THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XXXV | SEPTEMBER 1956 | NUMBER 5

Copyright 1956, American Telephone and Telegraph Company

Electronics in Telephone Switching Systems

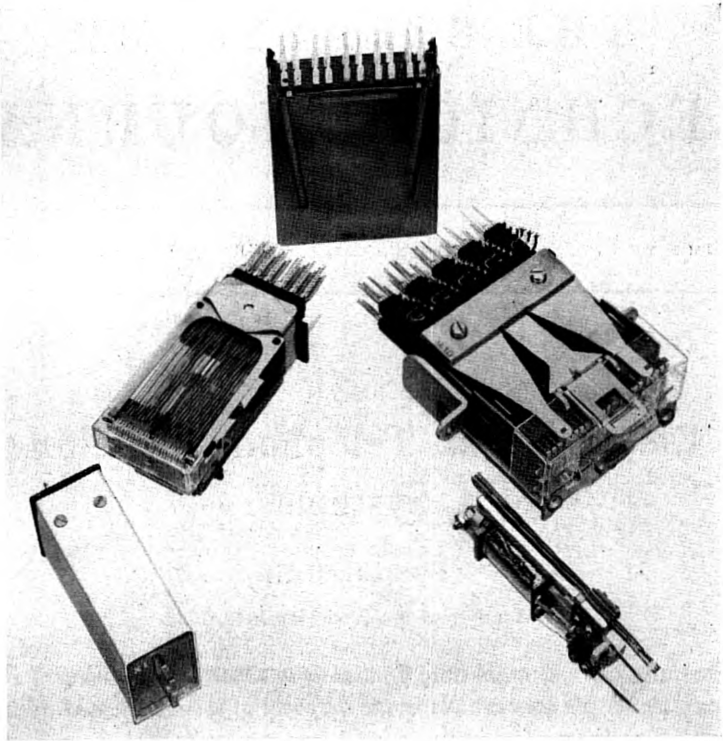
By A. E. JOEL

(Manuscript received March 18, 1956)

In recent years a number of fundamentals has been discovered through research which place new tools at the disposal of the circuit and system designers. Examples of this "new art" are concepts such as information theory, dealing with the quantization and transmission of information, and solid state principles from which have developed the transitor and other devices. This paper surveys certain new art principles, techniques and devices as they apply to the design of new telephone switching systems.

Over the past forty years a great background and fund of knowledge has developed in the field of telephone switching. Constant improvement in available devices has resulted in increasing the scope of their application. The field has almost reached a point of perfection as an art and is now rapidly entering a more scientific era.

The tools of the present day telephone system design engineer are well known and some are illustrated in Figure 1. These are the relay and the various forms of electromechanical switching apparatus. But over the years, while the art employing these tools was developing, the field of electronics has also been developing. Its applications were most needed when dealing with its characteristics of sensitivity rather



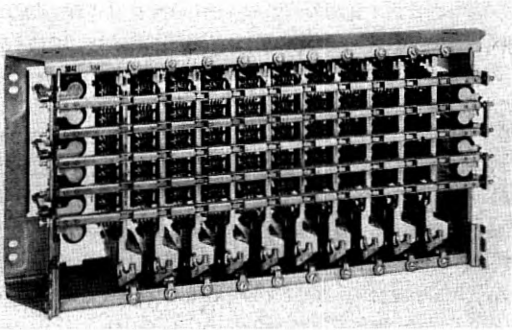


Fig. 1 — Typical telephone relays and switches.

than speed. Even in the telephone switching field, this property of electronics has made its inroads to provide us with better signaling and more accurate timing.

It was not, however, until World War II that the speed advantages of electronics were exploited. This exploitation came primarily in the quantizing of information, both in transmission and information processing equipment. In the latter field new digital computers made their appearance. These machines brought forth the development of new forms of electronic devices, most important of which are those classified as "bulk memory" devices.[1] Later in this paper the characteristics of many of these devices will be discussed in more detail.

In the post-war period the exploitation of another phase of electronics developed from research in semiconductor devices. The transistor is perhaps the best known invention to emerge from these investigations. The impact of the application of semiconductor devices is yet to be felt in the electronics industry and it will most likely find greatest application in the information processing field and in communications generally.

Before one may understand and appreciate the impact electronics will have on the design of new telephone switching systems it is necessary to consider the question: "What is a Telephone Switching System?" By evolution it is now generally recognized that the central office portion of a telephone switching system consists of two principal parts and certain physical and operational characteristics of these parts. These parts, as illustrated in Figure 2, are the interconnecting network, or conversation channel, and its control.

In some switching systems, particularly those of the progressive

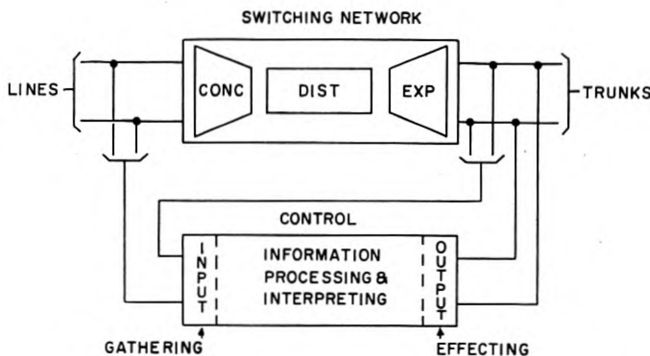


Fig. 2 — Principal parts of common control telephone switching system.

direct control type, such as the step-by-step system, these parts are inexorably integrated. But in the modern systems they have largely been separated. For purposes of the following discussion this type of system, viz., common control, will be assumed. The bulk nature of the electronic memory devices makes them more readily adaptable to systems of the common control type, where the control functions consisting of the receipt, interpretation, and processing of input signals and the effecting of output signals may be concentrated.

INTERCONNECTING NETWORK

In electromechanical switching systems the interconnecting network is composed of crossbar switches or other electromechanical devices. Each connection through the network is physically separated in space from the others and hence the type of network can be called generically a "Space Division" type of network. Such networks are subdivided functionally. First there is the concentration stage where active lines are separated from those not being called or served at a particular time. Next there is distribution stage where interconnection of active lines and trunks is accomplished. Finally there may be an expansion stage where active call paths are connected to selected destinations.

In electronic switching systems three classes of switching networks have been described.[2] These are:

a. "Space Division" similar to the space division for electromechanical apparatus except that electronic devices such as gas tubes are employed in place of mechanical contacts as the crosspoint element.[7]

b. "Time Division" where calls are sampled in time, each one being given a "time slot" on a single channel.[3,4]

c. "Frequency Division" such as employed in carrier systems where each call is modulated to a different frequency level on a single transmission medium.[5,6]

Thus in electronic switching the interconnecting networks derive their basic characteristics from the known methods of telephone transmission. Since transmission techniques are used it is generally not feasible to pass direct current signals through such networks. Also certain ac signals such as 20-cycle current now used for ringing are of such a high power level that they would overload the electronic switching devices employed. For this reason it appears that to accomplish switching with an electronic interconnecting network a change is required in the customer's apparatus to make it capable of responding to a lower level ac for the call signal. Telephone sets with transistor amplifiers and an acoustical horn are being developed. (See Fig. 3.) Interrupted

tones in the voice frequency range can be used effectively to call the user to the telephone.[8]

As in most electromechanical switching networks, the concepts of connecting successive stages of switching devices (stages to perform the functions of concentration, distribution and expansion) to form the network also apply. Since there is more than one method of interconnection, the successive stages of a network may employ different switching techniques — electronic, electromechanical, or both. In electromechanical switching, different devices may also be used in different stages.

In electromechanical space division networks certain types of crosspoints are more adapted to common control operation than others. Systems with electromechanical selector switches most generally are set progressively. In systems with relays or relay-like crosspoints all crosspoints involved in a connection may be actuated simultaneously. In either case the switching device, or the circuit in which it is used, has a form of memory. This memory, shown as a square labeled M in Fig. 4, may be the ability of a selector to remain mechanically held in a particular path connecting position or in a locking or holding circuit associated with a crosspoint relay or crossbar switch magnet.

To minimize the time consumed by the common control elements, simultaneous operation of relay or relay-like crosspoints is most desirable. However, this type of network requires a grid of link testing and control leads such as shown in Fig. 5 for a typical stage of a crossbar switching network. In a network of this type the calling rate capacity is limited by the slow actuating speed of the electromechanical relay or switch. Efficient network configurations can be devised for

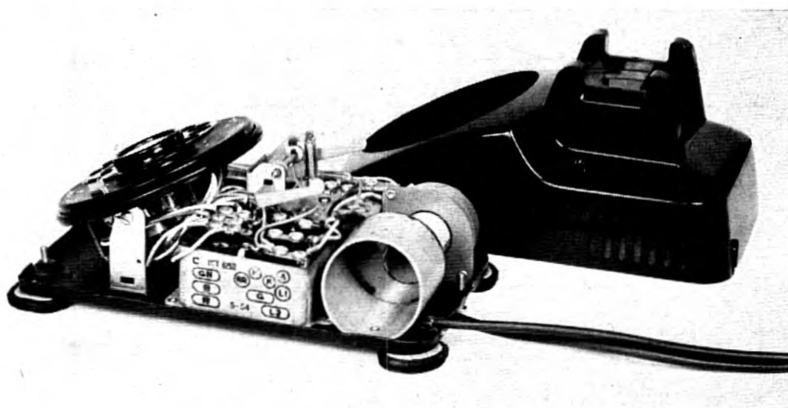


Fig. 3 — Tone ringer telephone set.

large capacity. To set up connections at a high rate in such a network requires a plurality of controls each capable of operating on all or part of the network. In any case, the controls function in parallel on the network because of the speed considerations.

With electronics applied to space division switching networks, two improvements over the operation of relay type space division networks may be achieved. First, the speed of operation of the crosspoint elements may be made high enough so that only one control is needed to operate on networks of the size now requiring a plurality of controls. Second, the properties of proposed electronic crosspoint elements are such that the principle of "end-marking" may be employed.

In contrast to the grid of testing and actuating wires required in electromechanical versions of space division networks, the electronic space division switching network requires only the selectors at each end of a desired network connection to apply the marking potentials. (This is what is meant by "end-marking"; see Fig. 6). The electronic cross-

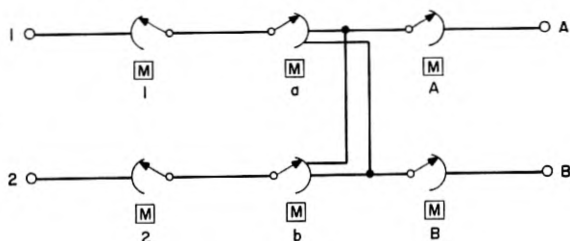


Fig. 4 — Space division switching.

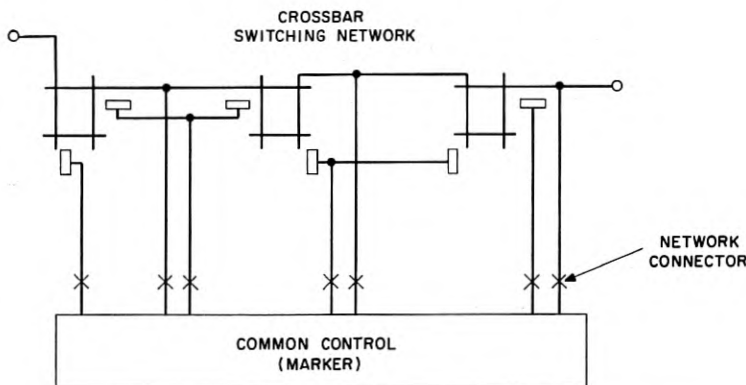


Fig. 5 — Typical common control of a crossbar switching network.

point element will be actuated if the link to which it connects is idle. Eventually all available paths between input and output will be marked. Means must be provided for sustaining only one of the possible idle paths. Here the memory property of the crosspoint device takes over to hold the path until it is released by release marks or removal of the sustaining voltages. So it may be seen that in space division networks the memory requirements must be satisfied the same as in electromechanical networks.

Multiplexing and carrier transmission systems[9] employ time and frequency division but the physical terminals at both ends of a channel for which the facilities are derived have a one-to-one correspondence which can only be changed manually. In a switching system means must be provided to change automatically the input-output relations as required for each call. Here the need arises for a changeable memory for associating a given time or frequency slot to a particular call at any given time. At some other time these points in time or frequency must be capable of being assigned automatically to different inputs and outputs. For the period that they are assigned, some form of memory must record this assignment and this memory is consulted continuously or periodically for the duration of the call.

With time division switching this new concept in the use of memory in a switching network appears most clearly, see Fig. 7(a). To associate an input with an output during a time slot the memory must be consulted which associates the particular input with the particular output. To effect the connection during a time slot the input and output must be selected A memory is consulted to operate simultaneously high speed

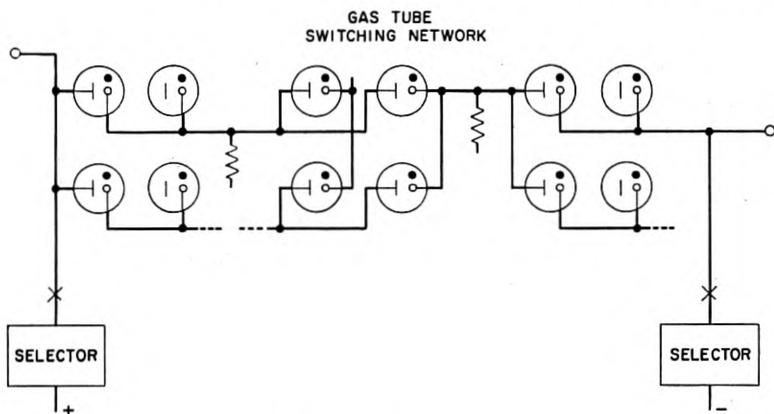


Fig. 6 — Typical "End Marking" control of a gas tube switching network.

selectors for both the input and output. Each selector receives information from a memory which actuates crosspoints to associate the input or output with the common transmission medium. The information from the memory which controls the selection process is known as an "address". The crosspoint is non-locking since it must open when the selector receives its next address. The individual memory of crosspoints for space division networks has thus been changed by time division to changeable memory, usually in the form of a coded address associated with each time slot. Furthermore since the successive addresses actuate the same selectors and hence may be held in a common high speed device, electronic bulk memory is ideally suited for this task. The memory must be changeable to allow for different associations of input to output at different times.

In frequency division the control characteristics of the interconnecting network require a modulation frequency to be assigned each simultaneous conversation to be applied within the bandwidth of the common medium. As shown in Fig. 7(b) the application of the modulation frequencies requires a separate selector for each input and output. These

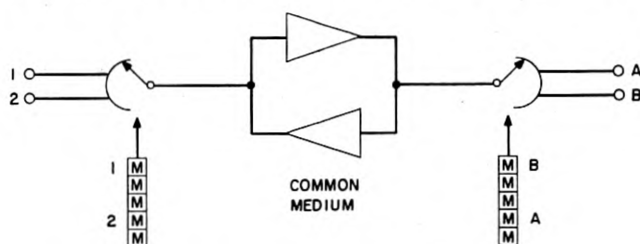


Fig. 7(a) — Time division switching.

selectors are nothing more than space division switching networks and therefore require memory in the switching devices whether they are electromechanical or electronic.

In addition to memory for associations within the switching network, selecting means are also needed to activate a terminal to be chosen in space division (e.g., Fig. 6), to place address information in the proper time slot in time division switching or to set the frequency applying switching network in frequency division.

CONTROL

The control of the switching system provides the facilities for receiving, interpreting and acting on the information placed into it. In par-

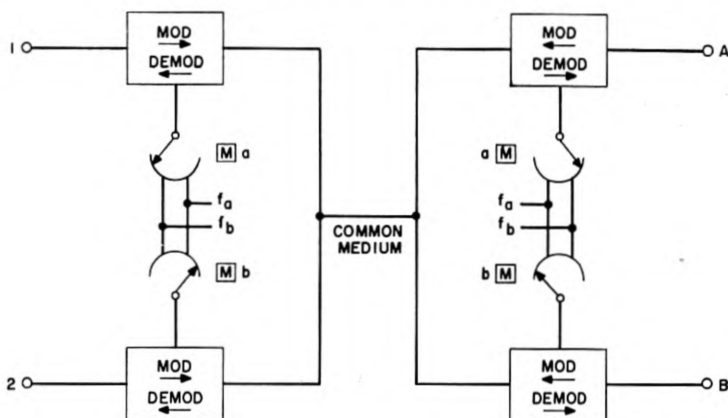


Fig. 7(b) — Frequency division switching.

ticular this is the address of the output desired. A service request detector (SR-D) is provided for each line or trunk.

In electromechanical systems these logic and information gathering functions are performed by relays or electromechanical switches. In order to keep up with the flow of information from a large number of customers, a number of register circuits must be provided to perform the same function simultaneously on different calls. Here information is being gathered on a "space division" basis and therefore a control switching network may be visualized as depicted in Fig. 8. The registers designated R-M constitute the memory used to store the input information as it is being received in a sequential manner from lines and trunks. As in the case of the conversation switching network, a space division control switching network has been used in electromechanical systems because the speed of these devices is not adequate to accommodate the rate at which information flows into the system. It is interesting to note in passing that in the step-by-step system the control and conversation switching networks are coincident. In the No. 5 crossbar system[10] the same network is used for both control and conversation on call originations but when so used the functions are not coincident, that is, the network is used for either control or conversation. In other common control systems, separate control networks known as "register or sender links" are employed.

When using relays to receive the information pulsed into the office by customers or operators a plurality of register circuits are needed. The number of the registers required is determined by the time required to actuate the calling device and for it to pulse in the information. The

registering function has two parts, one to detect or receive the information and the second to store it until a sufficient amount has been received for processing. The processing function is usually allotted to other circuits such as the markers in Crossbar systems.

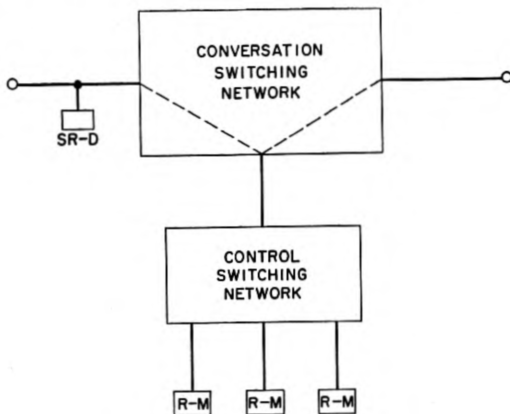


Fig. 8 — Control access.

Since the input of information to a switching system is usually limited to two conductors, a serial form of signaling is used. It would seem only natural that if a detector were fast enough it could function to receive the serial information in several simultaneously active inputs. Relays are not fast enough to do this, but high speed time sharing electronic devices have been designed to perform this information gathering function. Since it is a time sharing arrangement it is analogous to the time division switching. A time division control access as shown in Fig. 8 and 9 requires memory to control the time division switching function. Time sharing when applied to the gathering of information in telephone switching systems has been called "scanning". The individual register memories are still in parallel form because of the relatively long time required for sufficient information to be received before processing may start. Higher speed means for placing information into switching systems such as preset keysets is one way of reducing, if not eliminating, this need for parallel register storage in the switching system prior to processing. However, with this type of device one merely transfers the location of the storage from the central office to the customer's telephone set. The fundamental limitation is the rate at which a human being is able to transfer information from his brain into some physical representation.

Lower cost memory is a practical means for improving this portion

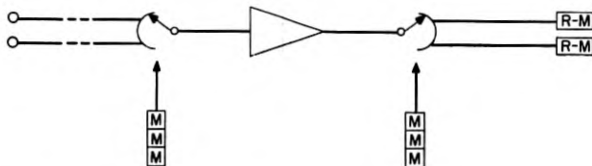


Fig. 9 — Time division control access with separate functional memory.

of the switching system. Many small low cost relay registers have been designed and placed into service.[10] Electronics, however, offers memory at one tenth, or less, of the cost per bit if used in large quantities with a common memory access control. New low cost bulk electronic memories are now available to be used in this manner. As shown in Fig. 10 the memory for the control of the time division control access network and the register memory may be combined in the bulk memory.

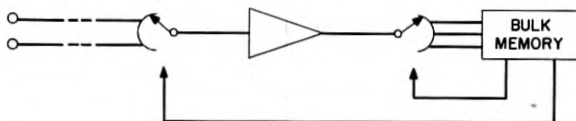


Fig. 10 — Time division control access with bulk memory.

Memory appears in the control portion of a switching system in many ways. Some are obvious and others are more subtle. Fig. 11 shows a typical electromechanical switching system, much like No. 5 crossbar and attempts to indicate various memory functions. First there is active memory designated A such as the call information storage A_2 whether in a register, sender or marker during processing. There is also certain pertinent call information storage associated with trunk circuits such as a "no charge class" on outgoing calls or the ringing code used on incoming calls. Another type of active memory A_1 has been mentioned in connection with switching networks to remember the input-output associations. In most electromechanical systems active memory has been emplemented with relays or switches.

Another form of memory is also employed in all telephone switching systems and much effort has been devoted to devising improved means for effecting this memory. This memory is of the type that is not changed with each call but is of a more permanent nature. Examples of this type of memory, which may be called passive memory, designated P, (Fig. 11), are the translations required in common control sytems to obtain certain flexibility between the assignment of lines to the switching network and their directory listing. These translations between

equipment numbers (network location) and directory numbers are required to direct incoming calls to the proper terminals (such as the number group frame in No. 5 crossbar, Fig. 12) and to provide on originating calls information for charging purposes (such as the AMA "Dimond" ring translator,[12] Fig. 13). Each of these translators for a 10,000 line office represent about 10^6 bits of information. Another use for passive memory is to translate central office codes into routing information. In local central offices this is also done by cross-connections as shown in Fig. 14.

Another form of passive memory is the punched card or tape. These have been used widely in telephone accounting systems. A step toward electronic memory is the card translator which provides routing information in the crossbar toll switching system[13] (see Fig. 15). Here the cards represent passive memory and are selected and read by a combination of electromechanical action and light beam sensing with phototransistor detectors. One such device equipped with 1,000 cards represents the storage of approximately 10^5 bits of information.

In all of the above types of passive memory limitations in the speed are involved in the choice of devices used within the memory or the access to it. This is one of the reaons these translators are subdivided so that the various portions may be used in parallel in order to satisfy the total information processing needs of the office.

A discussion of passive memory would not be complete without one further illustration, Fig. 16. This is a wiring side view of a typical relay circuit in the information processing portion of a switching system. It

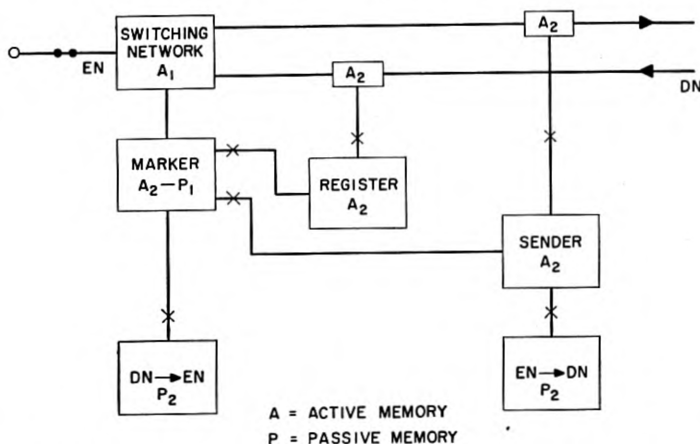


Fig. 11 — Memory in typical electromechanical switching system.

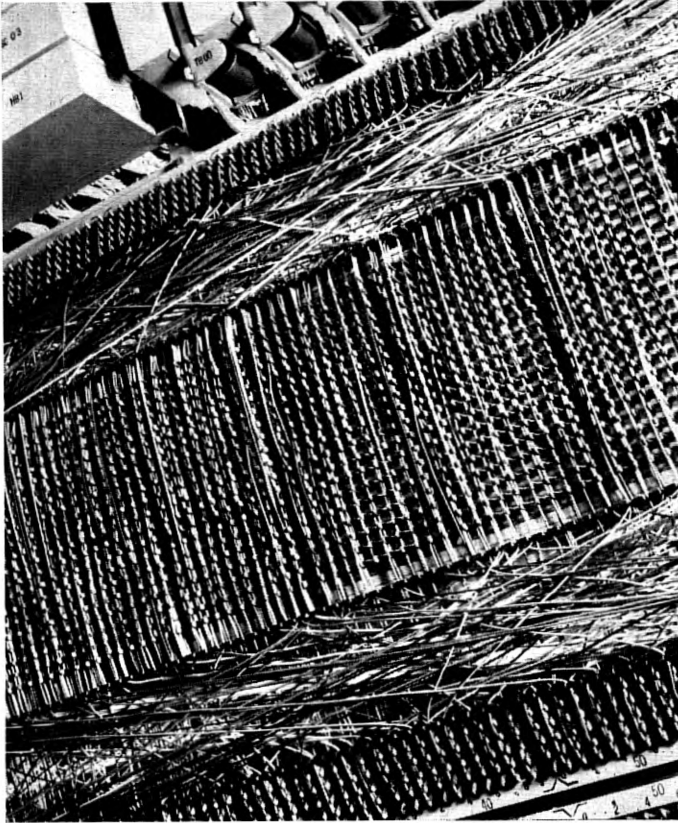


Fig. 12 — No. 5 number group.

could be any other unit, for example, a trunk circuit. The principal point is that each wire on such a unit is remembering some passive relationship between the active portions of the circuit, such as relays. This is the memory of the contact and coil interrelationships as conceived by the designer and based on the requirements of what the circuit is required to accomplish. It is the program of what the central office must do at each step of every type of call. Modern digital computers have been built with the ability to store programs in bulk memories for the solutions of the various types of problems put to them. It is conceivable that the program of a telephone contral office may also be stored in bulk memories to eliminate the need for much of the fixed wiring such as appears in relay call processing circuits.

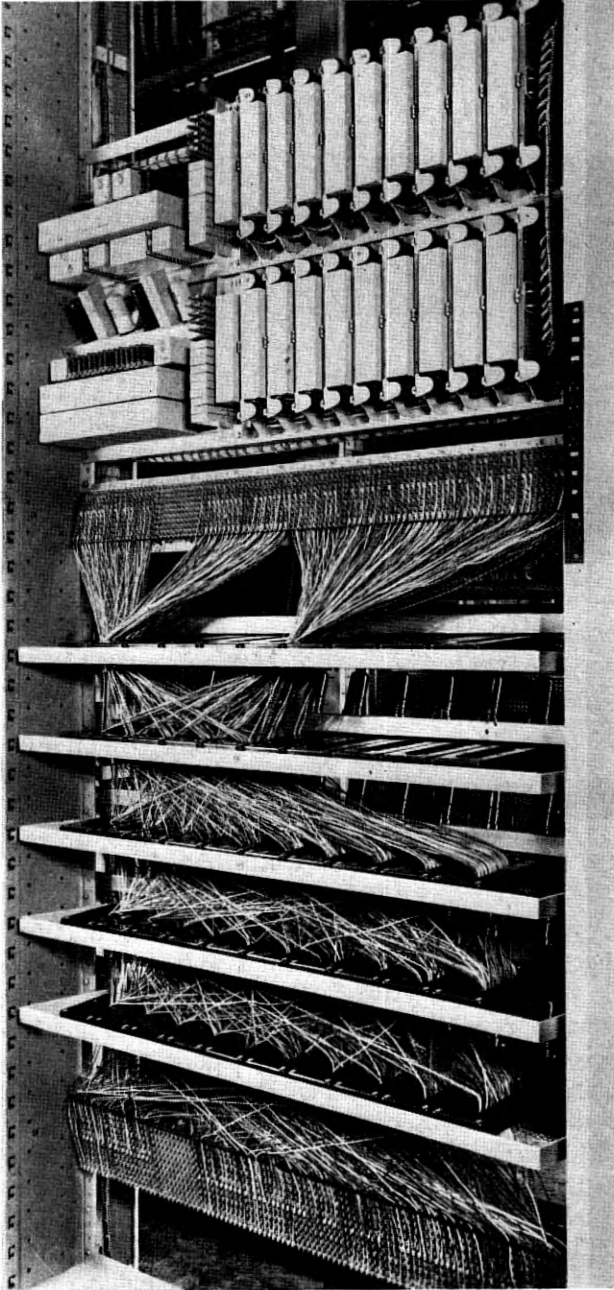


Fig. 13 — AMA translator.

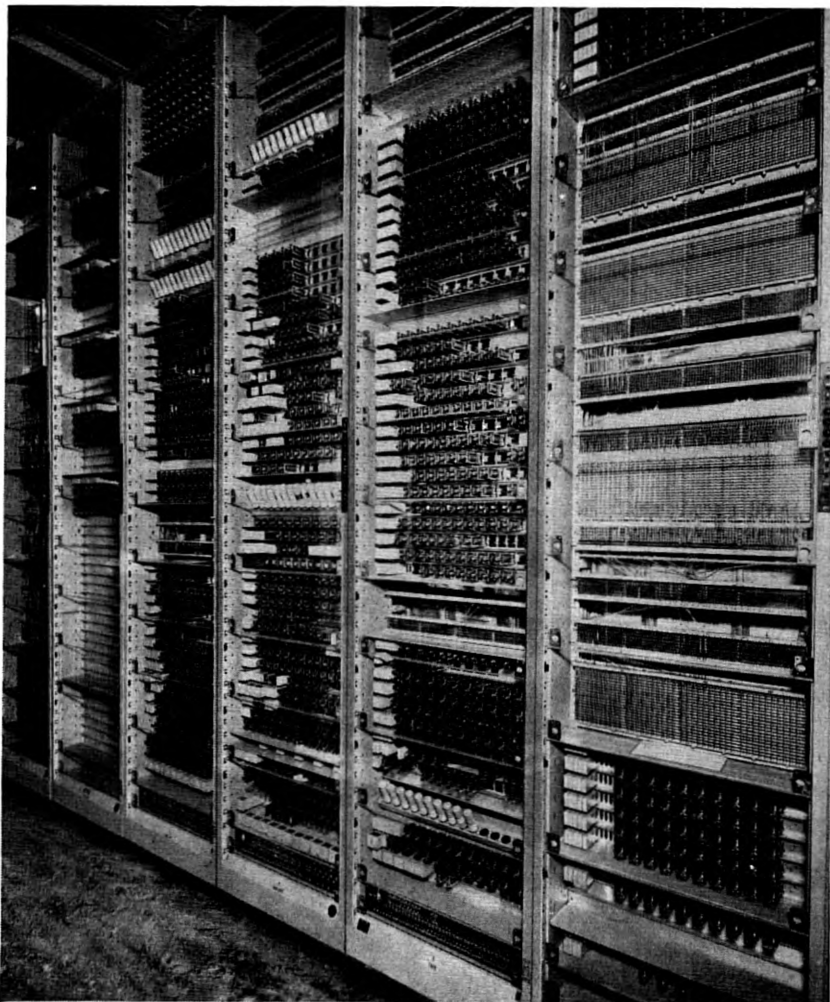


Fig. 14 — No. 5 route relay frame.

The form of memory available in electronics is considerably different from that which has been previously available. Electronic memory has been characterized as "common medium" or "bulk" memory. A single device is used capable of storing more than a single bit of information which is the limit of most relays or other devices capable of operating in a bistable manner. A number of different types of electronic bulk memories have been devised for digital processing. They differ appre-

Fig. 15 — No. 4A card translator.

ciably in physical form, each taking advantage of the phenomenon of some different area of the physical sciences — electrostatic, electromagnetic, optic. Magnetic tapes[14] and drums[15] (Fig. 17), cores[18] (Fig. 18), electrostatic storage in tubes[16, 17] (Fig. 19) and ferroelectrics[19, 20] (Fig. 20) and photographic storage[21] (Fig. 21) are available.

Several properites of these memory devices are of interest. Being electronic, the speed with which stored information may be read is of primary interest. This is known as "access speed". Another property of these common medium memory systems or devices is the ability to change what has been written. If the changes can be made rapidly enough they may be used in electronic systems in much the same manner as relays are used in electromechanical systems to process information. If the change must be made relatively infrequently, such as changing photographic plates, they may be used as substitutes for the type of memory in these systems which are provided by cross connections and wiring. The required fixed or semipermanent electronic memory may be characterized primarily by a high reading speed, large capacity, and the ability to hold stored information even during pro-

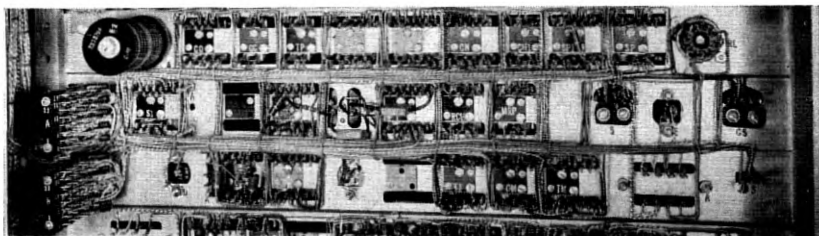


Fig. 16 — Wiring side of relay unit.

longed intervals of loss of power. The amount of memory is measured in terms of binary digits or "bits". The number of bits equivalent to single cross connection can be rather large. Therefore, electronic memory replacing fixed memory such as in the card translator in modern electromechanical systems should be high in bit capacity, from 10^5 to 10^7 bits for 10,000 lines.

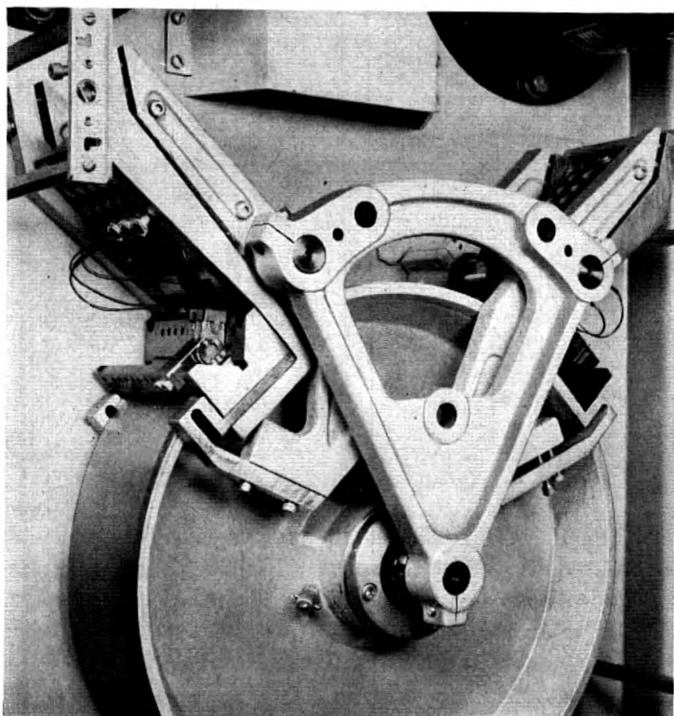


Fig. 17 — Magnetic drum.

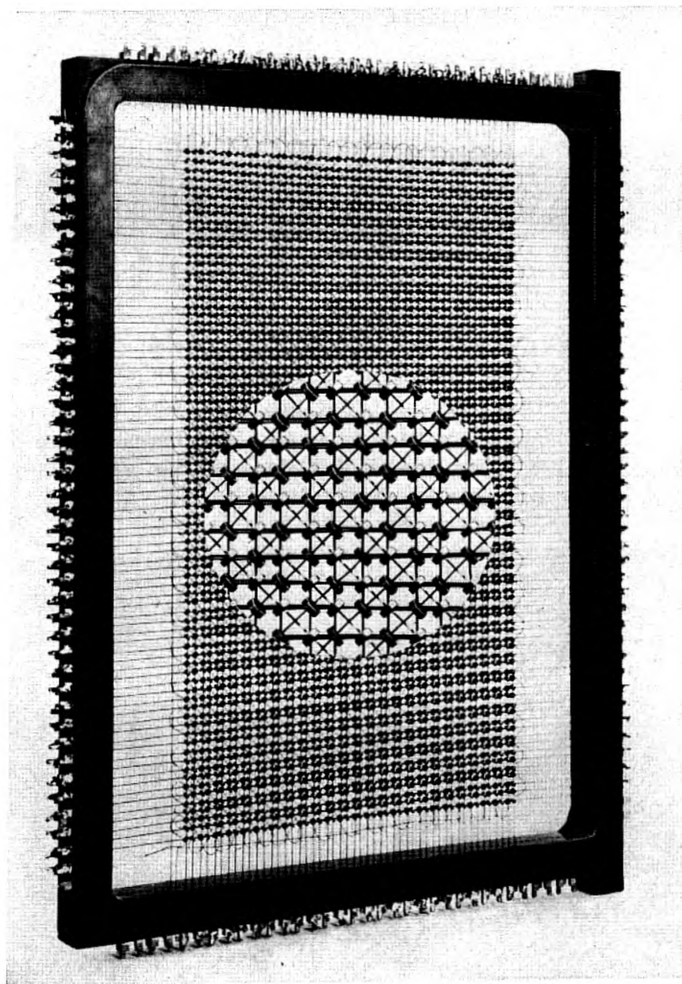


Fig. 18 — Magnetic core array (Courtesy of IBM).

One way in which electronic memory for various system applications may be evaluated is given by the chart of Fig. 22. This chart attempts to show, for the various forms of storage, the relation between the capacity in bits and cycle time, which includes access, reading and, if necessary, the regeneration time of the stored information. For sake of simplicity, ferroelectric and magnetic core memories have been combined as coordinate access arrays. Single bit electronic memory will be described in more detail later.

In the control portion of a switching system it is not only necessary to gather and store information but it must be interpreted and appropriate action taken. This function is called "processing". Processing circuits control the information gathering and storage functions and perform logical functions to produce the necessary flow of information. In the logic circuits of electronic systems, to keep pace with the time sharing nature of the information gathering function, the devices used must be several orders of magnitude faster than their counterparts, the relays, of the electromechanical system. The scanning and bulk memory

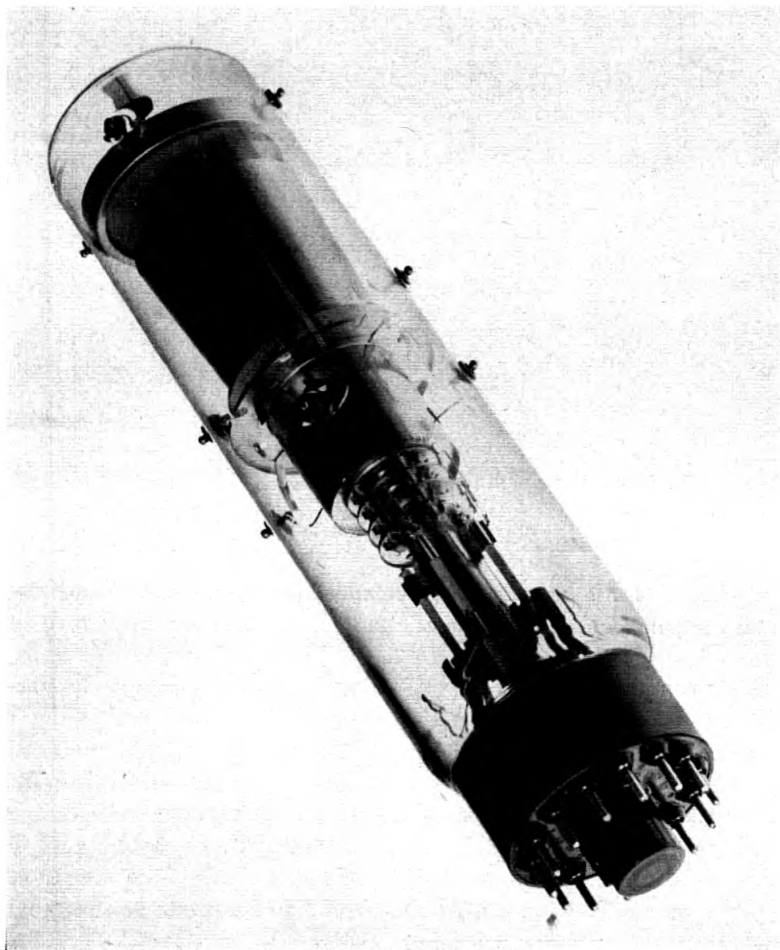


Fig. 19 — Electrostatic storage tube.

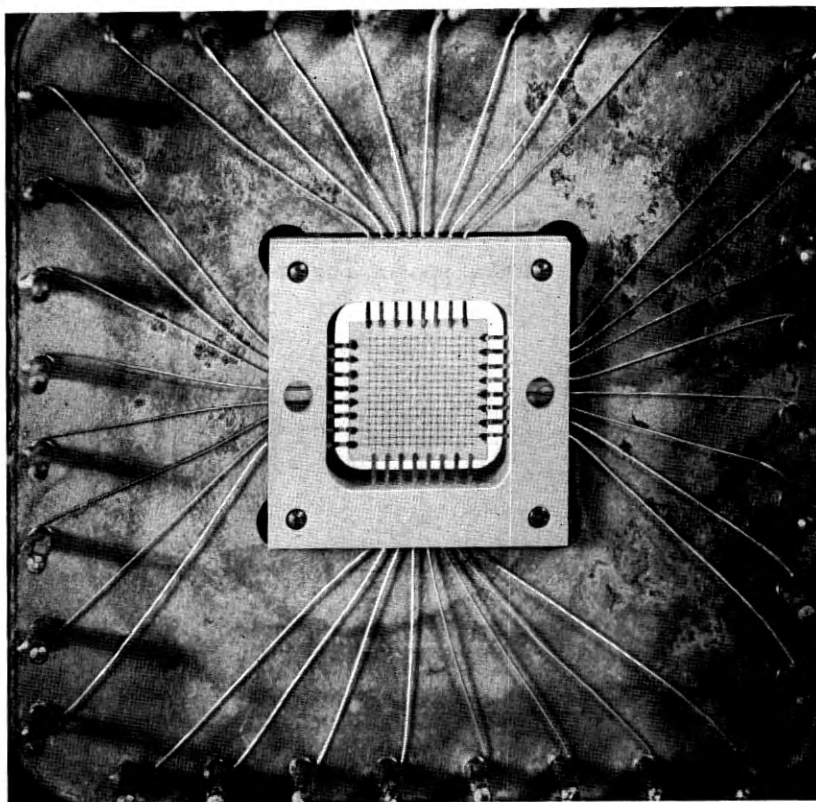


Fig. 20 — Ferroelectric array.

access speeds must be comparable in speed if they are not to become the speed bottleneck. All portions of the system must be in balance time-wise.

Devices and techniques for use in the design of high speed logic circuits are available.[22] With such devices information processing previously carried out by complex relay circuitry may be carried out in microseconds instead of milliseconds. Devices such as semiconductor diodes and transistors seem to be pointing the way to the future in performing these functions.[29] Previously, hot cathode tubes with high power consumption were needed to achieve the same functions at similar high speeds and for a long time this has been one of the greatest deterrents to electronic switching.

Semiconductor diode gate circuits are now quite familiar[23] and take

the place of the conventional make and break contacts in the electromechanical switching art (see Fig. 23 for the "AND" function). Magnetic core circuitry is also being exploited to perform high speed switching functions[24] (Fig. 24).

There are a number of differences between the circuit configuration used for relay contacts and diode or magnetic core gates for switching logic. When interconnecting such gates to realize complex logic functions other gates are required when circuit elements are placed in series or parallel, whereas in the wiring of relay contacts in series or in parallel no additional circuit elements are required (Fig. 25). Pulse signals passing through diode gate circuits are usually attenuated since the electronic device is not a perfect switcher (infinite impedance open circuit to zero impedance closed circuit). Some minute currents flow when open and some resistance is encountered when closed. Therefore, some amplification is needed at various places in logic circuits and this can be provided by transistor amplifiers. The use of transistors as the gating element eliminates this shortcoming by providing amplification in each gate (see Fig. 24). Transistors have also been successfully used in a new form of logic to provide relay contact like logic thus eliminating the need for gate elements to represent the series of paralleling functions[25] (see Fig. 26).

The processing of information usually requires a sequence of logic actions. To provide such sequences, momentary elements similar to locking relays but with microsecond action times are required. When this condition obtains a bistable or "flip-flop" circuit using transistors may be employed. Several forms of transistor circuits have been devised using either the Eccles-Jordan principle,[26] negative resistance properties,[27] such as achieved with a gas tube, or a regenerative approach.[28] Some suggestions have been made on the use of semiconductor diodes in special energy storing circuits to amplify pulses instead of the more conventional transistor amplifiers.[30]

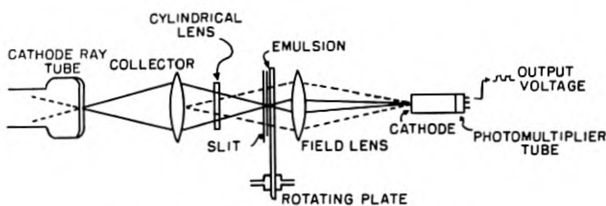


Fig. 21 — Photographic storage (from Proc. I.R.E., Oct. 1953).

EQUIPMENT CONCEPTS

In what has been said, consideration was given only to the concepts and circuitry of electronic telephone switching systems, but the things which the manufacturer and user come in contact with are the physical or equipment realizations of these concepts. One thing that is outstanding about the physical aspects of an electronic system is the large number of small components which are required. Fortunately, most of these components such as resistors, diodes, transistors, condensers, etc., are all of the same physical or similar mechanical design. From the manufacturer's point of view the problem then is to find the most economical way in which these many devices may be manufactured, assembled and tested, because of the large numbers required in a system. The basic solution appears to be: automatic production. This has led to the concept of small packages of components. These packages are the building blocks of a system and contain basic circuits which may be used repetitively. The trend in making such packages appears to be the use of printed wiring with automatic means of placing the components on the printed wiring boards.[31]

Despite the fact that there are large numbers of these small com-

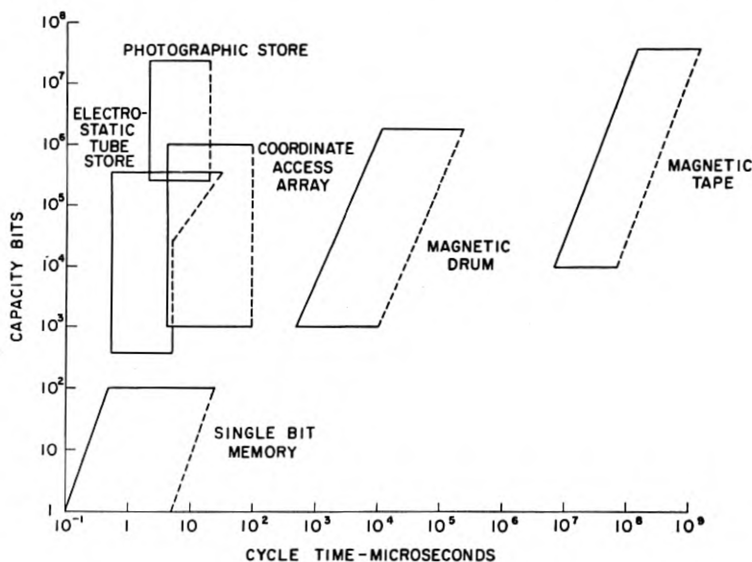


Fig. 22 — Memory system capabilities.

ponents required in electronic telephone switching they are small and when equipment using a multiplicity of printed wiring boards is assembled it takes on the aspect of a three-dimensional arrangement of components, with components mounted in depth as well as on the surface. This is in contrast to electromechanical systems where all components are generally mounted on a vertical surface. By using only one or two common control circuits of a given type (due to high speed) and

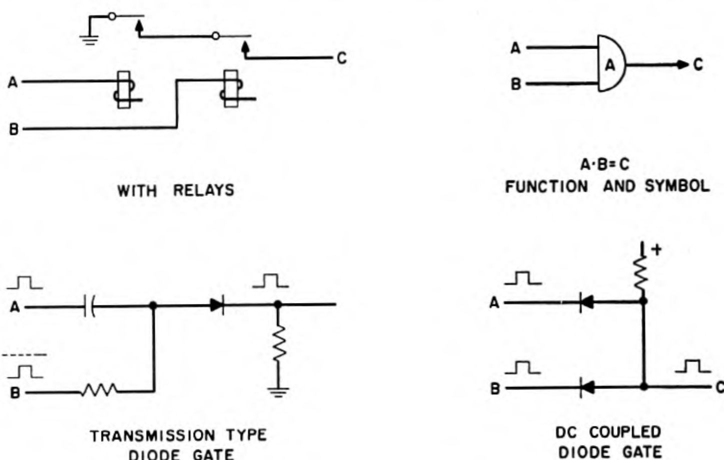


Fig. 23 — The "And" function.

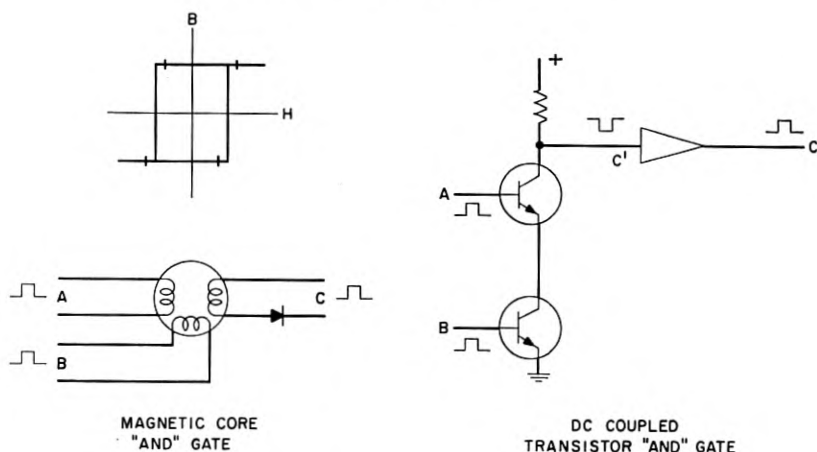


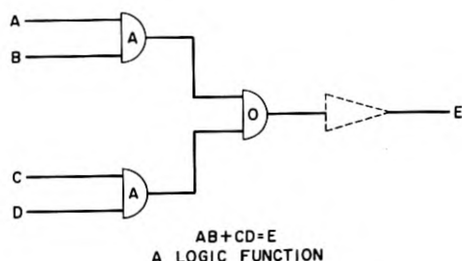
Fig. 24 — Other "And" circuits.

common medium bulk memory, fewer system elements are required which in the overall result in material space saving.

Another phase of the equipment aspects of electronic switching is that the devices require closer environmental control. Air conditioning appears necessary in early systems because of temperature limitations and other characteristics of some of the devices presently available. Also, vacuum tubes and other high power devices may develop objectionable hot spots in the equipment which make it advisable to exhaust hot air.

MAINTENANCE CONCEPTS

There is insufficient experience at this time to say what the maintenance problems of electronic telephone systems will be. Much has been written about the problems encountered in maintaining electronic



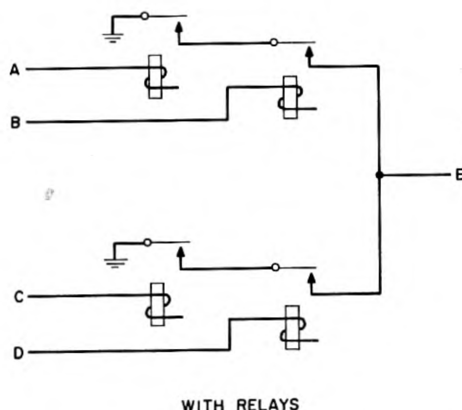


Fig. 25 — A logic function with relays.

computers; however, in designing a telephone system an entirely different philosophy must be pursued since it should not be necessary to have engineering caliber maintenance forces. At no time should the system be incapable of accepting and completing calls. This does not mean that portions of the system may not be worked on for routine or trouble maintenance.

A promising approach appears to be the use of marginal condition routine tests for detecting in advance components which are about to fail.[34] Automatic trouble locating arrangements may be devised for giving information as to the specific location of a package in trouble when it occurs.[35] This automatic trouble locator combined with the equipment concept of plug-in units means that service may be maintained without long interruptions. By designing devices which are reliable, employing them in a manner to give maximum service life and

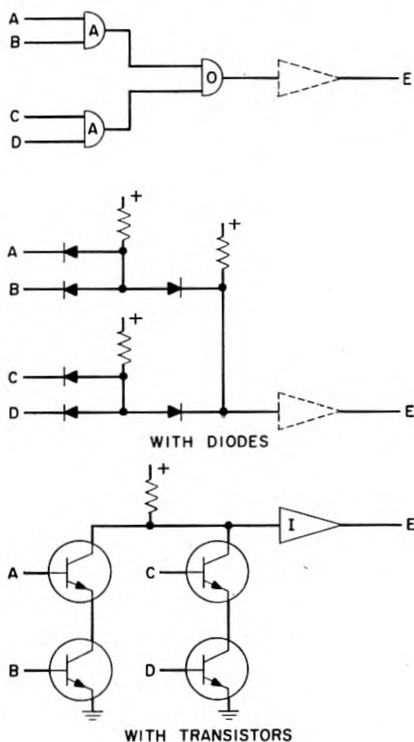


Fig. 26 — A logic function with diodes or transistors.

by judiciously introducing redundancy into the equipment, the chance of simultaneous failures of any two identical parts should be extremely improbable.[32] With automatic trouble locating, the maintenance forces will not be required to have a thorough understanding of the device characteristics and the circuitry used. Centralized repair of defective units as in modern telephone transmission systems[33] and perhaps even expendability of defective units are a distinct possibility.

As a result of some of these maintenance considerations it is quite likely that equipment in the future, besides being smaller and more compact, will appear more generally in enclosed low cabinets rather than exposed frames. The administrative control may be from consoles rather than vertical panels. More attention will be paid to appearance. The appurtenances, such as ladders required for high frames in electromechanical systems, may be eliminated.

Another change in concept which may come with electronics in telephone switching is the form of the power supply. Present day telephone systems use a centralized single voltage dc distribution system with reserve battery. The wide variety of devices and associated voltages, and the need for close regulation in some portions of electronic systems make a reliable ac distribution system with individual power rectifiers at the point of use appear quite attractive. To insure reliability of service the ac distribution must be continuous and not dependent directly upon the commercial sources.

There is no question that reliability is imperative if electronic switching systems are to survive among electromechanical systems which have achieved a high degree of reliability over a long period of years. The device reliability of the first electronic system may not be comparable since some of the components of the electronic switching systems will not in their initial applications be as reliable as the least reliable component in our present day systems. Reliability will be earned and this will probably require considerable effort. Even if initially some devices employed in electronic systems do not measure up to the present high standard which has been set, continuity of high quality service is a must. It is, therefore, necessary to design a system which will mask the shortcomings of any individual electronic component.[32] As their reliability is proven an optimum balance will be sought between system redundancy and component quality. Telephone engineers familiar only with the high degree of reliability of present day apparatus will have to accommodate themselves to the characteristics of new electronic devices.

REFERENCES

1. J. R. Eckert, A. Survey of Digital Computer Memory Systems, I.R.E. Proceedings, **41,** pp. 1393–1406, Oct., 1953.
2. T. H. Flowers, Electronic Telephone Exchanges, Proceedings I.E.E., **99,** Part I, pp. 181–201, 1952.
3. U. S. Patent 2,387,018.
4. U. S. Patent 2,490,833.
5. U. S. Patent 2,408,462.
6. U. S. Patent 2,379,221.
7. W. A. Depp, M. A. Townsend, Cold Cathode Tubes for Audio Frequency Signaling, B.S.T.J., **32,** pp. 1371–1391, Nov., 1953.
8. Tone Ringer May Replace Telephone Bell, Bell Laboratories Record, pp. 116–117, March, 1956.
9. W. R. Bennett, Time Division Multiplex Systems. B.S.T.J., **20,** p. 199, 1941.
10. F. A. Korn, J. G. Ferguson, No. 5 Crossbar Dial Telephone Switching System, *Elec. Eng.*, **69,** pp. 679–684, Aug., 1950.
11. J. W. Dehn, R. E. Hersey, Recent New Features of the No. 5 Crossbar Switching System, A.I.E.E. *Paper No. 55-580*.
12. T. L. Dimond, No. 5 Crossbar AMA Translator, Bell Laboratories Record, p. 62, Feb., 1951.
13. L. N. Hampton, J. B. Newsom, The Card Translator for Nationwide Dialing, B.S.T.J., **32,** pp. 1037–1098, Sept., 1953.
14. Review of Input and Output Equipment Used in Computing Systems. A.I.E.E. *Special Publication S53*.
15. Cohen, A. A., Magnetic Drum for Digital Information Processing Systems, Mathematical Aids to Computation. **4,** pp. 31–39, Jan., 1950.
16. M. E. Hines, M. Chruney, J. A. McCarthy, Digital Memory in Barrier Grid Storage Tubes, B.S.T.J., **43,** p. 1241, Nov., 1955.
17. M. Knoll, B. Kazan, Storage Tubes and Their Basic Principles, John Wiley & Sons, 1952.
18. M. K. Haynes, Multidimensional Magnetic Memory Selection System. Transactions of the I.R.E., Professional Group on Electronic Computers, pp. 25–29, Dec., 1952.
19. D. A. Buck, Ferroelectrics for Digital Information Storage and Switching, Report R212, M.I.T. Digital Computer Laboratories, June, 1952.
20. J. R. Anderson, Ferroelectric Materials as Storage Elements for Digital Computers and Switching Systems, Communications and Electronics, pp. 395–401, Jan., 1953.
21. G. W. King, G. W. Brown, L. N. Ridenour, Photographic Techniques for Information Storage, Proc. I.R.E., pp. 1421–1428, Oct., 1953.
22. Staff of Harvard Computation Laboratory, Synthesis of Electronic Computing and Control Circuits, *Vol. 27 of Annals of Harvard Computation Laboratory*, 1951.
23. B. J. Yokelson, W. Ulrich, Engineering Multistage Diode Logic Circuits, *Communications and Electronics*, pp. 466–474, Sept., 1955.
24. M. Karnaugh, Pulse Switching Circuits Using Magnetic Cores, Proc. I.R.E., **43,** pp. 576–584, May, 1955.
25. R. H. Beter, W. E. Bradley, R. B. Brown, M. Rubinoff, Surface Barrier Transistor Switching Circuits, I.R.E. Convention Record, Part 4, pp. 139–145, 1955.
26. R. L. Trent, Two Transistor Binary Counter, Electronics, **25,** pp. 100–101, July, 1952.
27. A. E. Anderson, Transistors in Switching Circuits, Proc. I.R.E., **40,** pp. 1541–1558, Nov., 1952.
28. J. H. Felker, Regenerative Amplifier for Digital Computer Applications, Proc. I.R.E., **40,** pp. 1584–1956, Nov., 1592.
29. J. H. Felker, Typical Block Diagrams for a Transistor Digital Computer, Communications and Electronics, pp. 175–182, July, 1952.

30. Promising Electronic Components — Diode Amplifiers, Radio Electronics, p. 45, Nov., 1954.
31. A. A. Lawson, Mass Production of Electronic Subassemblies, Electrical Manufacturing, **54**, p. 134, Oct., 1954.
32. C. J. Crevelens, Increasing Reliability by the Use of Redundant Circuits, Proc. I.R.E., pp. 509–515, April, 1956.
33. A. L. Bonner, Servicing Center for Short-Haul Carrier System, Communications and Electronics, pp. 388–396, Sept., 1954.
34. N. L. Daggett, E. S. Rich, Diagnostic Programs and Marginal Checking in Whirlwind I Computer, I.R.E. *Convention Record*, Part 7, pp. 48–54, 1953.
35. MAID Service for Computer Circuits, Automatic Control, p. 23, Aug., 1955.

Combined Measurements of Field Effect, Surface Photo-Voltage and Photoconductivity

By W. H. BRATTAIN and C. G. B. GARRETT

(Manuscript received May 10, 1956)

Combined measurements have been made of surface recombination velocity, surface photo-voltage, and the modulation of surface conductance and surface recombination velocity by an external field, on etched germanium surfaces. Two samples, cut from an n-type and a p-type crystal of known body properties, were used, the samples being exposed to the Brattain-Bardeen cycle of gaseous ambients. The results are interpreted in terms of the properties of the surface space-charge region and of the fast surface states. It is found that the surface barrier height, measured with respect to the Fermi level, varies from −0.13 to +0.13 volts, and that the surface recombination velocity varies over about a factor of ten in this range. From the measurements, values are found for the dependence of charge trapped in fast surface states on barrier height and on the steady-state carrier concentration within the semiconductor.

I. INTRODUCTION

This and the succeeding paper are concerned with studies of the properties of fast surface states on etched germanium surfaces. The experiments involve simultaneous measurement of a number of different physical surface properties. The theory, which will be presented in the second paper, interprets the results in terms of a distribution of fast surface states in the energy gap. The distribution function, and the cross-sections for transitions from the states into the conduction and valence bands, may then be deduced from the experimental results.

Early experiments[1] on contact potential of germanium, and on the change of contact potential with light, indicated that there are two kinds of surface charge associated with a germanium surface, over and above the holes and electrons that are distributed through the surface space-charge region. One kind of surface charge, usually called "charge

in fast traps" can follow a change in the space-charge region very fast in comparison with the light-chopping time used in that work (1/100 sec); the other kind, imagined to be more closely connected with adsorbed chemical material, can only change rather slowly. In a previous paper by the authors[2] it was pointed out that the Brattain-Bardeen experiments, taken by themselves, do not furnish unambiguous information concerning the distribution of these "fast" traps, but that such information might be obtained by performing, simultaneously, other measurements on the germanium surface. More recently Brown and Montgomery[3, 4] have provided a valuable tool in their studies of large-signal field effect; they point out that if, under given chemical conditions, it is possible to apply a field, normal to the surface, large enough to force the surface potential to the minimum in surface conductivity; then it becomes possible to determine the initial surface potential absolutely (provided certain considerations as to the mobility[5] of the carriers near the surface are valid).

This paper concerns studies of a number of physical properties that depend on the distribution and other characteristics of the surface traps or "fast" states. Measurements are reported of (i) the change of conductivity of a sample with field; (ii) the photoconductivity; (iii) the change of photoconductivity with field; (iv) the filament lifetime; and (v) the surface photo-voltage. Measurements were made in a series of gaseous ambients, first described by Brattain and Bardeen.[1] Evidence is presented to the effect that the variation in gas ambient changes only the "slow" states, leaving the distribution and other properties of the traps substantially unaffected. From measurements (i) to (iii) it is possible to construct the whole field-effect curve (conductance versus surface charge), even though the fields used were in general not large enough to reach the minimum in conductance.

Using the field effect data, values for the surface potential Y in units of kT/e could be obtained at each point, and also of the quantity $(\partial \Sigma_s / \partial Y)_{\delta=0}$, where Σ_s is the charge in surface traps, and the suffix $\delta = 0$ implies zero illumination. From measurements (ii) and (iv), the surface recombination velocity s could be deduced. (A more detailed study of photoconductivity in relation to surface recombination velocity will be reported at a later date.) Combined with the field effect data, this enables one to deduce the relation between s and Y.

Measurements of the surface photo-voltage may be presented in terms of the quantity $dY/d\delta$, where δ is equal to $\Delta p/n_i$, Δp being the density of added carrier-pairs in the body of the material, and n_i the intrinsic carrier density. The quantity $dY/d\delta$ is closely related to the ratio of the

change in surface potential produced by illumination of the surface to the change in the quasi-Fermi level for minority carriers. By measuring $dY/d\delta$ rather than dY/dL, discussed in Reference 2, the surface recombination velocity is eliminated from the surface photo-voltage data: the limiting values of $dY/d\delta$, after correction for the Dember effect, ought to be (p_0/n_i) and $-(n_i/p_0)$, no matter what the surface recombination velocity may be.

By combining this information with the field-effect data, one can deduce the quantity $(\partial\Sigma_s/\partial\delta)_Y$. This and the previous differential, deduced directly from the field-effect data, completely define the dependence of charge in surface traps on the two independent parameters Y and δ — that is, the dependence on chemical environment and on the bulk non-equilibrium carrier level.

The further interpretation of the quantities $(\partial\Sigma_s/\partial Y)_{\delta=0}$, $(\partial\Sigma_s/\partial\delta)_Y$ and s in terms of the distribution of surface traps is postponed to the succeeding paper. Here it is sufficient to say that the results are consistent with the assumption that the traps responsible for surface recombination are also those pertinent to the field effect and surface photovoltage experiments. Then the quantity $(\partial\Sigma_s/\partial Y)_{\delta=0}$ depends only on an integral over the distribution in energy of traps; $(\partial\Sigma_s/\partial\delta)_Y$ depends also on the ratios of cross-sections for transitions to the valence and conduction bands; and s depends in addition on the geometric mean cross-sections.

II. OUTLINE OF THE EXPERIMENT

The experiment is carried out with a slice of germanium, 0.025 cm thick, which is supported in such a way that there is a gap 0.025 cm wide between the slice and a metal plate. Substantially ohmic contacts are attached to the ends of the slice. Three kinds of experiment are now carried out:

(i) The conductance of the slice is modulated by illuminating it with a short flash of light; the subsequent decay of photoconductivity with time is studied, and the time-constant of the exponential tail measured.

(ii) A sinusoidally varying potential difference of about 500 volts peak-to-peak is applied between the metal plate and the germanium. Facilities are available for measuring the changes in conductance produced by the field. The sample is also illuminated with light chopped at a frequency different from that of the applied field. One measures: (a) the magnitude of the peak-to-peak conductance change in the dark; (b) the same in the presence of the light; and (c) the change in con-

conduction or the valence band. This in turn depends both on the surface potential and on the ratio of cross-sections for transitions to the two bands. For $Y \ll -1$, one expects $(\partial\Sigma_s/\partial\delta)_Y/(\partial\Sigma_s/\partial Y)_\delta$ to have the value $-\lambda^{-1}$; for $Y \gg +1$, the value $+\lambda$. These limiting values may be deduced by a somewhat general argument.[7]

At some intermediate value of surface potential, the above ratio must change sign. If the distribution of surface states in energy is known from the field effect measurements, then the value of Y at which the above ratio changes sign determines the ratio of cross-sections for those traps which are close to the Fermi level for that value of Y. By repeating the experiment for samples of differing bulk resistivity, it is then possible to determine whether the same ratio holds for the states at some different position in the energy gap.

III. EXPERIMENTAL DETAILS

Fig. 1 shows the experimental arrangements. The sample of germanium, of dimensions shown, was prepared by cutting, sandblasting, etching in CP4[1] and washing in distilled water. The exposed faces were approximately (100). The end contacts were made by sandblasting and soldering. The slots A, A' in the ceramic were incorporated in order to

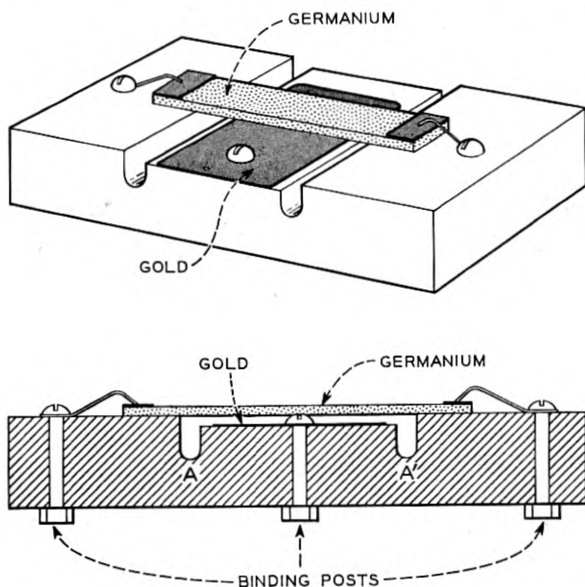


Fig. 1 — Experimental arrangement.

reduce the high field that would otherwise be present near the edges of the ceramic. The gold electrode was deposited by evaporation through a mask. Connections from the gold and from the ends of the sample were made to binding posts passing through the ceramic block.

The ceramic block was set into a metal box, divided into two compartments. In the upper compartment, which contained the sample, there were inlet and outlet tubes, to allow the gas to be changed. The lower compartment contained electrical components, which were thereby protected to a large extent from the changes in gas in the upper compartment. Facilities were available for the type of cycle of gas environment described by Brattain and Bardeen,[1] which cycle was found by them to produce reversible cyclic changes in surface potential. In the top of the box was a window, through which light could be shone onto the germanium either from a chopped or a flash source.

The electrical circuit is shown in Fig. 2. The condenser C_1 is that formed between the germanium and the gold, and has a capacity of about 2 $\mu\mu F$. Impedances Z_1 and Z_2 form a Wagner ground, which has to be balanced first. Then, by adjusting resistance R_1 and condenser C_2, one may obtain a balance in the case that there is no dc flowing through the sample. A current (determined by the battery B and the

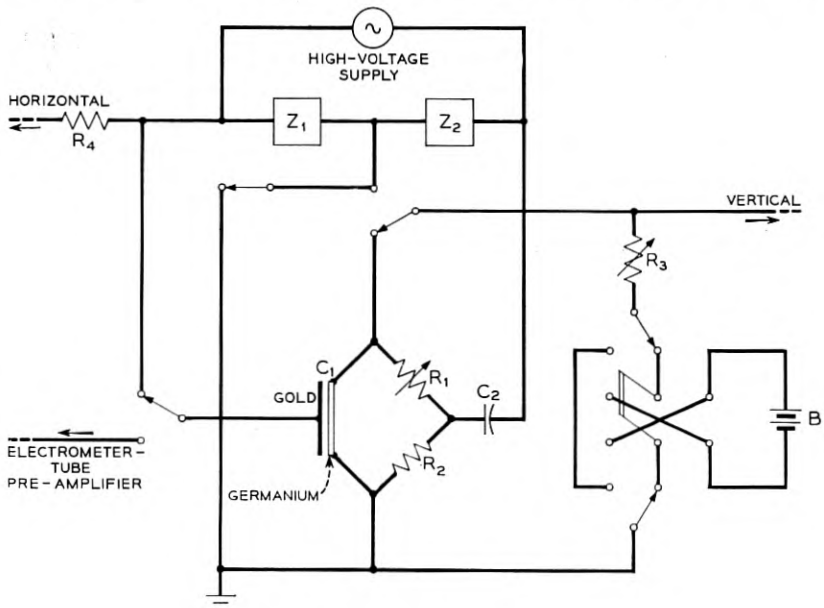


Fig. 2 — Electrical circuit.

resistance R_3) is now switched in, and the resulting off-balance (representing the field modulation of conductivity) presented on the vertical plates of an oscilloscope. The supply voltage is connected, via the high bleeder resistance R_4, to the horizontal plates. The frequency of the oscillator was chosen to be 25 cyc/sec, a value sufficiently low to obviate lifetime difficulties; the peak-to-peak swing was generally 500 volts.

During a field-effect measurement, the sample was also illuminated with light chopped at 90 cyc/sec. This had the result of causing to be presented on the oscilloscope screen a pattern such as that shown in Fig. 3. The lower tilted line represents the (dark) field effect curve; the vertical separation represents the photoconductivity, as modified by the applied field. Measurements were made of the mean vertical separation, and of the slopes of the upper and lower lines (by reading gain settings).

During a surface photo-voltage measurement, the gold electrode was disconnected from the high-voltage supply, and connected to a high-impedance detector, similar to that used in the work of Brattain and Bardeen.[1] A value for the chopped light intensity was chosen to give a contact potential change that was generally not more than 5 mV. A simultaneous measurement of the photoconductivity was also made.

The gas cycle was similar to that described by Brattain and Bardeen.[1] Some variations were made in it to try to spread out the rate of change with time so that the data could be obtained without large gaps. The cycle used was: (i) sparked oxygen 1 min, (ii) dry O_2, (iii) mixture of dry and wet O_2, (iv) wet O_2, (v) wet N_2, (vi) a mixture of dry and wet N_2, (vii) dry O_2, (viii) dry O_2, triple flow, and (ix) ozone normal flow. The normal rate of gas flow was about 2 liters per minute; the wet gas was obtained by bubbling through water (probably about 90 per

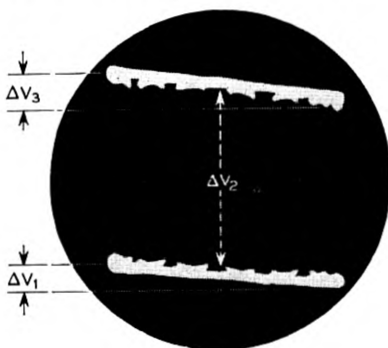


Fig. 3 — Picture of field effect-photoconductivity pattern, as observed on oscilloscope. Dark curve at the bottom.

cent r.h.) and the mixture of dry and wet was obtained by letting approximately one-half the gas flow bubble through H_2O. In carrying out the experiment, it was found convenient to carry out alternately a complete cycle of field effect and surface photo-voltage measurements. The values of the photoconductivity at equivalent points in successive cycles could be compared, in order to check that no systematic error was introduced by this procedure.

In addition to the foregoing, the following measurements were made:

1. All dimensions were determined.
2. The resistivity of the sample was found, and also the body lifetime, on another specimen cut from the same crystal.
3. The amplitude of the voltage swing was measured.
4. The amplifiers in the field effect circuit were calibrated.
5. The capacity of the germanium-gold condenser was determined (by a substitutional method). The value obtained was larger than that calculated from the parallel-plate formula, because of the edge effects.
6. A standard square-wave voltage was introduced into the surface photo-voltage circuit, in order to calibrate the high-impedance detector.
7. At several points in the cycle, the fundamental mode lifetime of the sample was determined by the photoconductivity decay method. This calibrated the 90 cyc/sec photoconductivity measurements, without the necessity for a knowledge of the light intensity.

IV. RESULTS

Measurements were made on two samples: one n-type, 22.6 ohm cm ($\lambda = 0.345$), the other p-type, 8.1 ohm cm ($\lambda = 17.7$). The body lifetime for both samples was greater than 10^{-3} sec, so that for slices of the thickness used (0.025 cm. or less), and for values of s in the range encountered, body recombination may be ignored.

Results of typical field-effect runs for the two samples are indicated in Tables I and II. The first column in each table gives the time in minutes from the beginning of the cycle at which the measurements were made. The second column shows the "effective mobility," $d\Delta G/d\Sigma$, obtained from the observed (dark) field effect signal voltage ΔV_1 (see Fig. 3) by use of the formula: $\mu_{\text{eff}} = w^2t^2\Delta V_1/I\rho_0^2CV_{\text{app}}$, where w is the width of the slice, t the thickness, I the dc flowing through it, ρ_0 the resistivity, C the capacity of the germanium-gold condenser, and V_{app} the voltage applied across it. The third column shows the mean value of $\delta(=\Delta p/n_i)$, obtained from the mean photoconductivity signal voltage

TABLE I — 22.6 OHM CM n-TYPE CYCLE 12.
RELATIVE LIGHT INTENSITY 0.082

Time min.	μ_{eff} $\frac{cm^2}{volt\ sec}$	δ	μ_{eff}^* $\frac{cm^2}{volt\ sec}$	s $\frac{cm}{sec}$
0	Sparked O_2			
1	Changed to dry O_2			
1.5	334	7.85×10^{-2}	520	90
2.5	344	6.9×10^{-2}	585	103
5.5	344	6.2×10^{-2}	595	114
6.5	344	6.07×10^{-2}	595	117
7.0	Changed to mixture of dry & wet O_2			
7.5	136	4.4×10^{-2}	520	161
8.5	84	4.02×10^{-2}	440	177
9.0	52	3.60×10^{-2}	270	196
10.0	Changed to full wet O_2			
11.0	−660	2.26×10^{-2}	−440	314
11.5	−890	1.74×10^{-2}	−780	408
12.0	−960	1.67×10^{-2}	−910	425
13.0	Changed to full wet N_2			
18.0	Changed to mixture of dry & wet N_2			
19.5	−1150	1.67×10^{-2}	−1060	425
20.5	−1050	1.74×10^{-2}	−960	408
22.5	−990	1.83×10^{-2}	−890	390
23.0	Changed to dry O_2			
23.5	−430	2.98×10^{-2}	−220	238
23.8	−290	3.2×10^{-2}	0	222
24.0	−84	3.81×10^{-2}	240	186
24.5	31	4.3×10^{-2}	310	165
26.5	146	4.7×10^{-2}	410	151
27.0	Tripled flow of dry O_2			
27.5	220	5.1×10^{-2}	450	139
29.5	260	5.7×10^{-2}	510	124
31.5	280	6.2×10^{-2}	510	114
35.0	Changed to ozone			
35.5	310	6.8×10^{-2}	510	104
37.5	320	8.2×10^{-2}	490	87

ΔV_2 by use of the formula: $\delta = wt\rho_i\Delta V_2/I\ell_{ill}\rho_0^2$, where ρ_i is the intrinsic resistivity and ℓ_{ill} the length of the illuminated part of the slice. The fourth column shows the apparent effective mobility in the presence of light,* obtained from the field effect signal voltage ΔV_3 in the presence of light, using the same formula as that giving μ_{eff}. The last column shows the surface recombination velocity, which is proportional to δ^{-1} for fixed light intensity, the constant of proportionality being determined by comparison with measurements of the fundamental mode lifetime.

The results of typical surface photo-voltage runs are shown in Tables

* One must be careful to avoid thinking of μ_{eff}^* as a true field effect mobility, since it is really a sum of two quite different components: the true field effect mobility μ_{eff}, and a term, proportional to thickness of the slice, arising from the photoconductivity.

TABLE II — 8.1 OHM CM p-TYPE CYCLE 5.
RELATIVE LIGHT INTENSITY 0.25

Time min.	μ_{eff} $\frac{cm^2}{volt\ sec}$	δ	μ_{eff}^* $\frac{cm^2}{volt\ sec}$	s $\frac{cm}{sec}$
0	Started sparked O_2			
1.0	Changed to dry O_2			
1.5	307	4.1×10^{-2}	490	503
3.5	318	3.2×10^{-2}	490	660
6.0	Changed to mixture of wet & dry O_2			
7.5	273	1.4×10^{-2}	376	1480
9.5	239	1.3×10^{-2}	320	1580
11.0	Changed to wet O_2			
11.5	94	1.2×10^{-2}	−194	1690
12.5	−200	1.1×10^{-2}	−230	1820
15.5	−216	1.2×10^{-2}	−285	1690
17.0	Changed to wet N_2			
18.5	−352	1.6×10^{-2}	−570	1310
22.0	Changed to mixture of wet & dry O_2			
25.0	−80	1.1×10^{-2}	−137	1820
26.5	0	1.2×10^{-2}	31	1690
27.5	3.3	1.3×10^{-2}	58	1630
28.0	Changed to dry O_2			
29.0	193	1.9×10^{-2}	330	1070
29.5	239	2.2×10^{-2}	400	1000
30.5	250	2.4×10^{-2}	420	873
33.0	Tripled flow of dry O_2			
33.5	296	3.2×10^{-2}	500	645
34.5	296	3.7×10^{-2}	525	560
36.5	296	4.2×10^{-2}	570	490
37.5	296	4.6×10^{-2}	570	455
38.0	Changed to ozone			
42.5	330	6.4×10^{-2}	535	323

III and IV. Values of δ were obtained from the photoconductivity signal, as before, taking the actual illuminated length as the length of the sample. In making use of the standard square-wave calibration for the surface photo-voltage measurement (Section III), it is necessary to allow for the fact that the measured capacity involves the whole length of the sample, plus end and side fringing effects, whereas the surface photo-voltage measurements involves only the illuminated length, plus the fringe effect at the *sides*.

The penultimate column in Tables III and IV shows the ratio of the change in contact potential, measured in units of (kT/e), to the added-carrier parameter δ, which was deduced from the photoconductivity. This is not yet, however, the true surface photo-voltage function $(dY/d\delta)$, since the observed change in contact potential includes also the Dember potential ΔV_i[10] which occurs between the illuminated and non-illuminated parts of the *body* of the semiconductor. The last column in Tables III and IV shows the true values of $(dY/d\delta)$, obtained by sub-

TABLE III — 22.6 OHM CM *n*-TYPE CYCLE 7

Time mins.	Relative Light Intensity	δ	ΔCP volts	$\frac{\beta \Delta CP}{\delta}$	$\frac{dY}{d\delta}$
Starting condition wet N_2					
6.5	2.25	0.36	6.5×10^{-3}	0.7	−0.10
11.5	2.25	0.34	1.0	0.115	−0.045
12.0	Changed to mixture wet and dry N_2				
12.5	2.25	0.32	2.2	0.27	0.10
13.0	2.25	0.34	3.5	0.40	0.23
13.5	2.25	0.35	4.6	0.51	0.34
14.5	2.25	0.38	6.8	0.70	0.53
15.5	0.56	0.10	3.1	1.2	1.03
17.5	0.56	0.11	3.7	1.33	1.16
18.0	Changed to dry O_2				
18.5	0.56	0.16	5.7	1.4	1.2
19.5	0.14	0.06	3.4	2.2	2.0
22.0	Changed to dry O_2 triple flow				
24.5	0.14	0.082	6.1	2.9	2.7

TABLE IV — 8.1 OHM CM *p*-TYPE CYCLE 8

Time mins.	Relative Light Intensity	δ	ΔCP volts	$\frac{\beta \Delta CP}{\delta}$	$\frac{dY}{d\delta}$
Starting condition wet N_2					
0	Changed to mixture wet and dry N_2				
0.5	0.14	0.011	-3.5×10^{-3}	−12.5	−12.6
1.0	0.14	0.0088	−2.1	−9.6	−9.7
5.0	Changed to mixture wet and dry O_2				
5.5	0.56	0.0275	−1.8	−2.6	−2.7
6.0	0.56	0.03	−1.45	−1.9	−2.0
7.5	0.56	0.0325	−1.16	−1.4	−1.5
9.5	0.56	0.035	−1.08	−1.2	−1.3
10.0	Changed to dry O_2				
10.5	0.56	0.044	−0.71	−0.63	−0.69
11.5	0.56	0.055	−0.49	−0.35	−0.41
14.0	Changed to dry O_2 triple flow				
14.5	0.56	0.0625	−0.32	−0.20	−0.26
16.5	2.25	0.28	−0.72	−0.10	−0.16
20.5	2.25	0.33	−0.42	−0.05	−0.11
30.0	Changed to ozone				
31.5	2.25	0.47	+0.47	+0.039	−0.023
32.5	2.25	0.53	+1.4	+0.103	+0.041

tracting from ($\beta\Delta$ c.p./δ) a Dember potential correction, given by $(b - 1)/(\lambda + b\lambda^{-1})$. (The boundaries of the illuminated region were sufficiently distant from the contacts for this formula to apply.)

Tables III and IV include only data from the second half of the cycle (wet $N_2 \rightarrow$ ozone), since the rate of change of Δ c.p. during that part of the first half in which dry oxygen was replaced by wet oxygen was too fast to follow.

The reproducibility of all the data from cycle to cycle was good. One surprising result is that the surface recombination velocity assumed its maximum value close to the "wet nitrogen" extreme for both p-type and n-type. This behavior is quite different from that reported by Brattain and Bardeen,[1] who found s to be constant within 20 per cent throughout the range and Stephenson and Keyes,[8] who found a maximum value sometimes at one end, sometimes at the other, and sometimes in the middle. There is quite good agreement on the other hand, with the results of Many et al.[12], who report a maximum in s near the wet end of the cycle. The result of Brattain and Bardeen is not understood at the present time, and is probably wrong. The differences between the present work and that of Stephenson and Keyes may be associated with differences in surface preparation.

V. ANALYSIS OF THE RESULTS

From now onwards we shall express all experimental and calculated quantities in terms of the following dimensionless ratios:

$$\bar{\Gamma}_p = \Gamma_p/n_i\mathfrak{L} \qquad \bar{\Gamma}_n = \Gamma_n/n_i\mathfrak{L}$$
$$\bar{\Sigma}_s = \Sigma_s/en_i\mathfrak{L} \qquad \bar{\Sigma} = \bar{\Sigma}_s + \bar{\Gamma}_p - \bar{\Gamma}_n \tag{2}$$
$$\overline{\Delta G} = \Delta G/en_i\mu_p\mathfrak{L}, \qquad \bar{\mu}_{\text{eff}} = \mu_{\text{eff}}/\mu_p\,, \qquad \bar{\mu}_{\text{eff}}^* = \mu_{\text{eff}}^*/\mu_p$$

where ΔG is the surface conductance, $\mathfrak{L}$ the Debye length for intrinsic germanium (1.4×10^{-4} cm), and μ_p is the mobility for holes (1800 cm^2v^{-1} sec^{-1}). Tables V and VI show values of the quantities we shall need, as functions of the surface potential Y, calculated from the theoretical considerations of Garrett and Brattain.[2] The surface conductance, and the differentials in the fifth and sixth columns, are evaluated for $\delta = 0$.

TABLE V — 22.6 OHM CM n-TYPE

Y	$Y - \ln\lambda$	$\bar{\Gamma}_p - \bar{\Gamma}_n$	$\Delta\bar{G}$	$\left(\frac{\partial(\bar{\Gamma}_p - \bar{\Gamma}_n)}{\partial Y}\right)_\delta$	$\left(\frac{\partial(\bar{\Gamma}_p - \bar{\Gamma}_n)}{\partial \delta}\right)_Y$
3	4.1	−10.3	17.5	−4.1	−1.3
2	3.1	−7.0	10.6	−2.6	−0.8
1	2.1	−4.9	6.2	−1.8	−0.4
0	1.1	−3.4	3.3	−1.3	0.0
−1	0.1	−2.3	1.45	−1.1	0.5
−2	−0.9	−1.2	0.36	−1.1	1.3
−3	−1.9	0.0	0.0	−1.4	2.7
−4	−2.9	1.7	0.65	−2.1	5.2
−5	−3.9	4.4	2.65	−3.5	9.4
−6	−4.9	8.9	6.8	−5.8	16.3
−7	−5.9	16.4	14.4	−9.5	27.7

TABLE VI — 8.1 OHM CM p-TYPE

Y	$Y - \ln \lambda$	$\bar{\Gamma}_p - \bar{\Gamma}_n$	$\overline{\Delta G}$	$\left(\frac{\partial(\bar{\Gamma}_p - \bar{\Gamma}_n)}{\delta Y}\right)_\delta$	$\left(\frac{\partial(\bar{\Gamma}_p - \bar{\Gamma}_n)}{\partial \delta}\right)_Y$
8	5.1	−8	9.8	−5.5	−87
7	4.1	−4	3.4	−3.1	−42
6	3.1	−2	0.8	−1.9	−19
5	2.1	0	0.0	−1.45	−8.2
4	1.1	1	0.25	−1.3	−3.4
3	0.1	2	1.3	−1.45	−1.5
2	−0.9	4	2.8	−1.75	−0.62
1	−1.9	6	4.8	−2.4	−0.21
0	−2.9	9	7.4	−3.4	0.0
−1	−3.9	12	10.9	−4.3	0.15
−2	−4.9	18	16.4	−6.4	0.31
−3	−5.9	26	25.0	−10.0	0.53

The first problem is the constructing, from the experimental results, of the curve relating $\overline{\Delta G}$ and $\bar{\Sigma}$. The experiments provide a series of pic tures like Fig. 3, each one corresponding to a different chemical environment, and so to a different Y. At each of two succeeding pictures of this sort one knows (i) the vertical displacement (photoconductivity) between the dark and light field effect curves; and (ii) the mean difference in the dark and light slopes, and hence the rate of change of photoconductivity with applied field, and therefore with $\bar{\Sigma}$. The problem is to deduce the horizontal displacement (in $\bar{\Sigma}$) between the two pictures.

A corrrection must first be made for the fact that the ambient changes $\bar{\Sigma}$ uniformly on both surfaces, whereas the applied field induces charge only on the lower surface, plus fringing effects.* The correction is applied by taking the difference in slopes ($\mu_{\text{eff}}{}^* - \mu_{\text{eff}}$), and multiplying this by (2/1.27), where the number 1.27 is deduced for the given geometry from the standard edge-effect formula.[9] This having been done, it is now possible to take the revised pictures and piece them together to form two smooth curves (Fig. 4). The process of assembling such a diagram determines the horizontal and vertical distances, and therefore the change of $\bar{\Sigma}$ and $\overline{\Delta G}$, between successive experiments.

This argument may be given analytically as follows. First notice that the photoconductivity voltage in the absence of field (ΔV_2 in Fig. 3) is proportional to $(1/s)$. The application of a voltage between the gold and the germanium induces some charge density Σ at each point on the germanium surface, Σ being (due to fringing effects) a complicated function of position. At each point $(1/s)$ is changed by an amount $\Sigma[d(1/s)/d\Sigma]$. This causes the photoconductivity in the presence of field

* We are indebted to W. L. Brown for bringing this to our attention.

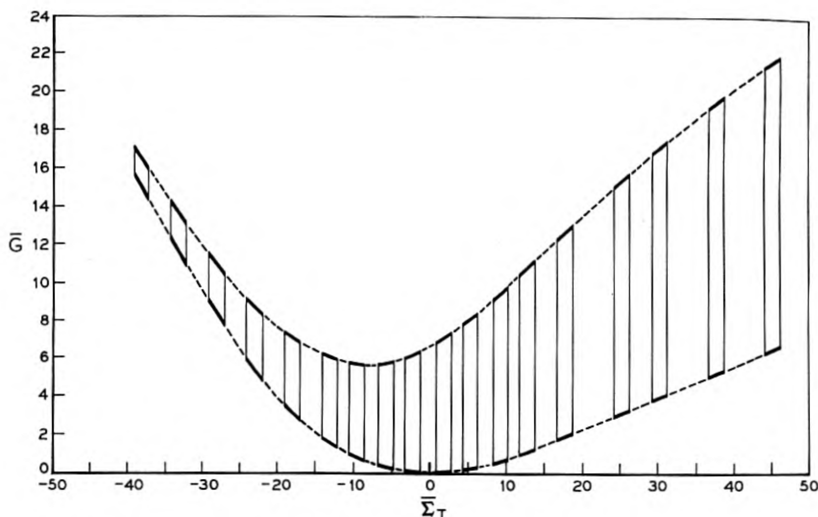


Fig. 4 — Construction of the curve relating $\Delta\bar{G}$ (surface conductivity, in units of $e\mu_p n_i \mathfrak{L}$) and $\bar{\Sigma}$ (surface charge, in units of $en_i\mathfrak{L}$).

to differ from that in zero field, and gives rise to the voltage difference $(\Delta V_3 - \Delta V_1)$ shown in Fig. 3. Expressing this difference in terms of the difference $(\mu_{\text{eff}}^* - \mu_{\text{eff}})$ between the apparent and true effective mobilities in the presence of light (see Section IV), one finds:

$$(\mu_{\text{eff}}^* - \mu_{\text{eff}}) = \left(\frac{w^2 t^2}{I \rho_0^2 C}\right) K \frac{d(1/s)}{d\Sigma} \left(\frac{C_{\text{unit}}}{2w}\right) \tag{3}$$

where K is the constant of proportionality between $(1/s)$ and the photoconductivity signal ΔV_2, and C_{unit} is the capacity per unit of the germanium-gold condenser in the illuminated region, which is 1.27 times the parallel-plate formula. From a series of measurements of $(\mu_{\text{eff}}^* - \mu_{\text{eff}})$ and ΔV_2 it is now possible to obtain $\bar{\Sigma}$ by graphical integration:

$$\Sigma = \left(\frac{1}{en_i\mathfrak{L}}\right) \left(\frac{w^2 t^2}{I \rho_0^2 C}\right) \left(\frac{C_{\text{unit}}}{2w}\right) \int \frac{d\Delta V_2}{\mu_{\text{eff}}^* - \mu_{\text{eff}}} \tag{4}$$

This and the graphical method are of course equivalent. It is worthwhile emphasizing again that either technique depends for its validity on the fact that the distribution of fast states is unaffected by the gas changes in the Brattain-Bardeen cycle, as shown in the experiments of Brown and Montgomery.[4] If, however, the assumption were too far from the truth, the fitting of *both* slopes in Fig. 4 would be impossible. The only

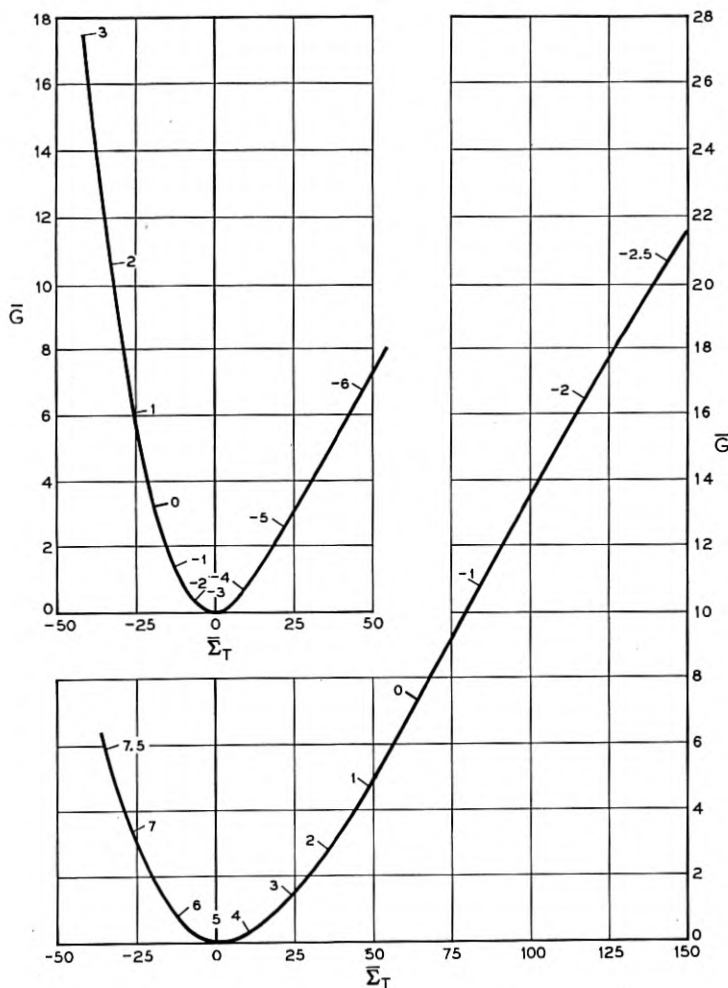


Fig. 5 — Curves showing $\Delta\overline{G}$ (surface conductivity, in units of $e\mu_p n_i \mathfrak{L}$) and $\overline{\Sigma}$ surface charge, in units of $en_i\mathfrak{L}$) for the 22.6 ohm-cm sample (upper curve) and for the 8.1 ohm-cm sample (lower curve). Values of Y, deduced from the surface conductivity, are indicated on the curves.

place at which fitting was at all difficult was at the extreme wet end. For most of the range, therefore, the method is at least internally consistent.

Fig. 5 shows the result of carrying out this procedure for the n and p-type samples. The data were averaged over a number of runs. The numbers appearing on the curves represent values of Y, obtained by reference to Tables V and VI.

From Fig. 5 one may now calculate[4] the changes occurring in $\bar{\Sigma}_s$, the (reduced) charge in fast states, since $\bar{\Gamma}_p - \bar{\Gamma}_n$ may be read from Tables V and VI, and $\bar{\Sigma}_s = \bar{\Sigma} - (\bar{\Gamma}_p - \bar{\Gamma}_n)$. Fig. 6 shows $(\partial\bar{\Sigma}_s/\partial Y)_\delta$ as a function of $Y - \ln \lambda$, calculated from the experimental results in this way. [The reason for plotting against $Y - \ln \lambda$ instead of Y is that this quantity represents the difference, in units of (kT/e), between the electrostatic potential at the surface and the Fermi level. In this way the effects of difference from sample to sample in the position of the Fermi level in the interior are eliminated.] Notice that the measurements of $(\partial\bar{\Sigma}_s/\partial Y)_\delta$ for the two samples have the same general shape, and that the turning points of the two curves occur at about the same value of

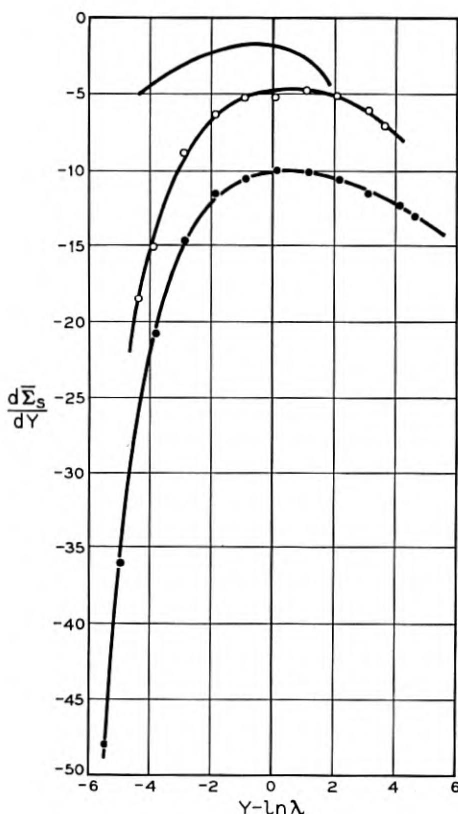


Fig. 6 — Differential charge in fast states versus surface potential. The graphs show $(\partial\bar{\Sigma}_s/\partial Y)$ plotted against $Y - \ln \lambda$. Dots: p-type; circles: n-type. A typical result of Brown and Montgomery, using 28 ohm-cm p-type germanium, is also shown.

$(Y - \ln \lambda)$. Fig. 7 shows the variation of surface recombination velocity with $Y - \ln \lambda$, using the experimental photoconductivity data and values of Y read from Fig. 5. The values of s have been divided by $(\lambda + \lambda^{-1})$, as indicated, since $s/(\lambda + \lambda^{-1})$ is expected to be the same, at a given value of $(Y - \ln \lambda)$, for all samples, so long as the distribution of fast states is the same. The agreement shown in Fig. 7 is probably closer than would be expected in the light of the experimental accuracy.

Fig. 8 shows the observed dependence of $dY/d\delta$ on $(Y - \ln \lambda)$ for both samples, using the data of Tables III and IV, and using the photoconductivity to determine, from Fig. 7, the value of Y at each point. On the figure the expected limiting values ($-\lambda$ and λ^{-1}) are shown for both samples. Of the four asymptotes, the higher limit of $(dY/d\delta)$ for the n-type sample is satisfactorily reached for large negative values of Y; the experimental values for the p-type sample appear to be approaching the expected limit for large positive values of Y, while the information regarding the approach to the two lower limits is too fragmentary to do more than show that the order of magnitude is as expected. Now taking the data shown in Fig. 8, making use of (1) and the calculations given in Tables III and IV, one calculates $(\partial\bar{\Sigma}_s/\partial\delta)_Y/(\partial\bar{\Sigma}_s/\partial Y)_\delta$. The values so found are plotted against Y in Fig. 9. Fig. 6, 7 and 9, showing the observed variation of $(\partial\bar{\Sigma}_s/\partial Y)_\delta$, s and $(\partial\bar{\Sigma}_s/\partial\delta)_Y/(\partial\bar{\Sigma}_s/\partial Y)_\delta$ with Y, furnish a complete description of the properties of the fast states at the

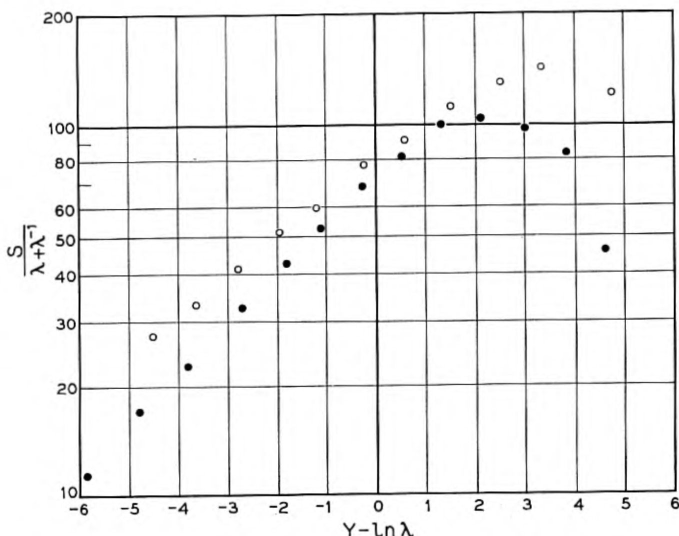


Fig. 7 — Surface recombination velocity versus surface potential. The curves show $s/(\lambda + \lambda^{-1})$ plotted against $Y - \ln \lambda$. Dots: p-type; circles: n-type.

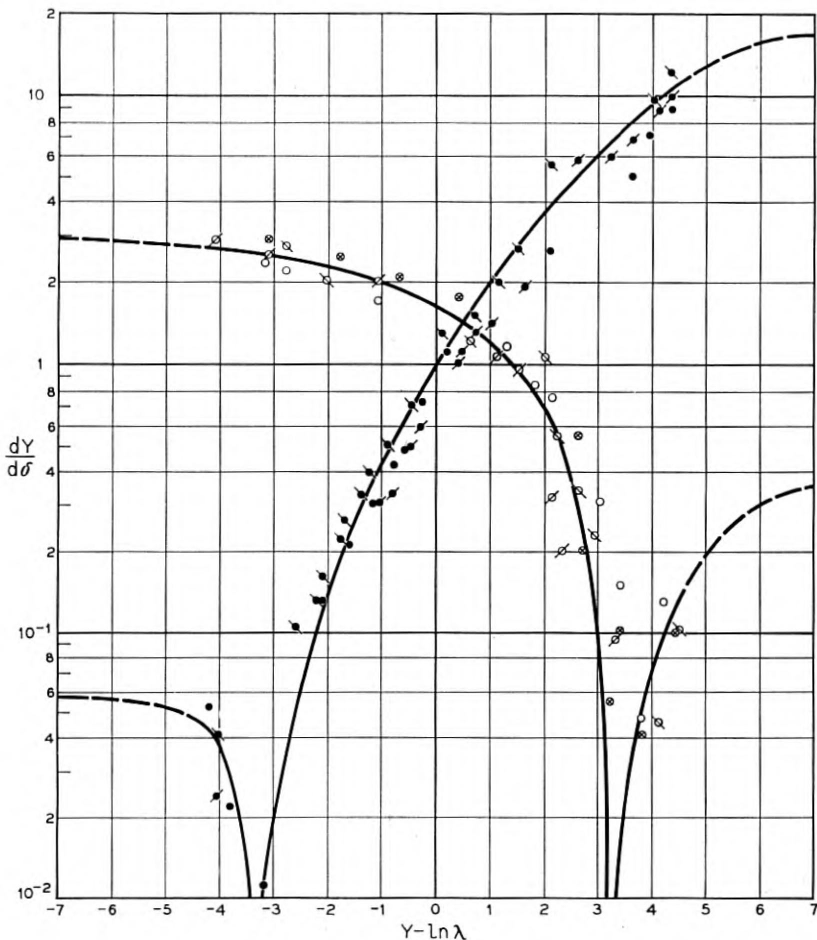


Fig. 8 — Surface photo-voltage (change in contact potential in relation to added carrier concentration). $dY/d\delta$ is shown plotted against $Y - \ln \lambda$. Dots: p-type; circles: n-type. Data from different runs are distinguished by modifications to these symbols. The left-hand branches denote absolute magnitudes, since the ratio is negative there. At the extreme left hand of the diagram, the fast states near to the Fermi level are in good contact with the valence band: at the extreme right hand, to the conduction band. The theoretical asymptotes (λ^{-1} to the left and λ to the right) are also indicated.

temperature studied. This is the basic information which any theoretical treatment must explain. In the succeeding paper this matter is discussed from the point of view of the statistics of a distribution of fast states, and information on the cross sections, as well as on the distribution itself, is derived from the data just presented.

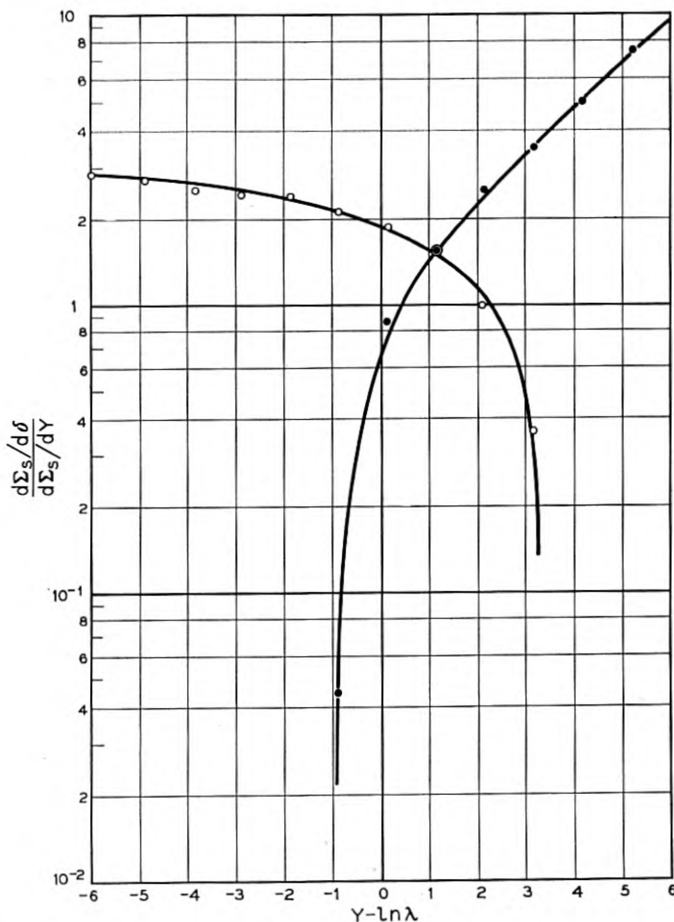


FIG. 9 — The function $(\partial\overline{\Sigma}_s/\partial\delta)Y/(\partial\overline{\Sigma}_s/\partial Y)_\delta$ plotted against $Y - \ln\lambda$. Dots: p-type; circles: n-type.

VI. FURTHER COMMENTS

The development given in the previous section has concerned particularly the properties of the fast states. As to the slow states, the experiments are much less informative. The variations of Y with gas are generally consistent with the variations of contact potential previously reported,[1] although the total range in Y (± 0.13 volt) is smaller by about a factor of 2 than that in contact potential found in the previous work. One must say that roughly half the change of contact potential is in V_B, (i.e., βY) and half in V_D, the potential drop across the ion layer.

It may be seen from the figures that it is the quantity $(Y - \ln \lambda)$, rather than Y, which appears to be characteristic of the point in the cycle reached. This property of a semiconductor surface, and possible reasons therefore, have often been discussed in the literature.[11] The total range of surface potential is illustrated in Fig. 10, which is drawn to scale, and also shows sundry other points of interest found in the present research. The potential diagrams for n-type and p-type are drawn with the Fermi levels aligned, to show the relation between the property $(Y - \ln \lambda) =$ const. and the frequently observed smallness of the contact potential difference between n and p-type germanium.

As to the reproducibility and accuracy of the work presented here, the following points may be of interest: (i) The measurements were repeated on another n-type sample of nearly the same resistivity as the one reported here, but cut from a different crystal. The results on this sample were indistinguishable, within the experimental error, from those found on the first n-type sample. (ii) If the sample was re-etched in precisely the same way as before, and the experiments repeated, the results were in good agreement with those obtained before. However, variations in the etching procedure sometimes gave quite different re-

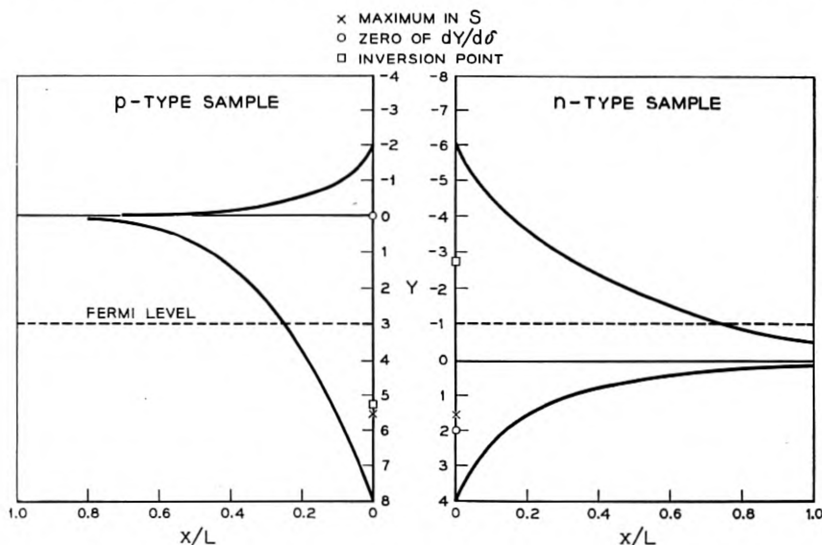


Fig. 10 — The shapes of the surface space-charge regions for the p-type and n-type samples in the extremes of gaseous environment. The two surfaces are to the center of the figure. The solid curves show the center of the gap (intrinsic Fermi level) plotted against distance, in units of an intrinsic Debye length. Also shown are the positions of the zeros of $(dY/d\delta)$, the maxima of s, and the minima of surface conductivity.

sults. We hope to discuss this at a future date. (iii) The accuracy of the measurements is not high. Some of the more directly-derivable quantities, such as s, should be known to 5 per cent, but a quantity like $(\partial\overline{\Sigma}_s/\partial\delta)/(\partial\overline{\Sigma}_s/\partial Y)$, which is only obtained after a long and elaborate calculation involving a number of corrections, is perhaps uncertain to 30 per cent.

VII. CONCLUSIONS

This paper has presented results of combined measurements of field effect, photoconductivity, change of photoconductivity with field, filament lifetime and surface photo-voltage, on slices of germanium. From the measurements, the surface potential Y has been found at each point, and the variations of the quantities $(\partial\overline{\Sigma}_s/\partial Y)$, s and $(\partial\overline{\Sigma}_s/\partial\delta)/(\partial\overline{\Sigma}_s/\partial Y)$ with Y determined.

It is a pleasure to record our thanks to W. L. Brown, for comments on field effect techniques and many stimulating discussions, to H. R. Moore, who constructed the high-voltage power supply, and to A. A. Studna, who assisted in the experiments. We are also grateful to C. Herring for comments on the text.

BIBLIOGRAPHY

1. W. H. Brattain and J. Bardeen, Surface Properties of Germanium, B.S.T.J., **32,** pp. 1–41, Jan. 1953.
2. C. G. B. Garrett and W. H. Brattain, Physical Theory of Semiconductor Surfaces, Phys. Rev., **99,** pp. 376–387, July 15, 1955.
3. W. L. Brown, Surface Potential and Surface Charge Distribution from Semiconductor Field Effect Measurements, Phys. Rev., **98,** p. 1565, June 1, 1955.
4. H. C. Montgomery and W. L. Brown, Field-Induced Conductivity Changes in Germanium, Phys. Rev., **103,** Aug. 15, 1956.
5. J. R. Schrieffer, Effective Carrier Mobility in Surface Charge Layers, Phys. Rev., **97,** pp. 641–646, Feb. 1, 1955.
6. C. G. B. Garrett and W. H. Brattain, Interfacial Photo-Effects in Germanium at Room Temperature, Proc. of the Conference on Photo Conductivity, Nov., 1954, Wiley, in press.
7. W. H. Brattain and C. G. B. Garrett, Surface Properties of Germanium and Silicon, Ann. N. Y. Acad. of Science, **58,** pp. 951–958, Sept., 1954.
8. D. T. Stevenson and R. J. Keyes, Measurements of Surface Recombination Velocity at Germanium Surfaces, Physica, **20,** pp. 1041–1046, Nov., 1954.
9. J. Clerk Maxwell, Electricity and Magnetism, 3rd Edition, **1,** p. 310, Clarendon Press, 1904.
10. W. van Roosbroeck, Theory of Photomagnetoelectric Effect in Semiconductors, Phys. Rev., **101,** pp. 1713–1725, March 15, 1956.
11. J. Bardeen and S. R. Morrison, Surface Barriers and Surface Conduction, Physica, **20,** p. 873, 1954.
12. E. Harnik, A. Many, Y. Margoninski and E. Alexander, Correlation Between Surface Recombination Velocity and Surface Conductivity in Germanium, Phys. Rev., **101,** pp. 1434–1435, Feb. 15, 1956.

Distribution and Cross-Sections of Fast States on Germanium Surfaces

By C. G. B. GARRETT and W. H. BRATTAIN

(Manuscript recieved May 10, 1956)

A theoretical treatment of the field effect, surface photo-voltage and surface recombination phenomena has been carried out, starting with the Hall-Shockley-Read model and generalizing to the case of a continuous trap distribution. The theory is applied to the experimental results given in the previous paper. One concludes that the distribution of fast surface states is such that the density is lowest near the centre of the gap, increasing sharply as the accessible limits of surface potential are approached. From the surface photo-voltage measurements one obtains an estimate of 150 for the ratio (σ_p/σ_n) *of the cross-sections for transitions into a state from the valence and conduction bands, showing that the fast states are largely acceptor-type. On the assumption that surface recombination takes place through the fast states, the cross-sections are found to be:* $\sigma_p \sim 6 \times 10^{-15}$ cm^2 *and* $\sigma_n \sim 4 \times 10^{-17}$ cm^2.

I. INTRODUCTION

The existence of traps, or "fast" states, on a semiconductor surface, becomes apparent from three physical experiments: measurements of field effect,[1] of surface photovoltage,[2] and of surface recombination velocity s. Results of combined measurements of these three quantities on etched surfaces of p- and n-type germanium have been presented in the preceding paper.[3] The present paper is concerned with the conclusions which may be drawn from these experiments as to the distribution in energy of these surface traps, and the distribution of cross-sections for transitions between the traps and the conduction and valence bands.

The statistics of trapping at a surface level has been developed by Brattain and Bardeen[2] and by Stevenson and Keyes,[4] following the work on body trapping centers of Hall[5] and of Shockley and Read.[6]

It is known that surface traps are numerous on a mechanically damaged surface[7] or on a surface that has been bombarded but not annealed;[8]

and that on an etched surface their density is comparatively low. It is also known that the available results cannot be accounted for by a single level, or even two levels, so that one is evidently dealing either with a large number of discrete states or a continuous spectrum. A given trapping centre is completely described by specifying: (i) whether it is donor-like (either neutral or positive) or acceptor-like (neutral or negative); (ii) its position in energy; and (iii) the values for the constants C_p and C_n (related to cross-sections) occurring in the Shockley-Read theory. In this paper we shall deduce what we can about these quantities, using the experimental results previously presented.

At the outset it must be admitted that it is by no means certain that the same set of surface states appear in the field-effect experiment and give rise to surface recombination. However, (i) it is found that such surface treatments as increase s also reduce the effective mobility in the field-effect experiment; (ii) any surface trap must be able to act as a recombination centre, unless one of the quantities C_p and C_n is zero;[9] and (iii) the capture cross-sections obtained by assuming that the field-effect traps are in fact recombination centres are, as we shall see below, eminently reasonable.

As to the nature of the surface traps, not too much can be said at the moment. The lack of sensitivity to the cycle of chemical environment used argues against their being associated with easily desorbable surface atoms; the intrinsically short time constants (Section 5) suggest that they are on or very close to the germanium surface. The possibility that the surface traps are Tamm levels[10] remains; or they could be corners or dislocations. However, the reproducibility with which a given value of s may be obtained by a given chemical treatment of a given sample, followed by exposure to a given ambient, suggests that there is nothing accidental about their occurrence.

II. STATISTICS OF A DISTRIBUTION OF SURFACE TRAPS

We start by quoting results from the work of Shockley and Read[6] and Stevenson and Keyes[4] on the occupancy factor f_t and the flow U of minority carriers (per unit area) into a set of traps having a single energy level and statistical weight unity:

$$f_t = (C_n n_s + C_p p_1)/[C_n(n_s + n_1) + C_p(p_s + p_1)] \qquad (1)$$

$$U = C_n C_p(p_s n_s - n_i^2)/[C_n(n_s + n_1) + C_p(p_s + p_1)] \qquad (2)$$

where the symbols have the following meanings:

n_s, p_s — densities of electrons and holes present at the surface

n_1, p_1 — values which the equilibrium electron and hole densities at the surface would have if the Fermi level coincided with the trapping level

$C_n = N_t v_{Tn} \sigma_n$; $C_p = N_t v_{Tp} \sigma_p$, where N_t stands for density of traps per unit area, v_{Tn} is the thermal speed for electrons and v_{Tp} that for holes, and σ_n and σ_p are the cross-sections for transitions between the traps and the conduction and valence bands respectively.

If we introduce the surface potential Y and the quantity δ, defined as $(\Delta p/n_i)$, where Δp is the added carrier density in the body of the semiconductor, we may write:

$$
\begin{aligned}
n_s &= \lambda^{-1} n_i e^{Y}(1 + \lambda\delta) \\
p_s &= \lambda n_i e^{-Y}(1 + \lambda^{-1}\delta)
\end{aligned}
\tag{3}
$$

where $\lambda = p_0/n_i$, p_0 being the equilibrium hole concentration in the body of the semiconductor. We further introduce the notation:

$$
\begin{gathered}
n_1 = n_i e^{-\nu} \qquad p_1 = n_i e^{\nu} \\
(C_p/C_n)^{\frac{1}{2}} = \chi
\end{gathered}
\tag{4}
$$

The quantity ν thus represents the energy difference, measured in units of (kT/e), between the trapping level and the centre of the gap;* and is positive for states below, negative for those above, this level. The parameter χ will be most directly associated with whether the state is donor-like or acceptor-like. If it is donor-like (neutral or positive), a transition involving an electron in the conduction band will be aided by Coulomb attraction whereas one involving a hole will not; so one would expect $\chi \ll 1$. For an acceptor-like trap, (neutral or negative) the contrary holds, and one expects $\chi \gg 1$.

Using (4), the occupancy factor (1) becomes

$$
\begin{aligned}
f_t &= \frac{\chi^{-1}\lambda^{-1}e^{Y}(1 + \lambda\delta) + \chi e^{\nu}}{\chi^{-1}\lambda^{-1}e^{Y}(1 + \lambda\delta) + \chi^{-1}e^{-\nu} + \chi\lambda e^{-Y}(1 + \lambda^{-1}\delta) + \chi e^{\nu}} \\
&= \tfrac{1}{2}\lambda^{-\frac{1}{2}}e^{-\frac{1}{2}Y}e^{\frac{1}{2}\nu} \operatorname{sech}\left[\tfrac{1}{2}(Y + \nu) - \tfrac{1}{2}\ln\lambda\right] \qquad \text{for } \delta = 0
\end{aligned}
\tag{5}
$$

Note that, in thermodynamic equilibrium, the occupancy factor does not depend in any way on the cross-sections, whereas for $\delta \neq 0$ it does, through the ratio χ.

* Strictly speaking, one should say "position of the Fermi level for intrinsic semiconductor" instead of "centre of the gap." These will fail to coincide if the effective masses of holes and electrons are unequal, as they certainly are in germanium.

Similarly, the flow of carrier-pairs to the surface (2) becomes:

$$U = N_t(v_{Tn}v_{Tp})^{1/2}(\sigma_n\sigma_p)^{1/2}n_i \frac{(\lambda + \lambda^{-1})\delta + \delta^2}{\chi^{-1}\lambda^{-1}e^Y(1 + \lambda\delta)\,\chi^{-1}e^{-\nu} + \chi\lambda e^{-Y}(1 + \lambda^{-1}\delta)\,\chi e^{\nu}} \tag{6}$$

which, for $\delta \rightarrow 0$, tends to the linear law $U = sn_i\delta$, where s, the surface recombination velocity, is given by:

$$s/(v_{Tn}v_{Tp})^{1/2} = N_t S_t$$

where

$$S_t = (\lambda + \lambda^{-1})(\sigma_n\sigma_p)^{1/2}/2[ch(\nu + \ell n\,\chi) + ch(Y - \ell n\,\lambda - \ell n\,\chi)] \tag{7}$$

The surface density Σ_s of trapped charge is given by:

$$\Sigma_s = N_t f_t \tag{8}$$

where f_t is the occupancy factor, given by (5).

Now let us turn to the question of a distribution of surface traps through the energy ν. Suppose that the density of states having ν lying between ν and $\nu + d\nu$ is $\bar{N}(\nu)\,d\nu$, expressed in units $(n_i\mathfrak{L})$. Then the total surface recombination velocity arising from all traps, and the total trapped surface charge density, are given by:

$$s/(v_{Tn}v_{Tp})^{1/2} = n_i\mathfrak{L} \int S_t(\nu)\bar{N}(\nu)\,d\nu \tag{9}$$

$$\bar{\Sigma}_s = \int f_t(\nu)\bar{N}(\nu)\,d\nu \tag{10}$$

where $S_t(\nu)$ and $f_t(\nu)$ are explicit functions of ν, given by (5) and (7). The limits of the integrals in (9) and (10) are the values of ν corresponding to the conduction and valence band edges; however, as we shall see, it is often possible to replace these limits by $\pm\infty$.

In summing up the contributions in the way represented by (9), we have implicitly ignored the possibility of inter-trap transitions, supposing that the population of each trap depends only on the rates of exchange of charge with the conduction and valence bands, and is independent of the population of any other trap of differing energy.

What kind of function do we expect $N(\nu)$ to be? Brattain and Bardeen[2] postulated that $N(\nu)$ was of the form of two delta-functions, corresponding to discrete trapping levels high and low in the band. This assumption is not consistent with the observed facts in regard to field effect, surface

photo-voltage, or surface recombination velocity. The general difficulty is that the observed quantities usually vary less rapidly with surface potential than one would expect. It is possible to fit the field-effect observations of Brown and Montgomery[11] with a larger number of discrete levels, but this would call for a "sharpening up" of the trapped charge distribution as the temperature is lowered, and this appears to be contrary to what is observed.* It is always possible that the surface is patchy, in which case almost *any* variation with mean surface potential could be explained. The simplest assumption, however, seems to be that $N(\nu)$ is a rather smoothly-varying function. All we need assume for the moment is that it is everywhere finite, continuous and differentiable. We may then differentiate equation (10) with respect to Y and δ under the integral sign, and get $(\partial\bar{\Sigma}_s/\partial Y)_\delta$ and $(\partial\bar{\Sigma}_s/\partial\delta)_Y$, the quantities for which experimental measurements were reported in the previous paper:[3]

$$\left(\frac{\partial\bar{\Sigma}_s}{\partial Y}\right)_\delta = \int \frac{\bar{N}(\nu)\, d\nu}{4\, ch^2[\frac{1}{2}(\nu + Y) - \frac{1}{2}\ell\text{n}\, \lambda]} \tag{11}$$

$$\left(\frac{\partial\bar{\Sigma}_s}{\partial \delta}\right)_Y = -\int \frac{\bar{N}(\nu)(\frac{1}{2}(\lambda^{-1} + \lambda)th[\frac{1}{2}(\nu - Y) + \frac{1}{2}\,\ell\text{n}\,\lambda + \ell\text{n}\,\chi] + \frac{1}{2}(\lambda^{-1} - \lambda))\, d\nu}{4ch^2[\frac{1}{2}(\nu + Y) - \frac{1}{2}\,\ell\text{n}\,\lambda]} \tag{12}$$

Notice that the expression in brackets in the numerator of (12) generally has the value λ^{-1} or $-\lambda$, except near the point $\nu = Y - \ell n\,\lambda - 2\ell n\,\chi$. This is indicative of the fact that, whatever the exact form of $\bar{N}(\nu)$, the ratio of $-(\partial\Sigma_s/\partial\delta)_Y/(\partial\Sigma_s/\partial Y)_\delta$ tends to these limiting values (λ^{-1} and $-\lambda$) for sufficiently large negative and positive Y respectively.

It may be verified from (7), (11) and (12) that $(\partial\bar{\Sigma}_s/\partial Y)_\delta$, found from the field effect experiment, depends only on $N(\nu)$; $(\partial\bar{\Sigma}_s/\partial\delta)_Y$, found from the surface photo-voltage, depends on $\bar{N}(\nu)$ and χ; while s, the surface recombination velocity, depends in addition on the geometric mean cross-section $(\sigma_n\sigma_p)^{1/2}$. Both χ and $(\sigma_n\sigma_p)^{1/2}$ might themselves, of course, be functions of ν. Thus relations (7), (11) and (12) are integral equations, from which the three unknown functions of ν may in principle be deduced from the experimental results. (Equation 11, in fact, may be solved explicitly. P. A. Wolff[17] has shown, however, that, to determine $N(\nu)$ unambiguously, it is necessary to know $(\partial\bar{\Sigma}_s/\partial Y)_\delta$ for all values of Y in the range $\pm\infty$.)

The foregoing considerations apply to "small-signal" measurements.

* There are some changes with temperature, but not what one would expect if there were only discrete surface states.

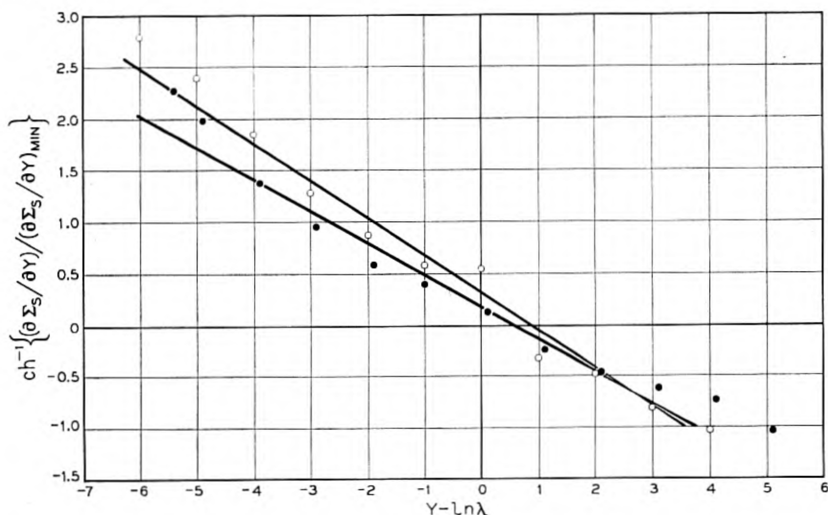


Fig. 1 — The fit between Equations (13) and (14) and the experimental data. The circles and dots give the experimental data for the n and p-type samples respectively and the solid straight lines represent Equations (13) and (14).

But it is also possible, once $N(\nu)$, χ and $(\sigma_n\sigma_p)^{1/2}$ are known, to calculate the expected behavior of the surface photo-voltage and surface recombination rate at high light intensities, and compare the answer with the experimental findings. We hope to discuss this matter in a later paper.

III. ANALYSIS OF THE EXPERIMENTAL DATA BY USE OF THE DELTA-FUNCTION APPROXIMATION

Let us first consider the interpretation of our field effect measurements by means of (11). We start by finding empirical expressions that describe the observed dependence of $(\partial\bar{\Sigma}_s/\partial Y)$ on Y (Fig. 6 of the preceding paper[3]). Except at values of $(Y - \ell n\ \lambda)$ close to the extremes reached one may fit quite well by a hyperbolic cosine function. Fig. 1 shows the function whose hyperbolic cosine is $(\partial\Sigma_s/\partial Y)/(\partial\Sigma_s/\partial Y)_{\min}$ plotted against $Y - \ell n\ \lambda$. From this figure we find:

22.6 ohm-cm n-type:

$$\left(\frac{\partial\bar{\Sigma}_s}{\partial Y}\right)_\delta = 4.5ch[0.36(Y - \ell n\ \lambda) - 0.8] \tag{13}$$

$$(\text{for } (Y - \ell n\ \lambda) > -4$$

8.1 ohm-cm p-type:

$$\left(\frac{\partial\bar{\Sigma}_s}{\partial Y}\right)_\delta = 9.7ch[0.31(Y - \ell n\,\lambda) - 0.5] \tag{14}$$

$$\text{for } 2 > (Y - \ell n\,\lambda) > -4$$

For values of $(Y - \ell n\,\lambda)$ less than -4, it appears that Σ_s is changing more rapidly with Y than is indicated by (13) and (14). We shall comment on this point later. Excluding this region, we note that in both cases the variation with Y is everywhere slow in comparison with e^Y, and proceed on the assumption that $\bar{N}(\nu)$ is a function of ν that varies everywhere slowly in comparison with e^ν. Then (11) indicates that there is one fairly sharp maximum in the integrand in the range $\pm\infty$, occurring at that value of ν which coincides with the Fermi level:

$$\nu = -Y + \ell n\,\lambda \tag{15}$$

The integral in (11) could be evaluated in series about this point (method of steepest descents). The zero-order approximation is got by replacing

$$\tfrac{1}{4}\,\text{sech}^2\,[\tfrac{1}{2}(\nu + Y) - \tfrac{1}{2}\ell n\,\lambda] \qquad \text{by} \qquad \boldsymbol{\delta}(\nu + Y - \ell n\,\lambda).$$

Later we shall proceed to an exact solution, and we shall find that this delta-function approximation is not too bad. From (11) we now find:

$$\left(\frac{\partial\bar{\Sigma}_s}{\partial Y}\right)_\delta \sim \int \bar{N}(\nu)\boldsymbol{\delta}(\nu + Y - \ell n\,\lambda)\,d\nu = \dot{\bar{N}}(-Y + \ell n\,\lambda) \tag{15}$$

This mathematical procedure will be seen to be equivalent to identifying $(\partial\bar{\Sigma}_s/\partial Y)_\delta$ with the density of states at the point in the gap which coincides with the Fermi-level at the surface. Using (13) and (14), one gets:

22.6 ohm-cm n-type:

$$\bar{N}(\nu) = 4.5\,ch(0.36\nu + 0.8) \tag{16}$$

8.1 ohm-cm p-type:

$$\bar{N}(\nu) = 9.7\,ch(0.31\nu + 0.5) \tag{17}$$

As we shall see in the next section, the exact solutions differ from (16) and (17) only in the coefficients preceding the hyperbolic cosines.

Turning to the surface photo-voltage measurements, we take (12) and again replace

$$\tfrac{1}{4}\,\text{sech}^2\,[\tfrac{1}{2}(\nu + Y) - \tfrac{1}{2}\ell n\,\lambda] \qquad \text{by} \qquad \boldsymbol{\delta}(\nu + Y - \ell n\,\lambda)$$

Using (15), one gets:

$$-\frac{(\partial\bar{\Sigma}_s/\partial\delta)_Y}{(\partial\bar{\Sigma}_s/\partial Y)_\delta} \tag{18}$$
$$= \tfrac{1}{2}(\lambda^{-1} + \lambda)\, th(-Y + \ell n\, \lambda + \ell n\, \chi) + \tfrac{1}{2}(\lambda^{-1} - \lambda)$$

This procedure, inaccurate as it is, has the advantage that no particular assumption need be made concerning the functional dependence of χ on ν, it being understood that χ in (18) has the value holding for $\nu = -Y + \ell n\, \lambda$. In particular, if Y_0 is that value of Y at which the ratio $-(\partial\bar{\Sigma}_s/\partial\delta)_Y/(\partial\bar{\Sigma}_s/\partial Y)_\delta$ changes sign,

$$\ell n\, \chi_0 = Y_0 - \ell n\, \lambda + th^{-1}[(\lambda - \lambda^{-1})/(\lambda + \lambda^{-1})] \tag{19}$$

From the experimental data, one finds, for the n-type sample, $\ell n\, \chi_0 \sim 2.4$ (at $\nu = -3.5$); for the p-type sample, $\ell n\, \chi_0 \sim 1.0$ (at $\nu = 1.9$).

In view of the approximations made, these estimates would not be expected to be more precise than ± 1 to 2 units. Notice that both values are positive, and that the difference between them is small in comparison with the difference in ν. This suggests that we start afresh with the assumption that χ is independent of ν, and work out the surface photovoltage integral exactly. This is done in the next section.

IV. EXACT TREATMENT FOR THE CASE $\bar{N}(\nu) = A\, ch\, (q\nu + B)$, WITH CONSTANT CROSS-SECTIONS

The results of the previous section suggest the procedure of assuming that $N(\nu)$ is of the functional form given by (16) and (17), and evaluating the integrals (9), (11) and (12) exactly. The integral for $(\partial\bar{\Sigma}_s/\partial Y)$, (11), depends only on the form of $N(\nu)$ and may be evaluated at once. To get $(\partial\bar{\Sigma}_s/\partial\delta)$, (12), one must know how χ depends on ν. On the basis of the work of the previous section, we shall suppose that χ is independent of ν. (Properly, we need only assume that χ varies with ν more slowly than e^ν. Since the function $th[\frac{1}{2}(\nu - Y) + \frac{1}{2}\ell n\, \lambda + \ell n\, \chi]$ has one of the values ± 1 everywhere except close to $\nu = Y - \ell n\, \lambda - 2\ell n\, \chi$, and since the denominator of (12) has a sharp minimum at $\nu = -Y + \ell n\, \lambda$, it follows that the region in which $(\partial\bar{\Sigma}_s/\partial\delta)_Y$ changes sign will be governed mainly by the value of χ at $\nu = -\ell n\, \chi$.) To get s [(9), using (7)], one must also assume something about the geometric mean cross-section, $(\sigma_n\sigma_p)^{1/2}$. In the absence of any information on this score, we shall assume that $(\sigma_n\sigma_p)^{1/2}$ also is independent of ν, and see how the computed variation of s with Y compares with the experimental results.

We assume:

$$\bar{N}(\nu) = A\ ch\ (q\nu + B) \tag{20}$$

and substitute in (11), (12) and (7). In view of the sharp maximum in the integrands of these expressions, it is permissible to set the limits which should correspond to the edges of the gap or of the state distribution equal to $\pm\infty$. The integrals are conveniently evaluated by the contour method (see Appendix 1) and yield the following results:

$$\left(\frac{\partial \bar{\Sigma}_s}{\partial Y}\right)_\delta = A\pi q \text{ cosec } \pi q\ ch\ [B - q(Y - \ell n\ \lambda)] \tag{21}$$

$$\left(\frac{\partial \bar{\Sigma}_s}{\partial \delta}\right)_Y = -A\pi q \text{ cosec } \pi q\ ch\ [B - q(Y - \ell n\ \lambda)] \times$$

$$\left[\tfrac{1}{2}(\lambda^{-1} + \lambda)\left(-\coth \mathcal{Y} + \frac{sh\ q\mathcal{Y}\ ch\ \mathcal{B}}{q\ sh^2\,\mathcal{Y}\ ch\ (q\mathcal{Y} - \mathcal{B})}\right) + \tfrac{1}{2}\ (\lambda^{-1} - \lambda)\right] \tag{22}$$

where

$$\left.\begin{aligned} \mathcal{Y} &= Y - \ell n\, \lambda - \ell n\, \chi \\ \mathcal{B} &= B - q\ \ell n\, \chi \end{aligned}\right\} \tag{23}$$

$$\frac{s}{(v_{Tn}v_{Tp})^{1/2}} \tag{24}$$

$$= \tfrac{1}{2}\ (\lambda + \lambda^{-1})(\sigma_n\sigma_p)^{1/2} n_i \mathcal{L}\ 2\pi\ A\ sh\ q\mathcal{Y}\ ch\ \mathcal{B} \text{ cosec } \pi q \text{ cosech } \mathcal{Y}$$

Comparing (21) with (15), we see that the delta-function approximation is in error to the extent that it replaces πq cosec πq by 1. With the value of q found experimentally, this is not too bad; we can now, however, by fitting the right-hand side of (21) to the experimental facts, (13) and (14), obtain exact solutions for $N(\nu)$:

22.6 ohm-cm n-type

$$N(\nu) = 3.6\ ch(0.36\nu + 0.8) \qquad (\text{for } \nu < 4)$$

8.1 ohm-cm p-type (25)

$$N(\nu) = 8.3\ ch(0.31\nu + 0.5) \qquad (\text{for } \nu < 4)$$

The question arises as to whether this solution for the distribution is unique. We have already pointed out that the mathematical methods fail if the distribution is discontinuous. It seems that (25) represents the only solution that is slowly-varying, in the sense used in the previous section; its correctness could presumably be checked by carrying out experiments at different temperatures. For $\nu > 4$, the above expressions

do not fit the observed facts, because, for $Y - \ell n \lambda < -4$, the charge in fast states is found to change more rapidly than is given by the empirical expressions in (13) and (14). The behaviour in this region is perhaps indicative of the existence of a discrete trapping level just beyond the range of ν which can be explored by our techniques. The observations (see Fig. 6 of preceding paper[3]) can be described by postulating, in addition to the continuous distribution of states given above, a level of density about 10^{11} cm^{-2}, situated at $\nu = 6$, or a higher density still further from the center of the gap. Statz et al,[13] using the "channel" techniques, which are valuable for exploring the more remote parts of the gap, have proposed a level of density $\sim 10^{11}$ cm^{-2}, situated at about 0.14 volts below the center of the gap ($\nu = 5.5$): this is not in disagreement with the foregoing.

In order to compare (22) with the experimental data derived from the surface photo-voltage, it is necessary to choose a value for χ. Fig. 2 shows the comparison with the results presented in the preceding paper. On the vertical axis, the values of $(\partial\Sigma_s/\partial\delta)/(\partial\Sigma_s/\partial Y)$ plotted have been divided by $(\lambda + \lambda^{-1})$, in order to show the n and p-type results on the same scale. (Note that the limiting values of this quantity should be $\lambda/(\lambda + \lambda^{-1})$ and $-\lambda^{-1}/(\lambda + \lambda^{-1})$, so that the vertical distance between the limiting values should be 1, independent of λ). The theoretical curves have been drawn with the value $\ell n \chi = 2.5$, in order to give best fit between theory and experiment at the points at which the ordinate changes sign. (It may be seen from the form of (22) that, with the actual value of the other parameters, the main effect of adopting a different value of $\ell n \chi$ would be to shift the theoretical curve horizontally, while a change of λ shifts it *vertically* without in either case greatly modifying its shape). The fit between theory and experiment is not quite as good as could be expected, even taking into account the rather low accuracy of the measurements. The variation of $(\partial\Sigma_s/\partial\delta)/(\partial\Sigma_s/\partial Y)$ with Y found experimentally seems to be rather slower than the theory would lead one to expect. The main points to make are: (i) the difference in Y between the zeros for the two samples (5.4 ± 1) is about what it should be (4.8) on the assumption that $\ell n \chi$ is the same for both samples and of the order of unity; and (ii) paying attention mainly to the zeros, the estimate $\ell n \chi = 2.5$ is likely to be good to ± 1.

Now let us consider the surface recombination velocity. Here we are on somewhat shakier ground, in that, in deriving (24), we have had to assume not only that χ is independent of ν, but $(\sigma_n \sigma_p)^{1/2}$ also. First we note from (24) that the maximum value of s should occur at $Y - \ell n \lambda = \ell n \chi$. Comparing with the experimental results given in the preceding paper,

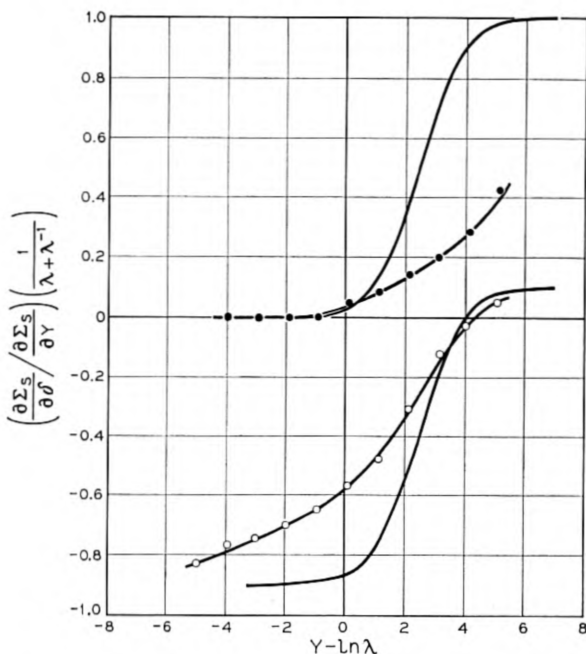


Fig. 2 — Experiment and theory for

$$\left[\left(\frac{\partial \Sigma_s}{\partial \delta}\right) \Big/ \left(\frac{\partial \Sigma_s}{\partial Y}\right)\right] \Big/ \frac{1}{\lambda + \lambda^{-1}}$$

Solid lines theory; circles and dots, with smooth curves through the points, represent experimental results for n and p-type samples, respectively.

we see maxima at $Y - \ell n\, \lambda = 2.0$ for the p-type sample, and 3.5 for the n-type sample. Both these values are within the limits to $\ell n\, \chi$ given in the previous paragraph, thus confirming the estimate made there. Fig. 3 shows a comparison between the experimental results and (24). The graph has been fitted horizontally, by setting $\ell n\, \chi = 2.5$, as found above; vertically, to agree with the mean value at that point. The agreement with experiment is reasonable, although again, just as in Fig. 2, the experimental variation of s with $(Y - \ell n\, \lambda)$ is rather slower than one would expect.

The fact that the experimental values, both of surface photo-voltage and of surface recombination velocity, vary more slowly than expected, is susceptible of a number of interpretations: (i) The deduced distribution of fast states might be wrong. However, the most likely alternative distributions — isolated levels, or a completely uniform distribution —

give (in at least some ranges of Y) a more rapid instead of a smoother variation of these quantities so long as the surface is homogeneous. (ii) The estimates of the changes in Y might be too large. It is unlikely that our calibration is sufficiently in error, and other workers have obtained results comparable to ours. The only possibility would be that the mobility of carriers near the surface is larger (instead of smaller, as found by Schrieffer) than inside — which seems quite out of the question. (iii) The ratio of capture cross-sections varies with ν. This, however, would only be in the right direction if one were to assume that the ratio χ *increases* with the height of the level in the gap — i.e., that the high states behave like acceptors, and the low ones like donors. While not quite impossible, this is an unlikely result. (iv) The surface is patchy. It is probable that a range of variation of two to four times (kT/e) in surface potential would be sufficient to account for the observed slow variation of surface photo-voltage and recombination velocity with mean surface potential. We have refrained from detailed calculations of patch effects, on the grounds that, without detailed knowledge of the magnitude and distribution of the patches, it would be possible to construct a model that could indeed fit the facts, but one would have little confidence in the result. The possibility of patches warns us to view with caution the exact distribution function deduced for the fast states. It would still be conceivable, for example, that one has but two discrete states, as originally proposed by Brattain and Bardeen,[2] and that the apparent existence of a band of states in the middle of the gap arises from the fact that there are always some parts of the surface at which the Fermi level is close to one or other of these states. Fortunately the conclusions as to the cross-sections are not too sensitive to the exact distribution function assumed.

Using the mean of the two coefficients in (25), substituting $n_i = 2.5 \times 10^{13}$ cm^{-3}, $\mathfrak{L} = 1.4 \times 10^{-4}$ cm, $(v_{Tn}v_{Tp})^{1/2} = 1.0 \times 10^7$ cm/sec, in (24), and using the experimental result (see Fig. 3) that $s_{\max}/(\lambda + \lambda^{-1}) = 1.2 \times 10^2$ cm/sec, one obtains $(\sigma_p\sigma_n)^{1/2} = 5 \times 10^{-16}$ cm^2. Now setting $(\sigma_p/\sigma_n) = \chi^2 \sim e^5 \sim 150$, one gets for the separate cross-sections:

$$\sigma_p = 6 \times 10^{-15} \text{ cm}^2$$

$$\sigma_n = 4 \times 10^{-17} \text{ cm}^2$$

There values appear to be eminently reasonable. Burton et al,[12] who studied recombination through body centres associated with nickel and copper in germanium, found $\sigma_p > 4 \times 10^{-15}$ cm^2, $\sigma_n = 8 \times 10^{-17}$ cm^2 for nickel, and $\sigma_p = 1 \times 10^{-16}$, $\sigma_n = 1 \times 10^{-17}$ for copper. The fact that

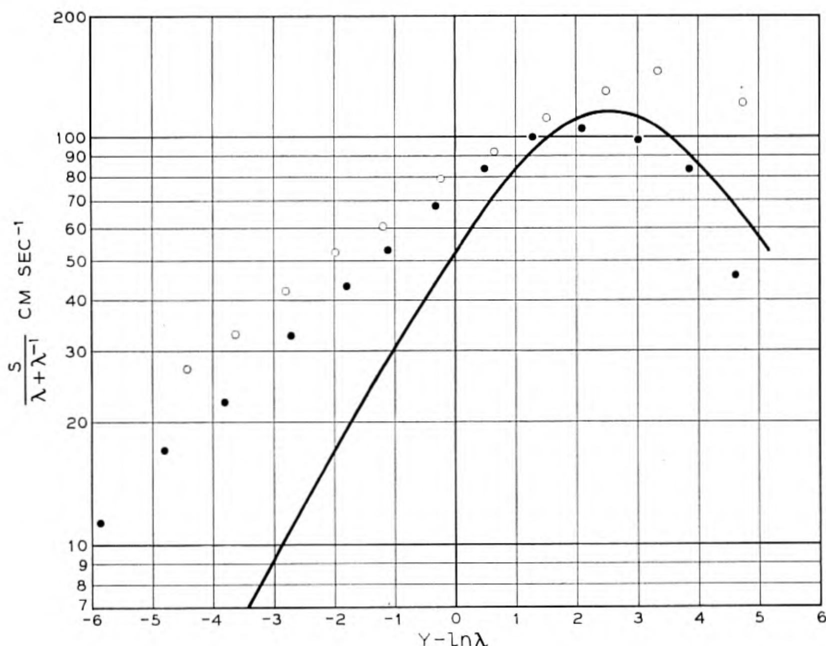


Fig. 3 — Experiment and theory for surface recombination. Solid curve theory circles and dots for n and p-type samples, respectively.

our estimates for σ_p and σ_n appear to be of the expected order of magnitudes lends strong support to the view that identifies the traps appearing in the field-effect and surface photo-voltage experiments with those responsible for surface recombination.

The result that $(\sigma_p/\sigma_n) = 150$ is good evidence that the fast states are acceptor-like. This statement must be restricted to the range $|\nu| < 4$; the states that are outside this range might be of either type. Also one might allow a rather small fraction of the states near the middle to be donor-type, without serious trouble; but the experimental results compel one to believe that most of the fast states within 0.1 volts or so of the centre of the gap are acceptor-like.

V. TRAPPING KINETICS

The foregoing considerations have concerned the steady-state solution to the surface trapping problem. If the experimental constraints are changed sufficiently rapidly, however, there may be effects arising from the finite time required for the charge in surface states to adapt itself

to the new conditions.[14] This section will concern itself with the trapping time constants (which are not directly related to the rate of recombination of minority carriers).

One case of trapping kinetics has been discussed by Haynes and Hornbeck.[9] A general treatment of surface trapping kinetics is necessarily quite involved, and will be taken up in a future paper. Here we shall restrict ourselves to giving an elementary argument relating to the high-frequency field effect experiment of Montgomery.[15] To simplify the discussion, we assume that the surface in question is of the "super" type; i.e., the surface excess of the bulk majority carrier is large and positive. At time $t = 0$, a large field is suddenly applied normal to the surface; the induced charge appears initially as a change in the surface excess of the bulk majority carrier; as time elapses, charge transfer between the space-charge region and the fast states takes place, until equilibrium with the fast states has been re-established. What time constant characterizes this process?

Take electrons as the majority carrier. Then the flow of electrons into the fast states must equal the rate of decrease of the surface excess of electrons. For a single level one may write:

$$\begin{aligned} U_n &= N_t v_{Tn} \sigma_n [(1 - f_t) n_s - f_t n_1] \\ &= -\dot{\Gamma}_n \end{aligned} \tag{26}$$

For a continuous distribution of levels, one can say that only those levels within a few times (kT/e) of the Fermi level at the surface will be effective, so that one may regard the distribution as being equivalent to a single state with $n_1 = n_i \exp(Y - \ln \lambda)$, which will be about half full. We assume further that the density of fast states is sufficient for the changes in Γ_n to be large in comparison with those in f_t, as is reasonable, having regard to the relative magnitudes of the measured values of $(\partial \Sigma_s / \partial Y)_\delta$ found in the present research, and of $(\partial \Gamma_p / \partial Y)_\delta$ and $(\partial \Gamma_n / \partial Y)_\delta$. Thus f_t may be treated as a constant in equation (26). Further, we may set $n_s = 4\Gamma_n^2 / n_i \mathfrak{L}^2$, as may be proved from considerations on the space-charge region.[16] Solving (26) with these conditions, one finds, for the transient change in Γ_n between the initial and the quasi-equilibrium state:

$$\Delta\Gamma_n \propto \left(1 - th \frac{t}{\tau}\right) \tag{27}$$

where

$$\tau = \lambda e^{-Y} \mathfrak{L} / [N_t v_{Tn} \sigma_n \sqrt{2} \sqrt{f_t(1 - f_t)}]$$

To clarify the order of magnitude of time constant involved, let us substitute $\mathfrak{L} \sim 10^{-4}$ cm, $N_t \sim 10^{11}$ cm^{-2}, $v_{Tn} \sim 10^7$ cm/sec., $\sigma_n \sim 10^{-15}$ cm^2, $f_t \sim 0.5$, $\lambda e^{-Y} \sim 1$. This gives $\tau \sim 10^{-7}$ sec, which suggests that one would be unlikely to run into trapping time effects in the field-effect experiment at frequencies less than 10 Mcyc/sec. This conclusion is consonant with the findings of Montgomery.[15]

Appendix 1

EVALUATION OF THE INTEGRALS IN SECTION 4

The integrals occurring in Section 4, giving the experimentally accessible quantities $(\partial\bar{\Sigma}_s/\partial Y)$, $(\partial\bar{\Sigma}_s/\partial\delta)$ and s in terms of the surface trap distribution and cross-sections, are conveniently evaluated by contour integration. In view of the general applicability of this method in dealing with integrals of the sort that arise from such a distribution of traps, we include here a short note on the precedure used. The integrals needed are:

$$I_1 = \int_{-\infty}^{+\infty} ch(cx + g) \operatorname{sech}^2 x \, dx$$

$$I_2 = \int_{-\infty}^{+\infty} th(x + b) \, ch(cx + g) \operatorname{sech}^2 x \, dx$$

$$I_3 = \int_{-\infty}^{+\infty} \frac{ch(cx + g)}{chx + chk} \, dx$$

To evaluate I_1, we evaluate $\int ch(cz + g) \operatorname{sech}^2 z \, dz$ around the contour shown in Fig. 4. The contributions from the parts $z = \pm R$ vanish in the limit $R \to \infty$, so that the integral has the value:

$$(1 - \cos c\pi) \int_{-\infty}^{+\infty} ch(cx + g) \operatorname{sech}^2 x \, dx - i \sin c\pi \int_{-\infty}^{+\infty} sh(cx + g) \operatorname{sech}^2 x \, dx$$

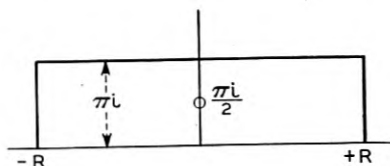


Fig. 4 — Evaluation of I_1.

The integrand has one pole within the contour, at $x = \frac{1}{2}i\pi$, at which the residue is $-c(\cos \frac{1}{2}c\pi\ sh\, g + i \sin \frac{1}{2}c\pi\ ch\, g)$. Multiplying by $2\pi i$ and equating the real part to that in the above expression, one obtains:

$$I_1 = \pi c \operatorname{cosec} \tfrac{1}{2}c\pi\ ch\, g$$

The same contour is used in evaluating I_2; there are now poles at $z = \frac{1}{2}i\pi$ and at $z = \frac{1}{2}i\pi - b$, and one obtains:

$$I_2 = \pi c \coth b\ ch\, g \operatorname{cosec} \tfrac{1}{2}c\pi - 2\pi \operatorname{cosec} \tfrac{1}{2}c\pi \operatorname{cosech}^2 b\ sh\, \tfrac{1}{2}bc\ ch(\tfrac{1}{2}bc - g)$$

To evaluate I_3, one integrates $\int [ch(cz + g)/(chz + chk)]\, dz$ around the contour shown in Fig. 5. There are poles at $i\pi \pm k$. Proceeding as before, one finds:

$$I_3 = 2\pi\ sh\, ck\ ch\, g \operatorname{cosec} \pi c \operatorname{cosech} k$$

Appendix 2

LIMITATION OF SURFACE RECOMBINATION ARISING FROM THE SPACE-CHARGE BARRIER

The question of the resistance to flow of carriers to the surface arising from the change in potential across the space-charge layer has been discussed by Brattain and Bardeen.[2] Here we shall recalculate this effect by a better method, which again shows that, within the range of surface potential studied, the effect of this resistance on the surface recombination velocity is for etched surfaces quite negligible.

Let I_p and I_n be the hole and electron (particle) currents towards the surface, and let x be the distance in a direction perpendicular to the surface, measuring x positive outwards. Then the gradient of the quasi-Fermi levels φ_p and φ_n at any point is given by:

$$\nabla\varphi_{\substack{p\\n}} = \mp (I_{\substack{p\\n}}/\mu_{\substack{p\\n}})/\binom{p}{n} \tag{1}$$

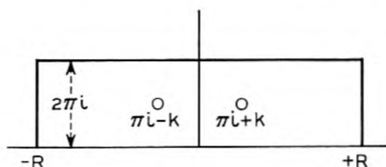


Fig. 5 — Evaluation of I_3.

Then the total *additional* change in φ_p and φ_n across the space-charge region, arising from the departure in uniformity in the carrier densities p and n, is:

$$\begin{aligned} \Delta\phi_p &= -\frac{I_p}{\mu_p}\int\left(\frac{1}{p}-\frac{1}{p_0}\right)dx \\ \Delta\varphi_n &= \frac{I_n}{\mu_n}\int\left(\frac{1}{n}-\frac{1}{n_0}\right)dx \end{aligned} \tag{2}$$

Suppose now that the true surface recombination rate is infinite, so that the quasi-Fermi levels must coincide at the surface, and:

$$\varphi_p + \Delta\varphi_p = \varphi_n + \Delta\varphi_n \tag{3}$$

These equations, together with the known space-charge equations,[16] complete the problem. Notice first, from (2), that $\Delta\varphi_p$ will be large only if there is a region in which p is small ($Y \gg 1$), while $\Delta\varphi_n$ is large only when, in some region, n is small ($Y \ll -1$). Introducing the quantity δ, approximating for δ small, equating I_p and I_n and setting the result equal to $sn_i\delta$, one finds:

$Y \ll -1$

$$s \to (D_n/\mathfrak{L})(\lambda^{1/2} + \lambda^{-3/2})e^{\frac{1}{2}Y}$$

$Y \gg 1$ (4)

$$s \to (D_p/\mathfrak{L})(\lambda^{-1/2} + \lambda^{3/2})e^{-\frac{1}{2}Y}$$

The coefficients $(D_n/\mathfrak{L})$ and $(D_p/\mathfrak{L})$ are of the order of 4×10^5 cm/sec. The most extreme case encountered in our work is that occurring at the ozone extreme for the n-type sample ($\lambda = 0.34$, $Y = -6$), for which the surface recombination velocity, if limited by space-charge resistance alone, would be about one-quarter of this (10^5 cm/sec). The fact that the observed surface recombination velocity is lower than that by more than two orders of magnitude shows that space-charge resistance is not a limiting factor in the present experiments. Equations 4 might well hold on a sand-blasted surface, however, where the trap density is much higher.

REFERENCES

1. W. L. Brown, Surface Potential and Surface Charge Distribution from Semi-Conductor Field Effect Measurements, Phys. Rev. **98,** p. 1565, June 1, 1955.
2. W. H. Brattain and J. Bardeen, Surface Properties of Germanium, B.S.T.J., **32,** pp. 1–41, Jan., 1953.
3. W. H. Brattain and C. G. B. Garrett, page 1019 of this issue.

4. D. T. Stevenson and R. J. Keyes, Measurements of Surface Recombination Velocity at Germanium Surfaces, Physica, **20,** pp. 1041–1046, Nov. 1954.
5. R. N. Hall, Electron-Hole Recombination in Germanium, Phys. Rev., **87,** p. 387, July 15, 1952.
6. W. Shockley and W. T. Read, Jr., Statistics of the Recombination of Holes and Electrons, Phys. Rev., **87,** pp. 835–842, Sept. 1, 1952.
7. T. M. Buck and F. S. McKim, Depth of Surface Damage Due to Abrasion on Germanium, J. Elec. Chem. Soc., in press.
8. H. H. Madden and H. E. Farnsworth, Effects of Ion Bombardment Cleaning and of Oxygen Adsorption on Life Time in Germanium, Bull. Am. Phys. Soc., II, **1,** p. 53, Jan., 1956.
9. J. A. Hornbeck and J. R. Haynes, Trapping of Minority Carriers in Silicon. I. P-Type Silicon. II. N-Type Silicon, Phys. Rev., **97,** pp. 311–321, Jan. 15, 1955, and **100,** pp. 606–615, Oct. 15, 1955.
10. Ig. Tamm, Uber eine mögliche Art der Elektronenbindung an Kristallober flächen, Phy. Zeits. Sowj., **1,** pp. 733–746, June, 1932.
11. H. C. Montgomery and W. L. Brown, Field-Induced Conductivity Changes in Germanium, Phys. Rev., **103,** Aug. 15, 1956.
12. J. A. Burton, G. W. Hull, F. J. Morin and J. C. Severiens, Effects of Nickel and Copper Impurities on the Recombination of Holes and Electrons in Germanium, J. Phys. Chem., **57,** pp. 853–859, Nov. 1953.
13. H. Statz, G. A. deMars, L. Davis, Jr., and H. Adams, Jr., Surface States on Silicon and Germanium Surfaces, Phys. Rev., **101,** pp. 1272–1281, Feb. 15, 1956.
14. C. G. B. Garrett, The Present Status of Fundamental Studies of Semiconductor Surfaces in Relation to Semiconductor Devices, Proc. West Coast Electronics Components Conf. Los Angeles, pp. 49–51, June, 1955.
15. H. C. Montgomery and B. A. McLeod, Field Effect in Germanium at High Frequencies, Bull. Am. Phys Soc., II, **1,** p. 53, Jan., 1956.
16. C. G. B. Garrett and W. H. Brattain, Physical Theory of Semiconductor Surfaces, Phys. Rev., **99,** pp. 376–387, July 15, 1955.
17. P. A. Wolff, Private communication.

Transistorized Binary Pulse Regenerator

By L. R. WRATHALL

(Manuscript received March 14, 1956)

A simple transistorized device has been constructed for amplifying and regenerating binary code signals as they are transmitted over substantial lengths of transmission line. By the use of simple circuitry, means are provided whereby the distortion in the output of one repeater due to low frequency cutoff is compensated in the next repeater. Furthermore, the repeater is effectively and simply timed from its own regenerated output. A brief discussion of the theory of the circuit is presented along with measured results and oscillograms showing its performance. The effects of extraneous interference on the production of errors in such a repeater are reported. These results are in substantial agreement with theory.

1. INTRODUCTION

Long distance communication using digital transmission is not new but was used by man in his earliest communication system. In fact, his first successful electrical system, the telegraph, made use of binary pulse codes. It was not until the invention of the telephone that the emphasis was shifted from the digital to carrier and voice systems. During recent years the development of new electronic devices and techniques have brought digital transmission into the picture again, and it now seems possible to use it not only for telephony but for television as well. Future systems will probably make use of the binary code, this choice being dictated by circuit simplicity and performance.

The fundamental requirement for perfect binary transmission is to be able to detect the presence or absence of a pulse in each of a regular set of discrete time intervals. From this requirement the principal advantages of such a system may be tabulated. First, a pulse can be recognized in the presence of large amounts of interference. Second, when a pulse is recognized it can be faithfully regenerated, suppressing the effect of the interfering noise to any desired degree. Third, simple high-efficiency non-linear devices such as multivibrators or blocking oscillators can be used to regenerate the pulses. The great disadvantage,

common to all pulse systems is the large bandwidth required for transmission.

On wire lines this large transmission band will create a number of problems. The phase-loss variations, crosstalk and temperature effects will be greatly increased over the transmission band as compared to that of the more conventional systems. It can be shown however that if the repeater spans are made sufficiently short these problems will largely disappear. Only rough equalization will be needed, crosstalk and temperature effects become negligible. Furthermore the repeater power requirements will be small and the circuitry comparatively simple, since only partial regeneration will be required. The problem remains to build a regenerative repeater so simple that it will be economically sound to use on short spans of line. The development of the transistor with its small size and low power requirements has made such a repeater feasible.

1.1 *Pulse Distortion Caused by Low Frequency Cutoff*

Since the frequency spectrum of a binary pulse train will extend down to and include dc, the ideal repeater should be able to handle the complete frequency band to avoid signal distortion. This would preclude the use of coupling transformers and condensers which attenuate the low frequencies and remove the dc. Practical considerations however dictate the use of these elements which means that the repeater will have a low frequency cutoff. The distortion of a binary pulse train produced by low frequency cutoff presents one of the most vexing problems the designer of a regenerative repeater must cope with. It produces what is probably the most potent source of intersymbol interference found in an average binary pulse communication system. This interference consists of a transient response whose effect may be appreciable far beyond the end of the pulse itself.

When a train of ideal flat top pulses with infinitely steep sides is applied to a load through a condenser or a transformer, the transient response persisting beyond the end of the pulse is an exponential and may be expressed as

$$T = kP_0e^{-bt} \tag{1}$$

The time t, is measured from the end of the pulse and the damping coefficient b is a function of the low frequency cutoff.* P_0 is the amplitude

* The value of b may be approximated by

$$b = 2\pi f_0$$

where f_0 is the frequency in cycles/sec at which the low frequency loss characteristic of the transformer is 6 db above that of the pass band.

of the pulse and k is given as

$$k = 1 - e^{-bt_p}$$

where t_p is the pulse duration. The sum of the transients of a sequence of pulses will shift the zero potential from the base of the pulse toward its average value as shown on Fig. 1(b). This phenomenon has been re-

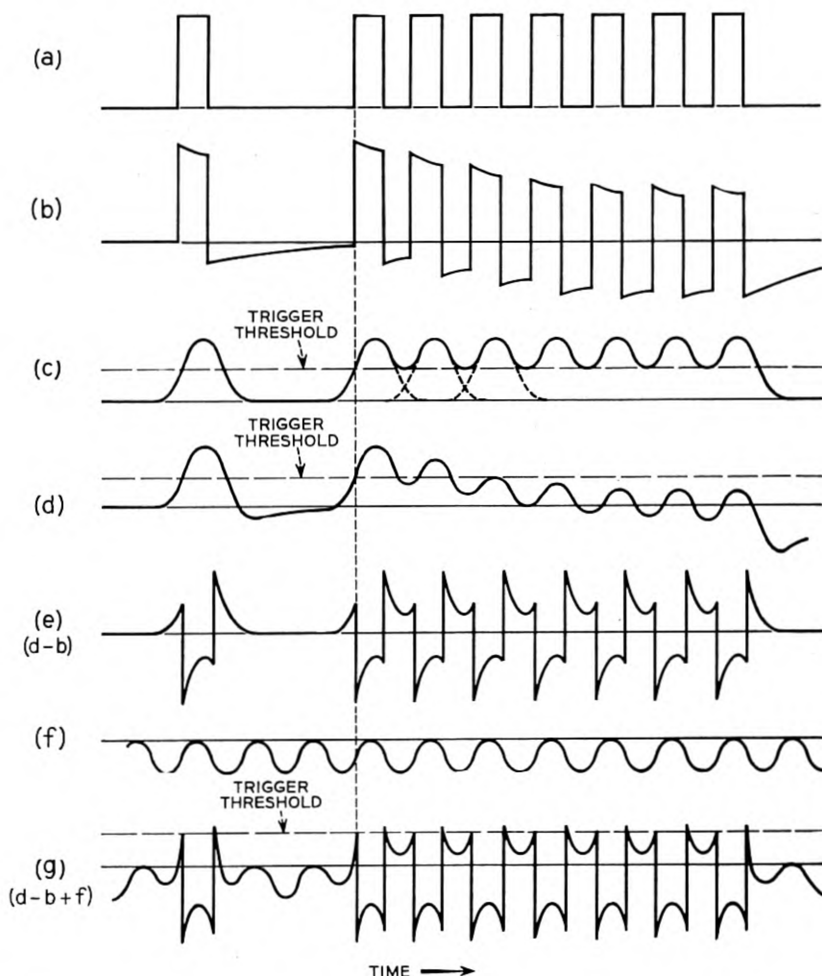


Fig. 1 — (a), a perfectly regenerated pulse train; (b) showing the effect of low-frequency cutoff; (c), showing (a) after passing over equalized line; (d), showing (b) after passing over equalized line; (e), effect of (d) minus (b); (f), inverted pedestal timing wave; (g), composite wave at input to repeater, namely, (d) minus (b) plus (f).

ferred to as "zero wander." In a regenerative repeater the trigger potential is tied to the zero level by a constant bias. Zero wander then will produce a changing bias which reduces the signal to noise margins of the repeater, or in some cases even prevents regeneration. Suppose, for example, a transmission line is equalized so the ideal pulse train shown on Fig. 1(a) will appear as Fig. 1(c) after being transmitted over the line. The individual pulses have widened until the envelope of a sequence of consecutive pulses shows as a ripple with a much smaller amplitude than the individual pulse. If the pulse train distorted by low frequency cutoff shown on Fig. 1(b) is transmitted over this line its output will appear similar to that shown on Fig. 1(d). The portion of the signal where the peak amplitude lies below the trigger threshold will not be regenerated.

1.2 *Compensation for Low-Frequency Distortion*

In the past many circuits have been devised to prevent zero wander, but none have been completely satisfactory. The repeater described in this paper effectively eliminates zero wander in a string of consecutive repeaters by means of a new and simple method. This may be better understood by referring to Fig. 2. Here are represented two successive repeaters of a transmission system. These repeaters have what appears as a conventional negative feedback loop consisting of a pair of resistors, R. The function performed by this feedback loop bears little if any resemblance to the negative feedback of linear amplifiers and is referred to as "Quantized feedback" in this paper.*

Suppose an isolated pulse of amplitude P_m is regenerated in repeater M and is applied to the line through its output transformer. The low frequency cutoff of this transformer will produce a transient response to the regenerated pulse as given in (1). A spectrum analysis of the transient tail shows that most of its energy occurs in the lower portion of the pass band of the equalized line. Consequently, it will be transmitted over the line to the next repeater with little if any frequency or phase distortion, but will be attenuated by a factor α. This transient at the input of the following repeater may be expressed as

$$T_M = \alpha k_M P_M e^{-bt} \tag{2}$$

where t is again measured from the end of the pulse. Suppose the regeneration of the pulse at the output of repeater N is delayed by time t_1

* A paper by Rajko Tomovich entitled "Quantized Feedback" was published in the I.R.E. Transactions on Circuit Theory. There are some fundamental differences in the meaning of the term, quantized feedback, as used in these papers.

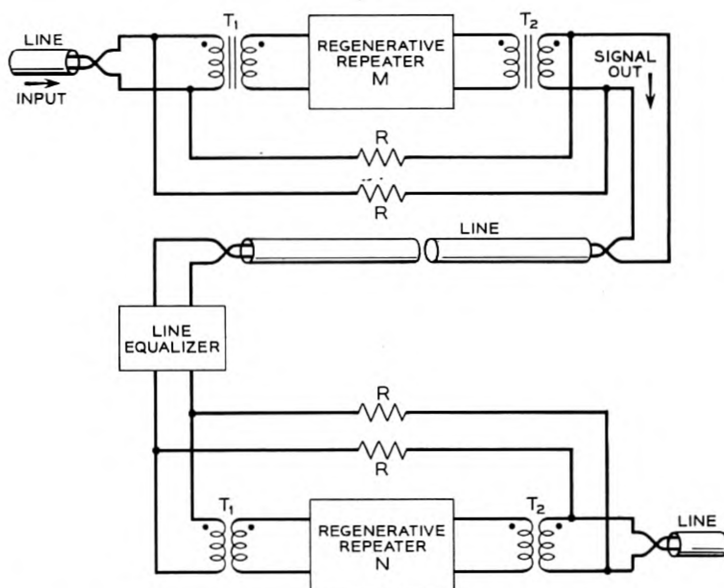


Fig. 2 — Block diagram of a section of equalized line and its terminating regenerative repeaters.

compared to the pulse at the input of the repeater. The transient response of the regenerated pulse after passing through its output transformer† will be

$$T_N = k_N P_N e^{-b(t-t_1)} \tag{3}$$

$$= k_N P_N e^{bt_1} e^{-bt} \tag{4}$$

If the transient (4) is attenuated by factor β and added in opposite phase to T_M through the feedback loop at the input of the repeater, their sum is

$$T_M - \beta T_N = \alpha k_M P_M e^{-bt} - \beta k_N P_N e^{bt_1} e^{-bt} \tag{5}$$

$$= e^{-bt}(\alpha k_M P_M - \beta k_N P_N e^{bt_1}) \tag{6}$$

This can be made equal to zero if

$$\alpha k_M P_M = \beta k_N P_N e^{bt_1} \tag{7}$$

which is accomplished by adjusting the value of β which represents the feedback attenuation introduced by resistances R. If the regenerated

† It is assumed that the electrical characteristics of the output transformers of all the repeaters are identical. In this case the damping coefficients will be identical for all the regenerated outputs.

output pulses of M and N are identical, then $P_M = P_N$ and $k_M = k_N$ and eq. (6) becomes

$$T_M - \beta T_N = e^{-bt} k_M P_M(\alpha - \beta e^{bt_1}) \tag{8}$$

This expression can be made equal to zero if

$$\beta = \alpha e^{-bt_1} \tag{9}$$

By this means zero wander produced in one repeater can be eliminated at the input of the next repeater. The low frequency distortion of one repeater corrects for the corresponding distortion produced in the previous repeater.

If the electrical characteristics of all the repeater output transformers are identical it is possible to completely remove the effects of the transient tails due to low frequency cutoff.* It is important however that t_1 should not be so large that the feedback pulse occupies the next timing interval. W. R. Bennett has shown that a similar cancellation of transients can be accomplished for more complicated types of low frequency cutoff characteristics. In this case the transient tails will be the sum of a number of exponentials having different amplitudes and damping coefficients. Here the quantized feedback must be provided by multiple loops, of greater complexity.

It may be disturbing at first to observe the resultant sum of the incoming signal and feedback as shown on Fig. 1(e). It should be noted however that the signal is not changed in any way until the repeater has triggered the regenerated pulse, and at the next time slot the tails have been cancelled, so that when the next pulse arrives it too will begin at the zero axis. Tails may also be produced by high frequency phase-loss characteristics. These however, may be removed by proper equalization.

1.3 *Timing In a Regenerative Repeater*

The binary regenerative repeater must not only regenerate the shape and amplitude of each individual pulse but it must also keep them in proper time sequence with other signal pulses. To accomplish this a suitable timing wave must be provided. This timing wave may be transmitted over separate pairs of wires or it may be derived from the signal. In the past it has been common to obtain a sine wave of the repetition

* It can be shown that, with reasonable differences in damping coefficients, quantized feedback will greatly reduce intersymbol interference even when considering a single pulse. If the contributions from all the transients of an infinite train of random pulses are summed, the resultant interference is further reduced and can be considered negligible.

frequency by exciting a high Q filter circuit from the received pulse train. Short timing pips generated from this wave are used to time the regenerated output pulses precisely. This procedure is far too involved to be used in a simple repeater. If less precision in timing is acceptable it may be accomplished with a minimum of circuitry by use of a sinusoidal wave derived from the repeater output. This is referred to in this paper as "self timing."

Self timing prohibits the use of short timing pips derived from the regenerator output. In this case most of the timing control would be exercised by the filter circuit and little, if any, by the input signal. The direct use of the sinusoidal output of this filter provides sufficient control by the input signal with only a small penalty due to less precise timing.* Self timing also sets certain requirements on the regenerator. If the timing wave is derived from an independent source it can be added to the signal in such a way as to act as a pedestal, lifting the signal above the trigger level. In such a circuit neither the signal nor the timing wave alone can trigger the regenerator. If the timing wave is derived from the output it is obvious that the signal alone must be able to trigger the regenerator, since the generation of a timing wave depends upon the signal triggering the regenerator. A timing wave derived by filtering the output of a random pattern of binary pulses will also have a varying amplitude which could cause variations in repeater noise margins. It is apparent then that self timing output cannot be used as a pedestal in a regenerator. All these objections can be overcome by the use of "inverted pedestal" timing.

Inverted pedestal timing is produced by tying the peaks of the timing wave having the same polarity as the signal pulses to a fixed level by means of a diode. This is illustrated on Fig. 1(f). The timing wave is added to the signal at the input so the sum of the signal, feedback and timing looks somewhat like the wave on Fig. 1(g). The effect of the inverted pedestal timing is to inhibit triggering except in the time interval near the peaks of the timing wave. This permits the signal to trigger the regenerator without a timing wave, yet allows timing control to be exercised as the amplitude of the timing wave builds up. With sinusoidal timing, noise often causes the regenerator to trigger either early or late, introducing a phase shift in the regenerated ouput which will be reflected in the timing wave. Since the timing wave is derived from the code pattern by a relatively high Q tuned circuit, the phase distortion of the timing wave from a shift of a single pulse will be small. With a random dis-

* E. D. Sunde, Self-timing Regenerative Repeaters (paper being prepared for publication).

tribution of noise the resultant phase shift of the timing wave will be negligible. If the interference has low frequency components, the phase shift of the timing wave may be appreciable but these are slow and consequently will not seriously effect the performance of the regenerator.

2.0 DESCRIPTION OF REPEATER CIRCUIT

The circuit diagram shown on Fig. 3 will aid in understanding the operation of the repeater. The incoming signal after being transmitted over the equalized line is applied through the input transformer T_1 to the emitter of transistor (1). The function of this transistor is to provide gain to the incoming signal. This amplified signal is applied to the emitter of transistor (2) through the blocking condenser C_2. The second transistor functions in a single shot blocking oscillator circuit being biased in the "off" condition through the resistance R_2. When the positive signal exceeds the trigger threshold, a pulse is regenerated by the blocking oscillator. During the pulse period a large emitter current flows through D_1 in the conducting direction. T_2 is the output transformer while transformer T_3 provides the essential positive feedback for the blocking oscillator.

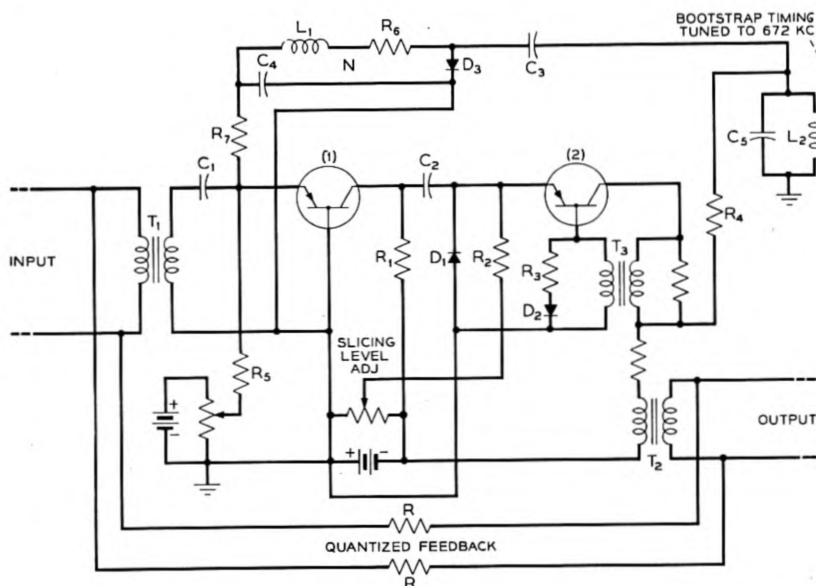


Fig. 3 — Circuit diagram of the regenerative repeater.

2.1 *Inhibiting in Blocking Oscillator*

The secondary of T_3 is connected between the transistor base and ground with the diode D_2 and resistor R_3 in series across it. The combination of diode and resistance across T_3 serves a very important function, the inhibiting of multiple triggering on a single input pulse. During the interval in which the pulse is regenerated a negative potential is applied between the base and ground. A current I_0 flows through the base of the transistor, the diode D_2 being poled to restrict the flow of current in R_3. At the end of the pulse the current I_0 in T_3 drops suddenly to a low value. This current change in the inductive winding of T_3 induces a relatively large potential across the base of the blocking oscillator. The impedance of D_2 becomes low and current flows in R_3 and T_3. The potential across T_3 decays exponentially and with proper circuit values will take the form of a damped cosine wave.

$$E = E_0 e^{-\alpha t} \cos \omega_0 t \tag{10}$$

where t is the time measured from the peak of the pulse. The values of α and ω_0 can be adjusted by varying the inductance the transformer and the capacity and resistance connected across it. E should become substantially zero at or near the next timing interval. The damping coefficient α should be sufficiently large to prevent an appreciable negative excursion of E since this will reduce the effective bias on the repeater and consequently its noise margins. This will be further discussed in the section on the measurements of errors.

2.2 *Quantized Feedback*

The quantized feedback is provided by coupling the input and output transformers by means of resistances R. The fed back pulse must be in the opposite phase compared to the input signal.

2.3 *Timing Wave Circuit*

The timing wave is derived by means of the parallel resonant tank circuit L_2C_5 which is tuned to the signal repetition frequency. The regenerated pulses are applied to this network through the relatively large resistance R_4. The amount of energy added to the network by each pulse as well as the amount dissipated in it is a function of Q. The higher the Q the smaller will be the variations of timing wave amplitude as the average pulse density of the signal train changes. This does not mean that the highest Q will be the most desirable for increased Q means larger,

more expensive coils. Higher Q's also produce greater variations in impedance and phase with small changes of resonant frequency which require much closer control of inductance and capacity with temperature. In the circuit described here the Q has a value of about 100 and its operation is quite satisfactory. The tank circuit is coupled through the small condenser C_3 to the diode D_3. This diode ties the positive peaks of the timing wave to ground as is required for inverted pedestal timing. The network N provides the timing delay needed for optimum repeater performance.

2.4 *DC Compensation in Timing Wave*

The timing wave amplitude from the tank circuit is insufficient to allow it to be applied directly to the emitter of the blocking oscillator. Consequently in the interest of circuit simplicity the signal amplifier is used for the timing wave as well. To avoid the complications introduced by dc coupled circuits when close bias tolerances must be maintained, the amplifier was coupled to the blocking oscillator by condenser C_2. This presents a problem as to how to neutralize the charge the dc component of the timing wave builds up on C_2. The means by which this is accomplished can be more easily understood by referring to Fig. 4.

In this figure the time constant of the feedback loop $R_0C_1R_1$, is made large so that substantially equal charges are added to C_1 by each regenerated pulse. In the timing loop this is also nearly true even though noise

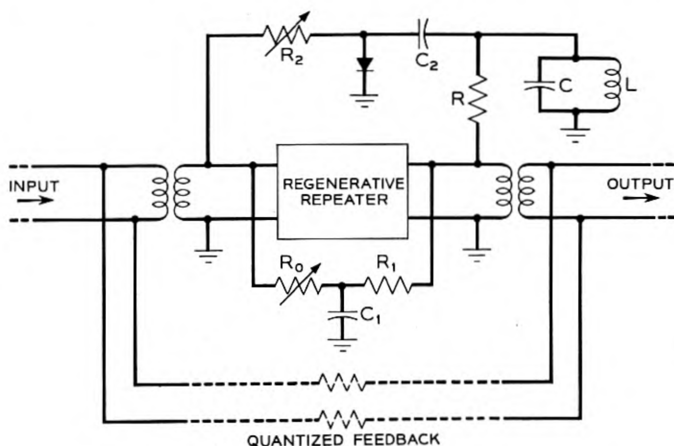


Fig. 4 — Method for maintaining the dc values of timing wave.

may change the phase of individual pulses. The change of amplitude of the sinusoidal timing wave in one pulse period will be

$$\Delta A_T = A_T[1 - e^{-\pi t_m/Q}] \tag{11}$$

where $Q = \omega L/R$ and t_m is the timing interval. In a similar manner the variation of the amplitude of the voltage across C_1 will be

$$\Delta A_C = A_C[1 - e^{-t_m/R_1C_1}] \tag{12}$$

If now R_1 and C_1 are adjusted until

$$\frac{\pi}{Q} = \frac{1}{R_1C_1} \tag{13}$$

and R_0 varied until the amplitude A_C is equal to the average value of A_T, the charge on the interstage coupling condenser should be effectively neutralized at all times. Since both loops are made up of passive elements with common inputs and outputs a single adjustment should suffice even though the pulse amplitude, width, or signal pulse density may vary.

In the repeater circuit shown on Fig. 3 this neutralizing principle is used but is more difficult to see. When a pulse is regenerated, a large emitter current flows in D_1, which produces a sharp negative voltage spike. This voltage adds a charge to C_2 which tends to neutralize the one the timing wave adds to it. The time constant of C_2 and its associated circuit may be made to equal the decrement of the tank circuit and the two amplitudes made equal by adjusting the level of the timing wave. By this means effective dc transmission of the timing wave is achieved through capacity coupling.

2.5 *Line Equalization*

The line equalizer is not essentially a part of the repeater itself. It is however so intimately connected with the repeater it is logical that they be considered together. One of the important equalizer requirements is simplicity, another, that the impedance seen from the repeater input shall be substantially constant over a relatively large frequency range. This latter requirement comes from the need of transmitting the feedback pulse around the feedback loop to the emitter of the first transistor without too much distortion. The equalizer is not used to equalize the low frequency losses of transformers but only the frequency characteristic of the line. The equalization must be such that the individual pulses are allowed to widen but not enough to cause inter-symbol interference.

A gaussian shaped pulse at the output of the line is one of the most economical to use and can have a maximum span of one timing interval at its base. However, in this case the envelope of a long consecutive sequence of such pulses will show substantially no ripple. It can be readily seen that in such a sequence the only timing control exercised by the input upon the timing wave comes from the first pulse. In the interest of better timing and consequently better repeater performance one should be content with narrower pulses at the repeater input. The resulting ripple of the envelope of a consecutive pulse sequence allows each incoming pulse some control over the repeater timing.

3.0 REPEATER PERFORMANCE

To check the performance of the regenerative repeaters a binary code generator was built having a nominal pulse repetition rate of 672 kc producing an eight digit code. Any code combination from the possible 256 can be selected or the code automatically changed at periodic intervals reproducing all possible codes in orderly sequence. Random codes may also be generated by making the absence or presence of a pulse

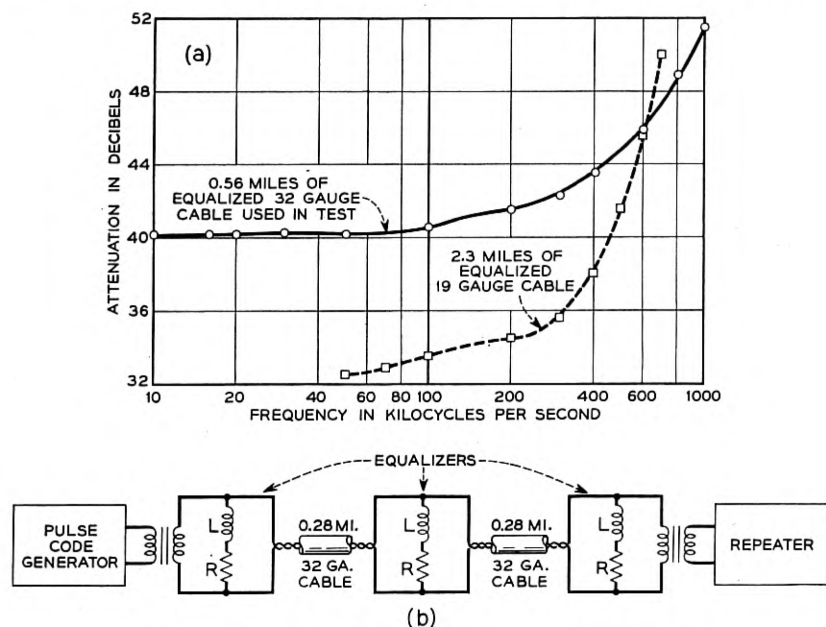


Fig. 5(a) — Equalized characteristics of 19 and 32 gauge line.
Fig. 5(b) — Block diagram of equalizer for 32 gauge line.

in any time slot dependent on the polarity of random noise. The output of the code generator was made substantially the same as the outputs of the repeaters both in shape and amplitude. Two types of transmission line were used, a line from a 51 pair 19 gauge exchange cable and a pair from a 32 gauge experimental cable. The nominal lengths of cable between repeaters was 2.3 miles for the 19 gauge and 0.56 miles for the 32 gauge cable. Fig. 5(a) shows the equalized characteristics for both these lines. The important differences between the two is a greater flat loss with a better high frequency characteristic for the 32 gauge cable. This was advantageous in the study of error production and consequently, the error measurements were all made with this cable. The 19 gauge characteristic represents about the maximum high frequency loss that can be tolerated by these regenerative repeaters.

The performance of the regenerative repeater circuit can best be shown by photographs taken from a cathode ray oscilloscope representation. Plate I shows the effect of the 19 gauge line equalizer. The output pulse (1) transmitted over the unequalized line has become very broad, extending over several timing intervals, which are indicated by small pips along the trace. The addition of the equalizer reduces the width of the received pulse (2) until it is somewhat narrower than the normal pulse interval of the code. Plate II shows a series of photographs taken of the input and output of a repeater with or without interference added at the repeater input.* A signal code at the input of the repeater is shown on (a) and its regenerated output on (b). A sinusoidal interference having a frequency of about 100-kc pictured on (c) is added to the signal as represented on (d). The regenerated output of input (d) is shown on (e). From these it can be seen that while interference does not change the pulse shape or size, it does produce a phase modulation.

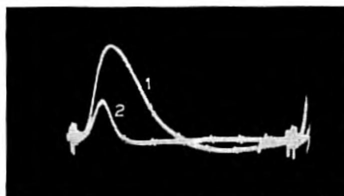


Plate I — Single pulse at output of 2.3 miles of 19 gauge cable. 1 — Unequalized. 2 — equalized.

* The input signal of this and some of the following photographs was taken with the repeater in an inoperative condition. This was done in order to avoid the resulting complexity that results when both the quantized feedback and timing wave are added to the combinations of incoming signal and interference.

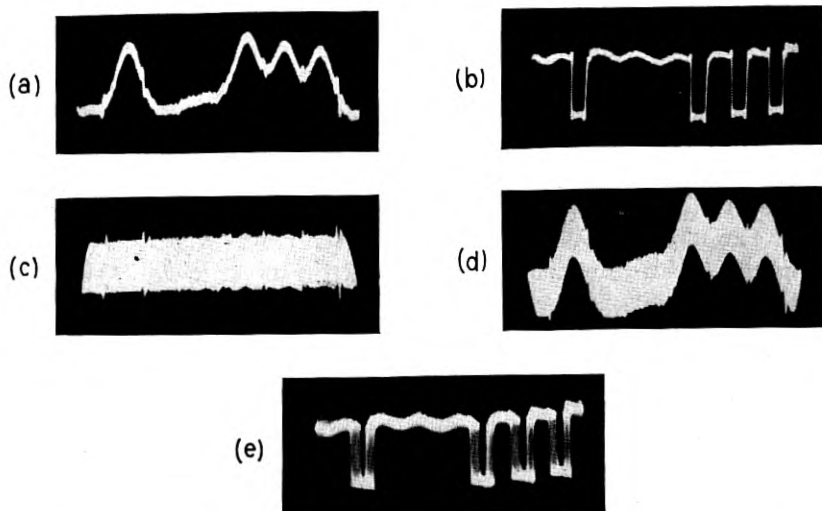


Plate II — (a), repeater input, no interference; (b), regenerated output with input (a); (c), sinusoidal interference; (d), repeater input, signal (a) plus sinusoidal interference (c); (e), regenerated output of (d).

3.1 *Performance of Repeaters in Tandem*

Plate III shows the results when certain phase modulated codes are transmitted through a series of repeaters in tandem. The regenerated signal from each successive repeater is transmitted over 2.3 miles of equalized 19 gauge line. One code which has two out of a possible eight pulses present has most of the phase jitter removed after passing through the three additional repeaters. The other fixed code shown contains four out of a possible eight pulses. The jitter is removed much more rapidly with this code, after passing through two repeaters it is regenerated almost perfectly. The reason for the difference in the regeneration of the two codes is variations in the amplitude of the timing wave. In any period of time the energy delivered to the tank circuit is proportional to the number of regenerated pulses in that interval. The amplitude of the timing wave for a fixed code with two pulses of the eight will be half the one produced by the code having four pulses out of eight present. The average number of pulses in a normal PCM signal will be half the maximum possible pulses. The timing wave should then average the same as that produced by the fixed code having four out of a possible eight pulses present. The phase jitter of the random code should be removed as quickly as it was with this fixed code. This is confirmed by

regenerating a noise-dictated random code having the same pulse density expected of a normal PCM signal. The results are shown on Plate III(c). After passing through two repeaters the jitter has been substantially removed as shown by the sharp vertical lines marking the pulses. The thickening of the horizontal lines are produced by transients produced by low frequency cut off distortion. In all these photographs the oscillograph synchronization was obtained from the code generator.

3.2 *Possible Effects of Line Temperature Variations*

The gain and phase characteristics of a particular wire transmission line is a function not only of its length but of temperature as well. To the first order approximation the effect of an increase in temperature may be considered as caused by an increase in the length of the line. In order to better understand the effect of temperature change on repeater performance the following steps were taken; The repeater was adjusted for optimum performance with 2.3 miles of line between it and the preceding repeater and then the length of the connecting transmission line was decreased by about 25 per cent. It was found that for the same interference on the input of the repeater no difference in the performance of the repeater was observed. Plate IV shows a fixed code signal after it has traversed 2.3 miles of equalized cable. Superimposed

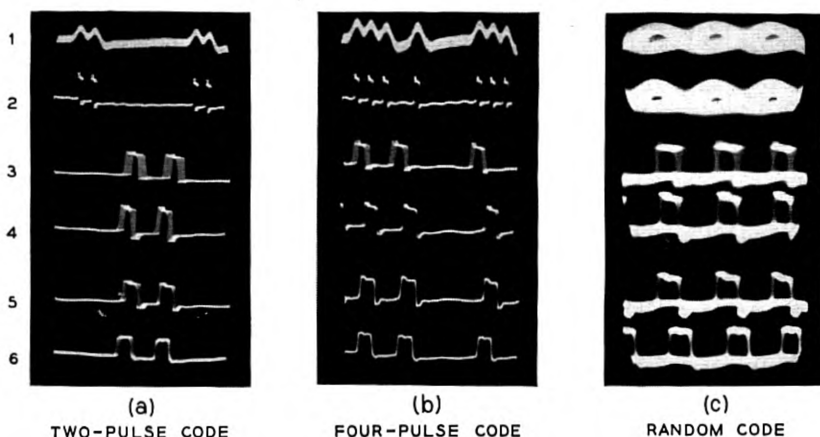


Plate III — (a), set code having 2 pulses out of possible 8; (b), set code having 4 pulses out of possible 8; (c), random code having an average of 4 pulses out of a possible 8.1 (a and b), input signal plus interference; 2 (a and b), regenerated output of 1; 3, expanded section of 2; 4, output of 2nd repeater; 5, output of 3rd repeater; 6, output of 4th repeater. 1(c), input signal alone; 2(c), imput signal plus interference.

on this is the same signal after traversing a 1.75 mile length of line and the same equalizer. Shortening the line results in the transmitted pulses having higher peak amplitudes and narrower widths. Faulty high frequency equalization of the shorter lengths produces the short tail following the pulse. It is interesting to observe that the transient tail due to the low frequency cut off has not changed appreciably as the line was shortened. This is to be expected since it can be shown that the energy of the low frequency cut off transient is concentrated in low frequency end of the transmission spectrum. In this region changes in the length of the line, or changes in the primary constants will result in inconsequential changes in attenuation and phase as is shown on Fig. 6. If the quantized feedback is adjusted for the worst condition, i.e., the highest temperature likely to be encountered, it will not need to be changed with lower temperatures.

4.0 ERROR PRODUCTION BY EXTRANEOUS INTERFERENCE

A knowledge of the performance of a regenerative repeater with various types and amounts of interference added to the input signal is important. Consequently a study of such errors produced in one of these repeaters was undertaken. Two general types of extraneous interference was used in this study. The first is impulse noise, the type which is produced by telephone dials, switches, lightning surges and crosstalk from other pulse systems. The second is sinusoidal noise, the type which come from power line or carrier crosstalk. This interference may affect the regenerated output in a number of ways. It may produce a phase shift or "jitter" in the output; cause a pulse to be omitted; or cause a spurious pulse to be inserted in the signal code. The phase jitter will be largely removed by timing regeneration in subsequent repeaters, but omission and most insertion errors will be carried through the remaining repeaters, causing distortion in the decoded signal.

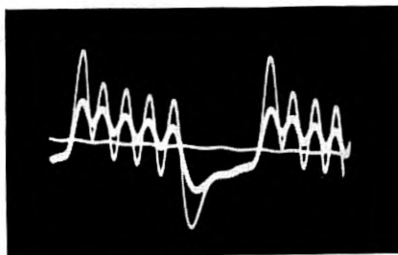


Plate IV — Superimposed picture of the outputs of 2.3 and 1.75 miles of 19 gauge cable with identical inputs.

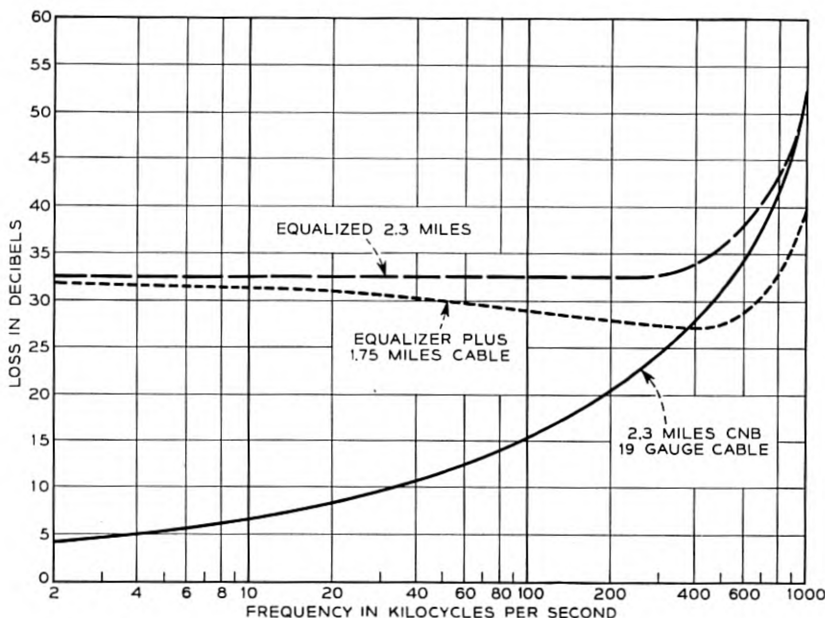


Fig. 6 — Effect of changing the length of 19 gauge line with fixed equalization.

4.1 *Description of Error Detecting and Counting Circuit*

An error detecting and counting circuit was built to count insertion and omission errors. This circuit (block diagram, Fig. 7) is a coincidence detector in which each pulse or space of the repeater input signal is compared to its corresponding regenerated output. As long as the two sources are the same, i.e., having corresponding pulses or spaces, there is no output from the detector. If the two differ the detector produces an output pulse which may be caused to actuate the counting circuit. The code generator as has already been described produces a number of different types of signal codes.

The output of the code generator is transmitted over 0.56 miles of equalized 32 gauge cable to the regenerative repeater under test. Interference is introduced at the repeater input when desired. A portion of the code generator output is differentiated and passed over a delay cable whose delay is substantially that of the section of 32 gauge line over which the signal is transmitted. This delayed signal is regenerated without error by the single shot blocking oscillator A. The width of the blocking oscillator pulses are adjusted to be about half of the total timing interval. The width of the pulses from the regenerative repeater

are likewise widened to a corresponding width by blocking oscillator B. Unfortunately a variable phase shift is introduced in the repeater output by interference and by variations in the timing wave amplitude and phase. This variable phase shift prevents perfect coincidence between the outputs of blocking oscillators A and B. An example of phase "jitter" caused by interference is shown on Plate V(a). To overcome this a sharp sampling pip, as shown on the same plate, is provided to enable the detection of the narrow region of coincidence between the two signals. These pips are generated from the repeater timing wave, hence they follow the timing wave phase variations. The regenerated signal pulses also follow the timing wave phase. If the sampling pulse is positioned to fall in the center of the regenerated pulses, it will tend to maintain that position as the timing wave changes.

The gates require a signal pulse and sampling pip to be present simultaneously before there can be an output. This output, then, will have substantially the same shape and position as the sampling pip. When a signal pulse is simultaneously applied to each gate the two outputs can be made to cancel when added in opposite phase as is done in T_1. If however there is a pulse on one gate and a blank on the other, an output pulse will be produced. The polarity of this pulse will depend upon which gate contains the signal pulse. Since the decade counter is

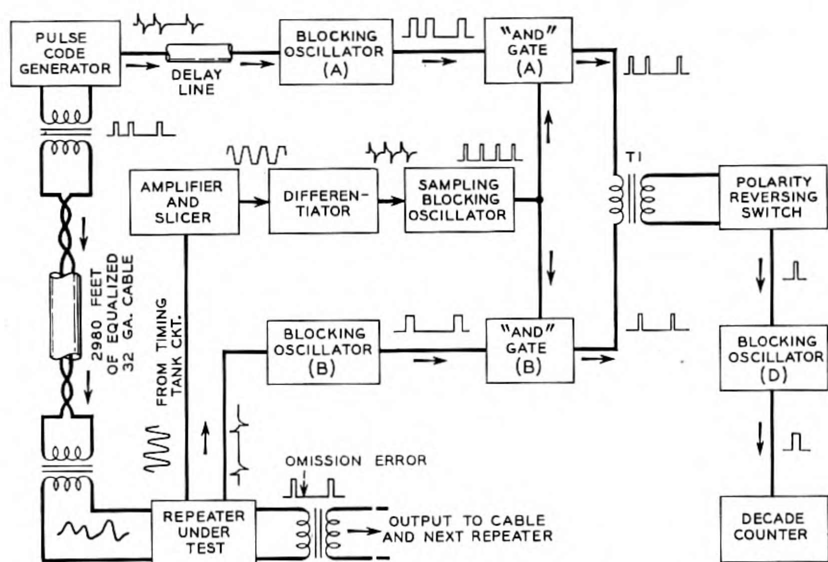


Fig. 7 — Block diagram of error detecting circuit.

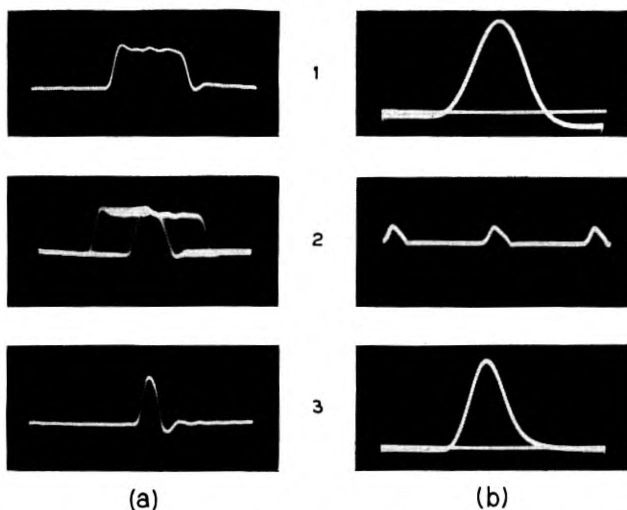


Plate V — (a) with 1, repeater output; 2, jitter on output pulse; 3, sampling pulse. (b) with 1, signal pulse at repeater input; 2, 672-kc timing pips; 3, interference input.

triggered by pulses of one polarity, the reversing switch permits the independent measuring of different types of errors. The counter used in this study has 9 decades capable of counting and recording $(10^9 - 1)$ errors at 10^6 counts per second.

4.2 *Discussion of Impulse Noise Generator*

A study of the noise in cable pairs leading from a central office indicate that impulse noise will cause much of the expected interference on pulse systems. In order to simulate the effect of this type of interference, a generator was built which produces uniformly shaped pulses over a wide range of rates. The polarity of these pulses can be reversed and their amplitude varied continuously from zero to a value exceeding the peaks of the signal pulses. These impulses were introduced into the center of a transmission cable through a high impedance. Plate V(b) shows photographs comparing the impulse with a signal pulse. The repetition rate for the impulse interference used in this investigation was 10^4/sec, which is low compared to the nominal pulse repetition rate of the signal $(6.72 \times 10^5/\text{sec})$. With the relatively large separation between interfering impulses, there is no measurable interaction between errors produced in the repeater. At the same time the impulse rate is high enough to get an excellent statistical distribution in the 10 second interval used in these measurements.

4.3 *Production of Impulse Errors — Nomenclature and Discussion*

To expedite the discussion of impulse errors, the following system of nomenclature is used. Any impulse having the same polarity as the signal pulse is designated as "plus." Those having the opposite polarity are "minus." Two types of errors are produced. First, a spurious pulse may be added to the regenerated signal; this is called an "insertion" error. Second, a signal pulse may be removed, which is called an "omission" error. A "plus insertion" error is a spurious pulse introduced by an impulse having the same polarity as the signal. A "plus omission" error on the other hand is pulse omitted because of a pulse of same polarity as the signal. A "minus omission" error is a pulse omitted because of an impulse having a polarity opposite to that of the signal.

A positive pulse, if large enough, can produce a spurious pulse at any instant of time not already occupied by a pulse. The only requirement for the production of such a pulse is that the sum of the impulse and timing wave exceed the trigger level.* On the other hand, a negative impulse cannot produce a spurious pulse but can only cause a signal pulse to be omitted. If a pulse is to be omitted the sum of its amplitude, the timing wave and the impulse must not exceed the trigger level. It would be expected that the number of plus insertion errors will exceed the minus omission errors. This follows from the fact that a spurious pulse may be produced at any point not already occupied by a pulse. On the other hand if a signal pulse is to be omitted the negative impulse must occur in the time interval occupied by the signal pulse. A positive impulse is indirectly responsible for the positive omission error. When a spurious pulse is produced a short interval of time ahead of a signal pulse, the latter may be removed by the inhibiting reaction of the spurious pulse. There is no apparent way in which a minus insertion error can be produced. This is confirmed by the fact that no error of this type was observed in this investigation. Thus we have three types of errors produced: plus insertion, minus omission and plus omission.

4.4 *Results of Impulse Interference Measurements*

Preliminary measurements of errors as functions of impulse amplitude were made using random code. These measured values, shown on Fig. 8 exhibit many of the expected characteristics. For example the insertion errors are more numerous than the omission and the threshold of the plus omission errors is considerably higher than those of the other two.

* The trigger level is normally considered to be the negative dc bias applied to the emitter of the blocking oscillator. There are however other components of the bias that will be discussed later.

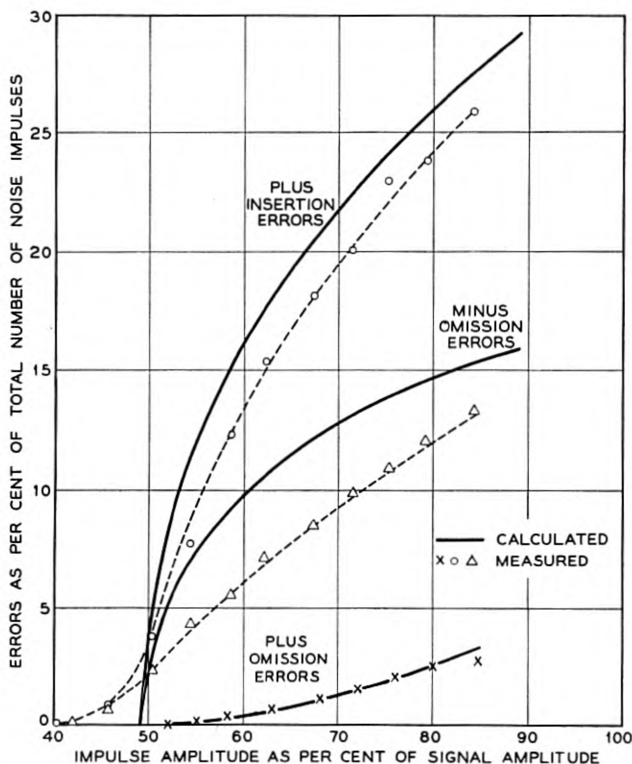


Fig. 8 — Repeater errors as a function of interference amplitude.

On the other hand there are some deviations from the simple theory of a perfect regenerator such as the low common threshold value of the plus insertion and minus omission errors. Some of the differences can be attributed to the extremely sensitive method of measuring errors. Here the maladjustments of timing tank circuit, quantized feedback amplitude as well as other factors which cannot be readily detected by other means are reflected as sources of error. However with care these errors can be made small and the measured values should follow the theoretical values reasonably well.

Most variations from theoretical values are due to changes in the effective bias caused by intersymbol crosstalk. This can be demonstrated by measurements made using set codes. In all these codes the number of pulses equaled the number of blanks but combinations varied from one to another. On Fig. 9 the omission errors are plotted for a fixed impulse amplitude as a function of the number of pulses which

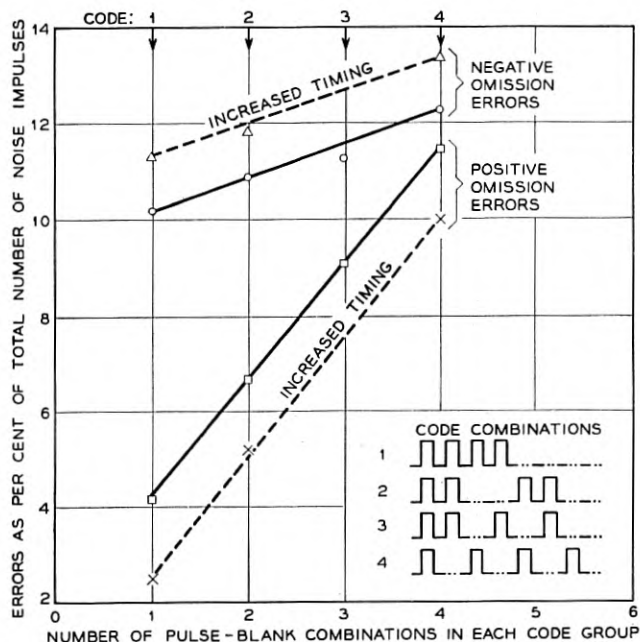


Fig. 9 — Repeater errors as a function of pulse distribution in code.

are followed by a space in the particular code. The codes used for various points on the abscissa are shown on the graph. The omission error curves plotted in this manner are linear. These data demonstrate that the presence of a pulse modifies the trigger level in the next timing interval. This is largely due to the negative excursion of the damped cosine voltage from base to ground in the blocking oscillator. On Fig. 10(a) is shown the circuit of the single shot blocking oscillator used in the repeater. With no timing an incoming signal must overcome bias V_{DC} to trigger the repeater. The solid curve on Fig. 10(b) shows the dc bias with the timing wave added at the blocking oscillator emitter. Fig. 10(c) shows the base voltage when a pulse is produced in the first timing interval. The pulse begins at t_0 and ends at t_1. As previously mentioned the sudden rise of the base and collector impedance coupled with the fall of the current in the transformer windings, produces an inductive voltage surge across transformer T_3 at t_1. The decay of this voltage surge can be controlled by the inductance of the transformer and the damping resistor R_b. This positive decay voltage across the base will inhibit the blocking oscillator from triggering. It is essential that this decay be adjusted so it will inhibit triggering until the following time slot. If

the decay transient is a damped oscillation and the base voltage passes through zero at the next normal triggering time, sufficient damping must be provided so the negative excursion is negligible. The dashed line shows how the effective bias at the emitter is modified by this voltage across the base.

Fig. 11 shows the measured values of plus insertion and minus omission errors for two set codes. These are plotted as functions of impulse amplitude. The first code has alternate pulses and blanks while the second consists of pairs of pulses separated by pairs of blanks. With these two curves the error threshold values may be determined from

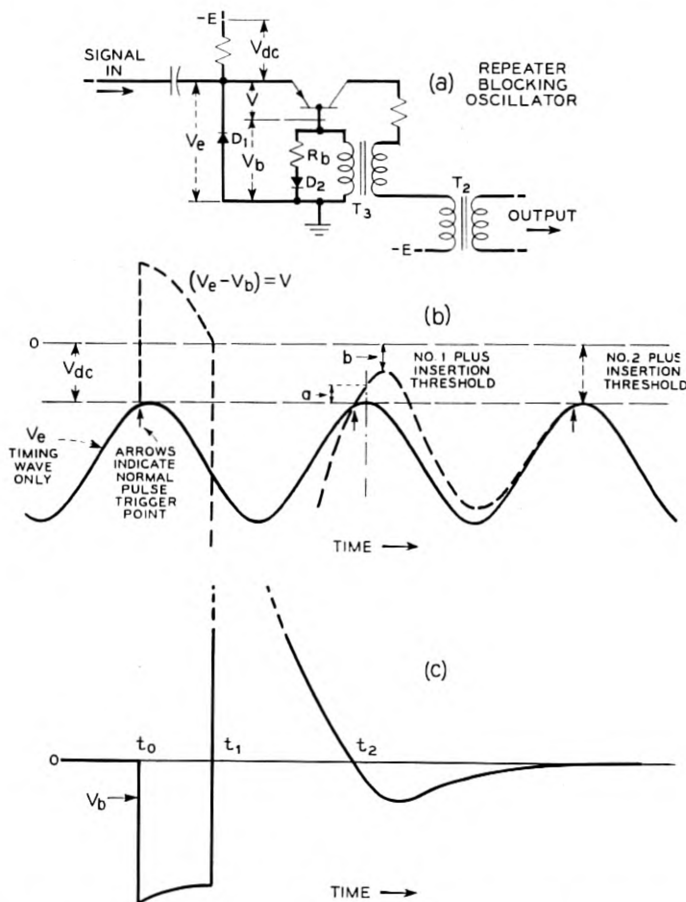


Fig. 10 — (a) Circuit diagram of blocking oscillator showing various components of the effective bias. (b) The effective bias as a function of time. (c) Inhibiting voltage V_b produced by a regenerated pulse.

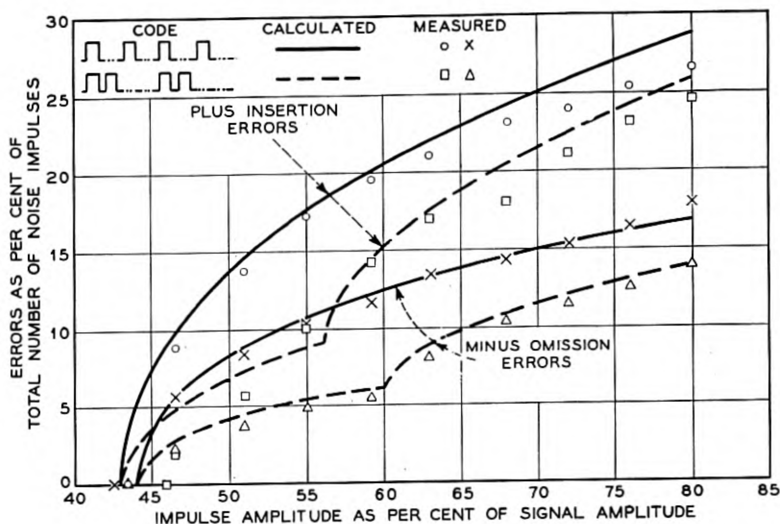


Fig. 11 — Calculated and measured repeater errors for two set codes.

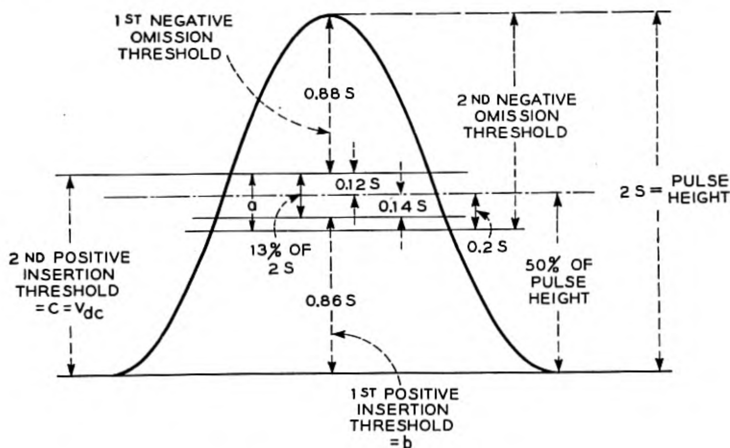


Fig. 12 — Bias levels used in calculating repeater errors.

the points of discontinuity. Fig. 12 illustrates these various error thresholds with reference to a signal pulse. Theoretical curves were plotted using these values and the observed values of timing and signal amplitudes as shown on Fig. 11. It can be seen that very good agreement exists between the measured and computed values.

The separate lower thresholds for insertion and omission errors may

be explained from Fig. 10(b). These are caused by the phase shift introduced by the inhibiting voltage to the effective bias compared to that of the timing wave. The omission thresholds are determined chiefly by the maximum signal amplitude. On the other hand the insertion thresholds are determined by the point of maximum trigger bias. There exists then two separate threshold values for a timing interval which follows a regenerated pulse. These values can be measured from points "a" and "b" on Fig. 10(b).

4.5 *Result of Sinusoidal Interference Measurements*

On Fig. 13 are shown the errors produced by sinusoidal interference. Here a 110-kc sine wave is added to the signal and the various types

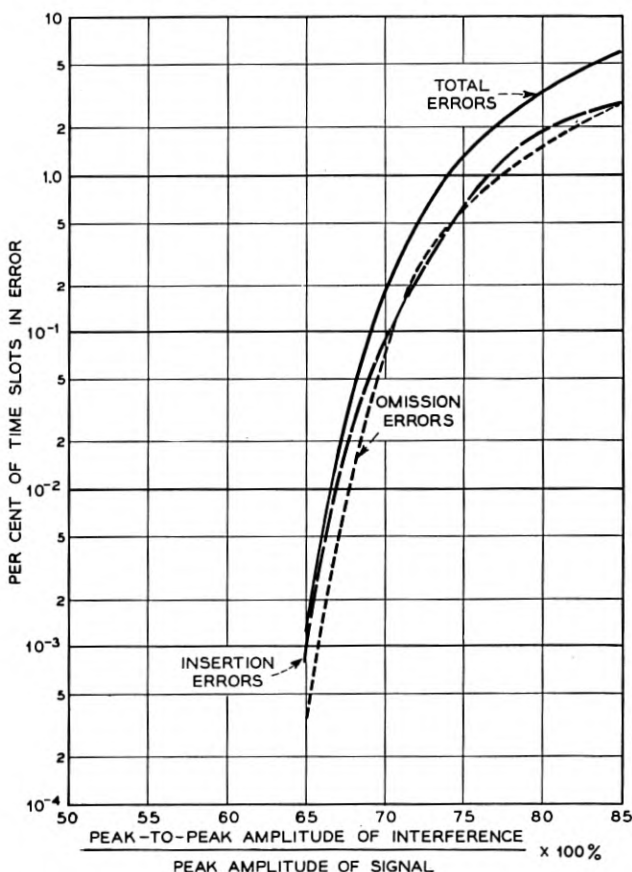


Fig. 13 — Repeater errors as a function of interferences level for sinusoidal interference.

of errors counted. Random code was used in this case and the repeater bias was adjusted to provide equal omission and insertion thresholds. The threshold for this particular case occurred when the peak to peak sinusoidal interference was 63 per cent of the signal amplitude. This is lower than the theoretical maximum which with a constant bias centered at the half amplitude point, would be 100 per cent of the peak to peak signal amplitude. For the bias conditions illustrated on Fig. 12, this percentage would be 86 per cent for the positive insertion threshold and 88 per cent for the minus omission. This becomes apparent when the negative and positive excursions of the interfering sine wave are considered as minus and positive impulses respectively. The remaining loss in the interference margins can easily be due to maladjustments of timing, quantized feedback or inhibiting.

When the frequency of the sinusoidal interference is varied, the number of errors for a constant interference voltage at the blocking oscillator emitter does not change appreciably. However, the input transformer and condenser coupling introduce a substantial frequency characteristic. This reduces considerably the errors caused by power line crosstalk. One of the striking things about the sinusoidal interference errors is the rate at which they increase above the threshold. For example, a change of 1 per cent of the interference amplitude can triple or quadruple the total number of errors.

5.0 SUMMARY

New techniques and devices now make it possible to build practical regenerative repeaters for use in digital transmission. Such a repeater which is suitable for a 12-channel, 7-digit PCM system, is discussed. Simple, inexpensive devices are used to eliminate the effects of distortion due to low frequency cutoff and to provide self timing for the circuit. Experimental evidence is presented which shows the repeater to function as expected.

ACKNOWLEDGEMENTS

I am deeply indebted to J. V. Scattaglia for his aid in this project and to the pioneering work of A. J. Rack on quantized feedback which was of great help in the development of this regenerative repeater. I also wish to thank W. R. Bennett, C. B. Feldman and Gordon Raisbeck for their aid and many valuable suggestions.

Transistor Pulse Regenerative Amplifiers

By F. H. TENDICK, JR.

(Manuscript received April 5, 1956)

A pulse regenerative amplifier is a bistate circuit which introduces gain and pulse reshaping in a pulse transmission or digital data processing system. Frequently it is used also to retime the pulses which constitute the flow of information in such systems. The small size, reliability, and low power consumption of the transistor have led naturally to the use of the transistor as the active element in the amplifier. It is the purpose of this paper to describe some of the techniques that are pertinent to the design of synchronized regenerative amplifiers operating at a pulse repetition rate of the order of one megacycle per second. An illustrative design of an amplifier for use in a specific digital computer is presented.

1. INTRODUCTION

A basic building block of many modern digital data processing or transmission systems is a pulse regenerative amplifier. The particular high speed transistor regenerative amplifiers to be discussed in this paper are intended for use in systems where the logic operations on the digit pulses are performed by passive circuits and the amplifiers are inserted at appropriate intervals to amplify, reshape, and retime the pulses. The design of these amplifiers for any specified system involves a knowledge of the environment of the amplifier in the system, a study of possible functional circuits which are combined to form an amplifier circuit, and the selection of a combination of these functional circuits to achieve the desired amplifier performance. Although a study of the functional circuits constitutes the major portion of this paper, the design of an amplifier for a particular digital computer is presented to illustrate the general design procedure.

One important way in which these amplifiers differ from many pulse amplifiers is that they must function properly under adverse conditions. That is, instead of merely expecting superior performance most of the

Fig. 2. In the following sections the relation between each of the above functional features and amplifier performance is discussed, various circuit configurations to achieve each function are investigated, and the interactions between the functional circuits are examined. The design of any particular amplifier then consists of a suitable selection of a transistor and functional circuits to achieve the desired amplifier performance.

2. TRANSISTOR PROPERTIES

In a regenerative amplifier the transistor operates as a switch with power gain. The "on" and "off" state usually are characterized, respectively, by high and low collector current levels, and changes of state are initiated by applied control signals. The performance items of interest are the power dissipation in the two states, the speed with which the transistor changes state, the amount of power gain available, and the attainable margins against false operation. The transistor parameters related to these items, as discussed below, are listed in Table I with typical values for several classes of transistors. Desirable and satisfactory values have been indicated in italics.

The power dissipated in a transistor in the "off" state is proportional to I_{co}, the collector current with the emitter open circuited, and to the collector supply voltage. This is wasted power and, since the minimum collector supply voltage usually is dictated by other considerations, a low I_{co} current is desirable to reduce standby power. Point contact units are relatively poor in this respect. In junction units the I_{co} power is almost negligible compared to other circuit standby power.

The power dissipated in a transistor in the "on" state is proportional to the saturation voltage between the collector and the common terminal.

Table I — Transistor Switching Properties

Switching Features	Point Contact Transistors (Low Resistivity Ge)	Junction Triode Transistors		
		Ge Grown	Ge Alloy	Si Grown
I_{co} at V_c = 10v	1500 μa	*5 μa*	*5μa*	*0.01 μa*
Collector to emitter saturation voltage at I_c = 10 ma	0.8 V	0.5 V	*0.05 V*	4 V
fα cut-off	*15 mc*	2 mc	4 mc	4 mc
Base resistance	*50 ohms*	500 ohms	*100 ohms*	500 ohms
Collector capacitance at V_c = 10V	*0.5 UUF*	10 UUF	20 UUF	10 UUF
Collector breakdown voltage	*40 V*	*100 V*	*35 V*	*100 V*
Punch through voltage	*no punch through*	*100 V*	*35 V*	*100 V*
Emitter breakdown voltage	*40 V*	5 V	*35 V*	1 V
Ratio of alpha at I_e = 10 μa to alpha at I_e = 1 ma	*3*	0.8	0.8	0.6

Again, this represents wasted power, but also important is the fact that it places an upper limit on the output power available from the transistor. Hence, it is desirable to have as low a saturation voltage as possible. Alloy junction transistors are especially good in this respect.

The speed with which a transistor changes state is principally a function of the alpha cut-off frequency (which should be high), base resistance, and collector capacitance (both of which should be low).[3, 4] Both the rise and fall times of the transistor response are greatly influenced by the associated circuitry; generally a blocking oscillator circuit yields the fastest response.

The amount of effective power gain available from a regenerative amplifier is influenced by two transistor properties. One property is the breakdown voltage, which may be the collector to base breakdown voltage or the collector to emitter punch through voltage (whichever is lower). This limits the output power by limiting the collector supply voltage. The other factor is the variation of alpha with emitter current, especially at low emitter currents. The minimum average emitter current required to initiate self-sustaining positive feedback determines the minimum input power. Point contact units are especially good in this respect in that alpha may approach ten at emitter currents as low as five microamperes. Junction units are poor since alpha generally decreases rapidly at emitter currents below one hundred microamperes.

Even though the attainable margins against false operation are largely a matter of circuit design, two transistor properties occasionally become important. In point contact units trouble with lock up in the "on" state may occur due to internal base resistance. Although this property of base resistance is exploited in negative resistance feedback circuits, it is undesirable in circuits where the feedback is obtained by external coupling. In grown junction units the emitter to base reverse breakdown voltage may limit the voltage margin against false triggering caused by noise or crosstalk. Normally it is desirable to have a one or two volt margin.

From the above discussion it can be seen that no one type of transistor is outstanding in all features. The choice of which unit to use in a specific amplifier depends upon the repetition rate, gain, and power requirements desired of the amplifier. Although the point contact type has the best overall performance of the types shown in Table I, it is quite possible that new types (such as PNIP or diffused triodes[13]) and improved designs of the present types will change the picture.

3. FEEDBACK CIRCUITS

The use of positive feedback in an amplifier results in high gain and short rise time. If the input circuit is isolated from the feedback loop by

a diode or large resistor, these effects are enhanced and the shape, duration, and amplitude of the output signal become independent of the input signal. These results are possible because once the circuit has been triggered and the feedback loop gain is greater than unity, the response proceeds independently of input conditions and is determined solely by the transistor and circuit parameters.

By definition a regenerative amplifier must have positive feedback sufficient to cause instability during the transition period between the "off" and "on" states. When investigating various circuits, it is necessary to eliminate circuits which are never unstable when a pulse is applied to the input circuit. If the circuit is unstable under either of the conditions shown in Fig. 3, sufficient instability is possible. However, if the circuit is stable, linear, and either the small signal open circuit voltage gain or the short circuit current gain is less than unity or negative at all frequencies, it is impossible to have instability. These latter conditions for instability often can be easily checked by inspection without tedious computation or experimentation.

This use of positive feedback requires that attention be given to its control. To be useful, the amplifier must be stable in one state and at least quasi-stable in the other state. The change from instability in the transition period to stability in the end states is accomplished by a nonlinear change in the gain or impedance of some element in the feedback loop. Usually the "off" state is made stable by causing the voltage and current conditions in the input circuit to reverse bias the transistor input. The "on" state may be made stable (or quasi-stable when there are reactive coupling elements in the loop) in several ways. For example, the transistor may be permitted to saturate when the desired pulse voltage is reached; a "catching" diode may be used to clip the pulse voltage at an appropriate level; or a current switch may be used to

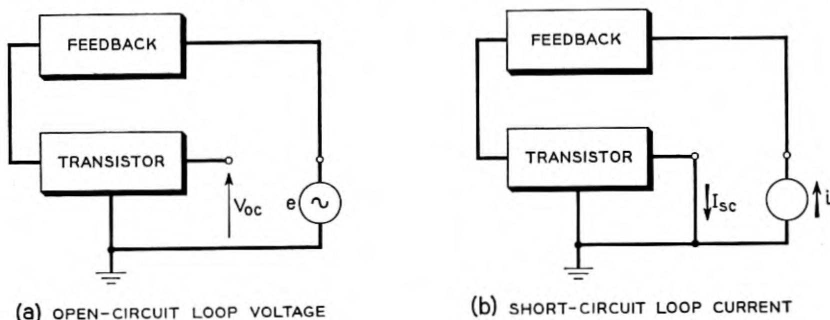


Fig. 3 — A check for instability.

introduce an impedance in the feedback loop at a predetermined current level.

The degree of stability of the amplifier in the "on" state may be thought of as the amount of power required to initiate the transition to the "off" state. During the early portion of the output pulse duration the degree of stability should be large, but near the end of the pulse duration it should be relatively small to make turn-off easier. Also, the degree of stability should not change over the range of output loading expected for the amplifier and should be effected without excessive wastage of pulse or supply power. These conditions are difficult to fulfill when the range of output load current may be as large as 20 to 1.

Three methods of obtaining positive feedback in transistor circuits will now be considered: (a) negative resistance feedback; (b) capacitor coupled feedback; and (c) transformer coupled feedback. Of these, transformer coupled feedback appears to be the best for most applications. It will be assumed that the type of feedback under discussion is the dominant or only type present; circuits employing more than one feedback mechanism generally violate the premise of simple circuitry and will not be discussed.

3.1 *Negative Resistance Feedback*

With the advent of point contact transistors a novel form of negative resistance was offered to circuit designers for use in positive feedback applications.[5] This negative resistance property occurs when the current gain of a transistor is greater than unity and the emitter and base small signal currents are in phase.* At first sight this property appears to lead to attractively simple regenerative amplifiers. However, as systems become more complex and, consequently, amplifier requirements more severe, the original simplicity often is lost due to the additional circuitry required to control the negative resistance. An example, shown in Fig. 4, is similar to a regenerative amplifier described by J. H. Felker.[2] The functional circuits are indicated by dashed outlines.

This amplifier operates at a one megacycle pulse repetition rate with one-half microsecond, three volt pulses. It is capable of driving from one to six similar amplifiers. The output pulse rise time is 0.05 microsecond, the average dc standby power is 33 milliwatts, only a few components operate at as much as half of maximum ratings, and the supply voltage margins† are greater than ± 15 per cent. Seven hundred of these ampli-

* Although point contact transistors are noted for this property, certain types of junction transistors also exhibit it. For example, see Reference 7.

† Supply voltage margins, the amount by which the supply voltage may be

fiers operated in a system for over 17,000 hours with a failure rate of slightly less than 0.07 per cent per thousand hours.

These features, however, are obtained at the expense of relative complex circuitry. This negative resistance type of high speed regenerative amplifier has the following inherent limitations.

1. The degree of stability in the "on" state depends critically on the collector current. In the example a dummy load must be strapped in when the amplifier drives less than four logic circuits.

2. A steering diode (D3) and a timing circuit diode (D1) have critical reverse recovery time[6] specifications.*

3. The requirements on transistor parameters (primarily the dynamic alpha versus emitter current and base resistance characteristics) are relatively critical.

4. A relatively large amount of synchronizing power is required.

5. With transformer output coupling (as discussed in Section 5.1) a large amount of the total standby power is absorbed by a circuit required to protect the transistor in case the timing voltage fails (In the example 21 milliwatts, or 64 per cent of the standby power, is absorbed by R3.)

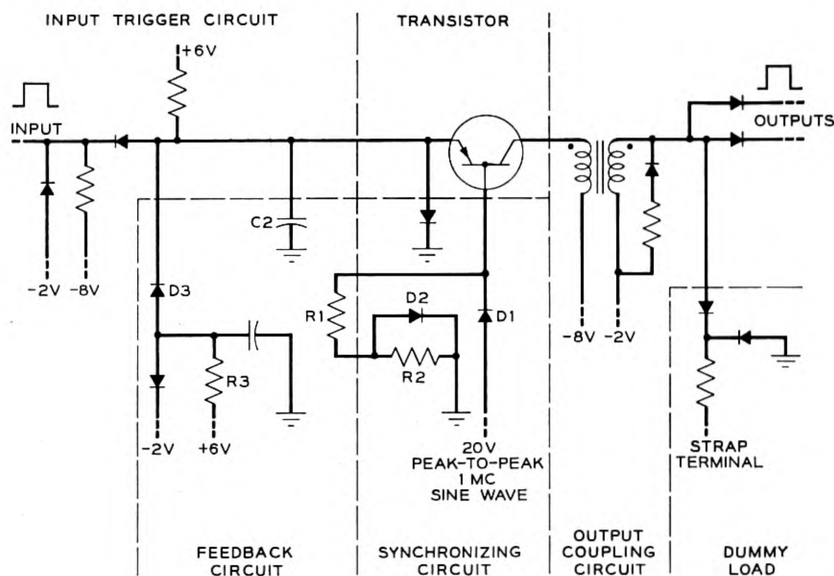


Fig. 4 — Negative resistance feedback amplifier.

varied without causing an operational failure, are an indication of the sensitivity of the amplifier to changes in component values.

* At lower pulse repetition rates this property may not be critical.

The use of an inductor, instead of a resistance, in the base lead does not appear to mitigate the limitations.

3.2 *Capacitor Coupled Feedback*

A second method of obtaining positive feedback is by external coupling through a capacitor or capacitor-resistor network. This method is seldom used for the principal feedback for reasons to be mentioned. Occasionally, in conjunction with some other type of feedback, it may be used to provide additional feedback during the rise time of an amplifier.

Since the voltage and current gain of a capacitor can not exceed unity, the open circuit voltage gain and the short circuit current gain of the rest of the loop (Fig. 3) must be greater than unity for instability. This criterion indicates that capacitor feedback is limited to point contact, or other transistors with an alpha greater than unity, or to a junction transistor in the common emitter configuration.*

A circuit with capacitor feedback around a short-circuit stable point contact transistor might take the form shown in Fig. 5. Although this type of circuit has the merit of simplicity, it has the following limitations:

1. The initial feedback current is highly dependent upon the incremental output load impedance. This may result in a failure to trigger when the load approximates a short circuit, as in the case of diode gates or a large stray capacitance.

2. The degree of stability in the "on" state is critically dependent on the load current and the collector supply voltage. Variations in either may cause a foreshortened output pulse or require an excessive timing signal current for turn-off.

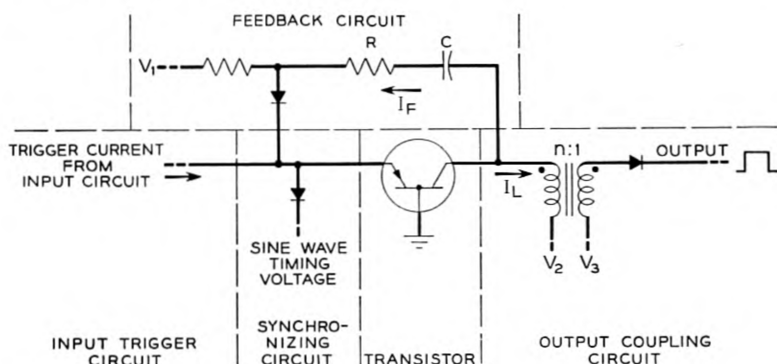


Fig. 5 — RC feedback amplifier.

* An inverting transformer is necessary with the junction transistor.

3. The necessity of a feedback circuit time constant equal to or shorter than the output pulse length results in a relatively low output power efficiency.

Due to the above considerations, capacitor feedback appears to be the least attractive type of feedback.

3.3 *Transformer Coupled Feedback*

A transformer appears to be the most convenient and versatile component for feedback coupling in a regenerative amplifier. The pertienent features* of a transformer are:

1. Current or voltage gain (impedance matching.) This feature permits full use of the power gain of the transistor, even if such gain be in the form of voltage or current gain only.
2. Bias isolation between circuit parts and the possibility of supplying dc voltage bias without the use of additional elements.
3. Phase inversion, if desired.

All of these features, conveniently combined in a transformer, provide great design freedom to meet specified circuit objectives. Since positive feedback is possible with any type of transistor (with power gain, of course), the choices of transistor and connection are determined by other circuit requirements.

The use of transformer coupled feedback yields the familiar blocking oscillator circuit. An important feature of this circuit is the fast rise time that is obtainable. Linvill and Mattson[4] have shown that a junction transistor with an alpha cutoff frequency of two megacycles may exhibit a rise time of 0.1 microsecond in an unloaded blocking oscillator with collector to emitter coupling, Fig. 6 (a). It can be shown that the same response may be expected with collector to base or base to emitter coupling, provided that the transformer turns ratio is modified, Figs. 6 (b) and 6 (c). When the circuit is providing useful output power into a load, a slightly different turns ratio would be used for optimum rise time, which may be appreciably slower than in the unloaded case. However, it should be noted that the foregoing gives no information about the initial response of the circuit from the time that the input trigger is applied until the output reaches ten per cent of its final value. In some instances this initial time, which is a complicated function of the transistor non-linearities, may be comparable to the output rise time.

In a blocking oscillator circuit with a fixed output load, the degree of stability in the "on" state decreases with time. The reason is that the

* The operation of a transformer over the non-linear portion of its magnetization characteristic is outside the scope of this paper.

voltage across the coupling transformer, which is approximately constant during the pulse duration, causes an increasing magnetizing current to be subtracted from the initial feedback. When the feedback current can no longer support the required output current, the circuit turns off. In a synchronized amplifier the value of the feedback transformer mutual inductance may be specified to give the desired degree of stability at the end of the predetermined pulse length. Thus, the least stable condition occurs at the end of the pulse duration and is under the circuit designer's control. At other times during the pulse duration the circuit is more stable, which reduces the possibility of premature turn-off.

Other considerations, such as stability variations with output current, power dissipation, and output voltage regulation, depend upon whether the output load is in series or in shunt with the feedback loop. Therefore, these considerations are discussed in connection with output coupling in

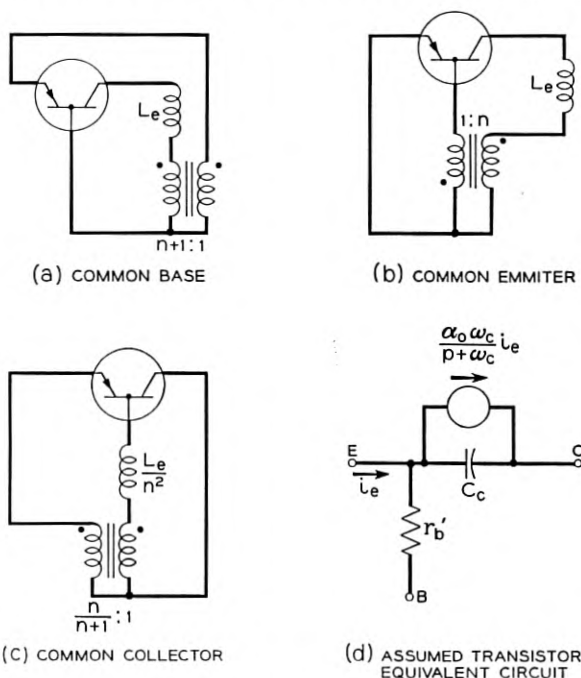


Fig. 6 — Transformer coupled blocking oscillator circuits.

Section 5.2. For constant voltage, variable current loads, transformer coupled feedback with the output load in series with the feedback loop results in low power dissipation, relatively small degree of stability variations versus output current variations, and non-critical components. The possible limitations are that transformers generally are more expensive than other passive components and are not as readily available in a variety of stock values.

4. INPUT TRIGGER CIRCUITS

The primary function of the input trigger circuit is to initiate the transition from the "off" to the "on" state when there is an input signal. At all other times the input circuit must provide a threshold or margin against false triggering due to noise or spurious disturbances.

Although the input circuit must supply sufficient energy to establish regeneration, it is unnecessary and undesirable that any additional energy be supplied. To do so reduces the gain of the amplifier, since gain may be defined as the ratio of the output power to the input power during one cycle of operation. Because regeneration makes the input and output power independent of each other, any reduction in input power results in greater amplifier gain.

In an amplifier with external feedback coupling it is possible, but not always practical, to have the input circuit trigger the transistor at the collector, base, or emitter terminal. The collector terminal seldom is selected because then the input circuit must supply energy to the output load as well as to the transistor. Also, the base is usually not used (except occasionally with negative resistance feedback) because extra components are required to steer the triggering energy into the transistor and it is diffiult to apply a timing signal.* However, the following discussion and the dc input characteristic of Fig. 7 (a) are equally valid for triggering at the base or emitter terminal of junction or point contact transistors which are short-circuit stable.

One of the simplest types of triggering circuits is shown in Fig. 7 (b). The voltage and current increments assumed necessary to initiate regeneration are designated V_t and I_t. Therefore, the required input signal voltage V_s and current I_s are:

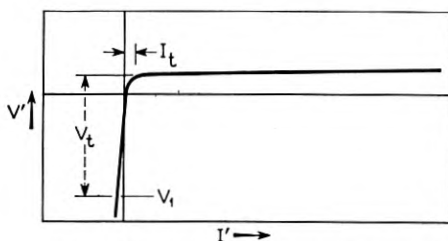
$$V_s \geqq V_t + I_t R_1 \tag{1}$$

$$I_s \geqq I_t \left(1 + \frac{R_1}{R_2}\right) + \frac{V_t}{R_2} + \frac{V_1 - V_2}{R_2} \tag{2}$$

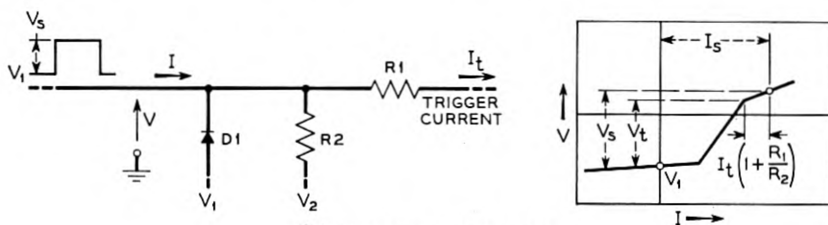
* Also, for junction transistors, about twice as much energy is required to trigger at the base as at the emitter.[4]

The purpose of diode D_1 is to provide a low impedance current threshold, the amount of current given by the last term of (2). This type of threshold is especially effective for preventing false operation from electrostatically induced crosstalk. Also, it allows a faster rate of discharge of stray capacity on the input terminal at the end of the input pulse period.

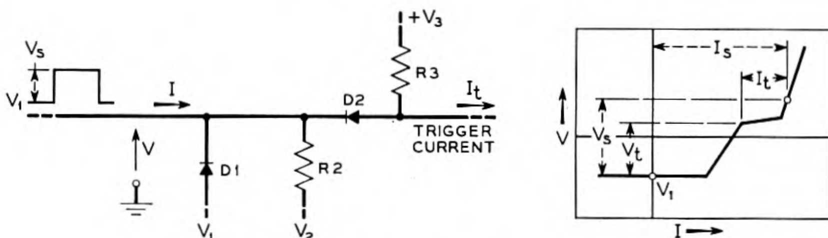
Although the circuit of Fig. 7 (b) is attractively simple, it is undesirably sensitive to variations in signal voltage. An increase in the input pulse voltage causes excessive triggering current and a decrease may easily result in failure to trigger. Since the circuit must be designed to operate reliably with the smallest expected input pulse, it is wasteful of input power with the average amplitude of input pulse.



(a) TRANSISTOR INPUT CHARACTERISTIC



(b) SIMPLE INPUT CIRCUIT



(c) ONE TERMINAL AND-TYPE INPUT CIRCUIT

Fig. 7 — Input trigger circuits.

The single terminal AND-type circuit[9, 10] Fig. 7 (c) has the desirable characteristics of the previous circuit, and is relatively insensitive to input signal variations. In this circuit the input pulse switches the current through R3 into the transistor input and then encounters the relatively high resistance R2, as compared to the parallel resistance of R2 and R1 in Fig. 7 (b). The blocking action of D2 thus reduces variations in the input signal current. However, R2, R3, V_3 and V_2 cannot be increased without limit to reduce the variations; the dc power dissipated in R2 and R3 would become excessive.

Another advantage of the AND-type circuit is that several inputs may be paralleled with a common R3 to provide an AND logic function as well as an input trigger function. This feature, when desired, saves components and does not reduce the gain of the amplifier.

When both the input circuit and the feedback circuit terminate at the same transistor input terminal, as is usually the case, some additional components are generally required to prevent one circuit from shunting the other circuit. To steer the trigger current into the transistor, a diode may be placed in the feedback path so that the diode is reverse biased except when there is feedback current. Similarly, a diode or a resistor may be placed in the input circuit so as to prevent the feedback current from flowing into the input circuit.*

Although the discussion has assumed positive polarity input pulses, the remarks apply equally well to negative pulses if the polarity of the diodes and the supply voltages are reversed.

It is recognized that the preceding remarks assume that the minimum triggering energy is known and that a step function of current or voltage is the optimum form of the triggering energy. Actually, until a study is made of the circuit and transistor parameters (including the non-linear aspects) that affect the initial triggering before the feedback is established, the design of an optimum input trigger circuit will remain an experimental art. Experience with the AND-type input circuit has indicated that appreciably more current is required to trigger junction units than point contact units.

5. OUTPUT COUPLING CIRCUITS

In addition to the obvious function of efficient power transfer from the amplifier to a load, the output coupling circuit is a convenient point at which to perform other functions, as for example, dc level restoration

* This precaution is not necessary if the transistor input exhibits appreciable negative resistance.

and pulse inversion. In a system of logic circuits interspersed with amplifiers at regular intervals, it is apparent that the dc level at similar points, such as the outputs of the amplifiers, must be identical if the amplifiers are to be interchangeable. Without some circuit or element to restore the dc level, the levels along the transmission path will monotonically decrease* due to the dc voltage loss through the logic circuits and across the transistor in the amplifier. The output circuit is one point where restoration of the dc level may be readily combined with other functions.†

In the following two sections three methods of output coupling are discussed and the interaction between the output and feedback circuits is considered.

5.1 *Output Coupling Elements*

Three types of coupling circuits are RC, transformer, and diode coupling. Each of these methods permits the dc level of the signal pulses to be corrected to a predetermined level. However, the restoration,‡ efficiency, and versatility characteristics of each circuit are quite different.

Although RC coupling is common in linear amplifiers, it is seldom used in transistor pulse amplifiers that operate at duty cycles near 50 per cent. The reason is that the time constants encountered do not permit both proper restoration of the capacitor and high efficiency of the output circuit. As indicated in Fig. 8 (a), the transistor is a low impedance in

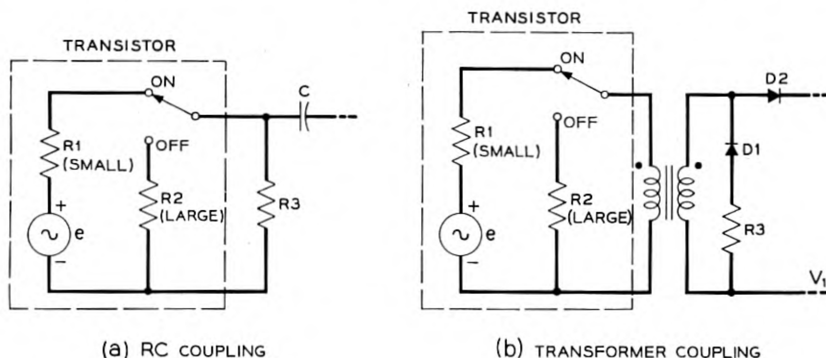


Fig. 8 — Reactive output coupling circuits.

* Decrease for positive pulses; increase for negative pulses.

† An exception, to be discussed, is diode output coupling where it is occasionally more convenient to correct the dc level in the input of the logic circuits or the amplifier.

‡ This refers to restoration of a reactive element (i.e., the return to a quiescent state) and is not to be confused with restoration of the dc level of a circuit.

the "on" state and a high impedance in the "off" state. Since C must be relatively large to make the voltage drop across it small during the pulse duration, R3 must be equal to or smaller than R1 for satisfactory restoration (50% duty cycle assumed). But then the current transmission efficiency of the coupling network is less than 50 per cent because generally R1 is smaller than the input resistance of the driven circuits during the pulse duration. Unless the pulse length is only a small fraction of the pulse repetition period, it is seldom possible to effect a suitable compromise. Also, it might be noted that variations of I_{co} current, which flows through R3, cause variations in the output pulse amplitude. Finally it is not possible to obtain pulse inversion.

A transformer coupled circuit, Fig. 8 (b), works efficiently with a transistor. Diode D2 isolates the transformer from the load and interlead stray capacitance during the interdigit period* so that the restoration time of the transformer is controlled by the value of R3. The restoration time is approximately proportional to the mutual inductance divided by the total shunting resistance. Diode D1 prevents R3 from shunting down the output during the pulse duration, thus permitting high output efficiency and proper restoration.†

As noted in Section 2, the maximum output power from the transistor is determined by the maximum collector voltage (as set by breakdown or punch-through) and the maximum collector current consistent with the permissible dissipation in the transistor. Usually this maximum voltage exceeds the desired amplifier output voltage and, occasionally, the maximum collector current is insufficient; in such instances a voltage step down is desirable. When the transistor is not required to operate at maximum power dissipation, it often is advantageous to balance the "off" and "on" power dissipation. An increase in the collector supply voltage increases the "off" power and decreases the "on" power (by decreasing the required collector current for the same output power). Thus the collector voltage may be adjusted to give the lowest total power dissipation consistent with the average duty cycle of the amplifier. The transformer turns ratio is specified to match the optimum collector voltage to the desired output voltage. Furthermore, I_{co} variations have negligible effect on the output voltage amplitude and pulse inversion (if desired, for example, for inhibition) is possible. For these reasons transformer coupling appears to give optimum output coupling performance.

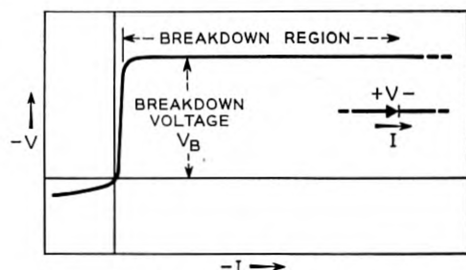
---

* The minimum time interval between the end of one pulse and the beginning of a succeeding pulse; for a 50 per cent duty cycle the interdigit period is equal to the pulse duration.

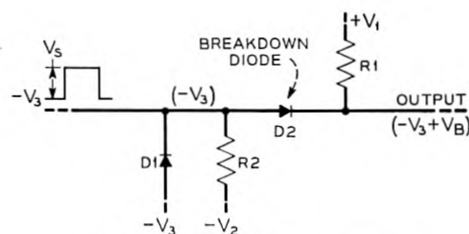
† Occasionally it is possible to specify the collector impedance, the transformer losses, and the reverse impedance of D2 so that D1 and R3 are not necessary.

A third method of coupling, which is attractive for systems using only AND- and OR-type logic, utilizes the reverse characteristic of a breakdown diode, Fig. 9 (a). The interesting feature of this diode is the sharp transition between the high and low incremental resistance regions of the reverse characteristic. With this diode it is possible to shift dc levels by an amount equal to the reverse breakdown voltage of the diode, as indicated in Fig. 9 (b). In the quiescent state D2 operates in the breakdown region and D1 serves to clamp the collector voltage at $-V_3$; during the pulse duration D2 operates in the high resistance portion of its reverse characteristic. If the driven circuit has a voltage threshold, like the transistor threshold in Fig. 7 (a), less than $-V_3 + V_B + V_s$ and $V_s < |V_B|$, the circuit operates like a normal AND-type circuit except for the dc level change. For this reason it is convenient with AND-OR logic circuits to include only D1 and R2 in the output circuit of the amplifier and use D2 and R1 as the AND input elements in the logic circuits.

The principal advantages of diode coupling are simplicity and the lack of an energy storage element. The limitations are that there is no opportunity to match transistor and output conditions, variations in



(a) BREAKDOWN DIODE V−I CHARACTERISTIC



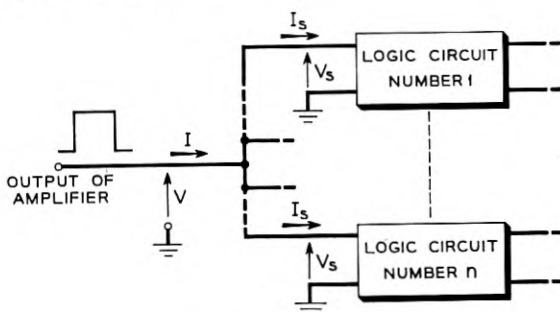
(b) COUPLING CIRCUIT UTILIZING A BREAKDOWN DIODE

Fig. 9 — Direct output coupling circuit.

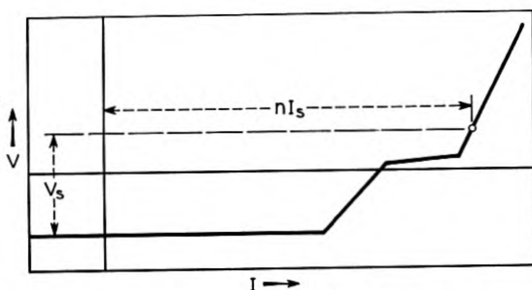
diode breakdown voltage reduce amplifier margins, and pulse inversion is not possible. For these reasons diode coupling has limited utility, but is attractive for some applications.

5.2 *Connection of Output and Feedback Circuits*

The performance of the amplifier is greatly affected by the method used to connect the output circuit and the feedback circuit together at the output of the transistor. Should these two circuits be connected in a shunt or a series fashion? Performance features, such as rise time, sufficient output voltage, degree of stability versus load current variations, and power dissipation directly depend upon this choice. With transformer output coupling, the choice always exists; with other types of output coupling the choice may or may not exist, depending upon the type of feedback coupling. The following discussion is in terms of transformer coupled output and feedback circuits and the general conclusions may be extended to other cases.



(a) CONNECTION OF AMPLIFIER LOAD



(b) V-I CHARACTERISTIC OF AMPLIFIER LOAD

Fig. 10 — Output load characteristic.

The principal factor that influences the choice of the output-feedback connection is the nature of the output load of the amplifier. In the majority of computer and switching systems the amplifier must drive a multiplicity of paralleled load circuits, as indicated in Fig. 10 (a). The input characteristic of each load circuit is assumed to be of the threshold type, like the AND-type input characteristic of Fig. 7 (c), which results in the amplifier load characteristic of Fig. 10 (b). During the initial portion of the rise time of the output pulse the incremental impedance is almost zero and during the remainder of the pulse duration it is relatively large. Due to the voltage threshold nature of the load, the amplifier load variations are current variations at a constant voltage. The minimum current is encountered in the system position where the amplifier drives the smallest number of logic circuits, often a single logic circuit; the maximum current is limited by the maximum output power of the amplifier. Although a desirable ratio of maximum to minimum current may be as high as 20:1, the amplifier is expected to exhibit optimum performance at any load current within this range.

The shunt connection of the output and feedback circuits is illustrated in Fig. 11.[12] Windings $1{:}n_1$ constitute the feedback coupling and $1{:}n_2$ the output coupling. The two circuits shunt each other in the sense that the ratio of the feedback to the output current is determined by the ratio of the impedance of these circuits as modified by the turns ratio of the transformer.

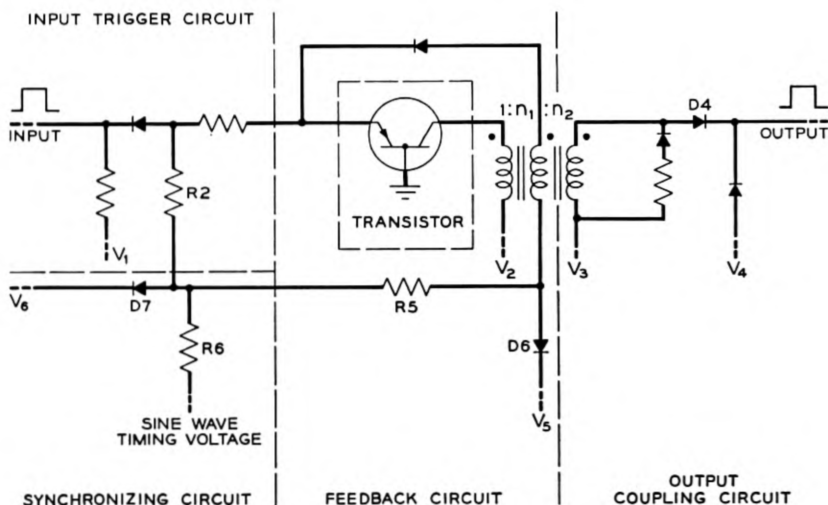


Fig. 11 — Shunt connection of output and feedback.

There are two limitations associated with this output-feedback connection. In the first place there is the possibility of insufficient output voltage, slow rise time, or complete failure of regeneration. This is caused by the shunt effect of the output load which places an almost zero initial incremental impedance across the feedback path. In order to overcome this limitation a current switch (R5 and D6 in Fig. 11) is used to obtain a low initial feedback impedance and the output diode (D4) is reverse biased so that the initial load impedance is large. The price paid is the undesirable power dissipation in the current switch. Moreover, stray capacity across the output terminal or a load current that exceeds the design value may still result in a long rise time, low output voltage, or regeneration failure.

The series connection of the output and feedback circuits is shown in Fig. 12. In this connection the output load is in series with the feedback loop. Thus, the transistor output current, feedback current, and output load current are all proportional to each other. This situation assures regeneration regardless of output load current variations.

The regeneration cycle of the series type amplifier is as follows. In the quiescent state diode D2 is reverse biased by V1 to prevent false triggering. After the arrival of an input signal, the timing signal voltage goes positive and steers the trigger current into the transistor. No

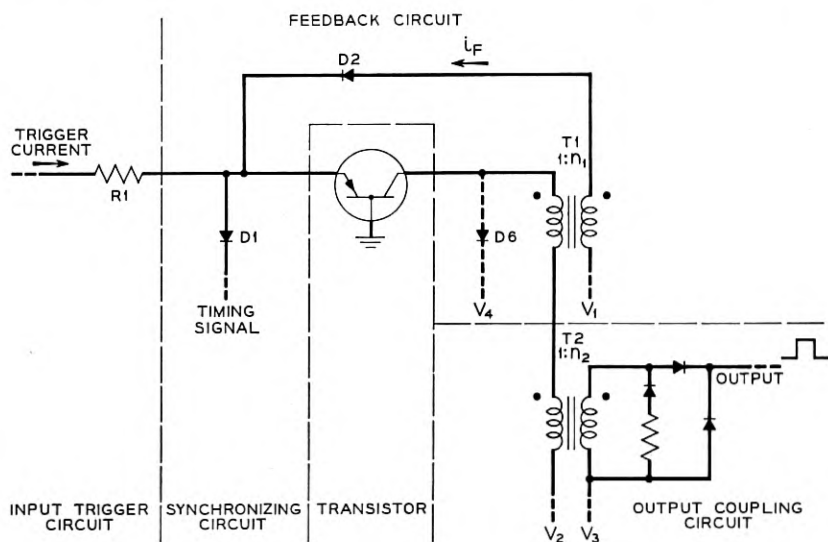


Fig. 12 — Series connection of output and feedback.

appreciable output current flows until the voltage across transformer T1 is sufficient to forward bias diode D2. Then both the feedback and output current build up simultaneously and rapidly since the turns ratio $1:n_1$ of T1 is selected to give a feedback loop gain greater than unity. When the sum of the voltages across the primaries of the feedback and output transformers almost equals the collector supply voltage, the transistor saturates and stabilizes the feedback loop. At the end of the pulse duration the timing signal voltage goes negative and robs current from the feedback loop, thus forcing the transistor out of saturation and causing the amplifier to turn off.

Because the feedback current is proportional to the output current during the rise time, the amplifier can deliver any value of load current up to the current corresponding to the maximum allowable collector current. Also, assuming that the leakage inductances of the transformers are small, a large stray capacitance across the output terminal does not appreciably degrade the rise time. Since a current switch is not necessary, the standby power dissipation in the feedback loop is negligible. These are the outstanding features of the series connection.

Two important performance considerations of the series type amplifier are the change in the degree of stability versus load current variations and the action of the amplifier when the timing signal fails. Both of these items may be controlled by the selection of suitable values for the turns ratio and the primary inductance of the feedback and output transformers.* In order to prevent burnout of the transistor in the event that the timing signal fails, the amount of excess feedback current must decrease during the pulse duration. Due to the low impedance of the feedback loop, this condition may be approximately† stated in terms of the primary inductances as:

$$\left|\frac{V_1}{n_1 L_1}\right| > \left|\frac{V_2 - \dfrac{V_1}{n_1} - V_{sat}}{L_2}\right| \left|1 - \frac{n_1}{\alpha}\right| \tag{3}$$

where V_{sat} is the collector saturation voltage and L_1 and L_2 are the primary inductances of T_1 and T_2 respectively.

The degree of stability in the series amplifier at the end of the pulse duration is proportional to the output load current. This situation may be seen more clearly if a "catching" diode (D6 in Fig. 12) is added to the

* If the transistor is not short circuit stable, it is also usually necessary to use a small resistance in series with the emitter.

† The principal approximation is that alpha is constant versus collector current. The value of alpha at the end of the pulse duration is a conservative value.

circuit to prevent saturation in the transistor.* Because the feedback loop gain, as determined by the alpha of the transistor and the turns ratio n_1, must be greater than unity for regeneration reasons, there will be current flow through D6 during the pulse duration. This current is proportional to the degree of stability. An increase Δi_{out} in the output current causes an increase of

$$\Delta i_c = \frac{\alpha n_2 \Delta_{i\,\text{out}}}{n_1} \tag{1}$$

in the collector current. Therefore, the current in D6 increases by an amount equal to

$$\Delta i_{\text{D6}} = \left[\frac{\alpha}{n_1} - 1\right] n_2 \Delta_{i\,\text{out}} \tag{5}$$

This variation in the degree of stability may be reduced by selecting α/n_1 close to unity and reducing n_2. However, since it is desirable to have α/n_1 much larger than unity for short rise time and since any reduction in n_2 increases the I_{co} standby power,† a compromise is necessary.

6. SYNCHRONIZING CIRCUITS

The majority of modern digital data processing systems employ coincidence gate circuits to perform the logical functions. In order to insure that digit pulses will coincide at the inputs to the logic circuits, it is convenient to synchronize the amplifiers. Usually a master oscillator, or "clock," produces the timing signals that are distributed to the amplifiers. The function of the synchronizing circuit in the amplifier is to turn on and to turn off the amplifier at predetermined time intervals in response to the clock signal.

In a regenerative amplifier there is always a small delay from the time triggering commences until the full output pulse is developed. Then there are variations in the transmission time to other amplifiers. For these reasons the clock signal must lag the input signal to the amplifier in order to maintain control of turn-on and to obtain a uniform pulse length from all amplifiers. Generally the time lag is one-fourth of the

---

* In an actual amplifier D6 is not required if the transistor saturation voltage is relatively constant versus collector current and the pulse fall time is not adversely affected by minority carrier storage in the transistor. Often the inductive "kick" of the transformers and the regenerative feedback are sufficient to make the minority carrier storage effect negligible. If D6 is used, its reverse recovery time may adversely affect the pulse fall time, thus nullifying its usefulness.

† The I_{co} standby power is proportional to V_2, which, for a given output voltage, is inversely proportional to n_2.

repetition period and, in such a case, the clock signal is made available in four phases.

Although the clock signal may have any one of a number of forms, a sine or a square wave are the most common forms. Usually a sine wave is preferred because it is simpler to distribute to a large number of amplifiers. Exceptions occur in cases where exceptionally precise timing is necessary, or the use of a square wave requires considerably less clock power. In the following discussion of where to synchronize, a square wave will do as well or better than the assumed sine wave. With either signal it is desirable to keep the clock power to a minimum.

If the synchronizing circuit is to be effective, the clock signal must be capable of accomplishing the following actions:

a. It must be able to hold the transistor in the "off" state in the presence of trigger current in order to control turn-on.

b. At the turn-on time it must rapidly inject the trigger current into the transistor.

c. At the turn-off time it must alter the conditions in the feedback loop in such a manner that the transistor turns off promptly.

In other words the synchronizing circuit must act like an inhibit logic circuit with the clock signal appearing as the inhibit signal during the interdigital period.

It is recognized that there are many amplifier configurations and several ways to synchronize each configuration. Generally it is preferable to synchronize at only one input terminal of the transistor or at only one point in the feedback circuit. A relatively complete discussion can be given with the aid of the following four examples.

A circuit that employs negative resistance feedback, such as in Fig. 4, requires a relatively large amount of clock power for synchronization. Because a capacitor (C2) is required on the emitter for regeneration,[2] the clock signal must be applied to the base of the transistor to control turn-on accurately. As far as turn-off is concerned, another clock signal might be applied to the emitter or to the current gate in the feedback circuit. However, this would result in additional components, a second clock signal 180° out of phase with the base clock signal, and approximately the same required clock power as if the base clock signal alone were used. Turn-off at the emitter is impractical due to the negative resistance characteristic. The power that the clock signal on the base must furnish is made up of two parts. One part is the average standby power that is absorbed every time the clock voltage is positive. It is composed of the I_{co} power supplied to the transistor plus the power dissipated in in R1 and R2. R2 and D2 serve to reduce the clock current in R1 and

R2, but the maximum value of R2 is limited by stray capacitance from the transistor base to ground.* The average clock standby power for this circuit (with a 10 volt peak clock voltage) is approximately 13 milliwatts. The second part of the clock power occurs at turn-off when the clock must supply approximately the full "on" state collector current. In this design the clock supplies about 20 milliamperes of current for 0.1 microsecond at voltages up to about 6 volts peak before the transistor turns off. Therefore, a negative resistance feedback circuit usually requires a relatively large amount of standby clock power continuously and a high peak clock power at turn-off. Also it should be noted that diode D1 must have a short reverse recovery time in order to prevent false triggering during the negative portion of the clock cycle.

A second example of synchronization is shown in Fig. 11. Here the clock signal is introduced in the feedback circuit to control turn-off. It is also applied to R2 in the input circuit so as to control turn-on. In this circuit most of the clock power is dissipated in R5 and R6 when the clock voltage is positive during the output pulse time slot (whether or not an output pulse is produced). Necessarily, this power is relatively large because the clock must supply the full amount of feedback current. Also, it is necessary to clip the positive peak of the clock voltage in order to prevent false triggering via R2 when there is no input pulse. A square wave clock signal would eliminate the need for R6 and D7, but would not change the power in R5. The average clock power in a typical circuit of this type is approximately 20 milliwatts, which is relatively large. The principal advantage of this method is that diode reverse recovery time is not a problem.

A third method of synchronization is to apply a square wave clock signal (a sine wave is not suitable in this case) between the base of the transistor and ground (for example, assume in Fig. 11 that R2 and R5 are returned directly to V6 and that the base of the transistor is the clock terminal instead of ground). Before turn-on the clock voltage must be more positive than the trigger voltage on the emitter. At turn-on the clock voltage drops rapidly to ground potential and triggering takes place. During the pulse duration the base current of the transistor is supplied by the clock source. At turn-off the clock voltage must rise rapidly several volts until D6 conducts and robs current from the feedback loop. The clock power required by this method is relatively large (order of 20 milliwatts) for point contact transistors because the base current of such units is large. In a junction transistor with alpha close

* The capacitance causes the base voltage to lag the clock voltage at turn-on if R2 is large, which degrades the timing.

to unity the base current is small and the required clock power may be as low as 3 milliwatts. However, it should be noted that this method of synchronization applies only to amplifiers with a gated feedback circuit (such as R5 and D6 in Fig. 11). In other circuits (Fig. 12, for example), a clock voltage applied to the base terminal of the transistor may never be able to turn off the transistor (the feedback current may actually increase instead of decrease). Thus, this method of synchronization is limited and is a low power method only when used with junction transistors.

A fourth synchronization method, which avoids the limitations cited in the previous examples, is illustrated in Fig. 12. The timing circuit is simply diode D1. The operation of the circuit, which is like an inhibit logic circuit, is as follows. When trigger current commences, the clock voltage is negative and D1 conducts the trigger current away from the emitter terminal. As the clock voltage rises positiveward, the emitter voltage follows until it reaches the threshold voltage of the transistor, usually ground potential. Then the trigger current flows into the transistor which turns on. As the clock voltage continues positiveward the emitter conduction clamps the emitter voltage so that D1 opens and the clock does not shunt the feedback path during the pulse duration. At the end of the pulse duration the clock voltage goes negativeward through ground potential and D1 becomes conducting. This action robs current from the feedback loop, thus causing the transistor to turn off. If no input pulse is present, D1 is always non-conducting and any small reverse leakage current is drained off through R1 (which is returned to voltage V1).

Because diode D1 is always non-conducting when no input pulse is present, the standby clock power is essentially zero. During a pulsing cycle the clock conducts only a small current before turn-on and only instantaneously at a low voltage at turn-off. Hence, the required clock power is usually less than two milliwatts.

It is important to note that the amplitude of the negative peak of the clock voltage usually should not be more negative than the quiescent bias voltage on the emitter. If it should be, D1 will conduct and, due to minority carrier storage, may cause false triggering when the clock voltage goes positive. The current through D1 at turn-off might have the same effect in the succeeding cycle except that the flyback voltage of the transformers during the interdigit period removes the minority carriers from both D1 and D2. Since D2 carries a larger current for a longer period than D1, the carriers are cleared from D1 first. It is then reverse-biased for almost one-half the repetition period before there is any chance

of false triggering. Hence, diode reverse recovery time is not a problem. However, D1 should have a short forward recovery time in order that turn-off will occur rapidly.

One possible limitation of this synchronization method is that a low impedance clock source is necessary. This is usually not difficult to obtain with a resonant circuit in the output of the clock signal source. Offsetting this point are the advantages of low clock power, esentially zero standby clock power, only one additional component, and no critical component tolerances.

7. ILLUSTRATIVE DESIGN

In the preceding sections the features of various configurations for the functional circuits of an amplifier have been described. The following discussion illustrates the application of these ideas to an amplifier design for use in a digital computer system. It is intended that the description of the design philosophy be sufficient to permit its application to other systems.

In the computer under consideration the amplifier is to be combined with a single level, diode logic circuit to form a logic network. The logic networks, together with delay lines, will be connected in appropriate arrays to perform the logic functions of the sytem, such as addition, multiplication, etc. Digital information is to be represented by one-half microsecond pulses and the amplifiers are to be synchronized at a one megacycle pulse repetition rate by a four phase sine wave master oscillator. Other system requirements are mentioned in connection with the selection of the corresponding functional circuit.

Since the amplifier is considered as a small system of functional circuits, it is necessary, as in most system designs, to re-examine, and possibly change, circuit choices as the design progresses. However, for the sake of clarity, the following discussion omits the re-examination and frequently refers to the final schematic shown in Fig. 13.

The first step in the design is to select the feedback configuration most suitable to the computer requirements. For this computer the dc and clock power are to be minimized and the amplifier should be able to drive from 1 to 12 logic networks. Miniaturization of the computer implies that there may be an appreciable amount of stray capacity across the amplifier output. These considerations suggest transformer coupled feedback connected in series with an output circuit. Since both positive and negative output pulses are to be required (one polarity for AND and OR logic and the other polarity for inhibition), transformer output coupling is indicated.

The next basic selection is the choice of an appropriate transistor. In this computer it is expected that pulses will occur in only about one third or less of the pulse time slots due to the nature of the digital information. In order to minimize the dc standby power an alloy junction transistor is a logical choice for this application because of the low I_{co} current. However, even with a junction unit possessing an alpha cut-off frequency of eight megacycles, it is difficult if not impossible to obtain acceptable gain and rise time with the desired output load current at a one megacycle repetition rate. If the rise time is improved by increasing the trigger current, the gain is decreased. The principal cause of the poor "gain-bandwidth" appears to be the depletion layer capacitance.[11] The difficulty can be overcome by selecting a point contact transistor. A particular germanium transistor coded GA-52996* appears to be suitable and has the following pertinent characteristics:

a. Collector capacitance less than 0.5 uuf.
b. Alpha cut-off frequency in excess of 80 mc.
c. Base resistance less than 100 ohms.

Since the alpha of this unit is greater than 2 at collector currents of the order of 10 ma, the common base connection will yield the greatest current gain. The disadvantage of a point contact unit, of course, is the I_{co} current. For this reason the amplifier will have to be designed to use the smallest possible collector supply voltage.

The point contact transistor, due to its high cut-off frequency relative to the amplifier pulse repetition rate and its high alpha at small emitter

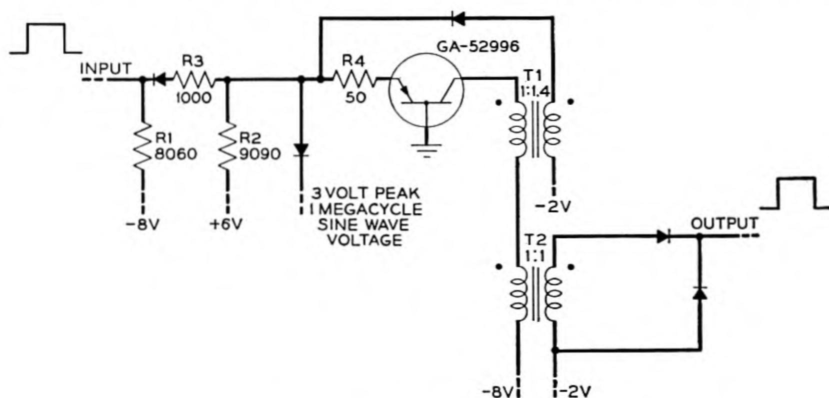


Fig. 13 — Illustrative design

* This is a relatively special unit especially suited for high speed switching applications.

currents,* permits the use of a simple input circuit. The AND type input circuit is suitable and desirable for another reason. When AND type logic is added to the amplifier, it may be paralleled with the basic input circuit and the input sensitivity of the complete network will be the same as for the amplifier alone. Other logic circuits will be added to an amplifier in a manner similar to that described by Felker[2] so that the input sensitivity will be reduced at most by the voltage drop across one series diode (approximately 0.3 volts).

The input pulse voltage and current requirements depend upon the voltage threshold necessary to prevent false operation and the minimum trigger current for reliable regeneration. A test of several sample transistors indicates that approximately 0.3-ma emitter current is required to trigger the transistor with an estimated collector supply voltage of 10 volts. The emitter breakpoint† voltage is found to vary between −0.25 and +0.25 volts. To allow for aging variations of the transistor and of R2, it seems reasonable to use a 6-volt source and R2 equal to 9090 ohms, which results in a trigger current a little more than twice the required minimum. Previous experience with computers of this type indicates that a 2-volt threshold will be sufficient to prevent false triggering. Thus, the secondary winding of the feedback transformer is returned to −2 volts and R1 is chosen to give a quiescent emitter voltage of −2 volts. With these considerations and an estimated voltage drop across R3, the input pulse amplitude is calculated to be 2.3 volts and 0.9 ma. Allowing 0.3 volts for a series logic diode, the minimum output voltage and current of the amplifier are 2.6 volts and 0.9 ma per driven network.

The selection of the collector supply voltage and the turns ratio of T2 depends upon the dc power dissipation due to I_{co} current and output voltage regulation versus collector current. For this transistor a unity turns ratio appears to represent a reasonable compromise. Then, by estimating the voltage drops across T1, T2, and the transistor, it is found that a collector supply voltage of −8 volts is sufficient to produce an output pulse voltage about 0.5 volt greater than the required miminum.

The next step is the selection of the turns ratio of T1 and the primary inductances of both T1 and T2. The two considerations involved are sufficient feedback with the minimum output current (the worst case with respect to feedback) and the maximum collector dissipation in the event that the clock fails. By means of the formulas and assumptions indicated in section 5, primary inductance values of 0.4 mh for T1 and

* Usually $\alpha > 4$ for $i_e = 0.5$ ma.
† The transition point of the emitter diode from cut-off to conduction.

0.2 mh for T2 together with a turns ratio of 1.4 for T1 are selected. Since the GA-52996 transistor is not quite short circuit staple, a 50-ohm resistor is added in series with the emitter. The excess emitter current at the end of the pulse duration is greater than 2 ma, thus assuring sufficient stability, and, if the clock fails, the amplifier will turn off by itself in approximately 7 μsec, at which time the instantaneous collector dissipation will be approximately 240 mw (considered to be a safe instantaneous dissipation for this transistor).

For low clock power and circuit simplicity the single diode synchronizing circuit is chosen. Although a peak clock voltage of 2 volts would normally be used (this value corresponds to the quiescent emitter bias voltage) it is found that the clock may be varied between 1 volt and 6 volts peak without a failure occurring. Therefore, the nominal clock voltage is set at a centered value of 3 volts peak. The dc level of the clock voltage is 0 volts, which approximately corresponds to the emitter break point voltage of the transistor. This concludes the basic selections in the design procedure.

The power dissipated in the amplifier is quite modest. In the quiescent state the amplifier absorbs only 0.2 mw average clock power and 30 mw dc power (this would be only 10 mw if the I_{co} power were negligible). When the amplifier is pulsing every microsecond the dc power is 50 mw and the averge clock power is 2 mw. Since the amplifier is so conservative of power, it is possible to use 4,000 networks in a computer and require less than 200 watts dc power.

One indication of the component sensitivity of a pulse amplifier is the magnitude of the supply voltage margins. In this amplifier the supply voltages may be varied, one at a time, over ± 12 per cent of the nominal values before a failure occurs. Generally margins of this magnitude under the worst conditions are considered sufficient to guarantee against failures caused by aging, or to insure that such failures will be indicated by routine checks before they occur. It is interesting to note that in a temperature test the amplifier continued to operate properly over a temperature range from -20 to $+80°C$. Even at $+75°C$ the supply voltage margins were 10 per cent or better.

8. SUMMARY

A method of analysis and design procedure have been presented in which a transistor regenerative amplifier is considered as an interconnected system of functional circuits. Each functional circuit may be evaluated or chosen in terms of the requirements of the complete digital system in which the amplifier is to be used. In general no particular cir-

cuit or collection of circuits can result in an amplifier suitable for use in every type of digital system. The use of an AND type input circuit, transformer coupled output and feedback circuits, and an inhibit type synchronizing circuit appear to be an optimum set of functional circuits to make up an amplifier for use in a synchronous digital computer system employing passive logic circuits. An illustrative design is presented for such an amplifier which operates at a pulse repetition rate of 1 mc, uses 12 components (none of which are especially critical), requires an average of 40-mw dc power and 1-mw clock power, is capable of driving from 1 to 12 similar amplifiers, and has voltage margins in excess of 12 per cent. Although the design philosophy was developed for this type of amplifier, it is believed that much of the philosophy is applicable to regenerative amplifiers for use in other digital data processing systems.

9. ACKNOWLEDGEMENT

The final design and the performance data of the illustrative amplifier are due to L. C. Thomas and H. E. Coonce. The author also wishes to express his appreciation for the many helpful and stimulating discussions with other colleagues, especially A. J. Grossman, T. R. Finch, J. H. Felker, and J. R. Harris.

REFERENCES

1. S. Greenwald, et al., SEAC, Proc. I.R.E., Oct., 1953.
2. J. H. Felker, Regenerative Amplifier for Digital Computer Applications, Proc. I.R.E., Nov., 1952.
3. J. L. Moll, Large-Signal Transient Response of Junction Transistors, Proc. I.R.E., Dec., 1954.
4. J. G. Linvill and R. H. Mattson, Junction Transistor Blocking Oscillators, Proc. I.R.E., Nov., 1955.
5. A. E. Anderson, Transistors in Switching Circuits, B.S.T.J., Nov., 1952.
6. T. E. Firle, et. al., Recovery Time Measurements on Point-Contact Germanium Diodes, Proc. I.R.E., May, 1955.
7. S. L. Miller and J. J. Ebers, Alloyed Junction Avalanche Transistors, B.S.T.J., Sept., 1955.
8. J. J. Ebers and S. L. Miller, Design of Alloyed Junction Germanium Transistors for High Speed Switching, B.S.T.J., July, 1955.
9. T. C. Chen, Diode Coincidence and Mixing Circuits in Digital Computation, Proc. I.R.E., May, 1950.
10. L. W. Hussey, Semiconductor Diode Gates, B.S.T.J., Sept., 1953.
11. J. M. Early, Design Theory of Junction Transistors, B.S.T.J., Nov., 1953.
12. Q. W. Simkins and J. H. Vogelsong, Transistor Amplifiers for Use in a Digital Computer, Proc. I.R.E., Jan., 1956.
13. M. Tanenbaum and D. E. Thomas, Diffused Emitter and Base Silicon Transistors, B.S.T.J., Jan., 1956.

Observed 5–6 mm Attenuation for the Circular Electric Wave in Small and Medium-Sized Pipes

By A. P. KING

(Manuscript received March 20, 1956)

At frequencies in the 50–60 kmc region the use of circular electric wave transmission can provide lower transmission losses than the dominant mode, even in relatively small pipes.

The performance of two sizes of waveguide was investigated. In the small size (7⁄16″ I.D. × 1⁄16″ wall) the measured TE_{01} *attenuation was approximately 5 db/100 ft and is appreciably less than that of the dominant mode. The measured attenuation for the medium sized (7⁄8″ I.D. × 1⁄8″ wall) waveguide was 0.5 db/100 ft which is about one-fourth that for the dominant mode.*

This paper also considers briefly some of the spurious mode conversion-reconversion effects over the transmission band and their reduction when spurious mode filters are distributed along the line. Allowance has been made for the added losses due to oxygen absorption when air is present.

INTRODUCTION

Since 5.4-mm dominant-mode rectangular waveguide has attenuations of the order of 60 db/100 ft, another transmission technique is required in applications which involve appreciable line lengths. Losses may be reduced by the use of oversize waveguide; some earlier work with dominant mode transmission in slightly oversize round waveguide (two or three propagating modes) has been reported.[1] The possibility of still lower losses exists with circular electric wave transmission in an oversize round waveguide. Miller and Beck[2] have computed the theoretical relative transmission losses of the TE_{01} and TE_{11} modes as functions of

[1] A. P. King, Dominant Wave Transmission Characteristics of a Multimode Round Waveguide, Proc. I.R.E., **40,** pp 966–969, Aug., 1952.
[2] S. E. Miller and A. C. Beck, Low Loss Waveguide Transmission, Proc. I.R.E., **41,** pp 348–358, March, 1953.

guide size and frequency. At 5.4 mm, a 7⁄16″ I.D. waveguide has an appreciably lower attenuation with the circular electric mode than with the dominant mode. A 7⁄8″ I.D. guide has a circular electric attenuation approximately one-fourth that of the dominant mode in the same pipe.

It is the purpose of this paper to present some experimental results which have been observed with circular electric wave transmission in the 5–6 mm wavelength region. The attenuation for three different lines and the transmission variations due to moding effects are reported. Allowance for the loss due to oxygen absorption has been included.

DESCRIPTION OF THE TEST LINES

The TE_{01} mode attenuation measurements were made on approximately straight runs of line ranging from about 100 to 200 feet in length. The copper pipe comprising these lines is believed to conform to the best tolerances and internal smoothness which are current manufacturing practice for waveguide tubing. The relative tolerances and their effect upon transmission are considered in a later section. Three kinds of copper line were measured: a waveguide of oxygen-free copper, one line of low phosphorous-deoxidized copper and one line of steel with a 20-mil low phosphorous-deoxidized copper inner lining. The oxygen-free high-conductivity-copper with its higher conductivity and somewhat greater ductility was chosen to provide comparative performance data with the low phosphorous-deoxidized copper which is commonly used in waveguide manufacture. A waveguide whose outer wall is constructed of steel to provide the necessary strength and wall thickness to support a very thin copper inner wall has the advantage that such waveguide requires less copper. This composite wall tubing was obtained to ascertain whether the tolerances and the nature of the inner surface would yield transmission data comparable to solid copper waveguide.

The lines were supported on brackets which were accurately aligned and spaced at 6-ft intervals. Although the brackets provided for an accurately straight line, the manufactured pipe was not perfectly straight but, in some samples, varied as much as 3⁄8″ in a 12-ft length. Installing the pipe on the brackets tended to straighten the line and reduce these variations to about half this amount. A general view of the lines is shown in the photograph of Fig. 1.

The sections of waveguide were joined together with a more or less conventional threaded coupling, but with one very important difference. The threads, which are cut at the ends of each section, are cut relative to center of the inside diameter and not the outside diameter. This is achieved by employing a precision pilot to provide a center for the cut-

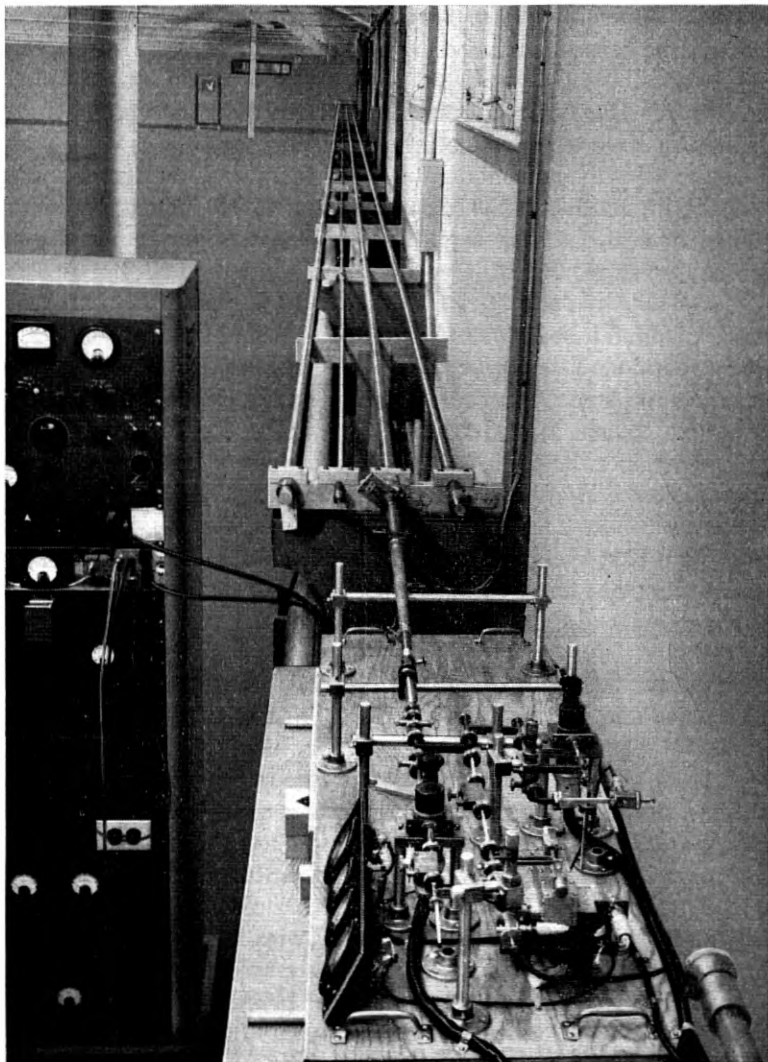


Fig. 1 — General view of the circular waveguide lines and the millimeter wave measuring equipment.

ting die. Since the internal diameter is made as precise as possible, the variations of outside diameter become a function of the tolerances of both the internal diameter and wall thickness and cannot be as precise as the inside of the pipe. Any thread cut relative to the outside diameter as in regular plumbing practice, will not, in general, be concentric to the

inside wall. To avoid an offset at the joint it is therefore important that the thread be centered relative to the inside diameter. After a section was threaded the ends were faced off to make the ends square and thus avoid any tilt between sections when the ends are butted together.

Of the two sizes tested the smaller diameter (7⁄16″ I.D. × 1⁄16″ wall) was chosen to provide a moderate line loss, while limiting the number of propagating modes. In the band of interest (5.2–5.7 mm) the theoretical TE_{01} wave attenuation is about 4 db/100 ft. The number of modes which can be supported at $\lambda = 5.2$ mm is limited to 12 modes and to only one of the circular electric modes. The higher order TE_{on} modes are beyond cut-off. These features limit the number of spurious modes and simplify the mode filtering problem. Furthermore, in this smaller sized waveguide, the associated components which may set up TE_{on} waves, for example conical tapers, need not be as long proportionately as in larger waveguides. The 7⁄16″ I.D. guide has the advantage of smaller size, lower cost and greater ease of transmitting TE_{01} through specially constructed bends. The attenuation of this smaller diameter guide is large enough that system requirements will usually restrict its usage to lengths of line of a hundred feet or so.

The larger size (7⁄8″ I.D. × 1⁄8″ wall) is exactly twice the diameter of the small size discussed in the preceding paragraph but has only one-tenth the attenuation, or about 0.4 db/100 ft. The low loss of this larger size becomes more attractive for runs as long as several hundred feet. This diameter guide will, of course, support more modes, 50 at $\lambda = 5.2$ mm; four of which are circular electric modes — TE_{01}, TE_{02}, TE_{03} and TE_{04}. Some of the disadvantages which accompany the increased diameter are: (1) greater care must be taken as to line straightness, (2) longer conical tapers are required when converting from one guide diameter to another, and (3) longer mode filters are required since the desired mode-filtering attenuations vary inversely with the filter diameter at a given frequency. Flexible spaced-disk lines employed as uniform bends for TE_{01} transmission require much greater bending radii than bends in the smaller diameter guide if the bend loss is to be kept proportionately low. This problem is considered in some detail in another paper.[3] With reasonable care the accumulative effect of these foregoing factors can be held to a reasonably low value. Expressed in terms of the ratio of measured to theoretical attenuation the values are, on the average, about 10 per cent higher in the 7⁄8″ I.D. waveguide than in the 7⁄16″ I.D. waveguide.

---

[3] A. P. King, forthcoming paper on bends.

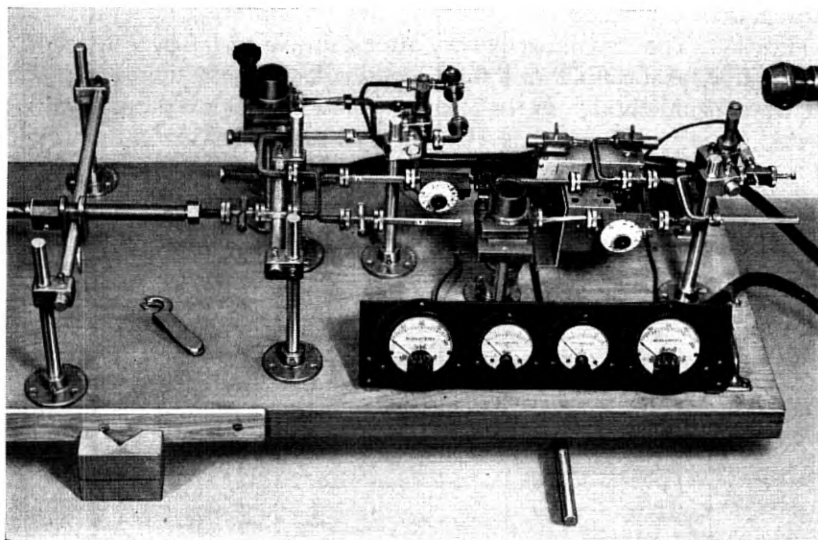


Fig. 2 — Waveguide portion of millimeter wave measuring set.

MEASURING PROCEDURE

With straight runs of round, TE_{01} waveguide lines whose length lies in the 100–200 ft range, it is convenient to make attenuation measurements on a round trip basis. This method has the advantage of convenience in that the attenuation can be measured directly by using a waveguide switch but has the disadvantage of requiring a careful impedance match of the measuring equipment to the line. Fig. 1 shows an overall view of the lines; Fig. 2 shows the arrangement of the 5–6 mm measuring set, and Fig. 3 shows a block diagram of the set-up employed.

This measuring set makes use of two klystrons developed by these laboratories.[4] The double detection receiver features a separate beating oscillator klystron which is frequency modulated and a narrow band (1.7 mc at 60 mc) IF amplifier. The resulting IF pulses are detected with a peak detector and then amplified to provide the usual meter indication. This method with its circuitry has been developed by W. C. Jakes and D. H. Ring,[5] and provides a greater amplitude stability than is possible with a cw beating oscillator.

In the waveguide schematic of Fig. 3 about a tenth of the power is

[4] E. D. Reed, A Tunable, Low Voltage Reflex Klystron for Operation in the 50–60 Kmc Band, B.S.T.J., **34**, p. 563, May 1955.
[5] W. C. Jakes and D. H. Ring, unpublished work.

taken from the signal oscillator to provide monitoring and wavemeter indication. The remaining power, after suitable padding, is fed into a 3-db directional coupler or hybrid junction 2. This junction is employed as a waveguide bridge so that, when arms A and B are properly terminated, no power flows in receiving arm C. Any reflection in line A will,

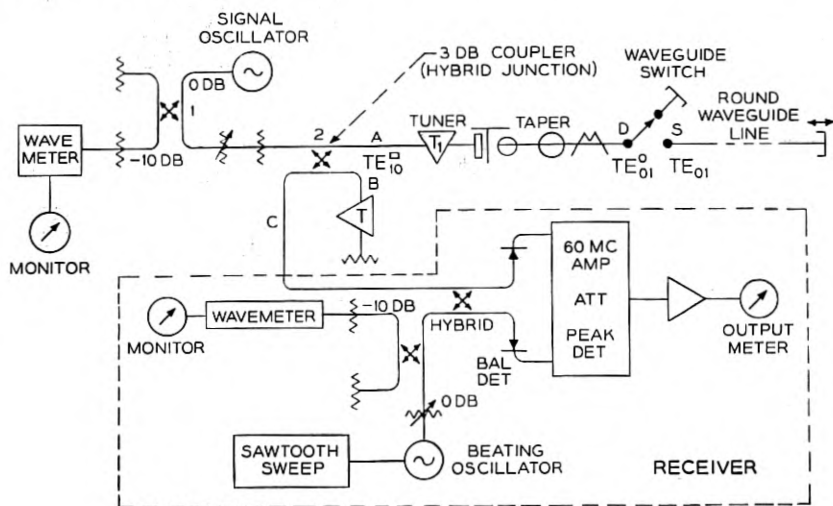


Fig. 3 — Schematic of measuring equipment.

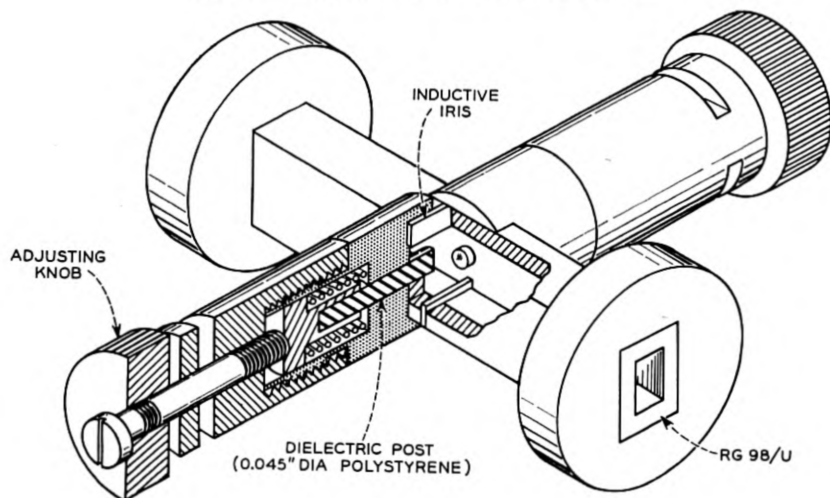


Fig. 4 — Structure of impedance matching tuner.

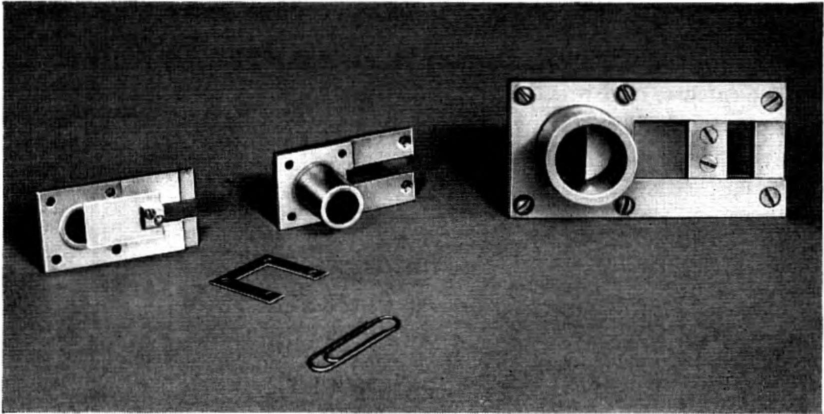


Fig. 5 — Structure of waveguide switch.

however, produce a power flow in the arm C to the balanced converter of the receiver and an indication in the output meter. So far this set is similar to a setup for measuring the round trip loss in a terminated waveguide system. The impedance of the $TE_{10}^{\square} \rightleftarrows TE_{01}^{\bigcirc}$ wave transducer,[6] taper section and mode filter connected as shown in the section A–D of Fig. 3 can be matched to the rectangular waveguide at A by an appropriate adjustment of the dielectric post tuner[7] T_1 whose structure is shown in Fig. 4. Under these conditions a conical taper termination placed in the round waveguide at D will again produce a balance and again no power will flow in arm C. A waveguide switch whose structure is shown in Fig. 5 is connected between the point D and the line under test. A movable short at the far end of the line completes the set-up.

With the impedances matched as described above, the only reflection which reaches the receiver will be from the far end of the line when the switch S is open or, when shorted, from the switch itself. The round-trip attenuation is the difference in attenuation measured for the two positions of the switch. By means of a movable short at the far end of the line, the line length can be varied to produce mode conversion and mode reconversion effects, and the resultant variation in TE_{01} mode transmission can be observed. This phenomena is described in some detail elsewhere.[8]

[6] Reference 2, page 354, Fig. 14.

[7] C. F. Edwards, U.S. Patent 2,563,591, Aug. 7, 1951. The millimeter tuner employs an adjustable dielectric post in place of a metallic tuning screw described in the patent.

[8] Reference 2, pp 356, 357.

LOSSES DUE TO OXYGEN ABSORPTION

In addition to the losses which result from imperfect conductivity, surface effects, and mode conversions, there is a very appreciable loss due to oxygen absorption when the guide is open to the atmosphere. In a waveguide the loss due to O_2 absorption is:

$$\frac{A}{\sqrt{1 - \nu^2}} \tag{1}$$

where

A is the absorption due to oxygen in the atmosphere

$\nu = \lambda/\lambda_c$

λ = free space wavelength

$\lambda_c = \frac{\pi d}{k} = \frac{\pi d}{3.83}$ = cut-off wavelength

d = internal diameter of waveguide

k = Bessel root for TE_{01} mode = 3.832

The loss due to absorption of oxygen which is present in the atmosphere (at approximately sea level) was obtained from the experimental data of D. C. Hogg.[9] The added loss produced by the presence

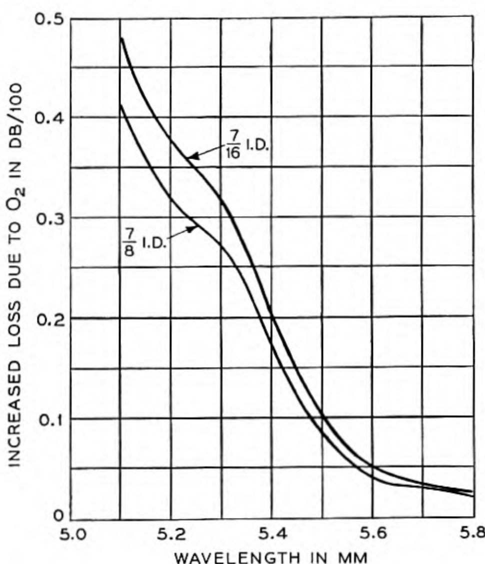


Fig. 6 — TE_{01} transmission loss in waveguides due to oxygen absorption.

[9] A. B. Crawford and D. C. Hogg, Measurement of Atmospheric Attenuation at Millimeter Wavelengths, B.S.T.J., **35**, pp. 907–917, July, 1956.

of oxygen in the waveguide in terms of (1) is plotted in Fig. 6. It will be noted that this loss becomes very appreciable at the short wavelength end of the band. At $\lambda = 5.2$ mm this loss is in the 0.3–0.4 db/100 ft range. For the larger size waveguide line (7/8″ I.D.) the loss due to O_2 is approximately equal to the theoretical wall losses; for the smaller size lines this amounts to about a tenth the wall loss. At the other end of the millimeter band the O_2 losses are very small, being in the 0.02 – 0.03 db/100 ft range at $\lambda = 5.7$ mm.

The relative effects of theoretical wall and expected oxygen absorption losses are shown plotted in Fig. 7. For the two sizes of waveguide the upper dashed curve represents the combined effect of these two factors and the lower solid line curve is the theoretical attenuation of the TE_{01} mode in empty pipe. The shaded area indicates the increase which is the result of oxygen absorption.

In order to minimize the transmission losses in any practical system it becomes desirable to exclude the presence of oxygen from the line, for example, by introducing an atmosphere of dry nitrogen. Since the ex-

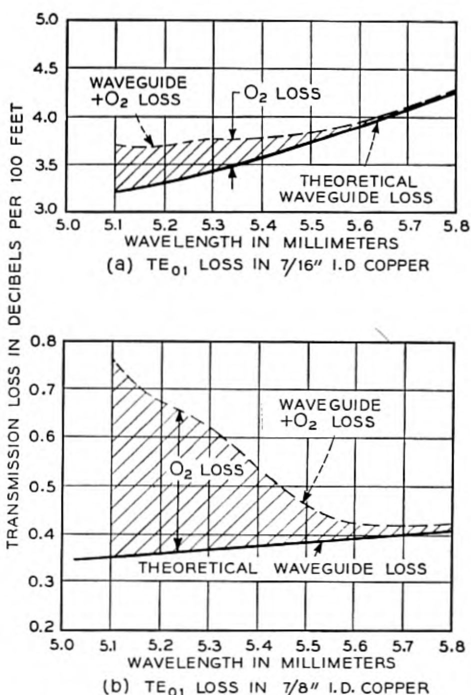


Fig. 7 — TE_{01} transmission losses.

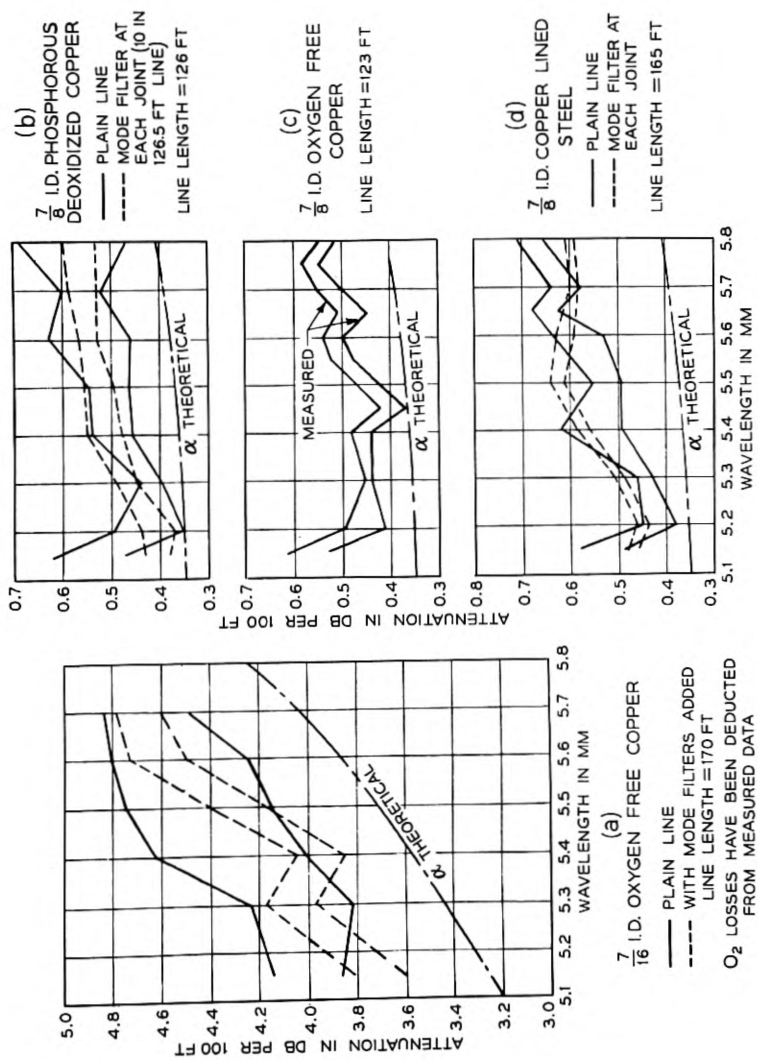


Fig. 8 — Measured and theoretical loss of the four lines studied.

clusion of oxygen was not very feasible in the experimental TE_{01} lines, the effects due to oxygen absorption were included in the measurements. However, in order to simplify the presentation of the attenuation data these absorption losses, as indicated in Figs. 6 and 7, have been subtracted from the measured data.

The measured attenuation of the four lines are shown in Fig. 8 as a function of wavelength (5.1–5.8 mm). In each case the dash-dot-dash lines represent the theoretical attenuation for copper. Each plot shows two solid lines which indicate the range of values measured over the mm band. The same range was observed either by varying the length of the line by means of a sliding piston at the far end of the line or by imposing a sweep voltage on the repeller of the signal klystron to produce a small frequency modulation. These variations in attenuation correspond to piston movements which are greater than a half wavelength and are due to the mode interference effects produced by spurious modes generated in the line. The resultant signal fluctuations which are due to mode conversion and reconversion effects have been described in considerable detail by Miller.[10]

Referring again to Fig. 8, the measured data shown by the solid lines, which are for a plain line without mode filters, indicates that the oxygen-free high conductivity copper line gave the lowest measured average attenuation as well as the least variation. The low phosphorous deox-

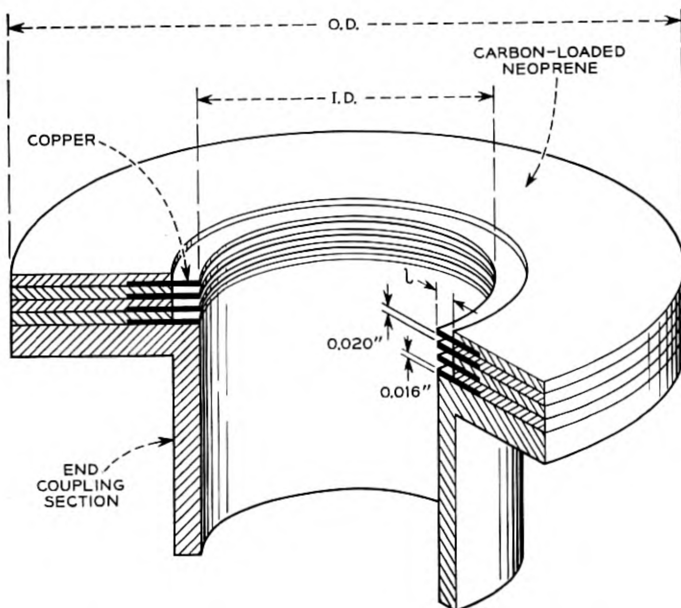
TABLE 1

	7/16″ I.D.	7/8″ I.D.	7/8″ I.D.	7/8″ I.D.
	OFHC Copper	OFHC Copper	Low Phos. Deoxidized Copper	Copper Lined Steel
Wall Thickness	1/16″	1/8″	1/8″	1/8″
α meas. (db/100 ft)	4.33 ± 0.24	0.47 ± 0.02	0.49 ± 0.05	0.52 ± 0.04
$\frac{\alpha \text{ meas}}{\alpha \text{ calc}}$	1.17	1.29	1.34	1.42
Average ovality				
A	1/1100	1/1100	1/1200	1/585
B	0.0004″	0.0008″	0.00075″	0.0015″
Maximum ovality				
A	1/730	1/875	1/875	1/290
B	0.0006″	0.001″	0.001″	0.003″
Maximum tolerance				
A	1/310	1/730	1/430	1/290
B	0.0014″	0.0012″	0.002″	0.003″

[10] S. E. Miller, Waveguide as a Communication Medium, B.S.T.J., **33,** pp. 1229–1247, Nov. 1954.

idized copper was next best while the steel line with a 20-mil inner copper lining was the poorest.

In the 7/16″ I.D. oxygen-free high conductivity copper line the measured attenuation was 17 per cent higher than the calculated value (see α meas/α calc in Table I). This higher loss is attributed to spurious mode conversion and to surface conductivity effects. In the 7/8″ line of the same material the α meas/α calc = 1.29 which is an increase of 12 per cent relative to the smaller waveguide. Since the 7/8″ diameter line supports about four times the number of modes of the 7/16″ diameter line, this increase in loss is attributed to mode conversion. In the other two 7/8″ diameter guides the added losses are believed to be increased mode conversion which results from the poorer dimensional tolerances. These data are listed in Table I together with dimensional tolerances. In this table α_{meas} is the measured attenuation averaged over the 5.2–5.7 mm band together with the variations shown in Fig. 8; α_{calc} is the average theoretical attenuation for standard (IACS) copper. The I.D. tolerances are listed in two sets of rows A and B; row A gives the fractional variation



TYPE	O. D.	I. D.	l
MEDIUM	2″	7/8″	0.080″
SMALL	1½″	7/16″	0.081″

Fig. 9 — Structure of spaced-disk mode filter.

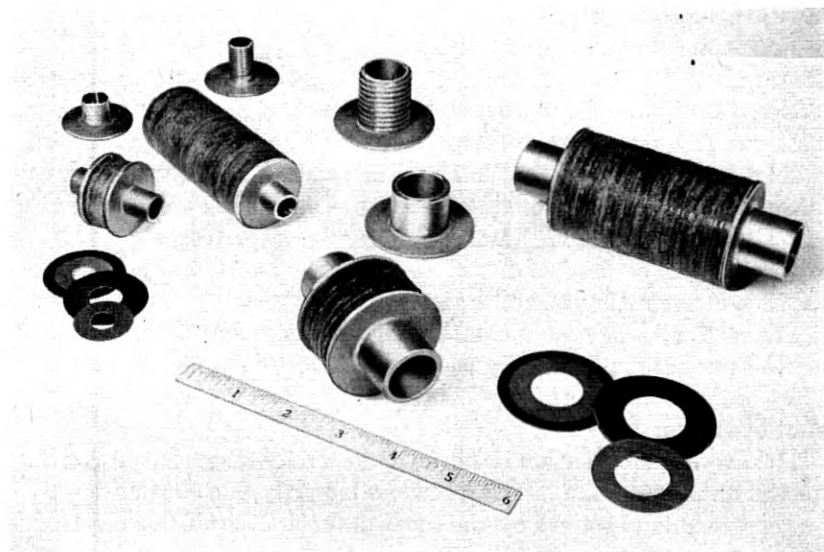


Fig. 10 — Mode filters.

relative to the average diameter and the rows marked B indicate the corresponding variations in inches. The average ovality gives the average difference between maximum and minimum diameters, maximum ovality the maximum difference in diameter and the maximum tolerance gives the maximum difference between diameter and ovality. These measurements have been limited to measuring at the two ends of each section of pipe. In spite of this small sampling the TE_{01} loss measurement appears to follow the I.D. tolerances quite well; the OFHC line shows both the lowest attenuation and the best tolerances.

Mode interference effects can be reduced considerably by increasing the loss to the undesired modes. This effect can be accomplished by modifying the structure so that the spurious modes are highly attenuated while the TE_{01} losses are increased only slightly. One way is to construct

TABLE II — AVERAGE PERFORMANCE OF TE_{01} WAVEGUIDES WITH MODE FILTERS

	7/16" I.D. OFHC copper	7/8" I.D. low phos. deoxidized copper	7/8" I.D. copper lined steel
α measured (average db/100 ft.)	4.24 ± 0.1	0.51 ± 0.025	0.56 ± 0.012
$\frac{\alpha \text{ measured}}{\alpha \text{ calculated}}$	1.16	1.39	1.52

the waveguide wall with a series of disks which are closely spaced as shown in Fig. 9 and the photograph of Fig. 10. The spacers serve a dual purpose; to hold the disks in alignment and to provide loss for the spurious modes. The circular disks provide the necessary continuity to support the TE_{01} and TE_{02} modes and the gaps introduce high resistivity to the longitudinal currents of the other modes. The spaced-disk filters, which were arbitrarily designed to provide a 10 db loss to the TM_{11} wave, were 1 ⅝″ and 3 ¼″ long for the $\frac{7}{16}$″ and ⅞″ waveguide sizes, respectively. In the experiments to be described, a mode filter was inserted at each joint of the line, at approximately 12-ft. intervals.

The measured attenuation data with mode filters at each joint of the various lines are indicated by the dashed lines of Fig. 8. As shown the effect of the mode filters is to reduce the TE_{01} loss variation by a factor of at least two.

The average attenuation is, however, generally somewhat higher than for the unfiltered lines. This higher loss is partly due to spurious mode power which is absorbed by the mode filter and is not reconverted to TE_{01} power and to a slight degree to the increased TE_{01} loss introduced by the mode filters. These results are shown in tabular form in Table II, where the nomenclature is the same as in Table I. Because of the excellent performance of the ⅞″ I.D. line (OFHC copper) by itself no measurements with mode filters were performed on this line.

CONCLUSIONS

The measured data presented above indicate the feasibility of realizing transmission losses as low as 0.5 db/100 ft. with the TE_{01} mode over distances up to several hundred feet. The transmission variations which occur over the frequency band are a function of the circularity or tolerances of the waveguide. In a particular line the variations can be reduced considerably by adding mode filters along the line. It is reasonable to expect that these variations can be reduced further by adding longer mode filters at the joints or adding more mode filters at shorter intervals along the line. Oxygen must be excluded from the line if the losses are to be a minimum.

ACKNOWLEDGMENT

The author wishes to thank J. W. Bell and W. E. Whitacre for their help in the measurements.

This study was carried out at Holmdel and was sponsored in part by a Joint Service Contract administered by the Office of Naval Research, Contract Nonr-687(00).

Automatic Testing in Telephone Manufacture

By D. T. ROBB

(Manuscript received May 8, 1956)

A general discussion is given on the philosophy behind the development of automatic test facilities and the relationship of this activity to product design and manufacturing engineering. A brief historical discussion of early automatic test machines used by the Western Electric Company leads to a summary of design considerations. These considerations are then illustrated by descriptions of the specific techniques used in three automatic facilities of considerable diversity.

INTRODUCTION

Many of the parts used in the telephone plant are made in such numbers that automatic shop testing of them is desirable. The cost of manual testing by suitable personnel is high, and its nature so repetitive and dull that accuracy suffers. Fortunately, in many cases the complexity of the test requirements has matched the state of the art and the business picture well enough to warrant the development of machine methods. It is our purpose in these articles to review the art as it has evolved in the Manufacturing Division of the Western Electric Company, and to describe some of the techniques. This is done with the hope that improvements or extensions to other testing or manufacturing problems may be suggested.

It should be emphasized that the developments treated here and in the other papers[3-9] have required cooperation among testing and manufacturing engineers in the Western and product design engineers in the Bell Telephone Laboratories. Modifications of design for Western's convenience, changed methods for translating basic requirements into manufacturing test requirements, informal Laboratories suggestions of approaches to manufacturing and testing problems, all are commonplace. The boundaries of the specialists' domains are readily crossed.

Testing is a process for proving something such as quality of a prod-

uct or accuracy of a computation. In one form or another, testing is essential in manufacture. It insures against further investment of effort in product found bad. More importantly, it provides information for the manual or automatic correction of earlier processes, to prevent manufacture of additional faulty product. Also, its techniques and devices are used in many applications where testing is not the object. Table I gives a listing of functions, with examples of some of our automatic means, that illustrates this. Of these, 1a, 4, and 5a are testing functions. The remainder are manufacturing processes.

TABLE I

Function	*Example*
1. Sorting, either a. sorting good from bad or	A network testing machine at Indianapolis.[1] A relay coil test set at Kearny.[2]
b. sorting into cells for selective assembly	A capacitor test machine at Hawthorne.[3]
2. Adjusting:	An adjusting machine for flat type resistors at Haverhill.[4]
3. Calibrating:	A calibrating machine for oscillator film scales at Kearny.[5]
4. Plotting data:	Continuous thickness test systems for alpeth and stalpeth cable sheath at Hawthorne and Kearny.[6, 7]
5. Operation of wired equipment, a. to verify accuracy of wiring or fulfillment of purpose, and b. to enable prompt location and correction of faults.	Cardomatic and tape-o-matic test sets for key telephone equipments and wired relay units at Hawthorne and Kearny.[8, 9]

GENERAL

The fundamental steps necessary to any testing operation are:

1. Putting the item to be tested in location;
2. Subjecting the item to a specified set of conditions;
3. Observing the results or the reaction of the item to the conditions;
4. Comparing the observed results to required results;
5. Deciding on the basis of the comparison what disposition to make of the item;
6. Indicating the disposition;
7. Making the disposition. (This may mean transportation, repair or adjustment.)

In purely manual testing all of these steps would be initiated by human

operators. In many cases it is feasible for all steps to be taken automatically. The bulk of our accomplishment in automatic testing, however, has been in steps 2 through 6. We do not ordinarily use "automatic" to describe rudimentary automaticity in combinations among steps 3, 4, and 5.

The present models of many of our machines have evolved from earlier models, either because of changed product or test requirements or through improved designs worked out for plant expansion or cost reduction. The names of engineers associated with the various developments mentioned are included in the references. About 1927 there were put in use at Hawthorne two machines, one for gaging a number of critical dimensions and performing a breakdown test on carbon protector blocks,[10] and the other for heat coils.[11] In the protector block machine the blocks follow a linear course drawn by an indexing chain conveyor through a number of positions where the various checks are performed. Failure of any block at a position causes a jet of air to blow the block into the opening of a chute which conducts it to a reject pan. Good blocks are delivered into a pan at the end of the run. The heat coil machine has an indexing turret over a ring of ports which open selectively to permit good or rejected coils to fall into chutes. The test parameters are three gaged dimensions and dc resistance.

In 1929 a machine with an indexing turret was put in use, testing paper capacitors for dielectric strength and leakage resistance,[12] and sorting them into 13 cells for capacitance grouped around a nominal 1 mf. The 13 cells correspond to 13 segments in a commutator disposed along the scale of a microfarad meter. For a given test capacitor, when the meter needle reaches its deflection a bow depresses it against the nearest segment, establishing a circuit through a relay. A system of relays then locks up and serves as a memory to operate a solenoid later when the turret has brought the capacitor to the point of disposition. Action of the proper solenoid causes the capacitor to be deposited in its cell. The cells are arranged as parallel files in a horizontal plane and, starting with the cells empty, the machine will in effect produce a stovepipe distribution curve. Capacitors from the middle cell and its upper neighbors may be used as 1 mf capacitors, and those from more remote cells combined, large with small, to make 2-mf capacitors.

Also in 1929 a turret type machine was first used for sorting mica laminations.[12] The sorting parameter was ac dielectric strength, the criterion being failure at 1760 volts r.m.s. The individual laminations were carried from position to position by vacuum fingers mounted on a turret. Again locking relays were used, in this case to operate a solenoid controlled

valve in the vacuum line at the right time in the turret indexing cycle to drop the laminations as class "A" or "B" mica.

Experience with these machines and with others that followed brought into being a more or less orderly body of knowledge as to what features are desirable and what constitutes good design in an automatic test machine.

If the machine is to have speed, reliability and long life, attention should be paid to the following matters:

1. *Reduction of the test process time to as low a figure as the capabilities and use of the product will permit.* Thus, if one of the requirements of a capacitor is a maximum limit on its leakage current measured after a charge time of 60 seconds, and if the materials and manufacturing process are such that a unit is surely good or bad after a 25-second charge, then the machine may be designed to charge for, say, 30 seconds. Frequently the only limitation is the speed of the machine itself. When this is true, it must be worked out so as to satisfy the needed production rate. Obviously the machine should satisfy the rate of the line it serves, or more than one machine should be provided.

2. *Rationalization of the number of test positions in the machine with the production rate and the total test process time.* This requires breaking the test time down into bits equal to the desired output cycle. In the example above, if the output needed is a capacitor every 5 seconds then the 30-second charge will have to extend over 6 positions.

3. *Ruggedness.* This must be stressed, even at the expense of space, power consumption, and dollars of first cost. If a project is large enough to justify automatic test facilities, then any down time associated with it will be expensive. A good mechanical design is essential.

4. *Provision of self-stopping and alarm features to serve in the event of certain types of failure.* A limited torque clutch in the main drive will prevent jamming and damage caused by parts getting into the wrong places, or in certain applications overload cutouts will suffice. Gong and lamp alarms are desirable to attract attention. The point is that allowance must be made for mishaps which, without precautions, could result in shutdowns of the equipment.

5. *Provisions of adequate checking for accuracy.* Accessible check points and suitable easy-to-use standards are essential. Checking intervals are determined by experience, but schedules should be laid out to cause as little interference with use as possible. Where practicable there may be means for self-checking in the regular operation of the machine. In this case, periodic checking of the checking devices themselves is necessary.

6. *Incorporation of features in the product and in the handling methods*

that will facilitate feed automatic testing. This requires the cooperation of the product design and product manufacturing interests. It is almost axiomatic that automation in manufacture requires special consideration in product design. Automatic testing imposes the same requirement. A notch or a lug may be needed for proper use of automatic feed devices, or terminals may have to be properly chosen. Again, the method of transport from the previous operation needs to be studied, rationalized, and fully agreed upon. If continuous conveyor transportation can be justified, so much the better. In the consideration of conveyor feed, the need for time flexibility must not be overlooked. It is important that provision be made for easy storage of product whenever the test machine is inoperative, lest a breakdown of this machine shut down the entire line.

7. *Arrangement of the events in the operating cycle in such a way that their sequence is reliably self determined.* This is comparatively straightforward when the programming is done by gear driven cams or other mechanical means. It requires care when switching logic is used. Switching engineers are familiar with the phenomena known as "relay races" and "sneak circuits." These have psychophysical analogies wherever humans and machines work together. The prevention of both the switching errors and their analogs is essential in automatic test set design. Interlocks must be provided against any conceivable mishap.

8. *Enough margin and design flexibility in electrical and mechanical parameters to cope with reasonable variations in product design.* Improvements are constantly being made in telephone apparatus and equipment, and these occasionally result in major redesigns or in entirely new systems. Also the need for adding new features to a historical complex of existing telephone plant causes the generation of an endless variety of special equipments. The product designer needs as much freedom as we can afford. There has to be enough flexibility in the costly automatic test sets to permit adaptation as new designs of product come along.

These considerations are in addition to the fundamental matters of personnel safety and comfort, motion economy, quietness and appearance.

While dealing with general considerations we must recognize one important difference between the product design and the facilities design problems. In product design there is a premium on optimization of parameters, or striving toward perfection. There is generally also opportunity for winning this premium on later tries even though the rush for first production may have denied it to us in the original design. In facilities design there is no such premium and frequently no such opportunity. While careful design is very important, the real premium here is on a

Fig. 1 — Network test position.

device that will do the required job and that can be put in use in time for early production. Once the facility is in use it may be starting on a productive life that will run thirty years or longer. The designer may think of countless ways to improve it or to redesign it completely. If his improvements or redesign can be proved in on a business basis, they may be undertaken. Sometimes they cannot be proved in. The evolution that has taken place in test set designs has been possible mainly because the customers have wanted newer products, or products delivered at a greater rate. Advancement has been attained under a compulsion to take each step quickly and surely. This has represented a real and continuing challenge to the test engineering force.

With these general considerations in mind the author has chosen three automatic testing devices of diverse character to discuss in some detail. The associated papers[3, 9] cover additional machines. The machines described illustrate in various ways the principles discussed above.

THE NETWORK TESTING MACHINE AT INDIANAPOLIS[1]

The 425B network[13] is used in the 500 series telephone sets to furnish the transmission link between the handset and the line. Its shop testing requires three tests for transmission, three for capacitance tolerance, three for leakage current, two for ac dielectric strength, one for dc dielectric strength and four for continuity. The rotating turret type test machine (Figs. 1 and 2) performs all these tests, applies a conditioning "burnout" voltage and counts and date stamps the good networks. Rejects from each test position are segregated in roller conveyors. In the rotation of the turret an empty test fixture is presented to the operator every 3¾ seconds moving from left to right. She must load each position, taking networks from the pans at her right; good networks, ejected automatically in a roller chute at the left, are hand loaded into the carriage fixtures of the overhead storage type conveyor, which pass within easy reach of the operator's left hand. The pans at the left are used to store good networks when the accessible fixtures of the overhead conveyor are full. The twelve roller conveyors for rejected networks are arranged along the sides of the machine, six on each side.

The turret contains forty test fixtures (Fig. 3 and 4) and the machine forty positions. The turret rotates continuously, causing eleven contact brushes associated with each fixture to pass against fixed commutator segments and a ground ring associated with the test positions. As each fixture advances past one test position a gear connected cam shaft rotates through a complete cycle. Seventeen switches are operated by the cams

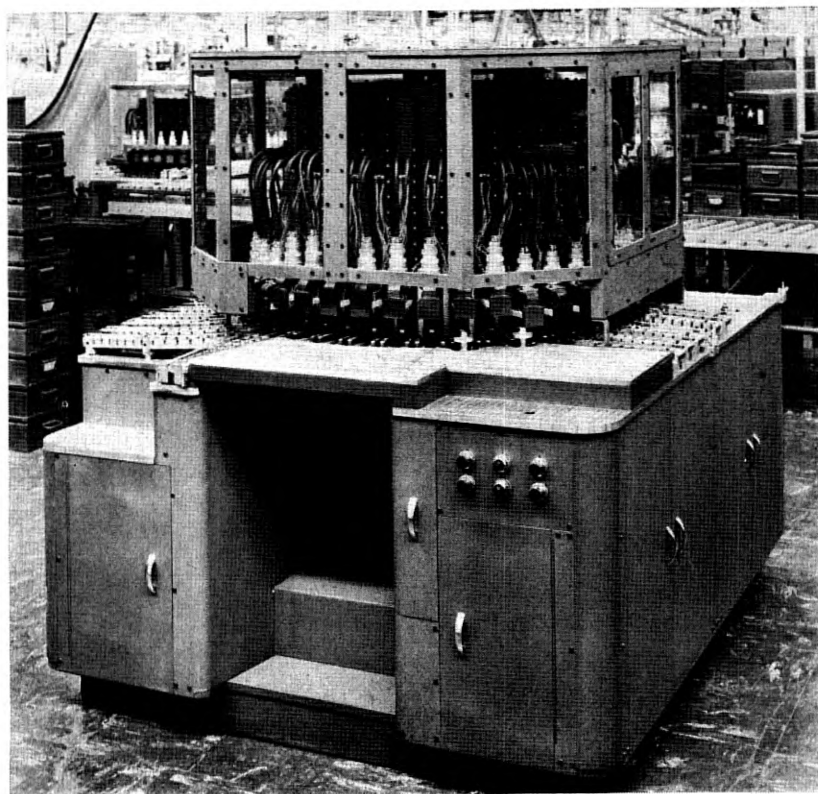


Fig. 2 — The network test machine.

to assure the proper sequence and timing of the conditioning and testing events occurring at the various positions. Table II shows the order of the positions and the approximate timing, with respect to cam rotation. The result of the test at each test position is remembered by a self-locking relay until the fixture comes just opposite the entrance to the corresponding rejection chute. At that instant a cam switch closes and causes rejection if the test result was a failure. Unloading into the rejection chutes is effected by compressed air operated cylinders as explained below.

The clamping movement of each fixture as it leaves the loading area (entering position 7) is driven by a helical spring which lowers the contact fixture over the terminals of the network, bringing spring loaded plungers into contact with the terminals. (See Fig. 3) At a rejection location a plunger rises, driven by an air cylinder under the control of a solenoid operated valve. The rising of the plunger first forces the fixture to

unclamp against the compression of the helical spring, and then operates an ejection arm which drives the network horizontally out of the fixture. The top rollers of several of these ejection arms can be seen in the fixtures at the front of the machine in Fig. 2.

The measuring circuits associated with the various test positions are straightforward. If there is a dielectric failure in one of the breakdown tests at position 8 or 9, the current through a relay coil in series with the

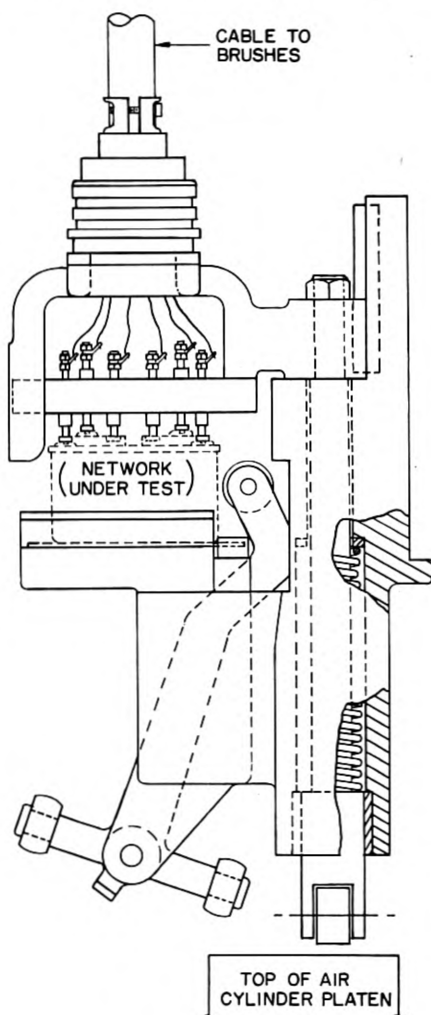


Fig. 3 — Test fixture, loaded.

test exceeds a predetermined value. This causes another relay to lock up and remember the failure until the network reaches the reject location.

In a typical transmission test position (Position 10, 35 or 36) a fixed-voltage, swept-frequency signal, 300 to 3,500 c.p.s., is impressed across two terminals of the network. The three tests are for transmission and short and long line sidetone with suitable terminations connected as in actual use. In each case the signal from two output terminals should be less than or greater than a specified value. This signal is amplified and fed

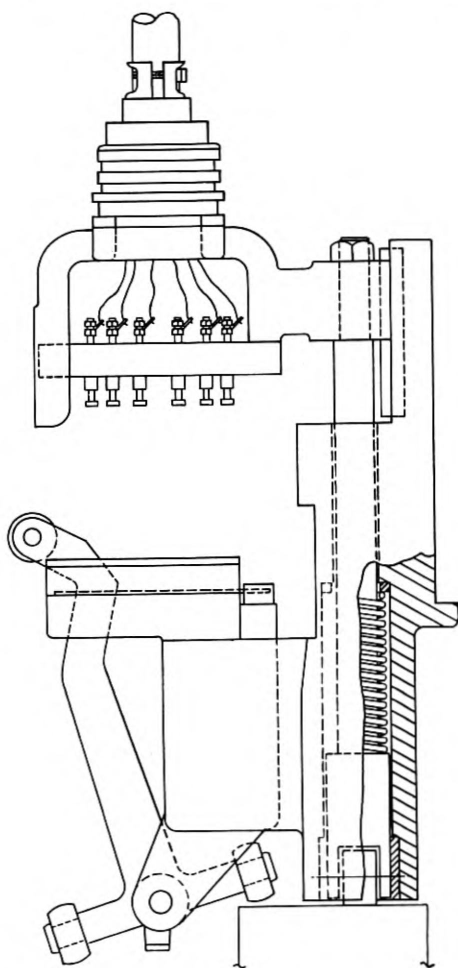


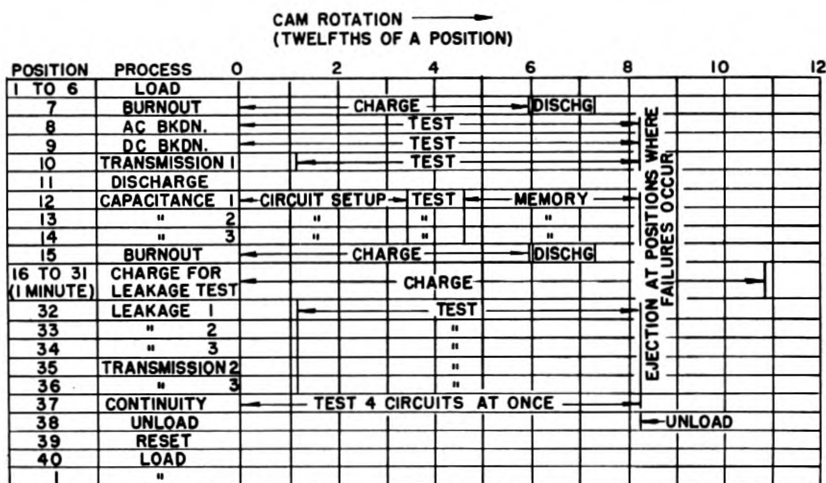
Fig. 4 — Test fixture, unloaded.

to a sensitrol relay, which is mechanically biased in amount and sense to correspond to the limit. If the sensitrol operates it prevents rejection.

In the three capacitance test positions 12, 13, and 14, capacitors in the networks are connected into a 60 c.p.s. comparison bridge. The output signal from the bridge is amplified and rectified, and impressed on a balanced dc amplifier which drives a sensitrol relay. If the bridge is out of balance (that is if the capacitance is greater or less than nominal) current flows in the relay, but always in the same sense. If the current in the relay exceeds an amount corresponding to either capacitance limit, rejection occurs. Determination of which capacitance limit was violated is done manually in a separate analysis of defects. It may be observed also that any rejection at the capacitance positions could have been caused by a loss unbalance of the bridge. If the conductance of the test capacitor were such as to cause this it would so appear in the separate analysis, mentioned above. The effect of any ordinary conductance deviation at 60 c.p.s. is negligible. Quality is protected by the fact that a conductance deviation could not cause an out-of-limit capacitor to be accepted.

Considerable pains are taken at each capacitance test position to prevent damage to the equipment from various kinds of mishaps. The sensitive winding of the sensitrol is short circuited at all times except for about 0.2 second when the actual test is performed. This prevents damage and erroneous rejections that would otherwise be caused by switching

Table II — Sequence of Events in Network Test Machine



transients from this and other circuits. During the short interval of actual test no other switching takes place in the machine.

A fixture that has no network because of rejection at an earlier test position or because of operator failure to load it, would cause open circuit in one bridge arm on capacitance test. Without intervention this would cause a violent unbalancing of the bridge, overloading of the detector system and possible damage to the sensitrol. Ordinary methods of limiting the overload signal would be only partially effective and would detract from the sensitivity. To forestall this trouble from empty fixtures, each capacitance test position is equipped with a microswitch which is operated by a dog at the bottom end of the ejection arm of any empy fixture (Fig. 3). When the microswitch operates it causes the bridge to be disconnected from the test leads and connected to a capacitor that is just out of limits, several tenths of a second before the removal of the short circuit from the sensitrol. Then when the test is made it results in a rejection.

There is also an interlock circuit which will stop the machine if a failure of the bridge and detector system causes an empty fixture not to show rejection. This serves as a random occasional check on the functioning of the circuit.

The conditioning of the three capacitors for the leakage current tests begins at position 16. Because of charging and absorption currents obscuring the effect of pure leakage, the test for leakage is made to an arbitrary current limit specified at one minute of charge. To insure that good units pass the test, it is desirable to use the whole minute. But if the leakage current reading is taken after more than a minute of charge, quality is jeopardized. Accordingly it is necessary to make sure that the charge is for a minute and no longer on each capacitor. Therefore, at position 16 the first unit is put on charge, at 17 the second, and at 18 the third. Then at position 32 the first unit is tested while the other two remain on charge. At 33 the first unit is discharged, the second tested, and so on.

The leakage test itself is made by measuring the voltage across a large resistor in series with the test capacitor and a dc voltage source. The energy in this signal is small and must be amplified before there is enough to operate a sensitrol. A dc amplifier with high input impedance is used for this purpose. In addition the mechanical bias of the sensitrol is kept small to increase sensitivity, and a carefully controlled dc biasing source is used to insure accuracy and stability.

At position 37 three capacitors and a coil winding are given a final check for continuity. The test of the winding is made by connecting it in series with a relay coil (say No. 1) and battery. If current passes, relay

No. 1 operates. The three capacitors are tested simultaneously by connecting each of them in series with an 8,000 c.p.s. source and detectors. The detectors consist of bridge type rectifiers and relays. If all of these three relays operate, a series connection through their closed contacts causes another relay to operate and lock up. Finally this relay when operated has open contacts in parallel with open contacts on relay No. 1, so that when the reject cam closes it finds an open circuit and rejection does not occur.

The reader may question the necessity for continuity tests on capacitors that have already been tested for capacitance. Perhaps the most convincing answer is that there is an occasional failure on the continuity test. Telephone apparatus is always exposed to more severe conditions in test than it will encounter in ordinary use. The leakage resistance charge and test operations and the transmission tests can on rare occasions cause the metallized connections at the ends of the capacitors to open. As the cost of making the final continuity test is vanishingly small, the additional insurance is economical.

The detail list of checking standards for this machine contains some twenty items. Most of them are modified 425B networks, specially arranged in one way or another to check certain functions of the machine. These are used right in the individual fixtures.

It is interesting to reflect on the labor saving virtues of this machine. The operator in one eight hour shift handles over five tons of networks. She does it easily and without fatigue. The testing would not be even attempted on a manual basis, because over and above multiple handlings, the added human effort of closing fixtures, operating switches and the like could not be tolerated.

In contrast to the multiposition set described above, it is instructive to consider two single position sets of diverse character. They are a relay coil test set and a film scale calibrating set.

THE RELAY COIL TEST SET AT KEARNY[2]

Coil assemblies for the U, Y and UA types of relays[14] are tested for dc resistance, direction of winding and breakdown before assembly into complete relays. Many thousands of the relays are used in any crossbar office. Minimum and maximum tolerance limits are placed on their winding resistances, to control cumulative current requirements and to insure a proper margin of relay operation. Each coil assembly, as presented to the test position, consists of a magnetic core, a solenoidal winding assembly and a terminal assembly. A winding assembly may have one, two or

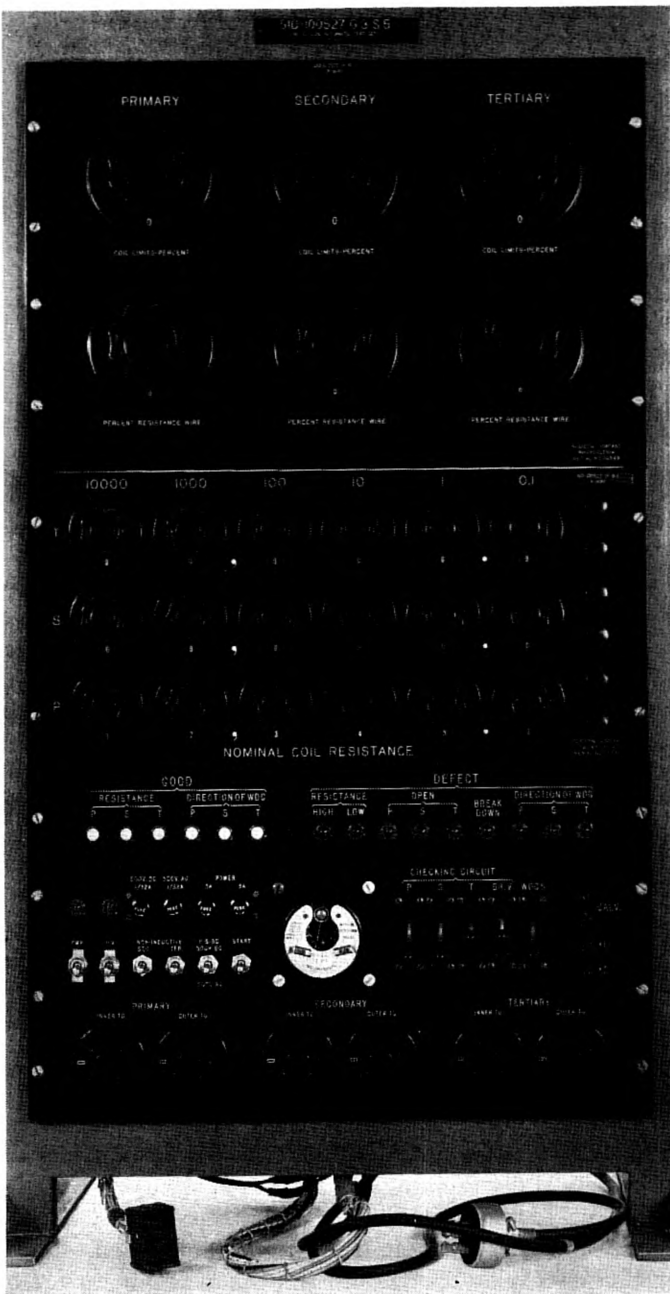


Fig. 5 — Relay coil test set control panel.

three windings (called primary, secondary and tertiary). The primary and secondary are wired to corresponding pairs of terminals on the terminal assembly, while the tertiary leads at this stage are not on terminals and must be connected to the test contact fixture by hand.

Direction of winding is important in the multiwinding coils because of external fields and the fact that the relays are required to respond to currents in more than one winding and the proper direction of flow in each, relative to the other, must be known. In some relays one or two of the windings may be noninductively wound, to serve merely as resistors. Also, many windings are wound part copper and part resistance wire to obtain the desired resistance without unnecessary increase in copper, inductance and response time. In such cases the percentages of copper and resistance wire are known. This is important because of the effect of temperature on the resistivity of copper. Resistance tolerances on the test windings are specified at 68°F, but shop testing is done at any value of room temperature. The effect of the difference on copper is serious enough to cause errors larger than some of the tolerances, and the effect on resistance wire may be neglected. Therefore, it is necessary to have the test set compensated for temperature in such a way as to allow for the proportions of copper and reistance wire.

The coil test set (Fig. 5) tests all windings for resistance and direction of winding and for breakdown to each other and the core. The maximum total test time for three-winding coils is less than 3 seconds under normal conditions. A borderline winding resistance will cause some delay. There are lamps to indicate the type of failure on a rejection. Other lamps indicate satisfaction of the requirements. At the completion of test on a good coil an "OK" lamp lights on the test fixture, so that the operator need look at the set itself only when there is a rejection.

Requirements data are stored in the set before a given code of coil is tested. The codes come to the set in batches, so that one setup will serve for a large number of coils. Three six-decade resistance standards are set to the nominal values for the respective windings. If there are fewer than three windings, a key is operated to disable bridges and furnish substitute continuity paths. The percentage tolerances for the windings are set on selector switches: ±1, 2, 5, 10 and 15 per cent tolerances are available. Also, the known percentages of resistance wire in the windings are set on selector switches in steps of 5 per cent from 0 to 100. Keys are operated to warn the set of noninductive windings and bypass the direction of winding circuits as needed.

Once a coil is placed and connected in the test fixture and the fixture closed by operation of a pedal, the test is automatic up to the point where

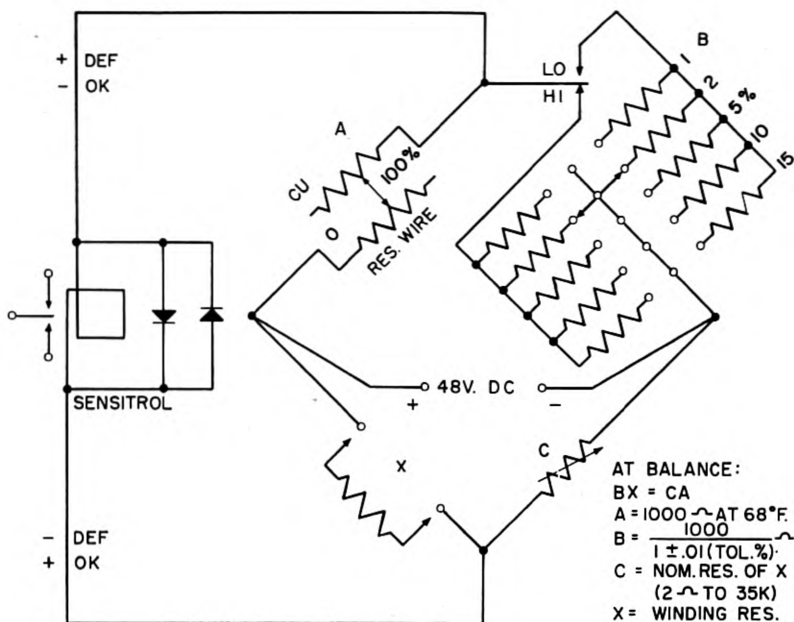
the operator must make disposition. The sequence of events within the set is controlled by a switching circuit containing thirty telephone relays, a sensitrol relay and two electron tubes. The sensitrol is used in succession to detect the existence and sense of unbalance of six dc bridge circuits (high and low limit for each of three windings). The operation sequence for primary windings is shown in Table III.

Fig. 6(a), shows schematically a typical bridge arrangement for testing a winding at one tolerance limit. A and B correspond to the ratio arms of an ordinary Wheatstone bridge, and are nominally 1,000 ohms each. The temperature compensation referred to above is obtained by including the same resistance percentage (within 2.5 per cent) of copper in the A arm of the bridge as there is known to be in the winding. Inspection of the bridge balance equation in Fig. 6(a) will show that an error in X could be compensated by a proportional error in either A or C. A is chosen as the compensating arm because of its simplicity. It has available twenty resistors of copper and twenty of low temperature coefficient resistance wire. Each resistor is 50 ohms, measured at 68°F. The selector switch is arranged so that the arm always has twenty resistors, the indicated percentage being resistance wire.

For proper compensation it is necessary that the A arm be as near ambient temperature and the temperature of the coils as possible. The di-

Table III — Sequence of Events in Test of Primary Winding for High Limit Resistance

"OK" LAMP	STEP	DEFECT LAMP
	POWER SWITCH CLOSED	
	SENSITROL RESETS AND HOLDS	
	OPERATOR CLOSES FIXTURE	
	FIXTURE START SWITCH CLOSES	
	CONTINUITY TEST – ALL WINDINGS	"P OPEN" ETC.
	"HIGH" B ARM CONNECTED TO PRI. BRIDGE	
	SENSITROL RESET RELEASED	
	SENSITROL OPERATES	"HIGH"
	"LOW" B ARM CONNECTED TO PRI. BRIDGE	
	BREAKDOWN TEST ON PRIMARY	"BREAKDOWN"
	SENSITROL RESETS AND HOLDS	
	SENSITROL RESET RELEASED	
"P RES. GOOD"	SENSITROL OPERATES	"LOW"
	DIRECTION OF WINDING DETECTOR ENABLED	
	D.C. POWER DISCONNECTED FROM PRIMARY BRIDGE	
	INDUCED VOLTAGE IN PICKUP COIL	"P DIR. OF WDG. DEFECT"
("OK")	SENSITROL RESETS AND HOLDS	
	"HIGH" B ARM CONNECTED TO SEC. BRIDGE	
	(SECONDARY TEST PROCEEDS; SIMILAR TO PRIMARY)	
	(FOR SINGLE-WINDING COIL, LAMP ON FIXTURE LIGHTS)	



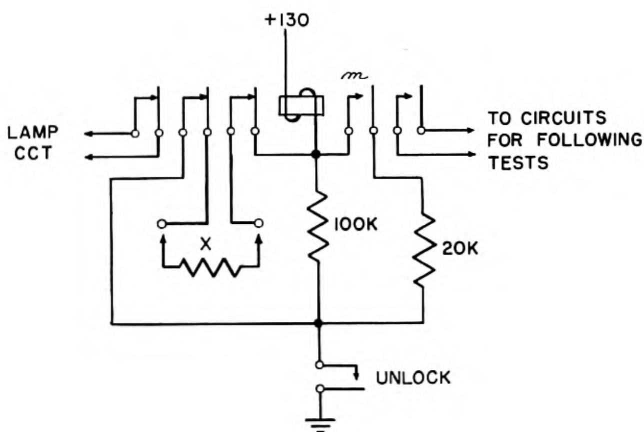


Fig. 6 — Circuits used in Relay Coil Test Set. (a), resistance bridge, simplified schematic; (b) continuity, simplified schematic.

vision into twenty resistors helps in this by maintaining high effectiveness of dissipation. In addition, the automatic switching circuits are arranged to keep the duty cycle of current in the bridge arms low.

The B arm of the bridge is selected by the setting of the percentage tolerance switch. Each resistor is used alone and consists of low temperature coefficient resistance wire as in standard bridge practise. The value of each resistor in ohms is 1,000 divided by one plus or minus the corresponding tolerance fraction. Thus, for ± 1 per cent tolerances the resistors are 1,000/1.01 (=990.0) and 1,000/0.99 (=1010.1), respectively. One setting of the switch indicates zero tolerance and is equipped with 1,000-ohm resistors to permit easy checking of the C arm precision.

The six-decade standard resistor in the C arm, which is set to the nominal value of the test winding, is of a high quality commercial type with a range of 0 to 40,000 ohms in steps of 0.1 ohm. Because the C and X arms may contain values as low as 2 ohms, no relay contacts are used in them. Relay switching is done in the A and B arms where the resistances are always of the order of 1,000 ohms and small variations in contact resistance are negligible. The more stable wiping contacts of selector switches do appear in the X arm. These switches permit any contact in the test fixture to be connected to any bridge terminal, to enhance flexibility.

A continuity test on all windings, before resistance test, is desirable for two reasons. The effect on the sensitrol of the severe bridge unbalance caused by an open winding would be life-shortening and is to be avoided if possible. Also, the result of the resistance test would only show high resistance, and separate analysis would be needed to reveal that a winding was open. The continuity test circuit in Fig. 6(b) was devised to prove continuity for windings having resistance values as high as 35,000 ohms. A relay (UA-104) was chosen which is sensitive enough to close a pair of "preliminary make" contacts (m) on 0.005 ampere, and which provides the number of other contacts needed to satisfy circuit requirements. When the test winding is connected at X, the currents through it and the 100,000 ohms combine to equal 0.005 ampere or more. This closes m, connecting the 20,000-ohm resistor in parallel with the 100,000 ohms, thus locking the relay and assuring that all the other contacts operate. In the act of proving continuity, the relay disconnects itself from the test winding and remains locked. The make contacts shown at the right end of the relay symbol are in series with similar contacts on the continuity relays for the other two test windings, and when all are closed they pass operating current to a relay which initiates the first resistance test (for primary high limit).

In the direction of winding circuit, Fig. 7(a), it is necessary to have a negative pulse from the pickup coil, in the test fixture, cause the 313CA

gas tube to fire and the relay to operate. The circuit is designed to handle a wide range of pulse amplitudes. The VR tube limits negative pulses to 90 volts to protect the 6AK5. The varistor dissipates positive pulses and prevents any false acceptance that might be caused by damped oscillations following a positive pulse. The 6AK5 furnishes the needed sensitivity for small pulses.

Occasionally a winding will have a value of resistance just equal to

Fig. 7 — Circuits used in relay coil test set. (a), direction of winding, simplified schematic; (b) anti-stall, simplified schematic.

its upper or lower tolerance limit. On the corresponding resistance test, the sensitrol will balance and not operate either way. Without an anti-stall device the test cycle would then be stalled until the balance failed. Current flowing through the A arm would eventually heat it up and vitiate the temperature compensation feature. The anti-stall circuit in Fig. 7(b) is essentially a slow release device to which external energy is interrupted at the same time as the sensitrol reset is released. Energy stored in the 50-mf capacitor prevents release of the relay for about 3 seconds, long after the bridge test is ordinarily finished. If at release the bridge is still balanced, a 50,000-ohm resistor is thrown in parallel with that ratio arm which will make the sensitrol accept the test winding.

A prominent and hitherto valuable feature of this test set is its adaptability to a large variety of coil assemblies. Some hundreds of distinct designs of product are presently accommodated. In the Kearny relay coil shop there are four sets of the design described here and four sets of earlier designs. It is possible that future development, if justifiable, will be directed toward greater automaticity for some of the simpler and more numerous product codes, with less emphasis on universal application.

THE CALIBRATING MACHINE FOR 56-A OSCILLATOR FILM SCALES[5]

Photographic films are used for the frequency scales of some oscillators to afford scale length and enhance readability. There have been several successive designs of film scale calibrators built and put in use at the Bell Telephone Laboratories and at Kearny. Some have been described in the literature.[16, 17, 18] One very early design is still in use on production at the Marion Shops in Jersey City. In its use, a calibrating run requires about an hour, and the possibility of frequency drift due to temperature variations makes the use of an air conditioned room essential. All of those used at Western, prior to the one described here, depended for accuracy on the film scale of a standard prototype of the oscillator to be calibrated. Using a frequency controlled servo linkage, the scale of the standard was reproduced photographically on the film of the product. Some of the prior art appears in the design of the new machine. In order to describe the principle clearly, it seems necessary to discuss some features which were previously covered, but which now are used in new ways.

The 56A is a heterodyne oscillator designed for use in the field testing of L3 installations.[15] It has a usable range of 50 kc to 10 mc. One component oscillator is fixed at or near 90 mc and the other may be varied between 80 and 90 mc by means of a tunable cavity. The calibrated portion of the 35-mm film scale geared to the cavity tuner is about 17

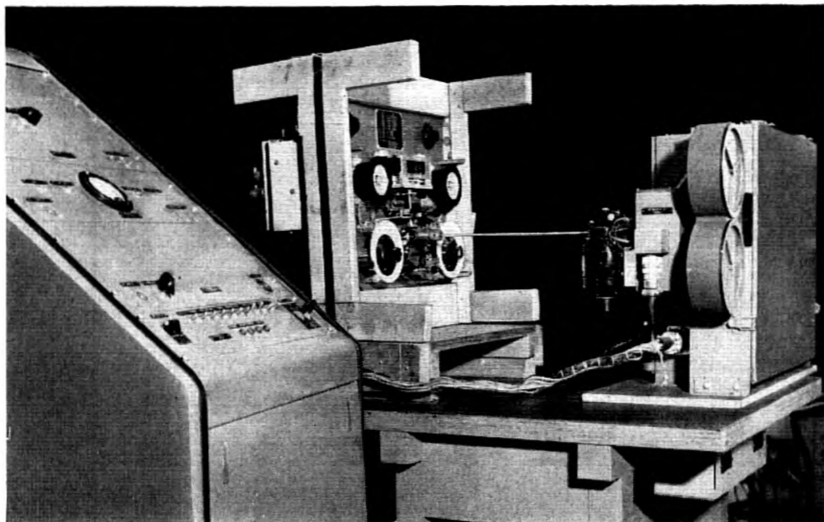


Fig. 8 — Film scale calibrator.

feet long. It has sprocket holes and is moved by a standard movie sprocket. The required precision of each calibration mark is ±2 kc. Two resonant devices are included in the circuit to permit checking and adjusting two widely separated points on the scale, 100 kc and 7,266 kc. Considering the output frequency as a function of scale setting, one of the two adjustments controls the lateral displacement of the curve and the other its average slope. By design the curve approaches linearity but not closely enough to permit less than a unique calibration for each oscillator manufactured.

Fig. 8 shows the machine which performs the calibration, with an oscillator connected, and the control cabinet. The oscillator is shown in its shipping frame. An unexposed photographic film to be calibrated is mounted in a camera so that it can be driven by a sprocket. The sprocket is connected by gears to a drive motor which also drives the take-up reel and, through a flexible shaft, the cavity tuner and sprocket in the oscillator itself. The gear arrangement is such that the peripheral speeds of the two sprockets are the same.

A positive master film is provided which has a scale similar to the one to be made for the product except that it is very precisely linear. A portion of the master is shown in Fig. 9(b). The master film passes over a sprocket which is driven by a servo motor. A lamp illuminates and shines through that portion of the master which is in front of an aperture at

any instant. An optical system, Fig. 9(a), produces on the unexposed film an image of the illuminated portion of the master. As the oscillator, its film, and the master advance, the markings on the master can be reproduced on the new film.

The problem in control is to cause each mark on the master film to pass the slit just as the oscillator goes thru the corresponding value of frequency. To do this we drive the oscillator and its scale together at a constant linear speed. The oscillator frequency increases steadily but not at a constant rate. Its rate of increase varies according to the law of its particular cavity. So our problem reduces to causing the master film to move according to that same law.

The method is to time the passage of known points in the oscillator frequency spectrum, and then to pace the movement of the master film to maintain precise correspondence. The pacing is done by detecting small differences in times of arrival at corresponding points and correcting the speed of the master film to keep successive differences small. Fig. 10 is a block schematic of the automatic control system. The varying oscillator output passes through multiples of 10 kc at a rate near five multiples per second. When it is compared in a balanced modulator with

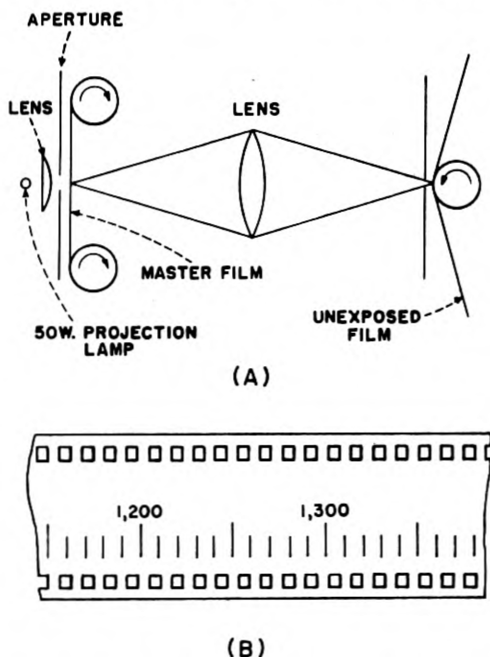


Fig. 9 — Film scale calibrator. (a), optical system schematic; (b) section of master film.

the fixed harmonics of a standard 10-kc signal, the first order difference frequency in the modulator output varies back and forth between 0 and 5 kc. It passes through the 2500 c.p.s. point twice per period of variation, or twice per 10-kc interval of the oscillator frequency.

The output of the modulator is sent through a narrow band amplifier which peaks at 2500 c.p.s. A burst of signal, therefore, leaves this amplifier twice per 10-kc interval. The bursts are further amplified and rectified and become pulses which time the progress of the oscillator through its spectrum. The pulses are impressed across the winding of a high speed relay, causing its contacts to close momentarily twice per 10-kc interval. During the instant when the contacts are closed they connect a particular value from a sawtooth voltage wave to a 0.1-mf capacitor.

The voltage of the capacitor biases the grid of a cathode follower tube, and the output voltage from this tube is fed to a servo system and controls the speed of its motor. Thus the motor runs at a speed determined by the voltage of the sawtooth at the instant when the relay contacts close. As the sawtooth itself is timed by the rotation of the servo motor, its voltage-time relationship is the device for pacing the master film. The sawtooth wave originates in the alternate shorting and charging of a 1-mf capacitor. Each tooth begins when a pair of shorting contacts is closed momentarily by a cam geared to the servo motor. After a discharge, the voltage on the 1-mf capacitor increases negatively as a practically linear function of time, with charging current flowing through a one megohm resistor. Thus the value of voltage transmitted to the 0.1-mf capacitor at the instant of closure of the relay contacts depends on the time elapsed since the most recent shorting of the 1-mf capacitor. Twenty volts at the input to the servo system corresponds to midvoltage of the sawtooth and to 3,600 rpm of the motor, which is the same as the constant speed of the motor driving the oscillator and undeveloped film.

If the characteristic of the oscillator causes a given 2,500-cycle point to occur early, the contacts of the relay will close at a higher positive voltage point on the corresponding sawtooth. The servo motor will start to speed up to make subsequent sawteeth start earlier than they otherwise would have. The motor will slow down if the 2,500-cycle points fall later and lower on the teeth.

Several design features in the system are of interest. The servo system was supplied by Industrial Control Company (SL-1035). It has a tachometer feedback in inverse sense to enhance system stability. The cam used to operate the shorting contactor and start the sawtooth is a small permanent magnet mounted on a wheel. The moving field causes the contactor to operate very briefly as the magnet swings past. The contactor itself is a Western Electric 222-A mercury switch, which has a

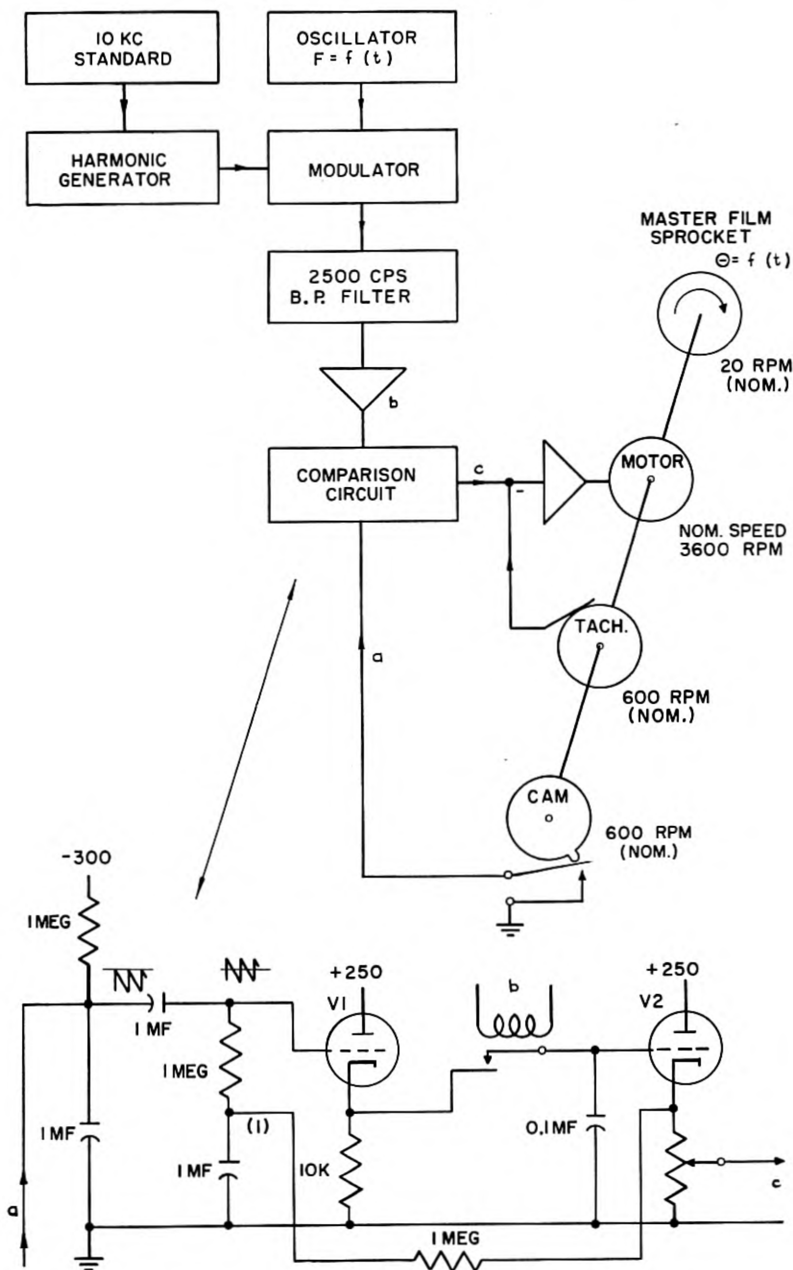


Fig. 10 — Block diagram of film scale calibrator with schematic of comparison circuit.

hydrogen atmosphere, high speed capability and high current capacity. The magnetic arrangement reduces shock torque loads on the servo motor, which might result from mechanical operation. The high speed relay which operates at the 2,500-cycle points is a Western Electric 275-B, chosen because of the speed required (about 10 operations per second).

The time comparison circuit has a small amount of long time constant positive feedback (shown at 1 in Fig. 10) to raise or lower the midvoltage of the sawtooth wave in cases of extreme correction and prevent the control point from slipping one or more teeth. In effect this supplies extra acceleration to the master film when needed.

There is also incorporated in the design an arrangement which permits an important variation in the method of use. A magnetic tape is driven by a sprocket which is geared to the main drive motor and moves with the oscillator drive. The magnetic head for recording on the tape receives its signal in the form of 2,500 c.p.s. bursts through an amplifier. These are the same bursts that time the progress of the oscillator through its spectrum. Thus it is possible to separate the function of calibration from that of printing the film scale. The calibration data on the oscillator is stored on the tape and may be checked for absence of abrupt departures from linearity before it is used to drive the servo and master film in an actual printing run. This eliminates some wastage of raw film. Also a recording (or calibrating) run is made without the servo linkage and can be made at twice the speed of a printing run. A 56A oscillator can be driven through its spectrum, 50 to 10,000 kc, in 100 seconds, allowing very little opportunity for temperature effects to change the check points. In fact no particular effort need be made to control the temperature beyond an ordinary warm up interval.

The control portion of the machine contains various circuits for convenience in setting up and starting the runs. For example one relay circuit under the control of a start button brings a fixed dc voltage into the servo loop, and automatically disconnects after a period long enough for the motor to reach approximately the right speed. A gear shift lever permits changing the ratios between the speeds for the recording run and the printing run.

It is doubtful that a calibration of the 56A oscillator could be performed by manual means. It has been estimated that even if possible, such a task would require more than a week of the most painstaking effort, under very carefully controlled conditions. By comparison, the calibrator requires one minute forty seconds to obtain the data, and three minutes twenty seconds to reproduce it. Development and checking of the exposed film takes about a day. Accuracy of the scales has always been well within the ±2-kc limit.

CONCLUSION

In this and the accompanying articles we have given a partial picture of the facilities for automatic testing in the Western Electric Company. At this writing several new machines are under development, and modifications are in progress extending the application of some of the present machines. There is a continuing search for new fields in which to apply these techniques. A staff portion of the manufacturing engineering force now devotes its full attention to automation techniques in general, keeps abreast of the field, bulletinizes important additions to the literature, lends assistance in the solution of problems, and develops specific applications. It is likely that the near future will see important extensions in the use of automatic test equipment.

ACKNOWLEDGMENTS

The author is indebted to the people cited in the references for information used in this article, and particularly to A. L. Bennett, J. Lamont and F. W. Schramm who furnished valuable comments on the early drafts.

REFERENCES

1. Developed by A. L. Bennett and C. R. Rasmussen.
2. The original design of automatic relay coil test set was developed at Hawthorne by R. W. Brown. The set discussed here is a Kearny modification developed by J. Lamont.
3. C. C. Cole and H. R. Shillington, page 1179 of this issue.
4. Developed by G. H. Harmon and A. E. Rockwood.
5. Developed by F. W. Schramm based on suggestions by T. Slonczewski, Bell Telephone Laboratories.
6. B. M. Wojciechowski, Continuous Incremental Thickness Measurements of Non-Conductive Cable Sheath, B.S.T.J., p. 353, 1954.
7. W. T. Eppler, Thickness Measurement and Control in the Manufacture of Polyethylene Sheath, B.S.T.J., p. 599, 1954.
8. A. N. Hanson, Automatic Testing of Wired Relay Circuits, A.I.E.E. Technical Paper 53-407, Sept., 1953.
9. L. D. Hansen, Tape Control, Automation, p. 26, May, 1956. Also see page 1155 of this issue.
10. Developed by C. F. Dreyer and A. W. Schoof.
11. Developed by L. H. Brown and N. K. Engst.
12. Information was supplied by C. A. Purdy.
13. A. F. Bennett, An Improved Circuit for the Telephone Set, B.S.T.J., p. 611, 1953.
14. Improved U, UA, and Y Type Relays, Bell Lab. Record, p. 466, 1951.
15. J. O. Israel, Broadband Test Oscillator for the L-3 Coaxial Carrier System, Bell Lab. Record, p. 271, July, 1955.
16. W. J. Means and T. Slonczewski, Automatic Calibration of Oscillator Scales, A.I.E.E. Miscellaneous Paper 50-80, Dec., 1949.
17. T. Slonczewski, A Servo System for Heterodyne Oscillators, A.I.E.E. Technical Paper 51-218, May, 1951.
18. F. W. Schramm, Calibrating Strip Type Dials, Electronics, pp. 102–3, May, 1950.

Automatic Manufacturing Testing of Relay Switching Circuits

By L. D. HANSEN

(Manuscript received May 18, 1956)

The large variety and quantity of shop-wired relay switching equipments produced by the Western Electric Company lead to the use of comprehensive and flexible manufacturing testing facilities to insure quality of product and to reduce costs. An older manual type test set is briefly described and used to illustrate the functions and operation of two automatic test sets designated as Card-O-Matic and Tape-O-Matic respectively.

INTRODUCTION

Early telephone central office installations were of the manual switchboard type which were relatively simple and required few relay circuits other than those located in switchboards themselves. Installation effort, in addition to actual erection of the switchboards, equipment frames, fuse boards and the like consisted largely of running and terminating the central office cabling. As the telephone art grew, both with the introduction of the dial telephones, and carrier and repeater equipments for long distance calls and the consequent need for interconnection of these various types of systems, a considerable variety of relay switching circuits was required.

To reduce the installation time and effort the practice of doing as much circuit wiring in the factory as possible was introduced. Relay switching units are now completely assembled, wired to terminal strips and tested in the shop. Since these are in effect working circuits the installation testing effort, after the connection of office cabling, consists largely of overall tests required to insure the proper functioning of the entire office.

Due to the wide variety and complexity of these units, many of which have optional circuit conditions that can be supplied on order and few of which have sufficient demand to justify specially designed high pro-

duction test sets for their exclusive use, adaptable manually operated test sets were first used. These sets required a high degree of flexibility in interconnecting the terminals of the circuit under test to those of the test set and in applying the proper potentials in sequence that would insure putting the circuit through its paces and checking that the switching functions are properly performed.

It should be stated here that since all apparatus components of these circuits such as relays, transformers, capacitors, inductors and resistors are tested and inspected for their respective electrical and mechanical requirements when manufactured, except in the case of some types of relays which require adjustment to meet their particular circuit requirements, the testing of switching circuits is largely confined to verification of the circuit wiring with normal voltages. Although marginal component tests are not normally applied, operation tests will, of course, detect defective apparatus components which cause malfunctioning of the circuit.

MANUAL TEST SET

Fig. 1 shows a representative manual type test set that was extensively used for wired relay unit testing before the introduction of the automatic test sets to be described later. On the left side is a pin jack field into which the numbered wires of the connecting cable can be individually plugged in order to connect the test set terminals to the proper terminals of the relay unit under test. The other end (not shown) of the cable is equipped with a contact fixture arranged to give quick electrical connections to the terminals of the wired relay unit. The plugging of the pins into the proper pin jacks is a feature needed to provide flexibility in a test set arranged to test many types of circuits and is a part of the setup operation for any one circuit. It is a slow and time-consuming operation since each lead has to be identified and plugged into the proper pin jack. The pin plug setup must be taken down and rearranged in order to test any other type of relay circuit.

The test set is equipped with signal lamps for visual response indications and manually operated keys for the use of the tester in performing the test operations. Separate power cords are plugged into power distribution jacks which supply the various potentials commonly used in telephone central offices.

After the initial setup the tester operates the numbered keys and observes the lamp signal responses in accordance with the chart clipped to the front of the test set. Failure to get a particular lamp indication

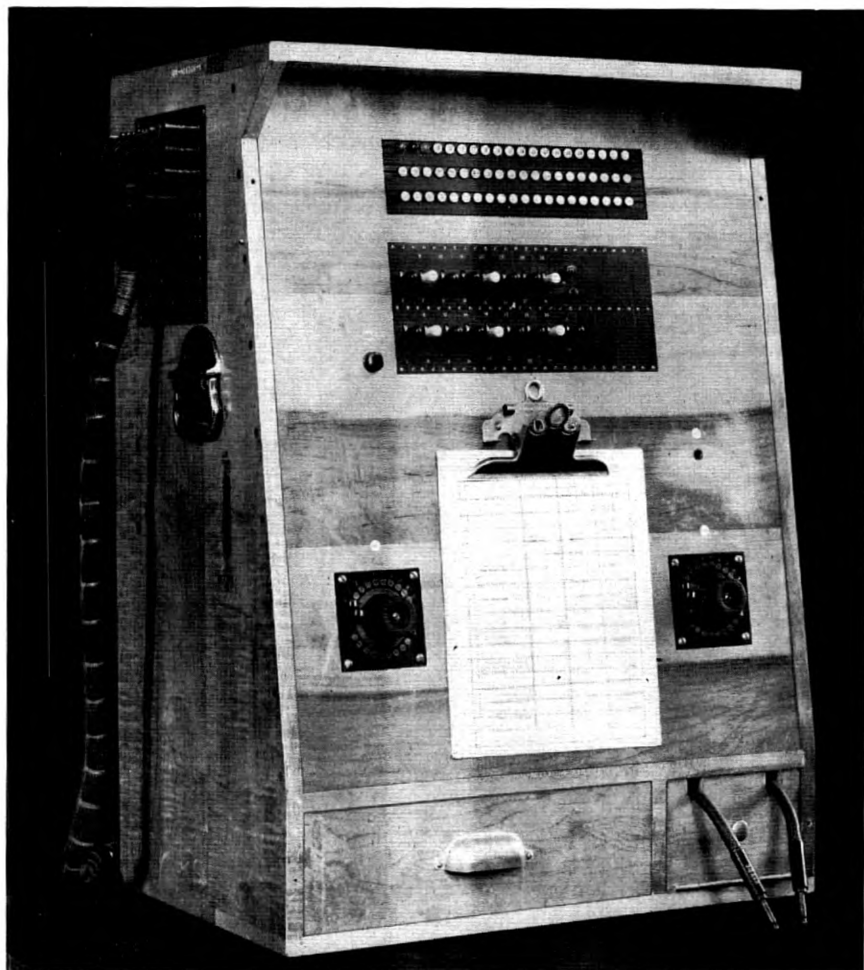


Fig. 1 — Manual wired unit test set.

requires that he analyze the circuit conditions and locate the cause of the trouble. Usually a circuit fault must be corrected before testing can proceed.

Fig. 2 shows a small portion of a simplified circuit test arrangement for such a manual test set. In this illustration a single key, when operated, supplies battery and ground potentials to the winding of a relay in the circuit under test. Assumption is made that the three relay contact terminals are wired directly to the relay unit terminal strip so that

they can be connected to ground and to battery through lamps for circuit closure indications. The switching functions of the relay can then be checked by operating the test key and observing that signal lamp (1) extinguishes and that (2) lights.

While such an arrangement can adequately test most switching circuits of any complexity by further extension of the basic scheme, when supplemented by internal circuit connections where necessary, the

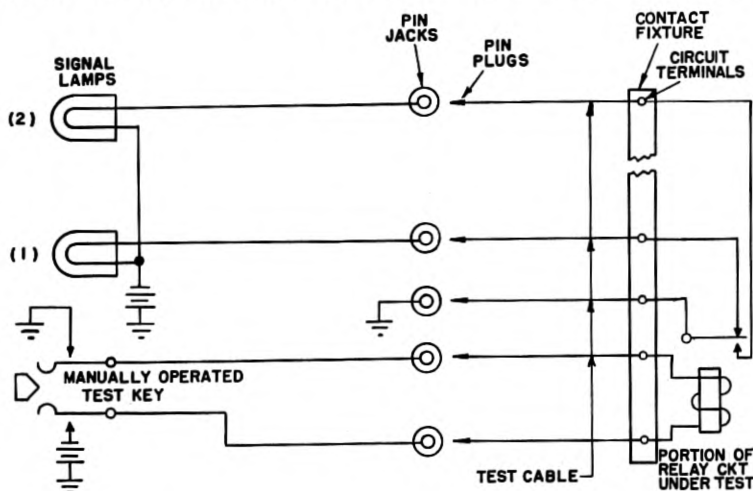


Fig. 2 — Simplified circuit sketch for manual test operation.

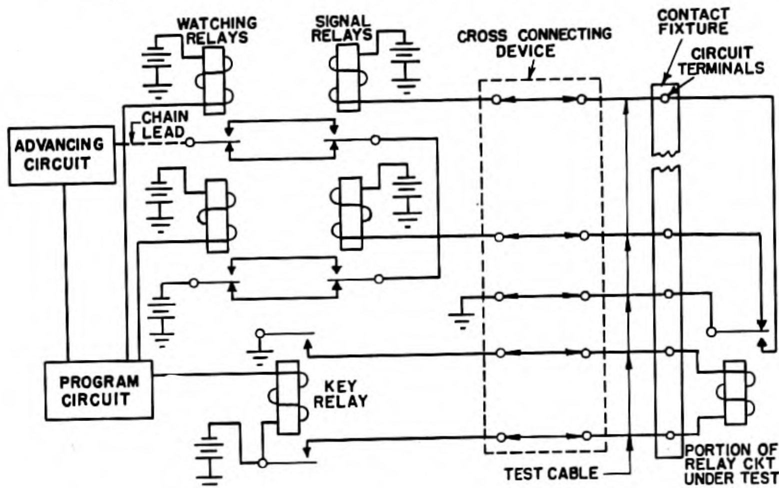


Fig. 3 — Simplified circuit sketch for automatic test operation.

system is at best a slow and laborious one which is subject to human error. Wages for testers are determined not primarily on their ability to operate keys and check the indications of lamps but on their skill in analyzing and clearing trouble conditions. If some quick and automatic means could be devised to make the initial cross connection setup, apply the potentials in the proper sequence under control of some programing device and check the circuit responses at each step a real advance in speeding up tests and reducing human errors would be accomplished Such an automatic set ideally should have improved response indications to aid the the tester in locating circuit troubles when the test set stops on the failure of meeting any test requirement.

THE AUTOMATIC TEST SET

The key and visual lamp indicating functions of the manual test set can be replaced by relays in an automatic test set which perform these operations if they are under control of suitable programing and advancing circuits as shown in Fig. 3. Here the "signal" relays operate through the contacts of the relay under test and their operating positions are checked by the "watching" relays whose contact closures must match those of the signal relays. The series path through the contacts of all signal and watching relays is called a chain lead. The program circuit establishes the positions of the watching relays to meet the expected conditions prior to operating the key relay and then any lack of continuity through the chain lead caused by failure to satisfy test conditions halts the progress of the tests under control of the advancing circuit. At this point additional contacts (not shown) on the signal and watching relays may be used to light signal lamps to convey information to the tester as to which portion of the circuit failed to operate properly.

For quick setup a pre-wired multi-contact adapter plug may be used as a cross-connection device to permit establishing the proper test connections to the unit under test. One will be required for each type of relay circuit to be tested. These, together with some means whereby the sequential operation of the programing circuit can be controlled, constitute the essential features of an elementary automatic relay switching circuit test set. How these basic features can be extended into practical embodiments will be explored further below.

THE CARD-O-MATIC TEST SET

Key equipment relay units are small switching circuits used as circuit building blocks to provide the desired optional features in conjunction with the key boxes or key-in-base telephones often seen in small

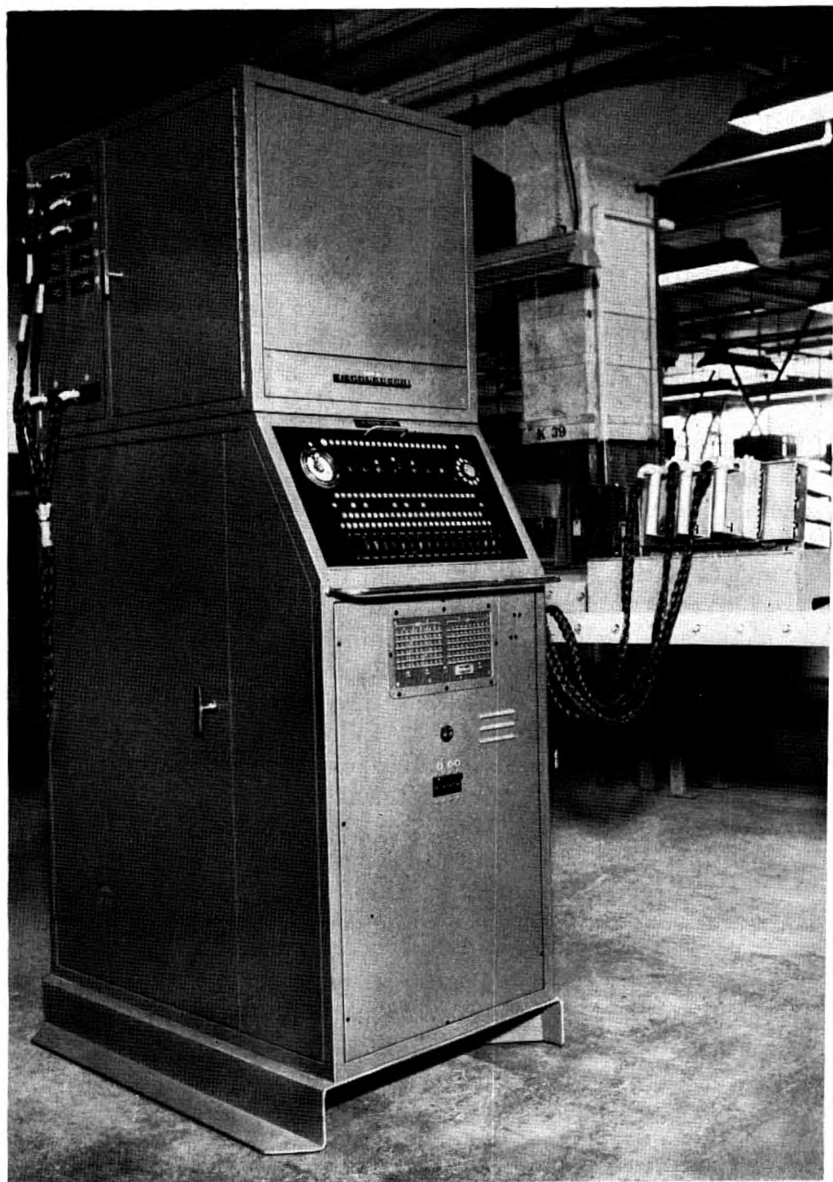


Fig. 4 — Card-O-Matic test set.

business offices to furnish the flexibility needed in answering and transferring calls. These systems are used where the number of telephones served does not warrant the use of a regular PBX switchboard.

These circuits are relatively simple but their large scale production warrants the use of high speed automatic test sets to perform the test functions and to indicate circuit trouble.

Fig. 4 shows the operating position of the Card-O-Matic* test set which was developed to test such unit assemblies. The keys shown are used to initiate and control the automatic operation of the test set and in trouble shooting. They are not to be confused with those that perform the actual testing functions described previously for the manual test set. The lamps provide indications of the progress of the tests and of the positions of the watching relays which are also needed to aid in determining the point of circuit failure. The meter type relay in the upper left corner of the operating panel provides a sensitive checking device for audio frequency tests through the voice transmission circuits. The telephone dial affords a simple means of generating any required number of pulses for operating stepping selectors on some types of units. The terminal field in the lower front of the cabinet gives the tester access to the circuit terminals of both the unit under test and the test set for his use in analyzing and locating faults. The upper cabinet was a later addition and contains the multi-contact relays needed to permit testing units with more than one circuit. The row of push buttons are used to select the circuit to be tested.

Fig. 5 is a rear view of the set that shows the perforated insulating card from which the set derives its name. The coded card controls the sequence of test operations and is hung on pins over the field of 1,000 spring plungers (20 × 50) as a part of the setup operation for a particular relay unit. Closing the door and screwing up the hand wheel, which is necessary to provide the force required to depress the plungers, will ground those which coincide with holes in that particular card.

Cross-connection setup of the test leads is achieved by the use of a plug-board such as is commonly used for quick change over on perforated card type business machines. Fig. 6 shows the plug board being inserted into the transport mechanism. The relatively large number of terminals are required because each of 60 test leads must be capable of being patched in to an equivalent number of terminals on a maximum of ten different circuits. Not all of our test sets are equipped with the upper cabinet since most key units have only one circuit and on these a simpler

* Patent No. 2,329,491.

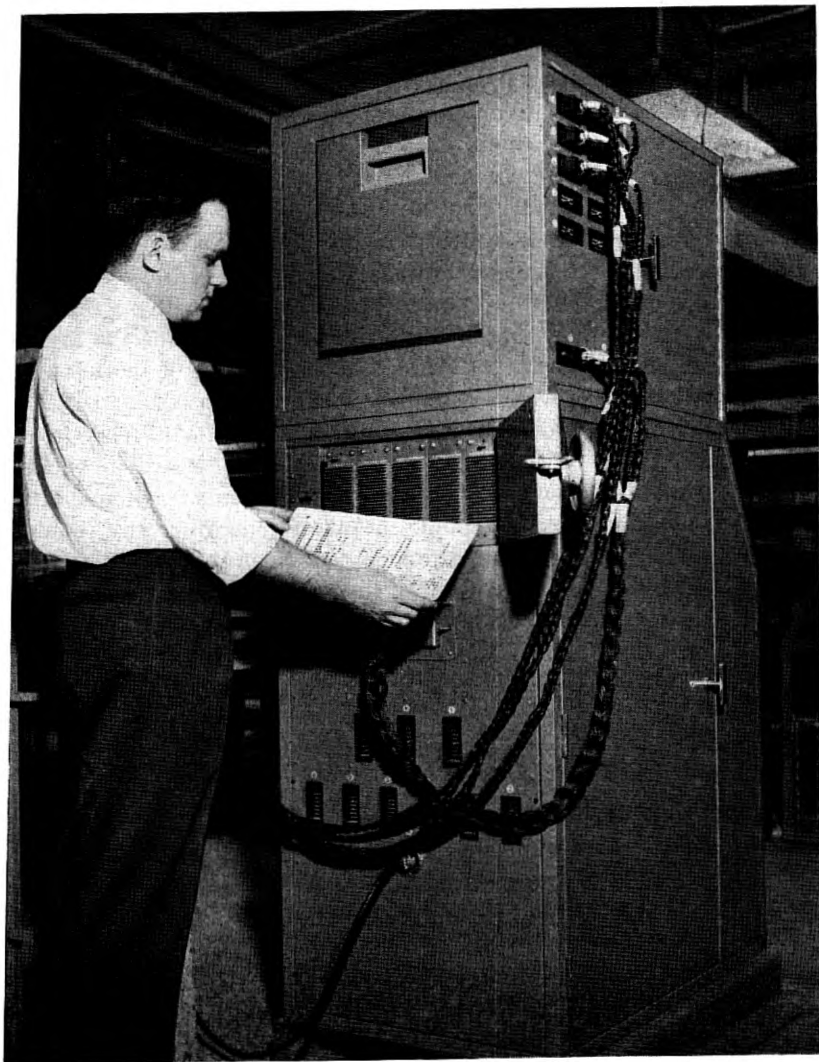


Fig. 5 — Rear view of Card-O-Matic test set showing insertion of perforated card.

cross connection fixture is plugged into the location where the lower end of the cable joining the two cabinets is shown terminated in Fig. 5.

A side view of the test set is shown in Fig. 7 to give an indication of the amount of switching equipment and wiring necessary for an automatic test set of this sort. The set is powered from a 120-volt 60-cycle

source from which are derived the 24-volt dc, 90-volt 20-cycle ringing current and 600-cycle audio tone supplies that are required. The test circuit features include tone transmission checking, dial pulsing, 90-volt 20-cycle ringing and ground and battery supplied either directly or under relay control. Other battery and ground relays are available for checking the response of the circuit under test.

These test features have been sufficient to perform operation tests on most relay units associated with key telephone systems. The test cycle is fast and the twenty test steps can be performed in approximately ten seconds. The lamp indications given when the test is interrupted by an open-circuited chain lead, convey information to the tester as to which test step is involved and when any pairs of signal and watch relays fail

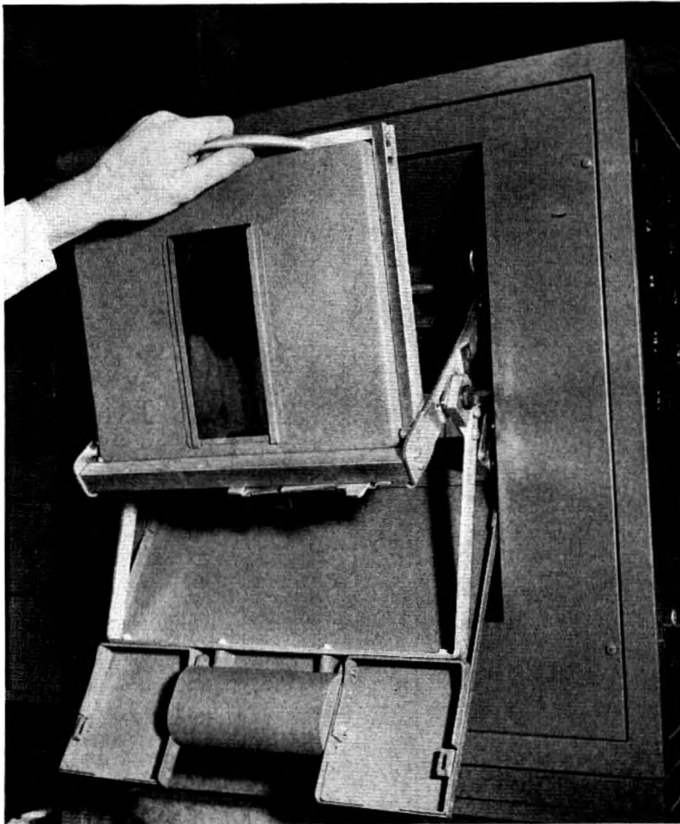


Fig. 6 — Insertion of cross-connection plug board into Card-O-Matic test set.

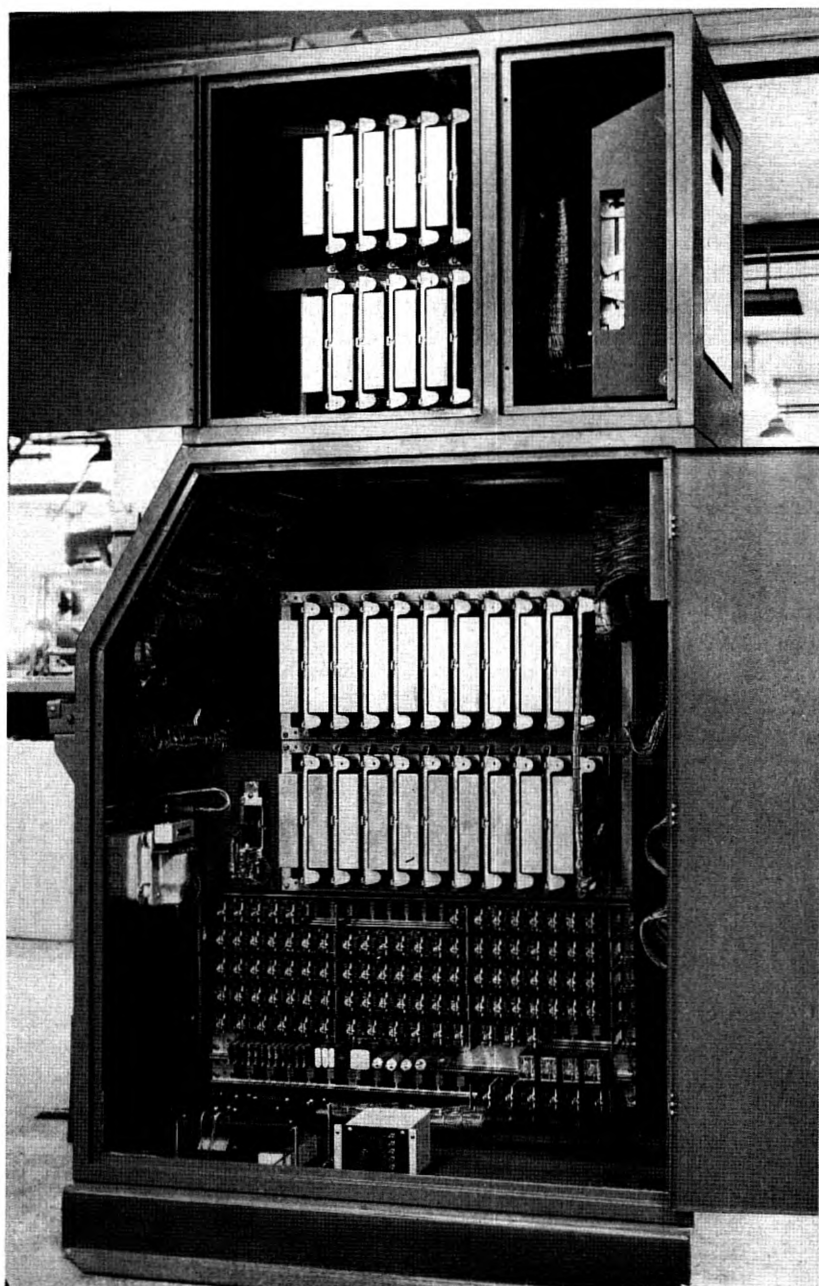


Fig. 7 — Interior of Card-O-Matic test set.

to match each other. Simplified circuit sketches which show the interconnection of test set and wired unit circuits are provided to enable the tester to determine quickly the cause of the failure.

The Card-O-Matic test set, while performing admirably on the relatively simple relay circuits within its range and capabilities, falls down on the more complicated relay switching circuits used in telephone central offices for several reasons. The most important of these are:

1. A fixed cycle within a maximum of twenty steps with any one coded card.
2. No provision for alternate or optional circuit conditions on a card.
3. The only power supplies provided to operate relays are negative 24-volt dc and 90-volt 20-cycle ringing whereas telephone office units frequently also require negative 48-volt and positive 130-volt dc as well as positive or negative biased ringing currents for party line ringing.
4. The increase of either test steps or features would increase the size of the perforated card beyond a practical size.

THE TAPE-O-MATIC TEST SET

The experience gained in the design and successful operation of the Card-O-Matic test set led naturally to the exploration of ways and means whereby a more versatile and comprehensive set could be devised. The five hole coded perforated teletype tape was selected as a cheap and flexible programing device. It afforded a means of providing a test cycle of any required length and, since the perforating and reading mechanisms were already available, it appeared to be nearly ideal for its purpose.

Consideration was given to the following desirable features all of which were incorporated in the design of the new set:

1. Provision for cross-connecting (under control of the coded tape) any test set circuit to any terminal of the circuit under test for as long as necessary and then disconnecting for reuse in later testing steps if required. This would greatly extend the range and capabilities of the set.
2. Provision for several power voltage sources which could be selected as required to meet the normal telephone office voltage requirements of the unit under test.
3. Provision for alternate or optional tests to be coded into the tape to meet the various circuit arrangements that may be wired into the unit as required by the Telephone Company who is our customer. Such optional test arrangements could be applied by the test set under the control of keys to be operated by the tester as part of the setup at the start of the tests.

4. Provision for stopping the test cycle to enable the tester to perform manual operations such as inserting a test plug in a jack on the unit or insulating relay contacts in order to isolate portions of the circuit for test simplification and to obtain a more detailed test.

5. Provision of improved lamp indications to aid the tester in clearing wiring faults or in locating defective apparatus. These would include the necessary information as to which test set circuits are connected to which unit terminals as well as which relays of the wired unit should be operated at that stage of the test cycle.

6. Provision for connecting several terminals of the unit under test together as a means of providing circuit continuity where required.

7. Provision for measuring resistance values of circuit components.

8. Provision for insertion of various resistors in battery or ground leads to control currents to desired values.

9. Provision for checking voice transmission paths through nonmetallic circuits such as transformers or capacitors.

10. Provision for measuring circuit operating times in steps of approximately 100 milliseconds.

11. Provision for sending and receiving dial pulses.

12. Provision for a single code for releasing all test connections and conditions previously established by the coded tape as a means of quick disconnect. This is in addition to the release of individual connections mentioned in (1) above.

13. Provision for audible and visual indications of completion of a successful test cycle.

Through the use of two letters (each of which has its own combination of the five holes) for each signal it was possible to obtain the over 500 codes required to control all test and switching functions even though the teletype keyboard has only 32 keys. The only Teletype transmitter (tape reader) available when the test set was first designed operated at a speed of 368 operations per minute and was arranged for sequential read out on two wires by means of a commutator. Conversion to five wire operation and removing the commutator permitted reading each row of holes simultaneously. The gearing was also changed to permit 600 operations per minute but even so the hole reading contact dwell time was increased from approximately 20 milliseconds to 70 milliseconds for more reliable operation with ordinary telephone relays.

The machine which was designated as the Tape-O-Matic* test set, is shown in Fig. 8 in operation on a typical wired relay unit mounted in its shipping frame. The contact fixture is attached to the unit terminal

* Patent No. 2,328,750.

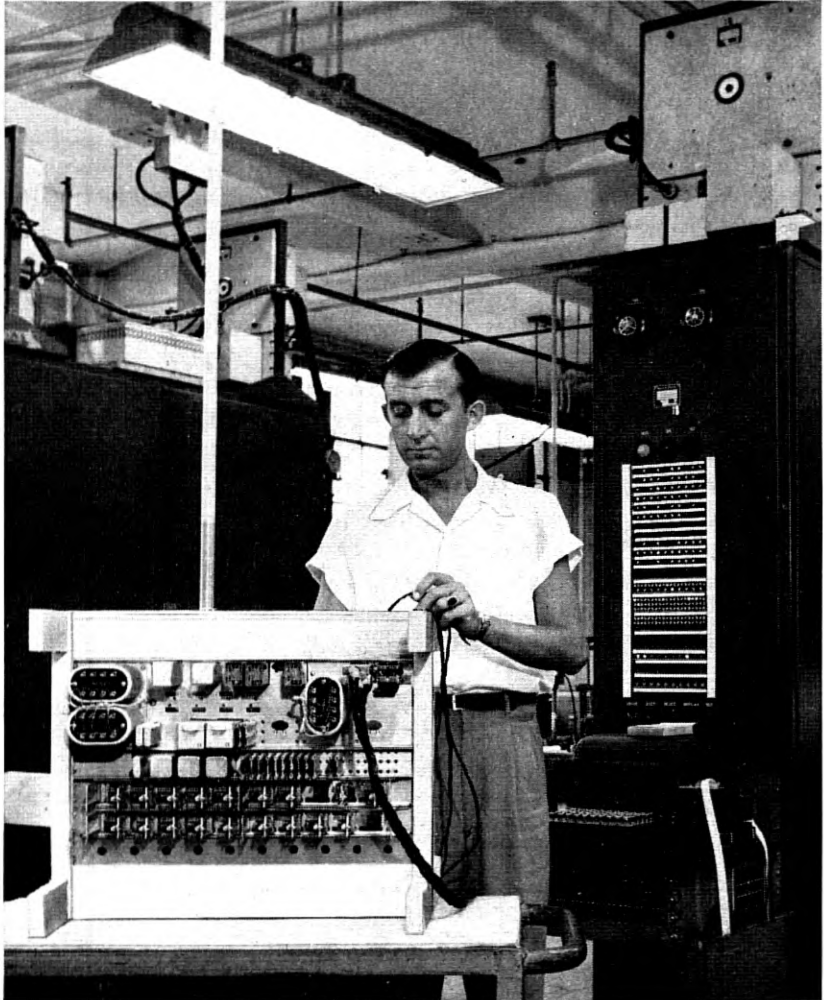


Fig. 8 — Tape-O-Matic test set in operation.

strip and cabled to a gang plug which in turn is plugged into a receptacle behind the operator. These leads are extended through a duct to the metal enclosure at the base of the set for entry to the test set proper. The coded tape is dropped into the receptacle at the side of the key shelf to which it returns after its traverse through the reader. A row of circuit breakers on the end of the key shelf control the application of

and provide protection for the various power supplies. Two of these supplies are mounted on the top of the set.

The rows of vertical push button keys on the key shelf afford the tester a means of determining (for trouble shooting purposes) the association (through lamp display signals) of the wired unit circuit terminals with those of the test set and the corresponding test voltages which are connected at that particular stage of the test. The lamp display panel also indicates which test set circuits are in use and through fast or slow (0.5 or 1 second) flashes whether the fault thus indicated is the result of a failure to meet either an expected condition or the occurrence of an unexpected condition. This feature is illustrated in Fig. 9 which shows one link of the chain leads which extend through all pairs of signal and watching relays for the check of satisfaction of all test conditions and the application of steady or interrupted ground to the associated test feature lamp. The operating condition of all test set key relays as previously established by the tape is also indicated by the display lamps. Another type of information obtained from the lamp display panel which is valuable to the tester in trouble clearing is the indication of the particular unit relays which should be operated at that part of the test cycle. By checking the lamps against the operated or non-operated position of the relays he can frequently localize the fault in a minimum of time.

As mentioned above an important part of the test set flexibility is the ability of the tester to set up the test set to test only those optional circuit arrangments which are provided in any particular unit ordered

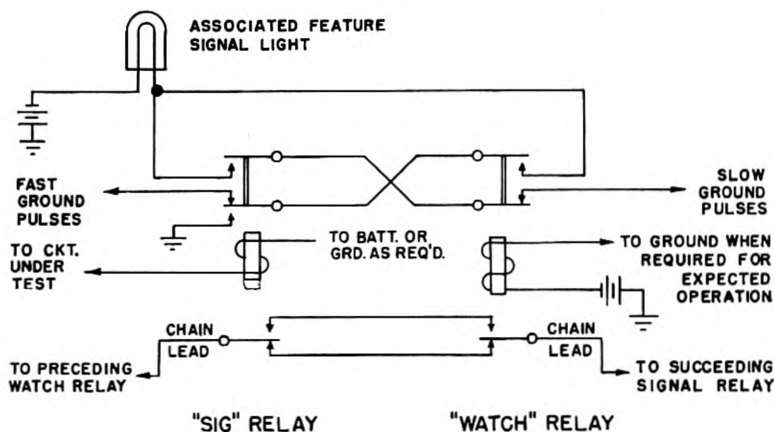


Fig. 9 — Chain circuit showing watching relay function.

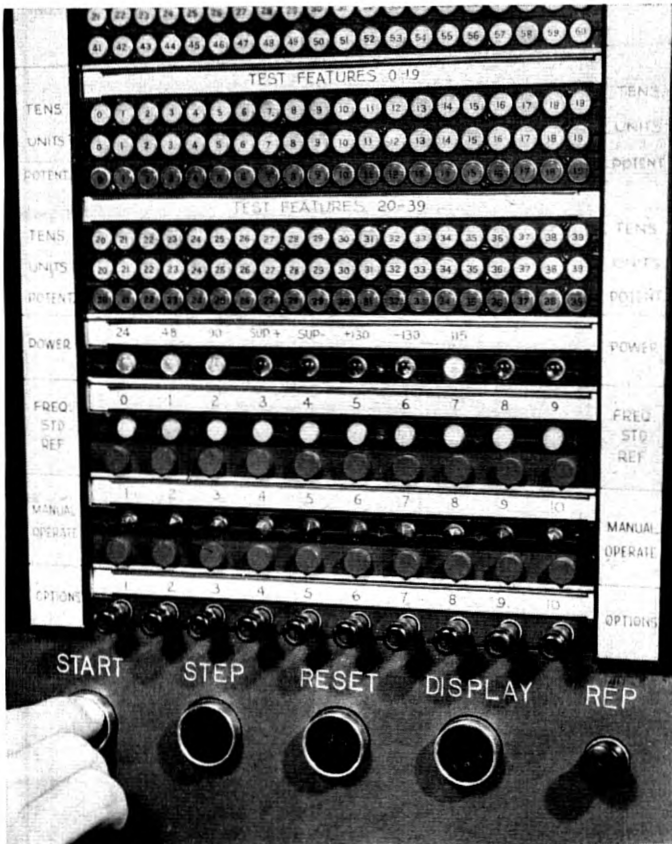


Fig. 10 — Lower portion of lamp display panel.

by the customer. Failure to provide this would result in fixed test cycles and many more tapes, which might be similar but varying only in regard to the options, would have to be prepared. Figure 10 shows the lower portion of the lamp display panel with the push-pull option keys on the bottom row. Directly above are the manual operation keys with their associated lamps which the tester must operate to cause the test set to resume the testing cycle after it has stopped for him to perform a manual operation.

A side view of the interior of the set is shown in Fig. 11. Two bays each facing the oppcsite direction from the other are housed within the cabinet and are used for mounting the crossbar switches and telephone type relays which are the principal circuit components. Two doors on

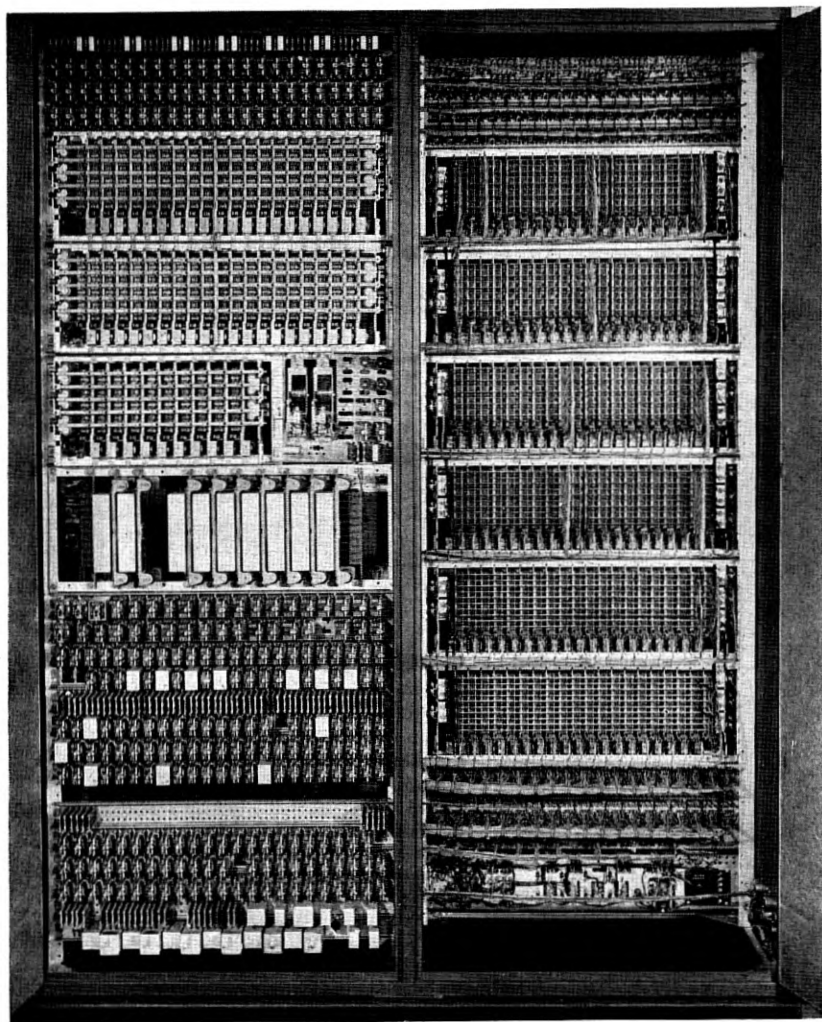


Fig. 11 — Interior of Tape-O-Matic test set.

each side give convenient access to all wiring and apparatus for maintenance purposes.

A fairly large portion of the mounting space is occupied by the crossbar switches which perform the functions of interconnecting the circuit terminals of the unit under test and those of the test set. They also connect the proper voltages to these circuits. The switching plan Fig. 12

shows in abbreviated diagramatic form that the unit terminals 0-99 appear on the horizontal inputs of the two 10 × 20 and one 10 × 10 switches that comprise the primary group. The horizontal multiple of these switches are split so that each section runs through five verticals to afford connection to each of the hundred unit terminals.

The vertical outputs of the primary switches are connected to the horizontal inputs of the two 10 × 20 secondary switches. The horizontal multiple of these switches are split so that each section runs through eight verticals. The verticals of the secondary switches are linked to the horizontals of the two 10 × 20 tertiary switches which have their

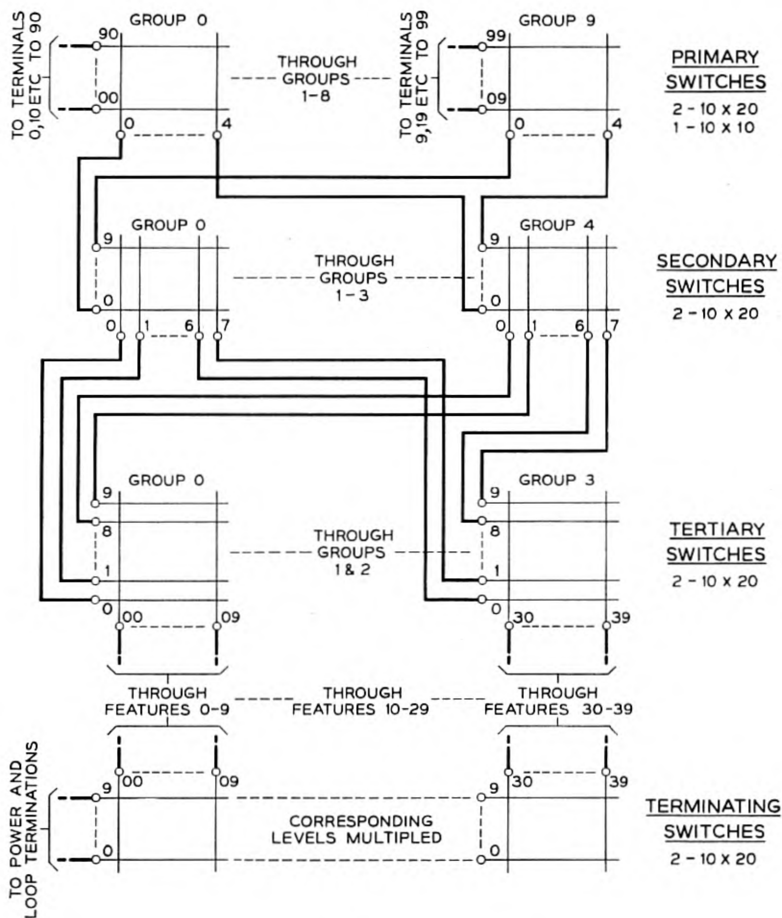


Fig. 12 — Switching plan.

multiple split into groups of ten. The 40 verticals of the latter are connected directly to the 40 test set feature circuits designated 0-39.

Two additional 10 × 20 crossbar switches perform the function of connecting any of the five power or five multiple terminations to any of the forty test set features. These terminations are comprised of 5 loops and one each of ground, negative 24 volts, negative 48 volts, 90-volt 20-cycle ringing current and positive 130 volts.

Thus it can be seen that, through proper operation of the primary, secondary, tertiary and terminating crossbar switch cross points, a path can be established from any circuit terminal to any test set feature and supplied with any of the available power or loop terminations. It is

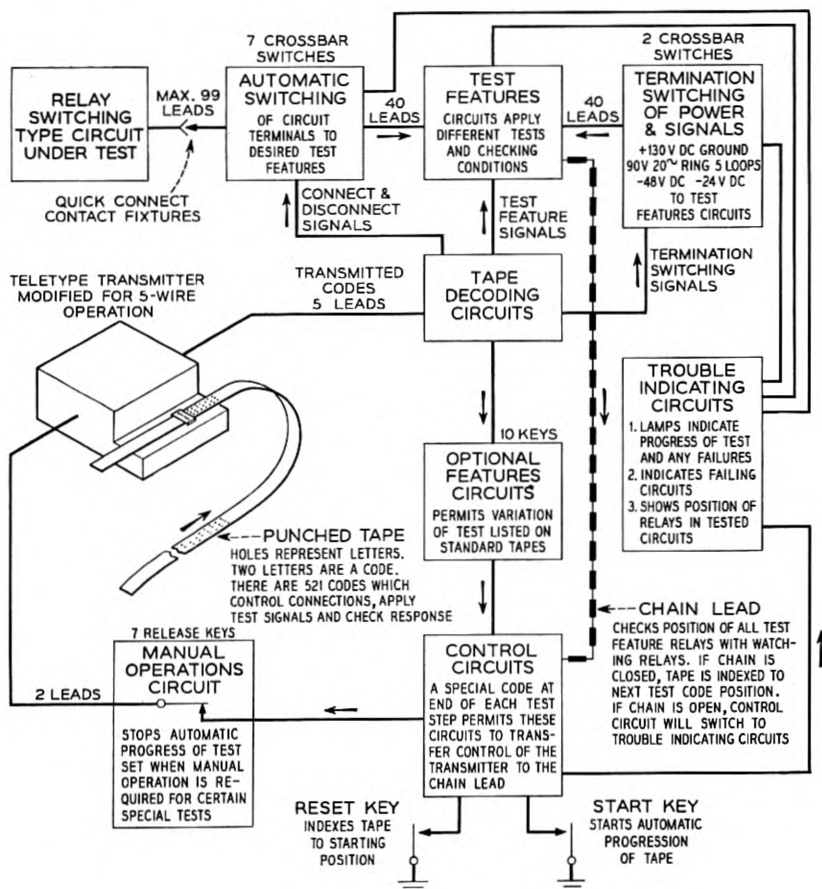


Fig. 13 — Block schematic.

also apparent that several paths can be found that will satisfy any one switched connection. Paths are assigned in sequence by a series relay loop circuit. The entry point on this circuit is changed periodically to distribute wear on the relays and switch cross points.

Although only one lead for the switched circuit is shown for each cross point in Fig. 12 there are actually four leads through corresponding pairs of contacts through each cross point. The remaining leads are associated with the holding and signalling functions of the switch.

The block schematic (Fig. 13) shows the principal functions which must be included in an automatic test set of this sort. A somewhat more detailed schematic is presented in Fig. 14 in order to show the functions of the forty test features 0-39. These are tabulated in Table I.

The coding of the two letter combinations in the tape must follow a definite sequence in order that the machine may recognize and act on the information it receives. This sequence is as follows:

1. Code FW to stop the tape at the end of the reset cycle after which tests will proceed when the start button is pressed. This is the first code on all tapes.

2. Codes to set up crossbar switches to connect each circuit terminal to its proper test set terminal and the proper termination. Knock down or release codes may also be sent.

3. Codes to operate or release "Key" relays. These relays are shown without windings in Fig. 14.

4. Codes to operate or release the watching relays associated with the "Signal" relays which are shown with windings in Fig. 14.

5. Codes to operate or release relays controlling the lamps associated with relays in the circuit under test to aid in trouble shooting.

6. Codes to delay the timing out interval up to a maximum of ten seconds.

7. Code FJ which checks the matching of all signal and watching relays through the chain circuit for satisfaction of all test conditions being applied.

In addition to the above, additional codes can be inserted after each FJ test signal to stop progress of the test to permit the tester to perform some required manual operation. After completion of this step he presses a button associated with that operation and the test proceeds. Option codes can also be inserted at the beginning and end of each testing step to permit bypassing of that part of the tape if the corresponding option keys are operated at the beginning of the test. A common knock down code FR can be inserted at any time to release all connections and relays for a quick disconnect and make all test set features available for

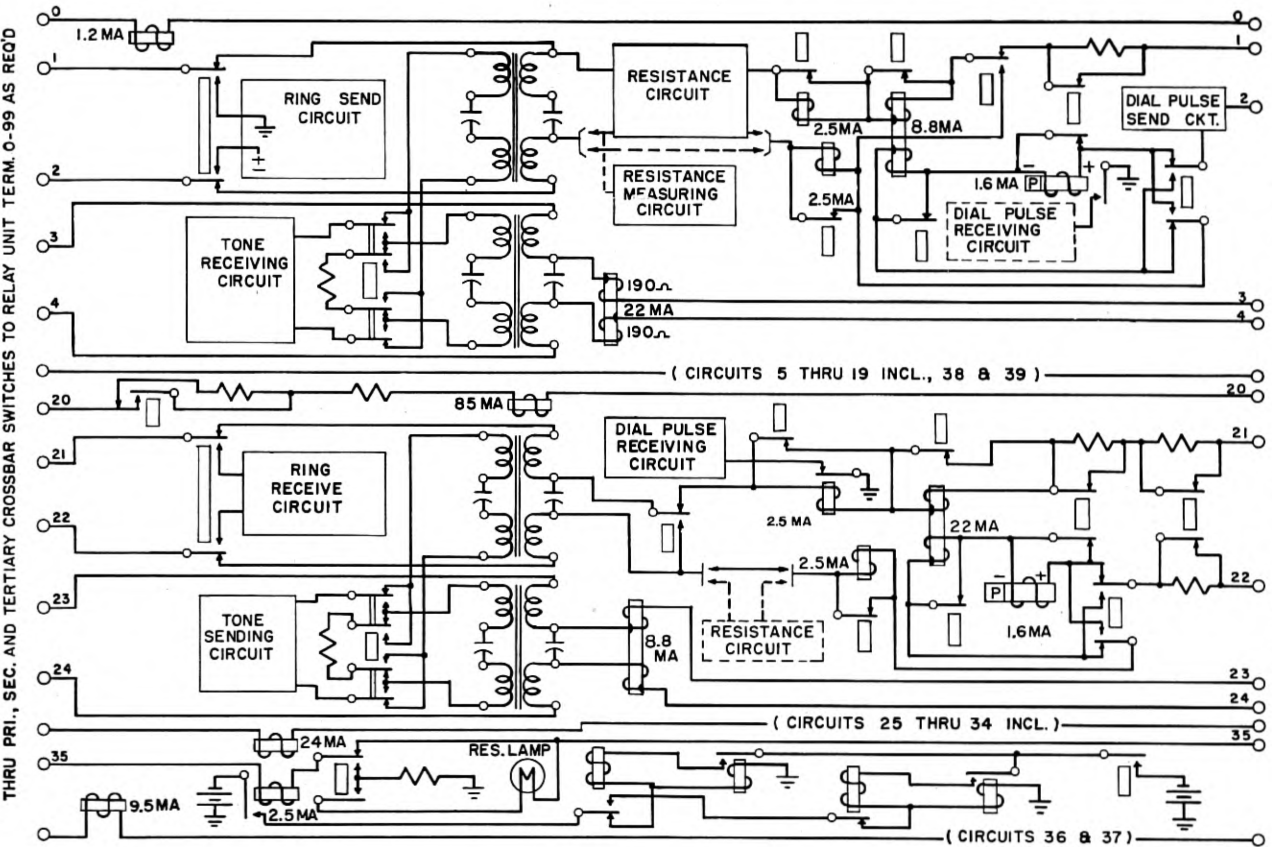


Fig. 14 — Functional schematic.

reuse. A final code SC must be put in every tape to operate the OK lamp and gong if a successful test cycle has been performed or conversely to indicate that the tape should be re-run if trouble has been found and cleared during the test cycle to be certain that no new faults have been introduced.

Preparation for testing a particular wired relay unit requires only the selection of a test cable one end of which is equipped with a suitable contact fixture for attachment to the unit terminal strip and the other with a gang plug for connection to the set. The proper tape is selected from a nearby file cabinet and inserted in the gate of the tape reader as shown in Fig. 15. The tape is stored in a cardboard carton $3\frac{3}{4} \times 4$

TABLE I

Feature Numbers	Description of Functions
0	High sensitivity relay circuit. Simulates 1,800-ohm sleeve circuit for busy test and general continuity through high resistance ciruits.
1 and 2	Simulates the distant tip and ring terminations of a subscriber or exchange trunk. Provides for ringing, tone receival, dial pulse sending, line resistance, high-low or reverse battery supervision, pad control, continuity, and resistance verification.
3 and 4	Auxiliary tip and ring circuit for holding, checking continuity, receival of tone on four wire or hybrid coil circuits. Loss range of less than 0.5 db, 0.5 to 1.5 db, 1.5 to 6 db and 6 to 15 db can be checked.
5 through 19	Direct connections for supplying any of the ten terminating conditions.
20	Simulates low or medium resistance sleeve circuits for marginal tests.
21 and 22	Simulates the local tip and ring terminations of a switchboard or trunk circuit. Provides for ringing and dialing receival, high-low reverse battery supervision, transmission pad control, tone transmission, continuity and resistance check by balance.
23 and 24	An auxiliary tip and ring circuit for holding, checking continuity, tone transmission on four-wire hybrid coil circuits.
25 through 34	Low sensitivity relay circuits for general continuity checking.
35	A circuit for checking balance on the (M) lead of composite or simplex signalling circuits and for checking receival of none, one or two pulses.
36 and 37	Medium sensitivity relay circuits for continuity checking.
38 and 39	Direct connections for supplying any of the ten terminations.

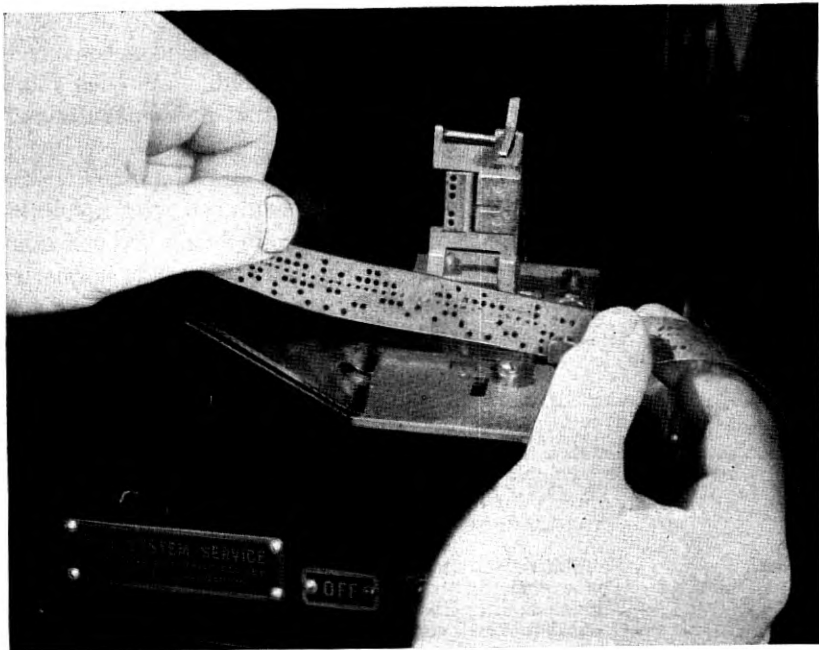


Fig. 15 — Perforated tape being inserted in reader.

inches in size, the label of which carries all pertinent information required for setup of the option keys, preliminary tests and manual operations during test. A separate 12 conductor cable equipped with individual test clips permits connection to internal parts of the circuit if needed for adequate tests. No other information than that on the box label, the circuit schematic and the lamp panel display is needed by the tester to operate the test set and to analyze and locate circuit faults when they occur.

With the tape inserted, the test connections established and any preliminary operations performed the tester has only to push the RESET button to index the tape to the initial perforation on the tape and the START button to initiate the test cycle. The set will continue to operate until either a circuit trouble is encountered or a manual operation must be performed. After a defect has been repaired, the automatic progression of the tape is again started by the momentary depression of the STEP button. When a manual operation is performed the tape is restarted by the momentary depression of the red button associated with the lighted manual operation lamp signal.

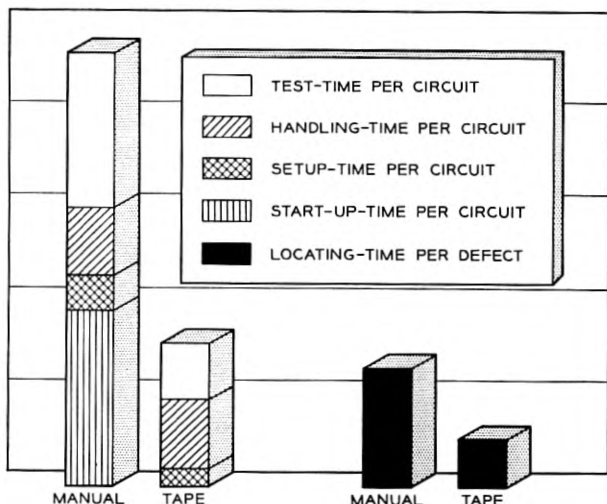


Fig. 16 — Comparison of manual and Tape-O-Matic test operation times.

As might be expected the easy setup, automatic testing and superior trouble indicating features of the Tape-O-Matic test set have materially improved the quality and reduced the testing time and effort required for wired relay units as compared to the older manually operated sets. The average time per circuit for six representative units are shown graphically in Fig. 16. One time consuming operation on manual testing is the start up time allowance for reading and understanding the written test instructions which has no counterpart in the Tape-O-Matic tests and this alone represents a sizeable gain. The handling time of the unit itself is the only operation which is not reduced in automatic testing.

HISTORY

The initial Card-O-Matic test set was installed in 1938 in the Western Electric, Kearny, New Jersey plant. Post war and subsequent expansions of production levels have necessitated construction of six more sets of improved design of the type described earlier in this article.

The first three Tape-O-Matic test sets were built in 1942 for the Wired Relay Unit Shop and additional sets have since been constructed to bring the number to twenty-six including six that are used in testing trunk units in the Toll Crossbar Shop. They have performed admirably with few changes from the initial design. They have been used to test well over a million wired units with a minimum of maintenance. This

may be accounted for, in part, by the fact that most of the component parts are telephone type apparatus designed for heavy duty use.

A maintenance feature is the use of 18 specially coded tapes which, together with a properly strapped input plug, permit the maintenance technician to obtain indications on the lamp display panel of the performance of the set.

Nearly three thousand tapes have been coded to date. Of these approximately two thousand are in active use on the many types of wired relay units made at the Kearny plant. More tapes are being added weekly as the Bell System telephone plant grows in size and complexity.

CONCLUSION

Automatic testing of wired relay switching circuits has been successfully applied to the manufacture of these equipments at the Kearny, New Jersey, plant of the Western Electric Company for a number of years. Even though the total production is large, manufacture is essentially of a job lot nature due to large number of types made and is further compounded by the optional circuit arrangments that may be ordered. The solution to the problem was found through provision of flexibility in programing and cross connection leading to quick setup, rapid testing and improved transmittal of essential information to the tester to aid him in clearing circuit faults.

Automatic Machine for Testing Capacitors and Resistance-Capacitance Networks

By C. C. COLE and H. R. SHILLINGTON

(Manuscript received May 8, 1956)

The modern telephone system consists of a variety of electrical components connected as a complex network. Each year, millions of relays, capacitors, resistors, fuses, protectors, and other forms of apparatus are made for use in telephone equipment for the Bell System. Each piece of apparatus must meet its design requirements, if the system is to function properly. This article describes an automatic machine developed by the Western Electric Company for testing paper capacitors and resistance-capacitance networks used in central office switching equipment.

INTRODUCTION

The capacitors discussed in this article are the ordinary broad limit units made with windings of paper and metal foil, packaged in a metal case. They include both single and double units in a package, connected to two, three, or four terminals. The networks consist of a capacitor of this same type connected in series with a resistor.

The testing requirements for capacitors include dielectric strength, capacitance, and insulation resistance. These same tests plus impedance measurements are specified for networks. In general, requirements of the kind involved here could be adequately verified by statistical sampling inspection. However, in equipment as complex as automatic telephone switching frames, even the minor number of dielectric failures that would elude a properly designed sampling inspection would result in an intolerable expense in the assembly and wiring operations. While engineering considerations thus called for a detailed inspection for dielectric breakdown, it was recognized that detailed inspection of the other electrical requirements could be obtained at no additional expense for labor with automatic testing machines.

DESIGN CONSIDERATIONS

In the development of this machine, the designer was faced with the same problems that obtain in the conception and design of any unit of complex equipment. These included the economic feasibility, reliability, simplicity, and versatility of such a machine.

Economic Feasibility

This can be determined by comparing the cost of performing the operations to be made by the proposed machine with the cost by alternative methods. Estimates indicated that the cost of the machines could be recovered within two years by the saving in labor that would be effected.

Reliability

Reliability has two connotations, (1) freedom from interruptions of production because of mechanical or electrical failure and (2) consistent reproducible performance. A rugged mechanical design combined with the use of the most reliable electrical components available is necessary. In addition, safeguards are required to protect the equipment from mechanical or electrical damage. To achieve consistent reproducible performance, it is important that testing circuits of adequate stability be used. Besides, it was recognized that each circuit should be so arranged that in case of a circuit failure, there would be immediate and positive action by the machine to prevent acceptance of defective product. All circuits are designed to provide positive acceptance. This means that the machine must take action to accept each item of product at each test position. In the case of the dielectric strength tests, a self-checking feature is included.

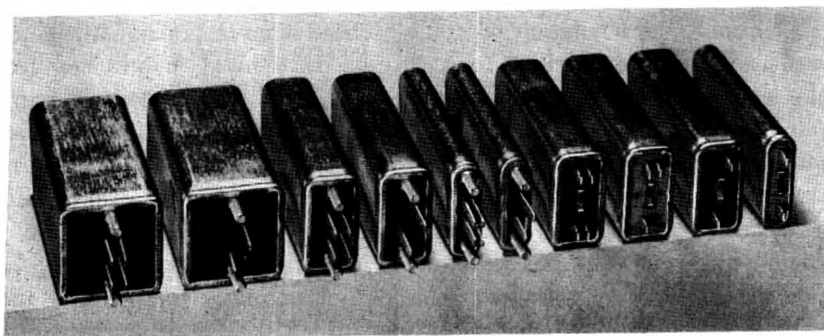


Fig. 1 — Types of capacitors and networks tested.

Simplicity

This type of equipment is operated by non-technical personnel. To minimize the possibility of improper operation of the equipment, it is important that adjustments and judgment decisions by the operator be minimized. From a production standpoint, it is important that the machine be designed to permit quick changes to handle the variety of product to be tested. All "set-ups" are made by the operator and the switching of circuits and changing of contact fixtures are simply and easily done.

Versatility

The product tested by this machine includes a variety of physical sizes and terminal arrangements with a wide range of electrical test requirements (Fig. 1).

a. *Physical Sizes.* The aluminum containers for this type of capacitors and R.C. Networks all have the same nominal length and width but are made in three different thicknesses.

b. *Terminals.* The product is made with terminals of two different lenths, two different spacings, and four different patterns connected in eight combinations. It is necessary to provide contact fixtures and switching facilities to handle all of these combinations.

c. *Electrical Tests*

(1) Dielectric strength tests are made between terminals, and between terminals and can, on single unit packages. Two-unit packages require an additional test between units.

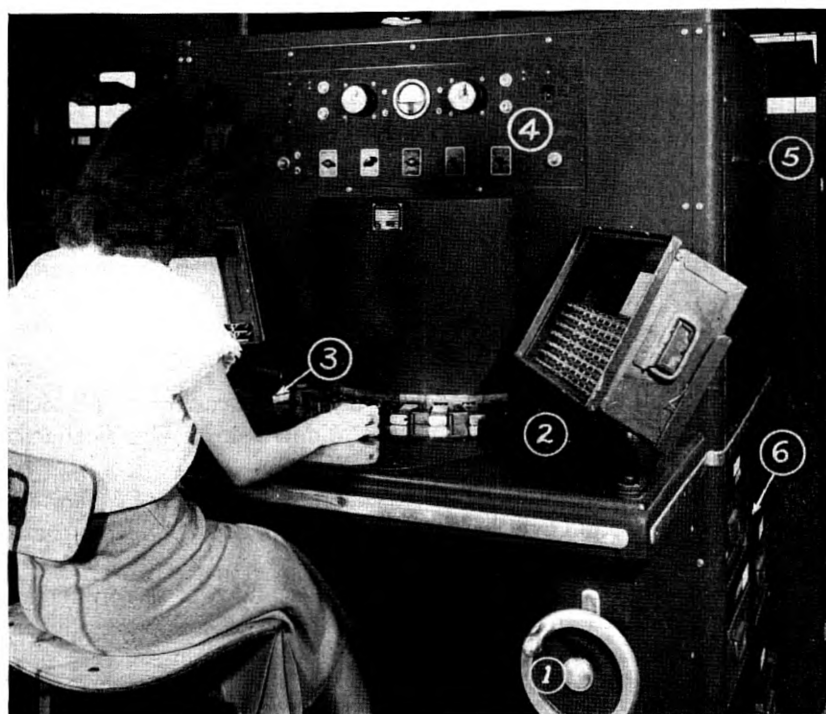
(2) Capacitance: The capacitance of the product to be tested ranges from 0.02 mf to 5.0 mf or any combination within this range in one- or two-unit packages with no series resistance in the case of capacitors, but with a series resistor from 100 ohms to 1,000 ohms in the case of networks. This problem is discussed in more detail in the description of the capacitance test circuit.

(3) Insulation Resistance: The minimum requirements vary from 375 megohms to 3,000 megohms.

(4) Impedance: The RC networks have impedance requirements at 15 kc that range from 100 ohms to 1,000 ohms.

MECHANICAL ASPECTS OF TESTING MACHINE

Packaging of the product precludes a magazine type of feed because the variety of terminal combinations associated with two-unit packages necessitates orientation in the contact fixtures that can not be done by



1. HANDWHEEL FOR POSITIONING TEST FIXTURES.
2. ROTARY FEED MECHANISM.
3. PRODUCT PASSING ALL TESTS EJECTED FROM FIXTURE.
4. INSULATION RESISTANCE TEST PANEL AND TERMINAL COMBINATION "SETUP" SWITCHES.
5. CABINET HOUSING TEST CIRCUITS.
6. CONTAINERS FOR REJECTED PRODUCT.

Fig. 2 — Testing machine in operation.

mechanical means. A turret type construction is used to permit one operator to perform both the loading and unloading operations.

Fig. 2 shows this machine in operation. The networks or capacitors are fed into the fixtures by an operator and as the turret carries the fixtures past the feed mechanism, rollers on the feed mechanism are synchronized with the fixtures and the roller forces the unit under test into the contact fixture against a spring loaded plunger to make contact with the fixture contact springs. Also, synchronized with the feed mechanism is the closing of the gripper hook on the bottom end of the can containing the unit under test.

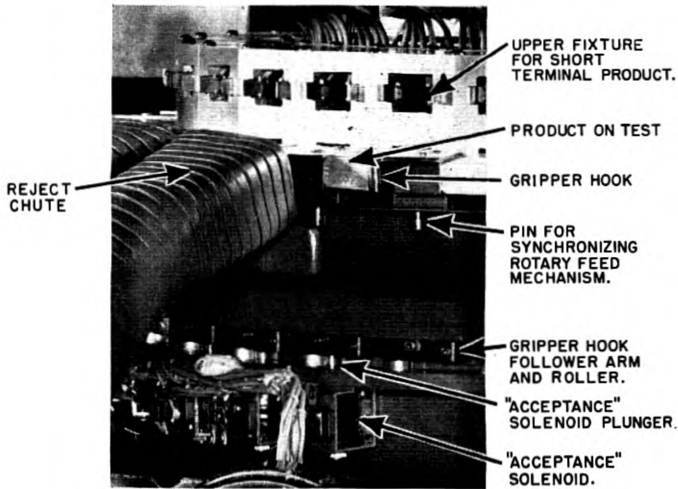


Fig. 3 — View of rejection and acceptance mechanisms.

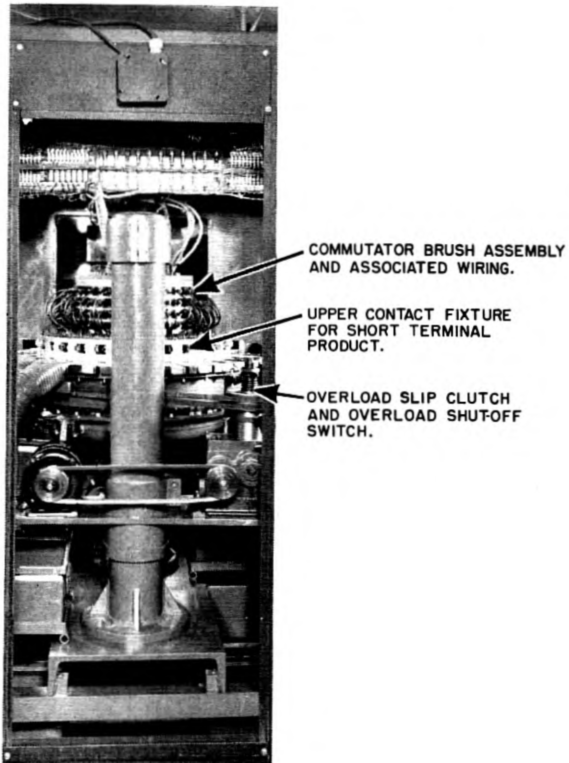


Fig. 4 — View of turret.

The acceptance or rejection of a unit under test at any one of the six test positions depends on whether the test on the unit energizes the "acceptance" solenoid associated with that test position. The gripper hook, which locks the unit under test in the contact fixture, is connected to a release shaft, follower arm, and roller (see Fig. 3). The roller rides in a track in which the plunger of each "acceptance" solenoid lies unless removed by energizing the solenoid from its associated test circuit. In the case of a defective unit, the acceptance solenoid is not energized and the roller in passing over the plunger of the "acceptance" solenoid trips the gripper hook and the spring loaded plunger in the contact fixture ejects the defective unit. Units that pass all tests are ejected on a turntable to the left of the operator from which they are stacked in handling trays by the operator.

The turret assembly includes the test fixtures, the gripper hooks and associated release shaft, follower arm and roller, and the brush assembly

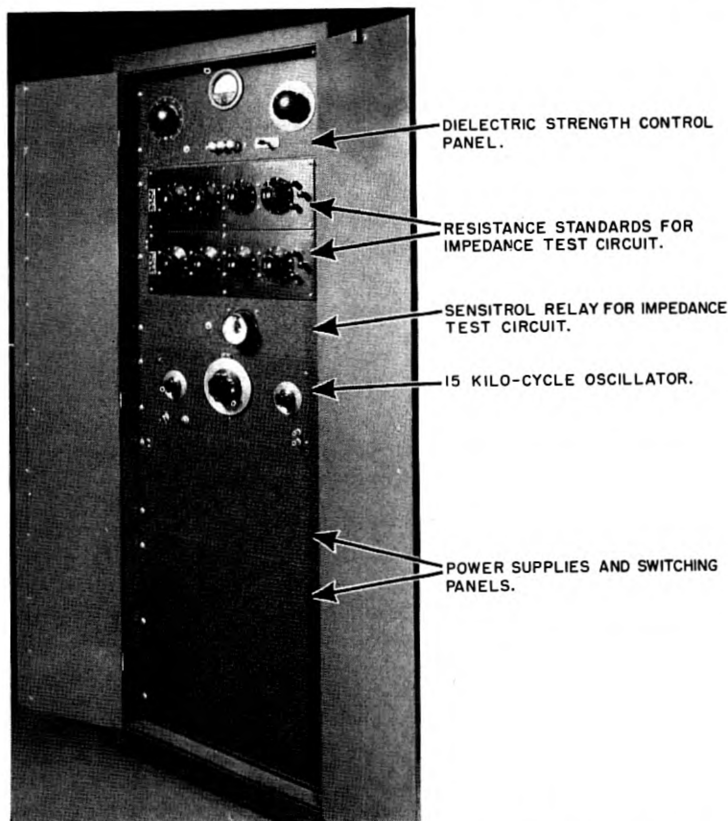


Fig. 5 — Control panels for dielectric strength and impedance tests.

connected to the test fixtures. The commutator is stationary and its segments are connected to the test circuit through permanent wiring. Fig. 4 shows the turret. Each fixture has two sections, one above the other, with the contacts wired in parallel. The lower section is designed for making contact to stud mounted units with long terminals and the upper section for strap mounted units with short terminals. To change the machine "set-up" from one fixture to the other, the turret assembly is raised or lowered by means of the hand wheel, shown on Fig. 2, located at the right of the operator. This feature was incorporated in this machine to facilitate rapid "set-up" which is essential for testing small lots. An overload clutch is incorporated in the driving mechanism to prevent mechanical damage to the machine in case of a "jam".

Fig. 5 shows the control panels for dielectric strength and impedance and Fig. 6 shows the control panels for the capacitance circuits.

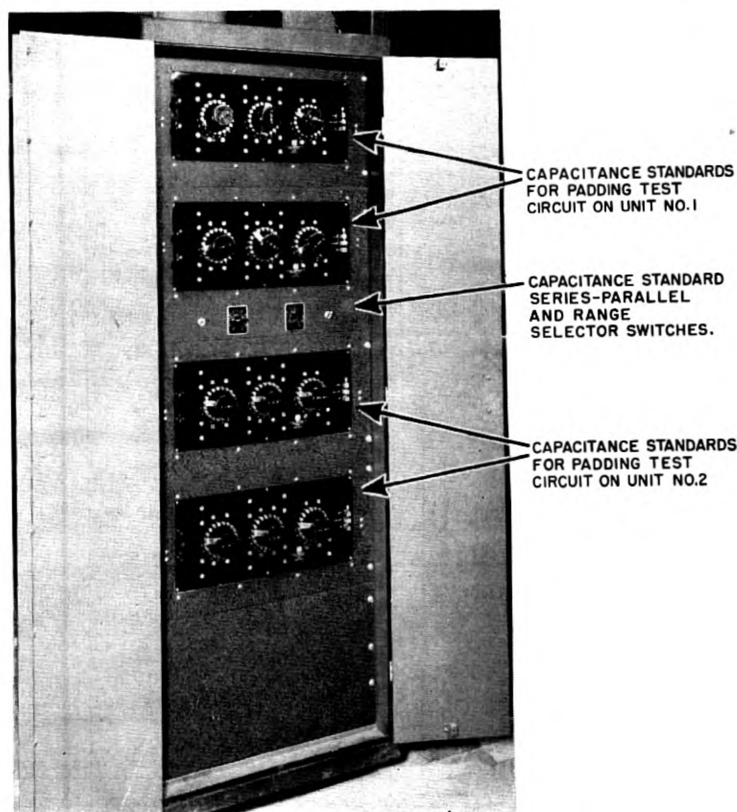


Fig. 6 — Control panels for capacitance circuits.

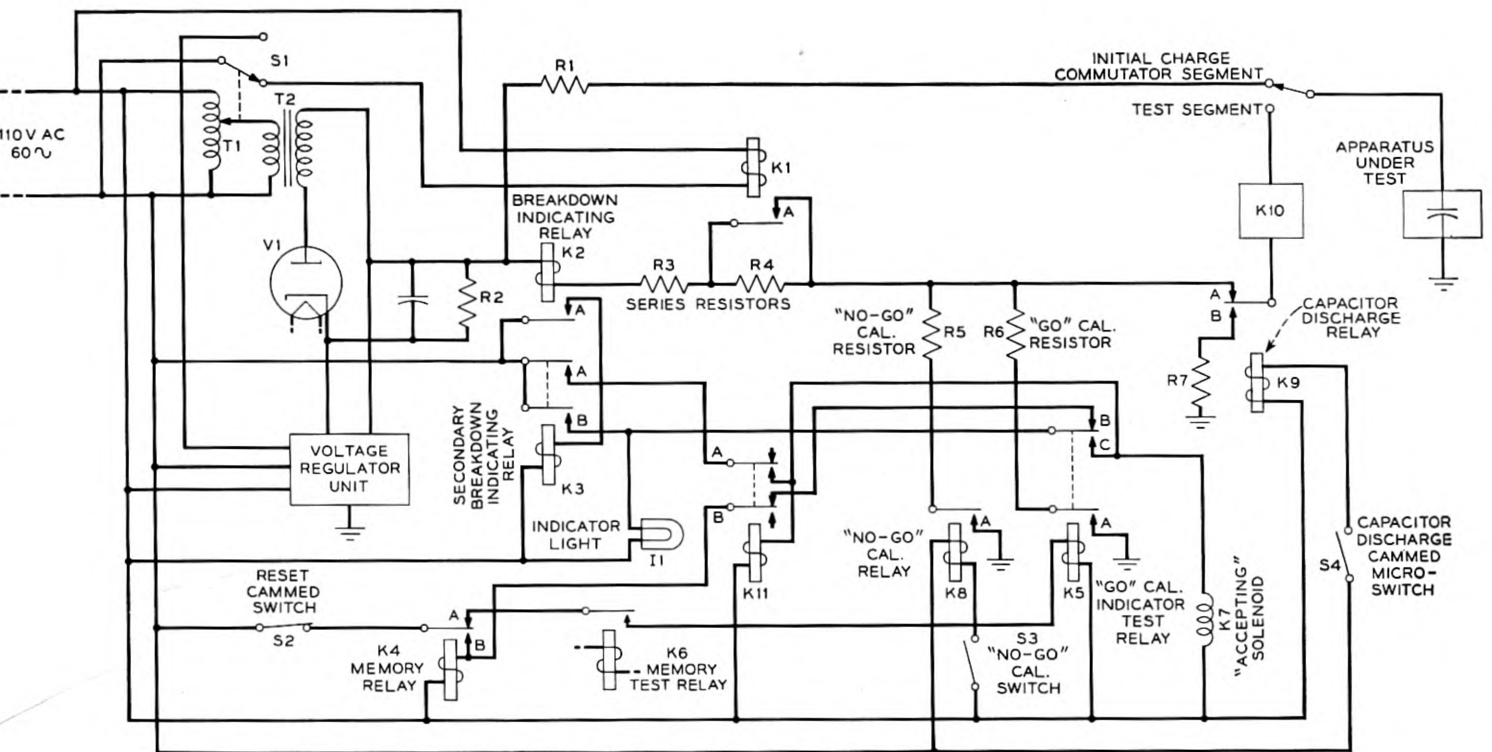


Fig. 7 — Simplified schematic of dielectric strength test circuits.

ELECTRICAL ASPECTS OF TESTING MACHINES

Tests are applied to the product in sequence during one revolution of the turret.

1. Dielectric strength test between terminals and can, and between terminals and studs.
2. Dielectric strength test between units in the same can when the can contains two units.
3. Dielectric strength test between terminals of each unit.
4. Impedance test.
5. Capacitance test.
6. Insulation resistance test.

Dielectric Strength Test Circuit Operation

Since the three dielectric strength tests are made on similar circuits, the operation of one of these circuits is described using the nomenclature and circuit designations shown in Fig. 7. A graphic interpretation of the circuit operations shown in Fig. 7 is given in Fig. 8.

The "heart" of each circuit is a calibrated current sensitive relay K2 that operates on minute values of current resulting when a defective unit under test attempts to charge on the "test" commutator position.

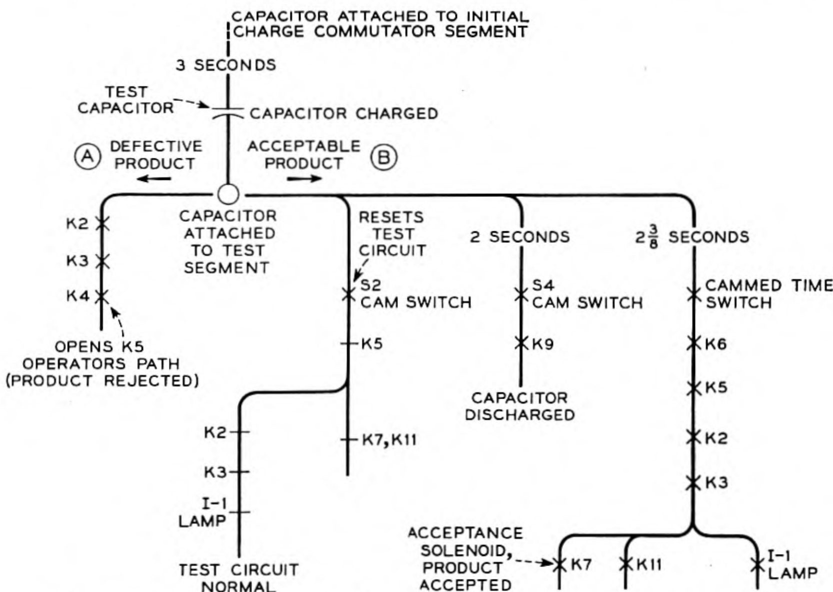


Fig. 8 — Sequence chart for dielectric strength test circuit operation.

Two commutator segments are required to make a dielectric strength test. These segments are known as "initial charge" and "test". After the unit under test has been charged at the test voltage for three seconds on the "initial charge" segment, it passes to the "test" segment in which the unit is again connected to the test voltage through relay K2, current limiting and calibrating resistors R3 and R4 and the contacts on the preset terminal selecting relays K10.

One of the two conditions (under heading A and B below) may be encountered in making this test and the circuit operation for each will be discussed separately.

A. Circuit Operation for Acceptable Product. An acceptable product retains the charge received on the "initial charge" segments and when this unit reaches the "test" segment, no further charging current of a magnitude great enough to operate relay K2 will flow through the unit. Two seconds after the unit under test has been connected to the test segment, a cammed timing switch S4 closes to operate discharge relay K9 to discharge the unit under test to ground through R7. The "self-checking" feature mentioned earlier in this article under "Design Considerations" functions as follows: After the unit under test has been on the "test" segment for approximately 2⅜ seconds, a cammed timing switch (not shown) closes the memory test relay K6 which in turn closes the "go" calibration indicator relay K5 and the "A" contacts on this relay grounds the high voltage test circuit through resistor R6. This resistor is of such a value as to permit sufficient current to operate relay K2. The contacts on relay K2 are not adequate to carry much current, so an auxiliary relay K3 is closed through contacts "A" on relay K2. Contacts "B" on relay K3 closes the indicator light circuit I1 and operates relay K11 and the acceptance solenoid K7. Contacts "A" on the same relay lock relay K11. The circuit is reset for the next unit to be tested by momentarily opening the reset cammed switch S2. Relay K11 was added to the circuit to eliminate a "sneak circuit" that occurred occasionally following the reset when relay K5 opened faster than relay K3. This would result in relay K4 operating to reject the next unit tested. Relay K1 is controlled by switch S1 operated by the manual control T1 on the test voltage power supply. The function of this relay is to add calibrating resistor R4 to the test circuit for voltages above 1,000 volts. Resistor R5, relay K8, and switch S3 control the manual calibrating "No Go" circuit for breakdown indicating relay K2.

B. Circuit Operation for Defective Product. Defective product will not retain the charge it received on the "initial charge" segment and when it reaches the "test" segment, current will flow through the breakdown

indicating relay K2 in an attempt to charge the defective unit, but this current will close relay K2 which in turn closes relay K3. This completes the circuit through the "B" contacts of relays K3, K5, and K11 to close memory relay K4. The closure of relay K4 prevents the memory test relay K6 from closing the "go" calibration indicator relay K5, thereby leaving contact "C" open on relay K5 and no power is applied to the "acceptance solenoid" K7 circuit, which rejects the unit under test.

IMPEDANCE — TEST CIRCUIT OPERATION

The impedance test is made with a 15-kc circuit (see Fig. 9). One arm of the circuit, composed of resistor R12 paralleled by capacitor C5 and the unit under test, is compared with another arm, composed of resistor R11, paralleled by capacitor C4 and either one of two resistance boxes, R13 and R14 respectively, representing maximum and minimum impedance limits. The detector consists of a balanced diode V2 with a 1–0–1 microampere sensitrol relay K24 connected between the diode cathodes. If the impedance of the unit under test falls within the limits for which the resistance boxes were set, the acceptance solenoid will be energized to accept the unit under test. A product outside the preset limits is rejected because the acceptance solenoid is not energized.

The circuit operation is discussed for the following four conditions under A, B, C, and D.

A. Impedance Test on Dual Unit Capacitors

This test is made on capacitors to prevent shipment of resistance-capacitance networks mislabeled as capacitors. Fig. 9 shows dual unit networks connected to the test terminals. Capacitors to be tested are connected to these same terminals. The greater than minimum test cutout relay K18 is preset closed by the switching circuit K23. The cammed memory reset timing switch S14 (normally closed) is opened momentarily to clear relay K19, K20, and K21 at the start of the test.

The sensitrol relay reset switch S16 is cammed shut momentarily to reset the contactor on the sensitrol relay K24. With relay K26 open, the "less than maximum" resistance box R13 is connected to the test circuit. If unit "A" of the dual unit capacitor under test is acceptable product, the contactor on sensitrol relay K24 will close on contact "A", which applies power to close and lock test No. 1 "less than maximum" memory relay K19. Cam operated switch S13 applies power to close relay K26 to connect the "greater than minimum" resistance box R14 into the test circuit. This resistance box is set on zero ohms when capaci-

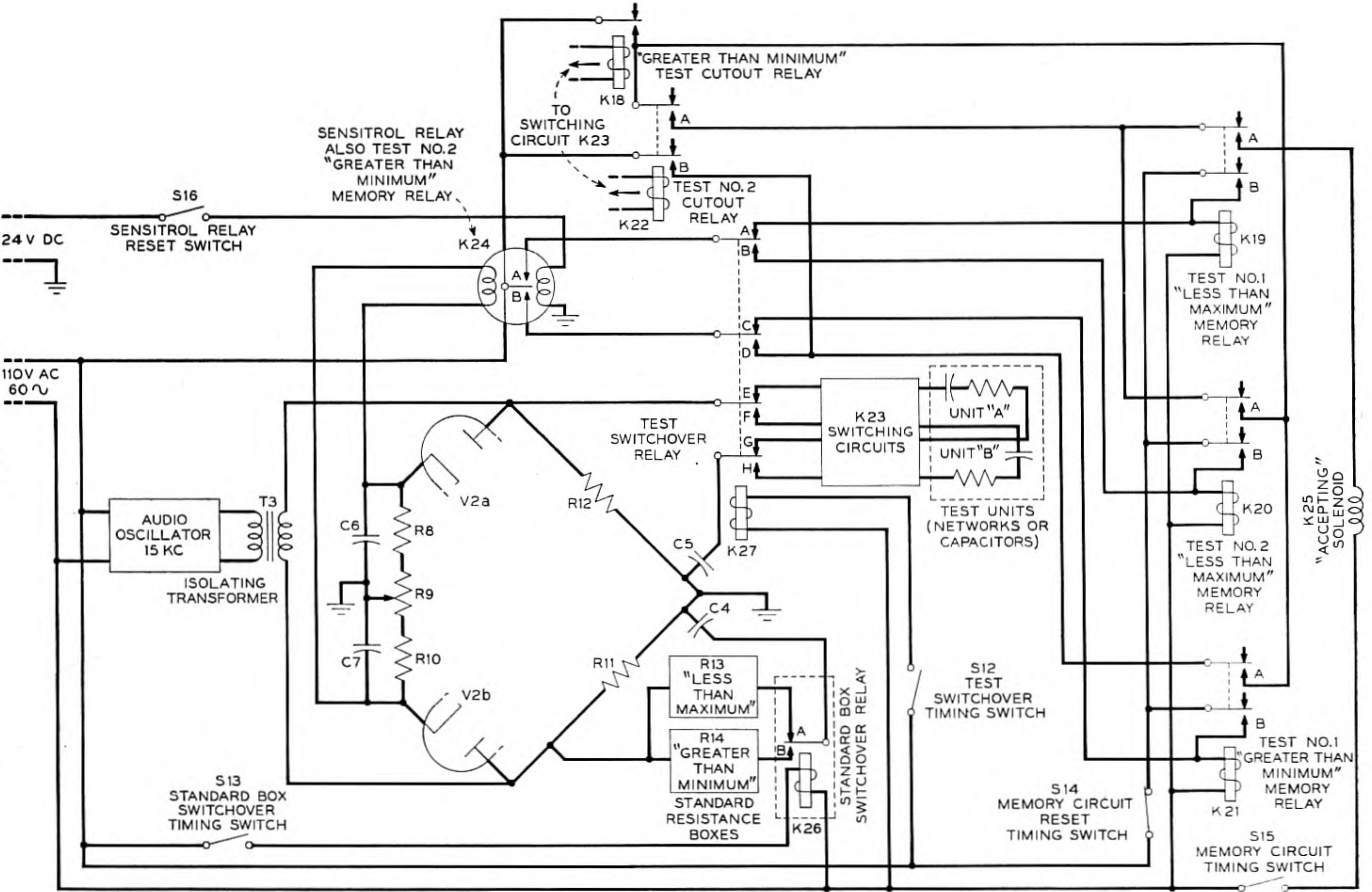


Fig. 9 — Simplified schematic of impedance test circuits.

tors are tested. Sensitrol relay reset switch S16 is cammed shut momentarily to reset sensitrol relay K24, after which the sensitrol relay contactor closes on its "B" contact, thereby applying power to close and lock test No. 1 "greater than minimum" memory relay K21.

Switch S13 is cammed open which opens relay K26 and connects the "less than maximum" resistance box R13 into the test circuit. At the same time switch S12 is cammed shut to close relay K27 which disconnects unit "A" from test and connects unit "B" to the test circuit. Switch S16 is cammed shut momentarily to reset the sensitrol relay contactor. If the unit "B" on test is an acceptable product, the sensitrol relay contactor will close on its "A" contacts and applies power to close and lock test No. 2 "less than maximum" memory relay K20 through contacts "B" of relay K27.

Switch S13 is cammed shut to close relay K26 and connect the "greater than minimum" resistance box R14 into the test circuit. Switch S16 is cammed shut momentarily to reset the sensitrol relay K24 after which its contactor closes on the "B" contact for acceptable product. Memory circuit timing switch S15 is cammed shut and power from one side of the 110 volt ac line flows through the acceptance solenoid, contacts "A" on relay K19, contacts "A" on relay K20, the closed contacts on relay K18 to the other side of the 110-volt ac line to close K25 and to accept the dual unit capacitor under test. The failure of either relay K19 or K20 to operate because of defective product tested opens the acceptance solenoid circuit and rejects the capacitor tested.

B. Impedance Test on Single Unit Capacitors

The impedance test on a single unit capacitor is identical with the testing of dual unit capacitors, except test No. 2 cutout relay K22 is preset closed and test No. 2 "less than maximum" memory relay K20 is not operated since only a single unit is tested.

C. Impedance Test on Dual Unit Networks

The impedance test on dual unit networks is identical with the test for dual unit capacitors, except the "greater than minimum" test cutout relay K18 is not preset closed and the resistance boxes R13 and R14 are set to represent maximum and minimum limits.

D. Impedance Test on Single Unit Networks

The impedance test on single unit networks is identical with the test of dual unit networks except test No. 2 cutout relay K22 is preset closed for the same reason given above for the test of single unit capacitors.

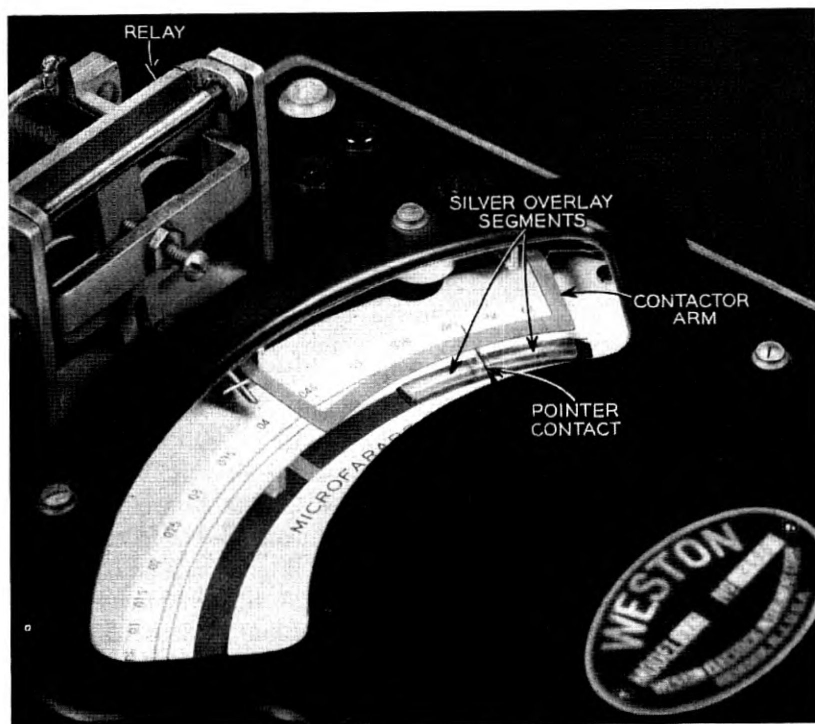


Fig. 10 — Details of modified microfarad meter.

CAPACITANCE TEST CIRCUIT OPERATION

The wide range of capacitance values to be measured, both with and without the series resistance in resistance-capacitance networks, and the one and two-unit construction of the product imposed limitations on the type of capacitance measuring circuits that could be used in this machine. The method selected consists of modified Weston Model 372 microfarad meters that automatically set up external circuits associated with the meters to accept or reject the product as determined by limits preset into the machine.

Two decade capacitance boxes, having a range from 0.001 to 1.0 mf in steps of 0.001 mf are connected in series or parallel with the capacitor on test to make the resultant capacitance fit the range of the meter and control the maximum and minimum limits. This procedure increases the number of capacitor codes that may be tested on a given meter. Capaci-

tors from 0.02 to 5 microfarads are tested on this machine to an accuracy of ±2 per cent.

The modified microfarad meters are equipped with two brass segments, covered with an overlay of silver (Fig. 10). These segments are mounted end to end in a predetermined cutout portion of the meter scale, representing maximum and minimum capacitance conditions. The physical distance between the adjacent ends of these two segments is as small as possible without the two segments touching. A small silver contact is mounted on an insulated portion of the meter pointer, directly above but not touching the segments while the meter pointer traverses its arc of rotation. The armature of the relay, mounted on the meter, actuates a contactor arm which forces the silver contact on the pointer down against the silver overlay segment, thus closing external circuits connected to the segments and contactor.

The testing machine is equipped with three ranges of the special microfarad meters as follows:

1. Suppressed scale from 1.2 to 1.8 mf, with the dividing point between the two segments at 1.60 mf.
2. Suppressed scale from 0.25 to 0.75 mf, with the dividing point between the two segments at 0.63 mf.
3. Suppressed scale from 0.051 to 0.075 mf with the dividing point between the two segments at 0.062 mf.

Two meters for each of the above ranges are necessary in each testing machine, one for each unit in a dual unit. Likewise, four capacitance boxes are necessary, two for each unit in a dual unit.

The discussion that follows, which is divided into two headings, A and B, is a detailed description of the capacitance test circuit. The circuit component designations are those shown in Fig. 11.

A. Capacitance Test on Dual Unit Capacitors or Networks

The cammed switch S5 is opened momentarily at the beginning of the test to restore the test circuit to normal; following this, the cammed switch S8 closes and operates relay K14, which applies power and closes the power supply circuit through the microfarad meters and the capacitor on test.

The capacitance decade box "less than maximum" C2 is shown in series with test capacitor No. 1 by the preset series-parallel switch S10, and in a like manner a capacitance box is connected in series with test capacitor No. 2.

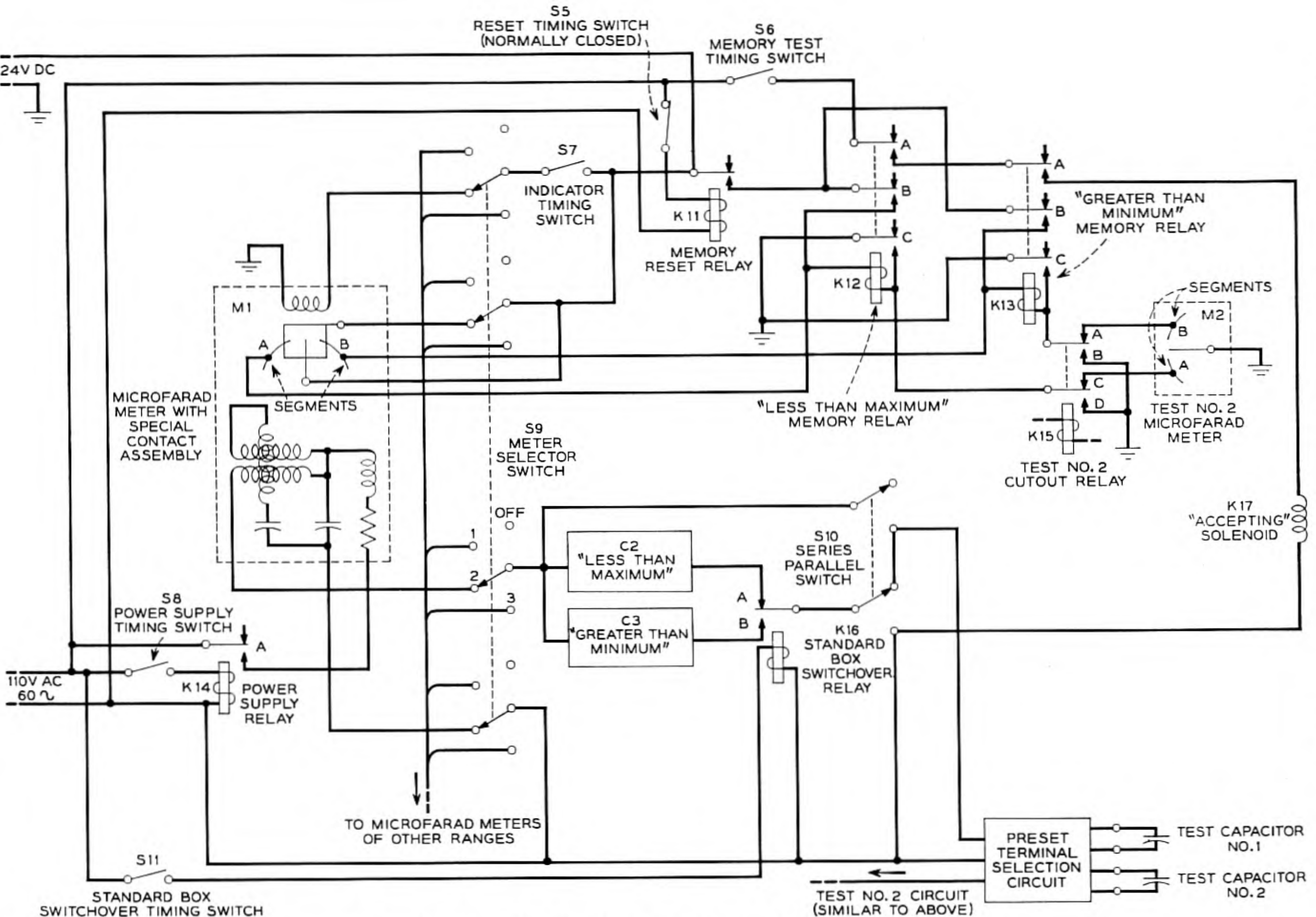


Fig. 11 — Simplified schematic of capacitance test circuits.

Note: The capacitance decade box for Test No. 2 is not shown in Fig. 11. Also, only the segments for M2 Test No. 2 microfarad meter are shown.

If capacitor units No. 1 and No. 2 under test are acceptable product, the pointer on the microfarad meters M1 and M2 will both swing to segment "A". The cammed switch S7 will close and energize the relay on the microfarad meters (not shown on meter M2) which will operate the meter contactor and close the circuit through segments "A" of meters M1 and M2 and apply 24 volts dc to close and lock the "less than maximum" memory relay K12.

The cammed switch S7 is opened to release the meter pointer from segments "A" on M1 and M2. Cammed switch S8 is opened momentarily to release relay K14 which removes the test voltage from the capacitors on test and from meter M1 and M2. During this interval cammed switch S11 is closed to energize relay K16 which connects the "greater than minimum" capacitance box C3 in series with capacitor unit No. 1 on test and meter M1. In a like manner a second "greater than minimum" capacitance decade box (not shown on Fig. 10) is connected in series with capacitor unit No. 2 and microfarad meter No. 2. If the capacitor units No. 1 and No. 2 under test are acceptable product, the microfarad meter pointers will swing to segments "B". The cammed switch S7 will close and energize the relay on the microfarad meters M1 and M2 which will operate the contactor that depresses the M1 and M2 meter pointers against segments B and closes and locks the "greater than minimum" memory relay K13. With relays K12 and K13 closed as described above, the cammed switch S6 is closed which operates the acceptance solenoid K17 through the "A" contacts on relays K12 and K13 to accept the dual unit capacitor under test.

It may be readily observed that in case either or both of the capacitor units on test are out of limits, the circuit will not close either or both relays K12 and K13, which would leave the acceptance solenoid K17 circuit open, and the product would be rejected.

B. Capacitance Test of Single Unit Capacitors or Networks

The capacitance test of single unit capacitors is the same as for dual unit capacitors, except test No. 2 circuit and test No. 2 microfarad meter M2 are not used. Test No. 2 cutout relay K15 is closed to apply ground to its contacts B and D.

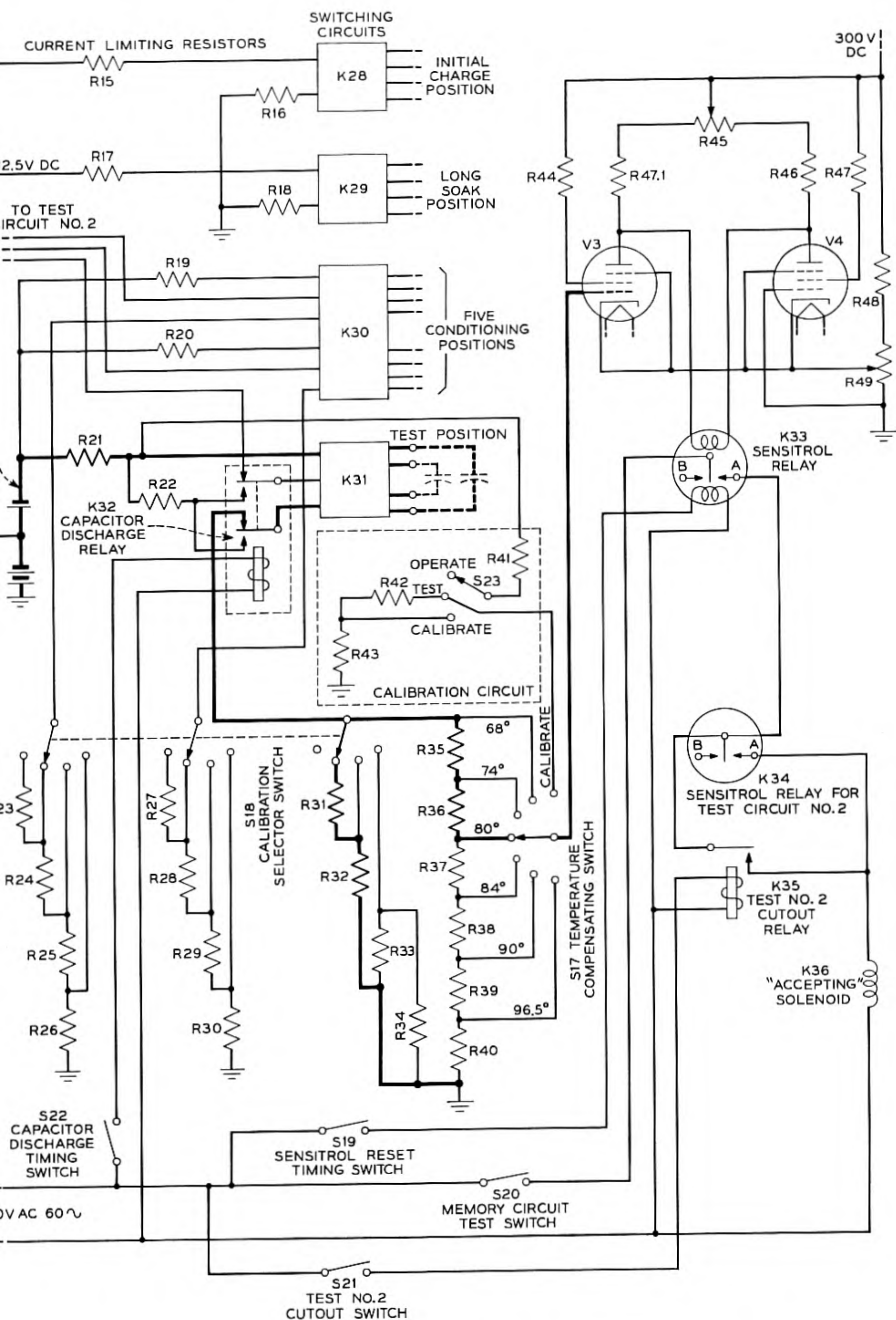


Fig. 12 — Simplified schematic of insulation resistance test circuits.

INSULATION RESISTANCE TEST CIRCUIT OPERATION

In general, the insulation resistance test consists of a charging period and a test period. The charging of the unit under test requires 10 positions or 30 seconds time to insure that the unit is thoroughly charged before it reaches the test position. At the test position the capacitor or network on test is connected to form part of a voltage divider in the grid circuit of a sensitive balanced detector. This sets up relays to accept or reject the unit under test depending on whether the unit meets the limits for which the circuit was preset and calibrated. Two insulation resistance circuits are required, one for each unit in a dual unit capacitor or network. A calibrating circuit is provided by switch S23 and resistors R41, R42, and R43.

The discussion that follows is a detailed description of the sequence of operation of the insulation resistance test circuit. The component designations are those shown on Fig. 12. The discussion is divided into two headings A and B as follows:

A. Insulation Resistance Test on Dual Unit Capacitors or Networks

The capacitor or network on test is automatically connected in succession to the INITIAL CHARGE POSITION, the LONG SOAK POSITION and FIVE CONDITIONING POSITIONS which assures that the acceptable product is thoroughly charged before it reaches the test position. The switching circuits K28, K29, K30, K31 switch S18, and the temperature compensating switch S17 are manual preset switch circuits for the particular code on test.

For the sake of simplicity, the balanced detector and the reset solenoid for sensitrol relay K34 for test circuit No. 2 are not shown. If the insulation resistance of the units on test meets the limits for which the circuit was calibrated and preset, the contactor on K33 and K34 both close on the "A" contacts. Switch S20 is then cammed closed to apply power through the "A" contacts on the sensitrol relays to energize the acceptance solenoid K36 to accept the units on test. At the close of the test, capacitor discharge timing switch S22 is cammed closed, thereby closing the capacitor discharge relay K32 which discharges the units on test before they are ejected as acceptable product. It may be readily observed from the schematic that a unit or units defective for insulation resistance will fail to close either or both of the "A" contacts on the sensitrol relays K33 and K34, which leaves the acceptance solenoid circuit K36 open, thereby rejecting the units tested.

B. Insulation Resistance Test on Single Unit Capacitors or Networks

The insulation resistance test on single unit capacitors or networks is the same as for dual units, except the second test circuit is not required and test No. 2 cutout switch S21 is closed to operate test No. 2 cutout relay K35 which eliminates the second test circuit and its sensitrol relay K34.

CONCLUSION

This machine has been in successful operation on a multishift basis for several years and has proven itself economically. Inspection of the product tested shows that the machine's performance, quality wise, is highly satisfactory. Difficulties that have been encountered were largely those associated with product handling, contact fixtures, etc. Machines of this type that are planned for the future will make use of circuitry developed since this machine was built, but many of the features described will be incorporated.

ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions to the development of this machine of G. E. Weeks of the Western Electric Company S. V. Smith and S. E. Frisbee of the Electric Eye Company.

A 60-Foot Diameter Parabolic Antenna for Propagation Studies*

By A. B. CRAWFORD, H. T. FRIIS and W. C. JAKES, JR.

(Manuscript received February 2, 1956)

A solid-surface parabolic antenna, sixty feet in diameter and of aluminum construction, has been erected on a hilltop near Holmdel, New Jersey. This antenna can be steered in azimuth and elevation and was specially designed for studies of beyond-the-horizon radio propagation at frequencies of 460 mc and 4,000 mc.

The electrical properties of the antenna and the technique of measurement are described; construction and mechanical details are discussed briefly.

INTRODUCTION

Studies in recent years have demonstrated that transmission of useful amounts of microwave energy is possible at distances considerably farther than the horizon.[1] The exact mechanism responsible is not as yet completely understood, although scattering by atmospheric irregularities seems to play a significant part. A program to study the nature of these effects has been started at the Holmdel Laboratory. An important and necessary tool for this work is a steerable antenna having unusually high gain and narrow beam width. Such an antenna has been built, and it is the purpose of this paper to describe its design and the methods used to measure its radiation properties.

DESCRIPTION OF THE ANTENNA

The antenna is a 60-foot diameter paraboloid made up of forty-eight radial sectors, each constructed of sheet aluminum. Each sector is held to the correct doubly-curved surface by reinforcing ribs, and all are fastened to a central hub eight feet long and thirty inches in diameter. During assembly, the axis of the paraboloid was vertical; thus no scaf-

* This work was supported in part by Contract AF 18(600)-572 with the U.S. Air Force, Air Research and Development Command.

[1] Proc. I.R.E., October, 1955, contains many papers by workers in this field.

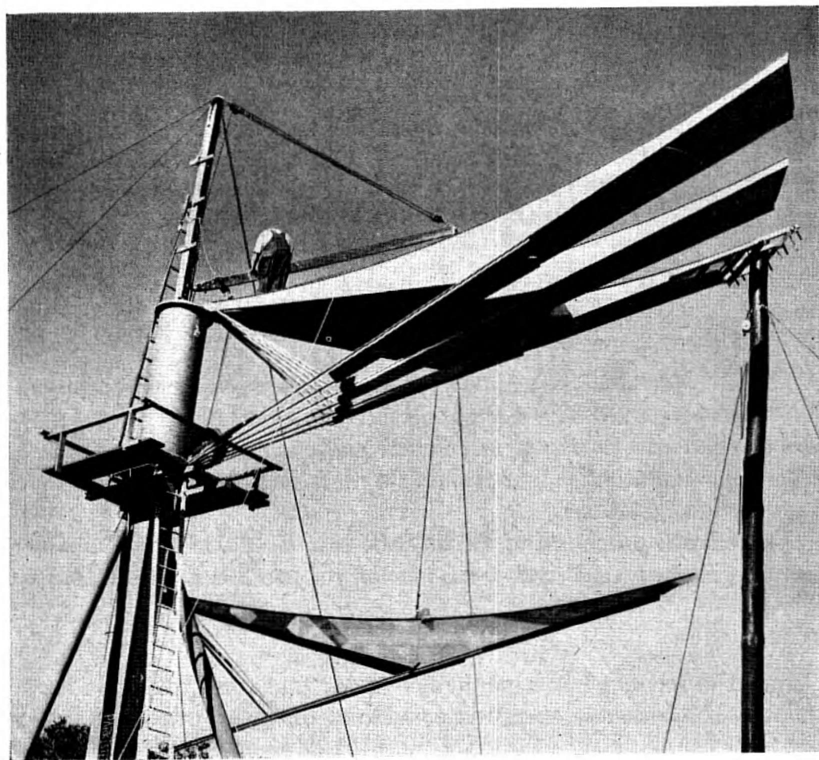


Fig. 1 — Fastening the radial sectors to the hub.

folding was required. Figs. 1 to 5 illustrate the paraboloid construction and support. The weight of the antenna is carried on a vertical column which is mounted on bearings to permit movement in azimuth. The column is held upright by a tripod structure. The central hub of the paraboloid is fastened to a steel girder which extends to the rear along the paraboloid axis and is pinned to a yoke carried by the vertical column, thus permitting movement in elevation. The antenna is scanned by two motors mounted on an A-frame and connected to the end of the axial girder by crank mechanisms. The total scanning range of the antenna is about 3° in both azimuth and elevation.

The antenna is designed for use at frequencies of 460 mc and 4,000 mc. The tolerance on the parabolic reflecting surface is set by the higher frequency and thus must be $\pm\frac{3}{16}$ inch to meet the usual $\pm\lambda/_{16}$ criteria. The focal length is 25 feet, so that the total angle intercepted by the paraboloid as seen from the focal point is 124°. Design of a feed horn for

this angle so that the illumination is tapered to −10 db at the edge of the paraboloid is not difficult; the horn used is diagramed in Fig. 6, with dimensions given in wave-lengths. The feed horn is mounted in a tripod support extending out from the front surface of the paraboloid. It is made strong enough so that two 460 mc horns can be mounted side by side.

The paraboloid itself weighs approximately 5½ tons; the frontal wind load for a 100 mph wind is about 40 tons. It is expected that winds of this force will be withstood.

The antenna is mounted atop Crawford Hill near Holmdel, New Jersey, at an altitude of 370 feet. It is aimed towards Pharsalia, New York, a distance of about 171 miles.

MEASUREMENT TECHNIQUE

The two important properties of the antenna which had to be determined before it could be put into use were its gain and radiation pattern at the operating frequencies of 460 mc and 4 kmc. It was also hoped to

Fig. 2 — Assembling the sectors.

Fig. 3 — The completed antenna.

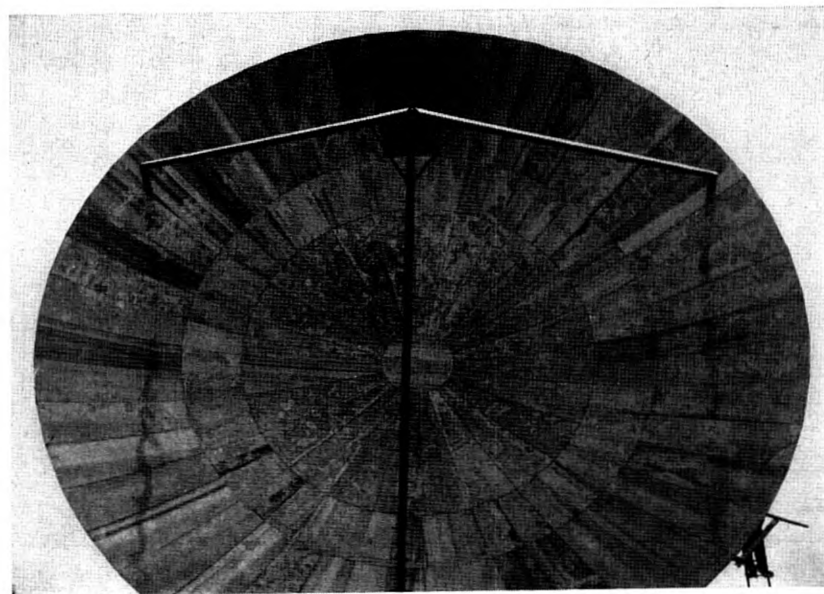


Fig. 4 — Front view of the paraboloid.

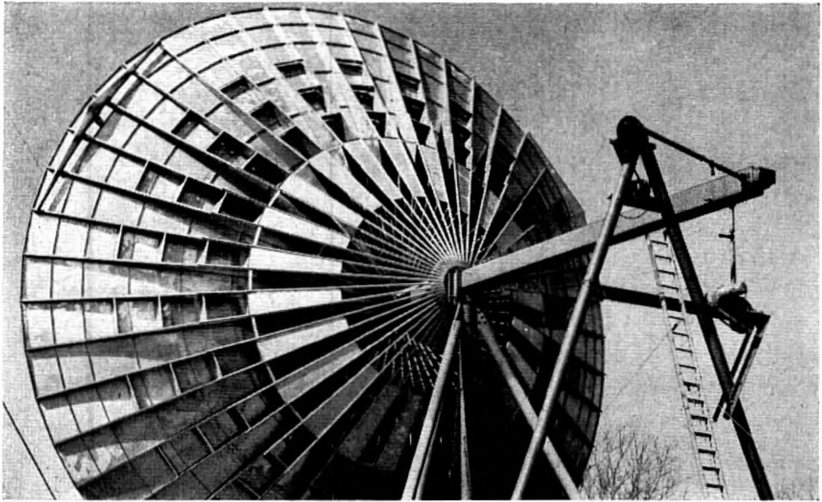


Fig. 5 — Antenna scanning motors.

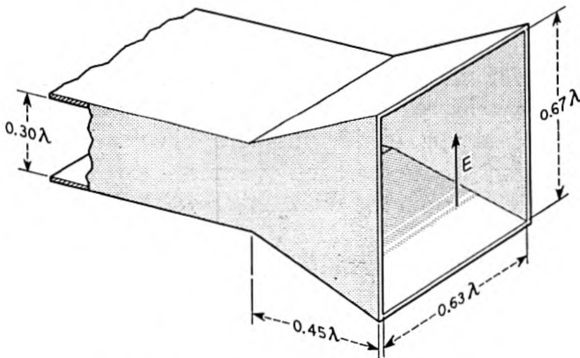


Fig. 6 — Feed horn dimensions.

measure these properties at 9.4 kmc to get some idea of how good the mechanical tolerances actually are.

The first requisite for making antenna measurements is a sufficiently uniform incident field. The source producing this field must be located at a distance of at least $2b^2/\lambda$, (b is the paraboloid diameter), which means a distance of about 0.6 mile at 460 mc, six miles at 4 kmc, and thirteen miles at 9.4 kmc. An obvious and convenient place for the sources was at Murray Hill, 22.8 miles away, which is on the transmission path to Pharsalia. Having located the sources at a suitable distance it was then necessary to test the incident field for uniformity. A 64-foot

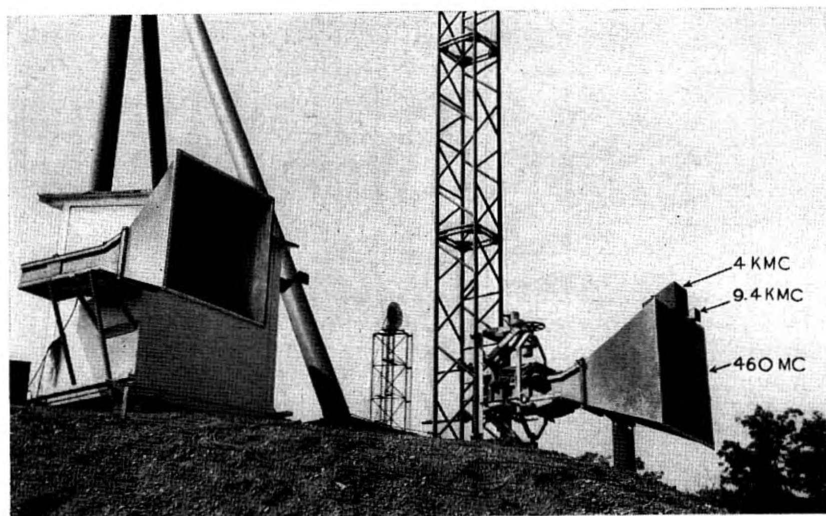


Fig. 7 — Height run tower with the three standard horns during preliminary studies of the incident field.

tower was used for this purpose, and the variation of the incident field with height was measured before the antenna was erected. Figs. 7 and 8 show a typical set up. Height runs were taken at intervals of 15 feet along a line normal to the direction of transmission in the plane which would eventually contain the antenna aperture. The results of these tests showed that the Murray Hill location was satisfactory for the 4 and 9.4 kmc sources, with ground reflections giving rise to ± 1 db variations with height. In each case several complete cycles occurred in the 60-foot height run so that an average signal level could be established with an accuracy of a few tenths of db.

However, at 460 mc the variation with height was about 5 db, and only a portion of one cycle was available, so that the average signal could not be determined. The solution was to bring the source to a location as close as possible to the effective ground reflecting surface. Such a location was found at the far edge of a large body of water lying in the path, and the source antenna was placed in a mobile truck 10 feet above the water and eight miles away. The resulting variation with height was now only about 1 db.

In all cases the variation of field at right angles to the direction of transmission was found to be no worse than ± 1 db; thus it was felt that suitable sources for test at all three frequencies were now ready.

The standard method of measuring the gain of a microwave antenna is to compare the signal received from the antenna to that from another antenna whose gain is accurately known. A pyramidal horn of about 20 db gain is usually used as the standard. Such horns are readily available at 4 kmc and 9.4 kmc, and, in principle, also at 460 mc. Under the present set up, however, the physical dimensions of the standard horn were limited by the necessity of raising the horn on a carriage attached to the 64-foot tower. The largest horn that could be so mounted had an aperture of 4 feet × 4 feet, or 1.8λ on a side at 460 mc. Since the gain of a horn of this small aperture size cannot be accurately calculated by the usual formulas a scale model was made and tested at 4 kmc. The result of this test showed that the actual horn gain was 15.05 db, which is about 0.4 db more than the calculated gain.

A typical gain measurement on the 60-foot paraboloid was thus made as follows:

1. The feed position and antenna orientation were adjusted to obtain maximum received signal level.

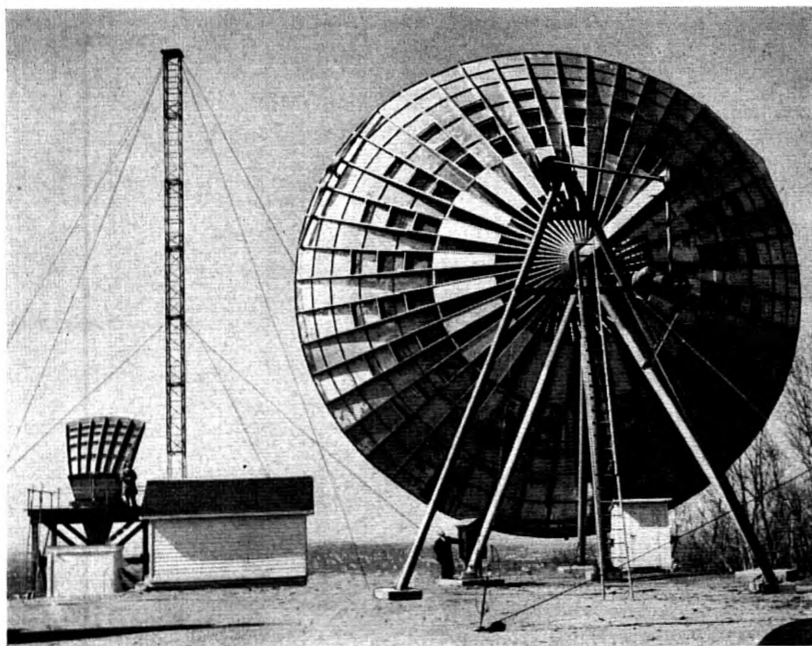


Fig. 8 — Position of height run tower during gain measurements.

2. The average incident field was determined by a height run with a standard horn.

3. The decibel gain of the antenna was then calculated by adding the db gain of the standard horn to the db difference in the signal levels determined in (1) and (2).

The problem of adjusting the 60-foot antenna for maximum received signal at 4 kmc and 9.4 kmc was complicated by the scintillations of the

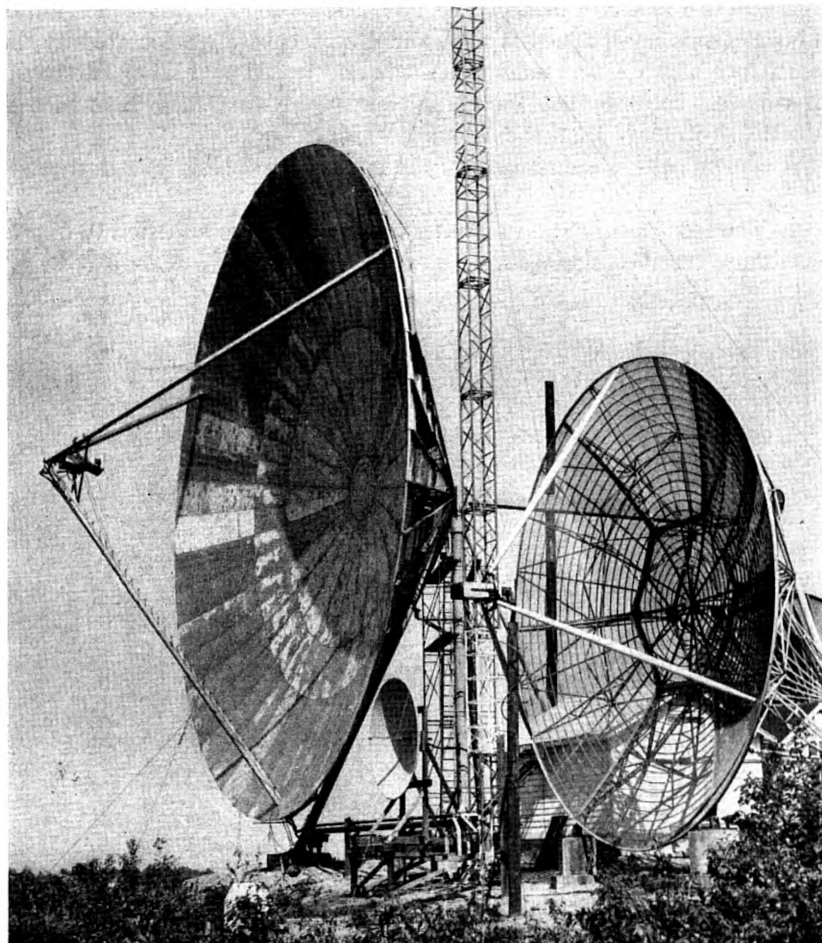


Fig. 9 — A view of the antennas at Crawford Hill used for beyond-the-horizon propagation studies and showing the 60-foot, a 28-foot and an 8-foot paraboliod, the latter between the two larger ones.

TABLE I

Frequency	Area Gain,* db	Gain, db Meas.	Ratio of Effective Area to Actual Area	3 db beam width		1st Minima	1st Minor lobes
				Calc.	Meas.		
460 mc	38.90	37.0 ± 0.1	0.65	2.35°	2.45°	—	—
3.89 kmc	57.44	54.6 ± 0.2	0.52	0.28°	0.3°	−33db	−23db
9.40 kmc	65.12	61.1 ± 0.5	0.40	0.12°	0.14°	−25db	−18db

* The area gain is defined as 10 log $\frac{4\pi}{\lambda^2}$, where A is the paraboloid projected area, 2,830 square feet.

incident field at these frequencies due to the remote location of the source. Accordingly, instead of adjusting the feed position for maximum signal level, it was adjusted to give vertical and horizontal radiation patterns having the best possible symmetry, deepest minima, and lowest minor lobes. It was then assumed that this was also the point of maximum gain. At 460 mc the scintillations were so small that the conventional technique of adjusting for maximum output was effective.

A double detection receiver was used for making all measurements. Signal level decibel differences were established by an attenuator in the intermediate frequency (65 mc) channel, and could be determined to an accuracy of ±0.02 db.

RESULTS

Carrying out the measuring procedure described above the results given in Table I were obtained. At 460 mc the restricted scanning range did not permit inspection of the minor lobes.

CONCLUDING REMARKS

The overall performance of this antenna is considered to be excellent. In general the radiation patterns are clean with satisfactory minor lobe structure. The good performance at 9.4 kmc (61 db gain) is particularly gratifying, since the mechanical tolerance of ±3/16 inch is equivalent to $\pm\lambda/7$ at this frequency.

As stated earlier, this antenna was designed to provide a research tool for propagation studies and thus has some features which are neither necessary nor desirable in an antenna intended primarily for communication use. A consideration of the problem of providing a sturdy 60-foot

antenna for fixed point-to-point service led to the square "bill-board" design* and antennas of this type are now in production.

ACKNOWLEDGEMENTS

The construction of the antenna described in this paper was carried out under the general direction of H. W. Anderson, Supervisor of the Holmdel Shops Department. The paraboloid was assembled in place by members of the Carpenter Shop supervised by C. P. Clausen. Daniel Beaton, of Lorimer and Rose, served in an advisory capacity on some features of the construction. Assistance in the design and testing of the antenna was given by many members of the technical staff.

---

* A picture and short description of this antenna appeared in Bell Laboratories Record, **34,** p. 37, Jan., 1956.

The Use of an Interference Microscope for Measurement of Extremely Thin Surface Layers

By. W. L. BOND and F. M. SMITS

(Manuscript received March 15, 1956)

A method is given for the thickness measurement of p-type or n-type surface layers on semiconductors. This method requires the use of samples with optically flat and reflecting surfaces. The surface is lapped at a small angle in order to expose the p-n junction. After detecting and marking the p-n junction, the thickness is measured by an interference microscope. Another application of the equipment is the measurement of steps in a surface. The thickness range measurable is from 5×10^{-6} cm to 10^{-3} cm.

INTRODUCTION

Extremely thin p-type or n-type surface layers can be obtained on semiconductors by recently developed diffusion techniques.[1, 2] Layer thicknesses of the order of 10^{-4} cm are currently used for making diffused base transistors.[3, 4] The thickness of the diffused layer is an important parameter for the evaluation of such transistors. Its measurement is facilitated by lapping a bevel on the original surface, thus exposing the p-n junction within the bevel where the thickness appears in an enlarged scale. With a sharp and well defined angle, one would obtain the thickness by the measurement of the angle and of the distance between the vertex and the p-n junction.

However, it is extremely difficult to obtain vertices sharp enough for measurements of thicknesses of the order of 10^{-4} cm. To avoid this difficulty, an interferometric method was developed in which the depth is measured directly by counting interference fringes of monochromatic light. The method can also be used for the measurement of small steps

[1] C. S. Fuller, Phys. Rev., **86**, p. 136, 1952.
[2] J. S. Saby and W. C. Dunlap, Jr., Phys. Rev., **90**, p. 630, 1953.
[3] C. A. Lee, B.S.T.J., **35**, p. 23, 1956.
[4] M. Tanenbaum and D. E. Thomas, B.S.T.J., **35**, p. 1, 1956.

and similar problems occurring, for example, in the evaluation of controlled etching and of evaporated layers.

PRINCIPLE[5]

A half-silvered mirror is brought into contact with a reflecting surface. If this combination is illuminated with monochromatic light, one observes interference fringes. Dark lines appear where the distance between mirror and surface is $n \times \lambda/2$, where n is an integer. Between two points on adjacent fringes the difference in this distance is therefore $\lambda/2$. Hence the fringes can be regarded as contour lines for the distance between the mirror and the surface under consideration. Since the mirror is optically flat, one can deduce the profile of the surface. Equidistant and parallel fringes, for example, prove the surface to be flat. By taking the profile across a bevel or a step, one is able to measure the depth of one part of the surface with respect to another optically flat part of the surface. The reflectivity of the crystal surface should be as high as possible, and that of the mirror should be of the same order. The fringes are then produced by the interference of several wave trains which make the fringes very sharp, and one can measure fractions of $\lambda/2$. With the equipment described here, one is able to interpolate to $1/10$ of $\lambda/2$ or less.

Since small linear dimensions are involved, this principle was adapted for use under a microscope. Hence, it is possible to measure small linear dimensions and the correlated depth simultaneously.

The measurement of small steps, or steps not too steep in an otherwise flat surface, can be done without altering the sample. For measurement on steep and high steps a bevel must be lapped on the sample.

For the measurement of p-type or n-type surface layers on semiconductors, it is essential to lap a bevel on the original surface. The p-n junction is thus exposed and can be found by an electrical method. After marking its position within the bevel, it is then possible to measure its depth with respect to the original surface by taking the profile across the bevel. The marking has to be such that it will be visible in the fringe pattern. By a proper adjustment of the optical flat, a fringe pattern can be produced in which the profile is easily interpreted and the depth measurement amounts to a counting of fringes.

PREPARATION OF THE SAMPLES

The method requires the use of samples with optically flat and highly reflecting surfaces with respect to which a depth can be measured. It is

[5] S. Tolansky, Multiple-Beam Interferometry of Surfaces and Films, Oxford, at at the Clarendon Press, 1948.

also advisable to use plane-parallel samples to facilitate the lapping of a bevel at a small angle.

For lapping. the sample is waxed with its back side to the face of a short steel cylinder. The face is cut at a small angle. Angles of 0.5° or 1.0° are practical. The cylinder is placed in a jig, in such a position that approximately half of the sample surface projects above the plane of the jig (Fig. 1). A short grind on a slightly rough glass plate using a fine abrasive with water gives usable bevels. For a shiny finish just the right degree of roughness of the glass is important. The use of a lapping machine with a vulcanized fiber plate and fine abrasive gives a better surface finish, but the ridge is not as sharp. A 0.5° bevel could be obtained only on a glass plate.

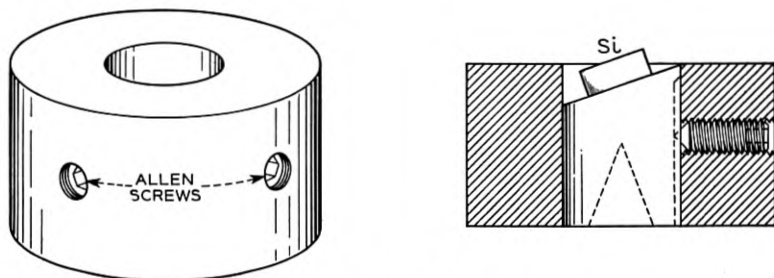


Fig. 1 — Jig for lapping a bevel.

MARKING OF p-TYPE OR n-TYPE SURFACE LAYERS

In a sample with a p-type or n-type surface layer the junction is exposed within the bevel. The next step is to detect and mark the junction.

The sample is fixed on a microscope stage which allows a micrometer controlled movement in two directions (Fig. 2 shows a Wilder micrometer cross slide). The sample is oriented in such a way that the ridge is parallel to one direction of movement (y-direction). One or two lines of aquadag are applied to the surface of the sample, perpendicular to the ridge. The aquadag should be diluted with water in such a proportion as to achieve a thin film which is non reflecting.

A needle is fixed to the base of the stage with a suitable linkage leaving a vertical degree of freedom. The needle is brought into contact with the surface of the sample outside the aquadag. Thus, the sample can be moved under the needle while the needle maintains contact. In a suitable electrical circuit, the needle serves as detector of the junction. The sample is moved in the direction perpendicular to the ridge (x-direction)

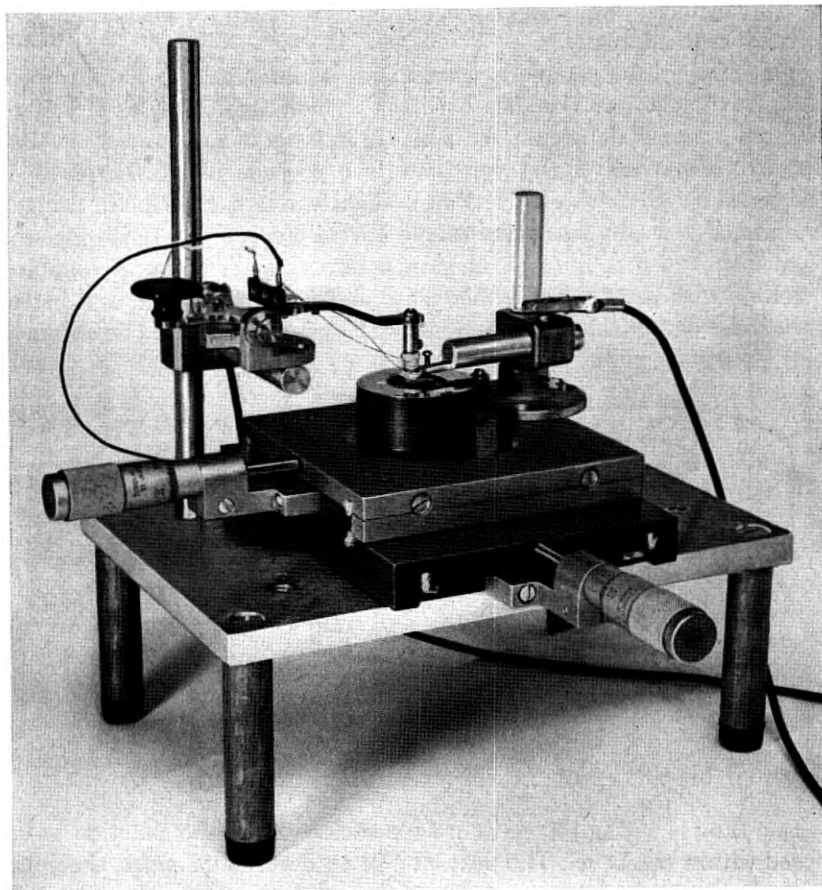


Fig. 2 — Apparatus for locating and marking p-n junctions.

until the needle rests on the p-n junction as seen by the electrical detector. By moving the sample in the y-direction the same needle scrapes a line through the aquadag. In this line, the reflecting sample surface is bared and thus, a reflecting line is produced within a non-reflecting surrounding and can be seen in the fringe pattern.

If the ridge of the sample is exactly lined up with the y-direction, the needle moves along the junction and the line in the aquadag indicates the position of the junction exactly. To minimize an error due to poor alignment, it is advisable to locate the junction close to the edge of the aquadag. By doing this on two different sides of the coating, the average

of both readings compensates the error. To obtain, however, the maximum of accuracy, one can locate the junction at any point. (See Fig. 3, Point A). Moving the sample in the y-direction scribes a line through the point at which the junction was found. A movement in the x-direction with the needle in the aquadag marks a point B on this line. The distance from this point to the junction can be obtained from the readings on the micrometer. Thus, the exact point at which the junction was located can be reproduced under the microscope.

DETECTION OF THE p-n JUNCTION

1. Thermoelectric Probe

The thermoelectric voltage[6] occurring between a hot and a cold contact to the sample, changes sign by crossing the junction with the hot contact. The advantage of this probe is that it does not depend upon the rectification properties of a p-n junction. The thermoelectric probe is most suitable for germanium since lapping across a p-n junction normally produces a "short" between the two regions. However, it is likely to give a p-reading on lightly doped n-material. It is therefore only usable on heavily doped layers, where the nearly compensated zone is very small. In the case of silicon, the junction normally maintains rectifying properties after lapping; thus, a photocurrent is present. This current is superimposed upon the thermocurrent. Therefore, the thermoelectric probe is only usable in the dark. The photoelectric method (see below) is more convenient for these cases.

The thermoelectric probe used, consisted of a commercial phonograph needle, which had a good hemispherical point and was surrounded by a piece of ceramic tubing carrying a heating coil. Between needle and sam-

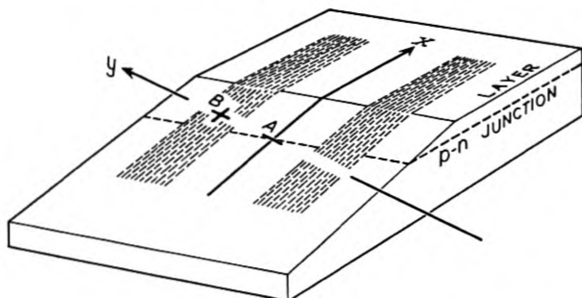


Fig. 3 — Schematic view of a scribed *p-n* junction.

[6] V. A. Johnson and K. Lark-Horowitz, Phys. Rev., **69**, p. 259, 1946.

ple a sensitive galvanometer is connected. It is unimportant whether the contact is made to the p-type or to the n-type side of the sample. The best results are obtained on freshly lapped and clean surfaces. It is, therefore, advisable to keep the sample on the steel cylinder while applying the thermoelectric probe.

When applying the probe, the sample is moved in the x-direction to the point of zero deflection on the galvanometer, whereby the point rests on the junction.

2. Point Rectification on the Surface

This test is also usable on p-n junctions which are not rectifying. With one fixed ohmic contact to the sample, the point rectification of the movable needle can be displayed on an oscilloscope. By crossing the junction with the needle, the characteristic changes from p-n to n-p. Thus, the needle again can be placed on the junction.

This test was applied on lightly doped Ge-layers. The oscilloscope pattern is not very definite, since on a lapped surface the point rectification is poor. However, with some experience the junction can be located. It is advisable to repeat the measurements several times. Boiling the sample in water before applying the probe improves the surface.

3. Photoelectric Probe

This method requires that the junction exhibit rectifying properties. It is most successfully applied to silicon. Between the needle and a contact to either the p-type side or the n-type side of the sample, a high impedance voltmeter is connected. While the sample is strongly illuminated, it is moved in the x-direction. When the needle crosses the junc-

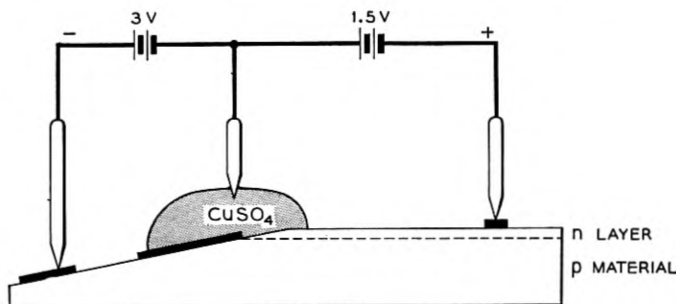


Fig. 4 — Arrangement for Cu-plating the p-type side of a p-n junction.

tion, a change in the photoelectric voltage occurs. For more careful measurements one might plot the photovoltage versus the x-coordinate in units of the micrometer. Such a plot allows an accurate location of the junction in these units. If the micrometer is set for this reading, the needle will rest on the p-n junction.

4. Potential Probe

This is another method for locating the junction where the junction is at least slightly rectifying. One needs two contacts to the sample, one on the p-type side and the other on the n-type side. When a current is passed through the sample in the reverse direction, the voltage between the needle and either contact shows a discontinuity at the junction. The voltage can be plotted in a similar way as described for the previous method, and thus the needle can be set on the p-n junction.

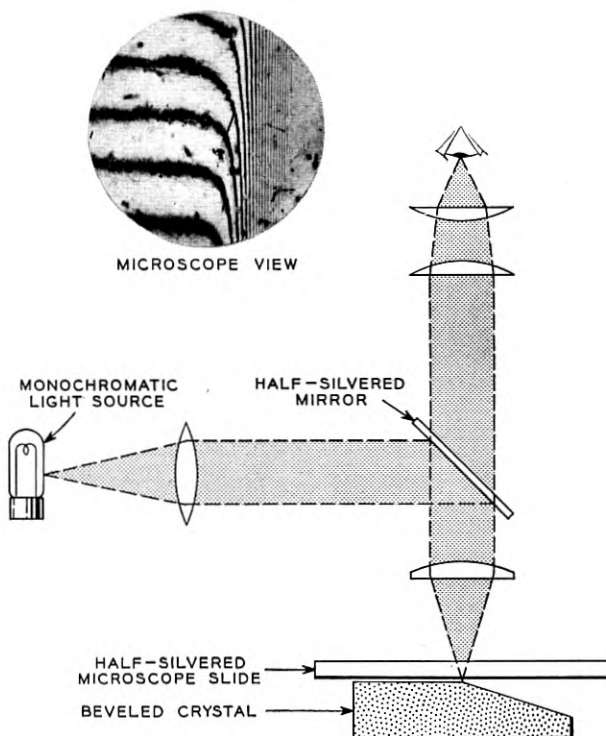


Fig. 5 — Diagrammatic view of the light path in the interferometer.

*5. Plating the p-side of the p-n Junction**

This method detects and marks the junction in one process without using the micrometer arrangement. Voltages are applied in such a way that only the p-type side is plated (See Fig. 4). Since the plating projects up and is not optically flat, it can be recognized under the interferometer. It has the advantage of showing the junction as a line. The disadvantage is that it is only convenient on rectifying p-n junctions (silicon), with n-type layers since the plating ought to take place on the body side of the p-n junction.

THE INTERFEROMETER

The main part of the interferometer is a microscope with illumination through the objective. As a source of monochromatic light, a sodium lamp for which $\lambda = 5.89 \times 10^{-5}$ cm is most convenient. The use of a

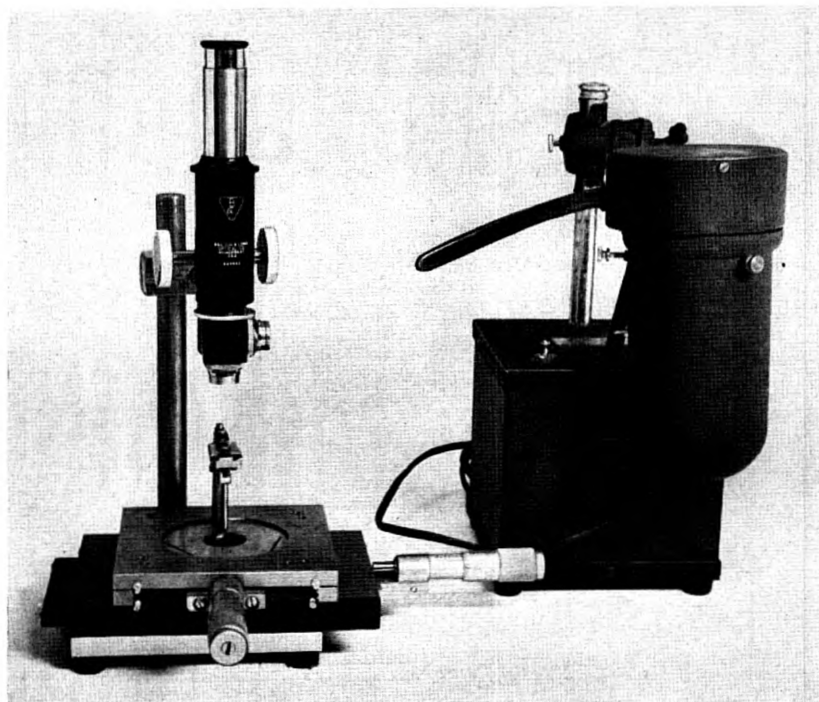


Fig. 6 — Interferometer with light source.

* This method was developed by N. Holonyak.

shorter λ would increase the resolution. However, a sodium lamp gives enough light that one can easily work in daylight.

The microscope is mounted above a micrometer cross-slide of the same kind as used in the procedure for marking the p-n junctions. The stage carries a special sample support. Fig. 5 gives a diagrammatic view of the light path in the interferometer. A normal microscope is used with an attachment carrying a semi-transparent mirror. Fig. 6 shows a photo of the complete arrangement, and Fig. 7 gives the details of the sample support.

The prepared sample is waxed to a microscope slide and covered by a half-silvered mirror. Both are placed on the adjustable lower jaw of the sample support. The lower jaw is raised so that the upper jaw presses against the mirror. In this position it is fixed by tightening the screw in the back. Thus the mirror and sample are in contact, and the fringes can be observed through the microscope. Three screws in the lower jaw make it possible to change the relative position of mirror and sample. Thus the fringe pattern can be adjusted to make it most suitable for the particular case.

THE MEASUREMENTS

The measurement of a layer thickness was chosen to demonstrate the principle of evaluating a fringe pattern. (See Fig. 8.) The first illustration

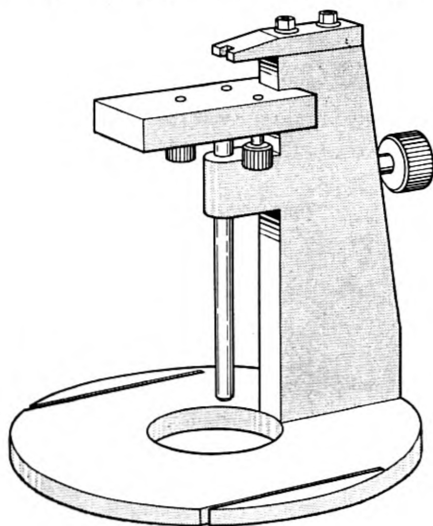


Fig. 7 — Sample support.

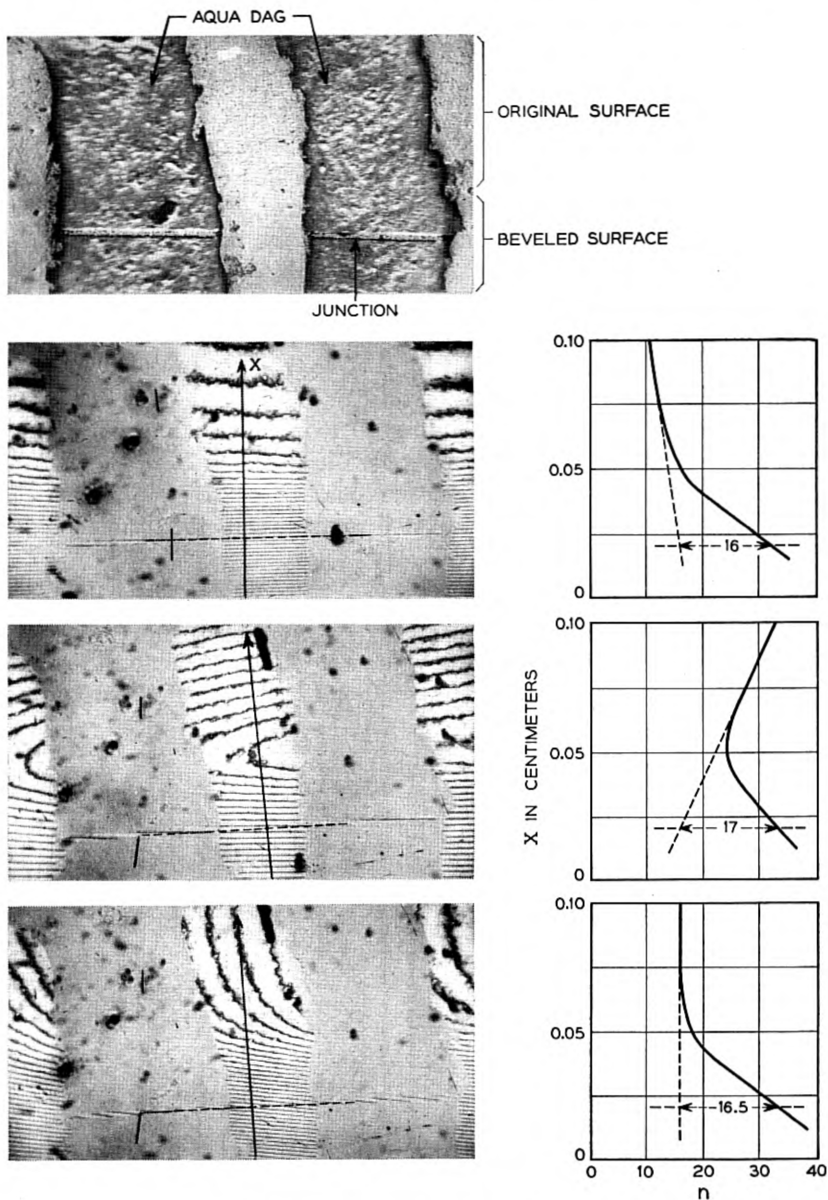


Fig. 8 — Evaluation of the interference fringe pattern on a scribed p-n junction.

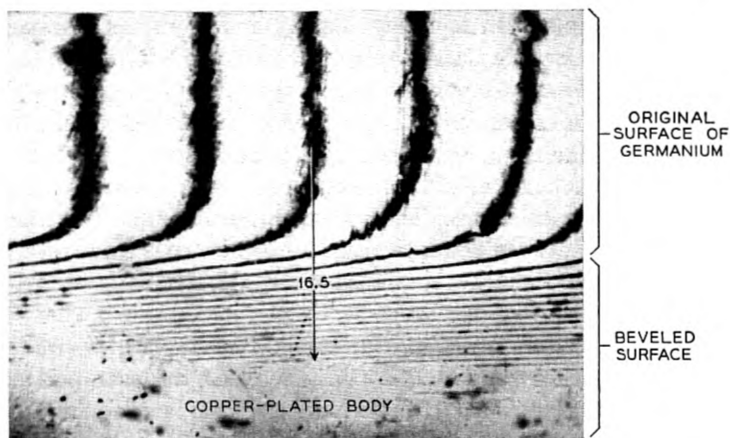


Fig. 9 — Evaluation of the interference fringe pattern on a Cu-plated p-n junction.

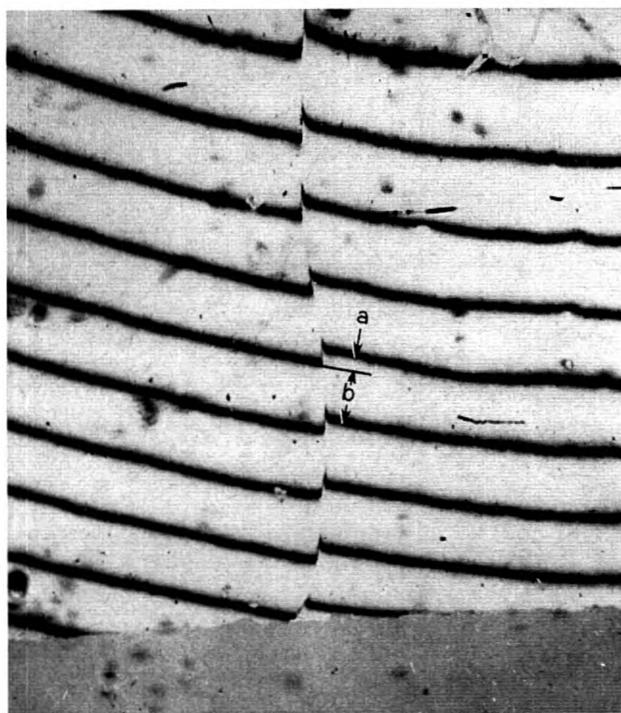


Fig. 10 — Evaluation of the interference fringe pattern of a shallow step in a surface.

shows a sample with aquadag coating and the lines marking the p-n junction. Under this, three typical fringe patterns obtained on this sample are presented. As pointed out in the beginning, one can regard the fringe pattern as contour lines for the distance between mirror and sample. The profile along any arbitrary straight line will show the structure of the sample surface. The profiles along the marked x-axes are shown to the right in each case. They were obtained by plotting the points of intersection between the n-th fringe and the x-axis against n. The original surface and the bevel (in this case 1°) are easily recognized. The dashed line is an extrapolation of the original surface. The vertical line marks the position of the p-n junction. The layer thickness is obtained as the difference in n at this point between the extrapolated original surface and the beveled surface. Note that the beveled surface need not necessarily be flat.

Fig. 11 — Evaluation of the interference fringe pattern of a steep and high step in a surface.

In the second case the fringes turn back. With a slightly different setting of the mirror the fringes could be almost parallel with the turning point outside the field of sight. The fringe pattern then resembles the first case, and therefore it might be easily misinterpreted.

The third case makes the plot unnecessary. The x-axis is chosen in such a way that it coincides with a fringe in the original surface. Thus, in the profile plot, the original surface is horizontal. Hence, the layer thickness can be obtained by counting the number of intersecting fringes between original surface and the p-n junction. This is, therefore, the most convenient setting of the mirror.

The noted number gives in each case the layer thickness in "fringes." All three cases are in essential agreement. The layer thickness in this case is

$$\Delta n \times \frac{\lambda}{2} = (16.5 \pm 0.5) \times \frac{5.89}{2} \times 10^{-5} \text{ cm}$$
$$= (4.85 \pm 0.15) \times 10^{-4} \text{ cm}$$

Fig. 9 gives the fringe pattern obtained with a silicon p-n junction marked by the plating procedure.

The evaluation of steps in a surface is shown for two cases. The very shallow step in Fig. 10 is an example in which fractions of $\lambda/2$ are to be measured. The step here is

$$\frac{a}{a+b} \times \frac{\lambda}{2} = 0.195 \times \frac{5.89}{2} \times 10^{-5} \text{ cm} = 5.75 \times 10^{-6} \text{ cm}$$

In Fig. 11 the step is so high and steep that it is impossible to correlate the fringes crossing the step. But with the aid of the bevel, seen in the lower part of Figure 11, a correlation is possible. The height of the step along the drawn line is

$$(25 - 12) \times \frac{\lambda}{2} = 13 \times \frac{5.89}{2} \times 10^{-5} \text{ cm} = 3.8 \times 10^{-4} \text{ cm}$$

The accuracy of the method depends mainly on the quality of the optically flat mirror since it serves as a plane of reference. A thin mirror is likely to be slightly bent under the pressure of the clamp. Therefore, it is advisable not to work with too high a pressure. For the measurement of layer thicknesses the quality of the original surface is also important. An accuracy of 5 per cent is easily obtained using half-silvered microscope slides for the mirror. These slides are essentially flat over the small region covered by the microscope.

Bell System Technical Papers Not Published in This Journal

ALBRECHT, E. G.,[5] DIETZ, A. E.,[1] CHRISTOFERSON, E. W.,[6] and SLOTHOWER,, J. C.[6]

Co-ordinated Protection for Open-Wire Joint Use — Minneapolis Tests, A.I.E.E. Commun. and Electronics, **24,** pp. 217–223, May, 1956.

ANDERSON, P. W.[1]

Note on Ordering and Antiferromagnetism in Ferrites, Phys. Rev., **102,** pp. 1008–1013, May 15, 1956.

ATALLA, M. M., see Preston, K., Jr.

BAKER, W. O., see Winslow, F. H.

BENSON, K. E., see Goss, A. J.

BENNETT, W. R. [1]

Techniques for Measuring Noise. Part III, Electronics, **29,** pp. 162–165, May, 1956.

BENNETT, W. R.[1]

Electrical Noise, Part IV. Design of Low Noise Equipment, Electronics, **29,** pp. 154–157, June, 1956.

BENNETT, W. R.[1]

Electrical Noise. Part V. Noise Reduction in Communication Systems, Electronics, **29,** pp. 148–151, July, 1956.

BENNETT, W. R.[1]

Methods of Solving Noise Problems, Proc. I.R.E., **44,** pp. 609–638, May, 1956.

---

[1] Bell Telephone Laboratories Inc.
[5] Northwestern Bell Telephone Company.
[6] Northern State Power Company, Minneapolis.

BOGERT, B. P.[1]

The VOBANC — A Two-to-One Bandwidth Reduction System. J. Acous. Soc., **28,** pp. 399–404, May, 1956.

BONNEVILLE, S., see Noyes, J. W.

BOYD, R. C.[1]

Objectives and General Description of the Type-Pl Carrier System, A.I.E.E. Commun. and Electronics, **24,** pp. 188–191, May, 1956.

BOYET, H., see Weisbaum, S.

BULLARD, W. R.,[4] and WEPPLER, H. E.[2]

Co-ordinated Protection for Open-Wire Joint Use — Present Trends, A.I.E.E. Commun. and Electronics, **24,** pp. 215–216, May, 1956.

CHYNOWETH, A. G.[1]

Surface Space Charge Layers in Barium Titanate, Phys. Rev., **102,** pp. 705–714, May 1, 1956.

CHYNOWETH, A. G.[1]

Spontaneous Polarization of Guanidine Aluminum Sulfate Hexahydrate at Low Temperatures, Phys. Rev., **102,** pp. 1021–1023, May 15, 1956.

CHYNOWETH, A. G.[1] and MCKAY, K. G.[1]

Photon Emission From Avalanche Breakdown in Silicon, Phys. Rev. **102,** pp. 369–376, Apr. 15, 1956.

DIETZ, A. E., see Albrecht, E. G.

DITZENBERGER, J. A., see Fuller, C. S.

DUDLEY, H. W.[1]

Fundamentals of Speech Synthesis, J. Audio Engg. Soc., **3,** pp. 170–185, Oct., 1955.

EBERHART, E. K.,[1] HALLENBECK, F. J.,[1] AND PERKINS, E. H.[1]

Circuit and Equipment Descriptions of Type-P1 Carrier System, A.I.E.E. Commun. and Electronics, **24,** pp. 195–204, May, 1956.

---

[1] Bell Telephone Laboratories Inc.
[2] American Telephone and Telegraph Company, Inc.
[4] Ebasco Services, Inc., New York.

ELLIS, H. M.,[7] PHELPS, J. W.,[1] ROACH, C. L.,[8] and TREEN, R. E.[7]

Co-ordinated Protection for Open-Wire Joint Use — Ontario Tests, A.I.E.E. Commun. and Electronics, **24,** pp. 223–236, May, 1956.

FULLER, C. S.,[1] and DITZENBERGER, J. A.[1]

Diffusion of Donor and Acceptor Elements in Silicon, J. Appl. Phys., **27,** pp. 544–553, May, 1956.

FULLER, C. S.[1]

Some Analogies Between Semiconductors and Electrolyte Solutions, Record of Chem. Progress, **17,** pp. 75–93, No. 2, 1956.

GARRETT, C. G. B., see Law, J. T.

GASTON, C. M.[1]

Stop Playing Hide-and-Seek with Engineering Drawings, Iron Age Magazine, **177,** pp. 100–101, May 17, 1956.

GAUDET, S., see Noyes, J. W.

GELLER, S.[1]

The Crystal Structure of Gadolinium Orthoferrite, $GdFeO_3$, J. Chem. Phys., **24,** pp. 1236–1239, June, 1956.

GILLEO, M. A.[1]

Magnetic Properties of a Gadolinium Orthoferrite, $GdFeO_3$ Crystal, J. Chem. Phys., **24,** pp. 1239–1243, June, 1956.

GILOTH, P. K.[1]

A Simulator for Analysis of Sampled Data Control Systems, Proc. Natl. Simulation Conf., pp. 21.1–21.8, Jan., 1956.

GOSS, A. J.,[1] BENSON, K. E.,[1] and PFANN, W. G. [1]

Dislocations at Compositional Fluctuations in Germanium-Silicon Alloys, Acta Met., Letter to the Editor, **4,** pp. 332–333, May, 1956.

HALLENBECK, F. J., see Eberhart, E. K.

HARROWER, G. A.[1]

Auger Electron Emission in the Energy Spectra of Secondary Electrons from Mo and W., Phys. Rev., **102,** pp. 340–347, Apr. 15, 1956.

[1] Bell Telephone Laboratories Inc.
[7] Hydro-Electric Power Commission of Ontario, Toronto, Ont., Canada.
[8] Bell Telephone Company of Canada, Montreal, Que., Canada.

HARROWER, G. A.[1]

Dependence of Electron Reflection on Contamination of the Reflecting Surface, Phys. Rev., **102,** pp. 1288–1289, June 1, 1956.

HOWARD, J. D., JR.[2]

Application of the Type-P1 Carrier System to Rural Telephone Lines, A.I.E.E. Commun. and Electronics, **24,** pp. 205–214, May, 1956.

HUTSON, A. R.[1]

Effect of Water Vapor on Germanium Surface Potential, Phys. Rev., **102,** pp. 381–385, Apr. 15, 1956.

KATZ, D.[1]

A Magnetic Amplifier Switching Matrix, A.I.E.E. Commun. and Electronics, **24,** pp. 236–241, May, 1956.

KOWALCHIK, M., see Trumbore, F. A.

LAW, J. T.,[1] and GARRETT, C. G. B.[1]

Measurements of Surface Electrical Properties of Bombardment-Cleaned Germanium, J. Appl. Phys., **27,** p. 656, June, 1956.

LEWIS, H. W.[1]

Two-Fluid Model of an "Energy-Gap" Superconductor, Phys. Rev., **102,** pp. 1508–1511, June 15, 1956.

LOGAN, R. A., see Thurmond, C. D.

LOZIER, J. C.[1]

A Study State Approach to the Theory of Saturable Servo Systems, I.R.E. Trans., **PGAC, 1,** pp. 19–39, 1956.

LUNDBERG, J. L.,[1] and ZIMM, B. H.[9]

Sorption of Vapors by High Polymers, J. Phys. Chem., **60,** pp. 425–428, Apr. 16, 1956.

MATTHIAS, B. T.,[1] and REMEIKA, J. P.[1]

Ferroelectricity in Ammonium Sulfate, Phys. Rev., Letter to the Editor, **103,** p. 262, July 1, 1956.

[1] Bell Telephone Laboratories Inc.
[2] American Telephone and Telegraph Company, Inc.
[9] General Electrical Research Laboratories.

McKay, K. G., see Chynoweth, A. G.

McLean, D. A.[1]

Tantalum Solid Electrolytic Capacitors, Proc. Natl. Conf. Aeronautical Electronics, pp. 289–294, May, 1956.

McSkimin, H. J.[1]

Propagation of Longitudinal Waves and Shear Waves in Cylindrical Rods at High Frequencies, J. Acous. Soc., **28,** pp. 484–494, May, 1956.

Noyes, J. W.,[8] Gaudet, G.,[8] and Bonneville, S.[8]

Development of Communications in Canada, Elec. Engg., **75,** p. 539, June, 1956.

O'Brien, J. A.[1]

Cyclic Decimal Codes for Analoge to Digital Converters, A.I.E.E. Commun. and Electronics, **24,** pp. 120–122, May, 1956.

Owens, C. D.[1]

Stability Characteristics of Molybdenum Permalloy Powder Cores, Elec. Engg., **75,** pp. 252–256, Mar., 1956.

Pearson, G. L.[1]

Electricity from the Sun, Proc. World Symp. Appl. Solar Energy, pp. 281–288, Book.

Perkins, E. H., see Eberhart, E. K.

Pfann, W. G.[1]

Zone Melting: A Fresh Outlook for Fractional Crystallization, Chem. & Engg. News, **34,** pp. 1440–1443, Mar. 26, 1956.

Pfann, W. G., see Goss, A. J.

Phelps, J. W.[1]

Protection Problems in Telephone Distribution Systems, Wire and Wire Products, **31,** pp. 555–596, May, 1956.

Phelps, J. W., see Ellis, H. M.

[1] Bell Telephone Laboratories Inc.
[8] Bell Telephone Company of Canada, Montreal, Que., Canada.

Pierce, J. R.[1]

Physical Sources of Noise, Pro. I.R.E., **44,** pp. 601–608, May, 1956.

Pomeroy, A. F.,[1] and Suarez, E. M.[1]

Determining Attenuation of Waveguide From Electrical Measurements on Short Samples, I.R.E. Trans. **MTT-4,** pp. 122–129, Apr., 1956.

Pondy, P. R.[1]

Dust-Lint Control in Tube Fabrication, Electronics, **29,** pp. 246–250, June, 1956.

Preston, K., Jr.[1] and Atalla, M. M.[1]

Transient Temperature Rise in Semi-Infinite Solid Due to a Uniform Disc Source, J. Appl. Mechanics, **23,** p. 313, June, 1956.

Prince, E.[1]

Neutron Diffraction Observation of Heat Treatment in Cobalt Ferrite, Phys. Rev., **102,** pp. 674–676, May 1, 1956.

Remeika, J. P., see Matthias, B. T.

Reiss, H.[1]

p-n Junction Theory by the Method of Functions, J. Appl. Phys., **27,** pp. 530–537, May, 1956.

Rice, S. O.[1]

A First Look at Random Noise, A.I.E.E. Commun. and Electronics, **24,** pp. 128–131, May, 1956.

Smith, D. H.[1]

Power Supplies for the Type-Pl Carrier System, A.I.E.E. Commun. and Electronics, **24,** pp. 191–195, May, 1956.

Suarez, E. M., see Pomeroy, A. F.

Theuerer, H. C.[1]

Purification of Germanium Tetrachloride by Extraction with Hydrochloric Acid and Chlorine, J. of Metals, **8,** pp. 688–690, May, 1956.

[1] Bell Telephone Laboratories Inc.

THURMOND, C. D.,[1] and LOGAN, R. A.[1]

The Distribution of Copper Between Germanium and Ternary Melts Saturated with Germanium, J. Phys. Chem., **60,** pp. 591–594, May, 1956.

THURMOND, C. D., see Trumbore, F. A.

TRUMBORE, F. A.,[1] THURMOND, C. D.,[1] and KOWALCHIK, M.[1]

The Germanium-Oxygen System, J. Chem. Phys., Letter to the Editor, **24,** p. 1112, May, 1956.

WEISBAUM, S.[1] and BOYET, H.[1]

Broadbank Non-Reciprocal Phase Shifts — Analysis of Two Ferrite Slabs in Rectangular Guide, J. Appl. Phys., **27,** pp. 519–524, May, 1956.

WEPPLER, H. E., see Bullard, W. R.

WINSLOW, F. H.,[1] BAKER, W. O.,[1] and YAGER, W. A.[1]

The Structure and Properties of Some Pyrolyzed Polymers, Proc. Conf. on Carbon, pp. 93–102, 1956.

WOOD, E. A. MRS. [1]

The Question of a Phase Transition in Silicon, J. Phys. Chem., **60,** pp. 508–509, Apr., 1956.

YAGER, W. A., see Winslow, F. H.

[1] Bell Telephone Laboratories Inc.

Recent Monographs of Bell System Technical Papers Not Published in This Journal*

BASHKOW, T. R.

DC Graphical Analysis of Junction Transistor Flip-Flops, Monograph 2615.

BECKER, J. A., see Rose, D. J.

BITTRICH, G., see Compton, K. G.

BOYET, H., see Weisbaum, S.

BRANDES, R. G., see Rose, D. J.

BRATTAIN, W. H., see Garrett, C. G. B.

COMPTON, K. G., EHRHARDT, R. A., and BITTRICH, G.

Brass Plating, Monograph 2467.

EGERTON, L., and KOONCE, S. E.

Structure and Properties of Barium Titanate Ceramics, Monograph 2517.

EHRHARDT, R. A., see Compton, K. G.

EIGLER, J. H., see Sullivan, M. V.

FRANCOIS, E. E., see Law, J. T.

FULLER, C. S., see Reiss, H.

GARRETT, C. G. B., and BRATTAIN, W. H.

Some Experiments on, and a Theory of, Surface Breakdown, Monograph 2589.

HAGELBARGER, D. W.

SEER, A Sequence Extrapolating Robot, Monograph 2599.

* Copies of these monographs may be obtained on request to the Publication Department, Bell Telephone Laboratories, Inc., 463 West Street, New York 14, N. Y. The numbers of the monographs should be given in all requests.

Haynes, J. R., and Westphal, W. C.

Radiation Resulting from Recombination of Holes and Electrons in Silicon, Monograph 2622.

Herring, C., and Vogt, E.

Transport and Deformation-Potential Theory for Many-Valley Semiconductors, Monograph 2596.

Herring, C., see Vogel, F. L., Jr.

Kleimack, J. J., see Wahl, A. J.

Koonce, S. E., see Egerton, L.

Law, J. T., and Francois, E. E.

Adsorption of Gases on a Silicon Surface, Monograph 2600.

Lewis, H. W.

Superconductivity and Electronic Specific Heat, Monograph 2597.

Logan, R. A.

Thermally Induced Acceptors in Germanium, Monograph 2601.

Lundberg, J. L., see Zimm, B. H.

May, J. E., Jr.

Low-Loss 1000-Microsecond Ultrasonic Delay Lines, Monograph 2584.

Mendel, J. T.

Microwave Detector, Monograph 2602.

Paterson, E. G. D.

An Over-all Quality Assurance Plan, Monograph 2630.

Pomeroy, A. F., and Suarez, E. M.

Attenuation of Waveguide from Electrical Measurements on Short Samples, Monograph 2625.

Press, H., and Tukey, J. W.

Power Spectral Methods of Analysis and Application in Airplane Dynamics, Monograph 2606.

READ, W. T., JR., see Vogel, F. L., Jr.

REISS, H., and FULLER, C. S.

Influence of Holes and Electrons on Solubility of Lithium in Silicon, Monograph 2603.

ROSE, D. J., BECKER, J. A., and BRANDES, R. G.

On the Field Emission Electron Microscope, Monograph 2588.

SUAREZ, E. M., see Pomeroy, A. F.

SULLIVAN, M. V., and EIGLER, J. H.

Electrolytic Stream Etching of Germanium, Monograph 2595.

THOMAS, D. E.

Tables of Phase of a Semi-Infinite Unit Attenuation Slope, Monograph 2550.

TUKEY, J. W., see Press, H.

VAN UITERT, L. G.

High-Resistivity Nickel Ferrites-Minor Additions of Manganese or Cobalt, Monograph 2594.

VOGT, E., see Herring, C.

VOGEL, F. L., JR., HERRING, C., AND READ, W. T., JR.

Dislocations in Plastic Deformation, Monograph 2616.

WAHL, A. J., and KLEIMACK, J. J.

Factors Affecting Reliability of Alloy Junction Transistors, Monograph 2604.

WESTPHAL, W. C., see Haynes, J. R.

WEISBAUM, S., AND BOYET, H.

A Double-Slab Ferrite Field Displacement Isolator at 11 kmc, Monograph 2605.

ZIMM, B. H., and LUNDBERG, J. L.

Sorption of Vapors by High Polymers, Monograph 2573.

Contributors to This Issue

W. L. Bond, B.S. 1927 and M.S. 1928, Washington State College; Bell Telephone Laboratories, 1928–. Mr. Bond has conducted investigations in the mineral field including studies of the piezoelectric effect in minerals and similar studies of synthetic crystals. He has designed optical, X-ray, and mechanical tools and instruments for the orientation, cutting and processing of crystals. Mr. Bond also served as consultant on quartz crystals with the War Production Board. He is a member of the American Physical Society, and of the American Crystallographic Association.

Walter H. Brattain, B.S., Whitman College, 1924; M.A., University of Oregon, 1926; Ph.D., University of Minnesota, 1929. Honorary D.Sc. Portland University, 1952, Whitman College and Union College, 1955. Radio section, Bureau of Standards, 1928–29. Bell Telephone Laboratories, 1929–. Co-inventor with Dr. John Bardeen of point contact transistor. Primary activity at Laboratories in semi-conductors. Research in field of thermionics, particularly electronic emission from hot surfaces. Frequency standards, magnetometers and infra-red phenomena. Studied magnetic detection of submarines for National Defense Research Committee at Columbia University, 1942–43. Visiting lecturer at Harvard University, 1952–53. Author of numerous technical articles. Recipient of John Scott Medal, 1955, and Stuart Ballantine Medal of Franklin Institute, 1952. Fellow of American Physical Society, American Academy of Arts and Sciences and American Association for the Advancement of Science. Member of Franklin Institute, Phi Beta Kappa and Sigma Xi.

C. C. Cole, B.S. in E.E., State College of Washington, 1923; U. S. Navy 1917–1919; Western Electric Company 1923–. His first assignment was in manufacturing development on paper and mica capacitors. Other assignments include manufacturing development on loading coils, quality control, and inspection development laboratory. During World War II he handled the design and construction of testing facilities for various defense projects. Since World War II he has been engaged in

inspection methods development and in the development and design of testing facilities for telephone apparatus and cable. Member of Sigma Tau and A.I.E.E.

ARTHUR B. CRAWFORD, B.S.E.E. 1928, Ohio State University; Bell Telephone Laboratories, 1928–. Mr. Crawford has been engaged in radio research since he joined the Laboratories. He has worked on ultra short wave apparatus, measuring techniques and propagation; microwave apparatus, measuring techniques and radar, and microwave propagation studies and microwave antenna research. He is author or co-author of articles which appeared in The Bell System Technical Journal, Proceedings of the I.R.E., Nature, and the Bulletin of the American Meteorological Society. He is a Fellow of the I.R.E. and a member of Sigma Xi, Tau Beta Pi, Eta Kappa Nu, and Pi Mu Epsilon.

HARALD T. FRIIS, E.E., 1916, D.Sc., 1938, Royal Technical College (Copenhagen); Engineering Department of the Western Electric Company, 1919–1924. Bell Telephone Laboratories, 1925–. Dr. Friis, Director of Research in High Frequency and Electronics, has made important contributions on ship-to-shore radio reception, short-wave studies, radio transmission (including methods of measuring signals and noise), a receiving system for reducing selective fading and noise interference, microwave receivers and measuring equipment, and radar equipment. He has published numerous technical papers and is co-author of a book on the theory and practice of antennas. The I.R.E.'s Morris Liebmann Memorial Prize, 1939, and Medal of Honor, 1954. Valdemar Poulson Gold Medal by Danish Academy of Technical Sciences, 1954. Danish "Knight of the Order of Dannebrog," 1954. Fellow of I.R.E. and A.I.E.E. Member of American Association for the Advancement of Science, Danish Engineering Society and Danish Academy of Technical Sciences. Served on Panel for Basic Research of Research and Development Board, 1947–49, and Scientific Advisory Board of Army Air Force, 1946–47.

C. G. B. GARRETT, B.A., Cambridge University (Trinity College), 1946; M.A., Cambridge, 1950; Ph.D., Cambridge, 1950. Instructor in Physics, Harvard University, 1950–52. Bell Telephone Laboratories, 1952–. Before coming to the Laboratories, Dr. Garrett's principal research was in the field of low-temperature physics. At the Laboratories he has been engaged in research and exploratory development on semiconductor surfaces and, for the past year, has supervised a group working in this field. He is the author of "Magnetic Cooling" (Harvard

University Press, 1954). Senior Scholar of Trinity College, Cambridge, 1945. Twisden Student of Trinity College, 1949. Fellow of Physical Society (London). Member of American Physical Society.

L. D. HANSEN, B.S., Montana State College, 1924; Western Electric Company, 1924–. Mr. Hansen joined the Equipment Engineering Organization at the Hawthorne Plant of The Western Electric Company in Chicago in 1924 where he was engaged in preparation of telephone central office specifications. He transferred to the Kearny, N. J., Plant in 1928 where he was promoted to section chief in 1929. He transferred to the Engineer of Manufacture Organization in 1930 and worked on carrier and repeater test development and methods until 1941 when he was promoted to Department Chief in charge of wired switching apparatus and equipment test set development and methods.

WILLIAM C. JAKES, JR., B.S.E.E., Northwestern University, 1944; M.S., Northwestern, 1947; Ph.D., Northwestern, 1948. Bell Telephone Laboratories, 1949–. Dr. Jakes is engaged in microwave antenna and propagation studies and holds a patent in microwave antennas. He is the author of chapter in antenna engineering handbook (McGraw-Hill). Member of Sigma Xi, Pi Mu Epsilon, Eta Kappa Nu, I.R.E. and Phi Delta Theta.

AMOS E. JOEL, JR., B.S., Massachusetts Institute of Technology, 1940; M.S., M.I.T., 1942; Bell Telephone Laboratories, 1940–. Mr. Joel is Switching Systems Development Engineer responsible for coordinating the exploratory development of a trial electronic switching system. Prior to his present position he worked on relay engineering, crossbar test laboratory, fundamental development studies, circuits for relay computers, preparation of a text and teaching switching design, designing AMA computer circuits and making fundamental engineering studies on new switching systems. He holds some forty patents. Member of A.I.E.E., I.R.E., Sigma Xi and Association for Computing Machinery.

ARCHIE P. KING, B.S., California Institute of Technology, 1927. After three years with the Seismological Laboratory of the Carnegie Institution of Washington, Mr. King joined Bell Telephone Laboratories in 1930. Since then he has been engaged in ultra-high-frequency radio research at the Holmdel Laboratory, particularly with waveguides. For the last ten years Mr. King has concentrated his efforts on waveguide transmission and waveguide transducers and components for low-loss circular

electric wave transmission. He holds at least a score of patents in the waveguide field. Mr. King was cited by the Navy for his World War II radar contributions. He is a Senior Member of the I.R.E. and is a Member of the American Physical Society.

D. T. Robb, B.S., University of Chicago, 1927; Western Electric Company, 1927–. Mr. Robb has been concerned with measurement and testing problems throughout his career. In the electrical laboratory at Hawthorne Works, Chicago, he specialized in ac standardization. Later he worked on the development of shop test methods and test sets. In 1944 he transferred to take charge of radar test engineering at the Eleventh Avenue Plant of Western Electric in New York City. In 1946 he supervised the engineering of the standards laboratory at Chatham Road Plant in Winston Salem, N. C. Currently, he has charge of transmission test set development and test set design at Kearny Works, N. J.

Harry R. Shillington, B.S. in E.E. Iowa State College, 1937; Long Lines Department of the American Telephone and Telegraph Company, 1928–1932; Western Electric Company, 1937–. Mr. Shillington's first assignment was that of product engineering on panel dial equipment. During World War II and the Korean War he was engaged in test engineering on various defense projects. He is presently concerned with the development of special test facilities for telephone apparatus. Member of Eta Kappa Nu and Tau Beta Pi.

Friedolf M. Smits, Dipl.Phys. and Dr.Rer.Nat., University of Freiburg, Germany, 1950; research assistant, Physikalisches Institut, University of Freiburg, 1950–54; Bell Telephone Laboratories, 1954–. As a member of the Solid State Electronics Research Department of the Laboratories, Dr. Smits has been concerned with diffusion studies of germanium and silicon for semiconductor device applications. He is a member of the American Physical Society and the German Physical Society.

Frank H. Tendick, Jr., B.S.E.E., 1951, University of Michigan; Bell Telephone Laboratories, 1951–. Mr. Tendick was first engaged in work pertaining to the synthesis of networks employed in the L3 coaxial cable system. Later he engaged in the design of transistor networks for digital computers. More recently, he has been associated with exploratory studies of submarine cable systems. He is a member of the I.R.E. Mr.

Tendick also belongs to four honor societies, Tau Beta Pi, Eta Kappa Nu, Sigma Xi and Phi Kappa Phi.

LEISHMAN R. WRATHALL, B.S., 1927, University of Utah. Mr. Wrathall did another year of graduate work at the University of Utah and joined Bell Telephone Laboratories in 1929. For many years he was primarily concerned with studies of the characteristics of non-linear coils and capacitors. During World War II non-linear coils were used extensively in radar systems, and his work in this field was intensified. Later he was occupied with general circuit research. He is now engaged in studies of conductor problems, particularly digital repeaters, as a member of the Transmission Research Department at Murray Hill.