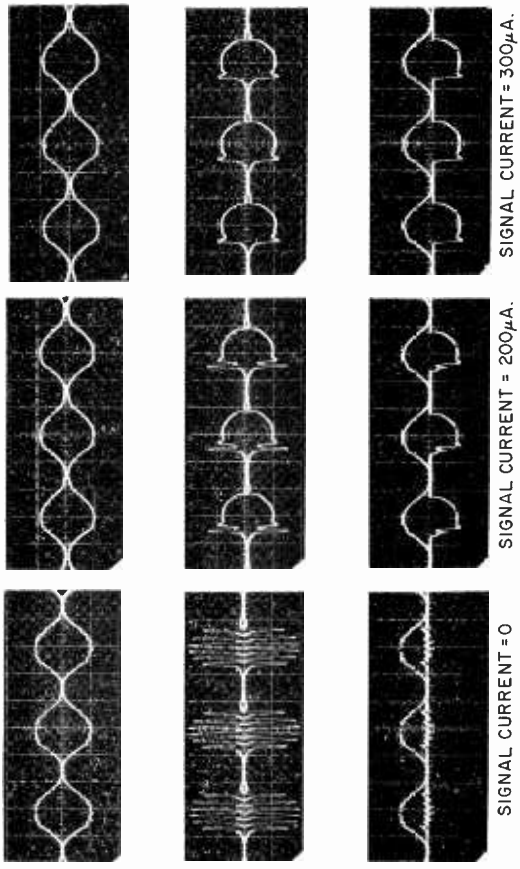


Fig. 16—Wave forms of experimental unstable circuits.

quencies. The Appendix contains a complete description of the switching of an unstable circuit driven at 50 mc. Increasing the diode capacitance to 50 pf makes the circuit conditionally stable and the 2-ma input current is sufficient to break the initial oscillation (Figure 17). Unfortunately, the initial oscillatory transient has a long period because of the large capacitance.

If the input current to an unstable circuit is insufficient to prevent the voltage on one diode from exceeding the peak voltage, then both diodes will switch and mode-1 oscillations will start immediately.

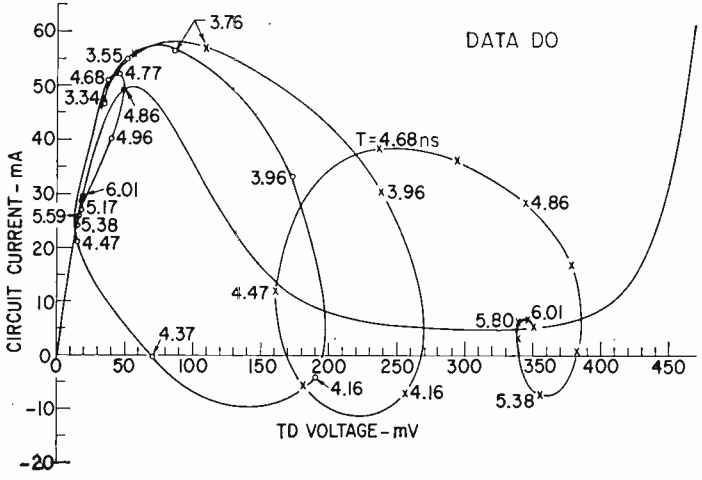


Fig. 17—*V–I* characteristics of conditionally stable circuit with large capacitance.

Figure 16 shows the oscilloscope waveforms of the output and the voltage on the two diodes for a low-frequency circuit designed to be unstable. Two output conditions were possible: (1) no output for a small or zero input but a large oscillatory voltage on the diodes as shown, or (2) a correct unipolar output corresponding to the polarity of a large input current.

c. Effects of Load Reactance

The presence of appreciable shunt-load capacitance will have the effect of reducing the difference between the voltages on the two diodes prior to switching and will necessitate increased input current to prevent both diodes from switching. Mode-1 oscillations are more likely with large load capacitance.

A series load inductance will tend to isolate the load resistance from the circuit during switching and thus reduce the input necessary to prevent both from switching. If both the source resistance and load resistance are small, however, load inductance will increase the possibility of mode-2 oscillations and the consequent mode-1 oscillations. Load reactances should therefore be held to the absolute minimum.

Summary

Although tunnel diodes can have inherent switching speeds of less than 0.1 nanosecond, balanced-pair circuits cannot be operated reliably at a comparable repetition rate because of the oscillation stability problem. Even perfectly balanced circuits require a reasonable percentage of the peak current as input to insure correct switching and the required amount increases with frequency. The output current is also limited by the stability problem, since the source resistance must be appreciable to avoid oscillations. The effect of unbalance is to reduce the effective input current. Bias voltage and phase imbalance are large factors in unbalancing the circuit as are tunnel-diode peak current and capacitance match.

Imbalance in the lead inductance is not of primary importance provided the input current is sufficiently large to inhibit one diode from switching. If both diodes switch, a large inductance imbalance might cause the circuit to lock in the wrong state, but this is unlikely. The absolute value of lead inductance causes a stability problem. Even the best available tunnel-diode-case design has an inductance of 200 ph or more. This is sufficient to cause a circuit to be unstable if high-current diodes are used. Two diodes might be mounted so that the bias leads for alternate halves of the circuit couple. This reduces the effective lead inductance and materially improves the operation of the circuit, particularly where the input is not quite large enough to prevent the voltage on both diodes from exceeding the peak value (Figure 18).

The locking period extends from a time prior to the voltage on the diodes reaching the peak value until one diode has stabilized in the high-voltage state, the other in the low-voltage state. Removal of the input prior to the decay of the initial oscillatory transient may allow erroneous switching or continuous oscillations.

Instead of making the source resistance low to achieve a low output impedance, the stability considerations require an appreciable source resistance to prevent oscillations. This limits the available output current of high speed circuits to 50 per cent of I_p or less. By using

tunnel diodes with lower peak currents and lower capacitances, the stability problem is reduced and comparatively smaller source resistances can be used making it possible to obtain more output. Circuit shunt capacitance and diode fabrication difficulties set a lower limit on the diode capacitance and limit the operating speed of the circuit. It would appear that a 15-ma 5-pf diode might provide more gain at

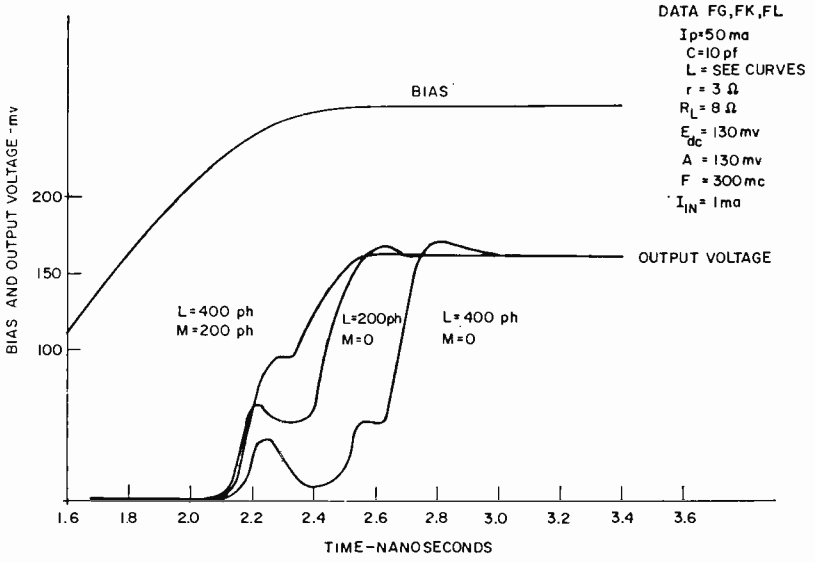


Fig. 18—Effect of mutual inductance between leads.

higher frequencies than a faster 50-ma, 10-pf diode. Rough calculations show that for the 15-ma diodes the capacitive current is decreased by a factor of 4, while for an equivalent circuit but with the same case inductance as for the 50-ma diodes, the available output is decreased by a factor of about 3. Therefore, the usable gain is actually increased at a given speed by using poorer quality diodes.

The stability problem is not limited to the balanced-pair tunnel-diode circuit but is generic to all tunnel-diode or negative-resistance circuits where an attempt is made to use one negative resistance device as the source impedance for a second. By removing the high-resistance ohmic load line normally associated with bistable circuits, power dissipation is reduced at the expense of stability. The use of the second active device increases the available negative conductance and increases the Q of the circuit, making oscillations quite probable.

It would appear that the circuit reactances and particularly the case and lead inductance limit the speed of many tunnel-diode circuits to repetition rates considerably below the rates predicted by the *RC* time constants of the diodes. In particular, the balanced-pair circuit seems to be limited to clock bias frequencies of 250 mc or below, although special low-gain applications might be possible well above 1 kmc.

ACKNOWLEDGMENT

The reported work was done while the author was a visitor at Cambridge University, Cambridge, England. I wish to thank Dr. M. V. Wilkes, Director, Mathematical Laboratory, Cambridge for making available the facilities of the Mathematical Laboratory and especially the Edsac II computer. I also wish to thank Dr. Erick Mutch for his assistance in making the arrangements for my visit and Dr. David Wheeler for advice on the use of the Edsac II computer. Many stimulating discussions with Dave Wheeler and Niel Wiesman helped in the interpretation of the computed data.

The visit at Cambridge University was made possible by the graduate study program of the David Sarnoff Research Center, RCA Laboratories, Princeton, N. J.

APPENDIX

Circuit Equations and Solution

The circuit equations for the balanced-pair equivalent circuit are

$$V_a = I_a r + L_1 \frac{dI_a}{dt} - M \frac{dI_b}{dt} + \frac{1}{c_1} \int i_{c1}\, dt + (I_a - I_b)\, R_L,$$

$$V_b = I_b r + L_2 \frac{dI_b}{dt} - M \frac{dI_a}{dt} + \frac{1}{c_2} \int i_{c2}\, dt - (I_a - I_b)\, R_L,$$

where

$$i_c = I - f(V_D),$$

$$f(V_D) = I_D.$$

For $V_D \geqq 0$,

$$I_D = A_1 V_D + A_2 V_D^2 + A_3 V_D \exp\left\{ -\left(\frac{2}{3}\right)\left(\frac{V_D}{V_P}\right)^{3/2} \right\} + A_4\, (\exp\{BV_D\} - 1),$$

for $V_D < 0$,

$$I_D = A_1 V_D + A_2 V_D^2 + A_3 V_D + A_4 \exp\{BV_D\} - 1),$$

where, for 50-ma diodes,

$$B = 35.130370,$$

$$A_1 = 39.597299,$$

$$A_2 = -78.3407,$$

$$A_3 = 1701.9337,$$

$$A_4 = 4.0408001 \times 10^{-6}.$$

The above expression[5] for the tunnel-diode characteristic gives the diode current in milliamperes. The differential equations are put in the form of a set of first-order differential equations and solved by the Runge–Kutta–Gill method. The solution is started with the currents and derivatives equal to zero. For the range of parameters of interest, the transient thus introduced dies out in a small fraction of a bias cycle. By starting on the negative-going portion of the bias voltage, a correct solution is obtained during the diode switching on the positive-voltage part of the bias waveform. On the Edsac II,* the solution time for one 150-mc bias cycle is approximately 3 minutes including line-printer output. All currents and voltages associated with each component were printed at frequent intervals during the switching process. The phase-plane plot of the output voltage was simultaneously displayed on the cathode-ray tube and photographed.

Calculation of the Capacitive Current

To predict the approximate capacitive current under balanced conditions so that an estimate of the necessary input current can be made as in the discussion of circuit switching behavior, the simplified circuit equation below can be solved.

$$i_c = C\frac{dV_a}{dt} - r\frac{dI_a}{dt} - L\frac{d^2I_a}{dt^2} - \frac{dV_o}{dt}.$$

This simplified equation ignores the effect of the mutual inductance M. The derivative of the output V_o, and the inductance term may also be neglected as small compared to the remaining two terms. This finally gives

---

[5] Developed by A. H. Simon, Mathematical Services, RCA Laboratories.

* Designed and built by Cambridge University, Cambridge, England.

$$i_c = C\frac{dV_a}{dt}\left[\frac{\dfrac{V_p}{I_p}}{\dfrac{V_p}{I_p}+r}\right],$$

where V_p/I_p approximates the resistance of the tunnel diode.

A value of input current twice this large will usually prevent one diode of the pair from switching initially. A smaller input current may be used in stable circuits, but the growth rate of the output will be dependent on the value of the input current.

Case Descriptions

The operation of high-speed diodes in an unstable circuit operated at moderately low frequencies is interesting because it illustrates certain important points. One point is that the characteristic curves of Figure 4 cannot be used to predict the stability of the circuit. This characteristic is not a true negative-resistance characteristic, but only the instantaneous difference of two separate negative-resistance characteristics. Secondly, the stability of a circuit may be such that although only one diode may switch initially, it is still possible that both diodes will switch and oscillations ensue for the remainder of the positive part of the bias cycle. This condition is difficult to predict exactly because it is the result of the one diode that switched, delivering energy to its own branch impedances plus the shunt impedance of the load resistance and the other branch of the balanced circuit containing the diode that did not switch. If the losses of this complex network are greater than the energy gained in the switching process, then the switching process is stable. If the net circuit resistance is negative over the switching period, however, oscillations will start in mode 2 and initiate oscillations in mode 1 as described previously. A 50-ma, 10-pf, 400-ph diode in a circuit with a source resistance of 1 ohm and a load resistance of 6 ohms operated at 50 mc with a 2-ma input illustrates these points. With a load resistance of 8 ohms, the circuit dissipation is large enough to insure correct switching even at 150 mc (Figure 15).

With a 6-ohm load resistance, however, the circuit becomes unstable and behaves at 50 mc as follows. When each of the bias voltages has increased to 100 mv, the voltages across diodes #1 and #2 are 46 and 52 mv, respectively. The capacitance current is only 0.25 ma and 0.45 ma

compared to the input-signal current of 2 ma. Because of the low bias frequency, the voltage drop across the lead inductance is less than 3 mv. As the bias voltage continues to increase, diode #2 switches. This causes its capacitive current to increase to a maximum of 35 ma, and the lead inductance voltage drop becomes —163 mv, which adds to the increased applied voltage of 113 mv. Because this inductive voltage is greater in the branch of the circuit containing the switched diode due to the greater rate of change of current, the circuit characteristic becomes unsymmetrical as shown in curve 2 of Figure 19.

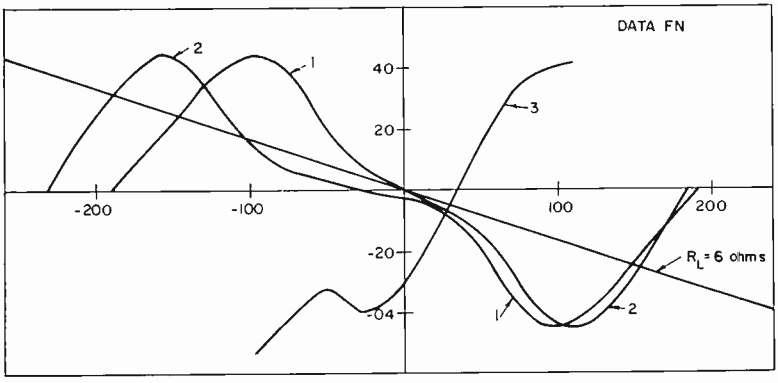


Fig. 19— Effects on I–V characteristic curve of various amounts of inductive voltage loss.

Curve 1 of Figure 19 is the characteristic expected for the maximum applied voltage under static conditions. Curve 2 is actually the characteristic at a time when the output voltage V_o (Figure 20) has reached a maximum value. The intercept of the 6-ohm load line and the characteristic curve 2 corresponds very closely to the 155-mv output voltage. At this time the applied bias voltage is only 114 mv, but the negative voltage drop across the lead inductances contributes 137 mv and 96 mv to the switched and unswitched branches, respectively. During the next 0.2 nanosecond, the current through the inductance reverses due to the cyclic exchange of energy between the inductance and capacitance, and the voltage across the inductance becomes a positive maximum. As the voltage drop increases in a positive sense, the voltage applied to the diodes decreases. Curve 3 of Figure 19, shows how the characteristic curve has already collapsed when the inductive drops are 88 mv and 24 mv, respectively, for the two halves of the circuit. At this point the bias voltage has increased to 120 mv. Mean-

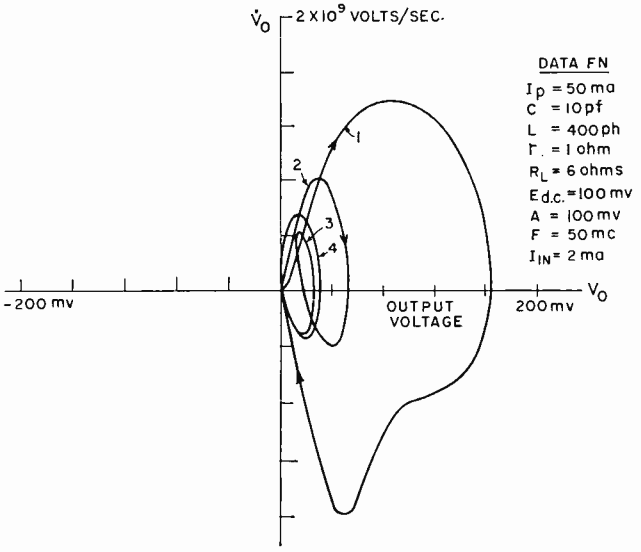


Fig. 20—Phase plane of output voltage for unstable circuit.

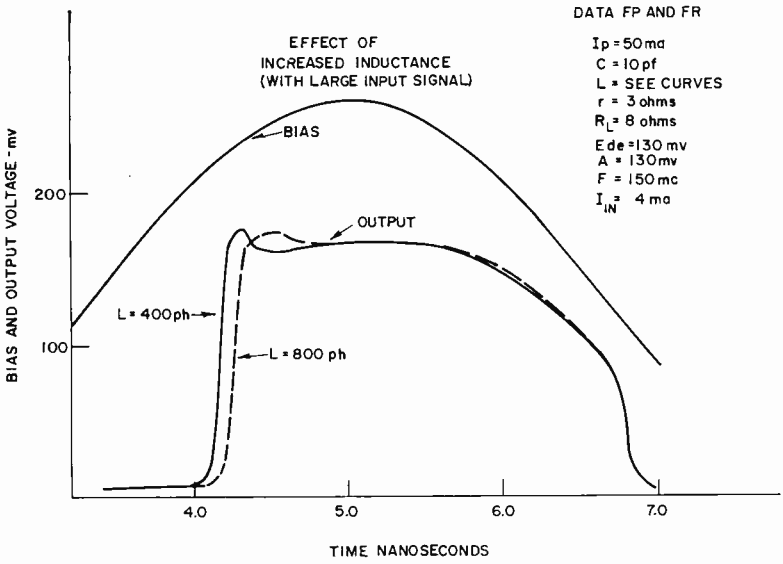


Fig. 21—Effect of increased inductance in conditionally stable circuits with large input signals.

while, the output is 84 mv (Figure 20) but decreasing rapidly as the voltage drop across the lead inductances increases to a maximum of 160 mv and 60 mv, respectively. The oscillatory exchange of energy continues and the potential across the lead inductance decreases toward a negative value. Now the potential across each of the diodes is less than the peak value but is increasing at a rate determined primarily by the oscillatory transient and not by the bias frequency. As the voltage across the diodes approaches the peak value, the capacitive currents are 2.3 ma and 3.0 ma for diodes #1 and #2, respectively. This current is larger than the 2 ma signal-current input and, consequently, both diodes switch. Oscillations now continue in mode 1, and the voltage on each diode traverses a limit cycle similar to that shown in Figure 6.

Obviously the operation of high-speed diodes at low bias frequencies is impractical unless precautions are taken to insure that the circuit is at least conditionally stable. This may be accomplished by increasing the source resistance. If the bias voltage is increased to offset the voltage lost in the increased source resistance, larger usable output currents are available with a higher source resistance since oscillations are prevented. The output impedance is, of course, higher and the output voltage more sensitive to variations in load impedance.

Additional Effects of Inductance and Capacitance

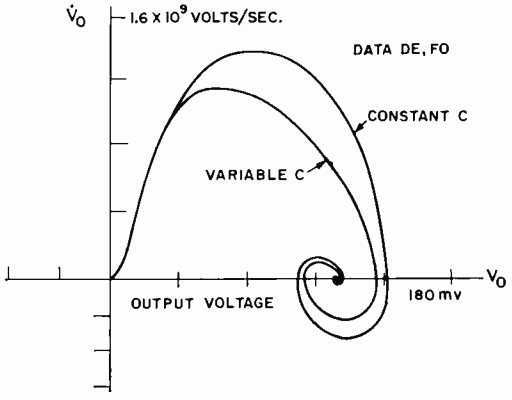
a. Inductance

The primary effect of lead inductance is to cause circuit instability. However, in circuits where large input signals might be available so as to minimize the stability problem, some lead inductance may improve the output waveshape. If there is zero lead inductance in a circuit, the output will rise rapidly during the time the diodes are switching, and then rise at a slower rate determined by the bias frequency. The effect of inductance is to cause the output voltage to increase to a greater value during the diode switching transient and then decay in an oscillatory manner. By the proper choice of inductance and source resistance, the inductive transient can be made to contribute to a relatively "flat top" output waveform (Figure 21).

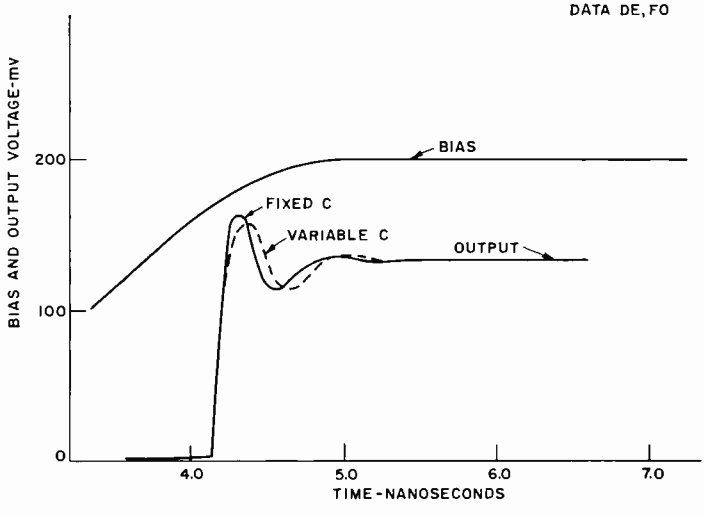
b. Effect of Variable Capacitance

Figure 22 shows the effect of a variable capacitance compared to a fixed capacitance of equal value at 300 mv. The capacitance variation was assumed to be the usual inverse square root relationship with a

contact potential of 0.5 volt. The effects of the variable capacitance on the initial switching and final output state are negligible, although there is a slight effect on the switching trajectory and output wave-shape.



(a)



(b)

Fig. 22—Effect of variable C (a) on phase plane of output voltage and (b) on output wave shape.

RETRIEVAL OF ORDERED LISTS FROM A CONTENT-ADDRESSED MEMORY*

By

Morton H. Lewin

Summary—*The problem of addressing an associative memory and reading out, in lexicographical order, all words which answer a given description is discussed. The solution to the problem involves generating an interrogation sequence which successively isolates each word in the list. It is shown that the efficiency of a given interrogation routine strongly depends on the amount of information furnished by the sensing elements used in the memory.*

A new "column-pair" sensing arrangement is proposed which results in an interrogation routine that retrieves an m*-word list in exactly* $2m - 1$ *accesses or read cycles. The time to isolate these words is independent of the number of bits per word. A proof of this is included along with a logic design for peripheral circuitry which mechanizes the routine.*

Introduction

THE ORDERED RETRIEVAL from a memory of a list of words answering a given description has received some attention recently.[1,2] This problem stems from the current interest† in "content-addressed" or "associative" memories, defined broadly as memories in which selection and retrieval of information stored is made not by specification of its physical location (address) but by specification of its character (contents). Also included in the definition is the requirement that the retrieval be accomplished without a time-consuming serial search of all words stored. The many uses to which a memory of this type can be put have been adequately described in the references cited.

The purpose of this paper is to describe an interrogation routine for an associative memory which allows retrieval of a list of words in

* Manuscript received 30 March 1962.

[1] E. H. Frei and J. Goldberg, "A Method for Resolving Multiple Responses in a Parallel Search File," *I.R.E. Trans. on Electronic Computers*, Vol. EC-10, No. 4, p. 718, December 1961.

[2] R. R. Seeber and A. B. Lindquist, "Associative Memory with Ordered Retrieval," *IBM Jour. of Research and Development*, Vol. 6, No. 1, p. 126, January 1962.

† A rather complete bibliography on this subject can be found in the two papers given in references 1 and 2.

a more efficient manner than has thus far been proposed. It is based on a new sensing arrangement for the memory, and allows the ordered retrieval of m words answering a given description in exactly $2m$-1 accesses or read cycles, thus averaging less than two accesses per word retrieved. An important feature of the routine is that the number of cycles is independent of the number of bits per word.

System Operation

Assume a memory which stores an array of words, each b bits in length, and a register of b interrogation drivers. Each interrogation driver is coupled to its corresponding storage element in every word via one or more drive ("bit") lines threading the memory. A driver is capable of three driving states, "0," "1," and "Ø" or "don't care." The pattern of states in the interrogation driver register will be called the "interrogation word." A word in the memory "matches" the interrogation word if there are no disagreements between its bits and the corresponding interrogation states. (By definition, a Ø interrogation state does not disagree with either 0 or 1 stored.) Those words which match the interrogation are "selected" or "isolated" words. The first part of any cycle, then, involves activation of the interrogation drivers with some pattern of 0's, 1's, and Ø's and selection of those words in the memory which match the interrogation word.*

As a simple example, suppose we have 10-bit words and desire to find all words which have 0 in the third place, 1 in the fifth place, and 1 in the eighth place. An interrogation of ØØ0Ø1ØØ1ØØ selects that set of m words answering this description. It is now necessary to read each of these words out, one at a time. This can be done by going through a sequence of interrogations which successively isolates each of the words originally selected.

The bits specified originally (three in this case) can be called "tag" bits or "descriptors." If their associated interrogation drivers remain fixed in the original state and only those drivers which started in the Ø state are allowed to change, the set of words isolated on any interrogation is a subset of the set of words selected on the first cycle. The generation of an interrogation sequence which appropriately changes the states of the latter set of drivers to permit the individual words selected to be read out in a relatively small number of read cycles is discussed in more detail below.

---

* Several physical realizations of this type of operation have been described in the references cited.

SENSING ARRANGEMENTS

An interrogation sequence is generated based on a set of rules by which the interrogation pattern for a given cycle depends on the pattern and the sensed results of the previous cycle. One effectively generates a "decision tree" in this manner. Considering only the set of interrogation drivers which started in the $\emptyset$ state, one desires to sequentially convert these to 0 and 1 drivers in such a way that the ambiguities resulting from multiple responses in the memory are resolved and each of the words originally selected is isolated and read out individually. To obtain an ordered list of words (i.e., lexicographical order), the conversion of $\emptyset$ drivers proceeds from one side to the other (assume from left to right) and then back again according to the prescribed set of rules.

The value of a given interrogation routine depends primarily on the speed with which all answers are retrieved and on the amount of hardware necessary to implement it.

Frei and Goldberg[1] have postulated a very simple sensing arrangement, namely, a device which indicates, for each cycle, whether or not one or more words have been selected. An interrogation routine using this sensing device has the following characteristic. First, in order to read out an answer, all interrogation drivers must be in 0 or 1 states and a YES indication must be given by the sense indicator. Suppose, on a given step, a single word has been selected but a large number of interrogation drivers are still in the $\emptyset$ state. One must spend the time necessary to convert each of these drivers to its proper state, corresponding to the word selected, before this answer can be read out. Second, there can be many interrogation cycles which isolate no words (i.e., a NO indication from the sense device).

Seeber and Lindquist[2] describe a more-detailed system in which the sensing device has three possible outputs corresponding to "no match," "one match," and "more than one match." With the additional hardware corresponding to one word-driver per word and b sense amplifiers (b is the number of bits per word), a considerable amount of time can be saved. This can be explained by noting that whenever a "one match" indication is received, the appropriate word-driver is activated and the whole word is read out (detected by sense amplifiers) immediately. This occurs even though some interrogation drivers are in $\emptyset$ states. An interrogation routine using this system has the following characteristics. First, suppose, on a given step, *two* words have been selected but a large number of interrogation drivers are in the $\emptyset$ state. Assuming the $\emptyset$ drivers are converted from left to right, suppose these two

words only differ from each other in one of the rightmost bit positions. One must spend the time necessary to convert each of the $\emptyset$ drivers to its proper state, corresponding to the identical bits of the two words selected, before the ambiguity can be resolved and an answer read out. Second, there can again be many interrogation cycles for which no words are selected.

Utilizing either of the above methods, then, it is possible to have a number of successive cycles which each isolate the same word or words and it is possible to have many cycles which isolate no words. In addition, from the description given, it is clear that the number of steps is a function of the number of bits per word.

The interrogation routine discussed below includes a different sensing arrangement and has the following characteristics: First, whenever one answer is isolated, it is read out immediately. Second, no two successive steps can isolate the same set of words. Third, assuming one or more words correspond to the descriptors given, each interrogation must select at least one of these words. Fourth, the number of cycles to retrieve all answers is independent of the number of bits per word and can be derived exactly.

Double Sense Columns

Consider that there are b sensing devices for the memory, where b is the number of bits per word. Each sensing device is connected to a pair of sense lines (columns) which threads the memory, coupling it to its corresponding storage element in every word. Under certain conditions, these lines may be the same as the interrogation drive lines. The sensing device is capable of detecting the four output states which are listed and defined in Table I.

For the purposes of this paper, it is only necessary to assume that the above signals can be detected. Several physical realizations using cores, transfluxors, or cryogenic elements can be made to generate these signals. A particular example, utilizing cryogenic devices, is discussed in Appendix III.

The important feature to note is the ability to determine, for each bit position, whether all isolated words store the same bit in this position or not. An X sensed in a given position indicates a mixture of 1's and 0's in that position for all words selected.

Interrogation Routine

Consider that the descriptors have been given and the associated interrogation drivers assume their appropriate states for the first cycle.

In general, many words (say, m) are selected and X's are sensed in many bit positions. As stated before, the "tag" drivers remain in fixed states for all subsequent cycles. On the next cycle, if one of the $\emptyset$ drivers, corresponding to sensed X position, switches to the 0 state, only those words (say, m_1) which answer the original description and also have a 0 in this position will be selected. If this driver switches to the 1 state, a different subset of words (m_2) is selected. Clearly, subsets m_1 and m_2 are mutually exclusive with $m_1 + m_2 = m$. (The trivial case when two words stored are identical is ignored.)

The advantages of the column-pair sense arrangement are the following. First, during any cycle, it is not necessary to waste time

Table I—Sense Signals and Their Meaning

Name	*Column-Pair Signals*	*Meaning*
Sense "0"	0,1	All words selected have "0" in this position.
Sense "1"	1,0	All words selected have "1" in this position.
Sense "X"	1,1	Some of the words selected have "0" in this position. The others have "1".
Sense "Y"	0,0	No words were selected. (Used only on the first cycle. This gives the indication that "no word stored answers the description given".)

changing the state of a $\emptyset$ interrogation driver which corresponds to a position sensing a 1 or 0. By only converting those $\emptyset$ drivers corresponding to positions sensing an X, one is assured that on the next cycle a subset of the words selected during this cycle will be isolated. Thus, no two successive cycles can isolate the same set of words. Second, at every step one is assured that at least one word will be selected (assuming there are some to be retrieved). By sensing 1, 0, and X for each bit position, we have eliminated any guessing concerning which $\emptyset$ drivers should be changed. Further, any $\emptyset$ driver which was advanced (say, to the 0 state) in going in one direction (say, left to right) must, at some later time, be advanced again (to the 1 state) when coming back in the other direction (right to left).

A flow chart for the interrogation routine is shown in Figure 1 and an example comparing this routine with those reported earlier[1,2] is shown in Table II. In all cases, words are retrieved in lexicographical order.

The implementation of this routine does require the additional peripheral hardware necessary to give X indications at all bit positions

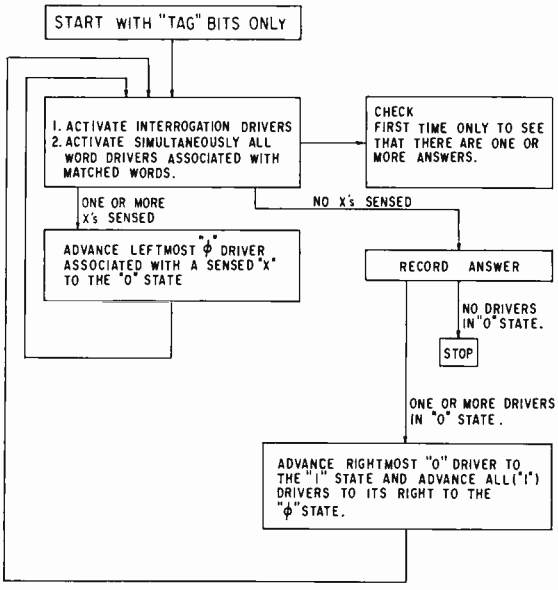


Fig. 1—Flow chart for interrogation routine (concerned only with interrogation drivers which began in the ϕ state).

Table II—Comparison of Retrieval Techniques

(a) *Words to be Retrieved In Order*

Word No.	Contents									
1	0	0	1	1	0	0	1	0	0	1
2	0	0	1	0	1	0	0	1	0	1
3	1	0	0	1	1	0	0	0	0	1
4	0	0	1	0	1	0	1	0	1	1
5	0	0	1	1	0	0	1	0	1	1

(b) *Interrogation Routine Proposed Here*

Cycle	Interrogation Driver States										Sensed Condition										Words Selected
1	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	X	0	X	X	X	0	X	X	X	1	1,2,3,4,5
2	0	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	0	0	1	X	X	0	X	X	X	1	1,2,4,5
3	0	0	ϕ	0	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	0	0	1	0	1	0	X	X	X	1	2,4
4	0	0	ϕ	0	ϕ	ϕ	0	ϕ	ϕ	ϕ	0	0	1	0	1	0	0	1	0	1	2 *
5	0	0	ϕ	0	ϕ	ϕ	1	ϕ	ϕ	ϕ	0	0	1	0	1	0	1	0	1	1	4 *
6	0	0	ϕ	1	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	0	0	1	1	0	0	1	0	X	1	1,5
7	0	0	ϕ	1	ϕ	ϕ	ϕ	ϕ	0	ϕ	0	0	1	1	0	0	1	0	0	1	1 *
8	0	0	ϕ	1	ϕ	ϕ	ϕ	ϕ	1	ϕ	0	0	1	1	0	0	1	0	1	1	5 *
9	1	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	1	0	0	1	1	0	0	0	0	1	3 *

* Denotes that word was read out on this cycle.

Table II (cont.)

(c) *Routine of Seeber and Lindquist*[2]

Cycle	Interrogation Driver States	Sensed Condition†	Words Selected
1	Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø	P	1,2,3,4,5
2	0 Ø Ø Ø Ø Ø Ø Ø Ø Ø	P	1,2,4,5
3	0 0 Ø Ø Ø Ø Ø Ø Ø Ø	P	1,2,4,5
4	0 0 0 Ø Ø Ø Ø Ø Ø Ø	0	None
5	0 0 1 Ø Ø Ø Ø Ø Ø Ø	P	1,2,4,5
6	0 0 1 0 Ø Ø Ø Ø Ø Ø	P	2,4
7	0 0 1 0 0 Ø Ø Ø Ø Ø	0	None
8	0 0 1 0 1 Ø Ø Ø Ø Ø	P	2,4
9	0 0 1 0 1 0 Ø Ø Ø Ø	P	2,4
10	0 0 1 0 1 0 0 Ø Ø Ø	1	2 *
11	0 0 1 0 1 0 1 Ø Ø Ø	1	4 *
12	0 0 1 0 1 1 Ø Ø Ø Ø	0	None
13	0 0 1 1 Ø Ø Ø Ø Ø Ø	P	1,5
14	0 0 1 1 0 Ø Ø Ø Ø Ø	P	1,5
15	0 0 1 1 0 0 Ø Ø Ø Ø	P	1,5
16	0 0 1 1 0 0 0 Ø Ø Ø	0	None
17	0 0 1 1 0 0 1 Ø Ø Ø	P	1,5
18	0 0 1 1 0 0 1 0 Ø Ø	P	1,5
19	0 0 1 1 0 0 1 0 0 Ø	1	1 *
20	0 0 1 1 0 0 1 0 1 Ø	1	5 *
21	0 0 1 1 0 0 1 1 Ø Ø	0	None
22	0 0 1 1 0 1 Ø Ø Ø Ø	0	None
23	0 0 1 1 1 Ø Ø Ø Ø Ø	0	None
24	0 1 Ø Ø Ø Ø Ø Ø Ø Ø	0	None
25	1 Ø Ø Ø Ø Ø Ø Ø Ø Ø	1	3 *

(d) *Routine of Frei and Goldberg*[1]#

Cycle	Interrogation Driver States	Sensed Condition‡	Words Selected
1	Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø	Yes	1,2,3,4,5
2	0 Ø Ø Ø Ø Ø Ø Ø Ø Ø	Yes	1,2,4,5
3	0 0 Ø Ø Ø Ø Ø Ø Ø Ø	Yes	1,2,4,5
4	0 0 0 Ø Ø Ø Ø Ø Ø Ø	No	None
5	0 0 1 0 Ø Ø Ø Ø Ø Ø	Yes	2,4
6	0 0 1 0 0 Ø Ø Ø Ø Ø	No	None
7	0 0 1 0 1 0 Ø Ø Ø Ø	Yes	2,4
8	0 0 1 0 1 0 0 Ø Ø Ø	Yes	2

† P = more than one match; 1= one match; and 0 = no matches.

To maintain consistency, 0 and 1 notation has been reversed from that discussed by Frei and Goldberg.

‡ Yes = one or more match; No = no matches.

Table II (cont.)

(d *cont.*)

Cycle	Interrogation Driver States	Sensed Condition‡	Words Selected
9	0 0 1 0 1 0 0 0 0 Ø	No	None
10	0 0 1 0 1 0 0 1 0 Ø	Yes	2
11	0 0 1 0 1 0 0 1 0 0	No	None
12	0 0 1 0 1 0 0 1 0 1	Yes	2 *
13	0 0 1 0 1 0 0 1 1 Ø	No	None
14	0 0 1 0 1 0 1 Ø Ø Ø	Yes	4
15	0 0 1 0 1 0 1 0 Ø Ø	Yes	4
16	0 0 1 0 1 0 1 0 0 0	No	None
17	0 0 1 0 1 0 1 0 1 0	No	None
18	0 0 1 0 1 0 1 0 1 1	Yes	4 *
19	0 0 1 0 1 0 1 1 Ø Ø	No	None
20	0 0 1 0 1 1 0 Ø Ø Ø	No	None
21	0 0 1 1 Ø Ø Ø Ø Ø Ø	Yes	1,5
22	0 0 1 1 0 Ø Ø Ø Ø Ø	Yes	1,5
23	0 0 1 1 0 0 Ø Ø Ø Ø	Yes	1,5
24	0 0 1 1 0 0 0 Ø Ø Ø	No	None
25	0 0 1 1 0 0 1 0 Ø Ø	Yes	1,5
26	0 0 1 1 0 0 1 0 0 Ø	Yes	1
27	0 0 1 1 0 0 1 0 0 0	No	None
28	0 0 1 1 0 0 1 0 0 1	Yes	1 *
29	0 0 1 1 0 0 1 0 1 Ø	Yes	5
30	0 0 1 1 0 0 1 0 1 0	No	None
31	0 0 1 1 0 0 1 0 1 1	Yes	5 *
32	0 0 1 1 0 0 1 1 Ø Ø	No	None
33	0 0 1 1 0 1 Ø Ø Ø Ø	No	None
34	0 0 1 1 1 Ø Ø Ø Ø Ø	No	None
35	0 1 Ø Ø Ø Ø Ø Ø Ø Ø	No	None
36	1 Ø Ø Ø Ø Ø Ø Ø Ø Ø	Yes	3
37	1 0 Ø Ø Ø Ø Ø Ø Ø Ø	Yes	3
38	1 0 0 Ø Ø Ø Ø Ø Ø Ø	Yes	3
39	1 0 0 0 Ø Ø Ø Ø Ø Ø	No	None
40	1 0 0 1 0 Ø Ø Ø Ø Ø	No	None
41	1 0 0 1 1 0 Ø Ø Ø Ø	Yes	3
42	1 0 0 1 1 0 0 Ø Ø Ø	Yes	3
43	1 0 0 1 1 0 0 0 Ø Ø	Yes	3
44	1 0 0 1 1 0 0 0 0 Ø	Yes	3
45	1 0 0 1 1 0 0 0 0 0	No	None
46	1 0 0 1 1 0 0 0 0 1	Yes	3 *
47	1 0 0 1 1 0 0 0 1 Ø	No	None
48	1 0 0 1 1 0 0 1 Ø Ø	No	None
49	1 0 0 1 1 0 1 Ø Ø Ø	No	None
50	1 0 0 1 1 1 Ø Ø Ø Ø	No	None
51	1 0 1 Ø Ø Ø Ø Ø Ø Ø	No	None
52	1 1 Ø Ø Ø Ø Ø Ø Ø Ø	No	None

and to detect the leftmost X position. A per-bit logic design to mechanize this routine is given in Appendix II. The shortening of interrogation time in many cases justifies the hardware added.

Fortunately, the routine is also amenable to an explicit calculation of the number of cycles necessary to retrieve a given number of words. In Appendix I, it is shown that $2m$-1 cycles are necessary to retrieve m words, independent of the number of bits per word.

ACKNOWLEDGMENTS

The author is indebted to J. A. Rajchman, H. R. Beelitz, and R. O. Winder for interesting discussions regarding this work.

APPENDIX I—
PROOF THAT RETRIEVAL OF m WORDS TAKES $2m$-1 ACCESSES*

On the first cycle, all interrogation drivers, except those representing tag bits, are in the $\emptyset$ state. For the remainder of this discussion, the reader should consider only the string of bits corresponding to the $\emptyset$ drivers.

Assume the number of accesses (cycles) to reach the first answer is s. On the s^{th} interrogation, $s - 1$ of the original $\emptyset$ drivers are in the 0 state. The rest are still in the $\emptyset$ state. At this time no X's will be sensed, the first answer will be recorded (leaving m-1 words to be retrieved), and the rightmost 0 interrogation driver will be advanced to the 1 state for the next cycle. Note that

$$s \leq m, \tag{1}$$

since each successive interrogation must reduce the number of words selected; thus the equals sign applies only when each successive cycle only reduces by 1 the number of words selected.

On the $(s + 1)^{st}$ interrogation, then, we have $s - 2$ "0" drivers (label their positions $P_1, P_2, \ldots P_{s-2}$) and one "1" driver on the right (label its position P_{s-1}). From this time until the time driver P_{s-2} is advanced to the 1 state, all X's sensed will appear to the right of position P_{s-1}. Hence, for this entire period the drivers in positions P_1 to P_{s-1} *do not change*. During this time the routine is tracking down that subset of words selected on the $(s + 1)^{st}$ interrogation. Assume this number of words is m_1. It is clear that during this period the routine is solving exactly the same type of problem as originally specified,

* Refer to Figure 1 for description of the interrogation routine.

but on a smaller scale. That is, it is retrieving serially a specific subset, m_1, of the original set of m words, where the "tag" bits for this smaller problem now include the original descriptors *plus* the bits associated with the drivers in positions P_1 to P_{s-1}. After this period is over, $m_1 + 1$ words have been read out, driver P_{s-2} is advanced to the 1 state, all drivers to its right return to the $\emptyset$ state and a different subset of m_2 words is selected. Again, the routine works on this new problem until all m_2 words have been found. Then driver P_{s-3} is advanced to the 1 state to isolate a new set of m_3 words, and so on.

Evidently, then, the problem of retrieving a set of m words breaks down into:

(1) s accesses to retrieve the first word, and

(2) $s-1$ similar problems with $s-1$ mutually-exclusive subsets of the m words, the first containing m_1 words, the second m_2, etc., where

$$m_1 + m_2 + \ldots + m_{s-1} = m - 1. \tag{2}$$

Suppose it takes $2m_i - 1$ accesses to retrieve any set of m_i words. If this is true, the total number of accesses to retrieve the m words is

$$s + \sum_{i=1}^{s-1} (2m_i - 1).$$

Substituting Equation (2) above, we find that

$$\text{Total number of accesses} = s + 2(m-1) - (s-1) = 2m - 1.$$

Thus, if the subproblems satisfy the rule to be proved, so does the main problem. We have, therefore, satisfied the first part of an inductive proof (i.e., if it is true for $n-1$, it is true for n).

It now remains simply to consider the smallest subproblems, namely those with 1, 2, 3, . . . answers, to satisfy ourselves that these do indeed satisfy the rule. Since every subproblem with more than one answer is always reducible to subsubproblems, after sufficient reduction one must reach the more trivial subsets.

It will now be shown that Table III holds. Recall from Expression (1) that $s \leq m$. Also, of course, $s > 1$ if more than one word is originally selected. The table enumerates *all possible conditions* for the cases $m = 1$, 2, 3, and 4. The total number of accesses is s plus the number to isolate m_1, m_2, m_3, etc.

Obviously, if $m = 1$ a unique selection is made and only one access

is needed to retrieve the answer. If $m = 2$, the first interrogation isolates two words. Adding one 0 driver gives the first answer and converting this to a 1 driver gives the second answer. This amounts to a total of three accesses for $m = 2$. For $m = 3$ or higher, different "routes" to the answers occur, depending on the number of cycles before the first answer is isolated. For $m = 3$, there are only two routes. If, on the second interrogation, an answer is found, the third interrogation will involve converting the one 0 driver to a 1 driver. This will isolate the other two words to be found. Thus, the third interrogation is exactly the same as the first interrogation to a memory with two answers. Since it has already been shown that $m = 2$ takes 3 cycles, the total needed for this "route" of $m = 3$ is 5 cycles.

Table III

m	s	m_1	m_2	m_3	Total Number of Accesses	m
1	1	0	0	0	1	1
2	2	1	0	0	3	2
3	2	2	0	0	5	3
3	3	1	1	0	5	3
4	2	3	0	0	7	4
4	3	2	1	0	7	4
4	3	1	2	0	7	4
4	4	1	1	1	7	4

Considering the other $m = 3$ route; if the second interrogation does not give an answer, the total number of words addressed in that cycle *must* be two. Therefore, on the next cycle, another 0 driver is added to isolate *one* of these two words. On the fourth cycle, this driver becomes a 1 driver and the *other* of these words is retrieved. Finally, on the fifth cycle the first 0 driver added is converted to a 1 driver to isolate the third word. Again five cycles are needed.

One can go through all possible conditions for $m = 4$ in exactly the same manner. These are listed in Table III. Evidently, as m increases the number of "routes" becomes excessive, hence this inductive proof.

To summarize, it has been shown that:

(1) A given retrieval problem, using the rules given, always reduces to a group of similar problems on a lower level.

(2) If retrieval of m words takes $2m - 1$ accesses for the $(n-1)^{st}$ level, the same is true for the n^{th} level.

(3) The rule is true for the lowest levels, $m = 1$, 2, 3, and 4.

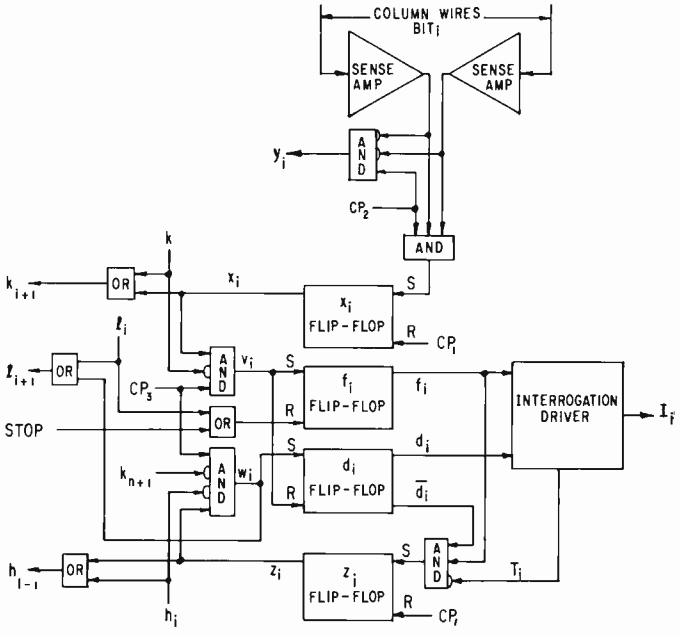


Fig. 2—Logic diagram for single bit.

Hence, the rule is proved. It is interesting to note that the number of accesses required does *not* depend on the number of bits per word, the number of X's originally sensed, the total number of words, etc., but only the number of words selected by the first interrogation.

Appendix II—Mechanization of Flow Chart

A "per-bit" logic design for the peripheral electronics is shown in Figure 2. The state of the interrogation driver, I_i, is determined from the following truth table:

f_i	d_i	I_i
0	0	$\emptyset$
0	1	$\emptyset$
1	0	0
1	1	1

If this bit is a "tag" or "descriptor" bit, $T_i = 1$ and the state of the interrogation driver is fixed. The method of accomplishing this is omitted from the diagram so as to keep only the essentials. Reference is made only to "untagged" bits in the remainder of this discussion.

The timing for each cycle involves resetting the x_i and z_i hold flip-flops (during clock pulse 1, CP_1), strobing the sense amplifiers (during CP_2), and setting up the f_i and d_i signals for the next cycle based on the sensed results of this cycle (during CP_3).

$x_i = 1$ (after CP_2) indicates a sensed X for this cycle. $k_i = 1$ denotes the fact that one or more X's were sensed to the left of bit i. Thus, $v_i = 1$ (during CP_3) denotes the leftmost sensed X. This signal advances the interrogation driver from the 0̸ state to the 0 state for the next cycle.

$z_i = 1$ (after CP_1) signifies a 0 interrogation driver (untagged). $\overline{k_{n+1}} = 1$ indicates that an answer has been reached. $h_i = 1$ denotes the fact that one or more 0 interrogation drivers are present to the right of bit i. Thus, $w_i = 1$ (during CP_3) denotes the rightmost (untagged) 0 interrogation driver. This signal advances the driver to the 1 state and propagates to the right, down the l OR gate chain, converting all drivers back to the 0̸ state.

The termination of an interrogation is detected by the condition $\overline{h_0}\,\overline{k_{n+1}} = 1$.

Since the "per-bit" circuits are interconnected with transfer signals h, k, and l, one can imagine all circuits connected in a "ring," with the chain broken at one point. By varying the position of this break, one can vary the "category" represented by the leftmost bits. By doing this, it is possible to order the information retrieved with respect to any desired criterion. This is also pointed out by Seeber and Lindquist.[2]

Appendix III—Cryogenic Memory

The READ portion of the cryogenic circuitry for a memory bit, discussed by Seeber and Lindquist,[2] is reproduced in Figure 3. Storage is obtained by a current steering arrangement (heavy lines) such that current in the left path (solid line) represents a 1 stored, while current in the right path (dashed line) represents a 0 stored. To read, the word-driver is activated and current is put on the READ line. If a 1 is stored, this current is diverted in the upper path, due to the action of cryotron A, making sense line (a) resistive, via the action of cryotron C. If a 0 is stored, the current takes the lower path, due to the action of cryotron B, making sense line (b) resistive, via the action of cryotron D. Sense lines (a) and (b) couple to the same bit in all words of

the memory. As proposed by Seeber and Lindquist, only one READ line is activated at any one time, this occurring only after a "one match" indication has been received. The current source I_0 is steered down the appropriate sense line to the sense register.

Assume a memory, containing an array of these cells, which has

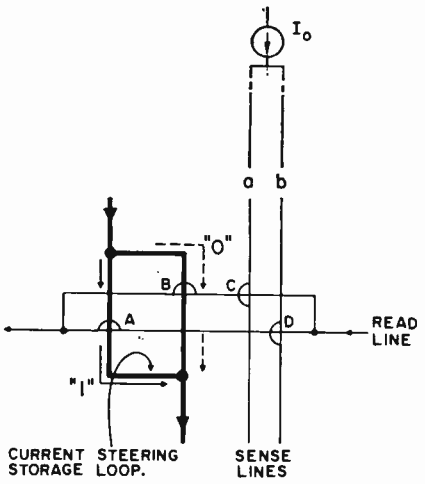


Fig. 3—READ portion of cryogenic memory cell of Seeber and Lindquist.[2]

provision for detecting whether any *individual* sense line is in the superconducting (S) state or the normal (N) state. For example, it is possible to send individual currents down all sense lines and detect whether or not a voltage is developed across each line. Other more-sophisticated techniques may be possible. Assume further that, after each interrogation, *all* word-drivers associated with matched words are unconditionally activated simultaneously. ("Unconditionally" in this case means irrespective of the number of words selected on that interrogation.*) Considering a given pair of sense lines (a) and (b), if SN is sensed we know that all the words selected have a 0 in this bit position, while if NS is sensed, all the words selected must have a 1 in this position. For these cases, the N line merely becomes more resistive as more selected words are added. More often, however, some of the words selected will have a 0 in this bit position while the others will have 1. In this case, *both* sense lines will become resistive(i.e., NN

* Seeber and Lindquist do not suggest operation in this manner. Modifications, including separate current sources for each sense line and different sense and drive logic, would be necessary to realize this mode of operation.

sensed). Note further that an SS sensed in *any* bit position must mean that no words were selected. (This indication is only useful at the start to detect whether or not there are any words answering the descriptors given.) It can be seen that the sense conditions listed in Table I apply (with N corresponding to 1 and S corresponding to 0).

Clearly, other physical realizations, which include simultaneous excitation of all words selected and provision for sensing all combinations of column-pair signals apply to this discussion.

THE SPACE-CHARGE-NEUTRALIZED HOLLOW CATHODE*

By

A. L. Eichenbaum

RCA Laboratories
Princeton N. J.

Summary—The space-charge-neutralized hollow cathode is a new source of electron beams of either very high current densities at moderate temperatures or very low temperature beams at moderate current densities. The basic structure consists of an electron-emissive cylinder containing cesium vapor. The cesium serves two purposes according to which mode is used.

For high current densities the cesium is contact ionized at the cathode surface and the ions are driven off together with thermionic electrons; thus the space-charge-limitation problem that retarded the development of conventional hollow cathodes is alleviated. Continuous current densities of 50 amperes/cm² at cathode temperatures of 2123°K have been obtained.

For low-temperature beams the cesium atoms are adsorbed on the cathode and lower its work function sufficiently to allow electron emission to occur at reduced temperatures. Current densities of 0.1 ampere/cm² with 0.030-inch-diameter beams have been obtained at cathode temperatures as low as 750°K.

These values are indicative of the state of the art, not of the ultimate capabilities of such cathodes.

Introduction

In order to extend the useful range of beam-type microwave devices to higher frequencies and higher power levels, beam current densities above those obtainable by conventional techniques are required. Also, a need exists for cathodes capable of delivering the low-density beams required for low-noise tubes but with unusually low temperatures.

Several authors[1-4] have investigated hollow cathodes as suitable

* Manuscript received 2 April 1962. A part of this work was presented at the Electron Tube Research Conference, Seattle, Washington, June 1960.

[1] H. M. von Foerster and H. S. Wu, "Thermodynamics and Statistics of the Electron Gas, 1 Electrical Properties of a Stable Spherical Electron Cloud," Contr. No. N6-ori-71 Task XIX, T. R. 3-1, El. Eng. Res. Lab., Univ. of Ill., Oct. 1950.

[2] D. F. Holshouser, "General Problems of Broadband Amplification in the Microwave Frequency Range," Contract same as in ref. 1, T. R. 13-3, Univ. of Ill., July 13, 1951.

[3] M. L. Babcock, D. F. Holshouser and H. M. von Foerster, "Diode Characteristics of a Hollow Cathode," *Phys. Rev.*, Vol. 91, p. 755, 1953.

[4] C. Epstein, C. Fried, L. D. Smullin, "Internally Coated Cathodes," Quarterly Prog. Rep., Res. Lab. of El., M.I.T., Oct. 15, 1952.

beam sources for filling these needs. The cathodes described consisted of chambers which were coated internally with an alkaline earth metal oxide, and closed except for a small exit port. These cathodes had the attractive features of being immune to ion bombardment, of affording a simple means for obtaining hollow beams, and of possessing a very rapidly increasing current–voltage characteristic. More importantly, hollow cathodes offered promise of providing very high density electron beams because of the large ratio of emitting area to exit-hole area. It was soon found, however, that at normal operating temperatures current densities in excess of 10 amperes/cm^2 could not be extracted because of space-charge limitation.[5] Furthermore, strong aberrations and astigmatism were observed in beams from such guns when used in picture tubes.[6]

In the space-charge-neutralized (SCN) hollow cathode described in this paper, the space-charge limitation problem is greatly alleviated through the introduction of positive ions. Cesium vapor is contact ionized at the walls of the uncoated cathode. When the cathode is sufficiently hot the cesium ions evaporate along with thermionic electrons. At lowered cathode temperatures, cesium is adsorbed on the cathode surface and lowers its work function; thermionic electron emission can thus occur at reduced temperature. Furthermore, because the hollow chamber is filled with a plasma, the extracted current rises extremely rapidly with increasing extractor voltage; thus lower extractor voltages may be used and transverse electron velocities, aberrations, and astigmatism are reduced.

Two modes of operation of the SCN hollow cathode are described—the high-current-density, moderate-temperature mode, and the moderate-current-density, low-temperature mode. Methods of space-charge neutralization, d-c operating characteristics, and certain avoidable space-charge instabilities are discussed. Finally, practical gun designs are considered.

General Description of an SCN Hollow Cathode

An SCN hollow cathode is shown schematically in Figure 1. It consists of an indirectly heated chamber which emits electrons from its inner surface; cesium ions are produced in the desired quantity by contact ionization at the electron emissive surface (or at a separate

[5] K. R. Brunn, *An Investigation of the Hollow Spherical Cathode*, Dissertation, Univ. of Illinois, June 1957.

[6] A. Sandor, "Emission From Miniature Hollow Cathodes," *Proc. Inst. Elec. Eng.*, Vol. 108, Part B, No. 37, p. 90, Jan. 1961.

surface within the hollow cathode). Cesium is chosen because of its low ionization potential ($V_i = 3.9$ electron volts). As is well known,[7] alkali-metal atoms of ionization potential V_i lose an electron when they contact a surface of work function $\phi_c > V_i$; they leave as positive ions if the cathode is sufficiently hot. At lower cathode temperatures the alkali-metal atoms are held on the emitter surface where they lower the emitter work function. For this case electron emission takes place

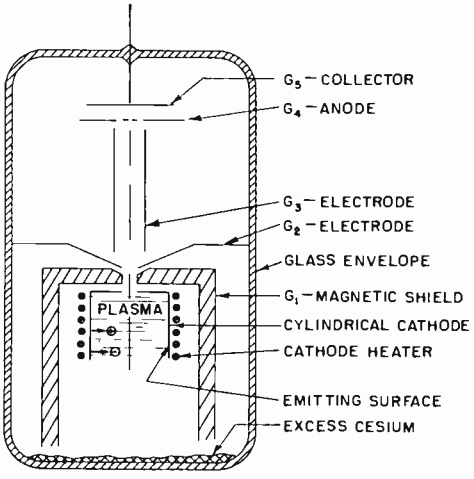


Fig. 1—Schematic drawing of a space-charge-neutralized hollow-cathode tube.

at reduced cathode temperatures. High-work-function materials such as tungsten, tantalum, and nickel are suitable for such cathodes.

The rate of electron emission and of ion generation is controlled by adjustment of cathode temperature and cesium vapor pressure. By proper choice of the ratio of emitted ion current to electron current, plasmas can be synthesized within the hollow cathode.[8] In sharp contrast to discharge plasmas, synthesized plasmas are characterized by low electron temperatures and by high degrees of ionization. For this

[7] A. von Engel and M. Steenbeck, *Electrische Gasenanthladunden*, Springer, Berlin, 1932.

[8] A. L. Eichenbaum, "The Space-Charge-Neutralized Hollow Cathode," *Electron Tube Research Conference*, U. of Washington, Seattle, June 29-July 1, 1960.

reason such plasmas have received considerable attention at these Laboratories[9-11] and more recently elsewhere.[12-14]

Data on ion and electron emission from tungsten in cesium vapor have been given by Taylor and Langmuir.[15] The electron emission from nickel in cesium vapor was studied by Hull[16] and from oxidized tungsten in cesium vapor by Kingdon.[17] The electron emission from various other metals in cesium was recently investigated by Houston.[18]

The qualitative behavior of the emission from high-work-function materials in cesium vapor can be understood from the Taylor–Langmuir plot[15] of electron emission from tungsten in cesium vapor shown in extended form[12] in Figure 2. The cesium atom arrival rate μ_n (atoms/cm^2/sec) is the parameter; the θ lines specify the fraction of the emitter surface covered with cesium. At sufficiently low emitter temperatures, cesium completely covers the surface ($\theta = 1$) giving an emitter work function of 1.8 electron volts. No ionization of impinging cesium atoms takes place at this low-work-function surface.

As the tungsten cathode temperature is increased the electron emission increases. For any given cesium vapor pressure, an emission peak is reached in the low-temperature region at $\theta = 0.55$ where a fairly continuous monolayer of cesium is still maintained on the surface. For higher cathode temperatures and lower θ values, a composite emitter surface, consisting of cesium-covered and bare tungsten

---

[9] W. R. Beam, R. C. Knechtli, R. W. Peter, Final Report Contr. No. DA36-030-sc-64443, Oct. 31, 1956.

[10] K. G. Hernqvist, M. Kanefski, F. H. Norman, "Thermionic Energy Converter," *RCA Review*, Vol. XIX, No. 2, p. 244, June 1958.

[11] A. L. Eichenbaum, K. G. Hernqvist, "Space Charge Instabilities in Synthesized Plasmas," *Jour. Appl. Phys.*, Vol. 32, No. 1, p. 16, Jan. 1961.

[12] C. V. Wilson, "Conversion of Heat to Electricity by Thermionic Emission," *Jour. Appl. Phys.*, Vol. 30, No. 4, p. 475, April 1959.

[13] J. M. Houston, "Theoretical Efficiency of the Thermionic Energy Converter," *Jour. Appl. Phys.*, Vol. 30, No. 4, p. 481, April 1959.

[14] R. C. Knechtli, J. Y. Wada, "Generation and Measurement of Highly Ionized Quiescent Plasmas in Steady State," *Phys. Rev. Letters*, Vol. 6, No. 5, p. 215, March 1961.

[15] J. B. Taylor and I. Langmuir, "The Evaporation of Atoms Ions and Electrons from Cesium Films on Tungsten," *Phys. Rev.*, Vol. 44, No. 6, p. 423, April 1933.

[16] A. W. Hull, *Conference on Gaseous Electronics*, Mellon Institute, Pittsburgh, Pa., Nov. 1949.

[17] K. H. Kingdom, "Electron Emission from Adsorbed Films on Tungsten," *Phys. Rev.*, Vol. 24, No. 5, p. 510, March 1924.

[18] J. M. Houston, "Thermionic Emission of Refractory Metals in Cesium Vapor," *Bull. Am. Phys. Soc.*, Ser. II, Vol. 6, p. 358, June 1961.

areas, is obtained. Reduced electron emission results since the average work function is increased. Contact ionization of cesium takes place at the bare tungsten areas. For $\theta = 0$, the electron emission takes place from bare tungsten.

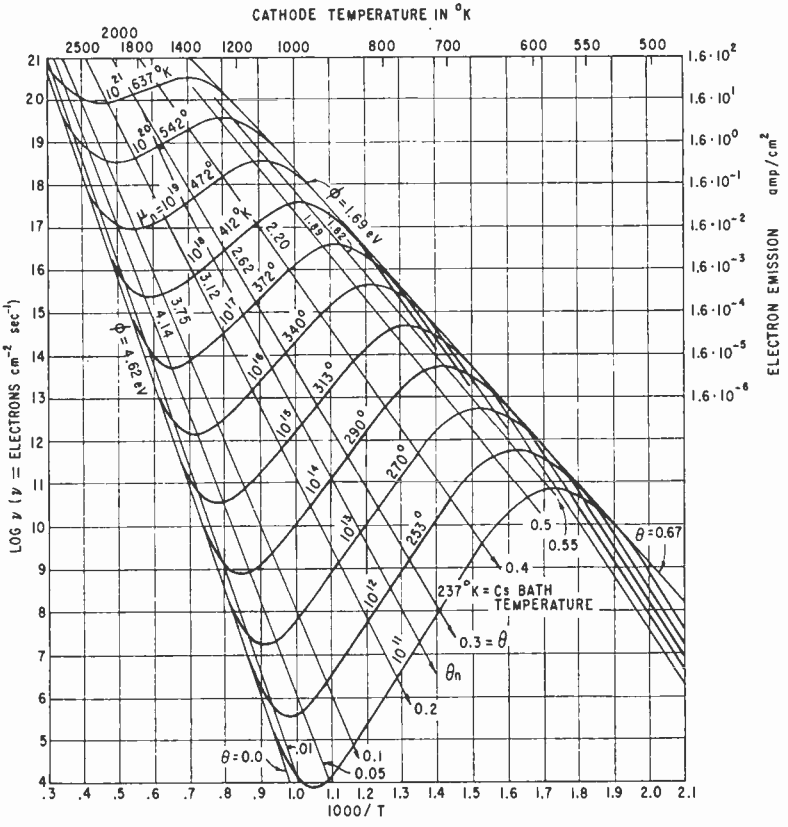


Fig. 2—Calculated field-free electron emission, ν, from a tungsten cathode in equilibrium with cesium vapor at the cathode temperature T. The bath-temperature and cesium arrival rate, μ_n, are given on each curve. Diagonal straight lines intersecting curves indicate the fraction of cathode surface covered with cesium and the effective cathode work function.

The conditions within an SCN hollow cathode employing a tungsten emitter in cesium vapor are described below for operation at conventional electron current densities but with reduced cathode temperatures, and for operation at high electron current densities but with conventional cathode temperatures.

Low-Cathode-Temperature Operation

In the low-temperature mode of operation the metal surface is not sufficiently hot to drive off enough ions for complete space-charge neutralization. The main role of the cesium in this case is to lower the emitter work function. As may be observed from Figure 2, at a given cesium pressure the peak electron emission in the low-cathode-temperature range occurs along the $\theta = 0.55$ line, where a work function of $\theta_c = 1.82$ volts is obtained.

For cathode temperatures below approximately 800°K, the electron emission is very small along this θ line. However, by making a hollow cathode with a large ratio of emitter area to electron exit aperture area, beams with useful electron current densities are obtainable at these low cathode temperatures.

High-Current-Density Operation

In this mode of operation, nearly complete charge neutrality must exist within the hollow chamber. The space-charge conditions within the chamber depend on the fraction of the emitted electrons which are extracted from the exit aperture. Since a beam of high electron density is sought, attention will be focussed on the idealized case where all of the emitted electrons are extracted; this implies proper space-charge neutralization and positive potentials on the beam-extracting electrodes.

It is convenient to discuss a simplified model wherein the emission takes place at the inner surface of a cylinder in the absence of any fields. In this model the emitted electrons move radially with r-m-s emission velocity to the axis and are extracted along the central axis. The ions, on the other hand, traverse the hollow cathode region with their r-m-s emission velocity from the cathode to the axis and continue to the opposite cathode surface. The ions thus make a double contribution to the ion space-charge density. Within such an idealized cathode the ion and electron space-charge densities at the cathode are

$$\rho_i = \frac{2J_i}{\bar{v}_i}, \qquad \rho_e = \frac{J_e}{\bar{v}_e}, \tag{1}$$

where J_i, $\bar{v}_i$ and J_e, $\bar{v}_e$, are the ion and electron current densities and r-m-s emission velocities. Space-charge neutrality requires that $\rho_i = \rho_e$ everywhere within the hollow; in particular, neutrality implies that at the cathode,

$$J_i = \frac{J_e \bar{v}_i}{2\bar{v}_e} = \frac{J_e}{2}\sqrt{\frac{m_e}{M_i}} \simeq \frac{J_e}{1000} \tag{2}$$

for cesium. For an incident particle flux of μ particles/cm^2/sec (ions and atoms) directed at a hot surface, a fraction, α, leave as ions. This fraction may be determined from the Langmuir–Saha equation;

$$\alpha = \frac{\mu_i}{\mu_i + \mu_n} = \left[1 + 2 \exp \left\{ (V_i - \phi_c) \frac{e}{kT} \right\} \right]^{-1}, \tag{3}$$

where $\mu_i + \mu_n = \mu$ is the sum of the ion and neutral particle fluxes, V_i is the ionization potential, and ϕ_c is the average work function of the cathode. The probability of ion formation due to an incident atom or incident ion is taken as equal in Equation (3). Thus using Equations (2) and (3), neutralization requires that the total flux of incident particles be

$$\mu = \frac{\mu_i}{\alpha} = \frac{J_i}{e\alpha} = \frac{J_e}{2e}\sqrt{\frac{m_e}{M_i}} \left[1 + 2 \exp \left\{ (V_i - \phi_c) \frac{e}{kT} \right\} \right]. \tag{4}$$

The ion current density is approximately equal to 1/1000th of the electron current density along the θ_n line shown in Figure 2. This line was computed for $J_i = J_e/1000$ from the data of Taylor and Langmuir and therefore satisfies the neutrality condition specified by Equation (2). For temperatures lower than those along this line, insufficient ionization prevents complete space-charge neutralization. For temperatures somewhat higher than those corresponding to this line, insufficient electron emission prevents space-charge neutrality. Neutrality can, however, be achieved along the $\theta = 0$ line of Figure 2 if a particle flux μ as specified by Equation (4) is provided.

For a certain required electron current density and charge neutrality, operation at a point along the θ_n line requires a lower cathode temperature compared to operation along $\theta = 0$ line. However, the required cesium pressure along the θ_n line is several orders of magnitude higher. In Figure 2, for example, for an electron emission of 10^{14} electrons/cm^2/sec, operation along θ_n requires a tungsten emitter temperature of 1100°K with a cesium arrival rate of 2.5×10^{16} atom/cm^2/sec. For the same electron emission density, operation along the $\theta = 0$ line requires a tungsten emitter to operate at 1750°K. For com-

plete space-charge neutralization inside of the idealized cathode, 10^{11} ions/cm^2/sec (i.e., $10^{14} \times 10^{-3}$) are required. Since nearly 100 per cent ionization of the arriving particles is obtained at this cathode temperature and $\theta = 0$, a particle arrival rate only slightly higher than 10^{11}/cm^2/sec is required.

In a practical space-charge-neutralized hollow cathode where both ions and electrons are produced at the emitter surface, it is impossible to avoid the formation of a sheath in front of the emitter.[11] As stated above, if it is desired to extract all of the emitted electrons, an excess of ion space charge in front of the emitter is desired. The potential rises over a distance of several Debye lengths in front of the cathode to approximately kT_e/e and stays at this value throughout the rest of the hollow cathode.[11] The electrons are therefore accelerated in going through this sheath while ions are retarded, and those with insufficient emission velocities are returned to the cathode.

The presence of an ion sheath in front of the emitter, recombination, collisions, external electric and magnetic fields, and the fact that the electron flow is three dimensional, modify the neutrality considerations of the idealized case discussed above. These modifications are, however, not expected to alter the behavior appreciably for low cesium vapor pressures and magnetic shielding of the cathode.

Experimental Results

Current–voltage characteristics for the low-temperature mode of operation were investigated on experimental tubes utilizing indirectly heated cylindrical hollow cathodes of the type illustrated in Figure 1. In one of the tubes a tungsten cathode was used and operated in cesium vapor at pressures between 10^{-5} and 10^{-7} mm Hg and with cathode temperatures near 750°K. At these cathode temperatures and cesium pressures, the operation was along the $\theta = 0.55$ line of Figure 2. For an electron exit-hole aperture diameter of 0.030 inch and a ratio of emitter to exit-hole aperture of 1000:1, I–V characteristics as shown in Figure 3 were obtained. Operation at slightly higher extractor voltages and cesium vapor pressures permitted beams of 100 ma/cm^2 to be drawn.

The electron temperature was measured by the retarding-field method. The velocity distribution of the emitted electrons was found to be Maxwellian, and the measured electron temperature corresponded to the cathode temperature. The low cathode temperatures, several hundred degrees lower than for conventional oxide cathodes with the same emission density, are important for low-noise microwave tubes and for other applications.

The high electron-current-density experiments were performed on a tube utilizing a tungsten cathode in cesium vapor; operation was along the $\theta = 0$ line of Figure 2, where electron emission from bare tungsten takes place. To reach high cathode temperatures, a directly heated cathode was employed with a ratio of emitter area to exit-hole area of 400:1. The electron exit aperture diameter was 0.030 inch. The current–voltage characteristics obtained with this gun are shown in Figure 4, with cesium vapor pressure as the parameter. It can be observed that as the cesium vapor pressure is raised (and the ion

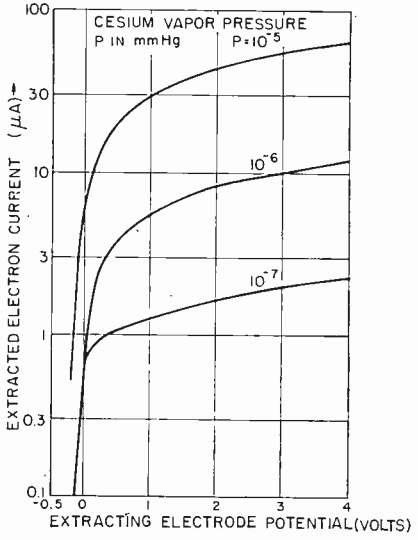


Fig. 3—The I–V characteristic of a cesium hollow cathode (low-temperature operation).

production, therefore, increased), the current extracted from the hollow cathode is increased. Because of improved space-charge neutralization in the latter case, the rise of I with V for low values of V is very steep. As may be observed, a continuous current density of 50 amperes/cm^2 was obtained with this gun at a vapor pressure giving $\mu_n \approx 10^{16}$ atoms/cm^2/sec.*

A center probe positioned on the axis of the hollow cathode was used to monitor the potential on the cathode axis. Space-charge instabilities and transitions between potential minimum and potential maximum states described in Reference (11) were studied with this

* Continuous current densities in excess of 40 A/cm^2 have also recently been obtained from such a cathode with a total beam current of 4 amperes, by T. Kuwobata of the RCA Electron Tube Division, Lancaster, Pa.

probe. Steady-state oscillations in the several-hundred-kilocycle range were detected with the probe for a narrow range of cathode temperatures near 1500°K. The oscillation amplitude constituted 0.1 per cent of the total current. Under normal operating conditions both in the low-temperature and high-density modes of operation, such oscillations did not occur.

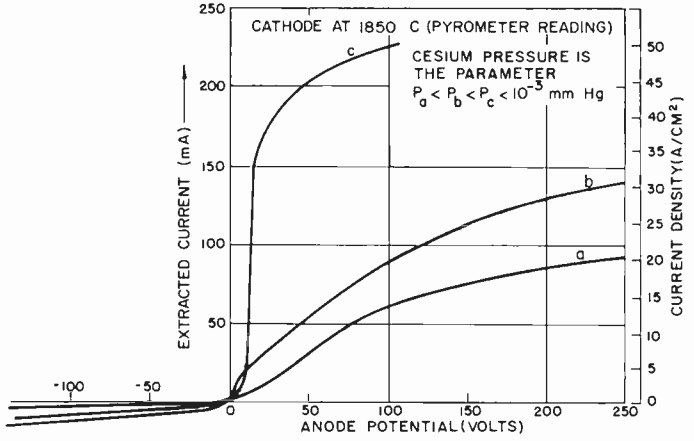


Fig. 4—The I–V characteristics of a cesium hollow cathode (high-current-density operation).

Practical Design Considerations

The practical utility of cesiated hollow cathodes as sources of low-temperature and high-current-density beams is contingent upon their meeting the following practical conditions. First, the cesium vapor should be supplied to the cathode chamber at a controlled rate to maintain equilibrium conditions. Second, the cesium vapor should be provided within the hollow cathode only when the tube is operating. Third, the cesium supply should be sufficient to give the hollow cathode a reasonable lifetime. Fourth, whatever cesium escapes from the chamber should be quickly trapped or recirculated to permit high-vacuum operation in the rest of the tube and to prevent formation of leakage paths. Finally, and very importantly, the vapor pressure and geometry must be such that no discharge occurs.

Most methods available for the control of the cesium vapor pressure do not satisfy the above requirements, and some have severe drawbacks. The techniques used in the present studies are described below.

For the low-electron-temperature mode of operation, a cesium vapor pressure of about 10^{-5} mm Hg is required within the hollow cathode chamber. The cesium vapor is conveniently generated in a container adjacent to the hollow cathode and is fed directly into the hollow cathode region only during tube operation. The cesium vapor which

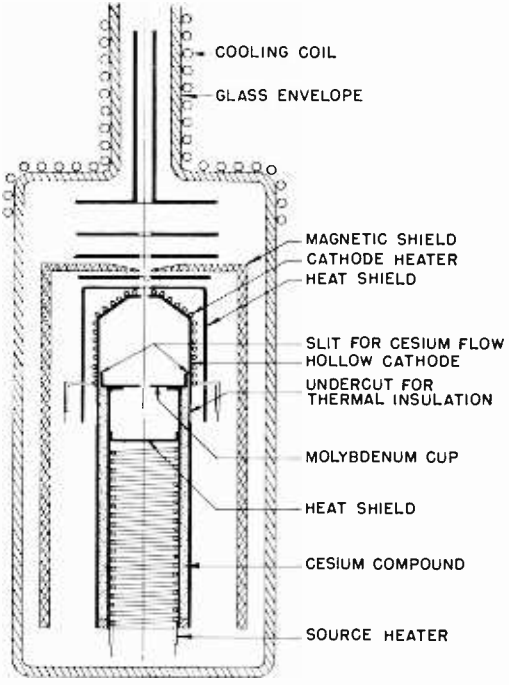


Fig. 5—Cross section of slit-feed hollow cathode.

escapes through the electron exit aperture is cold-trapped and prevented from penetrating into the high-vacuum regions of the device. A typical gun constructed with such a cathode is shown in Figure 5. A mixture of cesium chromate and silicon, which is inert at room temperature, is used as the source of cesium vapor.* The mixture is independently heated to about 700°C whereupon the following exothermic reaction takes place:

$$2Cs_2CrO_4 + Si \rightarrow 2CrO_3 + SiO_2 + 4Cs + \text{heat}.$$

* It is planned to publish details of this source at a later date.

The cesium vapor thus liberated enters the hollow cathode region through a circumferential slit formed between a cup and the cathode wall. (Alternatively, a porous tungsten plug has been used.) In the slit-feed arrangement the cesium vapor is directed toward the cylindrical cathode surface and away from the electron exit aperture; direct cesium escape is thus reduced. To maintain a desired equilibrium pressure, the rate of cesium supply must equal the rate at which cesium escapes through the electron exit aperture. The cesium supply rate is controlled by the temperature of the cesium source heater.

The cesium source container has one wall in common with the cathode. In order to provide thermal insulation between them, and thus permit them to operate at different temperatures, the cathode sleeve is undercut. For the same reason the cesium source heater does not extend to the slit cup.

The background cesium vapor pressure outside of the hollow cathode can be kept low by cooling a portion of the tube envelope where excess cesium is collected.

The scheme described above was found useful for low-electron-temperature, low-cesium-pressure operation. High-current-density operation, on the other hand, requires a different method of cesium control. In this case the required cathode temperature and cesium vapor pressure are considerably higher. The cesium container (containing in this case metallic cesium) must be located far from the hot cathode region to provide an independent cesium-feed-rate control. Some scheme for cesium recirculation may also be desirable.

The problem of discharge due to ionization of vapor outside of the hollow cathode is discussed in the appendix. The type of discharge discussed is that due to electron-current runaway arising from ions being swept back to the cathode. This type of instability has been treated in a somewhat different manner by Hernqvist[19] in connection with electron extraction from an arc-discharge plasma. Formulas were derived which give an estimate of the maximum allowable cesium vapor pressure, p_{max}, below which no discharge occurs. For the low-temperature, low-pressure mode of operation, Equation (10) in the Appendix applies.

As a typical example, let the voltage on the first electrode following the exit aperture be $V_1 = 3.9$ volts, which is also the ionization potential, V_i; let the second electrode, a distance $d = 1$ cm away, be at $V_2 = 1000$ volts. Then, since $(M/m)^{1/2} \doteq 500$ for cesium, Equation

[19] K. G. Hernqvist, "High Voltage Extraction From an Arc-Discharge Plasma," *RCA Review*, Vol. XXI, No. 2, p. 170, June 1960.

(10) gives $sp_{max} = 7.8 \times 10^{-2}$. The published[20] maximum value of the ionization yield for cesium, $s_m = 10$, is in doubt. Preliminary evidence obtained at these laboratories indicates that a value $s_m = 100$ is more appropriate. With this higher s_m value $p_{max} = 7.8 \times 10^{-4}$ mm Hg. Since in the low-temperature, low-pressure mode of operation the pressure within the hollow cathode itself is always below 10^{-4} mm Hg, no discharge would occur.

For the high-temperature, high-pressure mode of operation, typical values might be $\lambda = 0.1$ cm for the mean free path of the ions, $E_c = 1$ volt/cm for the mean field near the exit aperture, $V_1 = 4$ volts, $V_2 = 1000$ volts, $s = 100$ and $d = 1$ cm. In this case Equation (12) gives $p_{max} = 1.5 \times 10^{-4}$ mm Hg in the acceleration region. In this high-pressure mode of operation a typical pressure within the cathode chamber would be $p_0 = 10^{-2}$ mm Hg. However, by taking the first electrode to be, say, five exit-hole radii away from the exit hole, Equation (13) shows that the pressure is down by a factor of 100 from its value within the cathode chamber.

Appendix

In order to avoid excessive ionization and scattering of the electron beam, the pressure in the "high-voltage" region of the electron gun (above ionization potential of the gas present) must be sufficiently low. If the pressure is too high in the beam-accelerating region, ions which form there are accelerated back to the cathode where they cause an additional electron current to flow. This additional electron current, in turn, produces more ions, and so on. As a result the electron-extraction process can become unstable.

The vapor pressure below which stable electron flow is possible is estimated below. A model is used wherein ion generation takes place throughout the high-voltage region. The ion-generation rate is assumed to be low and the ions are swept out from the region in which they are formed rapidly so that no ion accumulation can take place.[21]

Consider the aperture of the hollow cathode to be followed by a beam-extracting electrode A_1 and a beam accelerating electrode A_2. Let the potential on the first electrode be V_1 at or *below* the ionization potential V_i; the second electrode a distance d from the first is at a voltage V_2 much above V_i. The ions produced between planes x and

[20] A. von Engel, *Handbuch der Physik*, Springer Verlag, Vol. 21, p. 508, 1956.

[21] A. v. Engel and M. Steenbeck, *Electrishe Gasentladungen*, Springer, p. 189, Berlin, 1934.

$x + dx$ due to an electron beam I_0 passing through a gas at pressure p give an incremental ion current dI_i which flows to the cathode;

$$dI_i = I_e p(x)\, s(V) dx, \tag{5}$$

where s is the number of ions formed by each electron per unit distance per mm Hg vapor pressure and $p(x)$ is the vapor pressure at plane x in mm Hg. In general, p is a function of position (x) and s is a function of voltage (V). In the beam-accelerating region, it is possible to express p and x as functions of voltage.

The incremental ion current, dI_i, arrives at the cathode and causes an incremental increase in the beam current;

$$d(\delta I_e) = N(V) dI_i, \tag{6}$$

where the weighing factor $N(V)$ is given by[22]

$$N(V) = \frac{2}{9}\left(\frac{M}{m}\right)^{1/2}\left(\frac{V_1}{V}\right)^{1/2}. \tag{7}$$

Combining Equations (5) and (6),

$$K = \frac{\delta I_e}{I_e} = \int p(x)\, s(V)\, N(V)\, dx$$

$$= \int p(V)\, s(V)\, N(V) \left(\frac{dV}{dx}\right)^{-1} dV$$

Assuming uniform vapor pressure in the interelectrode region, constant ionization probability and constant electric field, this equation becomes

$$K = ps\left(\frac{dV}{dx}\right)^{-1} \int_{V_i}^{V_2} N(V)\, dV. \tag{8}$$

[22] B. J. Thompson and D. O. North, "Fluctuations Caused by Collision Ionization," Part IV, *RCA Review*, Vol. V, p. 371, Jan. 1941.

The system is stable and no discharge occurs if

$$K < 1. \tag{9}$$

For the low-temperature, low-vapor-pressure mode of operation of the hollow cathode, the vapor pressure in the accelerating region is sufficiently low that ions formed there arrive at the cathode without suffering collisions. Ions therefore arrive with velocities corresponding to the voltage V at which they form. In this case $N(V)$ is as given by Equation (7); from Equations (8) and (9) one obtains as a stability criterion:

$$p_{\max} < \frac{dV}{dx}\left[\frac{4}{9}\left(\frac{M}{m}\right)^{1/2} sV_1^{1/2}\,(V_2^{1/2} - V_1^{1/2})\right]^{-1}. \tag{10}$$

For the high-density, high-pressure mode of operation the returning ions may suffer many collisions before they reach the cathode. Thus their velocity (or equivalent voltage) is determined by the mean free path λ and the electric field E_c within one mean free path of the cathode:

$$V = \overline{V} = E_c\lambda. \tag{11}$$

Combining Equation (11) with Equations (7), (8), and (9),

$$p_{\max} < \frac{dV}{dx}\left[\frac{2}{9}\left(\frac{M}{m}\right)^{1/2}\left(\frac{V_1}{\overline{V}}\right)^{1/2}(V_2 - V_1)\right]^{-1}. \tag{12}$$

In the high-pressure mode, for which Expression (12) applies, the vapor pressure in the acceleration region is due predominantly to the pressure of the atomic beam issuing from the hollow cathode. The background vapor pressure is presumed negligible because of efficient pumping and trapping. The pressure in the atomic beam at the first electrode, p_1, is related to the vapor pressure at the hollow cathode exit, p_0, by[23]

$$p_1 = p_0\left[\frac{0.213\,r_0^2}{l^2}\right]. \tag{13}$$

[23] N. F. Ramsey, *Molecular Beams*, Oxford Clarendon Press, 1956.

Here r_0 is the radius of the cathode exit aperture and l is the distance from the aperture to the first anode plane. If p_1 is maintained below p_{max} as calculated in Equation (12), then the region between the first and second anode with a pressure below p_1 will also satisfy Expression (12). The assumption made initially that the ion space-charge buildup is negligible in the accelerating region is fulfilled if $p < p_{max}/10$. For the case of $p = p_{max}$, the electron current can also become unstable due to ionic space-charge buildup.[21]

ACKNOWLEDGMENTS

The author wishes to thank S. Bloom and K. G. Hernqvist for valuable suggestions and discussions and M. Moi, who made valuable contributions in the design, construction, and testing of the experimental tube.

THE CALCULATION OF ACCURATE TRIODE CHARACTERISTICS USING A MODERN HIGH-SPEED COMPUTER*

BY

OTTO H. SCHADE, SR.

RCA Electron Tube Division,
Harrison, N. J.

Summary—The problem of space-current calculation is reviewed briefly to point out the inadequacy of the classic solution for modern close-spaced electron tubes. A more accurate solution for space current is obtained by calculation of the current distribution from the potential field in the grid–cathode space, corrected for space charge and initial electron velocities. A machine program set up on a large high-speed computer permits the calculation of complete voltage–current characteristics and their derivatives from basic input data. A brief account of the sequence of machine operations illustrates the large number of detailed steps and the flexibility of the program. The accuracy of the calculations can be judged by comparison of the computed and measured characteristics of a Nuvistor triode.

A NEW APPROACH TO SPACE-CURRENT CALCULATION IN ELECTRON TUBES

THE CLASSIC SOLUTION for the space current in a triode makes use of penetration factors ($D = 1/\mu$) derived from the electrostatic fields of the electrodes to establish a *mean equivalent diode potential,*

$$V_e = \frac{V_g + D_a V_a}{1 + D_a + \frac{4}{3} D_k},$$

in order that the space current may be computed from the Langmuir relations.[1-3] The effect of space charge in the cathode–grid space is

* Manuscript received 9 February 1962.

[1] A. Van Der Ziel, "Extension and Application of Langmuir's Calculations on a Plane Diode with Maxwellian Velocity Distribution of the Electrons," *Philips Research Reports*, Vol. 1, p. 97, Jan. 1946.

[2] W. R. Ferris, "Some Characteristics of Diodes with Oxide-Coated Cathodes," *RCA Review*, Vol. X, No. 1, p. 134, March 1949.

[3] P. H. J. A. Kleijnen, "Extension of Langmuir's (ξ, η) Tables for a Plane Diode with a Maxwellian Distribution of the Electrons," *Philips Research Reports*, Vol. 1, p. 81, Jan. 1946.

approximated by a 4/3 increase of the reverse penetration factor (D_k) of the cathode field through the grid. This solution is valid when the composite field and space current in the cathode–grid space are substantially uniform. The required assumptions, however, result in intolerable errors in modern close-spaced electron tubes, particularly at negative grid potentials.

It is fairly obvious that an accurate solution for the space current must be based on the detailed potential and current distributions in the grid–cathode space, and that space charge and initial velocity of the electrons must be taken into account.

A rigorous analytic treatment of the electrode flow in nonuniform fields altered by space charge is exceedingly complex and will not be attempted because it requires a prohibitively large number of successive approximations to obtain a simultaneous solution for current and charge distributions, space potentials, and electron trajectories with appropriate initial velocities in the grid–cathode space. A study of the problem indicated that the tracing of electron trajectories in the space beyond half of the cathode–grid distance is of secondary importance in accelerating fields and may be dispensed with in a solution for the total current drawn from the cathode. It is not necessary to know where the electrons pass through the grid plane, provided the current distribution can be determined at the half-way distance in the grid–cathode space.

The problem is thus reduced to a calculation of the effective field strength (including space charge) in sections (y) taken normal to the cathode, with the assumption of an independent parallel current flow in the second half of these sections. For a first approximation, this assumption can be extended to the full cathode–grid distance. It will be shown subsequently that this simplification leads to errors only at highly retarding field conditions, i.e., at small currents near current cutoff.

The Electrostatic Field in the Cathode–Grid Space

The first step required in an accurate solution for the space current is the calculation of the electrostatic potential field in the grid–cathode space. The extreme complexity of a general solution is avoided by limiting the calculation of potential fields to parallel-plane triodes having a grid of uniformly spaced parallel wires; cylindrical triodes or structures having nonuniform spacings can be converted into one or a number of equivalent parallel-plane triode sections by transformation or segmentation.

The electrostatic potential field in the grid–cathode space of a

parallel-plane triode is calculated from its geometric dimensions in normalized form using the grid-wire radius R as a unit. The electrode structure can be specified by the quantities X, Y, and C, where

$$X = d_1/R = \text{normalized cathode-to-grid-center distance,}$$

$$2Y = P/R = \text{normalized grid-wire pitch,}$$

$$C = d_2/R = \text{normalized grid-center-to-anode distance.}$$

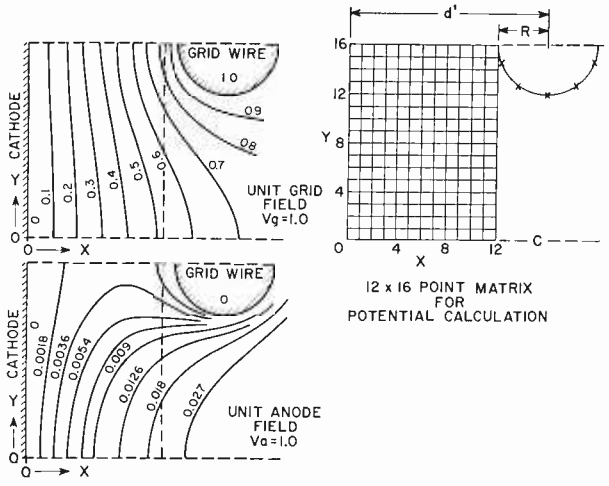


Fig. 1—Unit-potential fields in the cathode–grid space of a triode of Table I.

The field calculation is made in two parts: (1) the unit potential grid field, $v_{g(1)} = f(x,y)$; (2) the unit potential anode field, $v_{a(1)} = f(x,y)$, as shown in Figure 1. These fields are computed for a matrix of 204 points by applying one volt to the grid or anode with the remaining electrodes at zero potential. The total potential $v_{(x,y)}$ of a matrix point for any given combination of electrode potentials V_g, V_a is given by

$$v_{(x,y)} = v_{g(1)}V_g + v_{a(1)}V_a = f(x,y). \tag{1}$$

A typical set of unit potentials is shown in Table I.

Amplification Factors (μ)

Although the electrostatic amplification factor $\bar{\mu}_0$ is not required for the evaluation of equivalent potentials, it is very useful as a normal-

izing factor for the anode voltage, and is readily calculated from the unit potential functions.

The amplification factor μ at a point in the grid–anode space is, by definition, the potential ratio between grid and anode:

$$\mu = \frac{v_{g(1)}}{v_{a(1)}}. \tag{2}$$

The mean electrostatic value $\bar{\mu}_0$ in any plane x in the gride–cathode space (which is the object of normal μ calculations) is

$$\bar{\mu}_0 = \frac{\bar{v}_{g(1)}}{\bar{v}_{a(1)}}, \qquad x < x_c, \tag{3}$$

where $\bar{v}_{g(1)}$ and $\bar{v}_{a(1)}$ are the mean unit potentials and x_c is the plane just contacting the grid wires. The value $\bar{\mu}_0$ can therefore be calculated accurately from the potential data given in Table I. For the example used below, $x_c = 12$. For a given set of dimensions, the electrostatic (mean) value $\bar{\mu}_0$ is independent of the distance x from the cathode within the restriction $x < x_c$, although the variation of point μ-values becomes larger with increasing x values. The value $\bar{\mu}_0$, however, is not independent of the model dimension X, as generally assumed in μ calculations.*

Figure 2 shows the average potentials $\bar{v}_{g(1)}$ and $\bar{v}_{a(1)}$ from Table I as functions of x. The ratios $\bar{v}_{g(1)}/\bar{v}_{a(1)} = \bar{\mu}_0$ are seen to be constant for any value $x < 12$. For $x > 12$, some of the electrostatic flux lines in the space between grid wires do not penetrate the grid but terminate on the grid wires. They do not, therefore, contribute to the total field strength at the cathode.

Effective Potentials and Space-Current Distribution

The total field $v_{(x,y)} = f(x,y)$ is divided into 16 sections $v_{(y)} = f(x)$ which are treated as independent parallel triodes. The generally nonlinear electrostatic fields $v_{(y)} = f(x)$ of these triodes are replaced by linear functions representing "equivalent" parallel-plane diodes so that the space current $i_{(y)}$ can be computed from the Langmuir relations. The total current is the sum of all section currents.

The evaluation of equivalent diode potentials for calculation of the section currents (i_y) requires several steps. It is well known that the

* The majority of classic μ-calculations assume an infinite cathode–grid distance and give an asymptotic value.

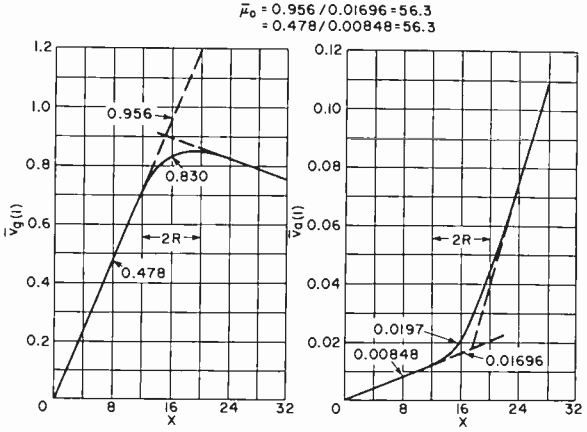


Fig. 2—Mean potentials $v_{g(1)}$ and $v_{a(1)}$ (from Table I) as functions of x.

Table I—Electrostatic Space Potentials (Unit Potentials) for $X/Y/C = 4/4/28$

		Anode Field $v_{a(1)} \times 10^2 = f(x,y)$, $\bar{\mu}_0 = 56.3$								
	$y =$	0	2	4	6	8	10	12	14	16
$x =$	2	0.300	0.294	0.276	0.248	0.214	0.180	0.149	0.128	0.121
	4	0.613	0.600	0.561	0.503	0.430	0.355	0.288	0.242	0.226
	6	0.952	0.931	0.868	0.771	0.650	0.521	0.406	0.326	0.298
	8	1.335	1.30	1.21	1.06	0.877	0.674	0.487	0.357	0.310
	10	1.780	1.74	1.61	1.40	1.12	0.809	0.511	0.300	0.227
	12	2.310	2.26	2.09	1.81	1.41	0.934	0.441	0.095	0.0
	14	2.960	2.89	2.68	2.32	1.80	1.09	0.229	0.0	0.0
	16	3.74	3.66	3.43	3.01	2.37	1.43	0.0	0.0	0.0

		Grid Field $v_{g(1)} = f(x,y)$								
	$y =$	0	2	4	6	8	10	12	14	16
$x =$	2	0.106	0.106	0.109	0.113	0.119	0.125	0.130	0.134	0.135
	4	0.209	0.211	0.217	0.226	0.238	0.250	0.262	0.271	0.274
	6	0.309	0.311	0.321	0.335	0.354	0.376	0.397	0.413	0.419
	8	0.403	0.407	0.419	0.440	0.468	0.501	0.537	0.566	0.577
	10	0.489	0.494	0.510	0.536	0.575	0.625	0.683	0.737	0.761
	12	0.564	0.570	0.589	0.622	0.670	0.739	0.831	0.939	1.0
	14	0.628	0.635	0.655	0.691	0.746	0.828	0.954	1.0	1.0
	16	0.679	0.685	0.705	0.741	0.796	0.877	1.0	1.0	1.0

electrostatic potential functions of a diode or triode are depressed by the negative space charge injected by an electron current, as shown in Figure 3. The resultant gradient changes manifest themselves as an increase in the effective grid–cathode capacitance which can be duplicated in a space-charge-free model by a reduction of the normal grid–cathode distance d to an equivalent shorter distance d'.

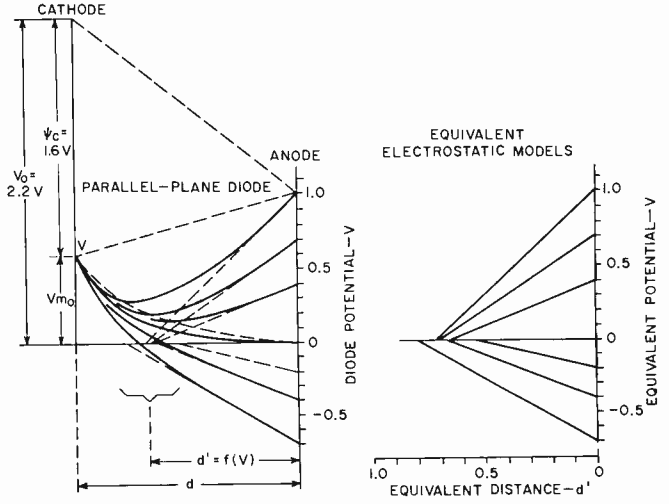


Fig. 3—Space-potential functions, $v = f(x)$, in a parallel-plane diode (left) and linear potential functions of an equivalent parallel-plane capacitance (right).

The distance d' for equal capacitance can be computed for a theoretically perfect parallel-plane diode or triode as follows. The potential gradient dv/dx at the anode of the actual potential function with space charge present is matched by the constant gradient line of a space-charge-free diode (broken lines). The intersection of the constant-gradient lines for two slightly different anode potentials occurs at the distance d'.[1] As can be seen, all intersections occur close to the potential level of the "exponential point" which is defined as that anode potential for which the anode gradient is equal to zero. The distance d' can therefore be determined with good accuracy from the intersection of one gradient line with the fixed potential level of the exponential point. The potential of the exponential point has, therefore, been selected as the reference potential for all calculations, i.e.,

$$V = 0 \text{ for } \frac{dv}{dx} = 0. \tag{4}$$

Thus the cathode of the equivalent space-charge-free diode model has the convenient potential $V = 0$. Its spacing d' is a function of current (see Figure 3) which can be derived from the Langmuir relations in normalized units, as follows:

$$\rho = \frac{d}{d'} = \xi_o \frac{I}{I_o} \frac{\gamma}{\eta_v}, \tag{5}$$

where $\xi_{(o)}$ = normalized distance d for $\eta_a = 0$,

I_o = current at the exponential point,

$\gamma = d\eta_a/d\xi_a$ = potential gradient* at the anode,

$\eta_v = 11.605$ volts = normalized diode potential (measured from the exponential point).

The reference ratio ρ_r is shown in Figure 4 as a function of the equivalent diode potential V_r for $T_k = 1000°\text{K}$, $I_s/I_o = 1000$, and $\xi_o = 2.509$. The scale factors for other temperatures and emission ratios are

$$\rho = \frac{\rho_r \xi_0}{2.509} \quad \text{and} \quad V = V_r \frac{T}{1000}. \tag{5a}$$

The space-charge correction of elemental sections in the actual triode requires that the cathode–grid distances d' of the electrostatic model used for calculations satisfy Equation (5). The simultaneous solution for d' and V can be obtained with adequate accuracy by a linear interpolation from the equivalent potentials V_d' calculated from two electrostatic models having the distance ratios $\rho_1 = 1.5$ and $\rho_2 = 1.8$, as indicated by the broken lines in Figure 4.

To apply this solution to practical triodes, it is necessary to convert their nonlinear potential function $V_{(y)} = f(x)$ (see Figure 5) to constant-gradient functions representing equivalent parallel-plane diodes. This conversion poses two problems: (1) where to locate the anode of the diode, and (2) how to replace the nonlinear function by an "equivalent" straight line. Basic considerations indicate that the anode of an equivalent diode should be located at the first "break point" of the electrostatic potential function, which for positive gradients occurs

* The negative gradients (γ) for retarding field conditions can be found in: G. Diemer and H. Dijkgraaf, 'Langmuir's ξ, η Tables for the Exponential Region of the I_a–V_a Characteristic," *Philips Research Reports*, Vol. 7, p. 45, Feb. 1952.

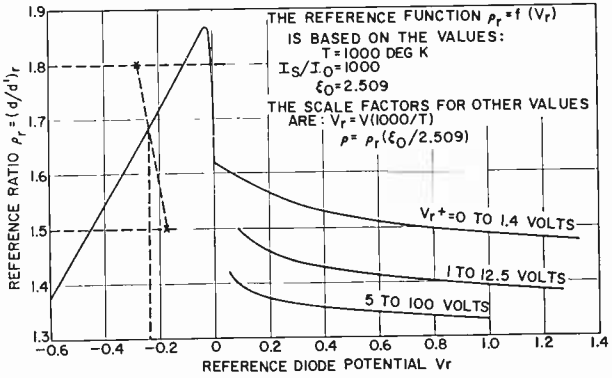


Fig. 4—Equivalent distance ratio (ρ_r) as a function of effective diode potential (V_r).

in the region between grid wires (see Figure 2). There are, however, potential functions exhibiting an electrostatic minimum ($y = 8$ in Figure 5), which should be considered as a location for the anode of the equivalent diode.

An extended numerical study indicated that a satisfactory equivalent can be obtained with a fixed anode location of the equivalent diode in the plane containing the grid-wire centers. For this condition, the function $V_{(y)} = f(x)$ must be limited to a range between $x = 0$ and

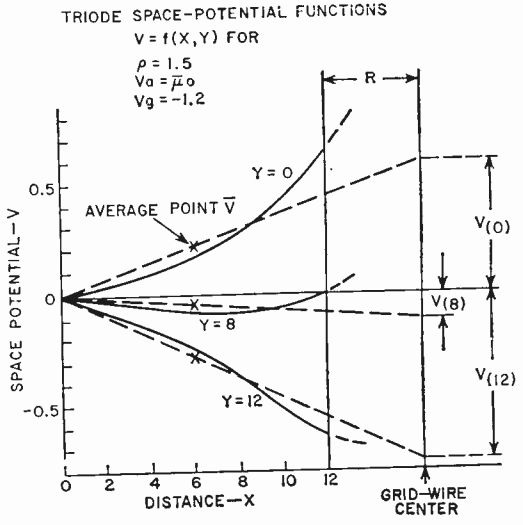


Fig. 5—Electrostatic space-potential functions of a triode.

$x = 12$ (Figure 1), which is the plane contacting the grid wires on the cathode side. The equivalent diode potential $V_{(y)}$ is extrapolated by drawing a straight line through the origin and the average point $\bar{v}$ of the truncated function $V_{(y)} = f(x)$, as illustrated in Figure 5. The correct combination of model distance d' and potential $V_{(y)}$ satisfying Equation (5) is obtained by computing $V_{(y)}$ for two models having the distance ratios $\rho_1 = 1.5$ and $\rho_2 = 1.8$ and using the linear interpolation indicated in Figure 4.

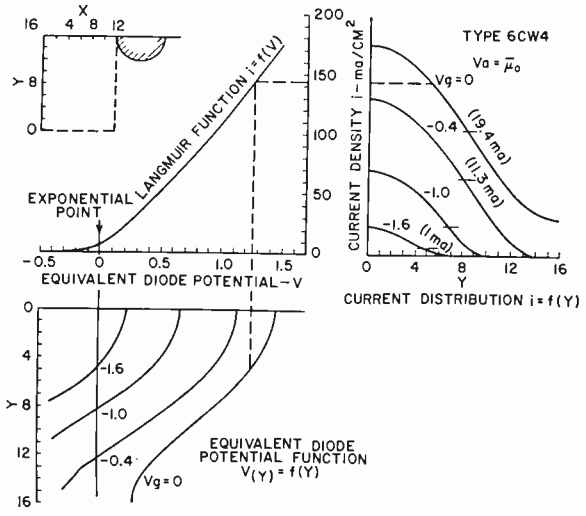


Fig. 6—Potential and current distribution for four grid potentials (V_g) of a Nuvistor triode at $V_a = \bar{\mu}_0$.

The diode currents and distribution $i = f(y)$ can then be computed from the Langmuir functions using the potentials $V_{(y)} = f(y)$, as shown in Figure 6 for the anode potential $V_a = \bar{\mu}_0$ and for four different grid-wire potentials of a 6CW4 Nuvistor triode. The computed mean current densities and total currents for the particular cathode area (0.185 square centimeter) are indicated.

The equivalent electrostatic potential fields used in the calculation are shown in Figure 7 for two operating conditions. The potential field in the actual cathode–grid space containing the space charge is obtained in first approximation by replacing the equivalent constant-gradient functions of Figure 7 by the corresponding space-charge functions (Figure 3), as shown in Figure 8. For an appraisal of electron trajectories, the reader may rotate the illustration 180 degrees. The space-charge "hill" in front of the cathode is built up by the

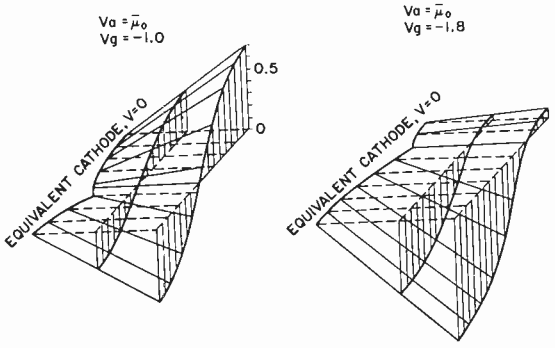


Fig. 7—Electrostatic models used for calculation of equivalent diode potentials of a triode.

large number of electrons which have insufficient initial velocities to overcome the retarding fields to the potential minimum or grid plane. The constant-potential contours marked —0.1 and —0.35 on the horizontal planes $V = 0$ indicate the potential barriers and boundaries of current flow, i.e., the regions where the computed current has decreased to 2 per cent of the highest current obtained at the center between grid wires ($y = 0$).

The areas A_o obtained by the assumption of an independent parallel current flow in the y-sections are clearly not sharp limits outside of which electrons from the cathode are turned back. The shaded vertical sections at various distances (x) from the cathode show a spread of the positive V-functions toward the cathode. The effective area A_e of current flow is, therefore, larger than A_o. The correction factor A_e/A_o

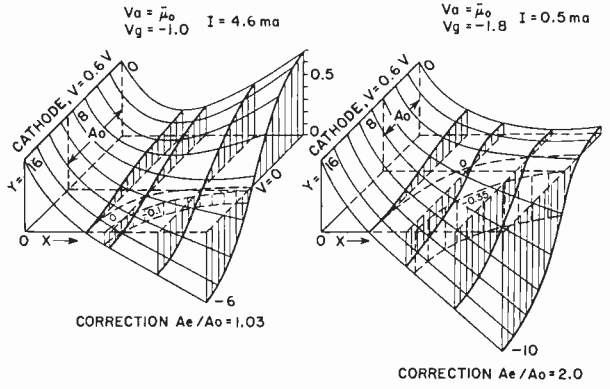


Fig. 8—Potential fields with space charge in the cathode–grid space of a triode constructed from Figure 7 and Figure 5.

for the effective cathode area is clearly a function of the potential modulation in the grid plane. It can be determined by calculation of electron paths or by comparing the computed currents with currents measured on carefully constructed triodes duplicating the constants used in the calculations. The correction becomes negligible when the amplitude of the positive portion of the potential modulation is high and when A_o extends over more than half the cathode plane ($y > 8$).

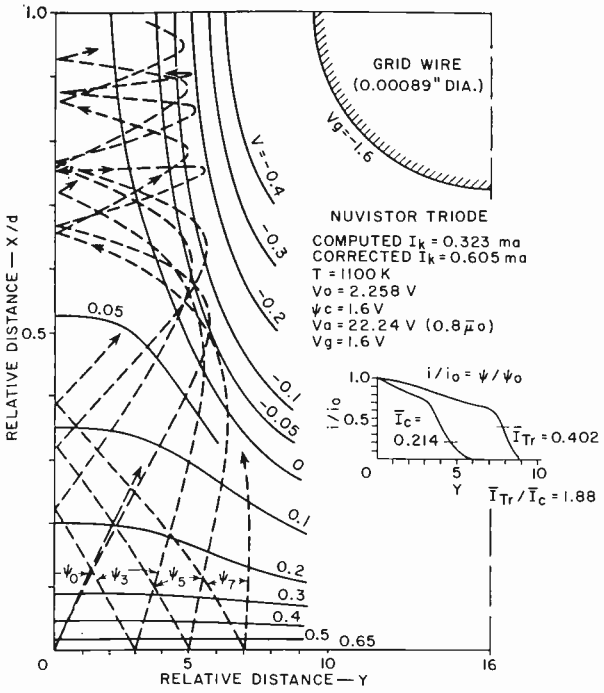


Fig. 9—Limiting electron trajectories and corrected current distribution (ψ/ψ_o) for a triode operating near current cutoff.

For low values of A_o and low positive potentials in the center between grid wires, however, the number of electrons "funneled" into the area A_o can represent a substantial percentage of the small computed current as indicated by the large correction factor $A_c/A_o = 2$.

A set of electron trajectories computed for a similar operating condition is shown in Figure 9. An initial velocity 0.2 volt in excess of the cathode-surface potential of 0.65 volt is assumed, so that no electrons can pass the —0.2-volt potential contours. The relative current distribution is given by the ratio ψ/ψ_o of the angles at the cathode

including all trajectories passing through the grid plane. The inset shows a comparison with the current distribution i/i_o computed without trajectories. The current computed from the trajectories is larger by the factor $\bar{I}_{T_r}/\bar{I}_c = 1.88$. The agreement with the measured current increase (a factor of $A/A_o = 2$) is a good indication that the discrepancies in the exponential region are caused to a large extent by the assumption of a laminar current flow.

In view of the good correlation between hand-calculated and measured currents for a large range of practical operating conditions, it was decided to dispense with a second approximation requiring the tracing of electron trajectories in the space-charge-corrected field and set up a computer program for machine calculation of triode voltage–current characteristics and their derivatives g_m, r_p, and μ, as outlined above. To obtain program flexibility, and to cover a wide range of parameter variations, it became necessary to combine and transfer earlier partial programs to a large computer (such as the RCA 601 or IBM 7090) having sufficient storage capacity for tables and data. Use of a large machine with its high-speed arithmetic circuits also reduces machine operating costs. The following brief description of the machine program set up for triode calculations illustrates the large number of operations and relative complexity required for the first-order solution of the problem discussed above.

Computer Program

The computer program which was set up is a numerical analysis program requiring floating-point and high-speed-arithmetic circuits. It includes instructions for calculating the electrostatic potential fields of a triode, and a generalized table of Langmuir relations computed for this purpose. The complete program contains approximately 5300 machine instructions and requires more than 5200 registers for storing tables and incidental information. In addition to performing the calculation of voltage–current characteristics and derivatives for parallel-plane triodes for all practical parameter values by the method outlined in the preceding section, the program contains transformation routines for calculating cylindrical structures, routines for comparing a number of designs, and routines for calculating nonuniform structures by segmentation. Also, an effort was made to keep the input simple, to keep the output understandable, and to provide for ease of computer operation.

The program-control codes and design data are entered into the computer core memory by means of a program tape and two punched

cards per tube design. The data form to be filled out by the engineer is shown in Figure 10.

Machine Program For Parallel-Plane Triodes

Calculation of Electrostatic Potential Tables (Subprogram)

The input data are the grid–cathode distance d_1 (x dimension), the grid pitch 1/T.P.I. (y dimension), the grid–plate distance d_2, and the grid-wire diameter $2R$. (The z dimension of the 3-dimensional spaces is uniform.)

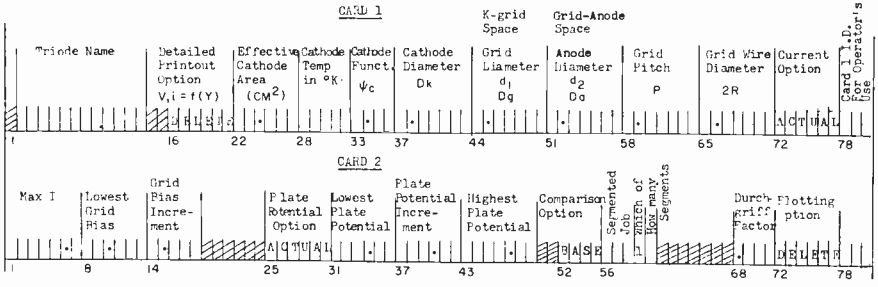


Fig. 10—Data form of input instructions for the machine calculation of triode characteristics.

The machine normalizes the dimensions using the grid-wire radius (R) as a unit, and solves the simultaneous equations set up for a system of unipole and multipole line charges located at the grid-wire centers and their image points on opposite sides of the anode and cathode planes to obtain constant potentials at ten points on the grid-wire diameter (indicated in Figure 1). With these line charges, the machine calculates the electrostatic unit potentials $v_{g(1)}$ and $v_{a(1)}$ for a matrix of 204 points in the grid–cathode space (x,y) (see Table I). This calculation is performed for two equivalent grid–cathode distances d_1' related to the real distance d_1 by the ratios $\rho_1 = d_1/d_1' = 1.5$ and $\rho_2 = 1.8$.

The output data from this part of the program are four potential tables (816 potential values) and the electrostatic amplification factors $\bar{\mu}_0$ calculated for the two distance ratios. The data are stored for use in the main program. Machine time for this sub-program is 12 seconds.

Calculation of Voltage–Current Characteristics (Main Program)

The input data are as follows;

(1) The potential tables and the value $\bar{\mu}_0$ for ρ_1 from the electrostatic field program.

(2) A fixed table of the space-charge functions η, $\xi = f(I_s/I)$ required for calculation of the voltage–current relation to parallel-plane diodes (modified Langmuir relations).

(3) A fixed table of values for the function $\rho_r = f(v_r)$ (Figure 4).

(4) The operating parameters specified for the particular case consisting of:

a) The cathode temperature, $T°K$, and work-function, ψ_c.

b) The cathode area, A, in square centimeters.

c) The plate-voltage and grid-voltage values, with specifications as to the characteristics for which plate current and derivatives (g_m, r_p, μ) are desired.

The machine calculation proceeds in the following manner (though not necessarily in the exact sequence given).

(1) The computer calculates the saturated emission current I_s from the equation $I_s = 120T^2\exp\{-(\psi_c 11600/T)\}$. The value $\psi_c = 1.6$ volts is used unless otherwise specified. (This step results in $I_s \simeq 7$ amperes per square centimeter for $T = 1100°K$; $I_s = 1$ ampere per square centimeter for $T = 1000°K$; etc.)

(2) The current I_0 and the corresponding minimum potential, v_{m_0}, at the exponential point ($dv/dx = 0$, $V = 0$) are calculated for the real grid–cathode distance, d_1, from the Langmuir relations and stored to establish the reference values ξ_0 and I_s/I_0 in the Langmuir table and the barrier potential $V_0 = (V_{m_0} + \psi_c)$ between cathode and grid plane, which is then at the reference potential, $V = 0$. (See Figure 3.)

(3) The ρ_r function (Figure 4) is multiplied by scale factors computed from Equation (5a).

(4) Starting with a low plate potential ($V_a = 0$, for example) and the first grid-potential value (e.g., $V_g = +0.6$ volt), the machine calculates 12 values of the space-potential function $v_y = f(x)$ with Equation (1) in the first elemental section of the cathode–grid space at $y = 0$ (see Figure 1). It then averages this function, extrapolates, and stores an equivalent-diode potential $V_{(y)}$ for the grid-plane distance d_1', as illustrated by Figure 5. The calculation is made with the unit potential values $v_{g(1)}$ and $v_{a(1)}$ from the potential tables for the distance ratio $\rho_1 = 1.5$. The calculation is repeated with the unit potentials for the ratio $\rho_2 = 1.8$, and a corrected equivalent-diode potential is interpolated from the two values $V_{(y)}$ and the stored function $\rho = f(V)$ (Figure 4). Having determined the equivalent-diode potential $V_{(1)}$

for the first elemental section, the machine calculates the current density $i_{(y)}$ from the stored table of space-charge functions and reference values computed in step 2. This process is repeated for the remaining 16 sections (y) of the cathode–grid space to furnish the potential and current distributions $v_a' = f(y)$ and $i = f(y)$, shown in Figure 6, for the particular electrode potentials (V_g, V_a). The total cathode current I_k is calculated by multiplying the mean value of the current-density function $i = f(y)$ by the cathode area.

(5) The calculations described in step 4 require less than 0.2 second and are repeated for all specified grid- and plate-voltage points in the tube characteristic. The total current values are stored in matrix form in the core memory. Tape storage of the corresponding distribution functions $i = f(y)$ and $V = f(y)$ for print-out is optional.

(6) The derivatives are calculated from increments in the stored current matrix ($g_m = \Delta I/\Delta V_g$, $r_p = \Delta V_a/\Delta I$, and $\mu = g_m r_p$) and stored on the output tape for print-out.

The *total machine time* for the calculation of four 140-point tables (including the time for the field calculation) is *approximately 34 seconds.* When a print-out of the potential- and current-distribution functions for each of the operating points is desired, the machine time is increased by the few extra seconds consumed in transferring the information onto the output tape. The final reduction of the taped information to typed data sheets is done by low-cost secondary equipment, as illustrated by Figures 11a to 11e. A page of the optional information showing the distribution function V, $I = f(y)$ for each operating condition is reproduced in Figure 11f.

Computed Data

The computed current (Figure 11b) is the cathode current $I_k = I_b + I_c$ given in milliamperes for the specified cathode area. (Print-out of current densities is optional.)

Computed Potentials (V_g, V_a) and External Voltages (E_c, E_b)

The electrode surface potentials V_g and V_a used in the calculations are measured from the exponential point $V = 0$, $dv/dx = 0$. The cathode has, therefore, the following positive potential:

$$V_o = \psi_c + V_{m(o)}\,, \tag{6}$$

where ψ_c is the cathode work function and $V_{m(o)}$ the minimum potential (Figure 3). The value of the potential barrier V_a is printed on the

62K80A606.69

THE INPUTS ARE AS FOLLOWS

HOT CATHODE DIAMETER = 0.06190 IN., GRID DIAMETER = 0.06699 IN., PLATE DIAMETER = 0.08000 IN.

PITCH = 0.00350 IN., WIRE DIAMTER = 0.00089 IN.

THE CATHODE AREA = 0.1650 SQ CM, CATHODE TEMPERATURE = 1100 K

THE CATHODE WORK FUNCTION = 1.60

THE COMPUTED VALUES ARE AS FOLLOWS

EMISSION CURRENT (I-S) = 6.937 AMPS/SQ CM.

I-INFINITY = 6.948 MILLI-AMPS/SQ CM., I-ZERO = 6.711 MILLI-AMPS/SQ CM.

V-ZERO = 2.25819 MU-ZERO FOR D/D IS 1.5 = 27.797 MU-ZERO FOR D/D IS 1.8 = 28.041

THE CORRECTED MU = 33.012

Fig. 11a—Title page of computer printed data matrices for a triode.

62K80A600.09

CURRENT IS GIVEN HERE IN MILLI-AMPS

BIAS	BEGINNING PLATE = -0. -0.	0.200	0.400	0.600	0.800	1.000	1.200	1.400	1.600	1.800	2.000	2.200
0.40	7.507	11.721	15.544	20.115	0.	0.	0.	0.	0.	0.	0.	0.
0.20	4.695	7.519	11.410	15.694	20.511	0.	0.	0.	0.	0.	0.	0.
0.	1.940	4.165	7.635	11.594	15.923	20.593	0.	0.	0.	0.	0.	0.
-0.20	0.130	1.625	4.429	7.906	11.857	16.252	20.951	0.	0.	0.	0.	0.
-0.40	0.	0.476	2.340	5.223	8.650	12.535	16.813	21.417	0.	0.	0.	0.
-0.60	0.	0.100	1.097	3.291	6.165	9.629	13.567	17.776	22.365	0.	0.	0.
-0.80	0.	0.	0.379	1.832	4.257	7.258	10.752	14.643	18.710	23.512	0.	0.
-1.00	0.	0.	0.099	0.903	2.736	5.293	8.324	11.925	15.654	20.159	24.752	0.
-1.20	0.	0.	0.	0.342	1.611	3.684	6.304	9.371	13.139	17.148	21.446	26.107
-1.40	0.	0.	0.	0.123	0.809	2.397	4.670	7.517	10.779	14.442	18.496	22.839
-1.60	0.	0.	0.	0.	0.327	1.426	3.403	5.751	8.637	12.018	15.775	19.875
-1.80	0.	0.	0.	0.	0.156	0.738	2.143	4.248	6.655	9.842	13.337	17.156
-2.00	0.	0.	0.	0.	0.	0.311	1.310	2.981	5.255	8.001	11.131	14.657
-2.20	0.	0.	0.	0.	0.	0.176	0.658	1.980	3.900	6.323	9.176	12.424
-2.40	0.	0.	0.	0.	0.	0.	0.304	1.227	2.769	4.886	7.429	10.391
-2.60	0.	0.	0.	0.	0.	0.	0.163	0.653	1.825	3.692	5.903	8.574
-2.80	0.	0.	0.	0.	0.	0.	0.	0.299	1.146	2.578	4.538	6.951
-3.00	0.	0.	0.	0.	0.	0.	0.	0.183	0.698	1.725	3.404	5.524
-3.20	0.	0.	0.	0.	0.	0.	0.	0.	0.296	1.072	2.391	4.269
-3.40	0.	0.	0.	0.	0.	0.	0.	0.	0.193	0.617	1.641	3.209
-3.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.296	1.023	2.271
-3.80	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.182	0.609	1.563
-4.00	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.297	0.986
-4.20	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.180	0.602
-4.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.300
-4.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.180

Fig. 11b—Currents, I_k (zeroes are printed for values not calculated).

6ZK9DA60G.89

GM IS GIVEN HERE IN MICRO-MHOS

BIAS	THE BEGINNING PLATE WAS -0.	0.200	0.400	-0. 0.600	0.800	1.000	1.200	1.400	1.600	1.800	2.000	2.200
0.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.20	15659.2	17966.1	19771.6	21391.5	0.	0.	0.	0.	0.	0.	0.	0.
0.	9785.4	14722.5	17452.5	19468.5	21134.5	0.	0.	0.	0.	0.	0.	0.
-0.20	0.	9148.2	13107.0	15929.1	18194.1	20145.3	0.	0.	0.	0.	0.	0.
-0.40	0.	3811.6	8328.5	11564.4	14178.4	16557.7	18608.1	0.	0.	0.	0.	0.
-0.60	0.	0.	5054.3	8351.9	10962.0	13191.7	15152.7	16933.9	0.	0.	0.	0.
-0.80	0.	0.	2496.0	5931.6	8624.4	10840.0	12783.4	14620.0	16277.7	0.	0.	0.
-1.00	0.	0.	0.	3848.5	6616.3	8935.3	10921.8	12679.9	14427.6	15908.1	0.	0.
-1.20	0.	0.	0.	1962.9	4816.2	7239.3	9261.1	11020.8	12686.8	14295.7	15639.9	0.
-1.40	0.	0.	0.	0.	3218.2	5644.8	7700.2	9551.9	11130.0	12776.4	14133.3	15579.9
-1.60	0.	0.	0.	0.	1633.2	4148.0	6304.6	8172.2	9610.4	11373.3	12896.9	14207.5
-1.80	0.	0.	0.	0.	0.	2786.9	4983.1	6924.1	8580.8	10091.1	11605.6	13045.9
-2.00	0.	0.	0.	0.	0.	1466.5	3699.1	5670.6	7367.6	8922.5	10404.3	11879.2
-2.20	0.	0.	0.	0.	0.	0.	2515.7	4384.8	6214.9	7788.0	9253.8	10666.0
-2.40	0.	0.	0.	0.	0.	0.	1264.1	3316.1	5135.4	6803.2	8182.2	9625.0
-2.60	0.	0.	0.	0.	0.	0.	0.	2320.6	4057.0	5769.5	7226.7	8597.9
-2.80	0.	0.	0.	0.	0.	0.	0.	1176.6	2994.0	4693.2	6247.8	7625.5
-3.00	0.	0.	0.	0.	0.	0.	0.	0.	2124.7	3766.2	5367.9	6727.7
-3.20	0.	0.	0.	0.	0.	0.	0.	0.	1112.7	2768.6	4407.5	5787.2
-3.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	1939.7	3420.3	4977.8
-3.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	1088.8	2576.4	4116.2
-3.80	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1814.3	3212.1
-4.00	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1072.7	2401.4
-4.20	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1715.2
-4.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1034.3
-4.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Fig. 11c—Transconductances, $g_{m(k)}$ (zeroes are printed for values not calculated).

62K80A6006.89

RP IS GIVEN HERE IN OHMS

BIAS	THE BEGINNING PLATE WAS -0.	0.200	0.400	0.600	0.800	1.000	1.200	1.400	1.600	1.800	2.000	2.200
0.40	0.	49.8	45.5	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.20	0.	54.7	48.9	44.9	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	62.6	53.6	48.2	44.5	0.	0.	0.	0.	0.	0.	0.
-0.20	0.	94.2	63.7	53.9	47.9	44.0	0.	0.	0.	0.	0.	0.
-0.40	0.	0.	84.3	63.9	54.7	49.0	45.0	0.	0.	0.	0.	0.
-0.60	0.	0.	125.8	78.6	63.0	54.6	49.1	45.2	0.	0.	0.	0.
-0.80	0.	0.	0.	103.1	74.4	61.6	54.2	49.0	45.1	0.	0.	0.
-1.00	0.	0.	0.	151.7	91.2	70.7	60.3	53.6	48.6	45.0	0.	0.
-1.20	0.	0.	0.	0.	119.7	83.8	67.9	59.2	52.8	48.2	44.6	0.
-1.40	0.	0.	0.	0.	175.9	103.1	78.1	65.7	57.8	51.8	47.6	0.
-1.60	0.	0.	0.	0.	0.	134.2	92.5	74.3	63.6	56.5	51.0	0.
-1.80	0.	0.	0.	0.	0.	198.8	114.0	85.3	70.9	61.7	55.1	0.
-2.00	0.	0.	0.	0.	0.	0.	149.8	101.4	79.7	68.1	60.1	0.
-2.20	0.	0.	0.	0.	0.	0.	221.7	124.6	92.1	75.8	65.6	0.
-2.40	0.	0.	0.	0.	0.	0.	0.	162.3	109.3	85.3	72.7	0.
-2.60	0.	0.	0.	0.	0.	0.	0.	243.5	135.7	98.1	80.5	0.
-2.80	0.	0.	0.	0.	0.	0.	0.	0.	175.5	117.9	91.5	0.
-3.00	0.	0.	0.	0.	0.	0.	0.	0.	259.4	144.1	105.3	0.
-3.20	0.	0.	0.	0.	0.	0.	0.	0.	0.	190.9	125.5	0.
-3.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	274.4	154.3	0.
-3.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	202.5	0.
-3.80	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	289.7	0.
-4.00	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
-4.20	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
-4.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
-4.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Fig. 11d—Plate resistances, $r_p/\bar{\mu}_0$ (zeroes are printed for values not calculated).

62K80A60G.89

AMPLIFICATION FACTOR OR MU

BEGINNING PLATE = -0.

BIAS	-0.	0.200	0.400	0.600	0.800	1.000	1.200	1.400	1.600	1.800	2.000	2.200
0.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.20	0.	1.0	1.0	1.0	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.9	0.9	0.9	0.9	0.	0.	0.	0.	0.	0.	0.
-0.20	0.	0.9	0.8	0.9	0.9	0.9	0.	0.	0.	0.	0.	0.
-0.40	0.	0.	0.7	0.7	0.8	0.8	0.8	0.	0.	0.	0.	0.
-0.60	0.	0.	0.6	0.7	0.7	0.7	0.7	0.8	0.	0.	0.	0.
-0.80	0.	0.	0.	0.6	0.6	0.7	0.7	0.7	0.7	0.	0.	0.
-1.00	0.	0.	0.	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.	0.
-1.20	0.	0.	0.	0.	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.
-1.40	0.	0.	0.	0.	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.
-1.60	0.	0.	0.	0.	0.	0.6	0.6	0.6	0.6	0.6	0.7	0.
-1.80	0.	0.	0.	0.	0.	0.6	0.6	0.6	0.6	0.6	0.6	0.
-2.00	0.	0.	0.	0.	0.	0.	0.6	0.6	0.6	0.6	0.6	0.
-2.20	0.	0.	0.	0.	0.	0.	0.6	0.5	0.6	0.6	0.6	0.
-2.40	0.	0.	0.	0.	0.	0.	0.	0.5	0.6	0.6	0.6	0.
-2.60	0.	0.	0.	0.	0.	0.	0.	0.6	0.6	0.6	0.6	0.
-2.80	0.	0.	0.	0.	0.	0.	0.	0.	0.5	0.6	0.6	0.
-3.00	0.	0.	0.	0.	0.	0.	0.	0.	0.6	0.5	0.6	0.
-3.20	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.5	0.6	0.
-3.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.5	0.5	0.
-3.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.5	0.
-3.80	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.5	0.
-4.00	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
-4.20	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
-4.40	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
-4.60	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Fig. 11e—Amplification factors, $\mu/\bar{\mu}_0$ (zeroes are printed for values not calculated).

V =	0	1	2	3	4	5	6	7
	8	9	10	11	12	13	14	15	16

62K80A566.89

GRID VOLTS = -3.20, PLATE VOLTS = 50.03 CURRENT PER SQUARE CENTIMETER = 5.7943 MILLI-AMPS.

V =	0.288	0.269	0.210	0.111	-0.003	-0.124	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	
I =	31.346	29.465	23.977	15.332	6.535	1.726	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	

GRID VOLTS = -3.40, PLATE VOLTS = 50.03 CURRENT PER SQUARE CENTIMETER = 3.5368 MILLI-AMPS.

V =	0.163	0.143	0.093	-0.001	-0.013	0.	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	
I =	19.772	18.035	13.029	6.614	1.825	0.	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	

GRID VOLTS = -3.60, PLATE VOLTS = 50.03 CURRENT PER SQUARE CENTIMETER = 1.6004 MILLI-AMPS.

V =	0.040	0.020	-0.003	-0.011	-0.468	0.	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	
I =	9.668	8.140	6.505	5.958	0.138	0.	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	

GRID VOLTS = -3.80, PLATE VOLTS = 50.03 CURRENT PER SQUARE CENTIMETER = 0.9827 MILLI-AMPS.

V =	-0.005	-0.007	-0.011	-0.280	0.	0.	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	
I =	6.342	6.246	5.955	0.351	0.	0.	0.	0.	
	0.	0.	0.	0.	0.	0.	0.	0.	

GRID VOLTS = -1.00, PLATE VOLTS = 55.59 CURRENT PER SQUARE CENTIMETER = 133.7954 MILLI-AMPS.

V =	2.236	2.216	2.157	2.060	1.925	1.755	1.551	1.317	
	1.059	0.780	0.487	0.179	0.	0.	0.	0.	
I =	306.461	303.018	292.896	276.311	253.789	226.051	194.072	159.126	
	122.729	86.311	52.858	21.157	0.	0.	0.	0.	

GRID VOLTS = -1.20, PLATE VOLTS = 55.59 CURRENT PER SQUARE CENTIMETER = 115.9254 MILLI-AMPS.

V =	2.079	2.059	1.999	1.900	1.762	1.588	1.379	1.142	
	0.876	0.586	0.280	-0.127	0.	0.	0.	0.	
I =	279.587	275.173	266.028	249.554	227.206	199.745	168.180	134.032	
	98.457	63.291	30.589	1.755	0.	0.	0.	0.	

GRID VOLTS = -1.40, PLATE VOLTS = 55.59 CURRENT PER SQUARE CENTIMETER = 99.9794 MILLI-AMPS.

V =	1.923	1.903	1.841	1.740	1.599	1.421	1.209	0.966	
	0.695	0.395	0.081	0.	0.	0.	0.	0.	
I =	253.431	250.016	240.003	223.643	201.504	174.586	143.583	110.195	
	75.668	42.193	11.784	0.	0.	0.	0.	0.	

Fig. 11f—Detail print-out of potential and current distribution $(V, I) = f(y)$ for a number of operating conditions.

title sheet (Figure 11a) for each particular case because it is a function of cathode temperature and distance (d_1) and must be known to establish the externally applied voltages E_c and E_b; these voltages are given by

$$E_c = V_g + (\psi_g - V_o), \tag{7}$$

and

$$E_b = V_a + (\psi_a - V_o), \tag{8a}$$

where ψ_g and ψ_a are the average work functions of the grid and anode surfaces, respectively. The plate potential V_a may be printed out in actual volts or in normalized units $V_{a(n)} = V_a/\bar{\mu}_0$ (optional). For normalized units, the voltage E_b is given by

$$E_b = \bar{\mu}_0 V_{a(n)} + (\psi_a - V_o). \tag{8b}$$

The work function ψ_g of the grid cannot be predicted accurately because it can be changed as much as 1 volt by processing and aging schedules. However, it is generally in the order of 2.25 ± 0.5 volt. The grid voltage for the example (Figure 11a), therefore, is $E_c = V_g + 0.06$ ± 0.5 volt. The additive factor for a given sample tube can be determined accurately by comparison of calculated and measured data, as will be discussed later. It may also be determined with fair accuracy by the difference $V_g - E_c$ from a measurement of E_c at a small grid-current value, as follows.

The grid potential V_g is negative for small grid currents. Because of the negative potential, the flux lines are divergent under the grid wires and the grid collects current from a cathode area smaller than the shadow area A_g of the grid. If this fraction is assumed to be $K = 1/3$ for $E_b = \bar{\mu}_0$ volt, the grid-current density may be approximated by

$$I_g = \frac{I_c}{KA_g} \simeq 3\frac{I_c}{A_g}. \tag{9}$$

For a grid current $I_c = 0.5$ microampere and a shadow area $A_g = 0.3A_k = 0.3 \times 0.185$ square centimeter, for example, the grid-current density is $I_g = 27 \times 10^{-6}$ ampere per square centimeter. The negative potential with respect to the exponential point ($V = 0$) is given by

$$V_g = -\frac{2.3T}{11600} \log \frac{I_o}{I_g}. \tag{10}$$

Fig. 13—Amplification factor of triodes.

For the above example and the values given on the data page, Figure

13 for the desired plate–grid distance, d_2, grid-wire diameter $2R$, and grid–pitch (1/T.P.I.). Calculate the factor

$$\alpha = 1 + \frac{1}{\bar{\mu}_0}\left(1 + 1.5\,\frac{d_2}{d_1}\right) \tag{11}$$

The required plate potential for the current $I_{b_{max}}$ at zero bias is

$$\left.\begin{aligned} V_{a(n)} &\simeq \alpha V \quad \text{(normalized)} \\ V_a &\simeq \alpha V \bar{\mu}_0 \text{ (in volts)} \end{aligned}\right\} \tag{12}$$

The cutoff bias at this plate potential is

$$V_{g(co)} \geqq -(\alpha V + 0.4). \tag{13}$$

Suitable intervals are $\Delta V_{a(n)} = V_{a(n)}/k$, with $k = 5$ and $\Delta V_g = V_{g(co)}/k$, where k may have a value between 5 and 15.

Example 1: $I_{b_{max}} = 100$ ma/cm^2, $d_1 = 0.0025$ inch, $d_2 = 0.007$ inch, $\bar{\mu}_0 = 70$. From Figure 12,

$$V \simeq 0.88 \text{ volt}$$

$$\alpha = 1 + \frac{1}{70}(1 + 4.2) = 1.074.$$

Hence $V_{a(n)} \simeq 0.94$ and $V_{g(co)} \geqq -1.34$ volts,

and therefore $\Delta V_{a(n)} \simeq 0.2$ and $\Delta V_g = 0.2$ volt.

The plate-voltage range obtainable on one printed page is 11 times $\Delta V_{a(n)} = 2.2\ \bar{\mu}_0$ volts = 154 volts. Shorter or larger ranges may be selected.

Example 2: $I_{b_{max}} = 600$ ma/cm^2, $d_1 = 0.005$ inch, $d_2 = 0.010$ inch, $\bar{\mu}_0 = 3$. From Figure 12,

$$V \simeq 10.2 \text{ volts}$$

$$\alpha = 1 + \frac{1}{3}(1 + 3) = 2.33.$$

Hence $V_{a(n)} \simeq 24$ and $V_{g(co)} \geqq -24.4$ volts.

For $k = 6$, $\Delta V_{a(n)} \simeq 4$ and $\Delta V_g \simeq 4$ volts.

The plate-voltage range printed on one page extends to 11 times $\Delta V_{a(n)} = 44\ \bar{\mu}_0$ volts = 132 volts.

Starting Values for V_a and V_g and Current Cutoff

A normal starting value for plate voltage is $V_a = 0$. The starting voltage, however, can be given any desired value. The starting point for the grid voltage can also be given any value, i.e., V_g may start with a positive value, at zero, or with a negative value. The grid bias is automatically increased by the specified increment, ΔV_g, until the current has decreased to the cutoff value. Although cutoff is usually set to $I = 1.5$ milliamperes per square centimeter, any other value may be specified.

It is obvious that the machine time is decreased substantially by eliminating the calculation of currents beyond a specified maximum value. Therefore, the program contains an instruction to subtract the value $\Delta V_g = \Delta V_a/\bar{\mu}_0$ from the starting grid-bias value (V_g) in a new plate-potential column when the first current value in the preceding plate-potential column exceeds the specified maximum. The starting bias V_g is thereby reduced progressively, as shown in Figure 11b.

Alternate Paths and Subroutines

Practical triodes depart more or less in their geometry from a perfect parallel-plane structure either by design (e.g., cylindrical structures or remote-cutoff designs) or because of unavoidable dimensional variations (eccentricity). It is therefore desirable to provide alternate paths in the flow of calculations containing subroutines for a conformal transformation of cylindrical structures to equivalent parallel-plane structures and subroutines for the additional segmentation and current addition required in the calculation of nonuniform structures varying in the z-dimension. The calculation of segmented structures requires a routine for establishing a common reference potential ($V = 0$) when the exponential points of the segments have different values. This routine is also needed to obtain a direct-reading comparison of currents and derivatives in consecutive cases in which the cathode–grid spacing and/or the cathode temperature are not alike.

Subroutine for a Direct Comparison of Computed Characteristics

When a number of triode characteristics are computed in sequence to explore variation of one or several tube parameters, a variation of grid–cathode spacing, d_1, or cathode temperature, T, causes a change of the potential barrier $V_0 = \psi_c + V_{m(0)}$ at the exponential point and thus of the reference potential $V = 0$ used in the calculations. A direct-reading comparison of currents and g_m values requires that the first one of several cases be made the reference case and that the grid and

plate potentials $(V_g, V_a)_n$ of all other cases, n, be shifted by an increment $\Delta V_0 = V_{0(n)} - V_{0(1)}$ to obtain operating points corresponding to equal external voltages (E_c, E_b). Such operating points are given by

$$\left.\begin{array}{l} V_g,\ V_a \text{ for the reference of "base" case,} \\ \text{and} \\ (V_g + \Delta V_0),\ (V_a + \Delta V_0) \text{ for all other cases, } n, \\ \text{with} \\ \Delta V_0 = V_{0(n)} - V_{0(\text{base})}. \end{array}\right\} \quad (14)$$

When instructed to "compare" a series of designs by crossing out the word "base" in the "comparison option" space (shown in Figure 10), in all but the first data forms the computer will automatically make the "base" design the reference case (its V_0 value) and will shift the grid-bias and plate-potential values of consecutive designs (n) according to the relationships given in (14). The operating points (V_a, V_g) specified for the voltage–current characteristics of a comparison series should, therefore, be alike or integral multiples, and the plate potentials must be requested in volts (not normalized) to accommodate changes in $\bar{\mu}_0$ values.

The actual values of V_0 are given on the title sheets for the different cases, and the currents and derivatives printed in corresponding V_g rows and V_a columns in the data tables refer to identical external voltages (E_c, E_b), although the printed grid- and plate-potential values are slightly different according to (14).

Subroutines for Triodes Having Cylindrical Elements

The elements of a coaxial structure have progressively different surface areas. For a given spacing of two elements, the capacitance of parallel-plane elements of equal area is smaller or larger than the capacitance of cylindrical elements, depending on whether the area A_1 of the inner cylinder (diameter D_1) or the area A_2 of the outer cylinder (diameter D_2) is used for reference. The capacitance ratios are given by

$$\frac{C_{\text{par}}}{C_{\text{coax}}} = \frac{D_1 \ln (D_2/D_1)}{(D_2 - D_1)}, \qquad A_1 = \text{reference area}, \quad (15)$$

$$\frac{C_{\text{par}}}{C_{\text{coax}}} = \frac{D_2 \ln (D_2/D_1)}{(D_2 - D_1)}, \qquad A_2 = \text{reference area}. \quad (16)$$

For electrostatic-field calculations, a cylindrical triode can be replaced by a parallel-plane triode model having a control grid of equal area, grid pitch (P_0), and wire diameter ($2R_0$), and equal grid–cathode and grid–plate capacitances. The spacings d_1 and d_2 of the equivalent parallel-plane model are therefore obtained by letting the capacitance ratios (Equation (15) or (16)) equal unity.

The grid–cathode distance of the equivalent parallel-plane model used for calculations of the potential functions of a coaxial triode having an internal cathode is thus obtained from Equation (16):

$$d_1 = \left(\frac{D_g}{2}\right) \ln \frac{D_g}{D_k}. \tag{17}$$

The grid–anode distance is obtained from Equation (15):

$$d_2 = \left(\frac{D_g}{2}\right) \ln \frac{D_p}{D_g}. \tag{18}$$

The computer performs this transformation when the data card (Figure 10) contains three values (cathode diameter, grid diameter, and plate diameter) specifying a coaxial structure instead of two values (cathode–grid space, d_1, and grid–plate space, d_2), indicating a parallel-plane structure.

The electrostatic values $\bar{\mu}_0$ and the unit potential functions for the equivalent distance ratios $\rho = d_1/d_1'$ are calculated by the computer from the transformed model dimensions. The calculation of currents, however, requires the numerical value of the actual or equivalent grid–cathode distance. The distance d_1 of the transformed model (Equation (17)) must therefore be scaled to duplicate the flux density obtained at the location of the equivalent cathode in the space-charge-free coaxial triode model.

It is evident from Figure 14 that the width Δy and flux density of a section (y) in a parallel-plane diode are constant for a given flux value at all distances (x) from cathode to grid, whereas the flux density in a corresponding pie-shaped section of a coaxial diode is inversely proportional to the radius length, having its lowest value on the mean grid diameter D_g used as reference surface for the electrostatic model. The equivalent cathode in the parallel-plane model is located at the distance $d' = d_1/\rho$. It occurs in the coaxial model on a cylindrical surface having the diameter D_ρ which must be made the reference

surface. The scaling factor derived from the logarithmic relations is given by

$$\frac{D_\rho}{D_g} = \left(\frac{D_k}{D_g}\right)^{1/\rho}.$$

The spacing of the equivalent diode for current calculations is, therefore,

$$d_{1(\mathrm{equ})} = d_1 \left(\frac{D_k}{D_g}\right)^{1/\rho}. \tag{19}$$

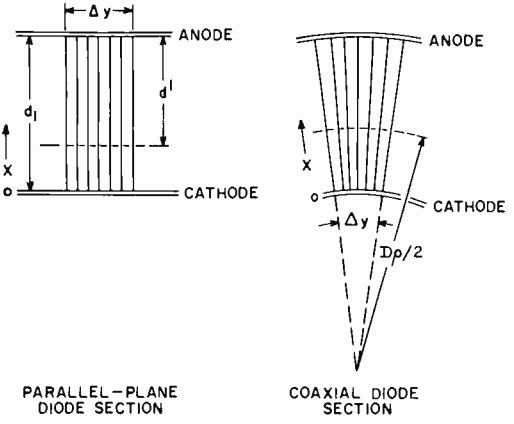


Fig. 14—Electrostatic flux lines in parallel-plane and coaxial diodes.

The computer is instructed to calculate $d_{1(\mathrm{equ})}$ from Equation (19) for the fixed ratio $\rho = 1.5$, although $d_{1(\mathrm{equ})}$ is actually a function of V in coaxial triodes. The error in the computed current resulting from this simplification does not exceed ±5 per cent for $(D_k/D_g) = 0.8$ and increases to ±10 per cent for $(D_k/D_g) = 0.65$. The correction for variable ρ values has been omitted in the present program because it is a second-order effect, affecting only a fraction of the current function, and most modern tubes have ratios $(D_k/D_g) \geq 0.8$. Inclusion of the variation $d_{1(\mathrm{equ})} = f(\rho)$ would result in a large increase in machine time.

Correction for Mesh Grids

A mesh grid can be regarded as two crossed parallel-wire grids located in the same plane (true mesh) or behind one another. Replace-

ment of the mesh-grid triode by an equivalent triode having a parallel-wire grid requires that the electrostatic values $\bar{\mu}_0$ be identical and that the potential field in the grid–cathode space be equivalent. A true mesh having the pitch values P_0 and P_1 may be replaced by a parallel-wire grid of equal wire diameter (2R) having a finer pitch P_{11} which can

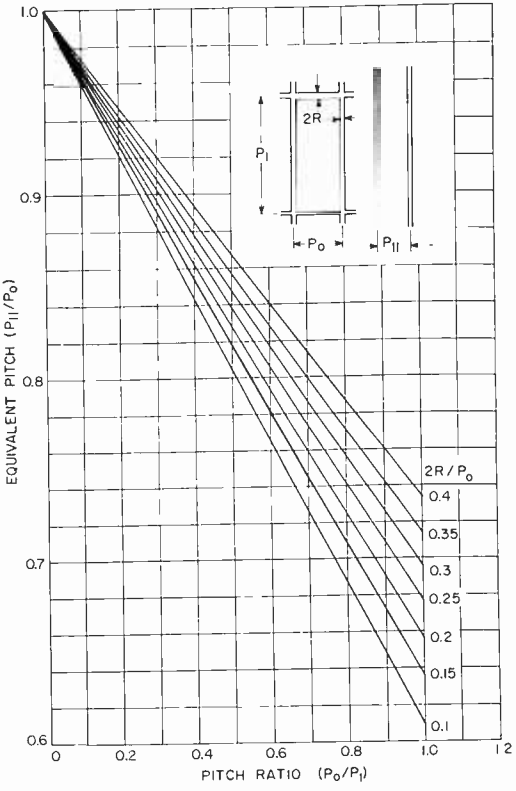


Fig. 15—Equivalent parallel-wire pitch (P_{11}) of true mesh grids.

be determined from Figure 15.* When the cross wires (P_1) are not in the same plane but are located on the plate side of the finer-pitch control wires (P_0) (e.g., the helix winding of a Nuvistor grid which is wound over the parallel cage-grid wires), their main effect is a decrease of the unit plate-field intensity as obtained with a larger grid–anode distance.

* Based on data given in, "Electrolytic Tank Measurements of Mesh-Grid Characteristics," by H. Hsu and C. E. Horton, *I.R.E. Convention Record*, Part 3, p. 114, 1957.

This type of mesh grid, therefore, can not be replaced by an equivalent parallel-wire grid of finer pitch, which would result in higher g_m values and a sharper cutoff, but it can be replaced by a μ correction factor and a parallel-wire grid having the actual pitch P_0 of the control wires. The reciprocal $\bar{\mu}$ correction factor or "D factor" (Durchgriff factor) is shown in Figure 16 as a function of the pitch ratio P_0/P_1 and the ratio $2R/P_0$ of wire diameter to pitch. (A mean value may be used when the wire diameters of P_0 and P_1 are not alike.) The computer uses the corrected value $\bar{\mu}_0/D$ in potential calculations when a D factor is specified in card 2 (Figure 10), and prints out both the uncorrected and the corrected $\bar{\mu}_0$ values, as shown in Figure 11a.

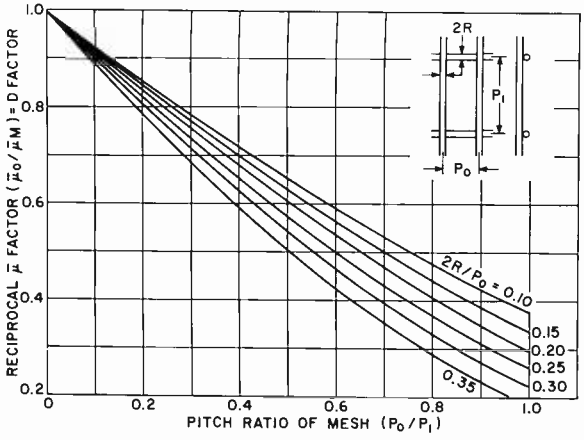


Fig. 16—D-Factor (reciprocal μ correction factor) of mesh grids.

Nonuniform Spacings or Parameters (Segmentation Routine)

The potential fields in structures having nonuniform spacings between elements are not uniform in the z dimension as in parallel-plane or coaxial structures; i.e., $v = f(x,y,z)$. Small variations Δd can be absorbed by the use of average spacing values; for larger variations, however, it is necessary to extend the segmentation of the grid–cathode field to the z dimension. A simple case is a parallel-plane structure having different element spacings on opposite sides of a planar cathode. It is obvious that the solution is obtained by two separate triode calculations (for the two halves) and subsequent addition of currents. A similar case is that of a cylindrical structure in which one of the elements is not coaxial.

Assume, for example, a cylindrical triode in which the grid axis

has a parallel displacement Δd. The sinusoidal spacing variations $\Delta s \sin \phi$ around the tube axis (z dimension) can be replaced by a stepwise variation. The step levels may be arithmetic mean values of the spacings ($s + \Delta s \sin \phi$) in these segments because the increments of s are generally small compared to s. The tube segments have equal cathode and plate diameters but different cathode areas and grid diameters.

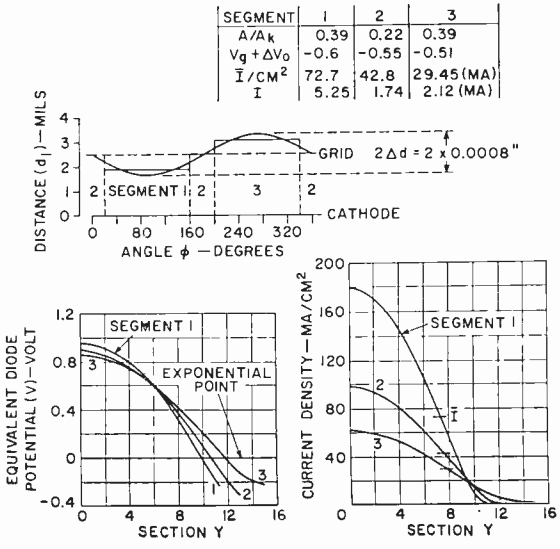


Fig. 17—Spacing variation (top), potential functions $V = f(y)$, and current distributions $I = f(y)$ of an eccentric triode.

The voltage–current characteristics of the triode segments are computed in the same manner as characteristics for direct comparison, so that the grid- and plate-potential shifts ΔV_0 caused by changes of the grid–cathode distance can be included; the computer can add currents from the segments directly into one current matrix when instructed that successive specifications are segments of one case (e.g., Figure 10, card 2, box labeled "segmented job").

Although three segments are usually sufficient for a good approximation by discrete steps, the program can accommodate as many segments as desired.

The optional print-out of current-distribution functions for all segments is available for a detailed study of the currents contributed to the total sum by the segments. Figure 17 shows the potential and current-density distributions obtained by a three-segment calculation

of an eccentric "coaxial" triode for one operating condition ($V_a = 80$ volts, $V_g = -0.55$ volt). A parallel displacement of the grid axis by $\Delta d = 0.0008$ inch is assumed; this displacement represents a 32-per cent eccentricity in grid–cathode spacing. The current densities in segment 2 are the same as obtained with a coaxial structure for which a total average plate current $I_b = 7.92$ milliamperes and a transconductance $g_m = 12560$ micromhos are computed. The eccentricity causes a 15-per cent increase in plate current ($I_b = 9.115$ milliamperes) and an 8.6-per cent increase in transconductance ($g_m = 13{,}665$ micromhos) because of the closer spacing and higher current densities in segment 1 (see Figure 17).

It is apparent that the segmentation routine can be applied to compute the characteristics of structures having several grid-pitch values (remote-cutoff tubes) or other dimensional variations, as well as emitter nonuniformities, such as nonuniform work functions or temperatures.

Correlation of Computed and Measured Data

Measurement of Voltage–Current Characteristics

The parameters of actual tubes are generally not a precise duplicate of the mathematical model. Cathode temperature, area, electrode spacings, and surface work functions may not be uniform or may deviate from assumed values, and the actual grid may not have exactly the electrostatic field and value $\bar{\mu}_0$ computed for an equivalent parallel-wire grid. The cathode temperature, for example, is a function of cathode current because of the heater power loss $P_c = V_0 I_k + 2kT/e$ caused by electron evaporation. A constant cathode temperature requires, therefore, an adjustment of heater power according to current; otherwise, the rate of change in current during a measurement must be faster than the thermal time constant of the cathode surface. The dynamic measurement of voltage–current characteristics with a cathode-ray curve tracer and camera equipped with optical graticule projection (eliminating parallax errors) comes close to meeting this requirement, and currents and voltages can be measured with a precision of $\pm$ 1 per cent for all ranges. The operating mode, i.e., number and value of bias steps and a-c plate-voltage scan, must remain fixed during a measurement to maintain a constant average current and a substantially constant cathode temperature. The current or plate-voltage sensitivity of the amplifiers may be varied to expand the scales (over-scanning). A large number of small grid-bias steps can be provided (without slowing down the repetition rate) by simple circuit

modifications to permit interlacing of two sets of bias lines by a double exposure. To obtain 22 grid-voltage steps $\Delta E_c = 0.2$ volt, for example, the step voltage is set to $\Delta E_c = 0.4$ volt and 11 steps for one exposure. A second exposure is then made with the starting voltage shifted by —0.2 volt. A photograph of an interlaced characteristic is shown in Figure 18.

When the cathode temperature in a high-perveance tube differs by $\Delta T = \pm 50°\text{K}$ from the computed value, the potential barrier V_0 changes by $\Delta V_0 \simeq \pm 0.1$ volt. The change in current can be counteracted by an

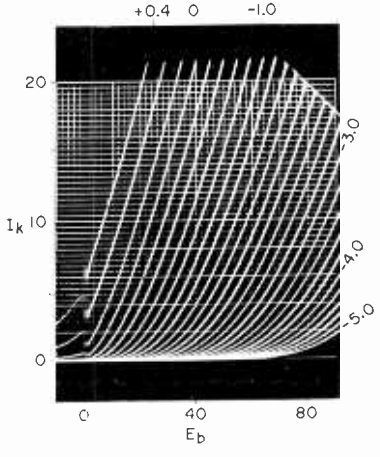


Fig. 18—Double-exposure photograph of voltage–current characteristics of a Nuvistor triode obtained by interlacing grid-bias lines.

opposite change of grid voltage $\Delta E_c \simeq \mp 0.1$ volt, but it is not completely counteracted for all current values, because of a residual difference in the voltage–current characteristic. This effect is illustrated in Figure 19 for a large temperature change $\Delta T = 200°\text{K}$ which causes a bias change $\Delta E_c = -0.4$ volt.

Analysis of Measured Characteristics

The parameters of sample tubes deviate more or less from the constant values postulated in the calculation of voltage–current characteristics. Grid and plate voltages differ from the computed potentials V_g and V_a by constants depending on processing and aging schedules. Dimensional deviations which occur during assembly or processing may change the perveance, grid-wire diameter, electrostatic value $\bar{\mu}_0$, and current scale; variations in thermal emissivity and work function

In the case where a horizontal shift does not produce an accurate match, the cathode area (A), the grid–cathode distance (d_1), or the cathode temperature may be in error. A vertical shift multiplies the current scale by some factor, thereby correcting accurately for an error in cathode area, and in good approximation for small errors in grid–cathode spacing (perveance). An error in cathode temperature is also corrected in first approximation, but a residual curvature error may remain in the grid-bias lines which follow a different power law, as shown in Figure 19. High coating or interface resistance, as well as pitch or spacing nonuniformities, can also prevent a good match of the computed bias-line curvature because they alter the power law of voltage–current characteristics. Errors of this type are indicated when a rotation of coordinates is required to match computed and measured characteristics.

Conclusions

Comparisons of computed data with values measured on carefully constructed electron tubes have shown excellent agreement of currents and derivatives over a wide range of voltages and currents. The computed current is in error in the exponential region (near current cutoff) because of the assumption of independent current flow in the y sections, as discussed in connection with Figure 9. It appears quite possible that a correction for the converging electron flow can be expressed as a simple function of the potential modulation in the grid–cathode space to provide higher accuracy in the cutoff region.

In the grid–current region (V_g more positive than approximately —0.4 volt), the constant-grid-voltage lines are shifted to slightly higher anode potentials ($\Delta V_a = 1$ to 3 volts). This error (see Figure 21) is caused by neglecting the space charge added by orbiting electrons, secondary emission, or electron reflection on the grid wires. The assumption of a uniform grid-wire potential may also cause deviations at low potentials because the work function (ψ_g) of the wires is reduced on only two-thirds of their surface (visible to the cathode) by deposits of emitter coating evaporated from the cathode, while approximately one third (facing the anode) retains the higher work function of substantially clean metal, as shown in Figure 22. The grid field is thus actually a dipole field:

$$v_g = f(x,y) = (V_g + a)\, v_{g(1)f} + (V_g - b)\, v_{g(1)b},$$

where a and b are constants specifying the difference $\Delta\psi_g$ from the

mean work function ψ_g for the unit potentials $v_{g(1)f}$ and $v_{g(1)b}$ of the front and back surfaces of the grid wires, respectively. The sum of the unit field components $v_{g(1)f} + v_{g(1)b}$ is equal to the unit field $v_{g(1)}$ for a uniform grid-wire potential ($a = b = 0$).

The "front" field of the grid wires is quite similar to the normal unipotential grid field of a slightly smaller grid wire, while the "back" field of the grid wires has the form of an anode field, as shown in Figure 22. In a future refined program, the two-unit grid fields could be computed by instructing the computer to assign zero potential to

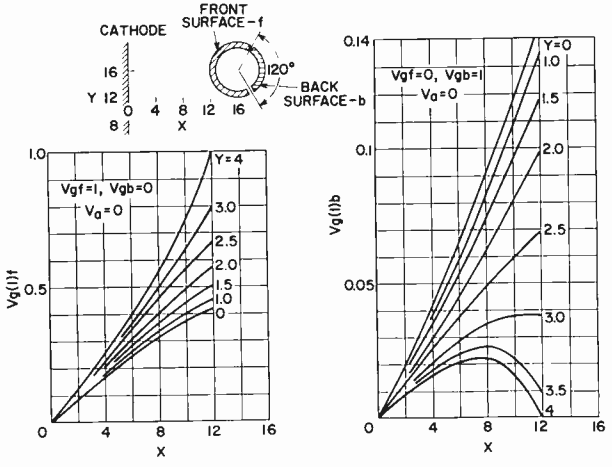


Fig. 22—Unit potential fields of a "dipole" grid in the model $X/Y/C = 4/4/28$. The sum of both fields is equal to the unit grid field given in Table I.

either four or six of the ten points on the grid-wire surfaces (see Figure 1) when solving the simultaneous equations for the line charges at the grid-wire center. The calculation of effective potentials in the plane of second and third grids may also be included for calculation of the cathode current in tetrodes and pentodes as a function of all element potentials.

It can be stated in summary that the accuracy of the machine-computed characteristics is far better than that obtained by other methods of calculation, and the experience gained thus far indicates that the machine program is a new analytic tool of great potential for the prediction, analysis, and optimization of electron-tube performance.

ACKNOWLEDGMENTS

The assembly and successful operation of the complete machine program for triode calculations is due to the able programming and painstaking efforts of Hugh Cort, Jr., of the RCA Data Systems group in Harrison. The instructions for calculation of the electrostatic potential fields of triodes and the table of space-charge relations based on the Langmuir relations were taken over from two earlier machine programs set up by W. A. Harris of Electron Tube Applications and H. W. Kirk of RCA Data Systems and Services, Electron Tube Division, Harrison. The author wishes to express his appreciation to these men from whose experience and helpful criticism he profited.

RCA TECHNICAL PAPERS†

First Quarter, 1962

Any request for copies of papers listed herein should be addressed to the publication to which credited.

"Characteristics of High-Conversion-Efficiency Gallium-Arsenide Solar Cells," A. R. Gobat, M. F. Lamorte, and G. W. McIver, *Trans. I.R.E. PGME* (January) 1962

"Measuring the Value of Information Services," J. Hillier, *Jour. Chem. Documentation* (January) 1962

"New Materials for Primary Batteries," G. S. Lozier, *Trans. I.R.E. PGME* (January) 1962

"Recent Developments in Maser Devices and Materials," H. J. Gerritsen, *Applied Optics* (January) 1962

"Relays Versus Solid-State Devices in Switching Circuits," D. B. Dobson and L. L. Wolff, *Electro-Technology* (January) 1962

"Symmetry of Transition Metal Impurity Sites in Crystals as Inferred from Optical Spectra," W. A. Weakliem and D. S. McClure, *Jour. Appl. Phys.*, Supplement (January) 1962

"Syllable Analyzer, Coder and Synthesizer for the Transmission of Speech," H. F. Olson and H. Belar, *Trans. I.R.E. PGA* (January-February) 1962

"Double Injection in Insulators," M. A. Lampert, *Phys. Rev.* January 1) 1962

"Thermal Conductivity of Ge-Si Alloys at High Temperatures," B. Abeles, D. S. Beers, G. D. Cody, and J. P. Dismukes, *Phys. Rev.* (January 1) 1962

"Circuits to Control and Protect High-Power Modulator Tubes," T. E. Yingst, *Electronics* (January 12) 1962

"New Ferroelectrics of the Tetramethylammonium-Trihalo-Mercurate Family," E. Fatuzzo, R. Nitsche, H. Roetschi, and S. Zingg, *Phys. Rev.* (January 15) 1962

"Nonlinear Optical Effects," R. Braunstein, *Phys. Rev.* (January 15) 1962

"Theory of Cyclotron Resonance Absorption by Negative-Mass Holes in Germanium," R. C. Duncan, Jr. and B. Rosenblum, *Phys. Rev.* (January 15) 1962

"Tunnel-Diode Tester," L. M. Zappulla, *Electronics* (January 26).. 1962

"Bandwidth of Hybrid-Coupled Tunnel-Diode Amplifier," R. M. Kurzrok, *Proc. I.R.E.* (February) (Correspondence) 1962

"Bibliography on Radio Frequency Interference," R. F. Ficcki, *Trans. I.R.E. PGRFI* (February) 1962

"Circuit Applications of the Field Effect Transistor," T. B. Martin, *Semiconductor Products* (February) 1962

"Heat Treatment of n-Type GaAs by Radiant Energy," T. S. Kinsel and T. E. Seidel, *Jour. Appl. Phys.* (February) (Letters to the Editor) 1962

"Majority-Logic Synthesis by Geometric Methods," H. S. Miller and and R. O. Winder, *Trans. I.R.E. PGEC* (February) (Correspondence) 1962

† Report all corrections to *RCA Review*, RCA Laboratories, Princeton, N. J.

"Medical Electronics," V. K. Zworykin, *Amer. Technion Soc.* (February) 1962
"Microwave-Carrier Modulation-Demodulation Amplifiers and Logic Circuits," W. Eckhardt and F. Sterzer, *Proc. I.R.E.* (February) 1962
"Printout System for the Automatic Recording of the Spectral Analysis of Spoken Syllables," H. F. Olson and H. Belar, *Jour. Acous. Soc. Amer.* (February) 1962
"The Stabilization of Germanium Surfaces by Ethylation, I. Chemical Treatment, II. Chemical Analysis, III. Electrical Measurements," J. A. Amick, G. W. Cullen, and D. Gerlich, *Jour. Electrochem. Soc.* (February) 1962
"A Tunnel–Diode-Tunnel–Rectifier, 15 Nanosecond Memory," M. M. Kaufman, *Solid/State/Design* (February) 1962
"Dielectric Breakdown in Cadmium Sulfide," R. Williams, *Phys. Rev.* (February 1) 1962
"Excitons at the L Absorption Edge in Zinc Blende-Type Semiconductors," M. Cardona and G. Harbeke, *Phys. Rev. Letters* (February 1) 1962
"Photoemission and Band Structure of the Semiconducting Compound CsAu," W. E. Spicer, *Phys. Rev.* (February 15) 1962
"Reflectivity of Gray Tin Single Crystals in the Fundamental Absorption Region," M. Cardona and D. L. Greenaway, *Phys. Rev.* (February 15) 1962
"Designing Tunnel-Diode Circuits Using Composite Characteristics," B. Rabinovici and J. Klapper, *Electronics* (February 16).... 1962
"Absorption Edge in Degenerate p-Type GaAs," I. Kudman and T. Seidel, *Jour. Appl. Phys.* (March) 1962
"Analysis and Synthesis of Dynamic Performance of Industrial Organizations — The Application of Feedback Control Techniques to Organizational Systems," R. B. Wilcox, *Trans. I.R.E. PGAC* (March) 1962
"Base-Layer Design for High-Frequency Transistors," H. S. Veloric, C. Fuselier, and D. Rauscher, *RCA Review* (March) 1962
"Broad-Band Parametric Amplifiers by Simple Experimental Techniques," B. B. Bossard and R. Pettai, *Proc. I.R.E.* (March) (Correspondence) 1962
"Character Recognition by Digital Feature Detection," I. H. Sublette and J. Tults, *RCA Review* (March) 1962
"Derivation of Ideal Electrode Shapes for Electrostatic Beam Focusing," W. W. Siekanowicz, *RCA Review* (March) 1962
"Design of Interstage Coupling Apertures for Narrow-Band Tunable Coaxial Band-Pass Filters," R. M. Kurzrok, *Trans. I.R.E. PGMTT* (March) (Correspondence) 1962
"Design for Product Improvement," H. E. Paschon, *Trans. I.R.E. PGPEP* (March) 1962
"Effect of Obstacle Profile on Knife-Edge Diffraction," M. P. Bachynski and Coauthor, *Trans. I.R.E. PGAP* (March) 1962
"Experimental Evidence of Landau Damping in Electron Beams," M. Caulton, B. Hershenov, and F. Paschke, *Jour. Appl. Phys.* (March) 1962
"Flux Reversal in Ferrite Cores Under the Effect of a Transverse Field," K. Li, *Jour. Appl. Phys.*, Supplement (March) 1962
"Heat Transfer in Coaxial Lines in Still Air," Z. M. Slusarek, *Trans. I.R.E. PGPEP* (March) 1962
"High Temperature Susceptibility of Garnets: Exchange Interactions in YIG and LuIG," P. J. Wojtowicz," *Jour. Appl. Phys.*, Supplement (March) 1962
"Minimum Size and Maximum Packing Density of Nonredundant Semiconductor Devices," J. T. Wallmark and S. M. Marcus, *Proc. I.R.E.* (March) 1962
"The Nature of One-ion Models of the Ferrimagnetic Anisotropy," P. J. Wojtowicz, *Jour. Appl. Phys.*, Supplement (March) 1962

"Pass-Band VSWR of Maximally Flat Band-Pass Filters," R. M. Kurzrok, *Trans. I.R.E. PGMTT* (March) (Correspondance) 1962
"Permalloy-Sheet Transfluxor-Array Memory," G. R. Briggs and J. W. Tuska, *Jour. Appl. Phys.*, Supplement (March) 1962
"Precipitates in Gallium Arsenide Single Crystals," G. Eckhard, *Jour. Appl. Phys.* (March) (Letters to the Editor) 1962
"Pulsed Radar Measurement of Backscattering from Spheres" S. B. Adler, *RCA Review* (March) 1962
"Requirements for Satellite Service," G. H. Brown, *Signal* (March) 1962
"Silicon Junction Nuclear Particle Detectors," R. L. Williams and P. P. Webb, *RCA Review* (March) 1962
"A Spectrophotometric Investigation of Dye-Sensitized Photoconductive Zinc-Oxide-Resin Layers Used in the Electrofax Process," E. C. Giaimo, *RCA Review* (March) 1962
"Stable Low-Noise Tunnel-Diode Frequency Converters," F. Sterzer and A. Presser, *RCA Review* (March) 1962
"Transistor Cutoff Frequencies," J. B. Fisher and A. Troiano, *Solid/State/Design* (March) 1962
"The Unit Cell Dimensions and Crystal Structure of $KBaPO_4$," C. W. Struck and J. G. White, *Acta Crystallographica* (March) 1962
"Value Engineering," A. D. Zappacosta, *Trans. I.R.E. PGPEP* (March) 1962
"Using Instability Characteristics of Semiconductor Plasmas," M. Glicksman, *Electronics* (March 9) 1962
"Paramagnetic Resonance of Ni^{2+} and Ni^{3+} in TiO_2," H. J. Gerritsen and E. S. Sabisky, *Phys. Rev.* (March 15) 1962

OTTO H. SCHADE was born and educated in Germany; he came to the United States in 1926. He joined the RCA Electron Tube Division at Harrison in 1931. Since 1938 he has specialized in television circuits, camera tubes, and picture tubes. From 1944 to 1957 he worked on a unified general method of image analysis and specification, including practical methods for measuring the "aperature effect" (square-wave and sine-wave response function) and fluctuation levels (noise) of optical, photographic, and electronic image-system components and the eye. For the past few years he has been responsibe for the thermal and electrical design of Nuvistor tubes. Dr. Schade has received numerous honors, including the Modern Pioneers Award of the National Association of Manufacturers (1940), the Morris Liebmann Memorial Prize of the Institute of Radio Engineers (1950) and a Fellowship (1951) from the Institute of Radio Engineers. In 1951 he was made a Fellow, and also was the first recipient of the David Sarnoff Gold Medal Award, of the Society of Motion Picture and Television Engineers. In June, 1953, he was invested with the honorary degree of Doctor of Engineering by Rensselaer Polytechnic Institute. In 1960, he received the Progress Medal Award of the Society of Motion Picture and Television Engineers for his outstanding technical contribution in the engineering phases of the motion picture and television industries.

HANOCH UR received the BS and Dipl. Ing. degrees from the Technion-Israel Institute of Technology, Haifa, in 1954 and 1955, respectively. From 1955 to 1956, he was a research assistant at the Computing Laboratory of Purdue University, Lafayette, Ind., and received the MSEE from Purdue in 1956. After a short period as a logical designer with Univac, St. Paul, Minn., he joined the Polytechnic Institute of Brooklyn, N. Y. as an instructor. He joined RCA Electronic Data Processing Division in 1959 where he has been working on tunnel-diode applications for high-speed computers. Mr. Ur is a member of Sigma Xi and the Institute of Radio Engineers.