Solderless Wrapped Connections

Introduction

By J. W. McRAE

In the telephone plant during the course of a single year, the operation of connecting a wire to a metal terminal is carried out approximately one billion times. Many of these connections are made in the factory. Others are made during the installation of equipment and a substantial number are made in the course of normal operation of the telephone plant. Successful functioning of the plant depends on trouble-free performance of each of these connections, most of which are now soldered in accordance with long-standing practice. Recently, a new technique for joining wires to terminals has been developed which will have important technical and economic advantages in the Bell System and which should have similar advantages in other fields.

The immediate need for a new connection arose with the development of the new wire spring general purpose relay.* In this relay the terminals appear in the form of closely spaced wires and the standard methods of applying connections were not satisfactory. Since the production schedule for these new relays required something like fifty million connections a year on relay terminals alone, an intensive effort was made to devise a satisfactory method for wiring. The first result was the development of a tool which could wrap a few turns of wire around the terminals of the relay, and do this efficiently on the closely-

---

* Keller, A. C., A New General Purpose Relay for Telephone Switching Systems. Bell System Tech. J., **31**, pp. 1023–1067, Nov., 1952.

spaced wire terminals.† It was soon found that similar tools could be used to advantage on existing types of terminals, and the Western Electric Company is now making extensive use of wrapping tools. Connections made with these tools are being soldered after wrapping.

In the meantime, further development work indicated the possibility that by wrapping the wire under tension around a properly shaped terminal, the need for soldering might be eliminated. Major economies appeared possible if such a solderless wrapped connection were applied extensively in wiring communication equipment. In addition, there would be freedom from trouble due to solder splashes in equipment and there would be an appreciable reduction in the consumption of tin for use in solder.

Three papers in this issue of the JOURNAL describe the present status of the program undertaken to exploit these possibilities. The paper on design indicates that practical tools for making solderless wrapped connections can be designed, built and used; the paper on analysis describes the basis for belief that the method is fundamentally sound; and the paper on evaluation indicates that the resulting connections are satisfactory for use in the telephone plant.

Other types of tools capable of cutting, skinning and wrapping the wire in one operation are under development, and the problems presented in adapting the basic techniques to other conductors, such as aluminum wire and stranded copper wire are being studied. In fact, a whole new area of development effort has been opened up.

Thus the work reported in the following papers is of interest from two points of view. On the one hand it is a record of progress in the continual quest for less expensive and more reliable equipment. On the other hand, it is an example of the broad effects which are often the result of development aimed at a specific problem. The search for a solution to the problem of making connections to a new type terminal has led to a new approach to the whole problem of wire-terminal connections.

† Miloche, H. A., Mechanically Wrapped Connectors. Bell Labs. Record, **29**, pp. 307–311, July, 1951.

Solderless Wrapped Connections

PART I — STRUCTURE AND TOOLS

By R. F. MALLINA

In the search for a better way of connecting wires to apparatus terminals a new joining method has been discovered. The new method not only eliminates soldering and its hazards but also reduces cost, improves quality and conserves space. In contrast to the solder joint which depends largely on human judgement and skill, the new connection is made with a calibrated tool. A degree of uniformity has been obtained which virtually eliminates the need for product inspection. The trend toward smaller apparatus and automation may now be further intensified due to the use of this new method of making electrical connections.

INTRODUCTION

Methods of joining wires to apparatus terminals for the purpose of electrical conduction can be broadly divided into two groups: solder connections and pressure connections. There are others such as welded and brazed connections; however, they are relatively few in number. The annual production of solder connections in the Bell System is estimated to be one billion. In television and radio manufacture the number of connections made per year is in the order of ten billion. Because of the high cost of manual soldering, the pressure connection is of great importance to the communication industry. One form of pressure connection — the solderless wrapped connection — will be described in this article.

In order to determine the technical and economic value of a new type of pressure connection it is necessary to compare it with those now accepted as good connections in the communication industry. A large portion of this article will, therefore, be devoted to the analysis of pressure connections some of which have been in use since the early development of the telephone.

525

WHAT IS A PRESSURE CONNECTION?

The chart Fig. 1 shows six typical pressure connections classified in terms of seven requirements. In this classification the screw connection, for example, meets the following requirements: large contact area, high contact force, great mechanical stability, long life, easy to disconnect. The space, however, which the screw connection occupies is large and its cost is high. In the history of electricity it is probably the oldest and best pressure connection. In the second column of the table in Fig. 1 is the plug connection. It is small in size, easy to disconnect, but has no large contact area, no high contact force, no long life, no mechanical stability and is not low in cost.

As will be shown later the solderless wrapped connection in Column 6 of Fig. 1 is indicated as meeting all seven requirements. Its main advantage over the screw connection is that it is low in cost and small in size.

CONTACT AREA

The effective contact area relative to the cross sectional area of the wire is of great importance since it controls the resistance of the connection. It must remain uniform in size, metallically bright and not be affected by temperature changes, vibration and handling.

Contact area is not easily defined. For example two flat metal surfaces having an area of one square centimeter each and brought into contact do not necessarily have a contact area of one square centimeter. If the force holding them together is small, only the high spots

	1	2	3	4	5	6
REQUIREMENTS	FAHNESTOCK CLIP	PLUG	CRIMP	WIRE NUT	SCREW	SOLDERLESS WRAPPED
1. LARGE CONTACT AREA			✓	✓	✓	✓
2. HIGH CONTACT FORCE			✓	✓	✓	✓
3. LONG LIFE			✓	✓	✓	✓
4. SMALL SIZE		✓	✓			✓
5. MECHANICALLY STABLE			✓		✓	✓
6. EASILY DISCONNECTED	✓	✓		✓	✓	✓
7. LOW COST						✓

Fig. 1 — Classification of pressure connections.

of the surfaces touch and large currents passing through such a connection may develop heat and melt the metal at the high spots.

CONTACT FORCE

To make the above mentioned area of one square centimeter effective for electrical conduction it is necessary to press the two metal parts together with a force so high that essentially all particles of the area are intimately interlocked and free from insulating impurities. If the pressure is high enough, the film which appears in the form of oxide on the terminal surface is crushed. In general it is assumed that in a good connection the contact force should be such that the contact area produced is equal to or greater than the cross sectional area of the wire. In screw connections, crimped connections and wrapped connections the contact area is normally a multiple of the wire cross sectional area. In plug connections, such as on vacuum tube sockets, the contact area is very small. In a Fahnestock clip for example, the contact area is about one quarter of the cross sectional area of the wire.

LIFE

If the electrical resistance of a pressure joint is to remain constant with time, it is the contact area which must remain substantially constant, but not necessarily the contact force. Once the metal particles are tightly interlocked a subsequent reduction in contact force within relatively wide limits does not change the electrical resistance. The resistance will increase only when the force is reduced to such a low value that vibration and handling cause partial separation of the contact area. In such a case two changes may take place:

1. The atmosphere may enter through the fringe of the contact area and a process of corrosion may begin.

2. The effective contact area may be reduced through dislodging some of the contacting particles.

In both cases the resistance is increased. Therefore, to produce a durable connection it is important to have a firm joint and one such that the atmosphere cannot enter the contact area. The term commonly used for such a joint is "gas tight."

ELASTIC RESERVE

The question now arises how much reduction in contact force can be tolerated before a joint loses its gas tightness? In all types of pressure
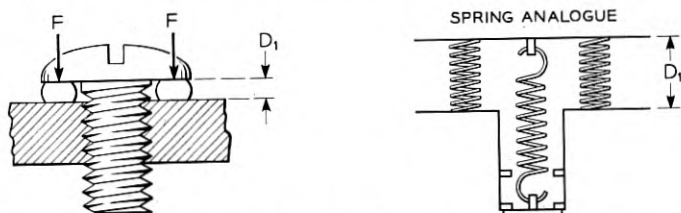
Fig. 2 — Screw tightened and wire compressed (with accompanying "Spring Analogue").

connections, the forces which hold the wire and terminal together are provided by the springiness or elasticity of the materials. The elasticity in a Fahnestock clip is quite apparent because of the long spring member. On the other hand, a screw connection, such as shown in Fig. 2, does not appear to have spring members of any kind. Analysis, however, shows that there is considerable elastic deformation. Most of this elastic deformation is in the elongation of the screw shank and there is also some bending in the screw head and some compression in the screw threads. When the screw is tightened, the wire which is interposed between screw head and nut is also elastically deformed. Since in most electrical connections the wire is a soft material such as copper or aluminum, it is nearly always compressed beyond the yield point and only the recovery of the overstressed material can be considered as elastic reserve.

To determine the usefulness of a connection and compare pressure connections of different kinds, the elastic reserve in the deformed wire and the deformed terminal must be measured or computed. Elastic reserve might be expressed either in terms of stiffness or the potential energy stored in the system. Stiffness S is defined as the ratio of the "applied force F to the elastic return D"; potential or elastic energy E is "one half of the product of the force F and the elastic return D." ($E = \frac{1}{2}FD$.)

Example: A wire is placed under a screw (Fig. 2) and compressed by the screw head to a thickness D_1. The screw is then loosened so that it just touches the wire. The wire now has expanded to a certain extent and its new thickness is D_2. (Fig. 3.) The difference ($D_2 - D_1 = D_W$) is the elastic return and the ratio $F/D_W = S_W$ is the useful stiffness of the wire.

A preferred way of expressing elastic reserve, is in terms of stored energy. (Strictly speaking the equation $E = \frac{1}{2}FD$ holds only for springs with constant stiffness. A round wire compressed by a screw head becomes stiffer as the compression increases). The distance, $D_S = D_3 - D_W$, is the elongation of the screw (see Fig. 3). The total energy

stored up is therefore the sum of the energy in the screw and wire ($E = E_s + E_w$).

Screws in terminal blocks are normally made of hard materials such as brass or phosphor bronze. Wires used for the interconnection of components are nearly always of a soft material and have a tendency to creep. If creep takes place in the wire during the many years a screw connection is in use, it is advantageous to have the loss of potential energy in the wire compensated for by the energy stored in the screw.
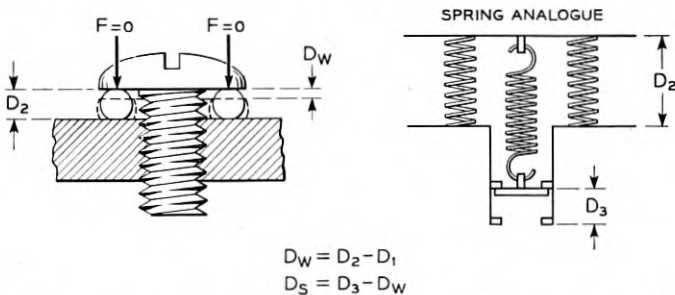

$$D_W = D_2 - D_1$$
$$D_S = D_3 - D_W$$

Fig. 3 — Screw loosened and wire not compressed. The recovery of the wire is denoted D_w. The recovery of the screw is $D_3 - D_w$ (with accompanying "Spring Analogue").

A screw, for example, made of soft copper would not be expected to make a lasting connection. If on the other hand the screw is made of a material which has little creep and much elasticity, such as brass or steel, it would act as a spring member and tend to keep the connection tight.

Several typical screw connections were measured to determine the elastic reserve. It was found that on an average the potential energy stored in the screw is about equal to that stored in the wire. Plastic flow of the wire creates an effective bearing area comparable to the area of the screw shank.

THE SOLDERLESS WRAPPED CONNECTION

The detailed analysis of the screw connection as an introduction to the solderless wrapped connection was necessary not only because the screw has such wide use as an electrical pressure connection but chiefly because of its proven value as a durable connection. When new types of pressure connections are put into large scale production, the question invariably arises, What is their life? While considerable analytical work has been done on the cold flow of metals under stress* and while certain

---

* See Part II.

theoretical predictions can be made on the durability of new connections, it affords additional satisfaction to be able to show that the solderless wrapped connection is in many respects similar in structure and performance to the conventional screw connection. If then this fact is supported by parallel analytical work, there should be little doubt that the solderless wrapped connection is a durable pressure connection.

THE RECTANGULAR TERMINAL

Generally speaking the terminal best suited for a wrapped connection is a terminal of rectangular cross section. It is an inexpensive terminal since it can be blanked from sheet stock or coined from round wire. It is ideally suited for a pressure connection because the edges produce a concentrated high pressure on the wire. The stress distribution in the wire produced by the terminal edges is shown diagrammatically in Figs. 4 and 5. If the wire is wound with high tension around the rectangular terminal, the terminal edges dig into the soft copper wire, crush and shear the oxide on both the wire and the terminal and form a large, intimate and metallically clean "gas tight" contact area. An indication of the high pressure is the crushing of the hard nickel silver terminal edge by the soft copper wire. A pattern of contact areas on both wire and terminal is shown in Fig. 6. Several turns of wire are required to preserve the high contact force. In general it is assumed that the first and last two edges around which the wire is wrapped do not contribute much to the joint as contact areas. A seven-turn wrapped connection on a rectangular terminal thus has six effective turns. Each turn contacts four edges or a total of twenty-four contact areas for six effective turns.
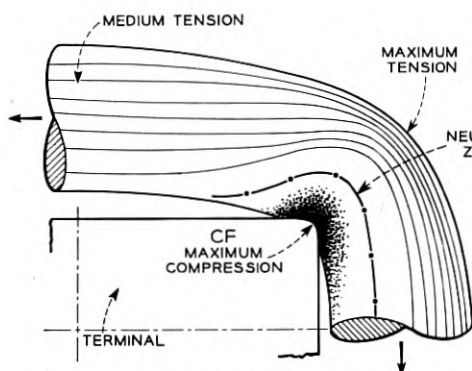

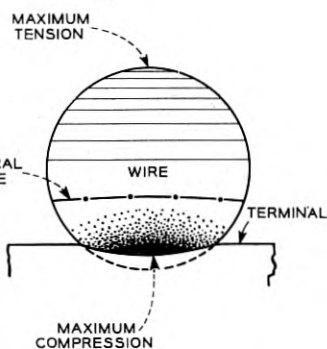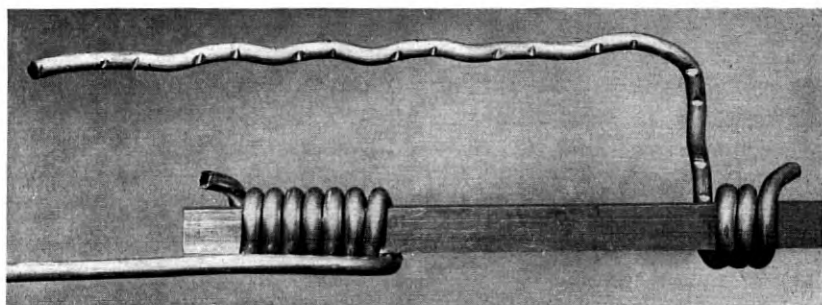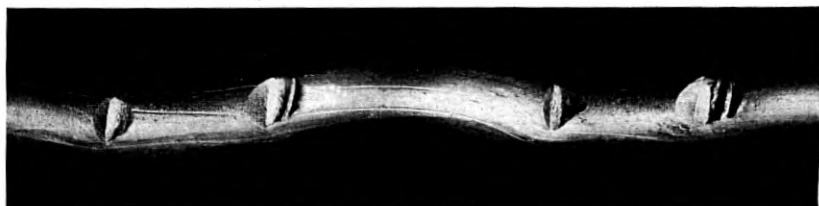FIG. 4 — Stress distribution along one-quarter turn of wire.

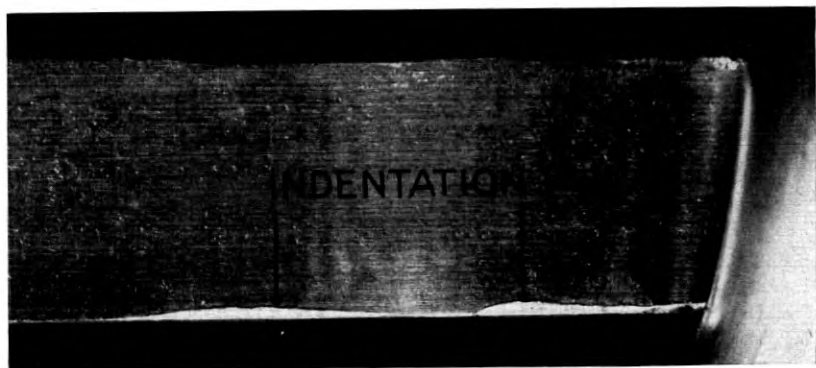FIG. 5 — Cross section through terminal edge showing stress distribution in the wire.

PATTERN OF CONTACT AREAS.

CONTACT AREAS ENLARGED.

THE SOFT COPPER WIRE CRUSHES THE HARD NICKEL SILVER TERMINAL EDGE
IN THE WRAPPING PROCESS.

Fig. 6 — Contact areas.

WHAT IS A GOOD CONNECTION?

The quality of a connection depends fundamentally on two factors: the contact area and the contact pressure. As long as there is sufficient pressure and the atmosphere cannot enter the joint, the connection is considered a good one. If, however, the elastic energy which holds the two surfaces together is small, various disturbances may cause a partial separation of the interlocking metal particles and thus effect a change in resistance. For normal telephone applications a good connection may, therefore, be defined as one which not only has sufficient contact area and contact pressure but which also has sufficient elastic reserve to maintain contact area and contact pressure throughout the desired life, which may be forty years or more.

The mechanical disturbances to which a connection may be subjected are: handling, vibration, temperature changes and cold flow.

HANDLING AND VIBRATION

The solderless wrapped connection is well protected from the point of view of handling and vibration. The locking effect on the rectangular terminal or a terminal having well defined edges does not permit loosening of the center turns from the terminal. In vibration tests where conventional soldered connections were compared with solderless wrapped connections, it was found that solderless wrapped connections outlast soldered connections. This is due to the fact that a sudden change in cross section from wire to solder lump localizes the stresses at a very small area. (See Figs. 7 and 8.) In the screw connection a similar condition exists where the wire emerges from under the screw head. In the
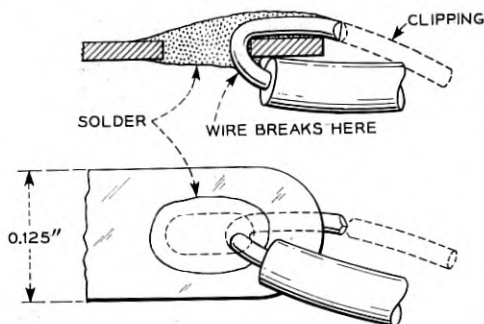

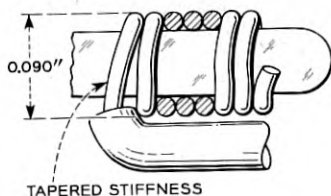FIG. 7 — Standard solder connection of U-relay.

FIG. 8 — Solderless wrapped connection of modified U-relay terminal.

solderless wrapped connection there is no sudden change in cross section and therefore no localization of stresses. The term commonly used to indicate the gradual change in rigidity of the wire as it approaches the anchoring point is "tapered stiffness." (See Fig. 8.)

HEAT AND COLD FLOW

When a pressure connection is subjected to high temperatures, which may be due to large current or to heat transfer from adjacent components, the pressure at the joint is relaxed. This is true in the solderless wrapped connection as well as in the screw connection. The same process of relaxation takes place in normal temperature with time. The relaxation of pressure with temperature and time will be shown in another part of this paper. Under ordinary conditions the relaxation of pressure in a solderless wrapped connection is not sufficiently large to indicate any change in resistance during a forty-year life. Furthermore, as Mason and Osmer point out in their paper, solid state diffusion takes place as time goes on. This process strengthens the joint mechanically and improves it electrically.

QUANTITATIVE EVALUATION OF ELASTIC RESERVE

Because of the above mentioned disturbances to which a pressure connection may be subjected, it is important to know how much elastic reserve is stored in a connection. If no potential energy were stored in the wire and in the terminal, no contact pressure would be produced. If little potential energy were stored in a connection, a slight change in temperature due to differential expansion of the metals would loosen the connection. The same would be true with vibration and handling.

A rough comparison with other pressure connections will serve to illustrate how much elastic reserve a solderless wrapped connection has to have in order to withstand the disturbances to which it may be subjected. The best known pressure connection, and the most universally used, is the screw connection. On a No. 4 screw (0.112″), the force exerted in clamping the No. 24 gauge (0.020″) wire is about 135 lbs. The elastic energy is stored by compressing the wire and by elongating the screw shank. Similarly, in a solderless wrapped connection as shown in Fig. 9 a total force of 90 lbs is exerted on the edges of the terminal (24 corners). Here the greater part of the energy is stored in the terminal which receives torsional as well as compressional stress from the tension in the wrapped wire. (See Figs. 10(a) and 10(b)).

Fig. 9 shows in diagrammatic form how the stored energy in a screw connection and a solderless wrapped connection compare. A typical solderless wrapped connection — seven turns of 20-mil copper wire wound with 1300 grams applied force on a 0.0148" x 0.062" nickel silver terminal — has approximately 3 mil pounds of stored energy E. (2.4 mil pounds are stored in the terminal and 0.6 mil pounds in the wire). The screw connection — No. 4-40 screw (0.112") tensioned to 135 lbs on 20-mil copper wire — has approximately 2.7 mil pounds of stored energy E. (Approximately half in the screw and half in the wire). In the screw connection the energy is about equally divided between the screw and the wire whereas in the solderless wrapped connection a larger part of the energy is stored in the terminal. This is advantageous since the hard materials of which terminals are generally made have less cold flow than copper. In a solderless wrapped connection, if the
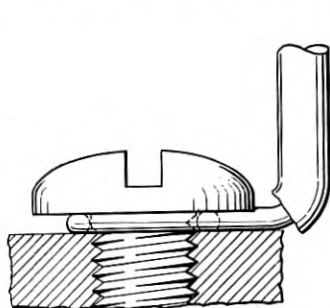
NO. 4-40 (0.112") BRASS SCREW
NO. 24 GA (0.020") COPPER WIRE
CONTACT FORCE 135 LBS

(a)   SCREW CONNECTION
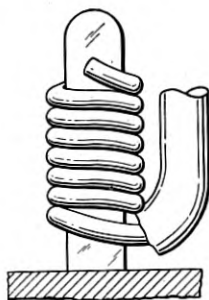
0.0148" x 0.062" NICKEL SILVER TERMINAL
7 TURNS NO. 24 GA (0.020") COPPER WIRE (1300 GR AF)
CONTACT FORCE 90 LBS (24 CORNERS)

(b)   SOLDERLESS WRAPPED CONNECTION

1.4 SCREW   1.3 WIRE

2.4 TERMINAL   0.6 WIRE

TENSION
18,750 PSI

COMPRESSION
18,250 PSI

$E = 1.4 + 1.3 = 2.7$

TORSION AND
COMPRESSION

TENSION
8500 PSI

$E = 2.4 + 0.6 = 3$
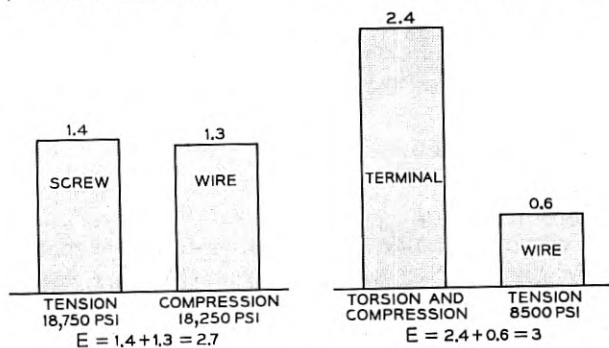
(c)   TOTAL ENERGY E (IN 10^{-3} INCH LBS)

Fig. 9 — Elastic energy stored in screw connection and in solderless wrapped connection.

terminal size is changed to 0.020″ x 0.062″ there is considerably less energy stored in the terminal and slightly more in the wire. Inasmuch as a screw connection in most cases depends on the human element, that is the amount of torque applied by the operator, it can be expected that some screw connections will be made with a force that may vary from 75 lbs to 150 lbs. The wrapped connection on the other hand, being made with a calibrated tool, can be expected to give substantially the same contact force at all times.

In order to understand more clearly how the wire and the terminal interact when they are under mutual stress and exposed to heat, the elastic deformation of the wire and the terminal must be analyzed. It has been shown in Fig. 4 that the wrapped wire on the four sides of the rectangle is under tension. This tension causes the terminal to twist. If instead of a helix the terminal were surrounded by a series of hoops,

NO CONTACT FORCE

TWIST ANGLE

CONCENTRATED CONTACT FORCE

STIFF TERMINAL

END TURN

INDENTATIONS AND CONCENTRATED CONTACT FORCE

COMPRESSION

CENTER TURNS

RELATIVELY SOFT WIRE

TENSION

END TURN

(a)
LONGITUDINAL TENSION IN A WIRE CAUSES TERMINAL TO TWIST

(b)
MODEL MADE WITH A RUBBER TERMINAL WRAPPED WITH RUBBER TUBING EMPHASIZES THE TWIST PRODUCED BY THE LONGITUDINAL TENSION IN THE TUBING
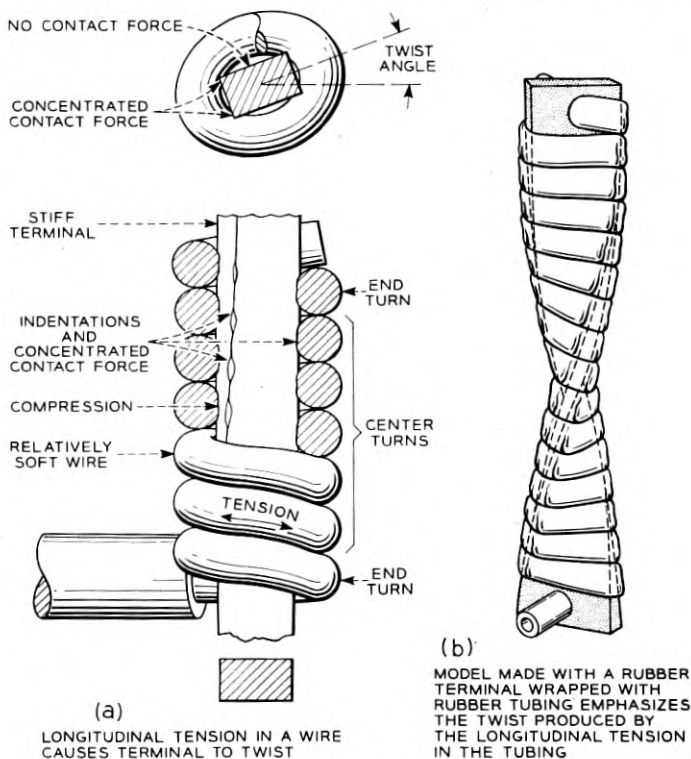
Fig. 10 — (a) Longitudinal tension in wire causes terminal to twist. (b) Model made with a rubber terminal wrapped with rubber tubing emphasizes the twist produced by the longitudinal tension in the tubing.
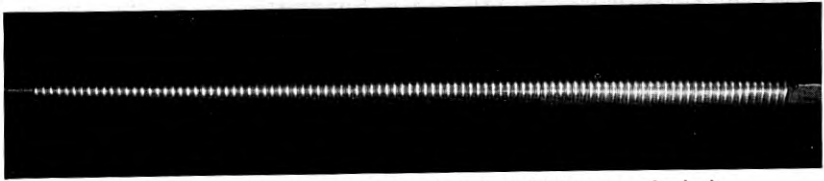
Fig. 11 — Twist of rectangular terminal (100 turns of wire).

the terminal would be compressed at the edges but the terminal would not twist. The terminal twist in the wrapped connection is therefore due to the fact that the terminal is surrounded by a helix and not by hoops. Figs. 10(a) and 10(b) show that a left-hand helix produces a right-hand twist in the terminal. As will be shown later, this visible deformation of the terminal is being used to determine the tension in the wire. The twist in the terminal of a wrapped connection with many turns can readily be seen in Fig. 11. For example an initial twist of 46° is produced in a nickel silver terminal 0.0148″ × 0.062″ wrapped with 100 turns of No. 24 (0.020″ dia.) copper wire with an applied force of 1300 grams.

One way to visualize the behavior of the wire and the terminal when wrapped under tension, exposed to time and heat and then unwrapped, is to represent the wire and the terminal by linear springs. This is shown schematically in Fig. 12. Position 1 represents both wire and terminal before wrapping. Position 2 represents the wire wrapped on the terminal. Position 3 is the same as Position 2 except that the wrapped terminal has been exposed at room temperature (20°C) for eight days. This causes the terminal twist to relax from 46° to 39°. Positions 2 and 3 are analogous to the wire under tension and the terminal under tor-

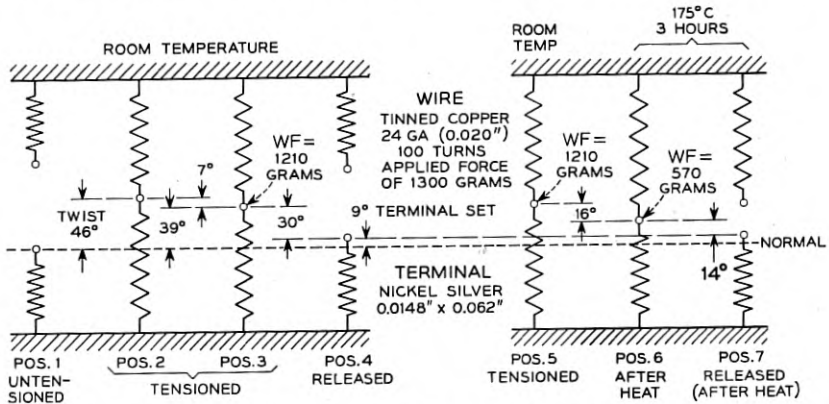Fig. 12 — Energy in unheated connection proportional to 30°. Energy in heated connection proportional to 14°.

sion. The force WF is the tension in the wrapped wire. This force can be determined by dividing the torque necessary to twist the terminal by the effective moment arm. Since the elongation of the wire cannot readily be measured, the terminal twist was chosen to determine the force exerted at the terminal edge. The 39° terminal twist shown in Posi-
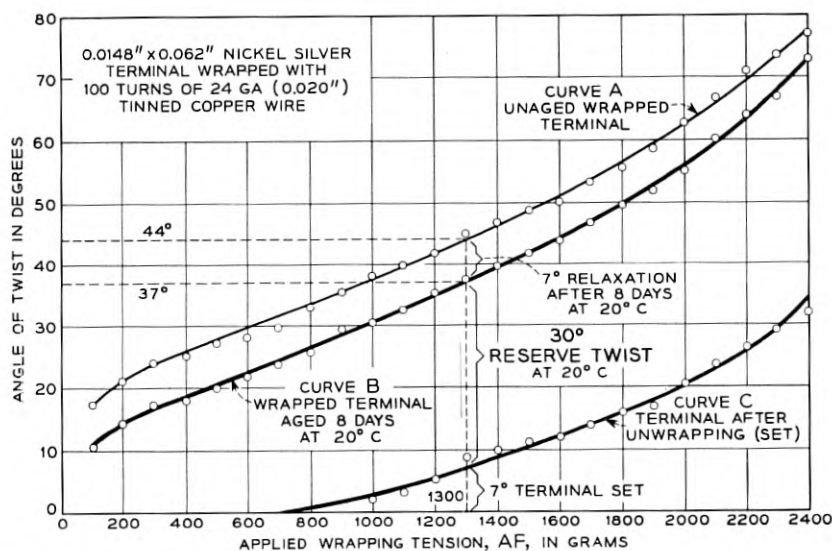
Fig. 13 — Angle of twist in terms of applied wrapping tension.

tion 3, however, cannot be used for determining the force since the terminal may be overstressed as is shown in Position 4. Instead of returning 39° the unwrapped terminal returned only 30°. In other words the terminal has taken a set of 9°. Fig. 12 illustrates the deformation of wire and terminal only for one value of applied tension, namely 1300 grams. If the angle of twist is measured for applied tension ranging from 100 to 2400 grams, a set of curves is obtained as shown in Fig. 13. Curve A shows the angle of twist immediately after the terminal is wrapped. Curve B shows the relaxation after eight days aging at room temperature. Curve C represents the terminal set. The value between Curves B and C is the elastic reserve. For 1300 grams applied tension the elastic reserve is expressed as 30° reserve twist.

Using the before mentioned ratio of torque and moment arm, the force WF can now be determined. The torque required to twist the terminal 30° is 37.2 inch grams. (See Fig. 14.) The effective moment

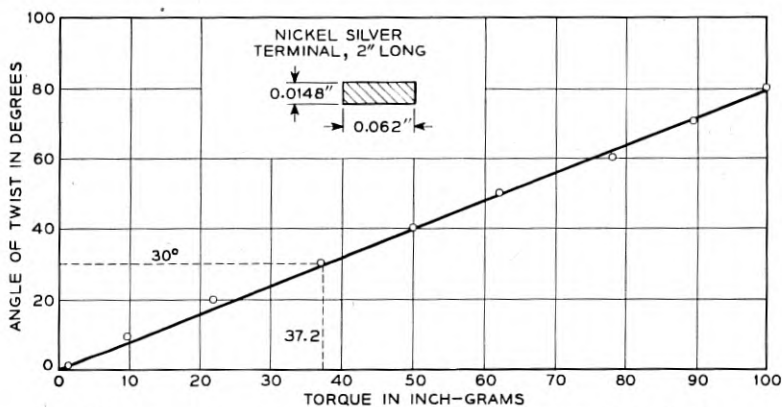Fig. 14 — Torque required to twist the unwrapped terminal.
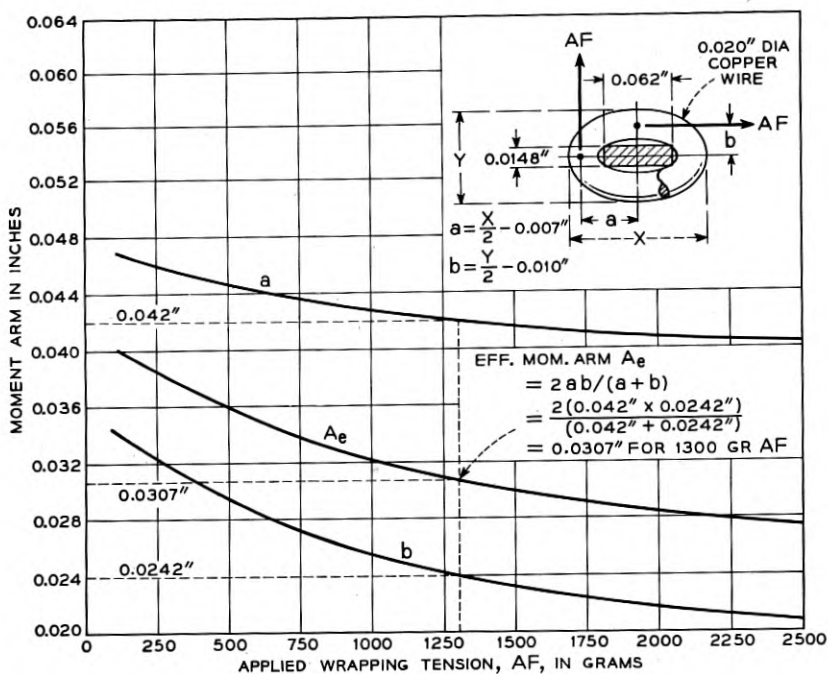
Fig. 15 — Moment arm in terms of tension.

arm A_e, which decreases a slight amount as the wrapping tension AF increases (see Fig. 15), is equal to $2ab/(a + b)$*. Therefore, $A_e = 2$ $(0.042'' \times 0.0242'')/(0.042'' + 0.0242'') = 0.0307$ in. Thus the tension in the wire WF $= T/A_e = 37.2/0.0307 = 1210$ grams. Using the recovery angle of the terminal as a measure of force, the tension in the wire can now be plotted in terms of angular twist. This is shown in Fig. 16. It should be noted that the tension in the wire WF (wrapped
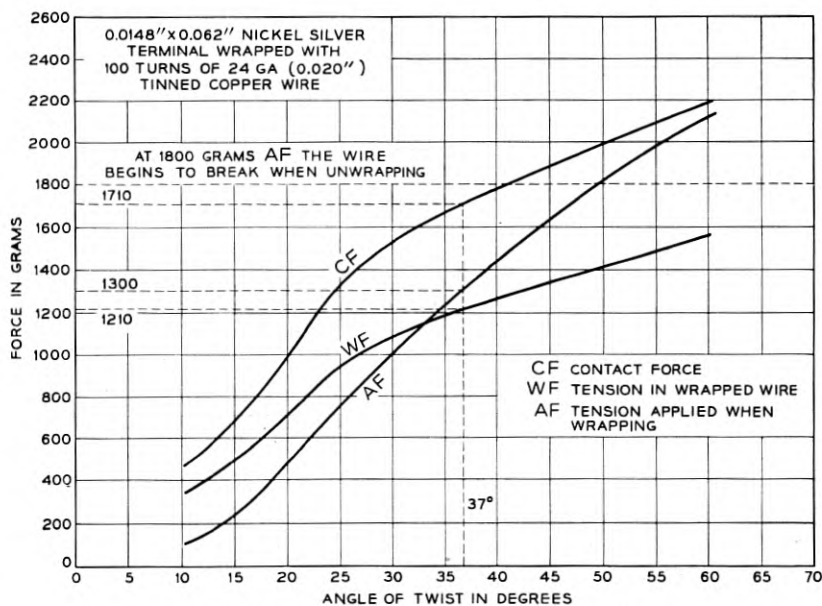
Fig. 16 — Forces in terms of angular twist.

force) is not directly proportional to the wrapping tension AF (applied force). The reason for this is that at low applied wrapping tension the bending of the wire around the corner of the terminal produces an additional increment of tension. For example at an angle of 15° the wrapped tension WF is nearly twice as high as the applied tension AF. The wrapped tension WF and the applied tension AF are about equal when the angle of twist is 33°. At higher values of applied tension, the wrapped tension increases at a much lower rate. This is caused by the terminal taking a set. (See Fig. 13.) At 1300 grams of applied tension, which is the recommended wrapping tension for No. 24 copper wire, the wrapped tension is 1210 grams.

---

* See Appendix I.
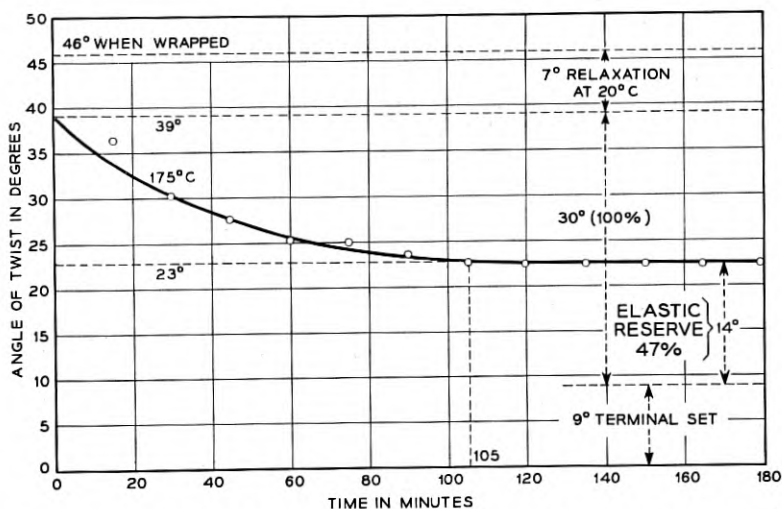
FIG. 17 — Relaxation of wrapped wire after heating for three hours at 175°C. The reserve twist of 14° is an indication of the tension left in the wire which amounts to 47 per cent of the original wrapped tension.

A severe test for a wrapped connection is heating to 175°C for three hours. This relaxes, about half the stress and is considered the equivalent of a 40-year life at 135°F. Position 5 in Fig. 12 is the same as Position 3, that is, the wire has just been wrapped onto the terminal and its tension is 1210 grams. If this connection is now heated to 175°C and the angle of twist noted every fifteen minutes, a curve is obtained as shown in Fig. 17. It should be noted that at 105 minutes the curve is for all practical purposes asymptotic at an angle of 23°. If the heated connection is cooled and unwrapped and the set in the terminal measured, it is found to be 9°. The 14° difference is a measure of the elastic reserve. This is 47 per cent of the wire tension before heating. The corresponding tension WF is then 570 grams. This process is illustrated in positions 6 and 7 of Fig. 12. Position 7 shows that the terminal set of 9° was the same as before heating.

A similar experiment was made with formex insulated wire wrapped on a nickel silver terminal. Instead of subjecting the connection to the heat of an oven, a high current was passed through the wire. Essentially the same curve as shown in Fig. 17 was obtained.

To further check the behavior of springs with complex elastic deformation such as in a wrapped connection on a terminal having edges, measurements were also made with simple helical springs tensioned

within the yield point. Two springs, one of nickel silver wire and the other of copper wire having a stiffness ratio of 5 to 6, were coupled in series (Fig. 18), tensioned to 30 grams and then heated to 173°C for two hours. The tension left after heating was 13 grams or 43 per cent of the original tension. The tension decay curve was similar to that shown for the wrapped connection. (See Fig. 17.) This test shows that in spite of the complex deformation of the wire at the corners of the terminal there is substantial agreement in results of the measurements obtained — namely 43 per cent remaining stress in the case of the helical spring and 47 per cent for the wrapped connection.

STRESSES IN THE FINISHED CONNECTION

Having determined the interacting forces in the solderless wrapped connection, the next questions of primary interest are — what are the stresses in various parts of the connection and what will happen to these stresses in forty years?

Most of the elastic energy stored in the wire is in the portion marked "Medium Tension" (Fig. 4). Here the stress is about 8,500 P.S.I. This assumes that a 20-mil wire is wrapped with 1300 grams applied force. The wrapped force or the useful force obtained from the elastic reserve is then 1210 grams.

The stresses at the corners are not easily determined because they are not uniform. This is shown in Fig. 5. As can be seen, the point of highest
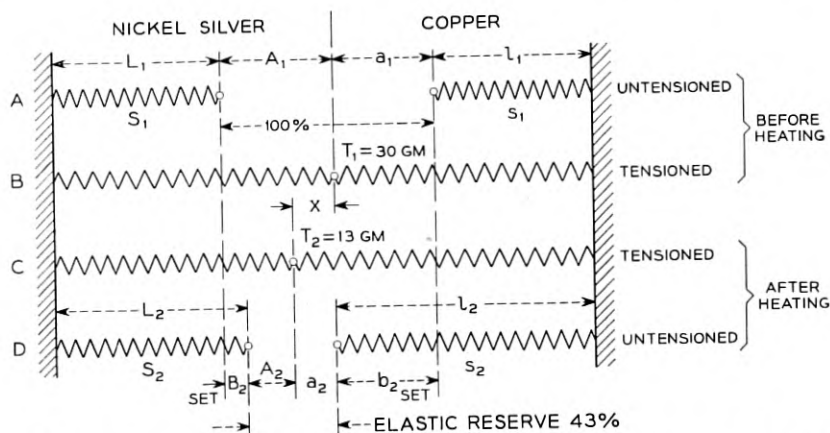
Fig. 18 — Tension T_2 in coupled springs after heating for two hours at 173°C is about 43 per cent of the original tension T_1. (Tension T_1 approximately 10 per cent below the yield point).
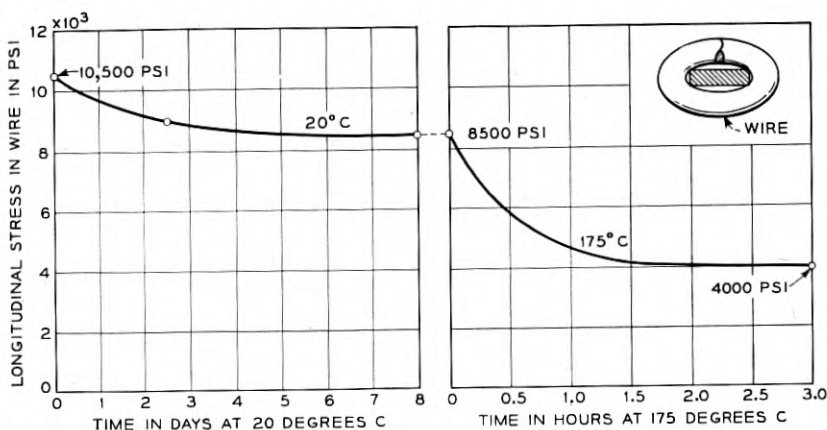
FIG. 19 — Stress relaxation in wire.

concentration is in the center of the contacting area. From this point to the periphery there is a pressure gradient which is similar to that of a circular compressed thin film of viscous material. At the boundary line the pressure is zero. The average pressure within the contact area is about 29,000 P.S.I. but the maximum stress in the center of the contact area may be as high as 100,000 P.S.I. The relaxation, that takes place in eight days at room temperature (Fig. 13), is assumed to be due to the very high initial stress in the center of the contact area which seeks equalization.
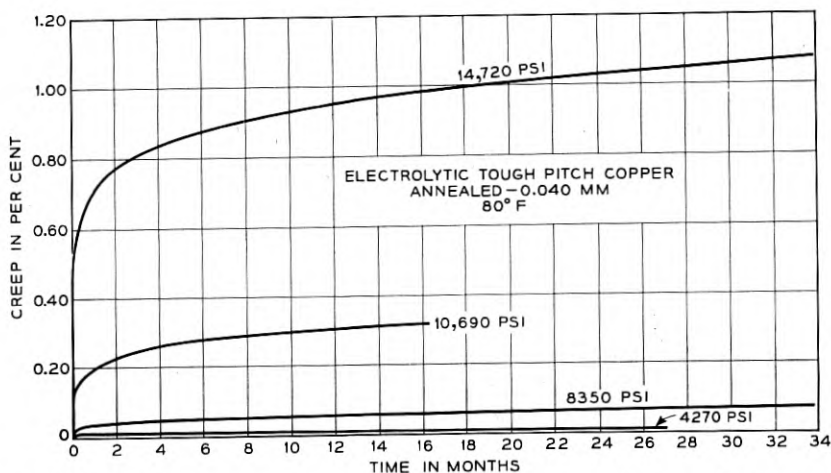
Fig. 20 — Creep curves of annealed copper for various stresses. (Courtesy of Chase Brass and Copper Co., Waterbury, Conn.).

Summarizing the stresses in the connection, one may therefore say that in the portion of the wire where most of the elastic energy resides, the stress after eight days is about 8,500 P.S.I. and at the points of contact 29,000 P.S.I. After forty years these stresses will be approximately 4,000 and 13,500 P.S.I., respectively or 47 per cent of the original stress. (See Fig. 19). As may be seen in the creep curves shown in Fig. 20, a stress of 8350 P.S.I. reaches a creep value of about 0.07 per cent in three years and from then on for all practical purposes ceases to creep.
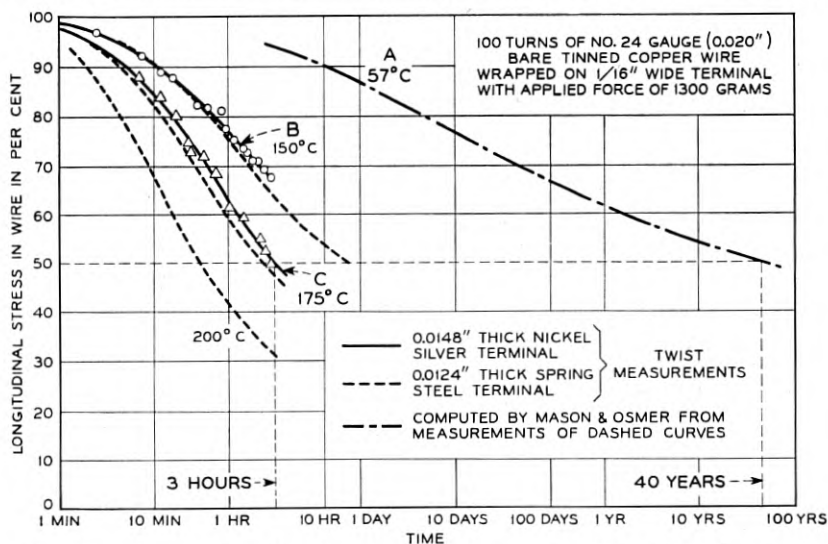
FIG. 21 — Stress relaxation in wire plotted on a logarithmic time scale.

The effect of time and temperature on the longitudinal stress in the wire can better be seen by curves plotted on a logarithmic time scale (Fig. 21). The initial stress after transient relief of three days is considered as 100 per cent. Curve A shows that at a temperature of 57°C (135°F) — which is the maximum temperature that solderless wrapped connections will be subjected to — the longitudinal stress relaxes approximately 50 per cent in 40 years. To reach the 50 per cent value at a temperature of 175°C takes approximately three hours (Curve C). Curves B and C show that the relaxation in the wire is essentially the same for either nickel silver or spring steel terminals.

METHOD OF WRAPPING

In nearly all soldered connections where a wire is to be joined to a terminal the procedure is as follows: The operator takes the skinned
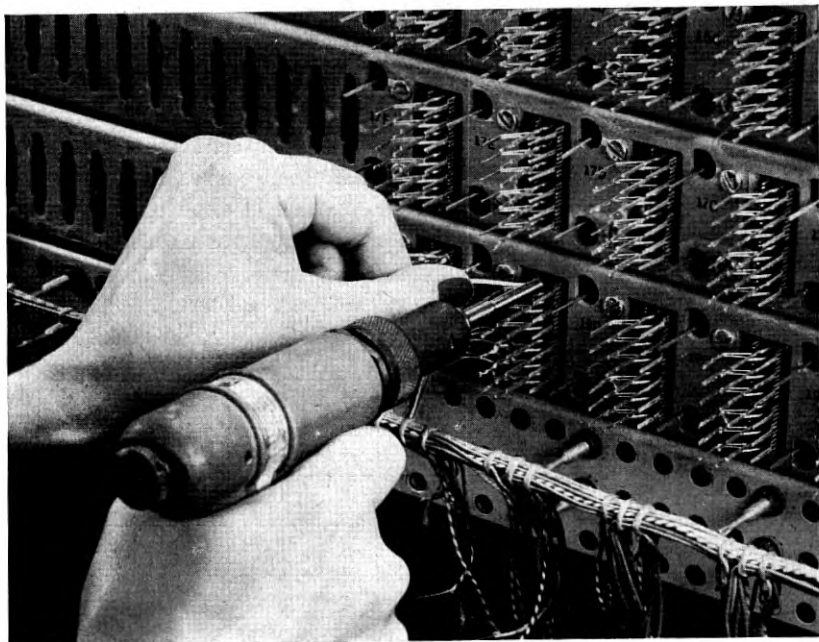
Fig. 22 — Air-driven wiring tool.

end of an insulated wire, hooks or threads it onto the terminal and applies solder. The hooking or threading is important because it is customary in production to attach several wires at first and then solder. The question now is: If a similar procedure is to be followed with a wrapped connection where the wire must surround the terminal with a high pressure, how do we produce that pressure?

In a screw connection a force is obtained by a high lever ratio. In a crimped connection the force is applied by heavy and powerful compression tools. These tools are not suitable for connecting wires to closely spaced terminals such as shown in Fig. 22. To produce high tension in the wire while it is being wrapped onto the terminal, a new method of tensioning had to be devised.

In the manufacture of helical springs it is customary to anchor the end of the wire in a hole in the arbor and tension the wire with a friction pad. By rotating the arbor a helical spring is produced. For closely spaced terminals this method is not practical as the wire cannot be fed tangentially to the terminal and the terminal cannot be rotated. A new

wire connecting concept was proposed whereby a rotating spindle housed a stationary terminal in an axial opening in the spindle and was provided with a second opening radially separated from the axial opening and arranged to accommodate a wire. When the spindle was rotated the wire was caused to form a spiral about the stationary terminal. One method involved anchoring the wire in the second opening and feeding the wire tangentially to the terminal as the spindle was rotated. Due to certain limitations inherent in tangential feed onto a stationary terminal an improved method was finally chosen. This is the axial feed method which is particularly adapted to wrapping closely spaced terminals of all cross sections. The operation of loading the wire and wrapping the connection is shown in Fig. 23. Position A shows the tool tip, Position B the bare wire 2 inserted into the feed slot 4, Position C the anchoring of the wire by bending it into the notch 5, Position D the terminal insertion and Position E the wrapping of the wire 2 by rotating the spindle 1 around the terminal 3. Position F is the finished connection. A more detailed drawing of the tool tip is shown in Fig. 24.

WRAPPING TENSION

The tension in the wire is produced by rotating the spindle 1 around the terminal 3 (Fig. 24) thus pulling the short skinner wire 2 out of the feed slot 4. In the process of pulling the wire out of the slot and wrapping it around the terminal each increment of the skinner wire length undergoes several bending operations. The first bending occurs at the edge R of feed slot 4 where the wire is bent through an angle of less than 90°. The second bending is the straightening out operation of the bent wire. The third bending takes place as the wire is wrapped around the terminal. All three bending processes contribute to the tension with which the wire is wrapped. The dimensions which control the tension and are therefore of engineering importance are the radius R at the tool tip (See Figs. 25 and 25(a)) and the wall thickness W (Fig. 24). The bending forces are inversely proportional to the respective bending curvatures and the frictional forces in turn are proportional to the bending forces. The tension imparted into the wire as it is wrapped around the terminal, however, is not only due to the friction alone but to the combined effect of friction and bending effort. If the wire were completely elastic and the friction zero, no tension could be produced. But there would be tension in the wire if the friction were zero and the wire only partly elastic such as copper wire. There also would be tension if a completely elastic wire would be pulled around an edge having friction.
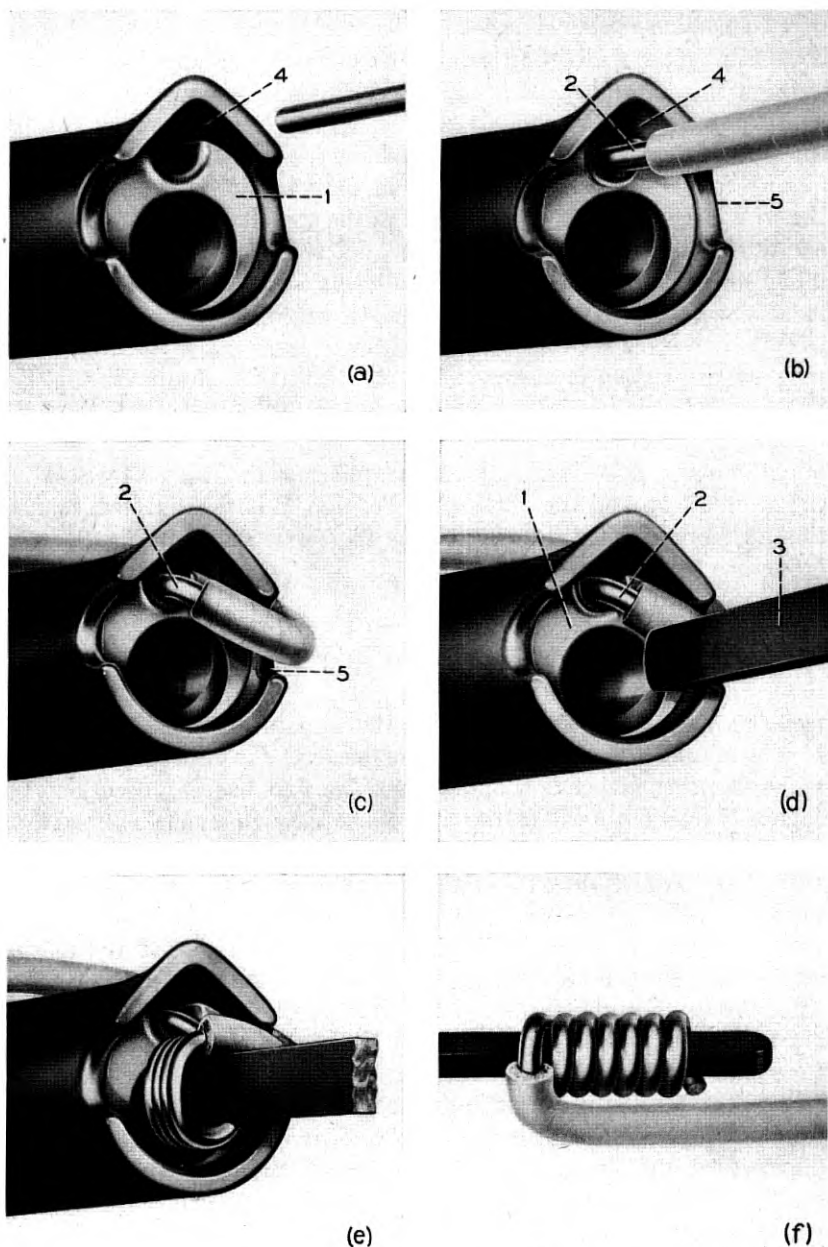
FIG. 23 — The wrapping process. A — Tool Tip. B — Wire Inserted. C — Wire Anchored. D — Terminal Insertion. E — Wire Wrapped. F — Finished Connection.
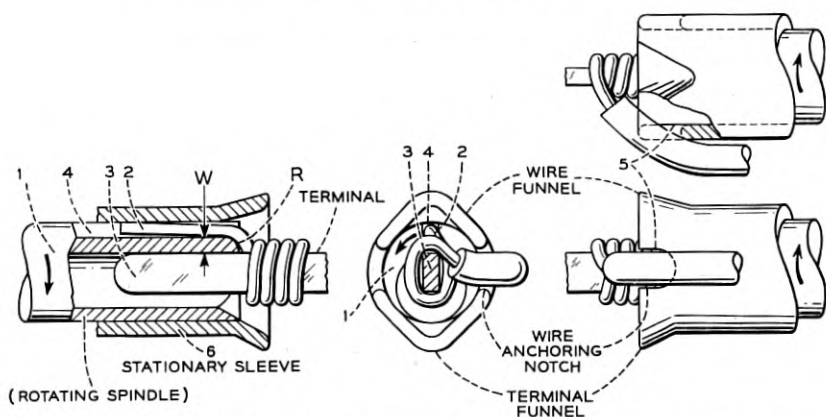
Fig. 24 — Method of wrapping the skinner wire on a rectangular terminal.

CLOSELY SPACED TERMINALS

When the terminals are closely spaced the stationary sleeve 6 and the anchoring notch 5 are used in order to anchor the first turn of wire to the terminal. (See Fig. 26.) However, when the terminals are not closely spaced the sleeve and notch are desirable but are not necessary since the insulated portion of the wire can be held by some means external of the tool at an angle of approximately 90° with respect to the tool spindle. The high acceleration of the wrapping motor produces a mass reaction of the wire leading up to the terminal. This counterforce coupled with a slight tension of the supply wire applied by the external
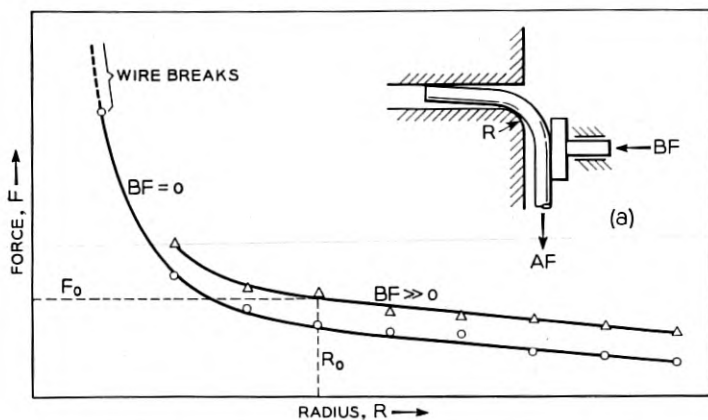
Fig. 25 — Wrapping tension is controlled by edge radius.
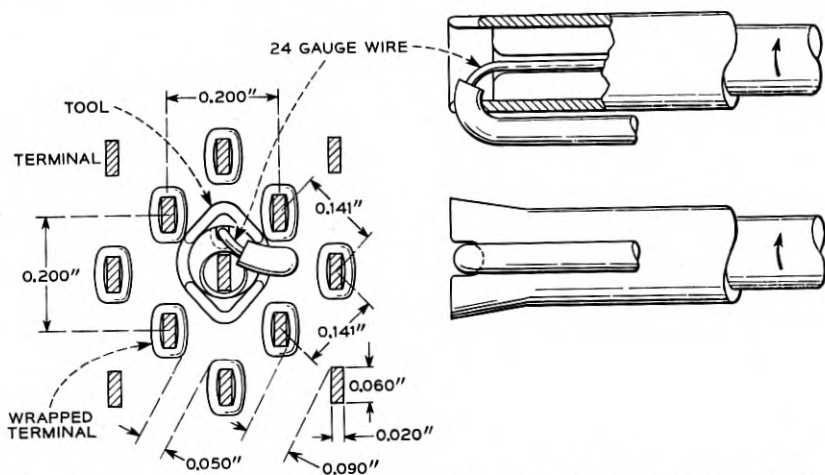
Fig. 26 — Space occupied by wrapping tool between closely-spaced terminals.

wire guiding means is sufficient to insure wrapping of the first turn. The following turns need no further anchor as the first turn locks the wire to the terminal.

The trend toward making circuit components smaller is now marked in all branches of communication engineering. With a tool tip such as
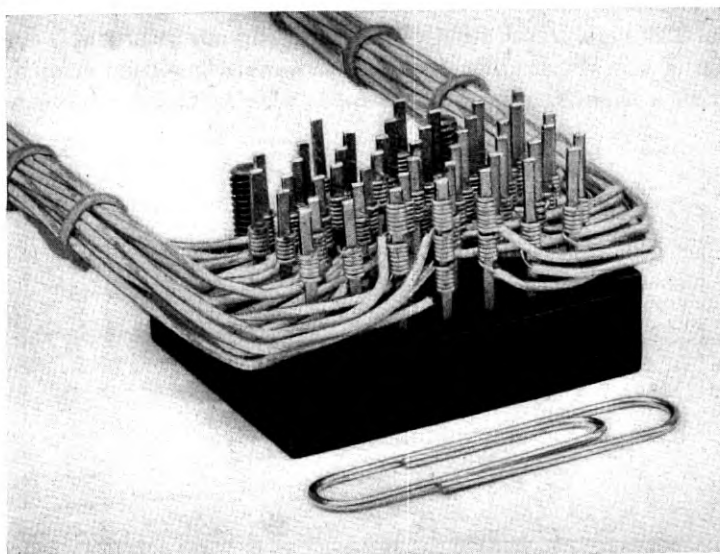
Fig. 27 — Forty-four point terminal block, 132 connection capacity, only 48 connections of 26 gauge (0.0159″) wire shown. Occupies 1⅛″ x ¾″ x 9⁄16″ or ½ cu in of space.
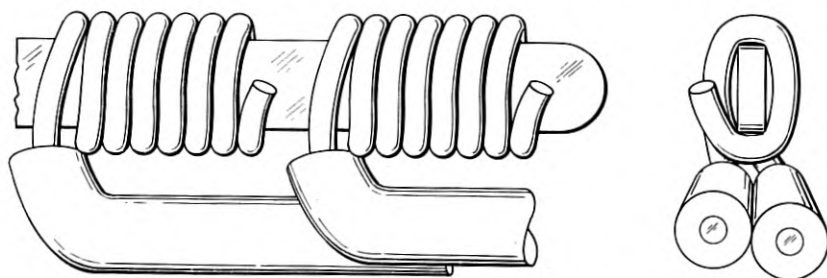
Fig. 28 — Double connection.

shown in Fig. 24, it is possible to wire apparatus having a terminal spacing as close as $2\frac{1}{2}$ times the terminal width. (See Fig. 26.) A terminal block $1\frac{1}{8}''$ by $\frac{3}{4}''$ by $\frac{9}{16}''$ having 44 terminals is shown in Fig. 27. The cables shown contain forty-eight No. 26 (0.0159" dia.) wires all wrapped on the terminals. Each terminal is capable of accommodating three wires or a total of 132 connections may be made in an area of less than one square inch. An enlarged view of a double connection is shown in Fig. 28.

REMOVAL OF CONNECTION

The solderless wrapped connection may be removed from its terminal by two methods. The most convenient method is by stripping. Two types of tools may be used for this purpose. The specially formed jaws of a pair of pliers are hooked in the back of the connection as shown in Fig. 29. By applying a force the connection may be stripped off. The other tool for stripping is shown in Fig. 30. The stripping force varies with the tightness of the wrapping and is plotted in terms of applied
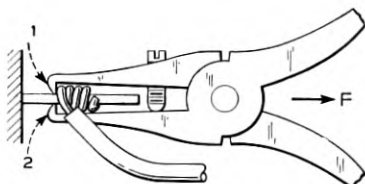
FIG. 29 — Stripping of solderless wrapped connection.
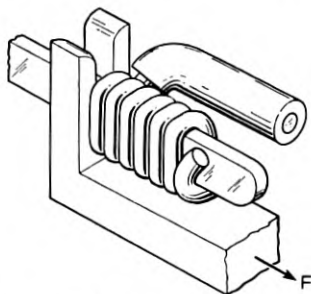
FIG. 30 — Stripping a solderless wrapped connection from a terminal.

wrapping tension in Fig. 31. Another method of removing a connection is by unwinding the helix. This may be done by using a pair of pliers as shown in Fig. 32. Either end of the wire may be used for unwrapping.

A terminal is not seriously damaged by stripping off a wrapped wire, however, the re-use of the stripped off wire is not recommended. A wire may be reconnected by skinning to the proper length and wrapping. When the wire is not sufficiently long to provide the necessary number of turns to insure a good connection, one or two turns may be wrapped and then soldered.

CONNECTION OF LARGE AND SMALL WIRES

There is no upper limit to the size of wire wrapped on adequately proportioned terminals. Connections have been made with both aluminum and copper wire over 200 mils in diameter with satisfactory results. The torque necessary to wrap large wire is considerable, since it increases with the third power of the diameter. A 20-mil wire requires a winding torque of 100 inch grams whereas, a 200-mil wire requires 100,000 inch grams (18 foot pounds). Wires as small as No. 39 (0.0035" dia.) may also be wrapped, however, the design of the wrapping tool must be changed slightly in order to facilitate the loading of the fine wire into the tool.

DIMENSIONAL RELATIONS

The data given in this paper refer only to No. 24 copper wire 20 mils in diameter. The terminal width most frequently used in conjunction with this size wire is about one-sixteenth inch or three times the wire diameter. The terminal width may also be twice the wire diameter or slightly less, however, the one-sixteenth inch size has been chosen for
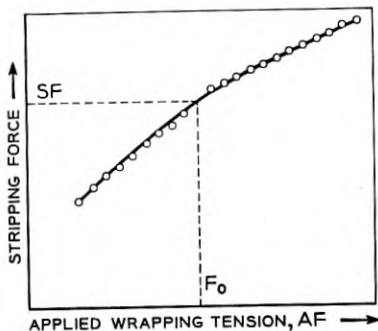
FIG. 31 — Stripping force in terms of applied wrapping tension.
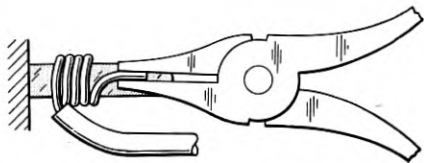
FIG. 32 — Unwinding wrapped connection with pliers.
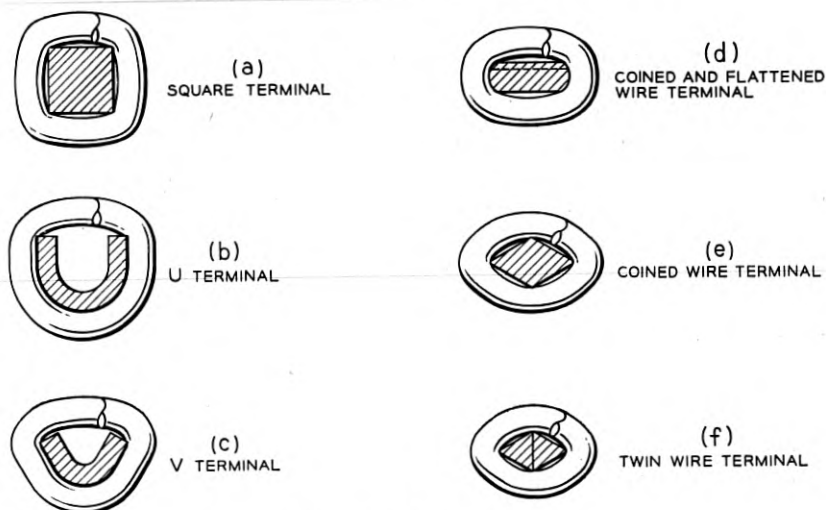
Fig. 33 — Various terminals.

good visibility. When smaller wires are used, the tendency is to make the terminal width greater than three times the wire diameter for better visibility. The terminal thickness depends to a great extent on the shape of the terminal. For a rectangular terminal the thickness may vary from three times to one-half the wire diameter. When the terminal thickness is less than one-half the wire diameter, the terminal may twist too much during the wrapping operation.

TYPES OF WRAPPED CONNECTIONS

The rectangular terminal is not the only terminal which lends itself to a good solderless wrapped connection. Any terminal offering one or more contacting edges substantially crosswise to the axis of the wrapped wire will make a good connection.

Since rectangular terminals of very thin material may twist excessively during the wrapping process the preferred shape is a U or V as shown in Fig. 33. These terminals are capable of storing even more elastic energy than a rectangular terminal of equal cross sectional area. The U and V terminals are particularly suited for vacuum tube sockets and thin relay springs.

Flattened or coined single wires as well as coined twin wires may be used as terminals for solderless wrapped connections. These are shown in Fig. 33.

Stranded wire connections have been made by laying the strands
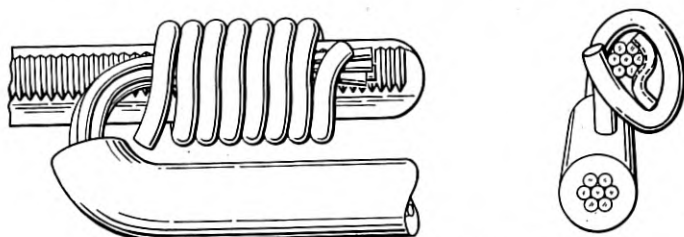
Fig. 34 — Stranded wire connection.

along the serrated edge of the terminal and then wrapping both the terminal and the strands with solid wire. This is shown in Fig. 34. A stranded wire may also be wrapped in the same manner as a solid wire. However, the strands of the wire to be wrapped must be dipped in pure tin in order to bind the strands into the equivalent of a solid wire. A test was made on a No. 24 (0.020″ dia.) wire having seven strands and three twists per inch. The preliminary resistance and aging tests have shown connections of this type to be good.

Enameled wire also has been wrapped. The tool for this purpose has a combination edge at the wrapping tip. Part of the edge has an arc for producing the wrapping tension and the other part has a scraping edge for removing the enamel from the underside of the wire as it is being wrapped onto the terminal.

	1 FAHNESTOCK CLIP	2 PLUG	3 CRIMP	4 WIRE NUT	5 SCREW	6 SOLDERLESS WRAPPED
	NO. 15	0.040″DIA. PIN	0.120″ DIA. x 0.112″LG.	0.350″DIA. x 0.550″ LG.	NO. 4-40 (0.112″)	0.0148″ x 0.062″ TERMINAL
CONTACT FORCE IN POUNDS	1.4	2.2	22	UNKNOWN	135	90
CONTACT AREA IN SQUARE INCHES	0.000079	UNKNOWN	UNKNOWN	UNKNOWN	0.0074	0.0031
CONTACT PRESSURE IN PSI	18,000	UNKNOWN	UNKNOWN	UNKNOWN	18,250	29,000
SPACE 10^6 CUBIC MILS	41	8.78	1.75	52.8	15.6	1.53
ELASTIC ENERGY IN MIL–POUNDS	21	UNKNOWN	UNKNOWN	UNKNOWN	2.77	3.05

Fig. 35 — Comparison of pressure connections for No. 24 (0.020″) wire.

EVALUATION

The solderless wrapped connection has been compared with other pressure connections. (See Fig. 35.) However, when compared with a soldered connection its advantages are as follows:

1. A substantial reduction in wiring defects in manufacture and in service because of:—

   a. Greater uniformity obtained with a calibrated tool.
   b. Less breakage of wires due to handling and vibration.
   c. No solder splashes.
   d. No clippings.
   e. No cold joints.
   f. No rosin joints.
2. Less expensive connection.
3. More compact connection.
4. More clearance between current carrying parts.
5. Easy to disconnect.
6. Saving of tin — a critical material.
7. No contact contamination from soldering fumes.
8. No damage to heat sensitive materials in circuit components.
9. No hazard from hot soldering iron.

SUMMARY

A good pressure connection depends on the amount of elastic energy which can be stored in the mutually stressed members, namely the wire and terminal. If the ratio of elastic energy to the size of the connecting members is favorable, and the contacting areas are sufficiently large, then the connection can be termed good. The solderless wrapped connection when properly proportioned not only meets these requirements, but is uniform in quality and low in cost.

APPENDIX I

EFFECTIVE MOMENT ARM FOR TORSION OF RECTANGULAR TERMINAL

In this appendix the relationship between the wrapped tension in the wire and the twist of the terminal will be analyzed. The structure of the solderless wrapped connection is equivalent to a terminal having springs attached between its edges as shown in Fig. 36. The springs are arranged in such a way as to form a helix of pitch p. Now let:

S_u = torsional stiffness of a unit length terminal

WF = wrapped tension

$2a$ and $2b$ be the sides of the rectangular cross section of the terminal.

$\ell_1 =$ lead of wire helix on long side ($2a$).

$\ell_2 =$ lead of wire helix on short side ($2b$).

$N =$ number of turns of wire

The classical formula for Torsion is

$$\theta = \frac{\text{Torque} \times \text{Length}}{S_u} \tag{1}$$

For an equivalent "spring" attached to the long side the axial length of the terminal is ℓ_1, the torque is WFb^* and hence the deflection $WFb\ell_1/S_u$; for the short side the length is ℓ_2, the torque is WFa^* and the deflection $WFa\ell_2/S_u$. For a complete turn of the helix the length of the terminal is therefore, $2(\ell_1 + \ell_2) = p$, and the torsional deflection is

$$\delta = \frac{2WF}{S_u} (b\ell_1 + a\ell_2) \tag{2}$$

It is logical to assume that

$$\frac{\ell_1}{\ell_2} = \frac{a}{b} \tag{3}$$
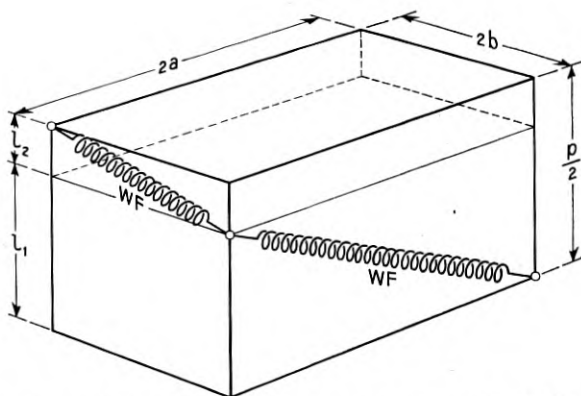
Fig. 36 — Equivalent structure of wrapped connection (only half turn shown).

* This is approximate and holds closely for $\ell_1 \ll 2a$ and $\ell_2 \ll 2b$

Substituting (3) into (2), the deflection per turn becomes

$$\delta = \frac{2ab}{a+b}\, p\, \frac{WF}{S_u} \tag{4}$$

Since all effective turns are similar the total deflection Δ for N turns is

$$\Delta = WF \left[\frac{2ab}{a+b} \right] \frac{(pN)}{S_u} \tag{5}$$

pN is the total length of the terminal. By equation (1) the effective torque must then be

$$\text{Torque} = WF \left[\frac{2ab}{a+b} \right] \tag{6}$$

and hence the effective moment arm, A_e

$$A_e = \left[\frac{2ab}{a+b} \right] \tag{7}$$

Solderless Wrapped Connections

PART II — NECESSARY CONDITIONS FOR OBTAINING A PERMANENT CONNECTION

By W. P. MASON and T. F. OSMER

In order to study the stresses and strains occurring in a solderless wrapped connection, a photoelastic technique using photoelastic bakelite and a photoplastic technique using polyethylene have been used. Polyethylene has a stress strain curve similar to a metal and can be used to investigate strains in the plastic region. Using these techniques, it is shown that the connection is held together by the hoop stress in the wrapping wire. In order to lock this in, a dissymmetry from a circular form has to occur. This may be in the direction of an oval shape or a square or rectangular shape. Sharp corners are preferred since a more definite contact area results. A number of rules are derived for constructing the most satisfactory solderless wrapped connection.

It is shown that the connection between the wire and terminal is intimate enough to permit solid state diffusion, but the strains are not high enough to cause cold welding of the connection. The life of the joint depends on the twin processes of stress relaxation and self diffusion. Stress relaxation occurs at a rate such that half the hoop stress is relaxed in 2500 years at room temperature. This loss of stress is compensated by the diffusion of one part of the joint into the other. Since the activation energies for stress relaxation and self diffusion are approximately equal for most metals, the two effects complement each other and produce a connection which should remain unchanged for times in excess of forty years under any likely ambient conditions.

INTRODUCTION

The solderless wrapped connection described in the paper by R. F. Mallina (see page 525) provides a very satisfactory and economical method for making connections with apparatus terminals when such

connections are properly made. Not all methods of wrapping or all types of terminals are equally satisfactory and it is the purpose of this paper to describe investigations that have been made to determine the necessary conditions for the best wrapped terminal. These investigations include a photoelastic investigation of the stresses in the terminal and a photoplastic investigation of the strains in the outside wrapping wire. A new photoplastic material, polyethylene, has been used which has a stress strain curve similar to a metal and a birefringence proportional to the strain. The use of this material makes possible the evaluation of strains in the plastic region and may find applications in other plastic flow problems such as the extrusion of metals.

Even after such terminals have been satisfactorily made, there remains the question of whether they will have sufficient life to satisfy the requirements of the telephone plant. A design objective for most relays and other switching apparatus of the telephone plant is an uninterrupted trouble-free life of forty years. Hence, unless the connections are to be the limiting factor in the maintenance of the equipment, they also should have a minimum life of forty years under the conditions for which the apparatus is designed. In order to investigate the probable length of life of such connections, theoretical and experimental work on stress relaxation in metals has served as the basis for calculations and tests. These have been extended to the materials and conditions of the wrapped solderless connection and the results indicate that the life should be adequate even under very severe ambient conditions.

PHOTOELASTIC ANALYSIS OF STRAINS IN TERMINALS OF THE SOLDERLESS WRAPPED CONNECTION

In studying the conditions necessary to insure a good solderless wrapped connection, it is desirable to know what strains occur in the terminals and in the wrapping wires and how these vary with the terminal shape, the winding force and other variables entering into the construction of the connection. While some of these strains can be surmised from the winding conditions and the shape of the terminals, it is difficult to obtain any quantitative results by calculations on account of the fact that the desirable terminal shapes are rather complicated and because a good many of the strains are in the plastic region.

To remedy this difficulty, use has been made of a photoelastic and a photoplastic technique. For the inside terminal, all the strains, except at the corners where the wires make contact with the terminals, are elastic and can be approximated with an ordinary photoelastic tech-
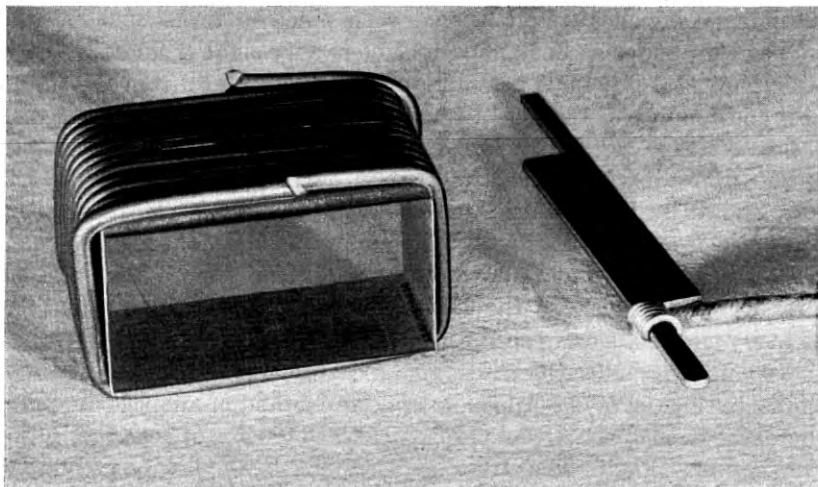
Fig. 1 — Photograph of photoelastic model and solderless wrapped connection.

nique using photoelastic bakelite. Fig. 1 shows one of the photoelastic models as compared with the metal solderless wrapped connection that it simulates. While the ratios of the wire diameter to the terminal dimensions are different in the model from those in the connection, considerable information can be obtained about strains in the terminals and wires from the photoelastic model.

The wrapping gun described in the previous paper puts a tension on the wire. To simulate the tension, the photoelastic model is placed in a chuck, one end of the wire is anchored to the chuck and an appropriate weight is applied to the other end of the wire. The specimen is then rotated by the chuck and a definite number of turns of copper wire are wound around the specimen. The extra wire is then clipped off and it is found that the wire tightly adheres to the terminal in the manner of a metal solderless wrapped connection. The specimen is then polished on the two ends up to the end wires and is put into the polariscope of a photoelastic analyzer. For obtaining the isochromatic lines, i.e., the lines occurring when the ordinary and extraordinary rays differ in path length by a half wavelength or some multiple of a half wavelength, the elements of the polariscope, as shown by Fig. 2, contain a quarter-wave plate before and after the specimen. These have the effect of making the plane wave from the polarizer circularly polarized and eliminate the isoclinic lines which mark the directions of the slow and fast axes of the material. Fig. 3 shows the isochromatic lines for a square

specimen 0.4 inches square wound with nineteen turns of 0.050-inch copper wire under a constant weight of twenty-eight pounds, which is a stress of 14,300 pounds per square inch. It is evident from the sharpness and number of the lines that can be seen that we are dealing with a case of plane stress that can be analyzed by the method discussed in the Appendix. Since the stress strain curve of copper wire has the form shown by Fig. 4 with a yield stress* of about 26,000 pounds per square inch and a breaking stress of about 34,000 pounds per square inch, the applied winding stress is about 42 per cent of the breaking stress. Fig. 4 shows also the recovery measured for copper wire. The recovery curves are quite accurately parallel to each other, but the larger the strain the smaller the percentage recovery. Using the isoclinic lines shown by Fig. 24 of the Appendix and the method of analysis discussed there, the stresses across and perpendicular to the line of "eyes" of Fig. 3 are shown by Fig. 5. The stress perpendicular to the line of eyes measures the total compressive stress put on the terminal by the hoop stress in the wire and from this measurement the average hoop stress remaining in the wire can be calculated as follows. The cross section of which
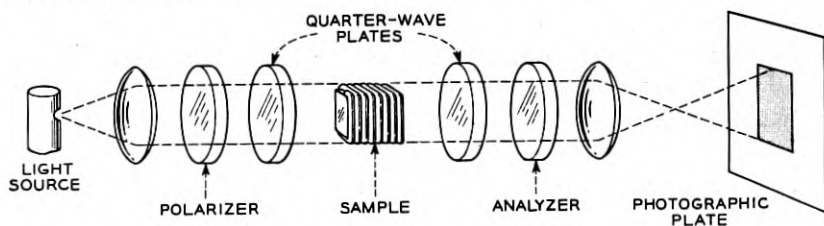
Fig. 2 — Elements of polariscope.

this force is applied is the width 0.4 inches by the length of the specimen 0.95 inches and hence the force applied by all the turns is

$$F = 0.95 \times 0.40 \times 2000 = 760 \text{ pounds.} \tag{1}$$

Since there were nineteen turns of wire wound around the specimen and each turn has two sides exerting a tension on the bakelite, the average tension remaining in the wire, required to balance the compressive stress, is

$$T = \frac{760}{2 \times 19} = 20 \text{ pounds,} \tag{2}$$

---

* In this paper the yield stress is taken as the point of greatest curvature of the stress-strain curve.
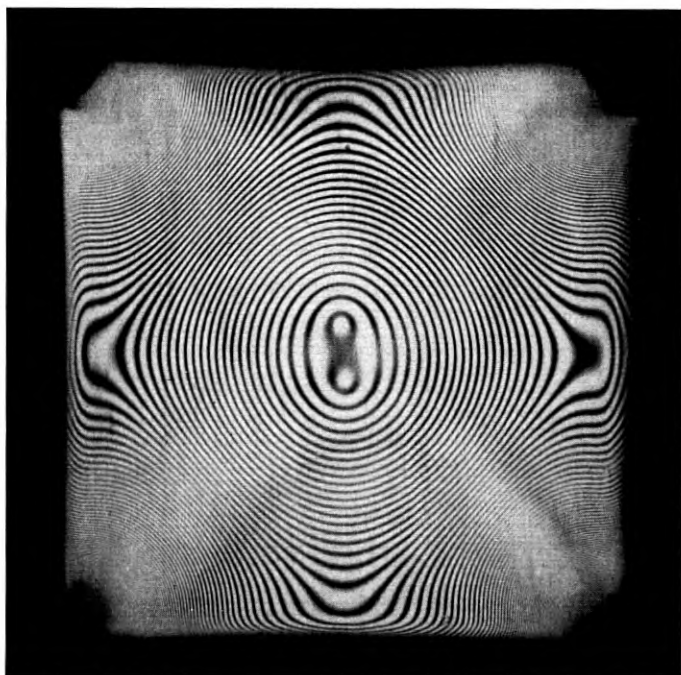
Fig. 3 — Isochromatic lines for a square model 0.4 inches on a side wrapped with nineteen turns of 0.050 inch copper wire with a constant load of 28 pounds.
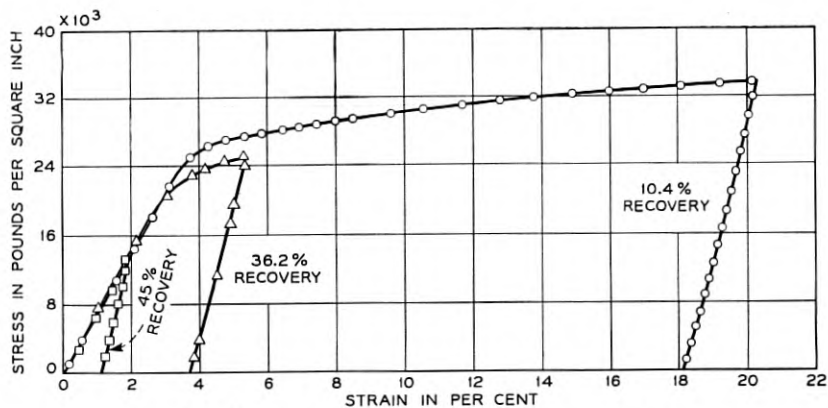
Fig. 4 — Stress-strain recovery curves for copper wire.
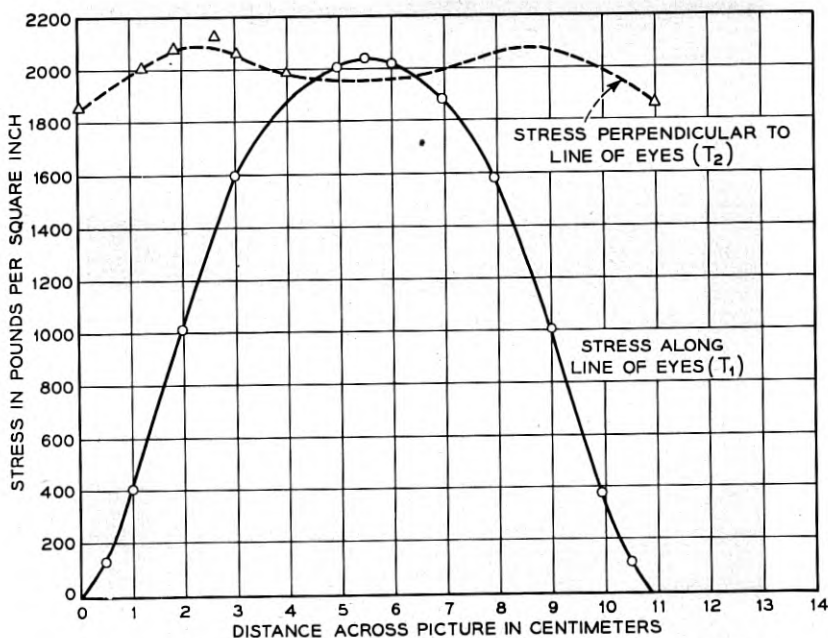
Fig. 5 — Stresses along X and Y directions for a square terminal.

or 71 per cent of the winding stress applied to the wire. The rest of the stress is lost in the contraction that occurs when the wire causes plastic flow to occur at the corners of the specimen. This plastic flow causes the terminal (both bakelite and metal) to flow around the wire and provides an air tight joint which is an essential requirement for a good solderless wrapped connection. That the stress in the terminal is high enough to do this can be seen directly from the photoelastic picture, Fig. 3. Counting the lines as far as they can be distinguished by a microscope, there are 72 lines from the eyes of the picture (which are isotropic points) up to 0.0525 inches from the geometrical corner lines. Using the stress constant of photoelastic bakelite which is 88 pounds per square inch per fringe per inch length along the optic path, this corresponds to a stress per unit area of

$$T = \frac{72 \times 88}{.95} = 6700 \text{ pounds per sq in.} \tag{3}$$

Since the total force put on by the wire is supported by successively smaller cross sections, as one approaches the corners the yield stress of bakelite of 15,000 pounds per square inch will be attained at a radius of

$$\left(\frac{0.0525}{x}\right) \times 6700 = 15{,}000 \text{ pounds/sq in. or } x = 0.0235 \text{ in.} \quad (4)$$

Hence, plastic flow should occur for about 23 mil inches into the plastic. Unwrapping the wires from the terminal, it is found that depressions of this order are cut in the terminals. Since one of the requirements of a better connection is that an air tight bond shall be formed between the wire and terminal, it is obvious that the terminal should have a low enough yield stress so that a sizable groove can be cut in it by the hoop stress of the wire. This rules out such terminals as hardened steel in the most satisfactory connections. It has been found that copper, brass, aluminum, soft iron and nickel silver are soft enough to meet this requirement. Any material with a plastic flow limit in compression, much lower than photoelastic bakelite, would probably have such a deep groove that it would be difficult to maintain the desired hoop stress.

Using the photoelastic technique as a tool, considerable data has been obtained on desirable shapes for the terminal and limiting winding stresses that can be used. One of the most used terminals is the rectangular terminal and Fig. 6 shows a photoelastic picture of a terminal 0.8 inches by 0.4 inches wound with nineteen turns of 0.050-inch copper wire with a winding load of 28 pounds. This figure is particularly easy to analyze for stress across a line half way down the long edge since the stress along this in, the direction of the line varies only a little. Hence,
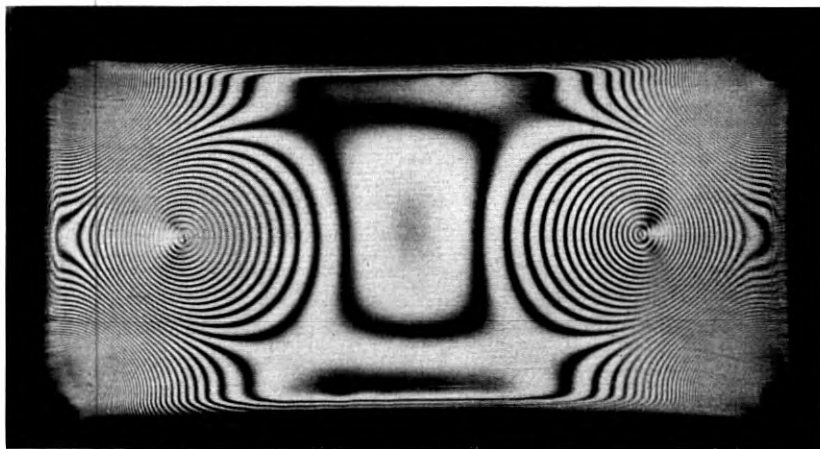
Fig. 6 — Isochromatic lines for a rectangular model 0.4 inches by 0.8 wrapped with nineteen turns of 0.050 inch copper wire with a constant load of 28 pounds.

one can count the number of fringes from the "eye" (which is an iso-tropic point) to the center of the specimen which in this case is eighteen fringes. Using the stress constant, 88 pounds per square inch per fringe per inch along the optic path, the compressive stress normal to the mid line is

$$\frac{18 \times 88}{0.95} = 1670 \text{ pounds/sq in.} \tag{5}$$

and the hoop residual stress per wire is

$$\frac{1670 \times 0.4 \times 0.95}{2 \times 19} = 16.6 \text{ pounds} \tag{6}$$

which is 60 per cent of the winding stress. Hence, the change in shape has not made any appreciable difference in the residual hoop stress. A specimen 0.4″ x 1.6″ was also tried and this had a residual stress of 56 per cent of the winding stress.

In order to determine the number of turns required to make a satis-factory joint, measurements were made of the residual stresses as a function of the number of turns with the results shown by Fig. 7. All of these experiments were made with the same weight, 28 pounds, which results in a stress of 14,300 pounds/square inch. Down to five turns, about 50 per cent of the winding stress is maintained. The results are consistent with assuming that the wire unwinds to the extent of two corners on each end while 60 per cent of the winding stress is main-tained in all the other turns. As seen from Fig. 8, this is what one might expect, for when the constant tension is released, recovery will cause the
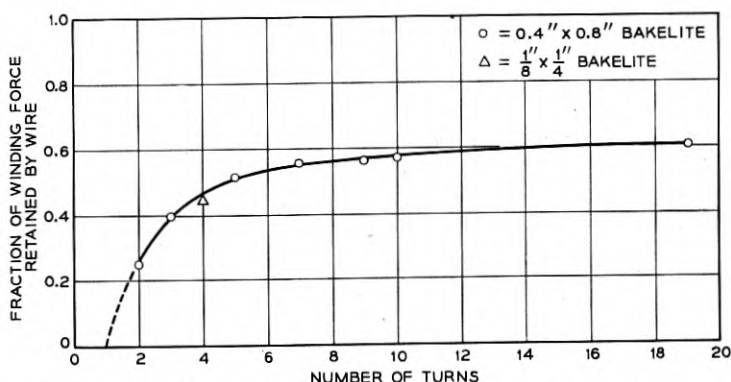
Fig. 7 — Residual hoop stress as a function of number of turns.

first corner to bend out and lose contact with the terminal. When the second corner attempts to unwind, it pulls the first corner up against the sample and no unwinding beyond the second corner can occur. In order to obtain the most satisfactory connections, at least five or six turns should be used.

The fact that the first two corners on each end do not make close contact to the inside terminal produces a very beneficial result when the connection is subject to vibration due to the handling and operation of
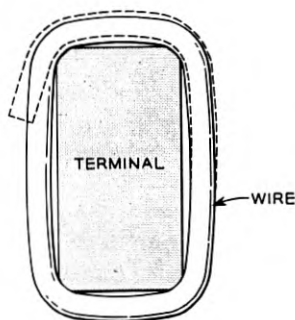
FIG. 8 — Locking in effect in a rectangular terminal which allows unwrapping at only two corners on each end.

a relay. In a soldered connection, large bending strains are caused by vibrations at the point of contact between the wire and the solder and in standard vibration tests when large 60-cycle vibrations are impressed upon the wires, the wires fatigue and break off at the point where the wire enters the solder in times in the order of fifty hours. Similar tests have been carried out for solderless wrapped connections and up to times of 2000 hours and longer no breaks have occurred. This is due to the fact that a bending strain is not enhanced by a sharp discontinuity as it is in the soldered connection and strains for a given vibration amplitude should be less than half as large as those for a soldered connection. Since the relation between fatigue and strain is such that a reduction of strain of two to one or more caused an increase in the number of cycles before breakage of factors of 1000 or more, the increased life under vibration for the solderless wrapped connection is not surprising.

Next a series of measurements were made on the value of the winding stress necessary to preserve a hoop stress in the wire. This was measured by winding wires with different weights on photoelastic samples and measuring the residual hoop stress by the technique described above. This was done for both copper and aluminum wires with the re-
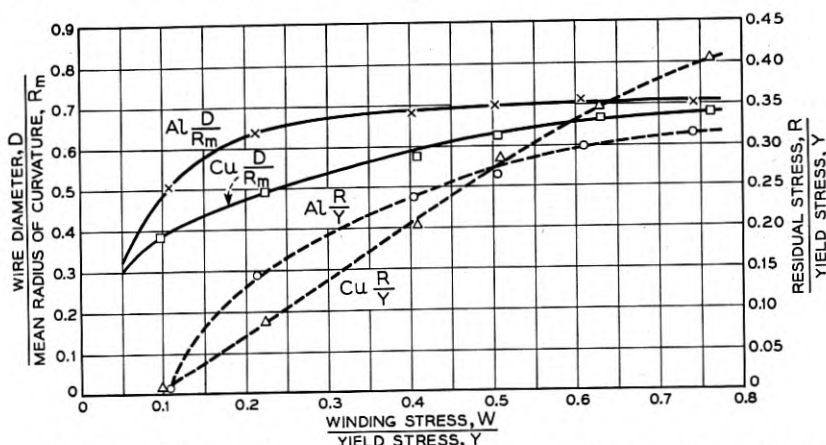
Fig. 9 — Relation between remanent hoop stress, ratio of wire curvature to wire diameter and the applied winding stress for copper and aluminum wire.

sults shown by the dashed lines of Fig. 9. To obtain a hoop stress greater than zero, the ratio of winding stress to yield stress must be greater than 0.1. To obtain a hoop stress that is 0.2 of the yield stress, a winding stress of 35 per cent of the yield stress has to be employed. This is the value recommended to give the most stable terminal. The winding stresses were carried up to 0.8 of the yield stress. At this stress the copper wire at the corners tends to draw down to a low value and may break when it unwinds and hence this is probably the upper limit for winding stresses. The radius of curvature of the middle of the wire, as it is bent around a corner was also measured from photographs similar to that shown on Fig. 6 and the ratio of the wire diameter to the mean radius of curvature is shown by the solid lines of Fig. 9. This is an alternate way of specifying the necessary winding force which may be useful for other shapes of terminals.

While square and rectangular terminals are very satisfactory shapes for the inside terminal, they are not the only ones that can be used. A number of coined and U shaped terminals are in general use as discussed by Mallina. In order to investigate the necessary requirements for such shapes, a number of experiments have been made on circular and elliptical terminals. When a wire under tension is wound around a rod of circular cross section, there are two sets of opposing stresses, one of which tends to make the helix smaller and the other to make it larger. As shown by Fig. 10, the tension in the wire tends to make the helix hug the cylinder while the bending strains introduced by the wrapping of the wire around the cylinder tend to make the helix open up when the constant stress is released. A number of experiments were made on

wrapping copper wire 20 mil inches in diameter on a steel cylinder having a diameter of 0.124 inches. In all cases even up to stresses of 80 per cent of the yield stress, the helix failed to grip the cylinder. Some further measurements were made with smaller sized inner cylinders down to 20 mil inches in diameter. At this small radius the wire barely gripped the cylinder and it took about 70 grams stripping force to pull the wrapped wire off the cylinder. As seen from Fig. 11, the normal force per unit length against the cylinder is balanced by the hoop force in the wire according to the equation

$$\theta F_N r = 2F_H \sin \frac{\theta}{2} \qquad \text{or} \qquad rF_N \doteq F_H . \tag{7}$$

The stripping force SF, when the wire does not dig into the terminal, should be equal to

$$SF = 2\pi r n F_N f = 2\pi n f F_H , \tag{8}$$

where n is the number of turns, f the coefficient of friction which is about 0.15 to 0.2 between metals. For a stripping force of 70 grams for 6 turns, the remaining hoop stress is equal to

$$F_H = 9.3 \text{ grams.} \tag{9}$$

Since the wire was wound with a 700-gram force, it is evident that only about 1 per cent is maintained in the wire, which is entirely inadequate.

In order to obtain a good wrapped connection with high hoop stress in the wires, some means has to be employed to eliminate the unwrapping effect of the strain due to bending. This can be accomplished by changing the shape of the terminal from a circular cylinder to a dissymmetrical shape. For then, as shown by Fig. 8, the tendency to un-
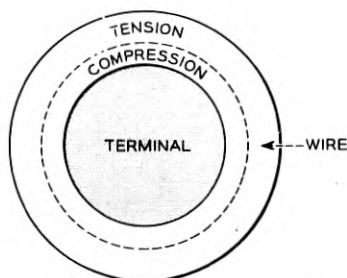
Fig. 10 — Strains in a wire wrapped around a circular terminal.
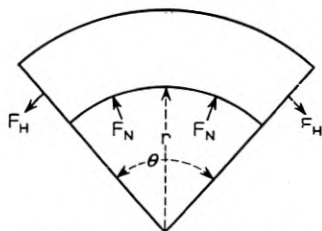
Fig. 11 — Relation between hoop stress and stripping force for terminal not indented.

wind is opposed by the locking in effect of the dissymmetry which in the most preferred types of terminals, as in Fig. 8, comprises abrupt changes in direction around the periphery of the terminal cross section. Some experiments were made with a cylinder whose cross section, as shown by Fig. 12, consisted of parallel sides with a semi-circle on each end. When the length A was twice the width, the top curve of Fig. 12 shows the stripping force as a function of the winding force. Assuming that most of the grip occurs on the two semi-circular surfaces, the hoop
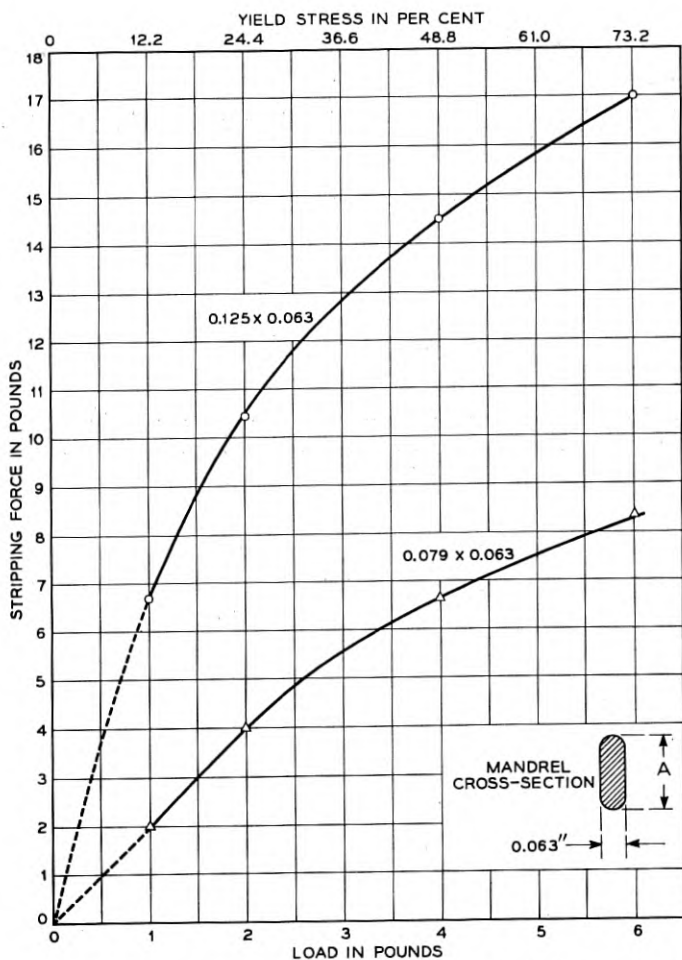
Fig. 12 — Stripping force as a function of winding load for an elliptical type terminal.
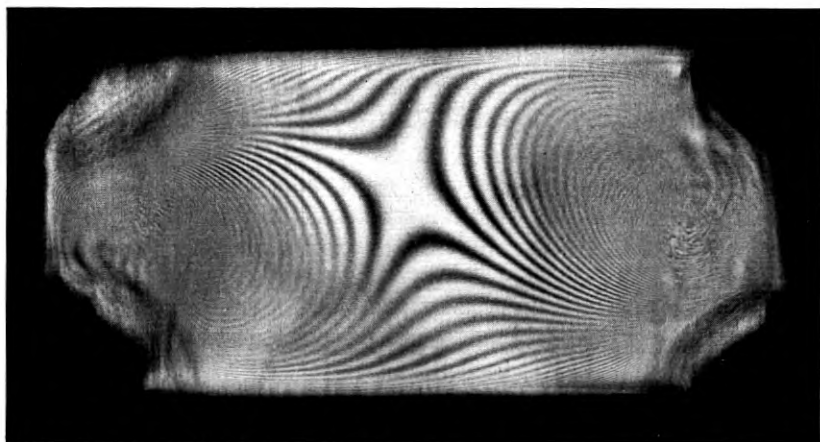
Fig. 13 — Photoelastic picture of a terminal when the outside wire diameter approaches the terminal diameter.

stresses indicated are in the order of 50 to 60 per cent of the winding stresses in agreement with the photoelastic experiments. Ratios of 5 to 1 on the length-width ratio were also tried with substantially the same result. To determine how much dissymmetry is necessary to lock in the bending stresses, a ratio of 1.25 to 1 of the length to width was tried with the result shown by the lower curve of Fig. 12. The stripping force for this ratio is about half that for the larger ratios. The conclusion is that the stripping force decreases as the symmetry increases but that a small deviation from circular symmetry is sufficient to lock in the bending stress. However, for the more satisfactory connections, other factors such as adequate intimate gas tight contact areas indicate that the required dissymmetry involves abrupt surface changes in the nature of edges having appreciable penetrating power with respect to the wire. Therefore, while a terminal like that of Fig. 12 may have the ability to lock in the bending stresses, it may not, for other reasons, be the better terminal to use.

All of the photoelastic and other types of stress measurements were made for terminals that have a large stiffness in torsion and as seen by the photoelastic pictures, the stresses are nearly plane stresses, i.e., they are all tensions, compressions or shears in a plane perpendicular to the axis of the connection. If, however, the torsional stiffness becomes low, another type of deformation can take place, namely, a twist of the whole terminal due to the torque put on by the helical form of the winding. Fig. 13 shows a photoelastic picture of the strain in the terminal when

the size of the outside wire is comparable to the size of the terminal
and it is obvious that a twist in the terminal is occurring. This is a case
of three dimensional stress rather than plane stress and cannot easily
be analyzed from the photograph. The photograph does, however, show
a twist of the terminal.

The twisting strain can be most easily analyzed by taking a long
section of terminal of low torsional stiffness, winding 100 or more turns
on the terminal and measuring the angle of twist as discussed in the
paper by Mallina. Calculations by Love[1] show that a twist in an ellipti-
cal section with its length along the Z axis introduces shearing strains
in the X, Z and Y, Z planes, i.e., $e_{ZX} = S_5$ and $e_{YZ} = S_4$ shearing strains
equal to

$$S_5 = -\tau \left(\frac{2a^2}{a^2 + b^2} \right) y; \qquad S_4 = \tau \left(\frac{2b^2}{a^2 + b^2} \right) x, \qquad (10)$$

where $2a$ is the diameter of the ellipse along the X direction and $2b$
the diameter of the ellipse along the Y direction and τ the angle of twist
in radians per centimeter. For the terminal with 100 turns of 0.020 mil
copper wire discussed by Mallina whose data are given by Fig. 13, 45°
angle of twist occurs in 2 inches giving a value of $\tau = 0.157$. This causes
a shearing strain of about 2 per cent in the worst case which is enough
to cause a considerable permanent set. While this twist is useful in
studying stress relaxation in the wire, it is undesirable for a solderless
wrapped connection to have too much twist since it may cause the ter-
minal to twist off in the winding process. According to the data of Fig. 13
of Mallina's paper, no terminal set occurs for nickel silver if the twist in
radians per centimeter is less than 0.09 which corresponds to a maxi-
mum shearing strain in the X, Z plane of 1.1 per cent. Hence in order
to avoid excessive permanent set and twisting off of the inner terminal,
the size and shape of the terminal should be controlled so that shearing
strains due to twisting should be less than 1 per cent. For standard
shapes such as rectangles and ellipses formulae are available to relate
the maximum strain to the dimensions of the terminals and the moment
due to the winding stress. For a given wrapping tension, this moment
can be calculated from Appendix I of Mallina's paper.

Summarizing the results of this section, the necessary conditions
that the terminal should meet are:

1. The wrapped connection is held together by the hoop stress in the
outside wrapping wire. This can be locked in if the terminal has a dis-
symmetrical shape in which the length-width ratio is 1.5 or greater,

[1] Love, Theory of Elasticity. Chap. XIV, p. 310, 4th Edition, Cambridge Uni-
versity Press.

or if some regular shape such as a square, rectangle or rhombus is used. Sharp corners are helpful since a sharp bend in the wire occurs around them.

2. The material of the terminal must be strong enough so that the wire will not deform or cut through the terminal but must be plastic enough so that an appreciable groove can be cut in it by the hoop stress of the wire, in order that an air tight connection shall be made. The most satisfactory metals are brass, copper, soft iron, nickel silver and aluminum.

3. In the most satisfactory solderless wrapped connection, the wrapping wire unwinds to the extent of half a turn on each end and about six turns or more are required to make a good connection.

4. To maintain sufficient hoop stress, the constant wrapping stress should be from 0.2 to 0.7 of the breaking stress of the wire.

5. Shearing strains in planes parallel to the axis of the terminal should not exceed 1 per cent in order to eliminate terminal set.

PHOTOPLASTIC ANALYSIS OF STRAINS IN WIRES OF A WRAPPED SOLDER-
LESS CONNECTION

All of the strains in the inner block or terminal are elastic except at the corners. Hence, for the interior terminal, photoelastic bakelite is a satisfactory material for strain investigations. However, the outer wire is necessarily stressed beyond its elastic limit and ordinary photoelastic techniques cannot be applied. In order to see if the wire strains could be studied with photoelastic bakelite, some time was spent in heating rods to a temperature for which they become elastic, winding them under a stress and cooling under the applied stress. Although the winding process was carried out successfully several times, the bakelite rod always broke on cooling. This appeared to be due to the fact that bakelite is nearly linear up to the breaking point, i.e., it suffers from brittle fracture and does not simulate a metal in this respect.

Some measurements had previously been made at the Bell Laboratories and in England[2] on polyethylene which indicated that it had properties similar to a metal in the plastic range. Stress-strain curves up to 15 per cent strain are shown by Fig. 14, and it is evident that on the ascending part, the curve is very similar to that for copper or soft iron. On the relief from stress, however, a considerably larger recovery is obtained than for a metal. This material is fairly transparent and the lower curve marked R/t shows the relative retardation for a 5461 A°

² Miss S. M. Crawford and Dr. H. Kolsky, Stress Birefringence in Polyethylene. Proc. Phys. Soc., Section B, London, **6,** Part 2, pp. 119–125, Feb. 1, 1951.
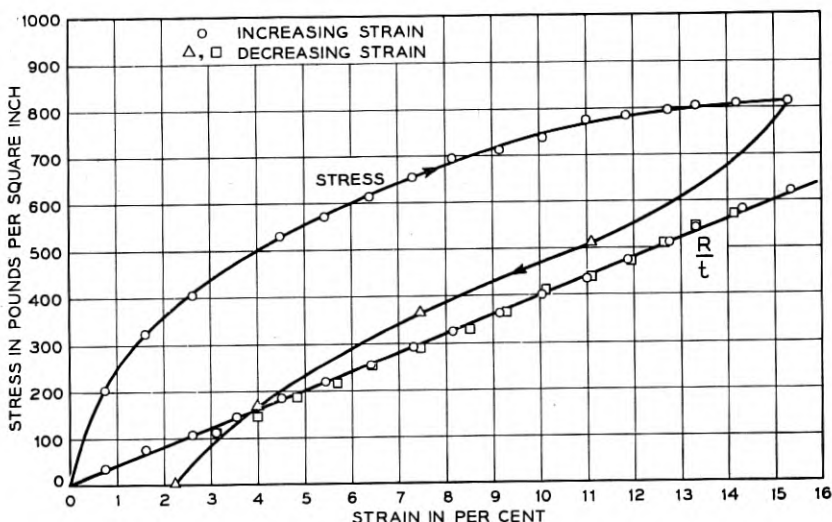
Fig. 14 — Stress and birefringence strain characteristic of quenched poly-ethylene.

mercury line. This relative retardation R/t is related to the number of fringes N and the wavelength λ by the equation

$$\frac{R}{t} = \frac{N\lambda}{t} \tag{11}$$

and hence for a 1 per cent strain there are 7.3 fringes for a 1-cm path length in the optic direction. It will be observed that for both increasing and decreasing strains, the birefringence is directly proportional to the strain. Hence, by using polyethylene it appeared possible to measure the strain even in the plastic region.

The first experiment tried was to wrap a square metal rod with a polyethylene "wire" one-sixteenth inch in diameter with a wrapping stress about half the yield stress. It was found, however, that the wire sprang off the metal rod when the constant stress was released. This is due to the fact that the polyethylene has considerably more re-covery than the metal wire and the dissymmetry is not sufficient to lock in the bending stress. This shows that one of the requirements of the wire is that the recovery shall not be too large.

If we are to use polyethylene as a photoplastic material, it is neces-sary to simulate the unloading curve as well as the loading curve. This can be done by heating up the polyethylene when it is wound under a

load, cooling under a load, and then removing the load at room temperature. Fig. 15 shows the loading and unloading curves for polyethylene as a function of temperature. The stress required to produce a given strain decreases very rapidly as the temperature increases, although the recovery remains about the same irrespective of the temperature. Suppose now that we apply a load and cool the polyethylene down to room temperature maintaining the strain. When the weight is taken off, the unloading curve will parallel that of the 20°C curve, and 42 per cent recovery will be obtained at 50°C and 12 per cent for 90°C.

In order to see if a wrapped joint would be simulated by this means, a one-quarter inch rod of polyethylene was heated up to 97°C, was wound with a one pound winding weight and was cooled to room temperature with the weight attached. This was sufficient to prevent the "wire" from unwrapping and when the weight was removed, the polyethylene "wire" gripped the metal closely and formed a bond similar to the wrapped connection. In order to analyze the strain, one has to cut a section through the center of the wire and put the section in the polariscope. First, a number of unstrained polyethylene samples were cut by various techniques and it was found that if they were cut with a
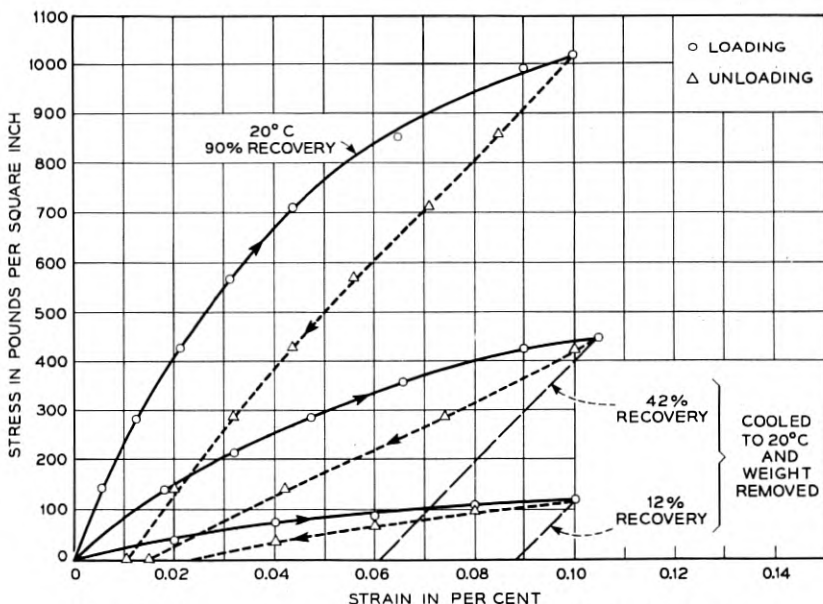
Fig. 15 — Stress strain recovery curves for quenched polyethylene.
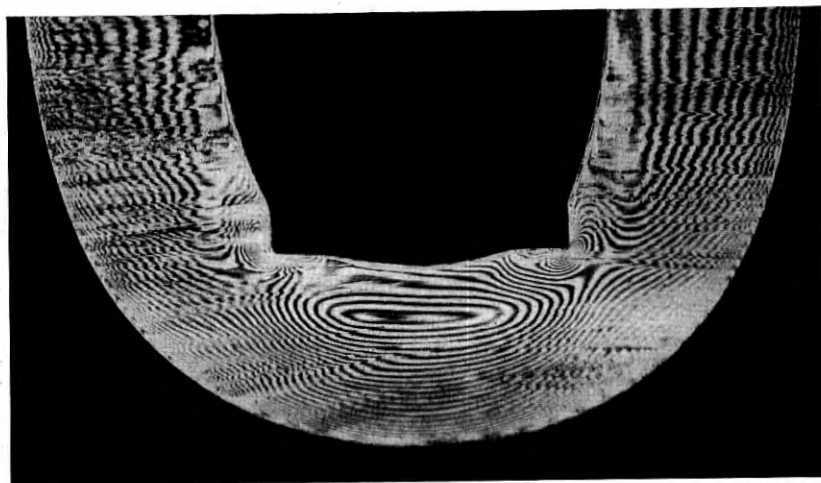
Fig. 16 — Photoelastic picture of a polyethylene "wire" wrapped at 97°C around a rectangular terminal and cooled off under the applied load.

jewelers saw with a good deal of set to the teeth, no strains were introduced in the sawing process. Using a jig with two parallel guides a sample 0.040 inches thick was cut through the polyethylene wire. Taking one in the center of the wrap, a photoelastic picture was taken with the result shown by Fig. 16 with an enlargement of one corner shown by Fig. 17.

The easiest parts to analyze are the strains in the two legs of the sample since the strains are similar to those for a bent section. The long leg which was twice as long as the short leg has its zero order fringe along the inside edge. There are twelve lines to the outer edge which corresponds to a tensile strain of 16 per cent with an average tensile strain of about 8 per cent which is about the strain caused by the loading weight. In the short leg the zero order fringe is the inside oval and the strain is about 11 per cent compressive at the inner edge of the segment and about 38 per cent tensile at the outer edge. These values are consistent with the radius of curvature that the wire is bent around for it can be shown that the strain in a wire of diameter d bent around a cylinder of diameter D without tension is equal to

$$S = \frac{2\rho}{D + d}, \tag{12}$$

where ρ is the radial distance measured from the center of the wire outward, D the diameter of the cylinder and d the diameter of the wire. If ρ is positive or the point is outside the center line, the strain is posi-

tive or tensile, while if ρ is inside the center line the strain is negative or compressive. From measurement of Fig. 16, it appears that the equivalent inner cylinder that corresponds to the radius of the center of the wire is about 3.0 times the wire diameter and hence the strain should be 25 per cent tensile at the outer edge and 25 per cent compressive at the inner edge. The addition of a tension of 13 per cent due to the winding force makes the outer strain 38 per cent and the inner one 12 per cent compressive which are close to the values found. This indicates that the tensile strain is somewhat higher in the short leg than in the long one. The same tensile strain occurs around the outside periphery of the wire opposite the corners of the terminal, but a very high

Fig. 17 — Enlargement of one corner of the photograph of Fig. 16.

point of compression develops just below the point of contact between the wire and terminal. The distribution of compressive and tensile strains in the wire is shown pictorially by Fig. 18. The distribution of strains in a metal wire can be considered to be quite similar except that the tensile strain due to winding should be only 2 to 3 per cent as seen from Fig. 4, while the strains due to bending may be even higher, for as seen from Fig. 9 the ratio of D/d may be in the order of 1 and the bending strains may be as high as 50 per cent.

In order to specify the properties that the wire must have in making a good solderless connection, experiments have been made on how much recovery can be tolerated in the wire. Since it is difficult to get a series of metal wires having different amounts of recovery, the technique was resorted to of heating polyethylene to a definite temperature, winding under a load equal to half the yield stress at the winding temperature, and cooling to room temperature under the load. As shown by Fig. 15, known recoveries can be obtained in this way. It was found that the largest amount of recovery that could be tolerated to make a joint at all was 20 per cent while a reasonable hoop stress was not obtained until the recovery was less than 10 per cent.

In summary, the necessary conditions that the wire should fulfill are:

1. Since strains of 50 per cent may be encountered in bending wires around sharp corners, wires should be used which have a large difference between the yield strain and the breaking strain. Copper, aluminum and soft iron are materials of this class while phosphor bronze and music wire are not as satisfactory.
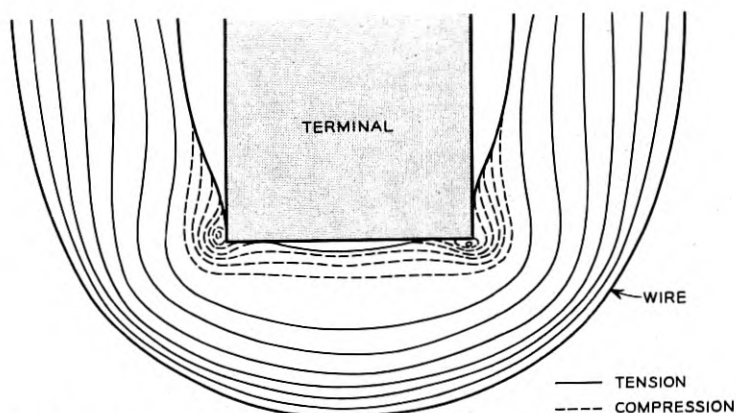
Fig. 18 — Distribution of compressional and tensile stress in the wire of a wrapped solderless connection.

2. The amount of recovery from strains of 20 to 50 per cent should not be greater than 20 per cent.

PERMANENCE OF WRAPPED SOLDERLESS CONNECTIONS

By the photoelastic, photoplastic and strain analysis of the previous two sections, it has been demonstrated that the wrapped solderless connection is held together by the hoop stress in the outside wire whose value is determined by the winding stress and the locking in effect dependent on a dissymmetry of the terminal. The high stresses cause plastic flow in the wire and terminal in such a manner that the two materials flow together and produce an intimate air tight joint. The intimate nature[3] of this contact has been demonstrated by dip coating nickel silver terminals with pure tin and wrapping them with cleaned bare copper wire. The wrapped terminals were then placed in a glass tube, evacuated, sealed off and heated for 400 hours to 180°C (37°C below the melting point of tin). The samples were then removed, mounted vertically, polished, etched and examined microscopically for distinguishing constituents. It is believed that if such a constituent appeared on the originally bare copper wire after such treatment, that the contact was sufficiently intimate to permit solid state diffusion. A section of the wire in contact with the corner is shown by Fig. 19. The copper is seen
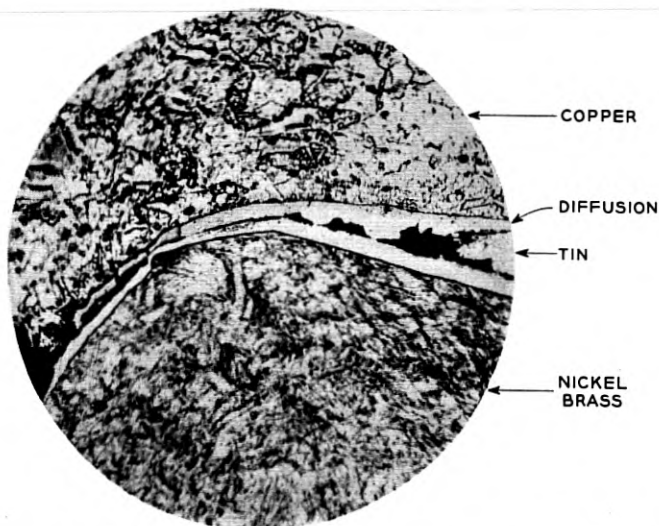
Fig. 19 — Solid state diffusion of tin into copper in a wrapped solderless connection.

[3] This experiment was conducted by G. S. Phipps.

to have a heavy layer of tin constituent at the contact surface. From the test, it can be concluded that wrapped connections are sufficiently tight to allow solid state diffusion and are therefore good electrical contacts.

On the other hand the contacts are not welded contacts such as occur when two pieces of aluminum are cold pressed together with strains in excess of 75 per cent. This is shown experimentally by the simple process of unwinding the wire from the terminal which takes place with no excess force when the wire is removed from a terminal corner. Since the strains at the points of contact do not exceed 30 to 40 per cent, one would not expect cold welding. It is possible, however, that in tinned terminals, some long time diffusion takes place at room temperature in the manner demonstrated at higher temperatures by the data of Fig. 19. This would occur very slowly and cannot be relied upon solely to maintain the contact.

Hence, it appears that long life in the connection depends on maintaining sufficient hoop stress in the wrapping wire to keep the elements of the connection sufficiently tightly pressed together so that no corrosion can occur in the connection in such a manner as to interrupt the electrical continuity. This is a problem in stress relaxation rather than creep. Stress relaxation is intrinsically a simpler phenomenon since the major fraction of stress that can be relaxed will be relieved through viscous flow in previously formed slip bands or along grain boundaries, and no generation of new slip bands is required. However, under ordinary creep conditions, an increase in stress is presumably attended by the generation of new slip bands. It appears likely then that stress relaxation phenomena even at quite high stresses should more nearly follow the conditions that have been established for low stresses than would be the case for creep phenomena.

A good deal of work has been done on stress relaxation at low stresses, particularly by Zener[4] and his coworkers, and this will be briefly reviewed. According to these studies, stress relaxation can be caused by several mechanisms including stress induced migration of impurities in the metal, viscous behavior of slip band material and the viscous behavior of grain boundaries. At the common junction between the two metal grains, there is an amorphous layer of material which acts as a viscous medium, i.e., if there is a shearing stress applied across it, the two grains will move with respect to each other with a velocity

[4] Zener, C., Elasticity and Anelasticity of Metals. University of Chicago Press, 1948. T'ing-Su Ké, Experimental Evidence of the Viscous Behavior of Grain Boundaries in Metals. Phys. Rev., **71**, No. 8, pp. 533–546, April 15, 1947. T'ing-Su Ké, Anelastic Properties of Iron. Tech. Publication No. 2370, Metals Technology, June, 1948.

$$V = \eta T/D, \tag{13}$$

where D is the thickness of the layer, η the coefficient of viscosity and T the shearing stress. Hence no matter how small the shearing stress, one grain will move with respect to the other in a finite time. The amount that the grains can move is limited by the necessity of making the grain boundaries fit. According to Zener, the situation is analogous to the case of a jigsaw puzzle in which the overall configuration possesses rigidity in spite of the fact that no shearing stress exists between adjacent pieces. Zener has calculated that the ratio of the relaxed stress to
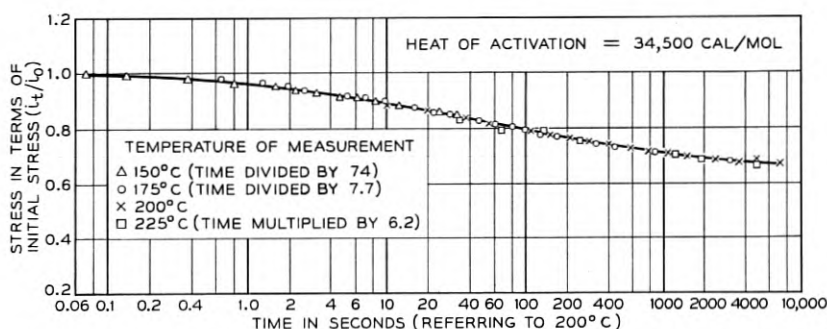
Fig. 20 — Stress relaxation in aluminum at three temperatures for strains less than 10^{-4} (after Ké).

the initial stress is equal to

$$\frac{T_R}{T_I} = \frac{\frac{1}{2}(7 + 5\sigma)}{7 + \sigma - 5\sigma^2}, \tag{14}$$

where σ is the value of Poisson's ratio. For values of Poisson ratio of from 0.25 to 0.5 this ratio lies between 0.595 and 0.76.

Fig. 20 shows measurements of stress relaxation plotted against time for aluminum for three different temperatures. These were obtained[4] by twisting an aluminum wire through a definite angle and observing the force required to hold it at this angle as a function of time and the temperature of the wire. All the curves can be made to coincide by multiplying the times by different factors. If we define the relaxation time τ as the time required to relax half of the variable component of stress, i.e. $\frac{1}{2}(1 - 0.67)$ of the stress, this relaxation time fits an equation of the form

$$\tau = Ke^{H/RT}, \tag{15}$$

where K is a constant, H an activation energy, T the absolute tempera-

ture in degrees Kelvin and R the Boltzman constant for one gram mole of the material. R is closely equal to 2 calories per degree K. Hence, if H is expressed in calories per gram mole and the value of K is obtained to fit equation (15), we have

$$\tau = 9.2 \times 10^{-15} \times e^{\frac{34,500}{2T}} .$$

The constant K is close to that given by the Langmuir-Dushman[5] theory

$$K = \frac{hN}{H} = \frac{6.62 \times 10^{-27} \times 6.06 \times 10^{23}}{34,500 \times 4.187 \times 10^{7}} = 2.7 \times 10^{-15}, \quad (16)$$

where h = Planck's constant equal to 6.62×10^{-27} ergs, N is Avogadro's number equal to 6.06×10^{23} and H is the activation energy expressed in ergs. Su Ké[4] has shown that the activation energy for grain boundary slip is essentially the same as for self diffusion and for creep.

Similar results have been found for α-brass and α-iron. These have activation energies shown by equation (17)

$$\alpha\text{-brass 41 kilocalories per mole,}$$
$$\alpha\text{-iron 85 kilocalories per mole.} \tag{17}$$

Although measurements have not been made for copper, the activation energy of self diffusion is about 57.2 kilocalories[6] per mole, but 39.9 kilocalories for the principle impurity silicon.

All of these measurements were made for strains under 10^{-4}, and the question arises as to whether these concepts are valid for the much higher strains experienced in the wrapped solderless joint. From the photoelastic pictures, Figs. 16 and 17, it is obvious that the greatest stress inhomogeneity occurs in the neighborhood of the corners and flow will take place in such a way as to relieve the high stress concentration. This will have the effect of making the terminal and wire mate even closer and may result in a slight transient lowering of the hoop stress. After the initial formation, however, it will be the long time relaxation of the hoop stress in the wire that determines the lasting quality of the joint.

As discussed previously, the twist that the terminal takes is determined by the mean value of the hoop stress in the wire, and any relaxation in this hoop stress can be studied by observing the angle of twist as a function of time and temperature. By using a long terminal wound

[5] S. Dushman and I. Langmuir, Phys. Rev. **20**, (1922) p. 113, 1922.
[6] Zenner, Elasticity and Anelasticity of Metals. Table 12, p. 98, Chicago University Press.

with 100 turns or more of copper wire, a twist of 50° or more can be obtained which is sufficient to measure.

In order to test first the relaxation in the copper wire alone, the inner terminal was made of clock spring steel 0.0124 inches by 0.062 inches. This was wound with 100 turns of 0.020 inch copper wire tensioned at 2.87 pounds (9,300 pounds per square inch). A twist of 25° was obtained which is sufficiently large to measure. If one observed the angle after transient creep has occurred, the angle decreased on the average about 17 per cent in the first month as shown by the circles of Fig. 22. The values agree with the solid curve which was established by relaxation measurements as a function of temperature as discussed in the next paragraph. At room temperature, further decreases in the angle of about 10 per cent were observed out to times in the order of a year.

If, however, the wrapped connection was heated up, a faster relaxation of the hoop stress occurred. Fig. 21 shows the ratio of angle measured to the initial angle as a function of time when the connection is subjected to a temperature of 200°C. As shown by the dashed line, which is a plot of the exponential equation

$$R = e^{-\alpha t} \text{ where } \alpha = 2 \times 10^{-4}, \tag{18}$$

this is not a single relaxation of the type found for grain boundary motion but is a sum of effects occurring with different activation ener-
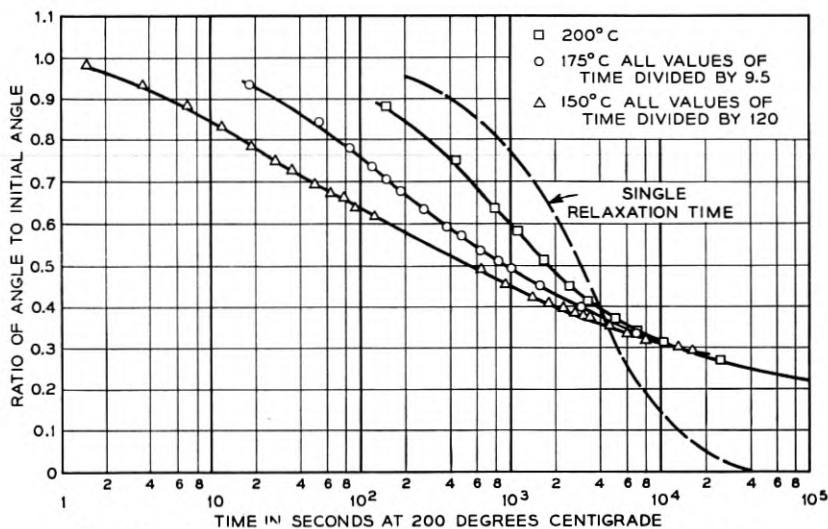
Fig. 21 — Ratio of angles of twist at time shown to initial angles.

gies. Furthermore, the change is not limited to two-thirds as in the case of grain boundary motion, but is much more tending toward a value of 0.2 for very long times. It appears that several processes are involved in addition to grain boundary motion. These are probably connected with slips along slip planes which occur with a lower activation energy when the stress is high. As the slip increases, strain hardening occurs with a resultant increase in activation energy until the activation energy of grain boundary movements is reached. For 0.3 relaxation and lower, the activation energy remains constant and equal approximately to 40 kilocalories per mole — the self-diffusion value — as can be calculated from the 150°C, 175°C and 200°C relaxation curves of Fig. 21. These curves show that the activation energy varies from 12.5 kilocalories at 0.9 relaxation, to 40 kilocalories for long time effects.

Similar measurements have been made on nickel silver terminals which are the terminals actually used and as seen from Fig. 21 of Mallina's paper, these agree quite well with those measured for spring steel terminals. The nickel silver terminals had the dimensions 0.0148 inches by 0.062 inches. A twist of 46° was obtained for 100 turns of 0.020 inch copper wire with 2.87 pounds winding stress applied. At this angle of twist, a permanent set of 19° occurred when the outside wire was unwound. If we subtract that value from the initial twist, the time-angle curve is very similar to that for spring steel and indicates that no additional relaxation occurs in the nickel silver terminal.

The conclusion from these experiments is that stress relaxation for the value of strain used in the wrapped solderless connections follows a similar activation energy pattern to that followed for smaller strains except that instead of a single process with a single activation energy we are dealing with many separate processes having an activation energy range from 12.5 kilocalories to 40 kilocalories. For each stage of the process a different activation energy is effective. For example, for a ratio of relaxed stress to initial stress of 0.9 the curves of Fig. 21 indicate an activation energy of about 12.5 kilocalories. With this value of activation energy the time required to relax this amount of stress at room temperature of 25°C (77°F) is 2.98×10^5 seconds or 0.0095 years as shown by Fig. 22. For a ratio of relaxed to initial stress of 0.8, the activation energy is 15.3 kilocalories and the time at room temperature is 0.126 years. Similar values can be calculated for the other relaxation ratios and the complete relaxation ratio and time curves are shown by Fig. 22 for temperatures of 77°F and 135°F. To reach a value of 0.5 of the initial hoop stress requires 2500 years at 77°F and about forty years at 135°F. The circles show the measured values at 77°F carried out in a tempera-
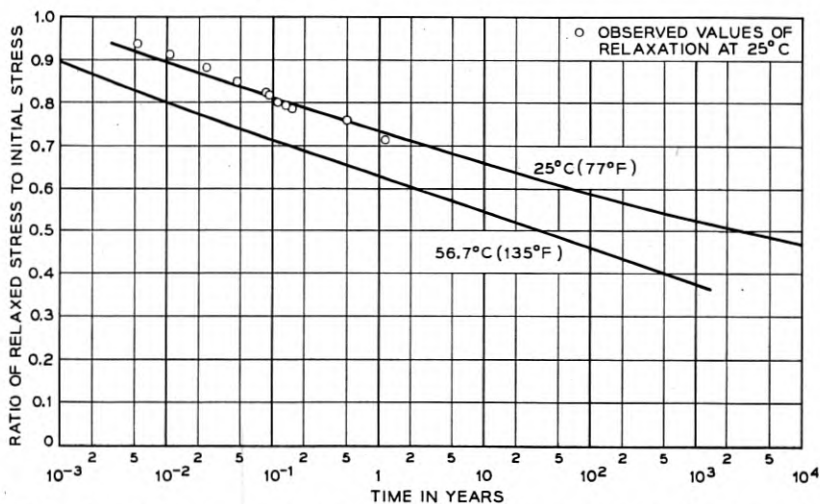
Fig. 22 — Aging at room temperature and at 135°F plotted as ratio of angle to initial angle as a function of time.

ture controlled air conditioned room and these agree well with the calculated values. 135°F is in general the maximum temperature that wrapped solderless connections will be subjected to. For cases of very high temperature it is planned to use copper covered soft iron wires for the wrapping wires since, as shown by Equation (17), iron has a much higher activation energy than copper and can be expected to maintain its hoop stress for forty years even under an ambient temperature of 200°C.

The question arises as to whether the hoop stress of a small part of the initial value, that has been shown to continue for a long time by the data of Fig. 21, is sufficient to maintain a good contact. This question may be important if aluminum is to be used as a wrapping wire since with an activation energy of only 34.5 kilocalories, 0.5 of the hoop stress will be maintained for forty years only for temperatures lower than 100°F. The problem then is whether corrosion can occur between the wrapping wire and the inside terminal when the relaxable component of hoop stress has been relaxed. A very sensitive test for this question is obtained by winding aluminum wire on an aluminum terminal since if any break occurs in the contact between the wire and the terminal, oxidation of the aluminum surface takes place very rapidly and should affect the resistance of the solderless connection. Accordingly, a number of aluminum-aluminum solderless connections were wound up and their resistances were measured. They were then put in an oven and heated to a temperature of 200°C for twelve hours, which was suffi-

cient to relax all the stress that can be relaxed. On remeasuring the resistance, it was found that there was no change within the experimental error of 1 per cent, which corresponds to a resistance of 3×10^{-5} ohms. A similar result is found by studying the corrosion of surfaces of copper and nickel silver in solderless wrapped connections when they are fully relaxed and subjected to a corrosive atmosphere as discussed in the paper by Van Horn.

The stripping force of the aluminum-aluminum connection subjected to a temperature of 200°C for 12 hours has actually been found to increase by a factor of about 2 which suggests that the two parts have diffused into each other and formed a permanent connection. This has been confirmed by the increase in force required to unwrap the wire. Since the activation energy of self diffusion is about the same as the activation energy for stress relaxation, then as the hoop stress is relaxed at high temperatures, solid state diffusion takes place and a diffusion joint is formed in aluminum. The same process, both for stress relaxation and diffusion, should take place at a much lower rate at lower temperatures and as the hoop stress is relaxed, a diffusion connection between the two parts is formed so that no decrease in the conductivity of the connection occurs and an actual increase in the strength of the connection results. The same process should result between any two materials provided the energy of self diffusion from one into the other is less or equal to the activation energy of stress relaxation for the weakest component.

It is planned to tin plate both terminals and wires for all of the wrapped solderless connections used in the telephone system. In order to find how the two processes of stress relaxation and self diffusion, which progress as a function of time and temperature, affect the two fundamental properties electrical conductivity and mechanical strength of the connections, a large number of connections were wound up and subjected to temperatures of 200°C for times corresponding to 0.9, 0.8 etc., of the initial stress as determined from Fig. 21. The electrical resistances of the connections were determined before and after the treatment and the stripping force was also measured. Within the experimental error the resistances of the connections remained the same while the average stripping force for twenty connections for each point are shown by Fig. 23. No significant change in the stripping force occurred out to values of stress relaxation less than 0.2 times the initial hoop stress. These experiments show that as the hoop stress is relaxed by time and temperature, self diffusion occurs between the two parts of the connec-

tion in such a way that the mechanical strength and conductivity are maintained unchanged with time.

In summary it has been shown that

1. The connections are sufficiently intimate to permit solid state diffusion but the strains are not high enough to cause cold welding of the connection.

2. The hoop stress relaxes as a function of time and temperature according to well known activation energy equations with an activation energy for copper wires on nickel silver connections varying from 12.5 to
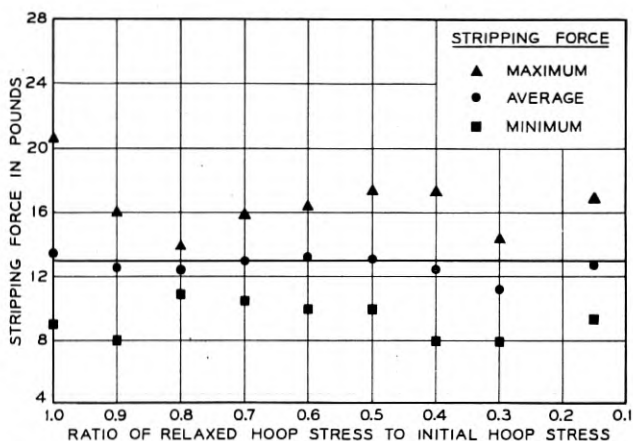
Fig. 23 — Stripping force plotted as a function of stress relaxation.

40 kilocalories. It requires about 2500 years to relax half the hoop stress at 77°F (25°C) and 40 years at 135°F.

3. The twin processes of stress relaxation and self diffusion occur in such a way as to maintain or increase the strength of the connection and to leave the resistivity of the connection unchanged with time.

4. The conclusion from all of these tests is that the life span of the solderless wrapped connection appears to be ample for meeting any of the likely requirements. This conclusion is reinforced by life tests of a large number of wrapped solderless connections in service that have been carried out over two years without a failure and by the long and satisfactory tests of the screw connection whose success also depends on the stress relaxation, diffusion processes. Such connections have been shown to be satisfactory over periods of time in excess of twenty years.

APPENDIX

METHOD FOR ANALYZING STRESSES AND STRAINS FROM PHOTOELASTIC PICTURES

The methods for analyzing photoelastic pictures are given in detail in a number of books and other publications[7] and only a brief summary of the method used here will be given.

When polarized light is sent normal to a plane of photoelastic material that is strained in the plane, the light is broken up into an ordinary and an extraordinary ray that travel with different velocities. It is shown[7] that the birefringence, which is defined as the difference between the two indices of refraction μ_1 and μ_2 (i.e. the index of refraction is the ratio between the velocity of light for one of the rays in a vacuum to the velocity in the medium) is given by

$$B = \mu_1 - \mu_2 = C'\sqrt{(T_1 - T_2)^2 + 4T_6^2}, \qquad (19)$$

where C' is a constant called the relative stress optical constant, T_1 is the tensional stress along the X axis, T_2 the tensional stress along the Y axis and T_6 the shearing stress in the XY plane. If we change the direction that we call the X axis until we reach the direction of maximum stress, the relations between this stress and the tensional stress at right angles to it are given by the equations

$$T_1' = T_1 \cos^2\theta + 2 \sin \theta \cos \theta T_6 + T_2 \sin^2 \theta,$$
$$T_2' = T_1 \sin^2\theta - 2 \sin \theta \cos \theta T_6 + T_2 \cos^2 \theta, \qquad (20)$$

where

θ is the angle between the X axis and the axis of maximum tension.

If θ is chosen so that T_1' is a maximum, we find

$$\tan 2\theta = \frac{2T_6}{T_1 - T_2} \qquad (21)$$

and

$$T_1' = \frac{T_1 + T_2}{2} + \tfrac{1}{2}\sqrt{(T_1 - T_2)^2 + 4T_6^2}$$
$$T_2' = \frac{T_1 + T_2}{2} - \tfrac{1}{2}\sqrt{(T_1 - T_2)^2 + 4T_6^2} \qquad (22)$$

[7] Coker and Filon, Photoelasticity, Cambridge University Press, 1931. M. Hetenyi, Handbook of Experimental Stress Analysis, John Wiley and Sons, Chap. 17, 1950. R. D. Mindlin, J. Applied Physics, April 1939, pp. 222–241 and May 1939, pp. 273–294. W. P. Mason, Electrooptic and Photoelectric Effects in Crystals, Bell System Tech. J., 29, pp. 161–188, April, 1950.

and hence

$$T_1' - T_2' = \sqrt{(T_1 - T_2)^2 + 4T_6^2}. \tag{23}$$

Hence, the birefringence is directly proportional to the difference between the principal stresses.

The retardation R is the difference between the path length for the fast and slow rays when they are transmitted through the thickness of the specimen. If h is the thickness of the specimen

$$h = v_2 t, \qquad h - R = v_1 t \tag{24}$$

and hence

$$\frac{R}{h} = \frac{v_2 - v_1}{v_2} \quad \text{and} \quad \frac{c}{v_1} \frac{R}{h} = \left(\frac{v_2 - v_1}{v_2} \right) \frac{c}{v_1} = \left(\frac{c}{v_1} - \frac{c}{v_2} \right) = B. \tag{25}$$

Hence, the retardation is given by

$$\frac{R}{h} = \frac{C'c}{v_1} (T_2' - T_1') = C(T_2' - T_1'). \tag{26}$$

If λ is the wavelength of light used for the measurement,

$$\frac{N\lambda}{h} = \frac{R}{h} = C(T_2' - T_1'), \tag{27}$$

or the number of fringes is related to the difference of the principal stress by

$$N = \frac{Ch}{\lambda} (T_2' - T_1'). \tag{28}$$

For photoelastic bakelite for the green mercury line 5461 Å, the fringe constant λ/C is 88 pounds per square inch/inch/fringe. Hence the difference between the principal stresses is given by

$$(T_2' - T_1') = \frac{N}{h} (88) \text{ pounds/sq in.} \tag{29}$$

The isochromatic lines such as shown on Fig. 3 are taken with quarter-wave plates in addition to the crossed polaroids. These lines give directly the multiple number of times the stress is greater than the value $(88/h)$ pounds per square inch. In order to know the exact multiple, one has to know the starting point or the points of zero stress difference, called isotropic points. These are determined by using white light and locating the gray fringes. The zero order fringe is gray because for this case, all the wavelengths of light are delayed the same amount and hence no

color effects appear. A first order fringe will have red (absence of violet) nearest the zero order fringe and violet (absence of red) furthest from the zero order fringe. High order fringes will not appear at all in white light since they are the resultant of a number of colors which add up to white light.

Having located the zero order fringe, a simple count will give the number of times the factor $(88/h)$ has to be multiplied by to obtain the number of pounds per square inch. This stress will, however, only be a stress difference and in order to resolve this into stresses along X and Y axes and a shearing stress in the XY plane, other information is necessary.

One part of the information is obtained when the isoclinic directions are obtained. These directions are the directions of the principal stress axes and these are obtained by taking out the quarter wave plates and rotating the axes of the polaroids (keeping them crossed) until the polarization axes coincide with the principal stress axes at any point. When this occurs, the picture will be black because if no model were there, the polarized light passed by the polarizer would be cancelled by the analyzer. If the principal axis of the stress ellipsoid coincides with the direction of polarized light from the polarizer, only one ray will be generated whose plane of polarization coincides with that of the polarized light and hence this will be cancelled in the analyzer. The isoclinic lines show up much better if a white light source is used. Hence, the isoclinic lines locate the direction of the principal axes of the stress ellipsoid. From equations (20) and (23) we have

$$T_6 = \left(\frac{T_1' - T_2'}{2}\right) \sin 2\theta, \qquad T_1 - T_2 = (T_1' - T_2') \cos 2\theta. \quad (30)$$

Hence, if we know θ (the direction of the principal stress axes with respect to the axes for which the stresses are to be analyzed) the shearing stress T_6 and the difference between T_1 and T_2 can be obtained from equation (30).

The other necessary relation can be obtained from the equilibrium stress relations that have to be satisfied by any stationary body, namely

$$\frac{\partial T_1}{\partial x} + \frac{\partial T_6}{\partial y} = 0; \qquad \frac{\partial T_2}{\partial y} + \frac{\partial T_6}{\partial x} = 0. \quad (31)$$

Integrating these equations

$$T_1 = T_{1_0} - \int \frac{\partial T_6}{\partial y}\, dx; \qquad T_2 = T_{2_0} - \int \frac{\partial T_6}{\partial x}\, dy. \quad (32)$$

For example, the isoclinic lines for the square plate of Fig. 3 are shown by Fig. 24. All the isoclinic lines converge on the positions of the "eyes" which is a characteristic to be expected of an isotropic point $(T_1' - T_2') = 0$. For the line through the center of the "eyes", which we designate as the X direction, the isoclinic line lies along the X axis and hence there is no shearing stress along this axis and $T_1 - T_2 = T_1' - T_2'$. At either edge the stress in the X direction is zero and hence at this point the stress T_2 (which is a compression) is obtained by counting the number of lines from the isotropic point to the edge (in this case 20) so that the stress at the edge is

$$T_2 = \frac{20 \times 88}{0.95} = 1850 \text{ pounds/square inch} \tag{33}$$

To determine the shearing stresses at any point, one needs to fit the isoclinic lines over the isochromatic lines. For example, for any point along the isoclinic line marked 15, the direction of the principal axis is
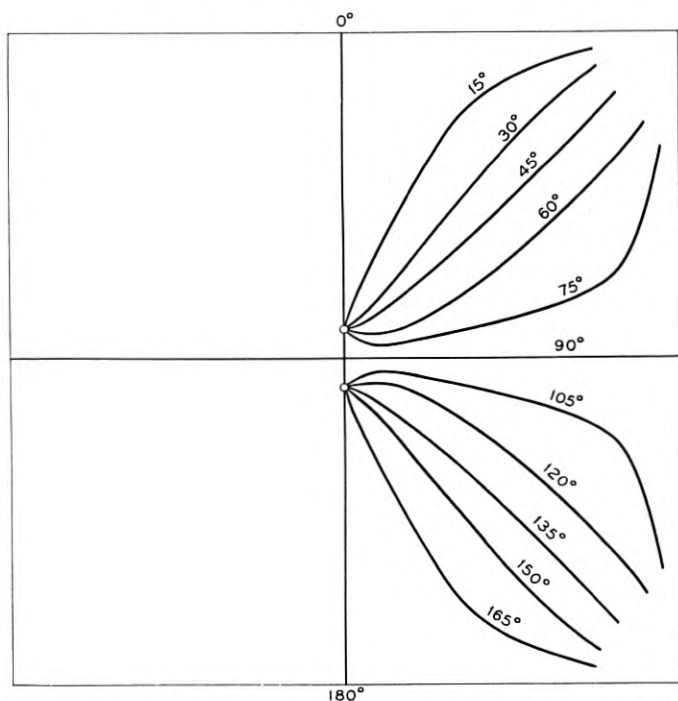
Fig. 24 — Isoclinic lines for a square terminal whose isochromatic lines are shown by Fig. 3.

15° from the X axis, since it was obtained by setting the analyzer and polarizer at 15° from the direction X. Counting the number of isochromatic lines from the "eye", the principal stress difference $T_1' - T_2'$ is known and using equation (30), T_6 is at once calculated. To obtain $\partial T_6/\partial y$, it is sufficient to divide T_6 by the distance y that the point is above the X axis. Proceeding this way, values of $\partial T_6/\partial y$ are plotted as a function of X. Then from the first of equations (32) with $T_{1_0} = 0$ at the edge, one can obtain T_1 by integrating $\partial T_6/\partial y\; dx$ over the X axis. The result is the curve marked T_1 of Fig. 5. Having T_1, T_2 can be obtained from the isochromatic lines which determine $T_2 - T_1$ and this is plotted on the upper part of Fig. 5.

Solderless Wrapped Connections

PART III — EVALUATION AND PERFORMANCE TESTS

By R. H. VAN HORN

In the development of solderless wrapped connections the basic require-ments of electrical and mechanical stability have been trans'ated into test requirements on laboratory samples of these connections and on the manu-factured product. These tests have shown that the connections can with-stand satisfactorily the effects of corrosion, humidity, vibration, and relaxa-tion. The effects of terminal dimensions, materials, corner sharpness. wrapping tool construction, etc. are noted.

INTRODUCTION

The previous two articles have described the fundamental considera-tions involved in solderless wrapped connections. A description of these connections together with a rather detailed explanation of the forces which maintain them has been presented. This third article discusses the results of a number of tests where the actual fabrication and use of these connections have been simulated. From these results it will be seen that reliable performance can be expected over the central office life of these connections, and that the variations permissible in their fabrication will privide sufficient margin to make that process easy to control.

GENERAL REQUIREMENTS

The minimum physical requirements which a solderless wrapped con-nection must meet are:

1. Intimate contact between wire and terminal.
2. The points of contact should be gas-tight to withstand corrosion.
3. The minimum dimension of the gas-tight area should be great enough so that it does not decrease appreciably during the expected

591

life (forty years) because of corrosion or because of relaxation of the internal stresses in the wire or terminal.

4. The sum of the areas of intimate contact should be equal to or larger than the cross-section of the wire to prevent local heating.

5. The connection should be mechanically stable so that forces applied to the connection during shipment, installation and subsequent maintenance activities will not dislodge the wire and break the points of intimate contact.

6. The wire should not be embrittled during the wrapping operation so that it will subsequently break due to vibration, handling, or unwrapping.

TRANSLATION OF GENERAL REQUIREMENTS INTO TEST REQUIREMENTS

In order to evaluate the feasibility of solderless wrapped connections, extensive development studies were necessary so that a good estimate could be obtained as to whether these connections would meet these general requirements under a wide variety of conditions and with sufficient margin to provide for ease of manufacture. It was necessary to translate these requirements first into a set of development test requirements and second into a set of shop inspection requirements. These two translations of requirements will not necessarily be the same although there will be a large degree of overlap. In either translation they can be broken down into two distinct areas. These two areas of tests cover those tests which are related to evaluating (1) life and electrical stability of the connection and (2) mechanical stability.

A great deal of engineering judgment was used in the translation of the physical requirements into inspection requirements and this judgment took into account the special nature of conditions to be encountered in telephone offices. Consideration was given to the methods of handling the wire, the manipulations of installation and maintenance men when working on central office equipment, the effect of vibration, the effect of variation in tool dimensions and the like. Furthermore, a good deal of knowledge of the corrosion and relaxation process had to be developed before it could be judged on the basis of accelerated tests that the connections might be expected to have a satisfactory field life.

There may be applications where the translation of the physical requirements into test requirements may be different, perhaps quite different, from the translation made for the telephone apparatus which was in mind during this investigation. The use of other kinds and sizes of wire or terminals may require an evaluation quite different from the one presented here. Nevertheless, the test requirements herein estab-

lished should have a wide application in many areas. In telephone practice they provide a reasonable latitude for variations in the process of making the connection, including tolerances in the parts, and at the same time guarantee a good product.

Most of the product tests which have been made so far apply specifically to connections which use Standard No. 24 tinned solid copper wire such as is used in 95 per cent of the telephone switching plant, and terminals having a rectangular cross-section punched from sheet nickel silver, brass or bronze and whose dimensions are one-sixteenth inch wide by the thickness of the stock (0.013″ to 0.062″). These terminals are typical of those which are or could be used on relays, switches, resistors, capacitors, terminal strips, etc. Studies with #20 and #22 wire have been made with results similar to those with #24 wire.

A similar investigation is under way on connections to terminals which are made from round silicon copper and nickel silver wire such as are used on the wire spring relay, and where the wire terminal had been prepared for connection by various treatments such as flattening, coining, serrating, annealing, etc.

AGING OF WRAPPED CONNECTIONS

Assuming that a connection can be wrapped with sufficient mechanical strength to withstand handling, vibration, etc., there appear to be two factors which might cause the connection to fail after a period of time. These factors are (1) relaxation of the internal stresses in the metals, and (2) corrosion of the metal surfaces. As Mr. Mason points out in his paper, it now appears that solid state diffusion of metal across the boundary between the wire and the terminal improves the connection as much or more than relaxation degrades it. Tests have been designed to relax the connections in a short time to the same degree that will occur in the normal forty-year life which is required. These tests will be described later.

Several investigators have studied the rate of surface corrosion in indoor atmospheres of metals such as are used in these connections.

Studies of corrosion* where oxidation is the primary factor indicate for example that the corrosion in zinc varies linearly with time and on copper it varies as the square root of the time exposed. The corrosion products are ZnO and Cu_2O. The corrosion rate for brasses fall between

---

* British Non-ferrous Metals Research Association, Investigation on the Atmospheric Corrosion of Non-ferrous Metals, First and Second Experimental Reports to the Atmospheric Corrosion Resistance Committee, May, 1924, and May, 1927, W. H. J. Vernon.

TABLE I—THICKNESS OF MATERIAL WHICH CORRODES IN FORTY YEARS

Zinc	0.00023″
Copper	0.000105″
Tin	Negligible

copper and zinc. There are very few data available on tin but its corrosion should be less than either copper or zinc. From the data presented the depth of metal which will corrode during a forty-year period in a central office for several metals is estimated* to be as shown in Table I.

Samples of copper bus-bar taken from telephone exchanges where they have been exposed for periods up to forty years show that the tarnish is primarily Cu_2O and the actual rate of corrosion may be appreciably less than that estimated by the above figures.

Thus the depth of metal corroded is small enough to neglect when the metals are subject to indoor atmospheric exposure. When dissimilar metals are joined in a connection there is the possibility of electrolytic corrosion in addition to atmospheric corrosion. The particular metals involved here, however, are relatively close to each other in the electromotive force series of metals so that it is expected that this effect will be negligible especially as there is no condensation on these connections.

It is therefore expected that the most important factor in the aging of these connections is the relaxation of stresses internal to the wire and the terminal rather than corrosion. The test procedures and results in the following sections reflect that view.

DEVELOPMENT TEST PROCEDURE — ELECTRICAL REQUIREMENTS

General Requirements 1 through 4 are considered together. A set of tests has been designed to evaluate the degree to which these requirements can be met.

Since the tests which have been devised are destructive it is necessary to check connections in production on a sampling basis.

In determining whether a connection meets the General Requirements 1 through 4, Test Procedure I as follows was devised.

Test Procedure I for Solderless Wrapped Connections

1. Check connection for insulating barrier film between wire and terminal.

2. Measure the variation of the resistance of connection while producing movement between the terminal and connecting wire.

---

* Unpublished memoranda, D. H. Gleason, Bell Telephone Laboratories.

3. Chill for two hours at 0°F.

4. Heat for three hous at a temperature (175°C) which will relax the stresses as much as will occur during the expected life at the normal central office operating temperature.

5. Expose the connection to a suitable agent which will discolor all the non gas-tight area.

6. Remeasure the resistance variation as in 2.

7. Unwrap the wire and estimate the size of the gas-tight areas.

Items 1 and 2 are intended to show whether initially there is the intimate contact between wire and terminal demanded by the General Requirement No. 1. At present, the multiple terminal banks on step-by-step switches are connected together with a clinched solderless connection. Based on experience with these connections, together with an estimate of the noise produced (See appendix A) a resistance variation in excess of 0.002 ohms is considered to indicate a poor connection. Items 5 and 7 will show the existence of the gas-tight areas demanded by General Requirement No. 2, and the size estimate from Item 7, the area of contact demanded by General Requirement No. 4 can be evaluated. Items 3 and 4 are intended to simulate most of the aging which will take place during the life of the connection. It is believed, as described earlier in this paper, that relaxation of internal stress is the chief factor in the aging of these connections. The chill at 0°F puts the maximum initial stress in the wire by shrinking it at extreme operating
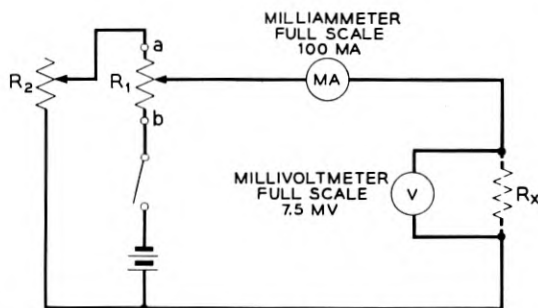
Fig. 1 — Test set for measuring resistance of solderless connection.

TEST PROCEDURE

(1) With R_1 set at "a" and switch SW closed, adjust R_2 so that the millivoltmeter reads 25 microvolts with R_x open.

(2) If millivoltmeter returns to zero with test connection across R_x, no barrier film is indicated and therefore connection is closed.

(3) All resistance measurements are made with R_1 set to give 100 milliamperes through R_x. Then R_x = millivolts/0.100.

(4) A variation of not more than 2 milliohms as the connection is moved or disturbed indicates a stable connection.

temperatures. The heating for three hours at 175°C produces the degree of relaxation expected over a 40-year period. By examining and estimating the size of the gas-tight areas after this process, the necessity for maintaining the intimate minimum gas-tight area demanded by General Requirement 3 is considered to be met. Item 6, the remeasurement of resistance is further confirmation of proper performance after relaxation.

For the measurement of resistance, and barrier films a circuit is used as shown on Fig. 1. With R_1 set at "a," R_2 is adjusted so that the voltage of about 25 microvolts is applied across R_x before the connection to be measured is inserted. This voltage is low enough to insure the absence of a film and at the same time gives a convenient reading on the test set. If the millivoltmeter drops to approximately 0 when the connection is inserted it indicates that no barrier film exists at the connection. The current is then increased to 100 milliamperes by means of the potentiometer, R_1, and the resistance is determined from the millivoltmeter reading. Since most of the measured resistance is in the wire rather than in the connection the important criterion of quality is the variation in resistance as the wire is moved relative to the terminal. If the variation in resistance of the connection does not exceed two milliohms when the wire is moved back and forth the connection is considered to be
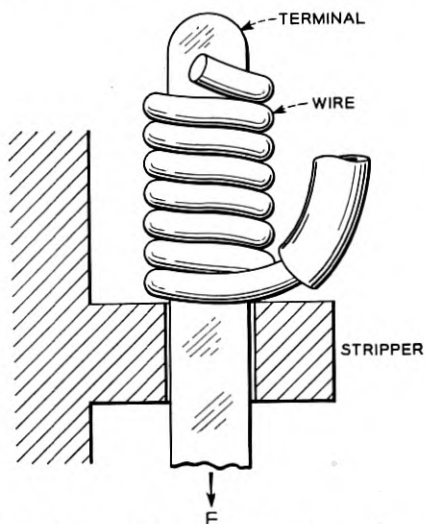
Fig. 2 — Stripping force test for solderless connection. (1) The connection shall consist of six turns minimum of which at least four are close wound. (2) The connection shall be capable of withstanding a stripping force, F, of at least 3000 grams applied as shown. (3) The wire shall be capable of being unwrapped from a terminal without breaking.

stable and the requirement for intimate contact between wire and terminal is considered met. In a typical local talking circuit this amount of resistance variation would correspond to a noise level of approximately -8 db where 0 db equals 10^{-12} watt and where anything less than $+26$ db would give no noise transmission impairment (See Appendix A).

DEVELOPMENT TEST PROCEDURE — MECHANICAL REQUIREMENTS

In determining whether a connection meets the General Requirements 5 and 6, Test Procedure II as follows was devised:

1. The connection shall be capable of withstanding a stripping force of 3000 grams applied as shown in Fig. 2.

2. The wire shall be capable of being unwrapped from a terminal without breaking.

These items are related directly to General Requirements 5 and 6. Experience has shown that under ordinary conditions of handling of connections as cables and frames are wired in the shop, or when a second connection is being made on an already wired terminal, a resistance to stripping in excess of about 3000 grams is required if the demands of General Requirement 5 is to be met. To insure adequate mechanical strength and current carrying capacity, a minimum of six turns is required. If the turns are close wound the strip-off force is readily met. If they are open wound they may strip off with a much lower force.

When the edge radius R (Fig. 3) of the wrapping tool is too small very high tension can be developed in the wire while wrapping. At very small radii the tension can be sufficient to break the wire. Although wrapping with high tensions in the wire produces a connection which will sustain very high stripping force, the wire is so embrittled that
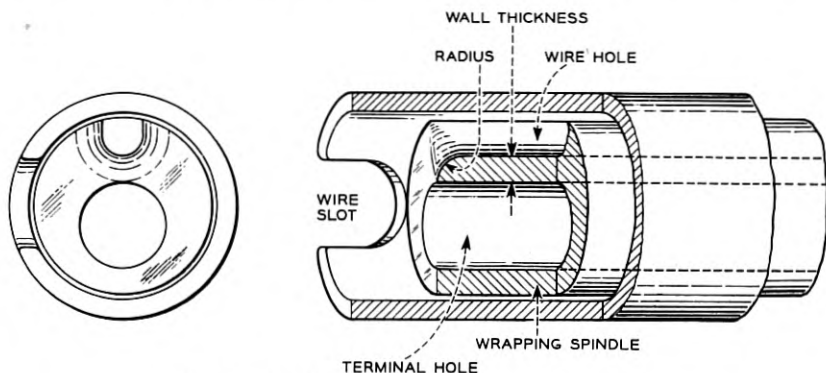
Fig. 3 — Solderless connection wrapping tool.

TABLE II—SOLDERLESS WRAPPED CONNECTION HUMIDITY AND CORROSION EXPOSURE TESTS

Test Specimens					Exposure							Result
Terminal Material	Terminal Size (Inches)	Wire Size	No. of Wraps	Average Stripping Force	No. of Samples	Vibration (Hours)	0° F. (Hours)	180° F. (Hours)	Corrosive Agent	Hours In Corrosive Agent (Hours)	Months at 85°F 90% Rel. Humidity (Month)	Number of Connections Where Resistance Exceeded 0.002 Ohms
Untinned brass	0.030 × 0.090	24 tinned	6	High†	24	50*	2	2	H₂S	½	21	None
	0.040 × 0.040	24 tinned	6	High	60	50	2	2	H₂S	½	7½	None
	0.020 × 0.060	24 tinned	6	High	60	50	2	2	H₂S	½	7½	None
	0.030 × 0.030	24 tinned	6	High	60	50	2	2	H₂S	½	7½	None
	0.040 × 0.090	24 tinned	6	High	48	50	2	2	H₂S	½	21	None
Brass-nickel plated and tinned	0.013 × 0.062	24 tinned	3	Low†	25	0	2	2	H₂S	½	9	25
	0.013 × 0.062	24 tinned	4	Low	25	0	2	2	H₂S	½	10½	5
	0.013 × 0.062	24 tinned	5	Low	25	0	2	2	H₂S	½	10½	1
	0.013 × 0.062	24 tinned	6	Low	25	0	2	2	H₂S	½	10½	None
Brass-nickel plated and tinned	0.013–0.031 × 0.047	24 tinned	6	High†	60	0	2	2	H₂S	½	15	None
	0.031 × 0.062	24 tinned	6	Low†	25	0	0	0	—	0	9	None
	0.031 × 0.047	24 tinned	6‡	Low	20	0	0	0	—	0	9	None
	0.031 × 0.062	24 tinned	6	5000 grams	25	0	0	0	—	0	9	None
	0.013 × 0.125	24 tinned	6	3100 grams	10	50	2	2	H₂S	½	2	None
	0.013 × 0.125	24 tinned	6	3400 grams	10	50	2	2	H₂S	½	2	None
	0.013 × 0.125	24 tinned	6	2100 grams	10	50	2	2	H₂S	½	2	None
	0.031 × 0.125	24 tinned	6	2600 grams	10	50	2	2	H₂S	½	2	None
	0.031 × 0.125	24 tinned	6	1800 grams	10	50	2	2	H₂S	½	2	None
Brass-nickel	0.020 × 0.063	24 tinned	6	7700 grams	50	50	2	2	H₂S	½	1½	None

Untinned nickel silver	0.035 × 0.125	22 tinned	4	4000 grams	5	0	2	2	HCl	⅔	10	None
	0.035 × 0.125	22 tinned	4	4000 grams	5	0	2	2	SO₂	⅔	10	None
	0.035 × 0.125	22 tinned	4	4000 grams	10	0	2	2	—	0	10	None
	0.040 × 0.125	22 tinned enameled	4	5400 grams	10	0	2	2	H₂S	½	6½	None
	0.040 × 0.125	20 tinned enameled	4	High†	10	0	2	2	H₂S	½	6½	None
Steel-tin plated	0.035 × 0.125	22 tinned	4	Low‡	50	0	2	2	H₂S	½	6½	None
Wire spring relay single wire	Flattened and solder dipped	24 tinned	6	3700 grams	25	0	2	2	H₂S	½	12	None
	Diamond shaped and solder dipped	24 tinned	6	4400 grams	25	0	2	2	H₂S	½	12	None
Twin wire	Solder dipped only	24 tinned	6	3700 grams	25	0	2	2	H₂S	½	12	None
	Diamond shaped and solder dipped	24 tinned	6	4400 grams	25	0	2	2	H₂S	½	12	None

* This was the time at which similar soldered connections began to show breakage.

† Estimated from wrapping tool dimensions.

‡ One turn insulated, five turns bare.
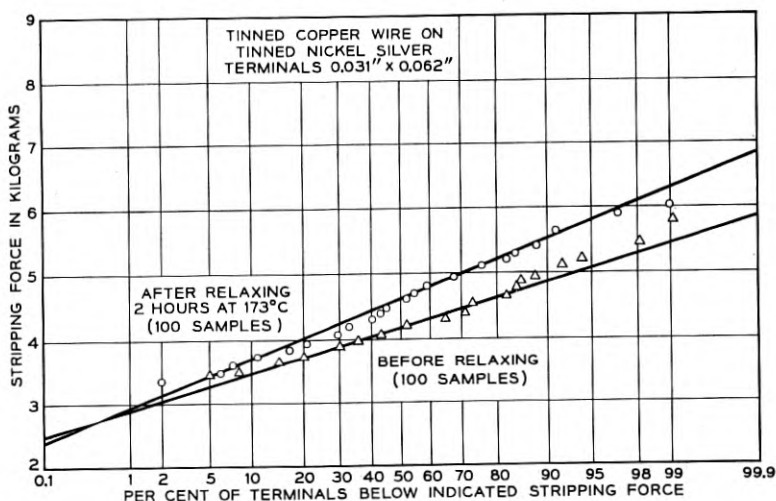
Fig. 4 — Effect of relaxation on stripping force.

under vibration and handling it may break easily. A practical and generally easily met requirement is that the connection be capable of withstanding unwrapping without wire breakage.

OTHER TESTS

A large amount of work has been done other than testing connections in accordance with the foregoing requirements. These tests (See Table II) include measurement of exposure to humidity, hydrogen sulphide corrosion, heating by internal electrical losses, vibration, etc. Many of these tests were made before any clear picture mechanism of deterioration of the connection was available so that they were not exposed to temperature high enough to cause very much relaxation. Therefore, there is no quantitative relation between these tests and the actual aging of connections, but they do indicate a general ruggedness which gives confidence in the reliability of the connections.

In measuring the effect of relaxation, connections have been tested in a preliminary way by subjecting them to sufficient heat sufficiently long to relax the stresses appreciably more than would be expected in a normal forty-year life. Upon unwrapping, the bright areas of contact were still present. Fig. 4 shows the result of a typical relaxation test on the stripping force. The rise in stripping force after relaxation is attributed to the solid state diffusion process described by Mr. Mason.

In order to compare how the solderless wrapped connection and soldered connections are able to withstand the effects of vibration over a period of time, several sets of terminals were mounted on a test table and wires from a cable were wrapped or soldered to the terminals. The cable was then oscillated through a small angle at 19 cycles per second and the elapsed time until lead breakage occurred was recorded. Some early similar tests indicated that soldered connections began to break at about fifty hours and the solderless connections lasted several times that long. In a more recent test where ten connections of each type were set up three of the soldered connections broke within six hours and all soldered connections broke within 1,000 hours. None of the wrapped connections had broken in that interval. In another test using ten soldered connections and twenty solderless connections, one of each broke at nineteen hours and after fifty-seven hours two more soldered connections broke with no further breakage of solderless connections. Further tests are being made of resistance to vibration under various conditions but the early indications are that the life of the solderless connections is as good as and perhaps better than that of the soldered connections.

The increased resistance of the wrapped connection to vibration breakage comes about because in the soldered connection the bending stresses are concentrated at the point of emergency of the wire from the solder, while in the wrapped connection the stresses tend to be distributed over the entire first turn of the connection. In order to obtain long life, it is important to avoid any exposed nicks in the wire at the first turn of the connection. Otherwise stress concentrations, perhaps as severe as those in soldered connections may be set up.

Incidentally, unless the wire has an enamel or similar coating it is unnecessary to clean the wire or the terminal before wrapping, since forces set up as the wire wraps on to the terminal shear off surface films and the resultant gas-tight areas are immediately established on clean surfaces. It is this intimate contact which also makes possible the diffusion process between the wire and the terminal.

A large number of parameters have an effect on the level of stripping force which a given connection will withstand. Among these are the number of wraps, tension on the wire during wrapping, sharpness of the corners of the terminals, the elastic properties of both the terminal and the wire, the size of the terminal, the hardness of the terminal and wire, the presence or absence of tin plating, taper of the terminal, aging, tool design, etc. Tests have been run showing the effect of variations in these parameters on the stability of the connections. Typical results of these
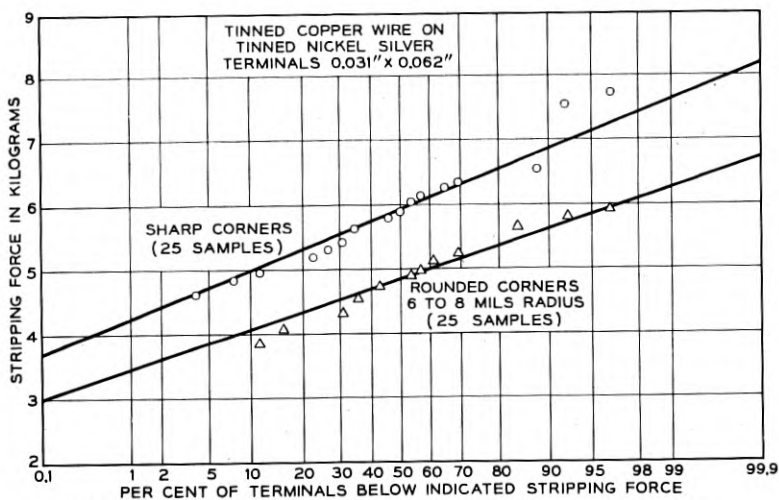
Fig. 5 — Effect of terminal edge sharpness on stripping force.
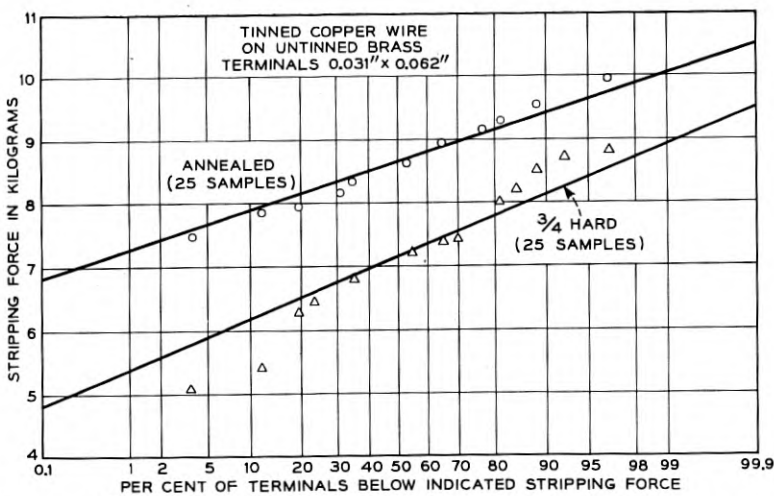
Fig. 6 — Effect of terminal hardness on stripping force.

tests are shown on Figs. 5 to 9 inclusive. These charts show a plot on arithmetic probability paper of stripping force against the cumulative per cent of the sample below the indicated stripping force. The straight lines shown there are actually drawn through the average and average minus 3σ points where σ is the standard deviation of the observed readings. In the cases shown, and in fact in practically all the tests made to date the results have come out so that the experimental values fit very well on the straight line drawn through these points, thereby indicating good normal statistical distributions of the data.
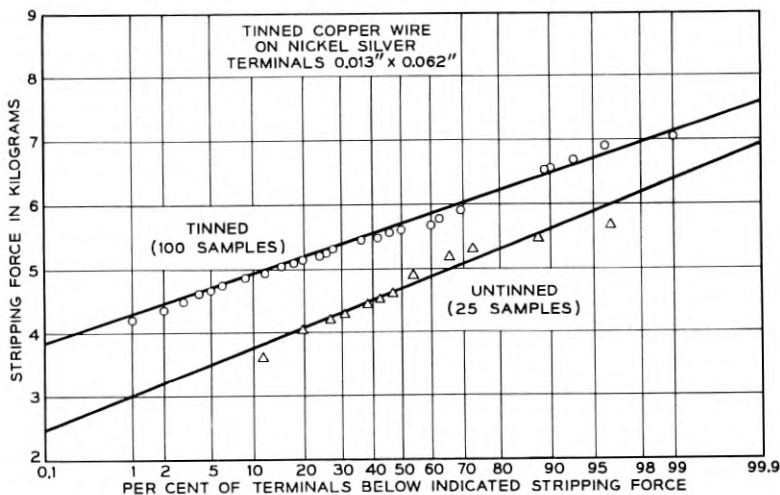
Fig. 7 — Effect of terminal plating (tin) on stripping force.

For example, Fig. 5 shows the effect of stripping force of rounding the corners of the terminal. These curves are for a typical tinned nickel silver relay terminal 0.031″ by 0.062″ in crosssection. The terminals represented by the upper curve were punched from sheet nickel silver stock and the corners were sharp as is usually the case where shearing operations are used. The lower curve represents connections where the corners were rounded off to a 0.006″ to 0.008″ radius. It is seen that on the average the stripping force using the sharp corners is about 1,000 grams higher than with the rounded corners, although both sets of terminals met the 3,000 gram requirement for mechanical stability.

The effect of terminal hardness on stripping force is shown in Fig. 6. Here brass terminals were used and it is evident that the softer terminals give considerably higher stripping force than the harder ones. Both
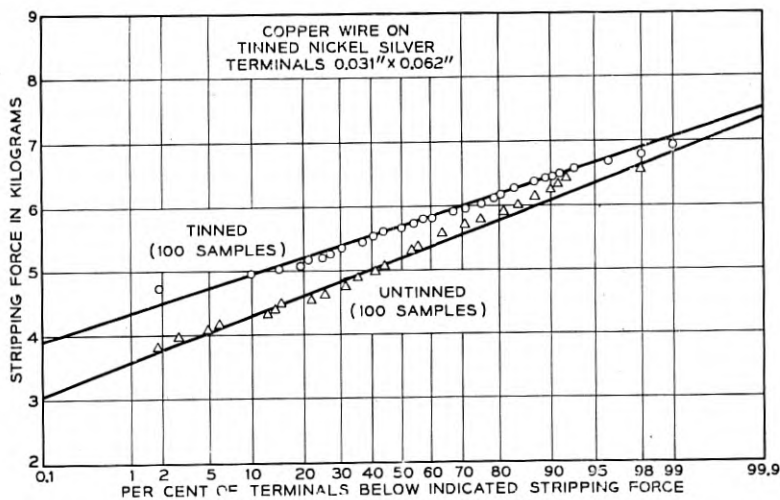
Fig. 8 — Effect of wire plating on stripping force.

curves run higher than the curves for the harder nickel silver in Fig. 5. Fig. 6 shows that the presence of tin plate on a 0.013″ by 0.062″ terminal increased the minimum expected stripping force by more than 1,000 grams. Figure 8 shows that on a 0.031″ y 0.062″ terminal that while the effect of the tin plate on the wire was to raise the stripping force as before, the increase was not as great as that in Fig. 7. It is also clear that the thicker terminal of Fig. 8 results in a higher stripping force than that for the corresponding thinner terminal of Fig. 7.

In analyzing the effect of tool variations it has been convenient to take a series of curves like those in Figs. 5 to 8, and from each curve select two points, namely one for the average stripping force and one for the average minus 3σ where σ is the standard deviation of the observed values of stripping force. Using these two points for each of several sets of curves a plot may be obtained of the average and minimum expected stripping force as one or another of the parameters of the tool design are changed. In this manner curves of the type shown on Fig. 9 were obtained. By preparing curves of this type for any given application, the permissible range, Z, of variation of tool radius, R, can be obtained. This range includes those values of R which are above that value which results in wire breakage and below that for which the minimum expected stripping force falls below the stripping force limit.

Table II summarizes the results of many of the tests which have been run on a large number of samples in which the materials and dimensions of the terminals, wires, tightness of wrap, etc., were varied. All of sample

connections originally met the initial test requirements outlined above for resistance variation. They were then subjected to the indicated exposures including being kept in the Humidity Room at 85°F and 95 per cent relative humidity for periods up to two and one-half years, following which they were remeasured. In no case was any barrier film found which did not break down at 25 microvolts nor was any resistance variation in excess of 0.002 ohms observed. Some of these terminals are shown in the photographs of Figs. 10 and 11. Upon unwrapping some of these connections there were still considerable areas which were bright (See Fig. 12), indicating that they were still gas-tight. This would seem to indicate a life expectancy of many years. Since connections
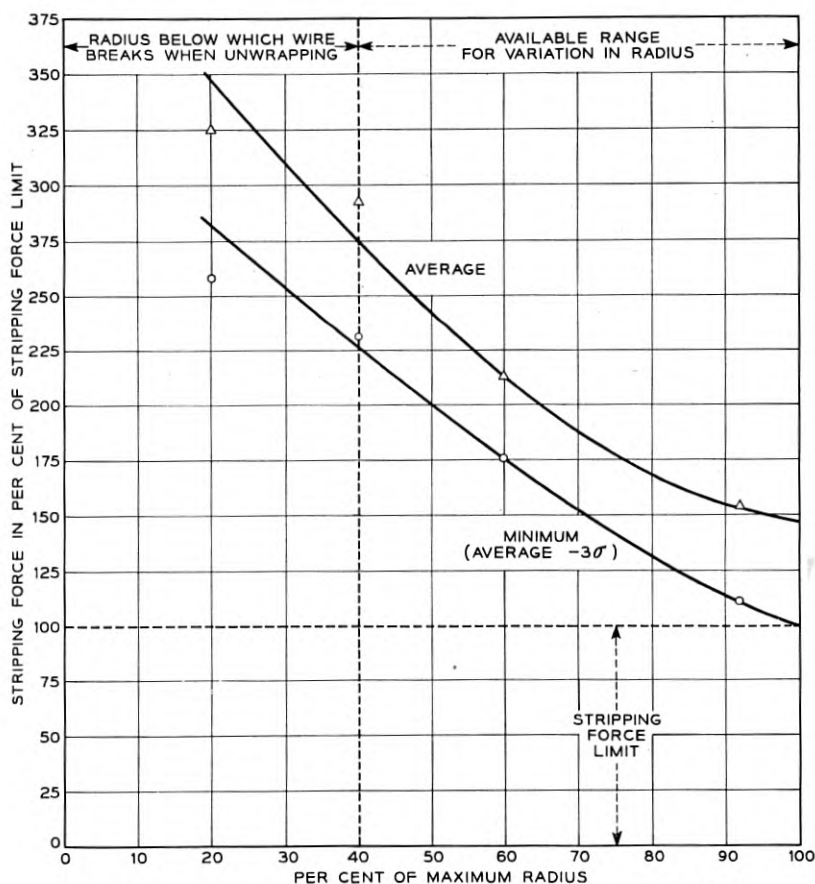
Fig. 9 — Stripping force as a function of radius of wrapping tool.
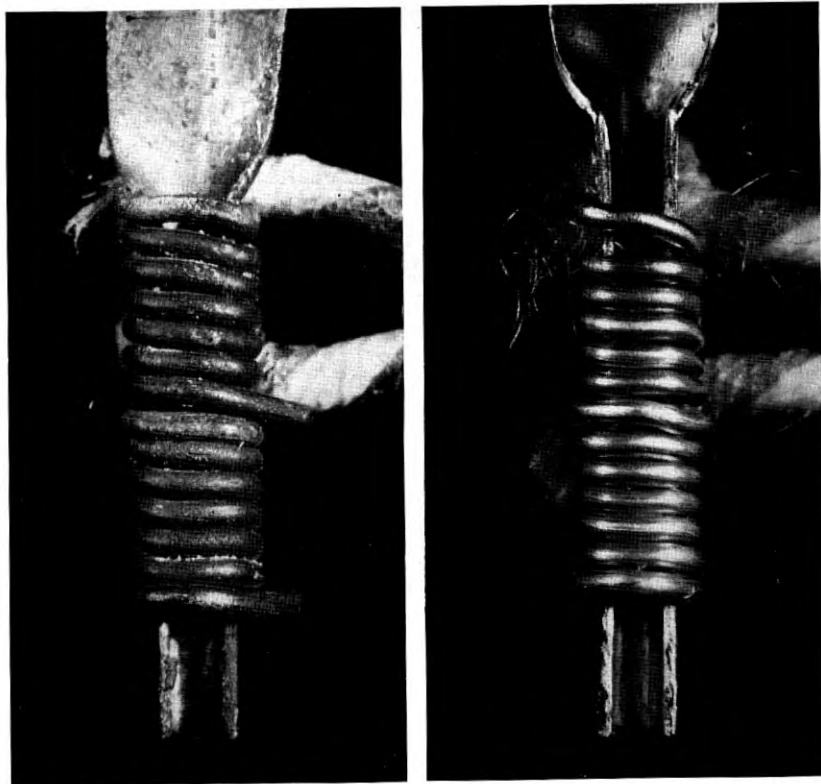
Fig. 10 — Terminal at left exposed to H₂S for 14 hours and then 30 months at 85°F 90%RH. The connections did not develop resistance variation in excess of the 0.002 ohm limit. Corresponding new terminal is shown at right.

which meet the 3,000-gram strip-off requirement appear satisfactory on the other tests, the strip-off test used on a sampling basis has been standardized as the shop control of the wrapping tool and indirectly of the quality of the connections themselves.

FIELD TRIALS

A number of equipment units using these connections is currently on field trial. At Tonawanda, N. Y., one trunk unit using wire spring relays of an early design and 300 solderless wrapped connections has been in service successfully since 1951. At Elmhurst, L. I., a sender frame comprising five senders with about 7,500 connections has been in use about a year and a half with no troubles reported to date. At Boston, Mass., an outgoing sender frame for the No. 4 crossbar office comprising three
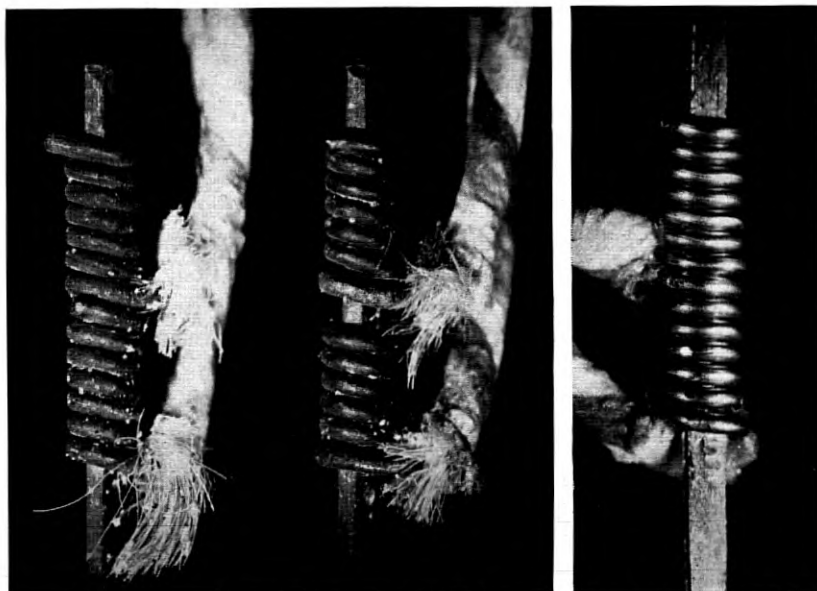
Fig. 11 — Terminals on left exposed to H₂S for one-half hour and 85°F 90 per cent RH for seven and one-half months without development of excessive resistance variation. Corresponding new connection is shown on right.
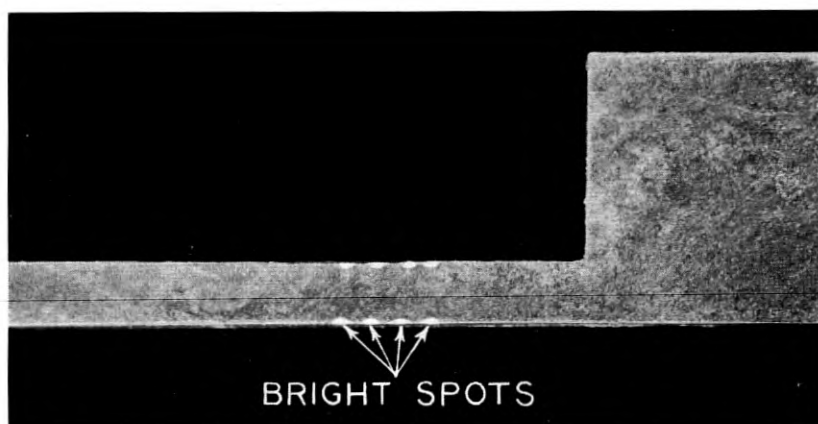
Fig. 12 — Corroded nickel silver terminal after unwrapping connection. Note the bright spots where the wire was in contact with the terminal.

senders with 6,000 solderless connections has been in service about nine months, again with no troubles reported.

In order to test these connections in a transmission circuit one channel of the K-1 carrier system between Newark and Atlanta was selected. Rather than wire a regular equipment unit with solderless wrapped connections, four units were built consisting of two groups of 320 connections in series. Each unit was exposed to 15 weeks at 85°F 90 per cent relative humidity and then inserted in the system. One unit was located at each of the following places: New York, Philadelphia, Baltimore, and Richmond. No troubles or adverse service reactions have been received to date, more than six months after installation.

SUMMARY

The tests described indicate that solderless wrapped connections are practical when wrapping No. 24 solid tinned copper wire on flat punched terminals of brass or nickel silver where the width is one-sixteenth inch and the thickness varies from 0.010" up to one-sixteenth inch. Similar tests with heavier wire, (No. 20 to No. 23) such as would be used in distributing frames and tests with No. 24 wire on flattened, coined, or otherwise treated wire spring relay terminals have also been successful. These connections are mechanically stable, and have less tendency to break due to handling and vibration than solder connections, and will

$$e = \frac{\sqrt{2}}{2} L (R_V)$$

(a) MEASURING CONDITION

$$V_A = \frac{2}{3} e$$

(b) SUBSCRIBER CONDITION

$$V_R = \frac{1}{2} V_B = \frac{1}{4} e$$
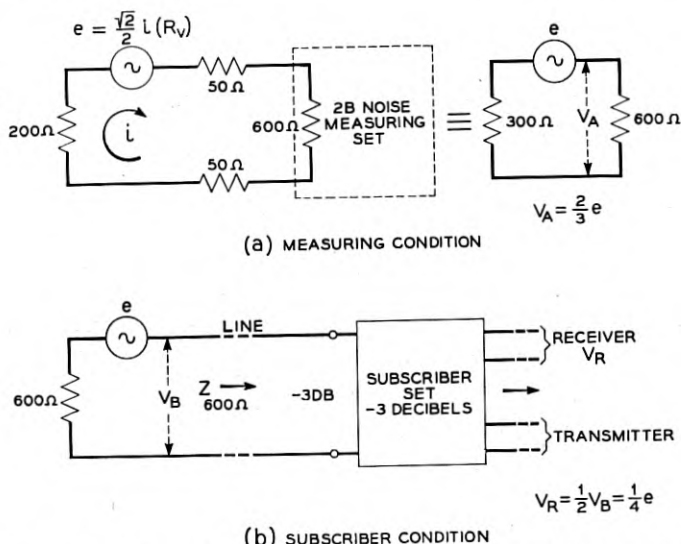
Fig. 13 — Comparison of noise measuring circuit with subscriber circuit.

probably have a central office life of more than fifty years without developing objectionable resistance variation. Tests on other terminals and using wires heavier than No. 20 and on wires smaller than No. 24 are to be made.

APPENDIX A

NOISE LEVEL ARISING FROM VARIABLE CONTACT RESISTANCE

The measurement of contact resistance noise in a telephone central office has been standardized, and the estimate of the effect of a given resistance variation has been made in accordance with the standard technique.

In this measurement, a 2-B noise measuring set is connected as shown in Fig. 13 (a) where the voltage, e, is that due to a variation in contact or connection resistance. If a dc current i is flowing and the resistance variation is $(\Delta R)_0 \sin \omega t$, then

$$e_{\text{eff}} = \frac{\sqrt{2}}{2} i \cdot (\Delta R) \tag{1}$$

and

$$V_A = \frac{2}{3} e. \tag{2}$$

In the subscriber condition (Fig. 13 (b)) there is a 3 db loss in the subscriber's loop and a 3 db loss at the subset so that the voltage at the receiver

$$V_R = \tfrac{1}{2} V_B$$

and

$$V_b = \tfrac{1}{2} e$$

It therefore follows that

$$\frac{V_A}{V_R} = \frac{2e}{3} \div \frac{1}{4} e = \frac{8}{3}$$

and

$$20 \log \frac{V_A}{V_R} = 9 \ db \text{ (to the nearest decibel).}$$

Extensive tests have shown that a level of 17 db at the receiver where 0 db $= 10^{-12}$ watts will produce no noise impairment of transmission. The corresponding level at the input to the noise meter in the measuring condition is $17 + 9$ or $+26$ db.

Currents in talking circuits may vary from 0.025 amperes in many central office circuits to approximately 0.150 amperes in short subscriber's loops. At a current of 0.025 and assuming $\Delta R = 0.002$ ohms the noise voltage will be

$$e = \frac{\sqrt{2}}{2} \times 0.025 \times 0.002 = 35.4 \times 10^{-6} \text{ volts}$$

and the voltage at the input to the noise measuring set will be $\frac{2}{3} e = 23.5 \times 10^{-6}$ volts. The power into the set will be

$$\frac{23.5 \times 10^{-6}}{600}$$

or 1×10^{-12} watts or 0 db. The noise measuring set is equipped with a weighting network which takes into account the relative interfering effect of different noise frequencies on received speech. For random noise, the effect of this network is to reduce the measured noise level by about 8 db as compared with the reading that would be observed without the network. Accordingly the noise produced by the above 23.5×10^{-6} volts would measure -8 db.

An Improved Circuit for the Telephone Set

By A. F. BENNETT

A telephone set known as the 500 type has been in production for Bell System use for some time.[1] The successful outcome of an intensive study has made it possible to simplify the circuit and some of the components of this set, and thereby to increase dependability and life and significantly to reduce the manufacturing cost. This change now has been completed and telephone sets embodying the necessary modifications are in production. This paper discusses some of the problems involved in this work and outlines the improvements which have been effected. Presented also is information concerning the superior performance of the 500 set over the preceding telephone set when used in noisy locations.

INTRODUCTION

One of the outstanding characteristics of the 500 type set is a 10 db increase in combined transmitting and receiving performance on long loops. This gain is equally divided between receiving and transmitting. This improvement has resulted largely from the use of a transmitter and a receiver which are not only more efficient, but also have better frequency response characteristics. To take full advantage of these transmission gains, two new elements were introduced in the original 500 set design. One of these elements was a better sidetone balancing circuit to offset the more sensitive transmitter and receiver, and maintain sidetone at a level no greater than it was with the previous design of set (known as the 302 type). The other was a tungsten filament and thermistor element to control automatically the transmitting and receiving levels so that the desired gains occur on a graduated basis as the loop length increases. This combination of filament and thermistor bead was called the transmission equalizer. While the transmission objectives were met with telephone sets having these elements, this additional component increased the manufacturing cost of the set appreciably, and therefore more economical means of attaining the desired results were sought. Such means have been found in the form of an arrangement in which a

611

pair of silicon carbide varistors serves for both the sidetone balancing and equalizing functions, and with certain novel modifications in the circuit of the set, yields essentially the same overall transmission performance as the original design at lower cost. The advantage is retained of a telephone set which from a transmission standpoint is universal in its application in the telephone plant. This universality avoids additional codes of set which require effort and expense to administer. Telephone sets having these features are designated as 500C (manual) and 500D (dial) types. The original design is known as the 500A (manual) and 500B (dial types).

GENERAL

When the sidetone, which is the sound level in the receiver caused by voltage developed in the local transmitter, becomes too high, the subscriber subconsciously lowers his talking level thereby reducing the level of outgoing transmission.[2, 4] To avoid this loss and make the higher efficiency provided by the new transmitter and receiver effective, a reduction in sidetone must be made. When the room noise level at the station is high, another undesirable effect of high sidetone occurs. This is the loss in effective receiving which results from the masking of incoming speech by the high level of sidetone noise in the receiver. This will be discussed in more detail later.

In developing the original design of 500-type set, several different equalizing means were considered, and the type involving a filament in series with the transmitter, and a thermistor bead in shunt with the receiver was selected as the most suitable with the means available at that time. Reduction in sidetone was realized by the use of a special balancing network which required an auto transformer. The set now being produced is based on a new circuit arrangement, utilizing a pair of semiconductors in the form of silicon carbide varistors, one of which has a résistance in a range which required development for this specific purpose.

CHARACTERISTICS OF SILICON CARBIDE VARISTORS

Fig. 1 shows the relation between the dc voltage and current for the two varistors used in the 500C and D sets. These varistors were coded 312D and 312E, but for brevity in this paper, they are referred to as V_1 and V_2 respectively. Also shown in Fig. 1 is the dc resistance of these varistors as determined by the ratio of applied voltage to current. It is
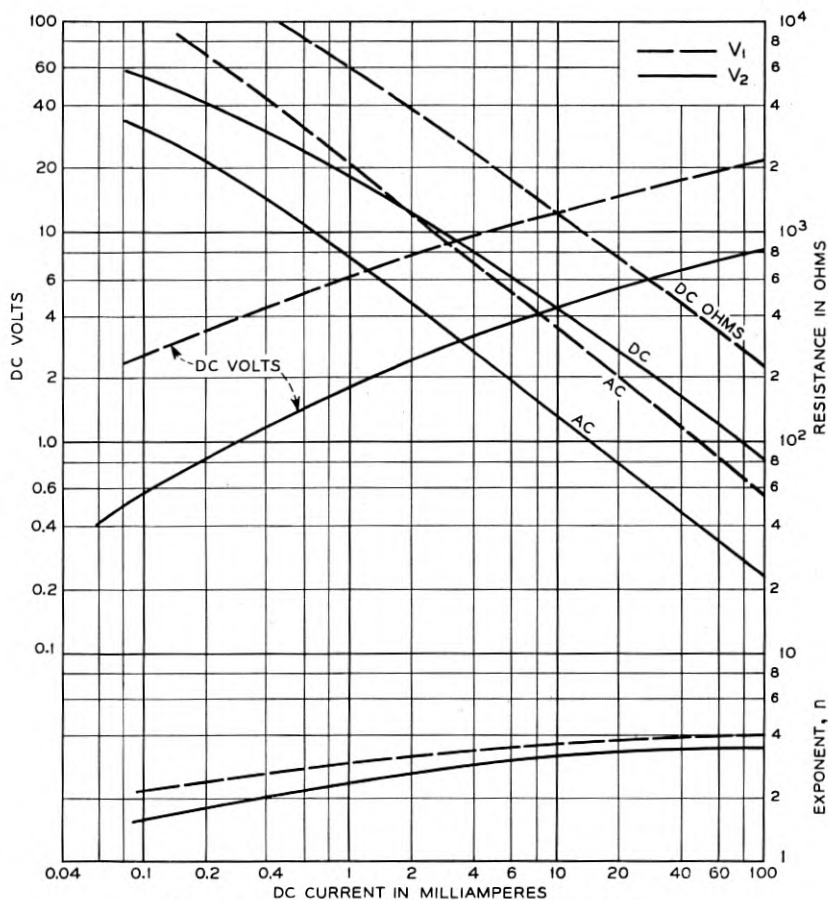
Fig. 1 — Characteristics of V_1 and V_2 varistors.

observed that the resistance decreases markedly with increasing voltage or current.

The direct current characteristics of a silicon carbide varistor may be expressed in a simple form by

$$I = AE^n, \tag{1}$$

where I is the direct current through the varistor and E is the applied dc voltage. When $n = 1$ the equation (1) reduces to an ohm's law relationship. A and n are generally regarded as constants when considered over a limited range in current. However, over the range shown in Fig. 1, it may be seen from the curvature of the voltage — current char-

acteristic that this is only a rough approximation and can be considered as valid only over a limited range in current. The value of the exponent n, which on a logarithmic scale is the reciprocal of the slope of the voltage-current characteristic and which also is plotted on Fig. 1, is seen to increase appreciably with increasing current. For example, n for the V_1 varistor covers a range from about 2 to 4.

In a telephone set the alternating signal currents are small relative to the direct currents. Therefore, in explaining how the silicon carbide varistor controls the transmission equalization and sidetone, it is necessary to examine how its ac resistance to small signals varies with direct current.

Equation (1) can be put in the form

$$E = \left(\frac{I}{A}\right)^{1/n}. \tag{2}$$

Differentiating this with respect to I we get for the ac resistance to small signals

$$\frac{dE}{dI} = \left(\frac{1}{A}\right)^{1/n} \frac{1}{n} I^{(1/n-1)} = \frac{1}{n} \frac{E}{I}, \tag{3}$$

where E/I is the dc resistance at the particular current at which n also is evaluated.

Thus, we see that the ac resistance of the varistor is the dc resistance divided by the exponent n, and further, that the ac resistance is con-
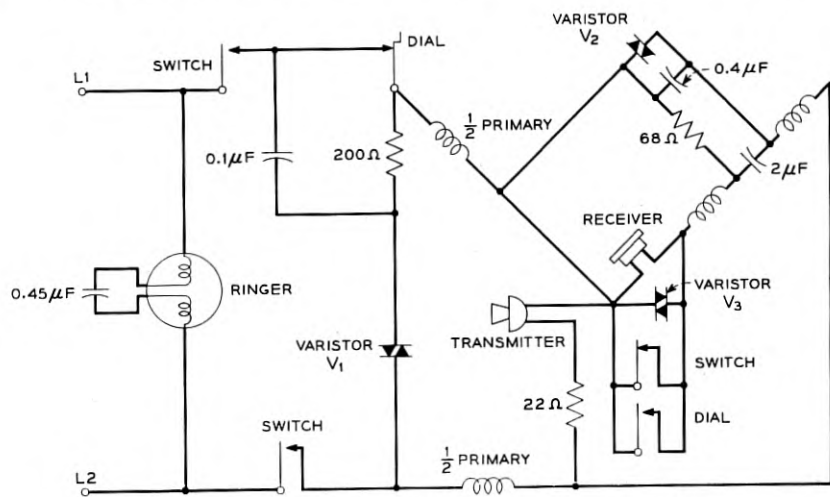
Fig. 2 — Circuit schematic of 500D telephone set.

trolled by the value of direct current. This is of great significance in the control of sidetone and transmission equalization, as will be apparent later.

Silicon carbide varistors have been used as circuit elements for various purposes for a number of years. In fact, the V_1 varistor was used in the original design of 500-type set to protect the tungsten filament against abnormal voltages. In this instance it was shunted directly across the filament, so that its inherent property of decreasing in resistance with applied voltage would serve to bypass current from the filament and thereby prevent damage when the applied voltage became too high. The requirements imposed by the transmission equalization and sidetone balance functions for the 500-type set are quite severe. A new varistor, the V_2, had to be developed which not only was much lower in resistance, but also maintained the n of the voltage-current characteristic at a high value. The successful outcome of this development and of the accompanying problems of design and manufacture, were of critical importance in making possible the simplification and economies offered by the 500D set.

TRANSMISSION EQUALIZATION

Fig. 2 is a circuit schematic of the 500D telephone set. This circuit is a modification of the circuit used in the 500B telephone set, which in turn is one of the Campbell anti-sidetone circuits[2] that has been standard in the Bell System for a number of years.

A simplified diagram of the transmission portion of the circuit is shown in Fig. 3, and it can be seen that there are three branches in the
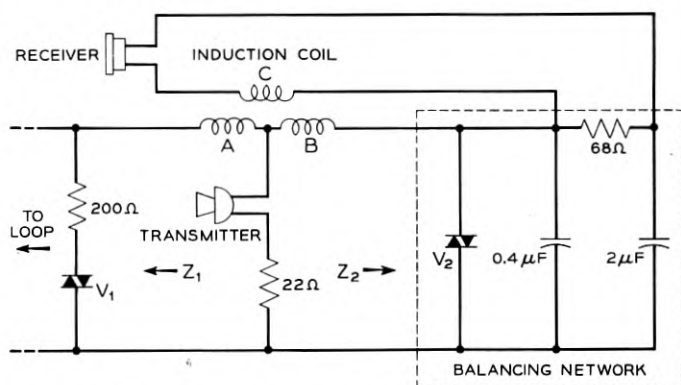
Fig. 3 — Diagram of transmission circuit of 500D telephone set illustrating side tone balance.

circuit which carry direct current. One is through the combination of the varistor V_1 and the 200-ohm resistance, which is bridged across the set. Another is through the winding A of the coil, the transmitter and 22-ohm resistance. The third is through winding B of the coil and the varistor V_2.

Fig. 4 shows the dc and ac resistance of the V_1 and V_2 varistors at various loop currents when connected in the telephone circuit in the normal way. The magnitude of the direct currents which flow through the varistors under these conditions differs from the currents shown in Fig. 1 and, therefore, the resistances are different. However, the inherent property of decreasing in resistance with increasing current is evident, as is also the fact that the ac resistance is much lower than the dc resistance. From this it is apparent that the varistors will introduce shunt losses and that these losses will be greater for ac than for direct currents because the resistance of the varistor is lower in the former case.

The variation in transmitting plus receiving loss with loop current is shown in Fig. 5. These curves are plotted in terms of the level of the 302-type set. If no equalization were provided in the circuit, the full transmitting and receiving gain offered by the new transmitter and receiver would be obtained at all loop currents. This condition is indicated by the broken line at the top of the diagram. With the equalization introduced by the varistors, the loss is graded from an insignificant
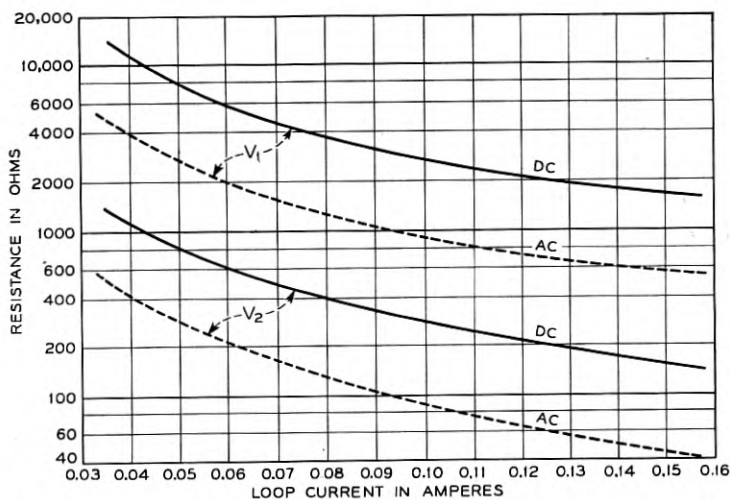
Fig. 4 — Characteristics of V_1 and V_2 varistors in circuit of 500D telephone set.
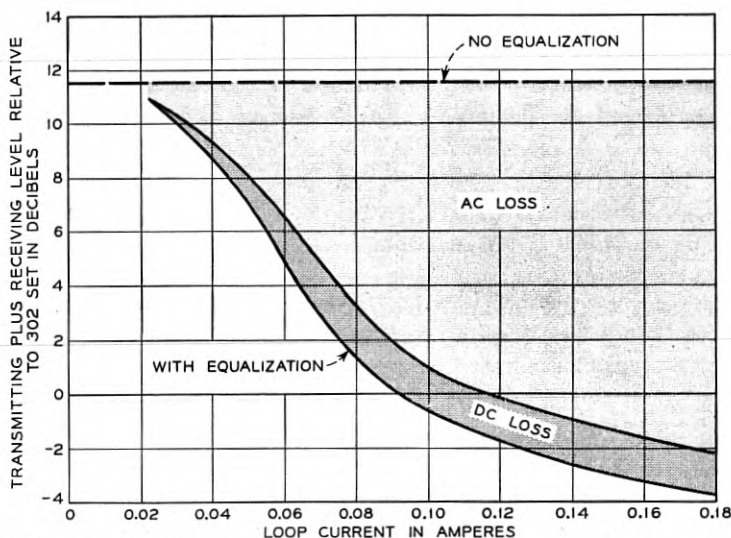
Fig. 5 — Level of 500D telephone set at various loop currents with and without equalization.

amount for long loops to a maximum for short loops. It is important to note that the greatest portion of this loss is the ac shunting loss introduced by the varistors.

SIDETONE BALANCE

Before considering the matter of sidetone balance, it will be helpful to examine some of the features of the telephone set circuit which are required for good receiving and transmitting performance. For simplicity, the induction coil is shown in a hybrid arrangement in Fig. 3. The speech currents received from the loop pass through the windings A and B of the induction coil which are so connected that they produce additive voltages in winding C which is connected to the receiver. These additive voltages would cause a resultant current to flow in the network resistance (68 ohms), were they not opposed by an approximately equal voltage 180° out of phase which results from the receiving current in winding B. Therefore, there is little or no voltage drop across the network resistance. Maximum receiving levels are thus obtained without appreciable power loss in the network resistance.

The transmitter of the set is made low in resistance because it is in series with the loop and influences the permissible loop range of the set from the standpoint of supervision. The impedance Z_1 looking toward

the loop is relatively high. To obtain acceptable transmitting levels, the low impedance of the transmitter must be approximately matched to the high impedance loop by providing a proper ratio of turns between windings A and B. Therefore, the impedance Z_2 looking toward the balancing network must be made low and since an important element of this network is the V_2 varistor, this had to have an unusually low resistance. This is the reason why it was necessary to undertake a special development to make available a silicon carbide varistor having suitable characteristics.

Sidetone would be eliminated if the vectorial sum of the voltages in the mesh which includes the receiver, and which arise as a result of speech and noise picked up by the transmitter, were zero. The complete elimination of sidetone is not desirable, but the objective is to keep it at a low level by a balance of opposing voltages. To achieve this result any voltage developed in the local transmitter is divided in the windings A and B so that the voltages induced in winding C are opposing. Furthermore, the voltage across the network resistance arising from current flowing in winding B is arranged to oppose this resultant of voltages induced in winding C. The overall effect of this balance is that the current in the receiver as a result of voltages developed in the transmitter is small. Also, this result gives maximum transmitting levels because there is little power loss in the receiver. However, to maintain the balance which gives low receiver currents the impedances Z_1 and Z_2, as affected by the transformation ratio of the coil, must be comparable. This is the key to good sidetone balance — that the impedances Z_1 and Z_2 be effectively matched both in magnitude and phase.

Now in the telephone plant, the impedance looking from the station toward the loop varies widely from one installation to another, and even from one call to another. The loop may be long or short, of small gauge or large gauge, and composed of cable or open wire or combinations of the two. Furthermore, the loop may be loaded or non-loaded and it may be connected through central office circuits of different characteristics to a trunk or distant loop and telephone set which also may cover a range in impedance. It is quite obvious that under these conditions the impedance looking toward the loop will not only vary over a wide range in magnitude, but may be inductive, or capacitive or essentially resistive.

If we assume that the V_2 varistor were not present, the impedance Z_2 looking toward the balancing network is not influenced by loop current. This is the situation that has obtained in balancing sidetone in preceding designs of telephone set. One of the impedances to be matched

was fixed and the other varied over a wide range. The sidetone balance under such conditions represented the best compromise that could be made among a large number of uncontrolled factors.

Let us examine next how the varistors affect this balancing problem and consider the influence of long and short loops. From what has already been shown, the varistor is a variable impedance element — that is, both its dc and ac resistance depend on the direct current. If the loop is long and composed of open wire the impedance looking toward it is high, perhaps 1200 ohms, and the direct current is low. The dc resistance of the V_1 varistor is then so high (approximately 10,000 ohms) that it has no appreciable effect on the impedance Z_1. Even the ac resistance of about 4,000 ohms does not have an appreciable shunting effect on Z_1. On short loops and local connections where the current is large and the impedance looking toward the loop may be of the order of 900 ohms, the dc resistance of 1,800 ohms has little effect. However, the ac resistance of 600 ohms does have an appreciable effect on Z_1, bringing about a better impedance match.

The impedance looking toward the balancing network and receiver without V_2 connected in the circuit is approximately 75 ohms, and is not affected by the loop current. If we go through the same process of examining the shunting effect of V_2 on the impedance Z_2, we see that at small loop currents both the dc and ac resistance of V_2 are too high to have any appreciable effect. At large loop currents the dc resistance of V_2 is under 150 ohms, but the ac resistance has dropped to well below 50 ohms and has a decided effect on the impedance.

Since the varistors are essentially pure resistances they tend to make the impedances Z_1 and Z_2 presented to the coil more resistive than they would be if the varistors were not present. This has the effect of reducing the difference between loop impedances which are capacitative, inductive, or resistive from the standpoint of phase angle and thereby improves the impedance match.

The overall effect of the automatic changing in resistance of the V_1 and V_2 varistors is to provide better matching and thereby reduced sidetone. This is shown in Fig. 6, where the loudness level of sidetone of the 500D set is plotted against loop current. It will be noted that with V_1 and V_2 open circuited, there is a continuing increase in sidetone loudness level with increasing loop current. When both V_1 and V_2 are connected in the circuit a decrease in sidetone level of 10 db or more is obtained at medium and large loop currents. The decrease is less at small loop currents. The close interrelation between the varistors and the circuit is well illustrated by the fact that if either the V_1 or V_2
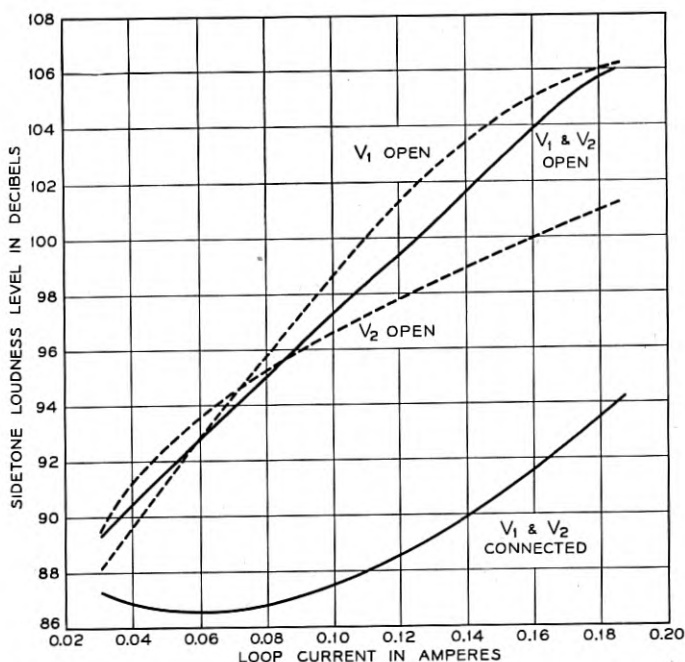
FIG. 6. — Effect of V_1 and V_2 varistors on sidetone loudness levels of 500D telephone set for average speech.

varistor is opened, almost all the effects of sidetone balancing are lost. It requires the presence of both V_1 and V_2 to obtain the impedance matching necessary for good balance.

From all that has been said, it should not be thought that the sidetone balance of the 500D set is perfect under all plant conditions. The impedance looking toward the loop varies so greatly in magnitude and phase that the best overall balance involves compromise. However, the balance provided is as good as with the original design of 500 set, which was considerably better than that attained by any previous design of commercial telephone set.

OTHER MODIFICATIONS

The introduction of the new equalization and sidetone controls in the telephone set has required a number of changes in other components of the set to realize all possible economies. Fig. 7 is a photograph of the transmission network (425A) which was employed in the earlier 500 A and B type sets. In the new arrangement shown in Figure 8
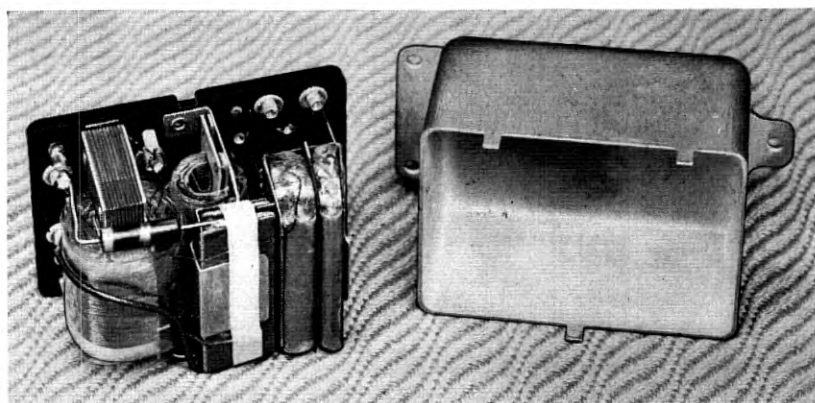
Fig. 7 — 425A Network used with early design of set — the 500A and B type.

the V_1 and V_2 varistors have been included in the network (425B), thereby providing mechanical protection for these elements. It has been necessary to design a new terminal block which is the plastic member to which the components of the network are mounted.

OVERALL RESULTS

By means of the V_1 and V_2 silicon carbide varistors and the rearrangements made in the circuit, the overall transmission performance of the modified set has been made substantially the same as that provided by the original design of set. The savings in manufacturing cost resulting from this modification are significant. Figure 9 shows on the left the
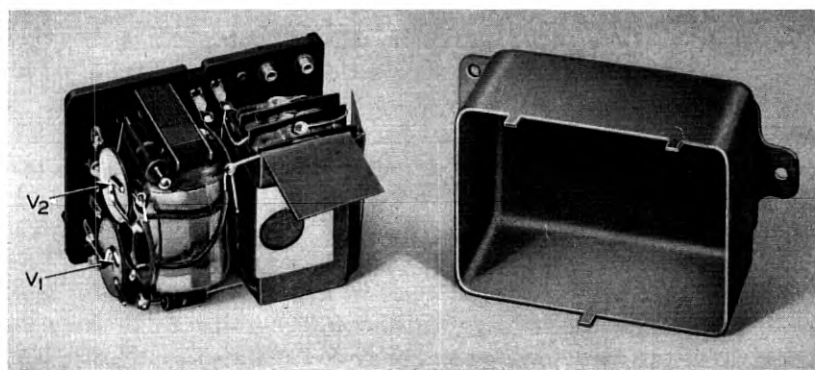
Fig. 8 — 425B Network used in present production set — the 500C and D type, showing the V_1 and V_2 varistors.
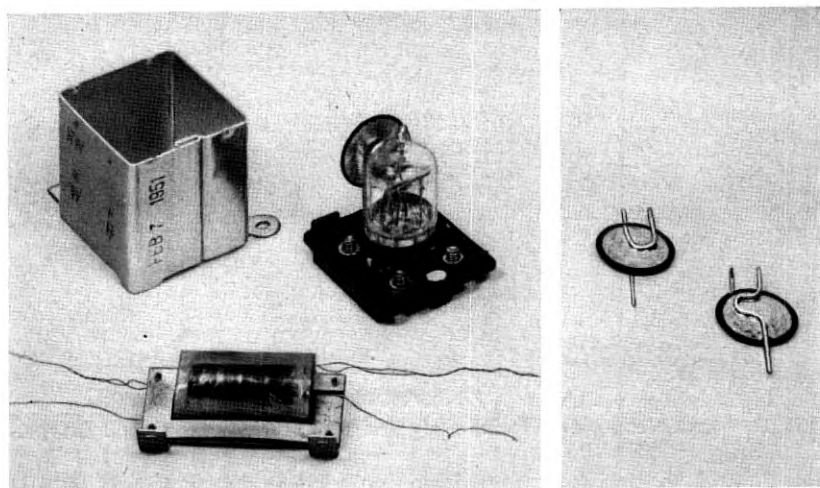
Fig. 9 — The 311A equalizer and autotransformer on the left are replaced in the 500C and D set by the V_1 and V_2 varistors on the right.

parts which have been eliminated — the 311A equalizer and auto-transformer. On the right are shown the V_1 and V_2 varistors which provide essentially the same performance as the replaced parts.

Another advantage of the change in design, though a difficult one to evaluate, is that of dependability. The silicon carbide element has an extremely long life and in this respect is inherently better than the type of equalizer which involves the use of a heated tungsten filament.

OPERATION IN NOISY LOCATIONS

Since the previous paper on the 500 set was published,[1] a study has been completed on the comparative performance of the 500-type set and its predecessor, the 302 type, in locations where room noise is severe. This is of particular interest, because the sidetone characteristics which have been discussed play an important part in transmission performance under such conditions.

For many years the securing of improved performance of the telephone set under noisy conditions has been a serious problem. It has been established that receiving impairment in a noisy location is far more limiting than the transmitting impairment resulting from the noise picked up at the noisy location and delivered to the distant party. This is due to the fact that the signal to noise ratio in the transmitting direction is improved by the natural tendency of the speaker to raise

his voice when the ambient noise level is high. A major factor in the impairment of received speech in the noisy location is the masking of received speech by noise picked up by the transmitter and delivered via the sidetone path to the receiver. Therefore, a telephone set which minimizes the masking sidetone noise during the listening interval should operate more satisfactorily in noisy locations. In many previous instances where the noise conditions were extreme, a "push-to-listen" switch was provided to cut off the local transmitter while listening and thus prevent the introduction of noise to the receiver via the sidetone path. With its higher efficiency and better sidetone balance, the 500-type set approaches the good performance of such a "variable" set under noisy conditions and provides this adequate performance to the user more conveniently and with the expenditure of less effort.

Laboratory tests have been made which permit appraisal of the merits of various telephone sets or modifications of them under high noise level conditions. These include talking and listening tests between two telephone stations under typical loop and central office conditions with an adjustable amount of typical room noise at the listening end. During the tests the length of trunk between the two stations was increased until the received speech was just sufficiently intelligible to permit carrying on a conversation. This threshold of intelligibility expressed in db of trunk loss was taken as a criterion of the performance of a given set.

Fig. 10 shows the results of the tests made with the 500-type set compared with the 302-type set, and with variations of the 500 type set. 500C- and D-type sets having silicon carbide varistors were used in this comparison, but in this respect the performance is no different than for the earlier design (500A and B). The curves shown are for the average of eight observers.

The abscissae of the curves are the noise levels at the listening end expressed in db above 10^{-16} watts per sq. cm. For reference purposes it should be kept in mind that the average noise in a fairly large business office is, on this scale, approximately 65 db. For further orientation as to the significance of the noise levels shown, it should be noted that when a level of about 100 db or higher is reached, it is extremely difficult to carry on a face-to-face conversation where no telephone is involved. The ordinates show the relative trunk loss in db over which it is possible, under the stated conditions, to just carry on a telephone conversation. Therefore, the larger the trunk loss, the better the performance.

It is seen from Figure 10 that the 500-type set is from 6 to 8 db better than the 302 set over the indicated range of ambient noise conditions.
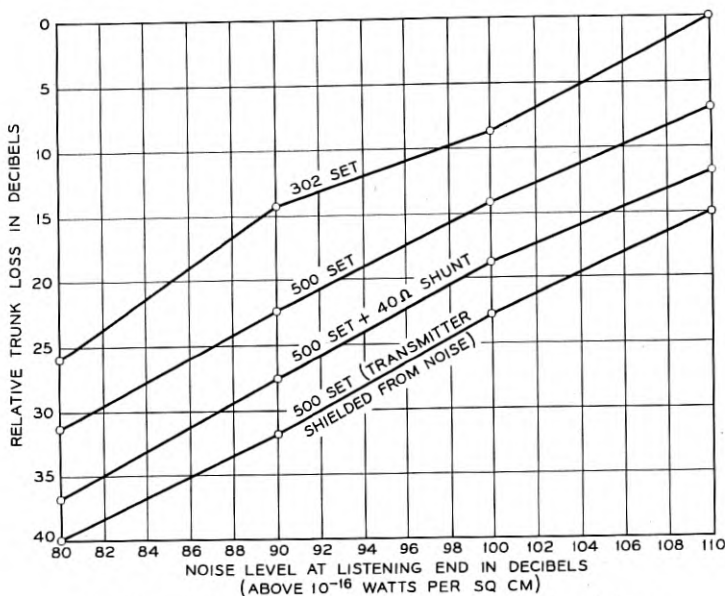
Fig. 10 — Effect of noise at listening end on trunk loss.

This improvement is attributed to the following:

1. The better sidetone balance of the 500-set circuit reduces the level of sidetone noise picked up by the transmitter and reproduced by the receiver, thereby providing a gain in signal to noise ratio.

2. The acoustic impedance looking from the ear canal into the U1 receiver[3] of the 500-type set is lower than the acoustic impedance of the HA1 receiver of the 302 set and, therefore, the acoustic noise pressure built up in the ear canal by leakage around the receiver cap is lower than when the HA1 receiver is used.

3. The output from the transmitter of the 500-type set under noise conditions contains less extraneous distortion products, giving it a character which is less disturbing than with the 302 set.

Included in the laboratory tests were experiments in which the output of the transmitter was intentionally lowered. It is known that a speaker in a noisy location instinctively raises his talking level. Since it is desirable to keep the sidetone level low to improve the signal to noise ratio, the output level of the transmitter at the listening end was reduced about 10 db by shunting a resistance of approximately 40 ohms across the transmitter. The sensitivity of the transmitter is reduced

and sidetone noise is directly lowered. The level of the received speech is unaffected. Therefore, the signal to noise ratio in the receiving direction is improved. The transmitting level is not too adversely affected, because the tendency of the subscriber to increase his talking level when the room noise level is high largely offsets the lower efficiency of the transmitter to speech sounds. These effects are borne out in the applicable curve of Fig. 10 which shows that an improvement of approximately 5 db in trunk results from the shunting of the transmitter.

Carrying this experiment further and acoustically shielding the local transmitter completely from noise, the lower curve of Fig. 10 was obtained. This is equivalent to shorting the transmitter as is done in the earlier "push-to-listen" types of telephone set, except that it does not introduce any electrical effects in the circuit. This shielding provides about 4 db additional discrimination against noise. This, then, indicates the limit to which we can go in reducing the effects of room noise on received speech by operating on the transmitter element alone.

From the data presented in Fig. 10 it is evident that the 500-type telephone set provides a significant improvement when the subscriber is carrying on a telephone conversation under noisy room conditions. While it has been indicated above that the greatest gain of the 500-type set over the 302 type under high room noise conditions is with respect to received transmission, it is also better in transmitting from noisy locations. This is because of the less disturbing character of the transmitted noise and because the signal to noise ratio is improved by having the transmitter closer to the subscriber's lips, which is a feature of the new handset.

Where noise conditions in the field are severe, the 500-type set will provide material improvement. Typical of such conditions are noisy business locations, telephone stations which are located in power plants and near loading ramps at airports. Where the noise is particularly severe, the provision of a 40-ohm shunt resistance on the transmitter of the 500-type set offers still further improvement and is recommended for application by the local telephone people.

CONCLUSION

The simplification of the 500 set made possible by the use of silicon carbide varistors, with its attendant reduction in manufacturing and maintenance costs and increase in life represents a significant step forward in station set design. A valuable characteristic of the 500-type set is the substantial improvement in transmission performance under conditions of severe ambient noise.

REFERENCES

1. Inglis, A. H., and W. L. Tuffnell, An Improved Telephone Set. Bell System Tech. J. **30,** pp. 239–270, April, 1951.
2. Gibbon, C. O., The Common Battery Anti-Sidetone Subscriber Set. Bell System Tech. J. **17,** pp. 245–257, April, 1938.
3. Mott, E. E., and R. C. Miner, The Ring Armature Telephone Receiver. Bell System Tech. J. **30,** pp. 110–140, Jan., 1951.
4. Inglis, A. H., Transmission Features of the New Telephone Sets. Bell System Tech. J. V. **17,** pp. 358–380, July, 1938.

Automatic Line Insulation Test Equipment for Local Crossbar Systems

By R. W. BURNS and J. W. DEHN

(Manuscript received February 9, 1953)

Moisture entering faults in the insulation of subscriber lines provides so-called "leakage" paths which reduce the insulation resistance. Testing the insulation resistance of all lines under the environmental conditions which tend to produce these leakages is a maintenance technique, relatively new, for detecting the insulation defects. The faults can then be corrected before they become serious enough to affect the customers' service. Subscriber reports are thereby reduced and the correction of the faults on a preventive maintenance basis tends toward a more uniform work load for the repair personnel. Rapid testing of the lines is necessary, otherwise the environmental conditions may change and the leakages will disappear without detection.

Rapid line insulation testing is practiced quite generally in all the switching systems throughout the Bell System, but the testing arrangements used are wholly or partially manually controlled in the testing and recording operations. While the benefits derived from rapid line insulation testing apply to all systems alike, this article is confined to a discussion of the entirely automatic testing and recording arrangements which are now being introduced in the No. 1 and No. 5 crossbar systems.

INTRODUCTION

The insulation resistance of subscriber lines is an important consideration in the design and operation of central office switching circuits. If the insulation resistance between the two conductors of the line, or the insulation resistance from the "ring" or "battery side" to ground, becomes low enough, the "leakage" current flowing produces the same effect as lifting the handset and failing to dial or to pass a number to the operator. This condition is called a permanent signal and the line is said to be "permanent." As long as the condition persists, the line is out of service both to incoming and outgoing traffic. Insulation resistance of a slightly

627

higher value will not cause a permanent signal, but its presence may interfere with other circuit functions, for example, dial pulses are distorted and a wrong number may be reached, the ringing signal may be tripped before the called party answers, or the switching circuits may fail to restore on hang-up of the receiver. If the insulation resistance is at least as high as the design limit, failures of the kind described above will not occur. The design limit for some switching systems used in the Bell System is 10,000 ohms; for others 15,000 ohms.

WHERE LINE LEAKAGES OCCUR

The outside plant distribution system for exchange lines usually consists of some underground cable, many miles of aerial cable to distribute the line conductors throughout the area and a small amount of open wire on the fringes. The insulation resistance of the cable conductors is normally quite high. If, however, a break in the cable sheath occurs moisture may enter and be absorbed by the paper insulation of the conductors near the break. This reduces the insulation resistance of the affected cable pairs. Sheath breaks may exist for a considerable length of time without reducing the insulation resistance sufficiently to cause any reaction on customer service. Eventually these sheath breaks will, if not detected and repaired, permit the entrance of sufficient water during a rain to reduce the insulation resistance to the point where permanent signals occur on several pairs. Then repairs must be made on an emergency basis to guard against a complete failure of all line circuits in the cable. One of the common causes of sheath breaks in some residential areas is gnawing by squirrels — squirrel bites.

Cable terminals are located on the poles or on the cable for making drop wire connections to the customers' premises. Binding posts mounted in a face plate within the terminal are connected to some of the cable pairs through a terminal cable stub joined to the aerial cable. Each cable pair is thus connected to binding posts in about four or five terminals on the average. When water or condensation wets the face plate, leakage currents will flow between the binding posts. If dirt and dust have accumulated on the face plates, the combined resistance of all leakage paths in parallel across the pair or to ground may become relatively low.

While open wire makes up only a small part of the outside plant circuits most areas have some lines containing from a few spans up to several miles of open-wire conductors. It is difficult to keep the open-wire plant in a condition so that it will be free from leakages under wet

weather conditions. Trees grow up into the wires and during rainy weather the branches often drop across the wires. When this occurs at many points on the open-wire run the combined leakage along the pair may become very low. In some localities, moss growing on the wires, salt spray or heavy fog conditions causing leakage at insulators are additional causes of low insulation resistance on open-wire lines.

Drop wire used to make the connection from the cable terminal to the customer's premises is subject to damage by abrasion and the insulation deteriorates from the effects of the weather. The old style of drop wire, a large amount of which is still in use in the plant, is insulated with a rubber compound and covered with a water proof cotton braid. When this braid protection is lost due to abrasion or effects of the weather after many years in service cracks form in the rubber insulation. Moisture enters these cracks in wet weather causing low resistance leakages. It is expected that the latest type of drop wire with the tougher neoprene covering will withstand the hazards for a longer period of time than does the old drop wire but undoubtedly the end of its useful life will ultimately be determinable by measurement of the insulation resistance under wet weather conditions.

Inside wiring on the customer's premises often remains in service for a long period of time and the insulation deteriorates over the years. If the insulation is in poor condition, inside wiring will develop leakages during periods of prolonged high humidity indoors which occur frequently during the summer months.

EFFECTS OF WEATHER ON LINE INSULATION RESISTANCE

When weather conditions are favorable — clear weather with low humidity — the insulation resistance of the line conductors in the outside plant is quite high compared to the design limit for central office switching circuits. When the plant becomes thoroughly wet from a hard rain the insulation resistance drops very considerably because of leakages which are in parallel all along the line. Fig. 1 shows this very clearly. The data for the curves in this figure were collected in a special study conducted in 1931 to determine the insulation resistance distribution of exchange lines in underground and aerial cable and open-wire plant under different weather conditions — dry, humid and wet. About 6,000 dial lines selected at random in twelve large cities in different parts of the United States were tested under different weather conditions and no repairs were made throughout the study period except to correct unsatisfactory service conditions. The tests were made with the voltmeter
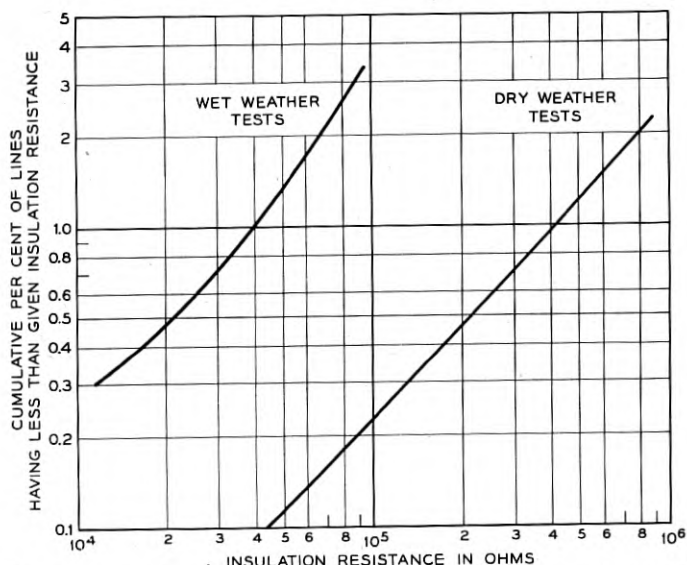
Fig. 1 — Wet weather versus dry weather tests — all types of Outside Plant combined.

test circuit of the local test desk which is used for testing subscriber lines reported in trouble. The speed at which lines can be tested from the test desk is necessarily slow because each number must be dialed or called individually therefore, only a small sample of lines was selected from any one office. Rapid line insulation test equipment was not in use in any area when this study was made.

The upper curve of Fig. 1 shows that about 0.3 per cent of the lines were below 11,000 ohms in wet weather while in dry weather only about 0.3 per cent of these same lines were below 140,000 ohms. Similarly, in wet weather two and one-quarter per cent of the lines were below 47,000 ohms but in dry weather only two and one-quarter per cent were below 900,000 ohms. The same wet weather test data summarized by types of outside plant is shown in the curves of Fig. 2. Lines in underground cable only, show the highest line insulation resistance. Those with aerial cable are next and those with some open wire have the lowest insulation resistance.

Fig. 3 shows the wet weather line insulation resistance distribution of 11,600 lines in an eastern city where the exchange outside plant had been thoroughly reconditioned prior to conversion from manual to cross-bar dial operation in 1940. The upper curve represents the approximate distribution before reconditioning and the lower curve represents the

distribution after the reconditioning was completed. The lower curve can be considered as representative of an outside plant in good condition as of the year 1940 when rapid line insulation testing was not yet used. Since that time, a considerable number of improvements leading to better insulation resistance characteristics have been made in outside plant items, such as drop wire and cable terminals, and a higher average insulation resistance would currently obtain. However, during wet weather, an undesirably large number of lines would still be closer than desired to the design limit.

MAINTENANCE WITHOUT RAPID LINE INSULATION TESTING

It may appear at first sight that the per cent of lines near or below the design limit is so small as to be unworthy of particular notice. However, an examination of this from a maintenance standpoint will prove otherwise. The testing of lines reported in trouble and dispatching of craftsmen on the outside to make repairs are handled from a local test center which on the average serves about 50,000 stations. There will be about six local test desk positions manned to do the testing and dis-
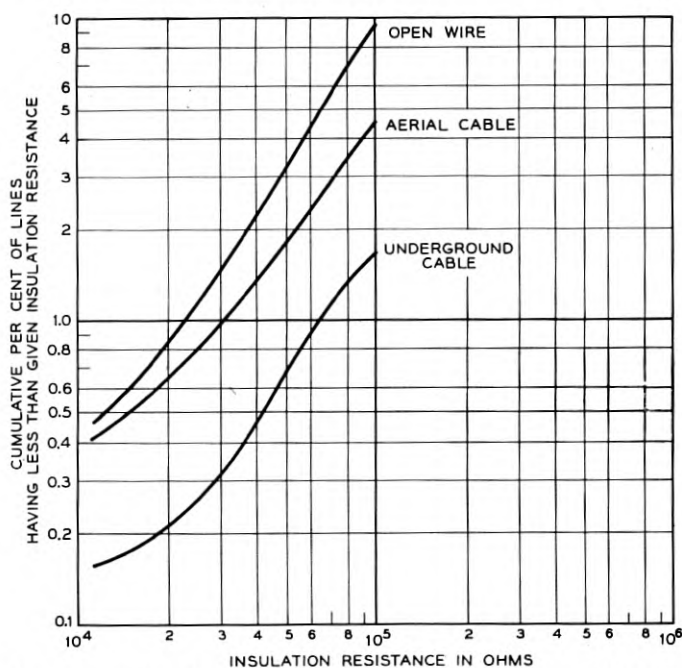

Fig. 2 — Wet weather tests — insulation resistance by types of Outside Plant.

patching work. Such a test center would handle at the current trouble rate about 120 reports on subscriber lines per day. If the outside plant is in a condition represented by the lower curve of Fig. 3 where 0.2 per cent of the lines are below 10,000 ohm resistance, this represents 100 additional stations in the region where service reactions may be expected from wet weather conditions. Consequently where these plant conditions obtain, there is usually a noticeable increase in subscriber's reports in wet weather and the repair load rises sharply. Without line insulation test equipment to test the lines rapidly while the low insulation resistance obtains, there is no way to pick out the worst lines from an insulation resistance standpoint after the weather has cleared. Visual inspections are costly and superficial indications do not always give reliable evidence of low insulation resistance. Consequently the large percentage of repairs to correct insulation defects are made as a result of subscriber reports and only a small per cent by routine preventive maintenance where rapid line-insulation testing equipment is not used.

MAINTENANCE WITH RAPID LINE INSULATION TESTING

The only experience to date with the fully automatic line insulation test equipment for crossbar offices is in the Media, Pa., No. 5 crossbar office. This test equipment has been in use for about one year and low insulation resistance cases recorded on each test have been investigated and repairs made. The condition of the outside plant in the exchange area is represented in Fig. 4. The curves in this figure are based upon the results of tests of the 4200 working lines in the summer of 1952 under very wet conditions of the outside plant.

The small percentage of lines below 40,000 ohms indicates that the poor insulation cases are being corrected well before reaching the stage where the customers' service would be affected. This is done with a minimum of maintenance effort as the test equipment spots the line automatically.

DETECTING SHEATH BREAKS

While sheath breaks in aerial cable are brought to light by rainy weather it is inadvisable to wait for rain to disclose them because of the possible serious effects on service and the need then for repairs on an emergency basis. Rapid line-insulation testing techniques have been very successful in disclosing sheath breaks in aerial cable. During the night the cable sheath cools and as the pressure inside decreases, outside air with a high moisture content enters the sheath break. The paper
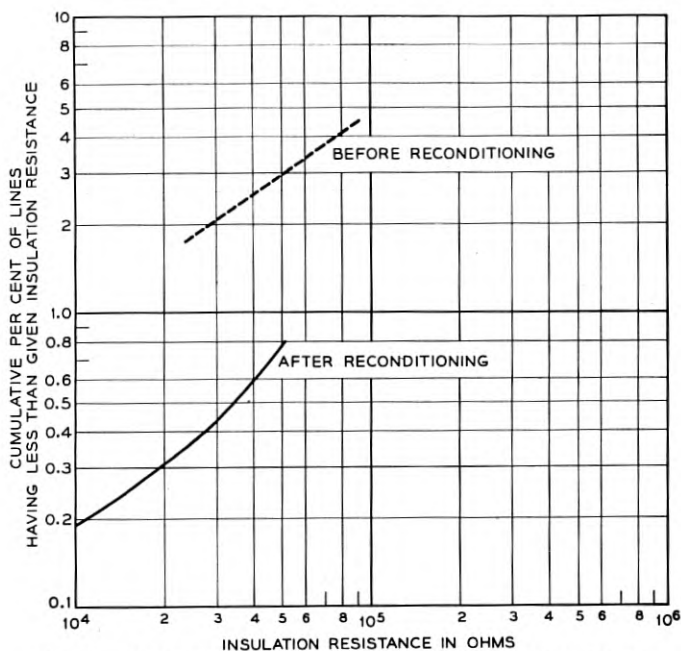
Fig. 3 — Wet weather tests — reconditioned Outside Plant.

insulation of one or more pairs near the break absorbs the moisture and the insulation resistance of the affected pairs is lowered. Tests made from about 3 A.M. to 5 A.M. are effective in detecting these leaks and by applying the latest fault locating methods, many of the sheath breaks can be located. The test equipment is so arranged that, for the most part, lower leak conditions on the line which may be present outside of the aerial cable will not register on these tests for cable defects.

BENEFITS DERIVED FROM RAPID LINE-INSULATION TESTING

With the aid of rapid line-insulation test equipment the maintenance personnel can correct the greater part of insulation defects on a preventive maintenance basis which results in a reduction of subscriber reports. Service to the customer is thereby improved, maintenance effort is reduced and repairs on an emergency basis often involving the expenditure of overtime are required less frequently. Rapid line-insulation testing is also of great value in rapidly determining the extent of storm or flood damage and makes it possible to direct immediate efforts where the greatest benefits will be derived.
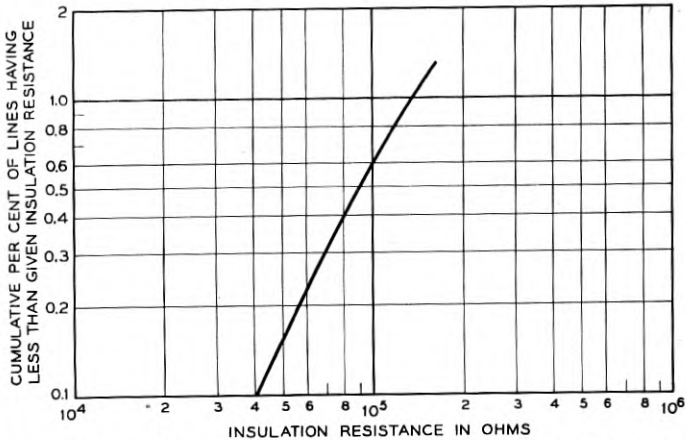
Fig. 4 — Wet weather tests — automatic line insulation testing equipment used.

CROSSBAR TEST EQUIPMENT — TYPES OF TESTS AND SENSITIVITIES

The crossbar line-insulation test equipment is arranged to make three different kinds of tests from the standpoint of the way in which the resistance measuring circuit (line-insulation test circuit) is connected to the line and to battery or ground. Each kind of test may be made in three different resistance ranges, hence there are nine different tests. When the test equipment is performing one of these tests a record is made of all lines which have an insulation resistance below the top limit of the particular test. Each test range is divided into three bands of resistance and the record indicates whether the insulation resistance lies within the first quarter (low band), the second quarter (medium band) or the upper half (high band) so that preference can be given to the worst cases in clearing the trouble. The test numbers, ranges and resistance bands are shown in the Table I.

SHORT AND RING GROUND TEST

The first kind of test is called "short and ring ground test." This test measures the leak in the way in which it affects the line circuits, pulsing circuits and supervisory circuits of the central office switching system. The line-insulation test circuit (LIT circuit) which is described later on is connected to central office battery potential and to the line in the manner shown in Fig. 5 (a). As indicated in this illustration, leakage resistance from ring to tip and from ring to ground are measured

TABLE I — CROSSBAR LINE-INSULATION TESTS

| Types of Tests and Test Numbers | | | Range | Resistance Bands* (kilo-ohms) | | |
Short and Ring Ground	Tip and Ring Ground	Foreign E.M.F.		Low	Medium	High
1	4	Not used	A	0–40	40–80	80–160
2	5	7	B	0–160	160–320	320–640
3	6	8	C	0–640	640–1250	1250–2500
Not used	Not used	9	D	0–1250	1250–2500	2500–5000

* The equipment is arranged so that by cross connection changes the bands of resistances in each range may be halved or doubled, if necessary, to meet local conditions.

in parallel. If the line-insulation resistance is more than the top limit of the test range, the test equipment proceeds to the next line. If, however, the resistance measured is less than the top limit the test equipment immediately retests the line with the central office ground removed as shown in Fig. 5 (b). This re-test measures the leakage from ring to tip. If leakage from ring to tip only is indicated the drop wire is probably the cause.

The No. 1 test is run under wet weather conditions to detect all lines which are below the top limit of this test (normally 160,000 ohms). If the maintenance force keeps these cases cleared out by running tests during every rain the wire plant can be kept in good condition. During long periods of dry weather all line insulation test indications previously recorded may have been investigated, then if a light rain occurs it may be desirable to run the No. 2 test to provide a satisfactory number of failure indications for subsequent investigation. The No. 2 test can also be used to detect leakages in inside wiring which occur frequently during the summer when houses are unheated and the inside humidity

(a)

FIRST TEST CONDITION MEASURES
COMBINED SHORT AND RING TO
GROUND RESISTANCE

(b)

RETEST AFTER FAILURE ON
COMBINED TEST MEASURES
SHORT ONLY
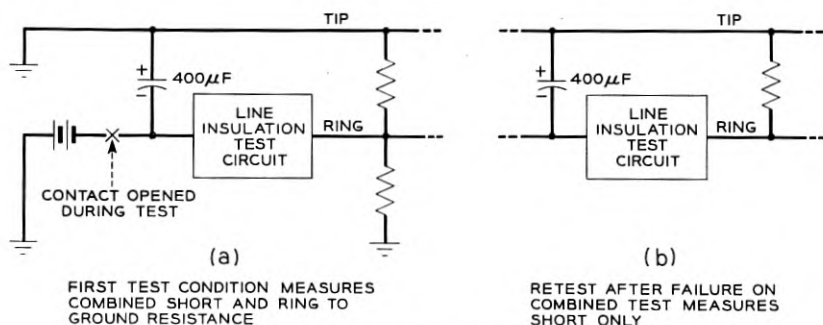
Fig. 5 — Short and ring ground test.

is very high. Tests to disclose insulation defects in inside wiring are made when the weather is clear with low outside humidity at which time leakages in other parts of the exchange plant will rarely be found.

TIP AND RING GROUND TEST

The "tip and ring ground" test can be made in three ranges of sensitivity as shown in Table I. The arrangement of the test circuit for this type of test is shown in Figure 6. The first test condition, Fig. 6 (a), measures the combined leakage resistance from tip and ring to ground. The tip and ring are connected together therefore leakage across the pair is not indicated. This eliminates practically all of the drop wire
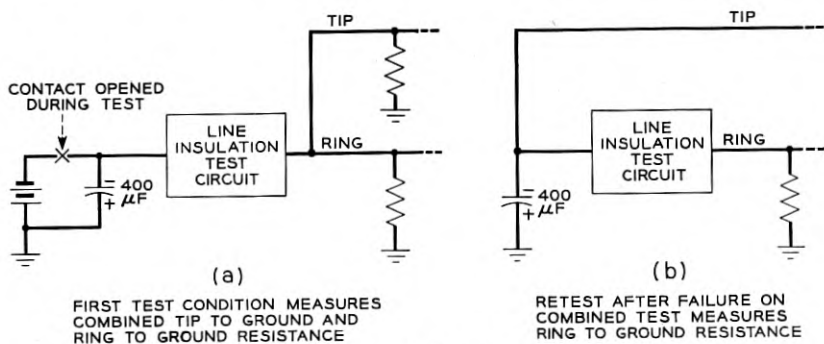
Fig. 6 — Tip and ring ground test.

leakages. If a failure is indicated on the first test condition, a re-test is made under the condition shown in Fig. 6 (b) to check only the resistance from ring to ground. If the ring tests clear it is known that the leakage is from tip to ground. This kind of test is of particular value in checking terminal face plate leakages which are predominately leakages from tip to ground. This type of test can also be used to check that tip conductors on party lines in message rate areas are free of low resistance grounds which might result in wrong party identification on calls made by the ring party.

FEMF TEST

The FEMF (foreign e.m.f.) test is used to measure leakages in cable to detect sheath breaks. These leakages are high resistance compared to other leaks which may be present across the pair connected for test or from the pair to ground. To make the latter ineffective so as not to

mask the higher resistance cable leakages the test circuit is grounded and connected to the line with the tip and ring short circuited as shown in Fig. 7. The test circuit measures the leakage current flowing from the pair connected for test to the ring conductors of other subscriber lines which are connected to battery potential in the central office. Leakages outside the cable will not cause leakage currents to flow through the line insulation test circuit. This test is called FEMF because it measures leak in the same way as does the test desk voltmeter when the FEMF test key is operated to connect ground instead of test battery to the voltmeter. Leakages as high as about 2 megohms can be successfully located after the cable has been identified by analysis of pairs affected by leaks.
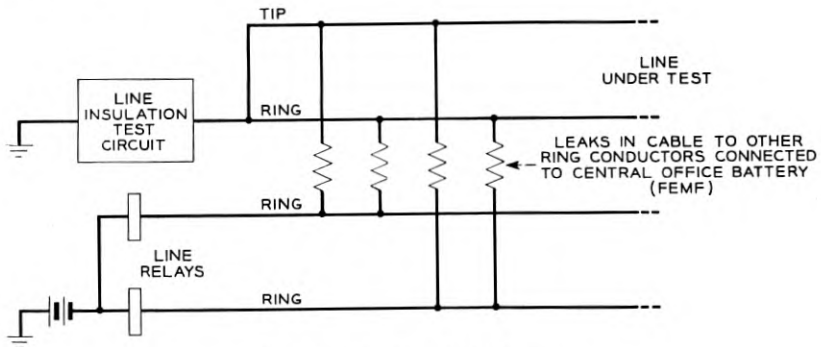
Fig. 7 — Foreign E.M.F. test.

TEST AND RECORDING CIRCUITS

The equipment for automatically measuring and recording line-insulation resistance consists of three principal parts. First, the device for measuring the low currents involved; second, the means for connecting this measuring device to each of the lines in rapid succession; and third, the apparatus required for recording the faults discovered.

The first of these, known as the line-insulation test circuit, is capable of detecting insulation resistance as high as ten megohms, with fast enough response, so that a satisfactory rate of line testing (about 10,000 to 12,000 lines per hour) is attainable. It is provided with a filter to attenuate both 25- and 60-cycle interference, so that leakage faults may be separated from other types such as induction from power lines. As shown in Table I, its sensitivity may be easily changed, so as to be compatible with atmospheric conditions, the kind of test to be made, and the condition of the outside plant. When a particular sensitivity or

"range" is chosen, the insulation fault detector will report all lines having lower insulation resistance than the chosen value, and as already stated it will further identify those lines having insulation resistance lower than one-half and one-quarter of this resistance.

This test circuit is connected to each of the lines in rapid succession by the line-insulation test control circuit. As shown in the block diagram of Fig. 8, the control circuit appropriates the control and testing wires between one of the markers (usually marker No. 2) and each of the line link frames. The arrangement shown is for a No. 5 crossbar office, but the No. 1 crossbar arrangement is similar in principle. By means of these appropriated connections the control circuit is able to select the

Fig. 8 — Block diagram of line insulation test control circuit connections.

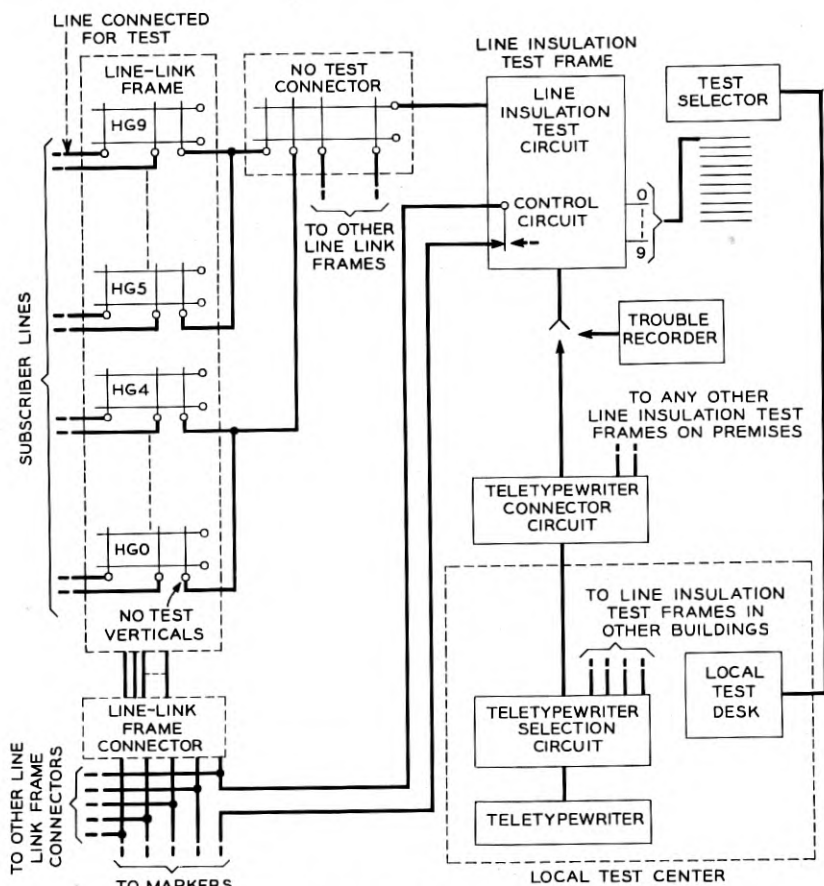lines and if the selected line is not busy its tip and ring conductors are connected to the insulation resistance measuring circuit (line-insulation test circuit). To establish this connection, the control circuit chooses an idle line link to connect the line vertical to the no test vertical and the test path is completed through the no test connector to the line-insulation test frame.

The operation of the control circuit may be started in the central office by operation of keys at the test frame, but it is more commonly started by remote control from the local test center. This permits insulation tests to be made even though the central office is unattended. When the control circuit is started, it must also be directed to make one of the three types of tests, and to choose one of the three test sensitivities for each of these types as shown in Table I. This is accomplished by operating one of nine keys at the test frame or by selecting the test trunk at the test center and dialing one of nine codes and then operating a key at the test desk. Either action chooses the type of test and the sensitivity, and causes the test circuit to start. A tenth key or a tenth code is used to stop the test before the end of a complete cycle, when this is required, so that the type of test or the sensitivity may be changed.

A pre-arranged regular pattern is followed in testing the lines. A line link frame is selected, a particular five lines are tested, and then the frame is released, and the next frame is selected, and another five lines are tested, and so on. Lines found busy, dial PBX trunks and line link frame terminations of intertoll trunks and similar circuits which would cause false operation of the test circuit are passed by without test. The line link frames are selected in order, beginning with the lowest numbered frames and continuing to the highest. When one cycle through the frames has been completed, another cycle will be started. However, on the next cycle, different groups of five lines will be tested. On the first cycle, the five lines in vertical group 0, horizontal group 0 will be tested, on the second cycle, the five lines of vertical group 1, horizontal group 0 will be tested, in each line link frame and so on until all lines in horizontal group 0 of all line link frames have been tested. On subsequent cycles the lines of other horizontal groups are tested in regular order.

Testing only five lines for each selection of a line link frame minimizes interference with traffic. Since one to two seconds is required for testing five lines, there will be only a slight delay to calls which require access to the frame. In addition, a heavy traffic load will cause the control circuit to stop insulation testing and restore to service the marker whose cabling it has been using. When the traffic has been reduced sufficiently, the test will be restarted automatically.

When a line fault is discovered, the control circuit makes use of either of two types of mechanism to record (a) the location of the line on the line link switches, (b) the type of test being made, and (c) a rough approximation of the insulation resistance. Since a substantial time is required to make this record, the line link frame is restored to service during this interval, to reduce interference with service. One of the above two devices is the trouble recorder, used only in the No. 5 crossbar offices. The control circuit connects itself to this machine in much the same way that a marker does, when it needs to record a trouble. Having made this connection, a card is perforated showing the line location and other data pertinent to the trouble indicated. The other device consists of teletypewriter equipment at the central office which transmits the required data to a teletypewriter page printer located at the local test center. The equipment at several central offices has common use of the same page printer at the test center and several test circuits in one building use the same teletypewriter transmitting equipment. This second recording arrangement is the more convenient of the two, since it produces the record at the place from which the activities of the outside plant repair force are directed. This arrangement is available for both No. 1 and No. 5 crossbar offices. A typical teletype record is shown in Fig. 9.

LINE-INSULATION TEST CIRCUIT

The insulation resistance measuring device is required to respond, not only to the very low current (five micro-amperes) obtainable with insulation resistance up to ten megohms, but it must also give good discrimination between resistance values in the order of 20,000 ohms. This is accomplished by desensitizing the measuring device by means of series and shunt resistors when a test using less than the maximum range is desired. A second requirement is that the measuring device be low in resistance so that the time constant of this resistance in combination with the line and filter capacity will be low enough to attain high testing speeds.

These leakage current amplitudes are so small that amplification is required in order to actuate the recording and control mechanisms of the measuring system. These small direct currents could be amplified by means of a dc amplifier. However, since it is easier to design and construct an ac amplifier of suitable stability, it is desirable to use a measuring device which has an alternating voltage output which varies with the direct current input.

A type of magnetic modulator, called a magnettor, which has these

desirable characteristics, is used as the current measuring instrument. It has a low impedance input circuit, in which the low amplitude dc leakage current flows. An alternating current of constant amplitude and frequency is supplied to separate windings of the magnettor, so that, as explained below, its output circuit delivers an alternating voltage which varies with the dc input. Fig. 10 shows the basic circuit, which operates as follows. Two identical windings, a and a′, and two other identical windings, b and b′, are wound on identical permalloy cores. Windings

LINE LINK FRAME

BAND ┐ ┌ VERTICAL GROUP

TEST NUMBER ──────── ┐ ┌────── HORIZONTAL GROUP

OFFICE ┐ │ │ ┌ VERTICAL FILE

```
0-82-0801-01
0-82-0704-10
0-82-0705-22
0-82-0007-21
0-82-0205-32
0-80-0406-63
0-82-0800-74
0-82-0601-70
0-80-0806-73
0-82-0100-83
```
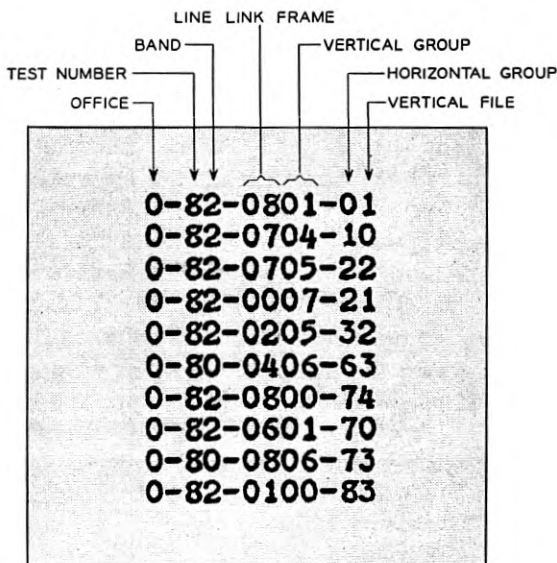
Fig. 9 — Teletype record of failures.

a and a′ are connected in series and supplied with an alternating current. The magnettor cores have a characteristic as shown in Fig. 11 (b), and the amplitude of the input voltage is great enough so that the core is driven to saturation on each half cycle. Fig. 11 (c) shows the resulting flat topped flux versus time curve*. Since the voltage induced in winding b is proportional to the rate of change of flux, it will have a wave form as shown in Fig. 11 (d). The voltage peaks occur when the flux rate of change is maximum, and during the "flat" intervals the induced voltage is small. Since the b and b′ windings are connected so that their output voltages are in opposition (see Fig. 11 (d)) the net output with no dc

---

* The wave shapes of Figs. 11 (c), 11 (d), 11 (e) and 11 (f) have been exaggerated to illustrate the action involved.

input will be zero. Manufacturing tolerances which produce dissimilarities in the windings and cores may cause a small output with zero dc input.

When the dc leakage current flows through windings b and b', a bias flux is established in each of the cores. This causes the shape of the flux versus time curve to be changed as illustrated in Fig. 11 (e). Since the ac input will saturate the cores, both half cycles of the wave will be flat topped, but the flat portion of one-half cycle will be increased and the other decreased. Also the steeper parts of the curve will be brought closer together on one-half cycle and further apart on the other.

As illustrated in Fig. 11 (e) the dc bias will have a different effect upon the flux in each of the two cores. In one core the bias aids the positive half cycle of the input current, and in the other the negative half cycle. The result is that the core reaches saturation more quickly and, therefore the flux curve is flatter on the half cycle which is aided by the bias.

This change in the shapes of the flux time curves produces corresponding changes in the shapes of the voltage-time curves of the output windings b and b'. These are shown in Fig. 11 (f). The peaks of the voltage curves occur at the points of maximum rate of change of the flux curves, and of course the flat portions of the voltage curves (near zero) are lengthened or shortened depending upon the flatness of the flux curve tops. This skewing of the two voltage curves prevents the output voltage cancellation which was obtained with no dc, and gives
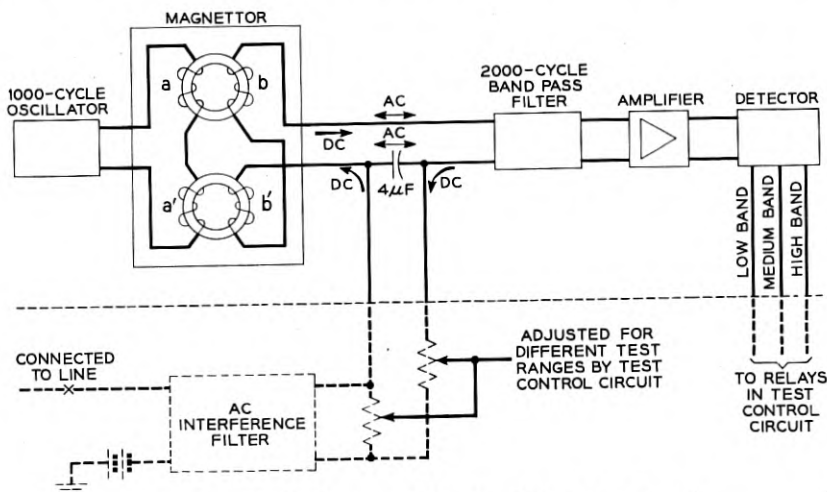
Fig. 10 — Block diagram of line insulation test circuit.

a resultant output which contains even harmonics of the input voltage. The second harmonic is selected by means of a filter for use as an indicator of the presence of leakage current.

The second harmonic is amplified by a three stage negative feedback vacuum tube amplifier, whose output is applied to the grids of three thyratrons. Adjustments are made, as described in the following, so that one of the thyratrons conducts when the insulation resistance is less than the value selected for the test, an additional thyratron conducts if the insulation resistance is one-half of this value or less, and all three conduct if the resistance is one-fourth or less of the selected value. Three relays, whose windings are connected in the anode circuits of the thyratrons, are actuated when the associated thyratrons conduct and cause a record of the trouble to be made, or cause the control circuit to select another line if the insulation resistance is above the range of interest.

In order to calibrate the detecting circuit, a test resistor of 160,000 ohms is connected to the magnettor by operation of a key. The amplifier gain is then adjusted so as to just cause conduction of the thyratron
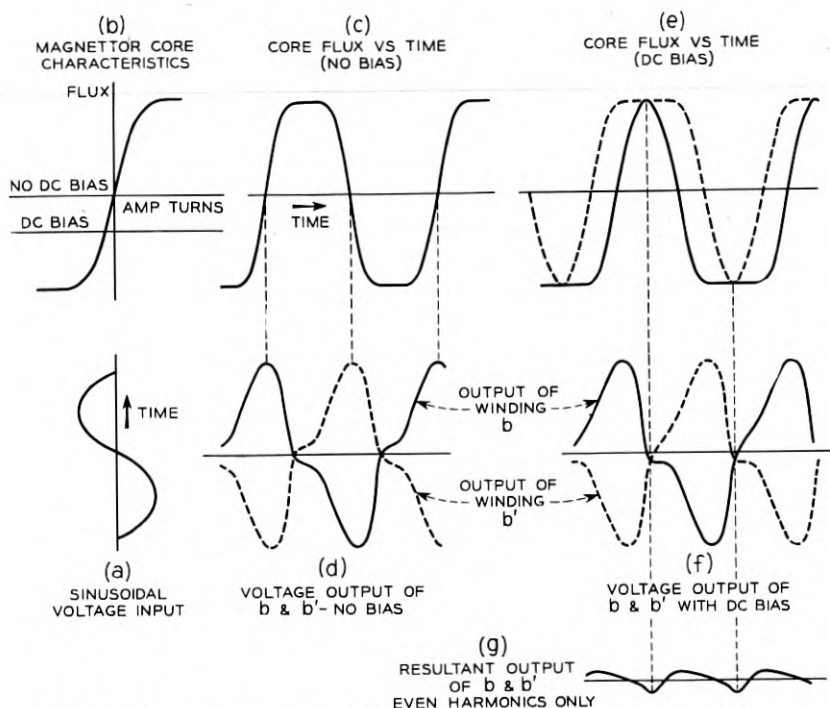
Fig. 11 — Graphical representation of voltage and flux in the magnettor.
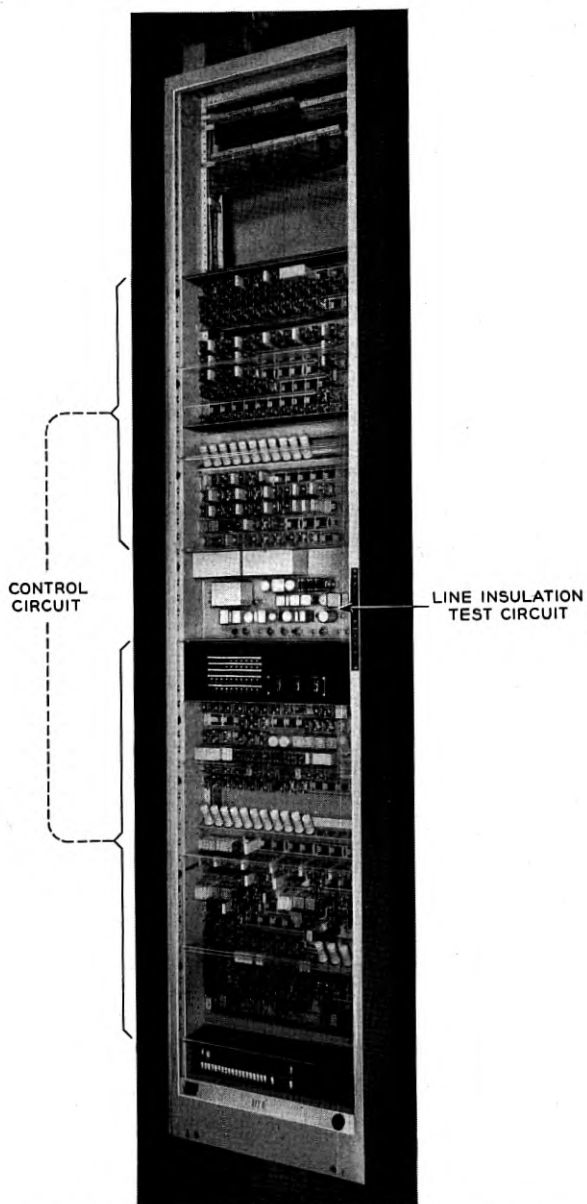
Fig. 12 — Front view of line insulation test frame.

which responds to the lowest insulation resistance value. Test resistors of 320,000 ohms and 640,000 ohms are then substituted and the grid bias of the other thyratrons adjusted so that they just conduct on the proper value of resistance. This procedure calibrates the device for one range, using a shunt and series resistor which corresponds to this range.

Facilities are provided for substituting suitable test resistors for checking the calibration of other ranges, with the shunt and series resistors for the range connected to the magnettor.

EQUIPMENT FEATURES

The apparatus components of the test control circuit and the line insulation test circuit are assembled and wired in an 11-foot bay, 23 inches wide. Fig. 12 is a front view of the test frame. Fig. 13 is an enlarged view of the line insulation test circuit equipment and the control panel which includes features for checking the accuracy of the test circuit and for calibrating it. When the teletype method of recording failures
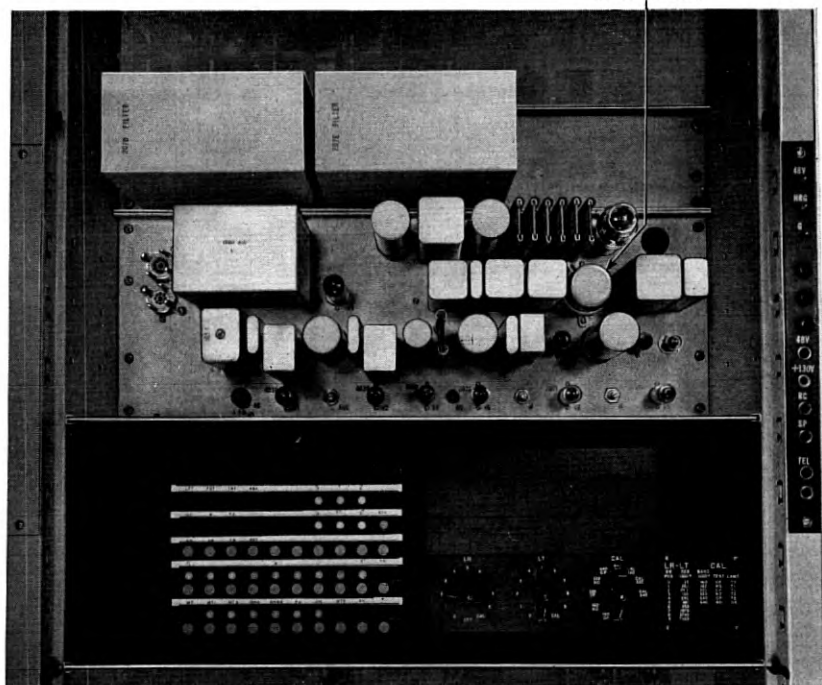
Fig. 13 — Front view of control and test panel (at bottom) and line insulation test unit (center) for No. 5 crossbar.

is used, the teletype transmitting equipment common to all line insulation test frames in a building is mounted in a separate 23-inch bay and occupies about one-third of the vertical space.

CONCLUSION

The initial installation of the automatic line insulation test equipment in a No. 5 crossbar office has been in operation for more than a year and the maintenance advantages of remote control of starting and automatic testing and recording have been fully demonstrated. It is expected that any future developments of line insulation test equipment undertaken for other switching systems will follow this same general pattern.

ACKNOWLEDGEMENT

Theory of Magnetic Effects on the Noise in a Germanium Filament

By HARRY SUHL

A magnetic field will influence the current noise in a germanium fila-ment. This fact bears out the hypothesis that at least part of the noise arises from minority carriers emitted in random bursts and recombining at the surfaces. A quantitative theory of this effect is given.

INTRODUCTION

In a series of fundamental experiments, H. C. Montgomery[1] has es-tablished that minority carriers play an important part in the current-noise associated with semiconductors. He found that on the one hand, the noise voltage is usually proportional to the biasing current, suggest-ing fluctuations in the conductivity, and hence the carrier concentration. On the other hand the spectrum of the noise suggested a rather coarse-grained time variation, not likely to be caused by fluctuations in the normal carrier density. One might conclude, therefore, that the noise is caused by a distribution of sources emitting or absorbing minority carriers in random bursts. Such carriers would be subject to the same laws of motion and of recombination as intentionally injected carriers. Montgomery was, in fact, able to verify that the noise along a filament showed marked correlation over a distance roughly equal to that through which minority carriers could drift in the biasing field before recom-bination.

W. Shockley has pointed out another corollary of this theory: A mag-netic field transverse to the filament should have a pronounced effect on the noise. This conclusion, too, Montgomery was able to verify experimentally.[1] His results are in good qualitative agreement with theory. Complete quantitative agreement was perhaps not to be ex-pected, since technical difficulties prevented attainment of the idealized conditions assumed by the theory. This paper gives an account of that

theory. On it are based the computed curves in Montgomery's paper showing the change in noise power with magnetic field.

To see how such a change comes about, we imagine the magnetic field applied normal to one pair of the long faces of a rectangular filament. This field, and the longitudinal drift current used to measure the noise, yield a sidewise thrust on the carriers, directed at right angles to the other pair of long sides. As a result the density distribution over the cross section is distorted, the minority carriers tend to accumulate near one of those sides, while the neighborhood of the opposite side is depleted. But for the usual conditions the recombination of carriers occurs mainly near the surface, and is proportional to their density there. Hence the magnetic field will change their lifetime.[2, 3] Clearly the amount of noise is dependent on the length of time carriers are able to contribute to the change in conductivity, that is, dependent on their lifetime. Therefore, the magnetic field should change the noise power. In simple extreme cases one can even make a semiquantitative argument for the maximum variation to be expected on the basis of such considerations.[1]

FORMULATION OF THE PROBLEM

In order to make an exact calculation, we require a few preliminaries: The conductivity g is supposed to undergo a small time-dependent fluctuation $\Delta g(t)$ about its mean value.

The fluctuation arises from certain sources each of which, for macroscopic purposes, may be considered to emit a noise-current $J(t)$ of minority carriers. Thus in a small time-interval dt' near t' the excess charge injected is $J(t') \, dt'$. This charge decays by recombination. Let $r(t - t')$ denote the fraction of carriers left over at time $t(>t')$. Then at time t there remains a charge $r(t - t') \, J(t') \, dt'$ of the original injection. Now provided the excess density is small compared with the mean density, $\Delta g(t)$ is proportional to the excess charge at time t, due to all the previous emissions added together. Therefore

$$\Delta g(t) \propto \int_{-\infty}^{t} r(t - t') J(t') \, dt'. \tag{1}$$

In practice we do not literally plot $\Delta g(t)$ as a function of t, but rather its frequency component $\Delta g(f)$ in a narrow range df of frequencies near f. In other words, we single out for observation the contribution to Δg from that part $J(f)$ of the injected current $J(t')$ which varies as $e^{-2\pi i f t}$ Suppose now that $1/f$ is large compared with the time over which $r(t)$

is appreciably different from zero (that is, let $1/f$ be much greater than the lifetime). Then, in the integral (i), $r(t - t')$ will have gone from unity to zero long before $J(f)e^{-2\pi ift}$ has changed appreciably from its value at $t' = t$. Therefore, for purposes of observation at frequencies much smaller than the reciprocal lifetime, we can rewrite (1) as

$$\Delta g(t) \propto J(t) \int_{-\infty}^{t} r(t - t')\, dt'$$

$$= J(t) \int_{0}^{\infty} r(t)\, dt. \tag{2)*}$$

The integral in (2) can be interpreted as the average lifetime of carriers. For, by definition, the rate of recombination at time t is $- dr(t)/dt$, so that $- (dr/dt)dt$ is the number of carriers recombining between time $t,\ t + dt$. Hence the average lifetime is

$$\tau = -\int_{0}^{\infty} t\, \frac{dr(t)}{dt}\, dt = -[tr(t)]_{0}^{\infty} + \int_{0}^{\infty} r(t)\, dt$$

$$= \int_{0}^{\infty} r(t)\, dt$$

since $tr(t) \to 0$ as $t \to \infty$.

If $1/f$ is not large compared with τ one cannot simplify the integral (1) in this way. One then has to consider separately each frequency component $\Delta g(f)e^{-2\pi ift}$ due to the current $J(f)e^{-2\pi ift}$. Then

$$\Delta g(f)e^{-2\pi ift} = J(f) \int_{-\infty}^{t} r(t - t')e^{-2\pi ift'}\, dt'$$

$$= J(f)e^{-2\pi ift} \int_{0}^{\infty} r(t')e^{2\pi ift'}\, dt'$$

or

$$\Delta g(f) = J(f)\tau(f)$$

where

$$\tau(f) = \int_{0}^{\infty} r(t')e^{2\pi ift'}\, dt'.$$

The calculation of $\tau(f)$ is more complicated than that of $\tau = \tau(0)$, and

---

* From here on the equality sign will replace the proportionality sign. The resulting change of units is of no consequence in the final results which are only concerned with ratios of conductivity modulations.

at the present time the experimental situation does not call for refinements of this kind. Therefore we shall restrict ourselves to the calculation of τ.

To evaluate τ it is not necessary to consider a time-dependent case at all. In our experiment, it is the mean square conductivity fluctuation which is actually observed. Hence from (2)

$$<\Delta g^2> \; = \; <J^2(t)> \tau^2.$$

If the emission processes are stationary in time, $<J^2(t)>$ is time independent:

$$<\Delta g^2> \; = \; <J_0^2> \tau^2.$$

Now τ can be written as

$$\int_{-\infty}^{t} r(t - t')\, dt',$$

which is simply the total concentration at the present time t due to a constant injection from time $-\infty$ to the present.

Therefore the problem is reduced to finding the total carrier concentration in the filament due to a distribution of sources of constant strength $\sqrt{<J_0^2>}$.

Let $w(x, y, z; x_1, y_1, z_1)$ denote the carrier concentration at x, y, z due to a steady unit source at x_1, y_1, z_1. Then the total carrier concentration is

$$\tau(x_1, y_1, z_1) \; = \; \int w(x, y, z; x_1, y_1, z_1)\, dx\, dy\, dz.$$

The reason for the dependence on x_1, y_1, z_1, is that the recombination process takes place largely on the surface. Therefore a source near the surface will yield a smaller concentration than one well inside the filament. (Volume recombination will be neglected throughout this paper.)

The mean square conductivity modulation due to many statistically independent sources at x_r, y_r, z_r ($r = 1, 2 \cdots$) is then

$$<(\Delta g)^2> \; = \; \Sigma \tau^2(x_r, y_r, z_r) \; <J^2(x_r, y_r, z_r)>.$$

The behavior of w is governed by the diffusion equation, subject to the boundary conditions expressing the recombination process, and subject to a suitable singularity at x_r, y_r, z_r, expressing the injection of a unit current. But in two and three dimensions the solution is not available in closed form, or at any rate not in terms of the elementary transcendental functions. The infinite series for the solution is not easy to

handle computationally. It is therefore desirable to simplify the experimental conditions to a point where the problem becomes almost one-dimensional. A solution for w can then be found in closed form.

Consider a very long uniform rectangular filament with one pair of sides very much wider than the other pair. Suppose that the y and z directions are respectively parallel to the wide sides and to the length of the filament (Fig. 1).

Consider sources located anywhere on a plane $x = \xi$, which is parallel to the wide sides of the filament. If the recombination properties of the filament are uniform in the $y - z$ directions, the lifetime due to a unit source anywhere in that plane is independent of the location of the source on that plane, and depends only on ξ. Hence the conductivity modulation due to sources of strength

$$J(\xi, y_r, z_r) \qquad (r = 1, 2 \cdots)$$

in that plane is simply

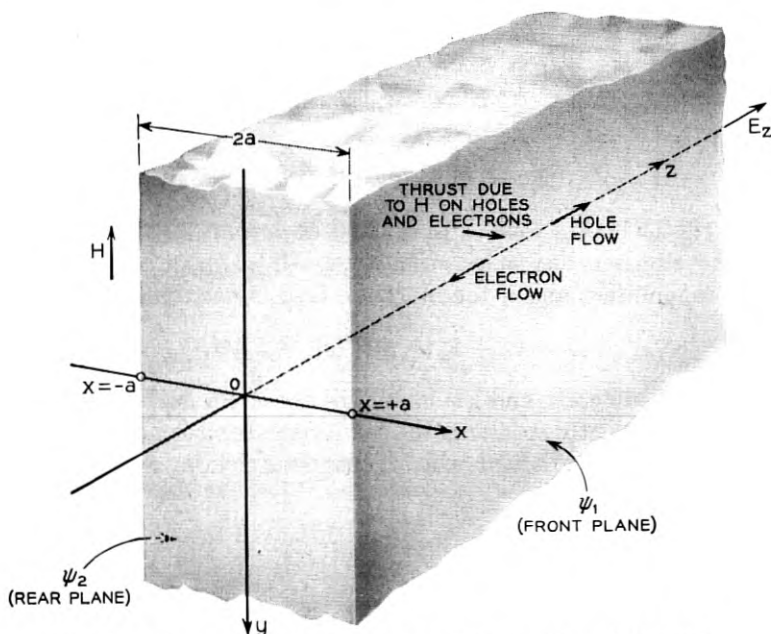$$\tau(\xi) \sum_r J(\xi, y_r, z_r)$$

Fig. 1 — Geometry of the filament, and disposition of the fields.

and is the same as that due to an infinite source of strength

$$\sum_r J(\xi, y_r, z_r)$$

uniformly distributed over the plane x-ξ.

But the density w due to such a source will be a function of x and ξ only, and will be the solution of the one-dimensional diffusion equation. Hence for a geometry approaching that of figure 1 sufficiently closely, the problem is one-dimensional.

The Evaluation of τ

We now have to write down the one-dimensional diffusion equation in the presence of a magnetic field along 0_y, which combines with the drift velocity of the carriers so as to force them towards one of the surfaces $x = \pm a$. If F_x is the effective field arising in this manner, and D is the diffusion coefficient, the equation is

$$D \frac{d^2 w}{dx^2} - \mu F_x \frac{dw}{dx} = 0, \tag{3}$$

which expresses the fact that the diffusion current $-Dq \, dw/dx$ plus the drift current $q\mu F_x w$ must be constant since the carrier density cannot build up indefinitely. μ is the mobility of the minority carriers: μ_n for electrons, μ_p for holes. As is shown elsewhere[2] the effective field F_x is given by

$$F_x = (\theta_n + \theta_p) E_z = \theta E_z,$$

where E_z is the biasing field causing the drift current, θ_n, θ_p are the Hall angles for electrons and holes, respectively. If μ_n, μ_p are the electronic and hole-mobilities, and if the magnetic field is not too large,

$$\theta = \theta_n + \theta_p = 10^{-8}(\mu_n + \mu_p)H,$$

where H is in oersteds, and the mobilities are in cm^2/volt-second, and θ is in radians. (Strictly speaking, the diffusion current is not in the direction of the density gradient when a magnetic field is present.[4] As the result mixed derivatives

$$\frac{\partial^2 w}{\partial x \, \partial y}$$

occur in the diffusion equation. But in the reduction to one dimension these terms integrate out. All that remains is a small correction to D,

negligible for ordinary values of H.) It is convenient to specify a dimensionless parameter in the same notation as H. C. Montgomery.

$$\Phi = \frac{2a\mu F_x}{D} = \frac{2a\theta E_z}{D/\mu}$$

By the Einstein Relation $D/\mu = kT/q$ this may be written

$$= \frac{2a\theta E_z}{kT/q}$$

where q is the absolute value of the electronic charge. Φ is the ratio of the voltage corresponding to the transverse field to the thermal voltage kT/q. In terms of Φ, equation (3) can be rewritten

$$\frac{d^2w}{dx^2} - \frac{\Phi}{2a}\frac{dw}{dx} = 0 \tag{4}$$

The integral of this equation has the form

$$w = Ae^{\Phi(x/2a)} + B \tag{5}$$

where A and B are two constants. Because of the existence of a singularity at $x = x_0$, say, the constants A, B take on different values for $x < x_0$ and $x > x_0$. To see what these values are, we first write the solution (5) in the form

$$w_1 = A_1 e^{\Phi(x-x_0)/2a} + B_1 \qquad x > x_0,$$

$$w_2 = A_2 e^{\Phi(x-x_0)/2a} + B_2 \qquad x < x_0.$$

At $x = x_0$ the density w must be continuous. Hence

$$A_1 + B_1 = A_2 + B_2 \tag{6}$$

Further, the discontinuity at x_0 must be such that the difference of the currents on the two sides is just unity, the strength of the injected current. Now the total current is

$$-D\left(\frac{dw}{dx} - \frac{\Phi}{2a}w\right)$$

(i.e., the diffusion current plus the drift current), and w is continuous at x_0. Hence the difference between the current on the two sides of x_0 is

$$1 = -D \lim_{h \to 0} \left(\left(\frac{dw_1}{dx}\right)_{x_0+h} - \left(\frac{dw_2}{dx}\right)_{x_0-h}\right) \tag{7}$$

$$= \frac{\Phi}{2a}D(A_2 - A_1).$$

So far we have two relations (6) and (7) between the four constants A_1, A_2, B_1, B_2. To determine these constants, we need two more equations. These are supplied by the boundary conditions. We assume that the recombination rate is proportional to the density at the boundaries ($x = \pm a$). The recombination then has the formal appearance of a current through the surface, the factor of proportionality s playing the part of a "recombination velocity."[2, 3] That current through the surface must equal the current arriving at the surface from the interior of the filament, under steady-state conditions. The boundary conditions are thus:

$$-D\left(\frac{dw_1}{dx} - \frac{\Phi}{2a}\,w_1\right)_{x=+a} = s_1 w(+a),$$

$$-D\left(\frac{dw_2}{dx} - \frac{\Phi}{2a}\,w_2\right)_{x=-a} = -s_2 w(-a),$$

where s_1, s_2 are the surface recombination velocities of the faces $x = +a$, $-a$ respectively. The minus sign on the right-hand side of the second of these equations is due to the fact that the recombination current at $x = -a$ is along the $-x$ direction. Defining the recombination parameters

$$\psi_{1,\,2} = \frac{s_{1,\,2}a}{D}$$

and substituting the solutions w_1, w_2 in the last two equations we obtain

$$B_1\frac{\Phi}{2} = \psi_1(A_1 e^{\Phi(1-x_0/a)/2} + B_1)$$

$$B_2\frac{\Phi}{2} = \psi_2(A_2 e^{-\Phi(1+x_0/a)/2} + B_2)$$

$$(8)$$

Equations (6), (7) and (8) suffice to determine A_1, A_2, B_1, B_2. We easily find that

$$A_1 = -e^{-\Phi(1-x_0/a)/2}\,\frac{2a}{\Phi D}\,\frac{\alpha_1(1 - \alpha_2 e^{\Phi(1+x_0/a)/2})}{\Delta}$$

$$A_2 = +e^{+\Phi(1+x_0/a)/2}\,\frac{2a}{\Phi D}\,\frac{\alpha_2(1 + \alpha_1 e^{-\Phi(1-x_0/a)/2})}{\Delta}$$

$$B_1 = -\frac{2a}{\Phi D}\,\frac{(1 - \alpha_2 e^{\Phi(1+x_0/a)/2})}{\Delta}$$

$$B_2 = -\frac{2a}{\Phi D}\,\frac{(1 + \alpha_1 e^{-\Phi(1-x_0/a)/2})}{\Delta}$$

$$(9)$$
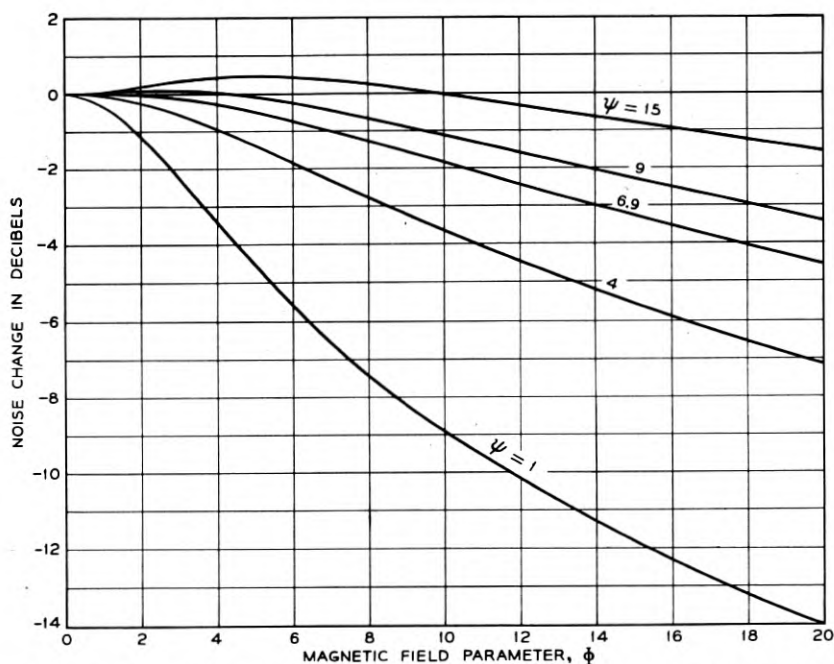
Fig. 2 — Noise change for a filament with equal surfaces as a function of Φ for various ψ. Surface generation.

where

$$\alpha_1 = \left(\frac{\Phi}{2} - \psi_1\right)/\psi_1 ; \qquad \alpha_2 = \left(\frac{\Phi}{2} + \psi_2\right)/\psi_2 ;$$

$$\Delta = \alpha_1 e^{-\Phi(1-x_0/a)/2} + \alpha_2 e^{\Phi(1+x_0/a)/2}.$$

Now we are in a position to calculate $\tau_H(x_0)$, the life-time in a magnetic field H (contained in Φ), due to an emission at x_0:

$$\tau_H(x_0) = \int_{-a}^{x_0} w_2(x, x_0) \, dx + \int_{x_0}^{+a} w_1(x, x_0) \, dx$$

$$= (B_1 + B_2)a + (B_2 - B_1)x_0 + \frac{2a}{\Phi}(A_2 - A_1) \qquad 10$$

$$\frac{2a}{\Phi}[A_1 e^{\Phi(1-x_0/a)/2} - A_2 e^{-\Phi(1+x_0/a)/2}].$$

Special Cases

With the help of (9) and (10) we can now examine special cases.

1. Surface Emission

When only the surface $x = +a$ is emitting, we set $x_0 = +a$ in (10) and get

$$\tau_H(+a) = 2B_2\eta + \frac{2a}{\Phi} A_2(1 - e^{-\Phi})$$

$$= \frac{4a^2}{\Phi D} (1 + \alpha_1) \frac{\left[1 - \alpha_2 \dfrac{e^{\Phi}}{\Phi} (1 - e^{-\Phi}) \right]}{\alpha_1 + \alpha_2 e^{\Phi}}$$

$$= -\frac{2a^2}{D\psi_1} \frac{\left(1 - \alpha_2 \dfrac{(e^{\Phi} - 1)}{\Phi} \right)}{\alpha_1 + \alpha_2 e^{\Phi}}.$$

Similarly, when only $-a$ is emitting, we get

$$\tau_H(-a) = -\frac{2\eta^2}{D\psi_2} \frac{\alpha_1 \dfrac{e^{-\Phi} - 1}{\Phi} - 1}{\alpha_1 e^{-\Phi} + \alpha_2}.$$

But in an actual experiment, both faces will be emitting, with mean square strengths $\langle J_1^2 \rangle$, $\langle J_2^2 \rangle$ say. The quantity that is then measured

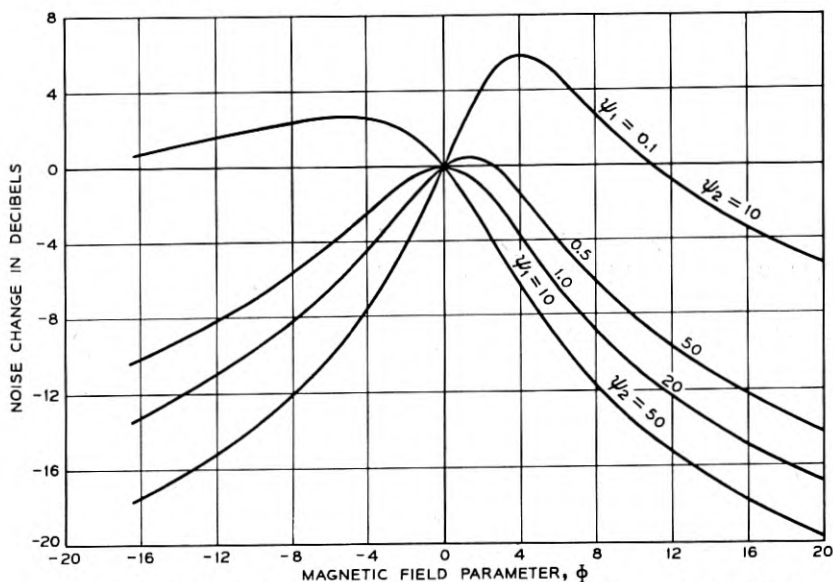Fig. 3 — Contribution to the noise change from a unit source at the plane $x = +a$.

is the ratio

$$N_H = \frac{\Delta g_H^2}{\Delta g_{H=0}^2} = \frac{\langle J_1^2 \rangle \tau_H^2(+a) + \langle J_2^2 \rangle \tau_H^2(-a)}{\langle J_1^2 \rangle \tau_{H=0}^2 (+a) + \langle J_2^2 \rangle \tau_{H=0}^2(-a)} \, . \tag{11}$$

Therefore we also need the lifetimes at zero H (that is, zero Φ). But at $\Phi = 0$, the τ's are indeterminate, and we therefore have to take limits. Expansion in powers of Φ shows that

$$\underset{\Phi \to 0}{lt} \; \tau_H(+a) = \frac{a^2}{D\psi_1} \frac{\left(1 + \dfrac{1}{\psi_2}\right)}{1 + \dfrac{1}{2}\left(\dfrac{1}{\psi_1} + \dfrac{1}{\psi_2}\right)} \, ,$$

$$\underset{\Phi \to 0}{lt} \; \tau_H(-a) = \frac{a^2}{D\psi_2} \frac{\left(1 + \dfrac{1}{\psi_1}\right)}{1 + \dfrac{1}{2}\left(\dfrac{1}{\psi_1} + \dfrac{1}{\psi_2}\right)} \, .$$

Thus we finally get

$$N_H = \frac{\Delta g_H^2}{\Delta g_0^2} = 4 \left(1 + \frac{1}{2}\left(\frac{1}{\psi_1} + \frac{1}{\psi_2}\right)\right)^2$$

$$\frac{\dfrac{\langle J_1^2 \rangle}{\psi_1^2}\left[\dfrac{1 - \alpha_2\left(\dfrac{e^\Phi - 1}{\Phi}\right)}{\alpha_1 + \alpha_2 e^\Phi}\right]^2 + \dfrac{\langle J_2^2 \rangle}{\psi_1^2}\left[\dfrac{\alpha_1\left(\dfrac{1 - e^{-\Phi}}{\Phi}\right) + 1}{\alpha_1 e^{-\Phi} + \alpha_2}\right]^2}{\dfrac{\langle J_1^2 \rangle}{\psi_1^2}\left(1 + \dfrac{1}{\psi_2}\right)^2 + \dfrac{\langle J_2^2 \rangle}{\psi_2^2}\left(1 + \dfrac{1}{\psi_1}\right)^2} \tag{12}$$

There remains one small difficulty in the way of comparing experiment with theory: We do not know $\langle J_1^2 \rangle$, $\langle J_2^2 \rangle$. As suggested by H. C. Montgomery, we are able to overcome this difficulty as follows: We first draw a number of curves of

$$\frac{\tau_H^2(+a)}{\tau_0^2(+a)} \, , \qquad \frac{\tau_H^2(-a)}{\tau_0^2(-a)} \; \text{versus } \Phi,$$

for various sets of parameters $(\psi_1 , \psi_2.)$ (See figures 3, 4). Then we contrive to match a linear superposition

$$c_1 \frac{\tau_H^2(+a)}{\tau_0^2(+a)} + c_2 \frac{\tau_H^2(-a)}{\tau_0^2(-a)}$$

where $c_1 + c_2 = 1$ to the experimental curve. This will be possible only for one particular set of values (ψ_1 , ψ_2). From c_1 and $c_2 = 1 - c_1$ and
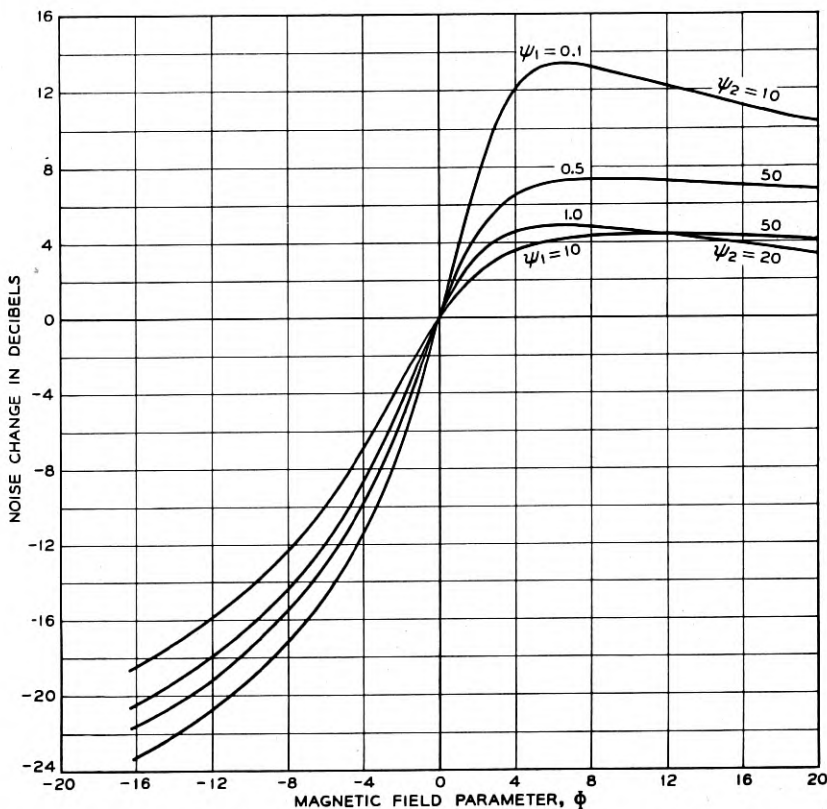
Fig. 4 — Contribution to the noise change from a unit source at $x = -a$.

from (ψ_1, ψ_2) we can determine the ratio of $\langle J_1^2 \rangle \langle J_2^2 \rangle$, for, having matched the experimental curve, we know that

$$c_1 \frac{\tau_H^2(+a)}{\tau_0^2(+a)} + (1 - c_1) \frac{\tau_H^2(-a)}{\tau_0^2(-a)} = \frac{\langle J_1^2 \rangle \tau_H^2(+a) + \langle J_2^2 \rangle \tau_H^2(-a)}{\langle J_1^2 \rangle \tau_0^2(+a) + \langle J_2^2 \rangle \tau_0^2(-a)},$$

which is satisfied for all Φ if

$$\frac{c_1}{\tau_0^2(+a)} = \frac{\langle J_1^2 \rangle}{\langle J_1^2 \rangle \tau_0^2(+a) + \langle J_2^2 \rangle \tau_0^2(-a)},$$

$$\frac{(1 - c_1)}{\tau_0^2(-a)} = \frac{\langle J_2^2 \rangle}{\langle J_1^2 \rangle \tau_0^2(+a) + \langle J_2^2 \rangle \tau_0^2(-a)}.$$

These equations are self-consistent and give

$$\frac{c_1}{1-c_1} = \frac{\tau_0^2(+a)\langle J_1^2\rangle}{\tau_0^2(-a)\langle J_2^2\rangle} = \frac{\langle J_1^2\rangle}{\langle J_2^2\rangle} \frac{\left(1+\dfrac{1}{\psi_2}\right)^2}{\left(1+\dfrac{1}{\psi_1}\right)^2} \frac{\psi_2^2}{\psi_1^2}.$$

Thus the match of theory to experiment actually provides information about the relative emissivities of the surfaces.

When the surfaces are equal, $\psi_1 = \psi_2 = \psi$; $\langle J_1^2\rangle = \langle J_2^2\rangle$ and the result 12 simplifies to

$$N_H = \frac{\left(1+\dfrac{1}{\psi}\right)^2 + \left(\dfrac{2}{\Phi} - \coth\dfrac{\Phi}{2}\right)^2}{\left(1+\dfrac{\Phi}{2\psi}\coth\dfrac{\Phi}{2}\right)^2}. \tag{13}$$

$10\log N_H$ for this case is plotted in figure 2.

It is interesting to see how N_H for equal surfaces varies for small values of Φ. $2/\Phi - \coth\Phi/2$ is regular at $\Phi = 0$ and varies as $-\Phi/6$ there. Hence the initial variation is as

$$N_{H\,\text{small}} = 1 + \frac{1}{4\left(1+\dfrac{1}{\psi}\right)^2}\left(\frac{\Phi^2}{9} - \frac{2}{3}\frac{\Phi^2}{\psi}\left(1+\frac{1}{\psi}\right)\right)$$

$$= 1 + \frac{\Phi^2}{12\left(1+\dfrac{1}{\psi}\right)^2}\left(\frac{1}{3} - \frac{2}{\psi}\left(1+\frac{1}{\psi}\right)\right).$$

Hence the curve rises or falls initially according as

$$\frac{1}{3} \gtrless 2\frac{1}{\psi}\left(1+\frac{1}{\psi}\right).$$

The noise therefore increases initially if

$$\sqrt{\frac{5}{12}} - \frac{1}{2} > \frac{1}{\psi} \qquad \text{or} \qquad \psi > 6.9 \text{ approximately}$$

and falls initially if

$$\psi < 6.9.$$

2. Volume Generation.

At first sight it may seem that if volume generation is considered, so should volume recombination (detailed balancing). This is not necessarily so, since we are not dealing with an equilibrium situation here.

Therefore the possibility of volume generation and surface recombination cannot be discarded.

This case is somewhat more difficult. Assuming all sources to be uncorrelated and uniformly distributed throughout the interval $(-a, +a)$, we have to square expression (10) and integrate it from $x_0 = -a$ to $x_0 = +a$ in order to find $\langle \Delta g_H^2 \rangle$. (We suppose that all the sources are of equal strength $\langle J^2 \rangle$). Substituting the values of the A and B from 9 and 10, we get, after some obvious cancellations

$$\tau_H(x_0) = \frac{2a}{\Phi D} [x_0 + Se^{-\Phi(x_0/2a)} + T]$$

where

$$S = 2a \frac{\left(1 + \frac{1}{2}\left(\frac{1}{\psi_1} + \frac{1}{\psi_2}\right)\right)}{\alpha_1 e^{-\Phi/2} + \alpha_2 e^{\Phi/2}},$$

$$T = \frac{2a}{\Phi} \left[\frac{\Phi}{2} \frac{\alpha_1 e^{-\Phi/2} - \alpha_2 e^{+\Phi/2}}{\alpha_1 e^{-\Phi/2} + \alpha_2 e^{+\Phi/2}} - 2 \frac{\alpha_1 \alpha_2 \sinh \frac{\Phi}{2}}{\alpha_1 e^{-\Phi/2} + \alpha_2 e^{\Phi/2}} - 1 \right].$$

Hence

$$\langle \Delta g_H^2 \rangle = \langle J \rangle \int_{-a}^{+a} \tau_H^2(x_0) \, dx_0$$

$$= \frac{4a^2}{\Phi^2 D^2} \left[\frac{2}{3} a^3 + 2T^2 a + \frac{8TSa}{\Phi} \sinh \frac{\Phi}{2} \right. \tag{14}$$

$$\left. + \frac{2aS^2}{\Phi} \sinh \Phi + \frac{8Sa^2}{\Phi} \left(\frac{\sinh \frac{\Phi}{2}}{\frac{\Phi}{2}} - \cosh \frac{\Phi}{2} \right) \right].$$

Before proceeding with this general case, we first consider the limiting case $\psi_1 = \psi_2 = \infty$, when $\alpha_1 = -1$, $\alpha_2 = +1$. Then

$$T = -a \coth \frac{\Phi}{2} \qquad S = \frac{a}{\sinh \frac{\Phi}{2}}$$

and

$$\int_{-a}^{+a} \tau_H^2(x_0) = \frac{8a^5}{\Phi^2 D^2} \left[\frac{1}{3} + \coth^2 \frac{\Phi}{2} - 6 \frac{\coth \frac{\Phi}{2}}{\Phi} + \frac{8}{\Phi^2} \right].$$

To find $N_H = \langle \Delta g_H^2 \rangle / \langle \Delta g_{H=0}^2 \rangle$ we need the limit of $\int_{-a}^{+a} \tau_H^2$ as $\Phi \to 0$. After some tedious algebra, we find this limit to be

$$\tau_{H=0}^2 = \frac{4a^5}{15D^2}$$

so that

$$(N_H)_{\psi_1, \psi_2 = \infty} = \frac{\int \tau_H^2}{\int \tau_0^2} = \frac{30}{\Phi^2} \left[\frac{1}{3} + \coth^2 \frac{\Phi}{2} - \frac{6 \coth \frac{\Phi}{2}}{\Phi} + \frac{8}{\Phi^2} \right]. \tag{15}$$

In the general case we can again take the limit of (14) as $\Phi \to 0$ in order to determine N_H, but this would be too tedious. Instead, we solve the diffusion equation directly when $\Phi = 0$. The equation is then simply

$$\frac{d^2 w}{dx^2} = 0$$

and the solution subject to the correct boundary conditions and allowing for a steady unit injection at x_0 is

$$w_1 = A_1(x - x_0) + B \qquad\qquad x > x_0$$

$$w_2 = A_2(x - x_0) + B \qquad\qquad x < x_0.$$

where

$$A_1 = -\frac{1}{D} \psi_1 \left[1 + \psi_2 \left(1 + \frac{x_0}{a} \right) \right] \bigg/ (\psi_1 + \psi_2 + 2\psi_1\psi_2),$$

$$A_2 = \frac{1}{D} \psi_2 \left[1 + \psi_1 \left(1 - \frac{x_0}{a} \right) \right] \bigg/ (\psi_1 + \psi_2 + 2\psi_1\psi_2),$$

$$B = \frac{a}{D} \left[1 + \psi_1 \left(1 - \frac{x_0}{a} \right) \right] \left[1 + \psi_2 \left(1 + \frac{x_0}{a} \right) \right] \bigg/ (\psi_1 + \psi_2 + 2\psi_1\psi_2).$$

We now have

$$\tau_{H=0}(x_0) = \int_{-a}^{+a} w \, dx = \int_{-a}^{x_0} w_2 \, dx + \int_{x_0}^{a} w_1 \, dx$$

$$= -\frac{1}{2} \left[2ax_0(A_1 + A_2) + \frac{a^2 + x_0^2}{D} \right] + 2aB.$$

From this we can compute

$$\int_{-a}^{+a} \tau_{H=0}^2(x_0) \, dx_0 \, ,$$
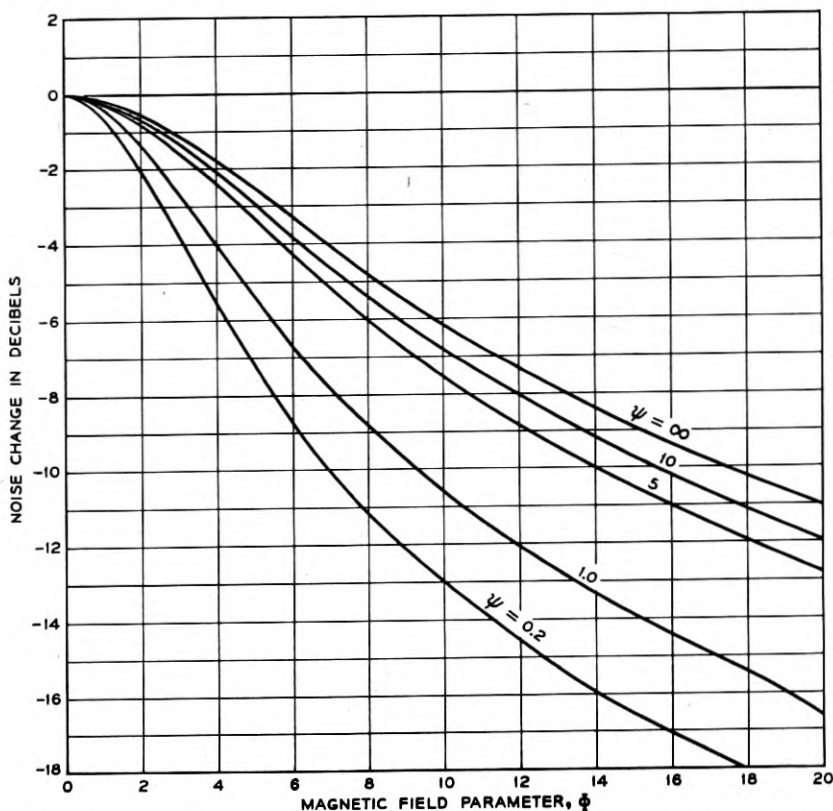
Fig. 5 — Noise change in a filament with uniform volume generation and equal surface recombination rates.

but the result is still rather complicated unless $\psi_1 = \psi_2 = \psi$. If we therefore restrict ourselves to that case we find

$$A_1 + A_2 = -\frac{\psi}{D(1 + \psi)} \cdot \frac{x_0}{a},$$

$$B = \frac{1}{2D}\left[\frac{1 + \psi}{\psi} - \frac{\psi}{1 + \psi}\frac{x_0^2}{a^2}\right]$$

and

$$D\tau_{H=0}(x_0) = a^2\left(\frac{1}{2} + \frac{1}{\psi}\right) - \frac{x_0^2}{2}$$

and

$$\int_{-a}^{+a} \tau_{H=0}^2(x_0)\, dx_0 = \frac{2a^5}{D^2}\left(\frac{1}{45} + \left(\frac{1}{\psi} + \frac{1}{3}\right)^2\right).$$

(Note that as $\psi \to \infty$, this quantity tends to $4a^5/15D^2$, as before in our limiting process.)

When $\psi_1 = \psi_2 = \psi$, S and T also simplify somewhat, and the result is

$$N_H = \frac{\langle \Delta g_H^2 \rangle}{\langle \Delta g_{H=0}^2 \rangle}$$

$$= 2 \frac{\frac{2}{3} + 2T^2 + \frac{8S}{\Phi}\left(\frac{2}{\Phi} - T\right)\sinh\frac{\Phi}{2} - \frac{8S}{\Phi}\cosh\frac{\Phi}{2} + \frac{2S^2}{\Phi}\sinh\Phi}{\Phi^2\left[\left(\frac{1}{\psi} + \frac{1}{3}\right)^2 + \frac{1}{45}\right]}$$

where now

$$S = \frac{1 + \dfrac{1}{\psi}}{\left(\dfrac{\Phi}{2\psi}\cosh\dfrac{\Phi}{2} + \sinh\dfrac{\Phi}{2}\right)},$$

$$T = \left(1 + \frac{1}{\psi}\right)\frac{\dfrac{\Phi}{2\psi} + \coth\dfrac{\Phi}{2}}{1 + \dfrac{\Phi}{2\psi}\coth\dfrac{\Phi}{2}}.$$

An alternative form is

$$N_H = \frac{4}{\frac{1}{45} + \left(\frac{1}{3} + \frac{1}{\psi}\right)^2}\left[\frac{1}{3\Phi^2} + \frac{4}{\Phi^3\Delta}\left(1 + \frac{1}{\psi}\right)\left(\frac{2}{\Phi} - \coth\frac{\Phi}{2}\right)\right.$$

$$\left. + \frac{\left(1 + \dfrac{1}{\psi}\right)^2}{\Phi^2\Delta^2}\left(\left(\frac{\Phi}{2\psi} + \coth\frac{\Phi}{2}\right)^2 + \frac{2}{\psi} - \frac{2\coth\dfrac{\Phi}{2}}{\Phi}\right)\right] \tag{16}$$

where

$$\Delta = 1 + \frac{\Phi}{2\psi}\coth\frac{\Phi}{2},$$

a result which correctly tends to (15) as $\psi \to \infty$. $10 \log N_H$ for various ψ is shown in Fig. 5.

ACKNOWLEDGMENTS:

independently calculated some special cases. Thanks are also due to H. C. Montgomery for many valuable discussions and to Mrs. C. A. Lambert for computing the curves.

REFERENCES

1. Montgomery, H. C., Electrical Noise in semiconductors. Bell System Tech. J., **31**, pp. 950–975, Sept., 1952.
2. Shockley, W., Electrons and Holes in Semiconductors, Chapter 3, Section 2, Van Nostrand.
3. Suhl, H., and W. Shockley, Phys. Rev., **75**, pp. 1617–1618, 1949.
4. Shockley, W., Electrons and Holes in Semiconductors, p. 299, Van Nostrand.

DC Field Distribution in a "Swept Intrinsic" Semiconductor Configuration

By R. C. PRIM

(Manuscript received January 15, 1953)

This paper contains an analysis of the dc field intensity distribution in an idealized one-dimensional n-intrinsic-p semi-conductor configuration biased in reverse. It gives some quantitative insight into the progressive penetration of the field into the intrinsic region as the magnitude of the bias voltage is increased.

INTRODUCTION

Possible applications have been suggested for semi-conductor configurations involving intrinsic regions adjacent simultaneously to n- and p-type extrinsic regions. The basic idea behind some of these proposals is that a suitably large reverse bias voltage (n-regions positive with respect to p- regions) will set up a substantial electric field in the interior of the intrinsic region. This field would sweep most of the mobile carriers out of the intrinsic material, producing a region of material ("swept intrinsic") supporting a large field and having a high resistivity.

This paper contains a dc analysis of an idealized one-dimensional swept intrinsic structure with abrupt transitions from strongly n-type to highly intrinsic to strongly p-type material. It gives some quantitative insight into the penetration of the electric field into the intrinsic region as the bias voltage is progressively increased.

FORMULATION OF PROBLEM

A one-dimensional structure will be considered having the distribution of excess of donor conceptration over acceptor concentration (N_d-N_a) shown in Fig. 1. It will be supposed that N/n_i and P/n_i are $\gg$ 1 and that a reverse bias voltage (Fig. 2) is applied between the bodies of the n- and p-type regions. (n_i denotes the thermal equilibrium con-
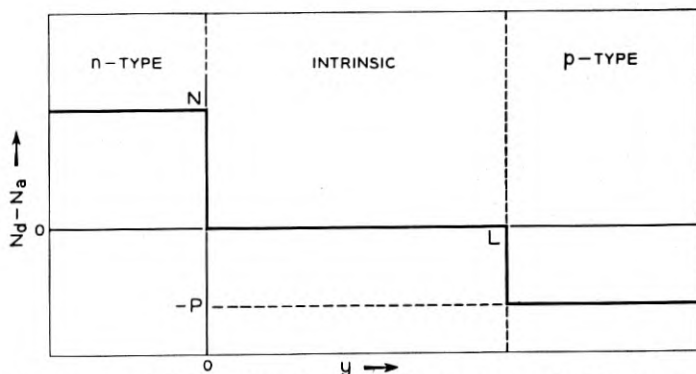
Fig. 1 — Assumed distribution of excess of donor concentration over acceptor concentration.

centration of mobile electrons — or of holes — in intrinsic material at the given temperature.)

The set of equations which (together with boundary conditions to be specified later) determines the electric field in the intrinsic region $0 < y < L$ is:

$$\frac{dE}{dy} = \frac{q}{\kappa} (p - n), \tag{1}$$

$$\frac{dp}{dy} = \frac{q}{kT} pE - \frac{i_p}{\mu_p kT}, \tag{2}$$

$$\frac{dn}{dy} = -\frac{q}{kT} nE + \frac{i_n}{\mu_n kT}, \tag{3}$$

$$\frac{d\psi}{dy} = -E, \tag{4}$$

$$\frac{di_p}{dy} = q(g - r), \tag{5}$$

$$\frac{di_n}{dy} = -q(g - r). \tag{6}$$

where E: electric field intensity, volts/m.

q : electronic charge magnitude, coulombs.

κ : absolute dielectric constant, farads/m.

p : hole concentration, m^{-3}.

n : electron concentration, m^{-3}.

k : Boltzmann's constant, joules/°K.

T : temperature, °K.

ψ : electric potential, volts.

i_p : hole current density, amps/m^2.

i_n : electron current density, amps/m^2.

μ_p: hole mobility constant, m^2/volt-sec.

μ_n: electron mobility constant, m^2/volt-sec.

g : rate of generation of hole-electron pairs, m^{-3} sec^{-1}.

r : rate of recombination of hole-electron pairs, m^{-3} sec^{-1}.

An order-of-magnitude comparison of the terms in (2) or (3) reveals that the currents probably have little influence on the field distribution. For example for

$$i_n \sim 5 \text{ amps/m}^2,$$

$$\mu_n kT = 1.44 \times 10^{-21} \text{ amp-m}^2,$$

$$N = 2 \times 10^{22} \text{ m}^{-3}, \text{ and } L = 10^{-3}\text{m}.$$

$$\int_0^L \frac{i_n}{\mu_n kT}\, dy \sim 3 \cdot 10^{21} L \sim 3 \cdot 10^{18} \text{ m}^{-3}$$

while

$$\left| \int_0^L \frac{dn}{dy}\, dy \right| = n(0) - n(L) \sim N = 2 \cdot 10^{22} \text{ m}^{-3}.$$

On this basis, the current terms in (2) and (3) can be omitted without serious error. No use then has to be made of (5) and (6), so the governing equations for the intrinsic region become:
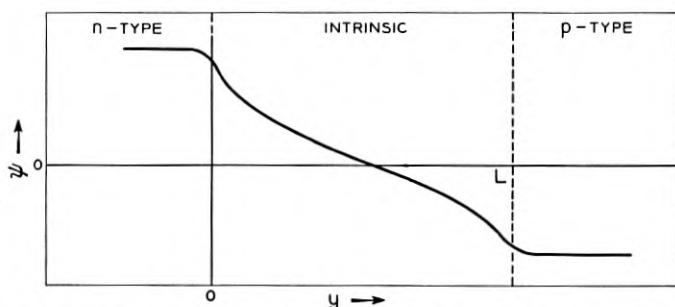
$$\frac{dE}{dy} = \frac{q}{\kappa} (p - n), \tag{1'}$$

Fig. 2 — Qualitative picture of potential distribution in reverse-biased *n-intrinsic-p* structure.

$$\frac{dp}{dy} = \frac{q}{kT}\, pE,$$ (2')

$$\frac{dn}{dy} = -\frac{q}{kT}\, nE,$$ (3')

$$\frac{d\psi}{dy} = -E.$$ (4')

Equations (2')–(4') will also be used within the n-type extrinsic region $(y < 0)$ and the p-type extrinsic region $(y > L)$. However, for the n-type region (1') is replaced by

$$\frac{dE}{dy} = \frac{q}{\kappa}\, (N + p - n),$$ (1'a)

where N denotes the excess concentration of ionized donor over acceptor centers for $y < 0$. Similarly, for the p-type region (1') is replaced by

$$\frac{dE}{dy} = \frac{q}{\kappa}\, (-P + p - n),$$ (1'b)

where P denotes the excess concentration of ionized acceptor over donor centers for $y > L$.

In order to solve the equation set (1')–(4') for the intrinsic region $0 < y < L$, it will be necessary partially to solve the sets (1'a) (2')–(4') and (1'b) (2')–(4') governing the two extrinsic regions because only in this way can a sufficient number of appropriate boundary conditions be imposed.

Deep inside the extrinsic regions the electric field intensity will be negligible and the mobile carrier concentrations will have their equilibrium values. This leads to the conditions

$$E = 0, \qquad np = n_i^2, \qquad n - p = N \text{ at } y = -\infty$$ (7)

and

$$E = 0, \qquad np = n_i^2, \qquad p - n = P \text{ at } y = +\infty.$$ (8)

It will further be supposed that there is no infinite charge concentration at the extrinsic-intrinsic interfaces so that the electric field intensity is continuous at the interfaces. The concentration of the (local) majority carrier will also be assumed continuous at an interface. In short,

$$E \text{ and } n \text{ are continuous at } y = 0$$ (9)

and

$$E \text{ and } p \text{ are continous at } y = L. \tag{10}$$

Finally, we choose the reference level for the electric potential in the intrinsic region so that

$$\psi = 0 \text{ for } n = p \tag{11}$$

and regard the potential at the interface $y = 0$ as a prescribed parameter,

$$\psi = \frac{kT}{q} \cdot U \text{ at } y = 0. \tag{12}$$

The two conditions (11) and (12) apply directly to the solutions for the intrinsic region. The conditions (7)–(10) indirectly imply the two additional restraints necessary to determine a unique solution of (1')–(4') in $0 < y < L$.

NORMALIZED VARIABLES AND EQUATIONS

It is convenient to introduce dimensionless normalized variables before proceeding further with the mathematical analysis. As reference voltage it is natural to adopt the *Boltzmann voltage*

$$\psi_B \equiv \frac{kT}{q}, \tag{13}$$

the voltage equivalent of the mean kinetic energy of an electron at temperature T. (At room temperature the Boltzmann voltage is about 1/40 of a volt.) As reference quantity for carrier concentrations we choose the geometric mean of the majority carrier excess concentrations for the two extrinsic regions. i.e.,

$$\text{reference concentration} = (NP)^{1/2}. \tag{14}$$

The reference voltage and carrier concentration having been so chosen, it is natural to select as reference length the *mean Debye length*

$$\mathcal{L} \equiv \left[\frac{kT/q}{2 \frac{q}{\kappa} (NP)^{1/2}} \right]^{1/2}. \tag{15}$$

This mean Debye length is related by

$$\mathcal{L} = (\mathcal{L}_n \, \mathcal{L}_p)^{1/2} \tag{16}$$

to the n-region and p-region Debye lengths defined respectively by

$$\mathcal{L}_n = \left[\frac{kT/q}{2\frac{q}{\kappa}N}\right]^{1/2} \tag{17}$$

and

$$\mathcal{L}_p = \left[\frac{kT/q}{2\frac{q}{\kappa}P}\right]^{1/2}. \tag{18}$$

We now use the reference quantities defined in (13)–(15) to introduce the *normalized distance*

$$\hat{y} \equiv \frac{y}{\mathcal{L}}, \tag{19}$$

the *normalized thickness* of the intrinsic layer

$$\hat{L} \equiv \frac{L}{\mathcal{L}}, \tag{20}$$

the *normalized concentrations* of positive and negative mobile carriers

$$\hat{p} \equiv \frac{p}{(NP)^{1/2}} \qquad \text{and} \qquad \hat{n} \equiv \frac{n}{(NP)^{1/2}}, \tag{21, 22}$$

the *normalized potential*

$$\hat{\psi} \equiv \frac{\psi}{\psi_B}, \tag{23}$$

and the *normalized electric field intensity*

$$\hat{E} \equiv \frac{E}{\psi_B/\mathcal{L}}. \tag{24}$$

In terms of these normalized variables, the governing equations [(1′)–(4′)] for the intrinsic region become:

$$\frac{d\hat{E}}{d\hat{y}} = \frac{1}{2}(\hat{p} - \hat{n}), \tag{25}$$

$$\frac{d\hat{p}}{d\hat{y}} = \hat{p}\hat{E} \tag{26}$$

$$\frac{d\hat{n}}{d\hat{y}} = -\hat{n}\hat{E}, \tag{27}$$

$$\frac{d\hat{\psi}}{d\hat{y}} = -\hat{E}. \tag{28}$$

For the n-type region, (25) should be replaced by

$$\frac{d\hat{E}}{d\hat{y}} = \tfrac{1}{2}(\Lambda + \hat{p} - \hat{n}) \tag{25a}$$

where

$$\Lambda \equiv \left(\frac{N}{P}\right)^{1/2}. \tag{29}$$

For the p-type region, (25) should be replaced by

$$\frac{d\hat{E}}{d\hat{y}} = \tfrac{1}{2}(-\Lambda^{-1} + \hat{p} - \hat{n}). \tag{25b}$$

FORMAL SOLUTION OF EQUATIONS FOR INTRINSIC REGION

Division of (25)–(27) by (28) yields, after evident rearrangements of factors,

$$\frac{d\hat{E}^2}{d\hat{\psi}} = \hat{n} - \hat{p}, \tag{30}$$

$$\frac{d \ln \hat{p}}{d\hat{\psi}} = -1, \tag{31}$$

$$\frac{d \ln \hat{n}}{d\hat{\psi}} = 1. \tag{32}$$

From (31) and (32) follow

$$\hat{p} = A e^{-\hat{\psi}} \tag{33}$$

and

$$\hat{n} = A_1 e^{\hat{\psi}}, \tag{34}$$

where A and A_1 are constants of integration. The condition (11) that $\psi = 0$ for $p = n$ implies that

$$A_1 = A.$$

Substitution of (33) and (34) into (30) yields

$$\frac{d\hat{E}^2}{d\hat{\psi}} = A(e^{\hat{\psi}} - e^{-\hat{\psi}}) \tag{35}$$

$$= 2A \sinh \hat{\psi}.$$

Integration of (35) now leads to

$$\hat{E} = [2A \, (\cosh \hat{\psi} + B)]^{1/2} \tag{36}$$

where B is another integration constant. Substitution of (36) into (28) in the form

$$\frac{d\hat{\psi}}{\hat{E}} = -d\hat{y}$$

yields after another integration

$$\int_0^{\hat{\psi}} \frac{ds}{(\cosh s + B)^{1/2}} = (2A)^{1/2}(C - \hat{y}), \tag{37}$$

where C is the fourth integration constant.

In order to express in terms of tabulated functions the relationship between $\hat{\psi}$ and $\hat{y}$ defined by (37) we shall consider two cases: $-1 < B \leq 1$ and $B \geq 1$. (It is not necessary to consider $B < -1$ because A is essentially positive [see (33)] so that $B < -1$ would imply an imaginary field strength [see (36)] at the plane in the intrinsic region where $\hat{\psi} = 0$.)

The changes of variable of integration

$$s = 2 \sinh^{-1} \cot \lambda \text{ for } -1 < B \leq 1,$$

$$s = 2 \sinh^{-1} \tan \theta \text{ for } B \geq 1$$

permit the carrying out of the integration indicated on the left side of (37). This gives

$$2^{1/2} \left(K\left[\left(\frac{1-B}{2} \right)^{1/2} \right] - F\left[\left(\frac{1-B}{2} \right)^{1/2}, \sin^{-1} \operatorname{sech} \frac{\hat{\psi}}{2} \right] \right) \tag{38a}$$
$$= (2A)^{1/2}(C - \hat{y})$$

or

$$\hat{\psi} = \ln \frac{1 + cn\left(\left(\frac{1-B}{2} \right)^{1/2}, K\left[\left(\frac{1-B}{2} \right)^{1/2} \right] - A^{1/2}(C - \hat{y}) \right)}{1 - cn\left(\left(\frac{1-B}{2} \right)^{1/2}, K\left[\left(\frac{1-B}{2} \right)^{1/2} \right] - A^{1/2}(C - \hat{y}) \right)} \tag{38b}$$

for $-1 < B \leq 1$

and

$$\frac{2}{(B+1)^{1/2}} F\left[\left(\frac{B-1}{B+1} \right)^{1/2}, \sin^{-1} \tanh \frac{\hat{\psi}}{2} \right] = (2A)^{1/2}(C - \hat{y}) \tag{39a}$$

or

$$\hat{\psi} = \ln \frac{1 + sn\left(\left(\frac{B-1}{B+1}\right)^{1/2}, \left(\frac{B+1}{2}\right)^{1/2} A^{1/2}(C - \hat{y})\right)}{1 - sn\left(\left(\frac{B-1}{B+1}\right)^{1/2}, \left(\frac{B+1}{2}\right)^{1/2} A^{1/2}(C - \hat{y})\right)} \tag{39b}$$

for $B \geq 1$.

Note:

$$F[k, \phi] \equiv \int_0^\phi \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}$$

is the Elliptic Integral of First Kind, usually tabulated for

$$0 \leq \phi \leq \frac{\pi}{2}, \qquad 0 \leq k \leq 1.$$

$$K[k] \equiv F\left[k, \frac{\pi}{2}\right]$$

is the Complete Elliptic Integral of First Kind.

$$sn[k, F] \equiv \sin \phi$$

and

$$cn[k, F] \equiv \cos \phi$$

are Jacobian Elliptic Functions associated with $F[k, \phi]$.

When A, B, and C are prescribed, (38b) or (39b) gives the normalized potential $\hat{\psi}$ as a function of the normalized distance $\hat{y}$. The normalized hole concentration $\hat{p}$, electron concentration $\hat{n}$, and field intensity $\hat{E}$ are then given as functions of $\hat{y}$ by (33), (34), and (36).

In a following section it will be shown that, subject to the reasonable assumptions

$$\Lambda e^{-U/2}, \ \Lambda^{-1}e^{-U/2} \ll 1, \tag{40a, b}$$

$$\hat{L} \gg \Lambda^{-1/2}, \ \Lambda^{1/2}, \tag{40c, d}$$

$$U < \frac{1}{20} \Lambda^{1/2}L, \frac{1}{20} \Lambda^{-1/2}\hat{L}, \tag{40e, f}$$

A, C, and B are given in terms of the normalized potential at $y = 0$, the normalized thickness of the intrinsic region, and the parameter $\Lambda \equiv (N/P)^{1/2}$ by

$$A \approx \Lambda e^{-U-1}, \tag{41}$$

$$C \approx \tfrac{1}{2}\hat{L} + e^{1/2}(\Lambda^{1/2} - \Lambda^{-1/2}), \tag{42}$$

and

$$\Phi(B) \approx \frac{1}{2}\left(\frac{\Lambda}{e}\right)^{1/2} \hat{L} e^{-U/2} \tag{43}$$

where

$$\Phi(B) \equiv \begin{cases} K\left[\left(\dfrac{1-B}{2}\right)^{1/2}\right] & \text{for } -1 < B < 1 \\[2mm] \dfrac{\pi}{2} & \text{for } B = 1 \\[2mm] \left(\dfrac{2}{B+1}\right)^{1/2} K\left[\left(\dfrac{B-1}{B+1}\right)^{1/2}\right] & \text{for } B > 1. \end{cases}$$

Thus $\hat{\psi}$ $\hat{p}$, $\hat{n}$, and $\hat{E}$ can be computed as functions of $\hat{\psi}$ for any choice of U, $\hat{L}$, and Λ consistent with (40).

DESCRIPTION OF SAMPLE FIELD DISTRIBUTION COMPUTATIONS

The results of the foregoing analysis were used to compute $\hat{\psi}(\hat{y})$ and $\hat{E}(\hat{y})$ for $\Lambda = 1$ (symmetric case, $N = P$) and the combinations of U and $\hat{L}$ indicated by the following table:

		$\hat{L}$	
	4,000	40,000	400,000
	7	X	X	X
	20	X	X	X
U	200	X	X	X
	2,000		X	X
	20,000			X

The results of these computations are presented in Figs. 3–5 (pages 675 to 677) as plots of $\hat{\psi}/U$ versus $\hat{y}/\hat{L}$, and in Figs. 6–8 (pages 678 to 680) as plots of $\hat{E}/(2U/\hat{L})$ versus $\hat{y}/\hat{L}$. ($2U/\hat{L}$ is the average reduced field strength for the intrinsic region.) Selected curves from the above figures are replotted on non-logarithmic axes in Figs. 9 and 10 (pages 681 and 682). The curves need only be plotted for $0 \le \hat{y}/\hat{L} \le \tfrac{1}{2}$ because, by symmetry for $\Lambda = 1$, $\hat{E}(\hat{L} - \hat{y}) = \hat{E}(\hat{y})$ and $\hat{\psi}(\hat{L} - \hat{y}) = -\hat{\psi}(\hat{y})$. The ordinates and abscissas of Figures 3–10 are given in terms of the original unnormalized variables by

$$\frac{\hat{\psi}}{U} = \frac{\psi}{\psi(0)}, \qquad \frac{\hat{y}}{\hat{L}} = \frac{y}{L} \qquad \text{and} \qquad \frac{\hat{E}}{2U/\hat{L}} = \frac{E}{2\psi(0)/L}.$$

($2\psi(0)/L$ is the average field intensity for the intrinsic region.)
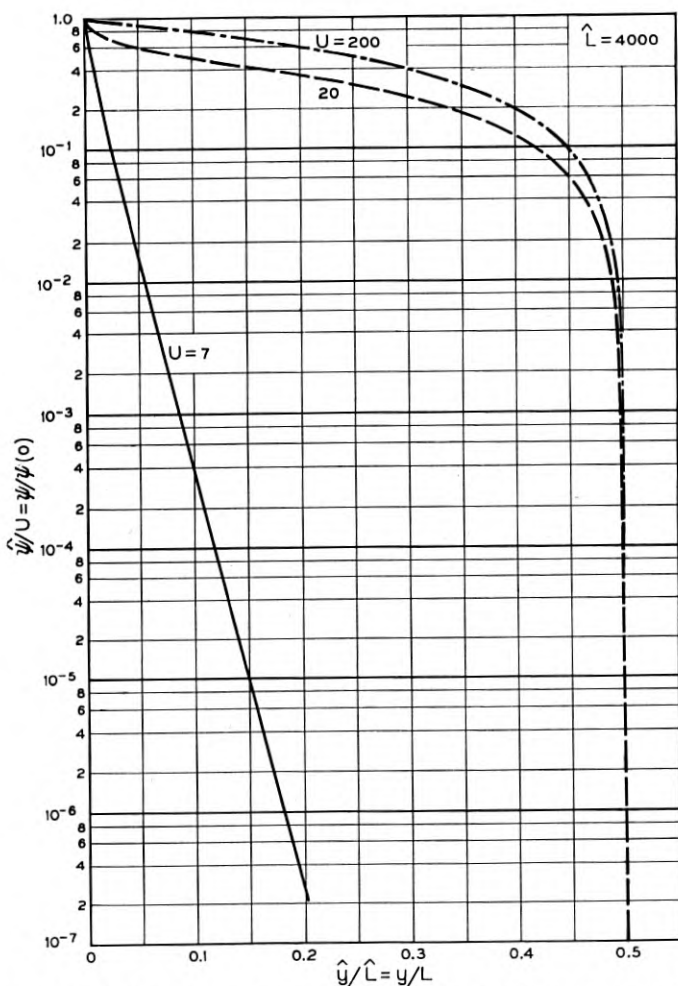
Fig. 3 — (Potential at point in intrinsic layer/Potential of *n-intrinsic* junction) versus (Distance from *n-intrinsic* junction/Intrinsic layer thickness) for $\hat{L}$ = (Intrinsic layer thickness/Mean Debye length) = 4,000 and several values of U = (Potential of *n-intrinsic* junction/Boltzmann voltage).
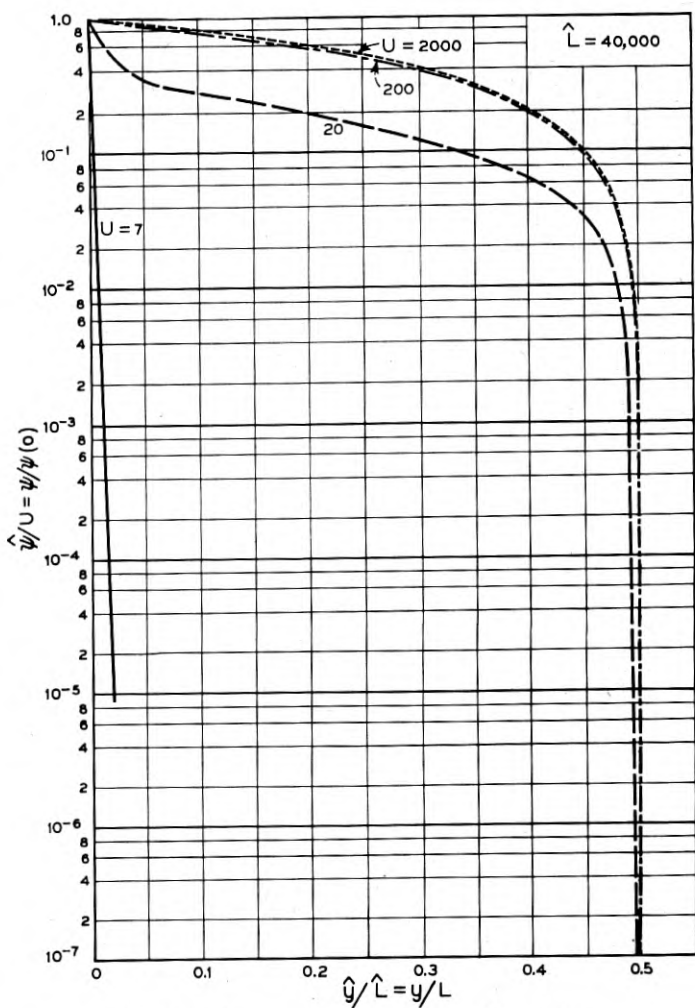
Fig. 4 — Same as Fig. 3 except for $\hat{L} = 40,000$.
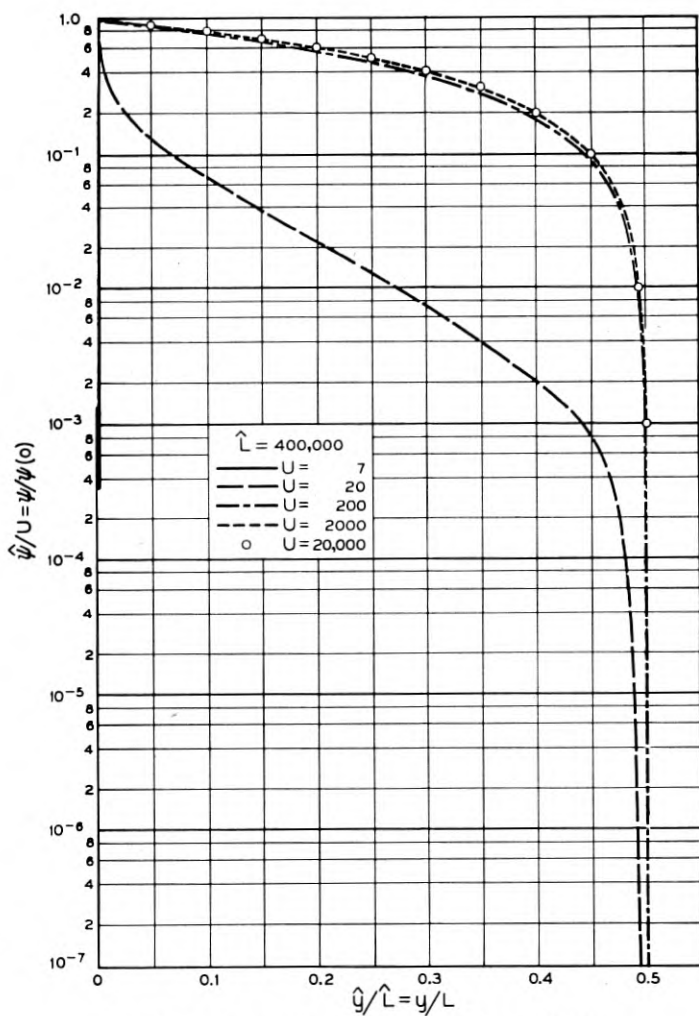
Fig. 5 — Same as Fig. 3 except for $\hat{L} = 400,000$.
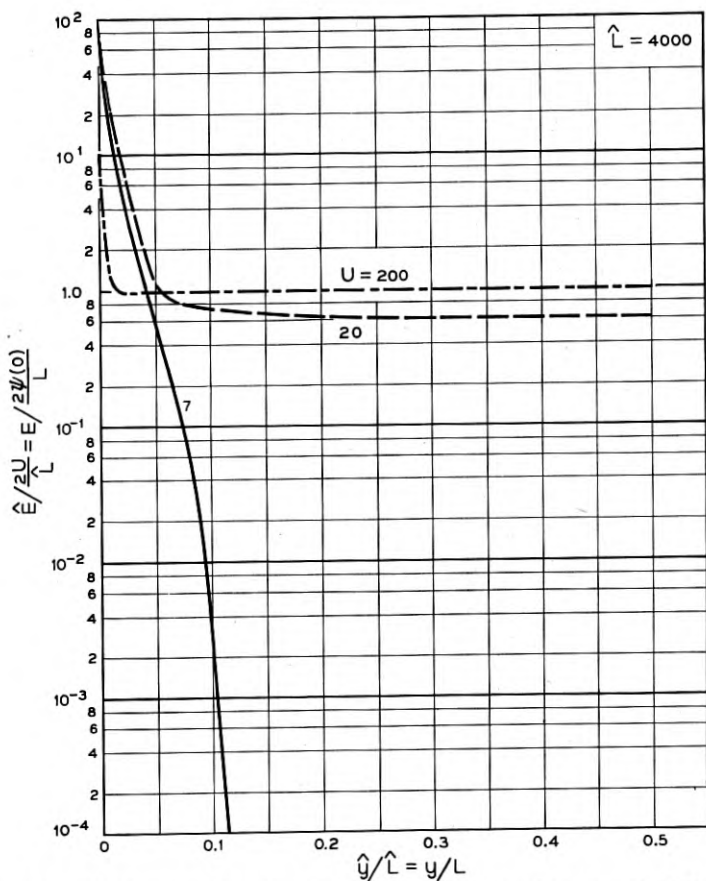
Fig. 6 — (Field intensity at point in intrinsic layer/Average field intensity over intrinsic layer) versus (Distance from *n-intrinsic* junction/Intrinsic layer thickness) for $\hat{L}$ = (Intrinsic layer thickness/Mean Debye length) = 4,000 and several values of U = (Potential of *n-intrinsic* junction/Boltzmann voltage).
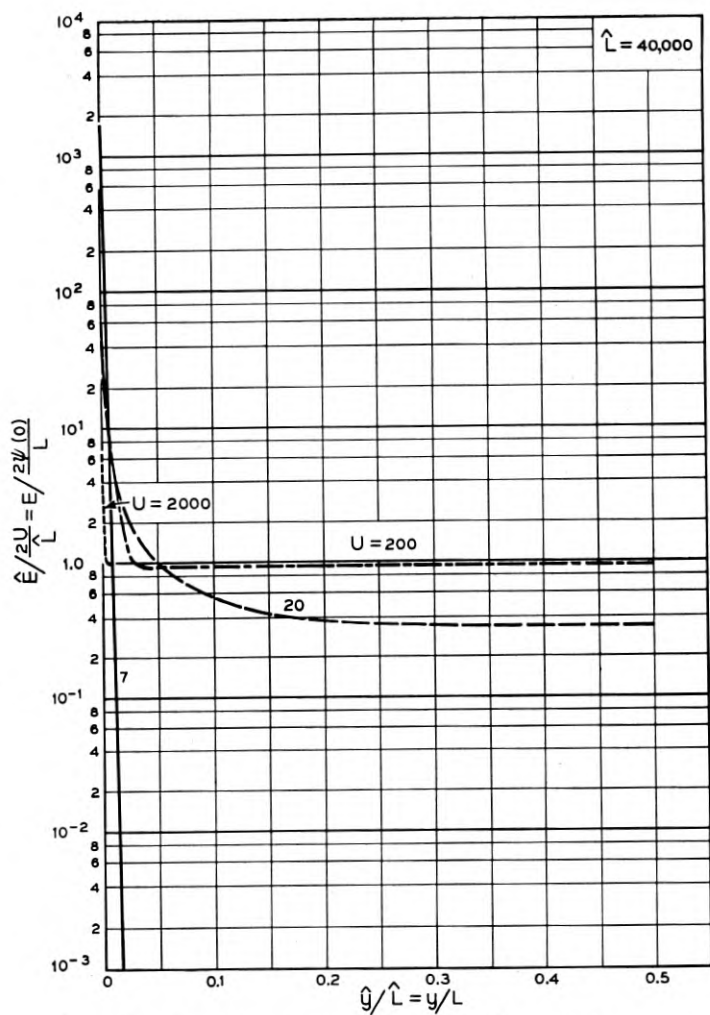
Fig. 7 — Same as Fig. 6 except for $\hat{L} = 40,000$.
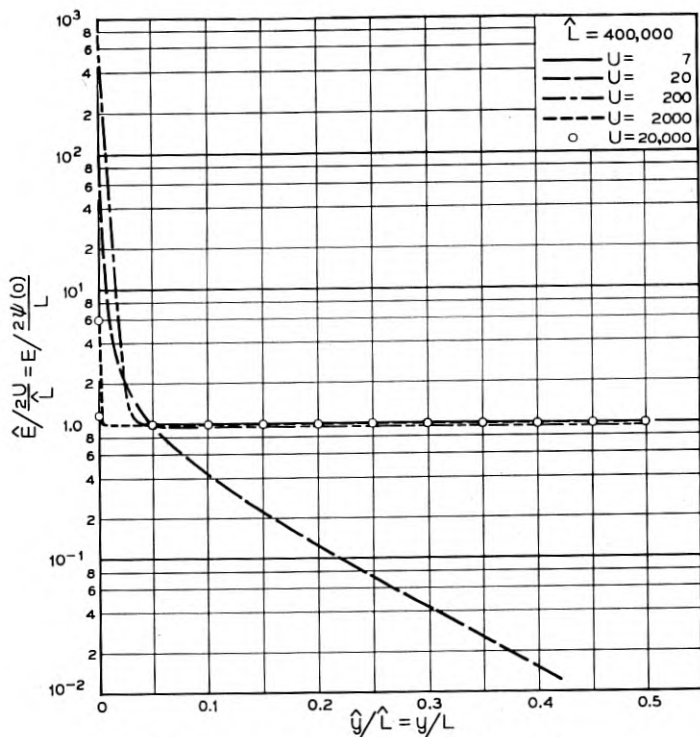
Fig. 8 — Same as Fig. 6 except for $\hat{L} = 400,000$.
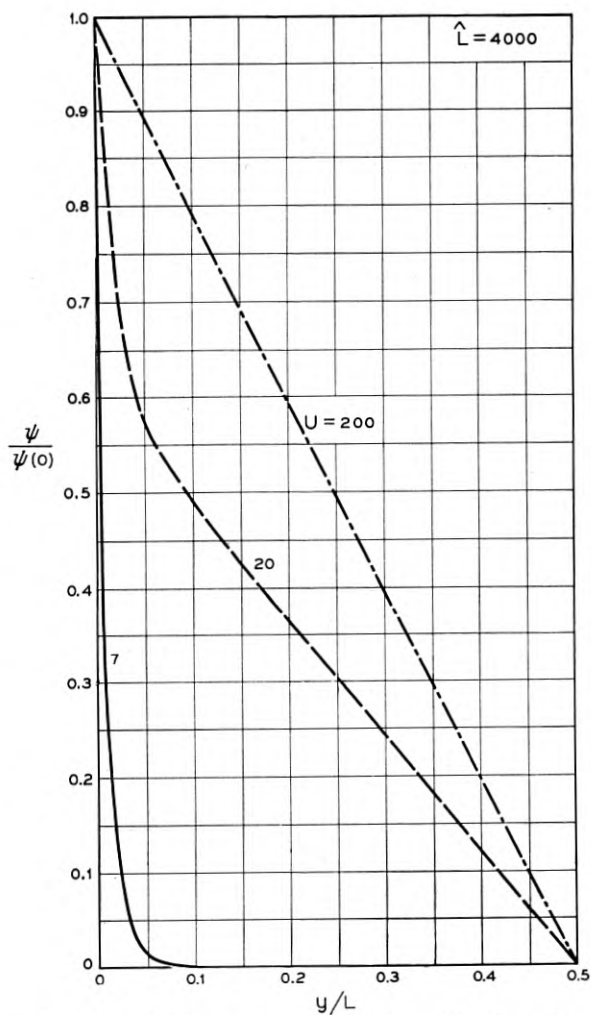
Fig. 9 — Same as Fig. 3 except with non-logarithmic potential scale.
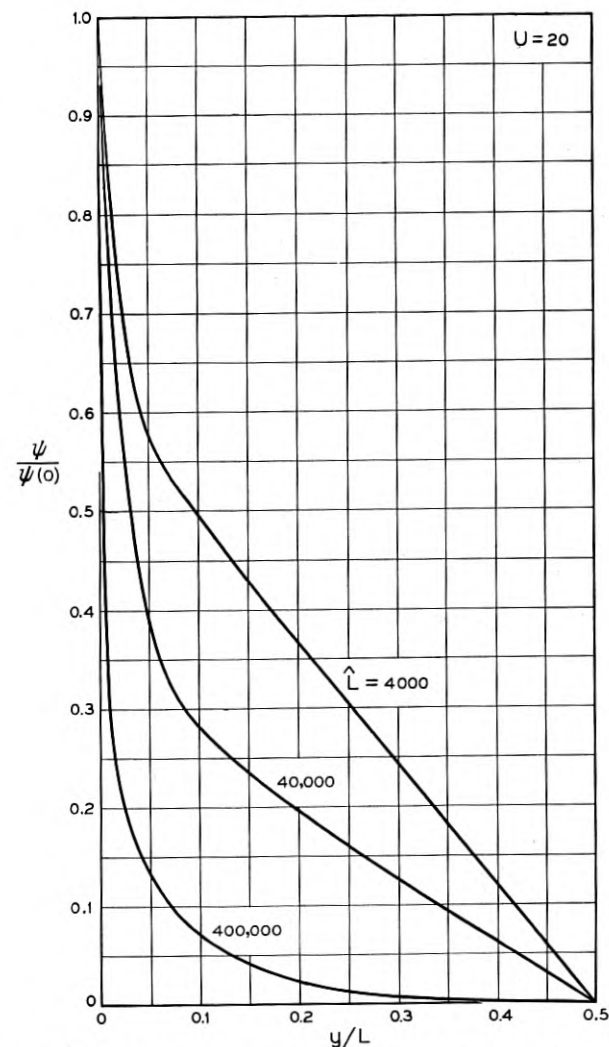
Fig. 10 — Similar to Fig. 9 except for $U =$ (Potential of *n-intrinsic* junction/ Boltzmann voltage) $= 20$ and several values of $\hat{L} =$ (Intrinsic layer thickness/Mean Debye length).

Because of the extreme values of the parameters encountered in these computations, it was necessary to make extensive use of the following expansions to supplement available tables of elliptic functions:

for $k^2 \ll 1$

$$K[k] = \frac{\pi}{2}\left[1 + 2\left(\frac{k^2}{8}\right) + 9\left(\frac{k^2}{8}\right)^2 + \cdots\right],$$

$$sn[k, v] = \sin v - \frac{k^2}{4}\cos v(v - \sin v \cos v) + \cdots$$

$$cn[k, v] = \cos v + \frac{k^2}{4}\sin v(v - \sin v \cos v) + \cdots$$

for $k' \equiv (1 - k^2)^{1/2} \ll 1$

$$K[k] = \ell n\frac{4}{k'} + \left(\ell n\frac{4}{k'} - 1\right)\frac{k'^2}{4} + \cdots$$

$$sn[k, v] = \tanh v + \frac{k'^2}{4}\operatorname{sech}^2 v(\sinh v \cosh v - v)$$

$$+ \cdots$$

$$cn[k, v] = \operatorname{sech} v - \frac{k'^2}{4}\tanh v \operatorname{sech} v(\sinh v \cosh v - v)$$

$$+ \cdots$$

Also useful were:

for $\phi \ll 1$

$$\ell n\frac{1 + \phi}{1 - \phi} = 2\phi\left(1 + \frac{1}{3}\phi^2 + \cdots\right)$$

for $\phi \gg 1$

$$\ell n\frac{1 + \phi}{1 - \phi} = \ell n\frac{2}{1 - \phi} - \frac{1 - \phi}{2} - \frac{1}{2}\left(\frac{1 - \phi}{2}\right)^2 - \cdots$$

In the determination of B from (43) the problem arises of solving the transcendental equation

$$k'K[(1 - k'^2)] = a \ll 1$$

where a is a given positive quantity. This equation must be solved by iteration or plotting and a reasonably good estimate of the root saves a great deal of labor. Making use of the approximation valid for $k' \ll 1$

$$K[(1 - k'^2)^{1/2}] \approx \ell n4/k'$$

leads to

$$k' \, \ell n 4 / k' \approx a.$$

Now set $k' \equiv 4e^{-r}$ to obtain

$$4re^{-r} \approx a$$

or

$$f(r) \equiv r - \ell n r - \ell n 4 / a \approx 0.$$

A rough approximation to r is obtained by neglecting $\ell n r$ in comparison to r, it is $r \approx \ell n \, 4/a$. This can be used to compute a much better approximation from Newton's formula:

$$r - r_0 \approx -\frac{f(r_0)}{f'(r_0)}.$$

The resulting useful formula is

$$r \approx \ell n 4 / a \left[1 + \frac{\ell n \ell n 4 / a}{\ell n 4 / a - 1} \right].$$

DISCUSSION OF FIELD DISTRIBUTION CURVES

For the symmetric case ($N = P$) for which numerical computations were made, the electric field is a minimum in the middle of the intrinsic region, and rises to symmetrical maxima at the extrinsic-intrinsic interfaces. For fixed $\hat{L}$ and relatively small U the field is very small except quite near the interfaces. That is, practically all the potential drop takes place in thin "space charge layers" near the intrinsic boundaries at $\hat{y} = 0$ and $\hat{y} = \hat{L}$. As U is increased (for given $\hat{L}$) the region of appreciable field strength increases in width and the minimum field at $y = L/2$ increases. As U is made very large the minimum field strength approaches the average field over the intrinsic region. In this limit the potential distribution approaches linearity across the intrinsic region.

For $N \neq P$, the qualitative behavior of the field is the same except that the minimum field ($\hat{\psi} = 0$) moves away from $\hat{y} = \hat{L}/2$ and the maxima at $\hat{y} = 0$ and $\hat{y} = \hat{L}$ are no longer equal. This minimum field (see (38b), (39b)) occurs at

$$y = C \approx \tfrac{1}{2} \hat{L} + e^{1/2} \left(\Lambda^{1/2} - \Lambda^{-1/2} \right)$$

Unless the asymmetry is very pronounced, this will not differ appreciably from $\hat{L}/2$ because of (40 c, d). It will be shown in a following section

that the normalized voltage at the p-intrinsic interface is given by

$$V \equiv -\hat{\psi}(\hat{L}) \approx U - 2 \ln \Lambda. \tag{44}$$

The average normalized field intensity over the intrinsic region is then given by

$$\hat{E}_{\mathrm{av}} \equiv \frac{U + V}{\hat{L}} \approx \frac{2U}{\hat{L}}\left(1 - \frac{\ln \Lambda}{U}\right). \tag{45}$$

For $\Lambda = 1$, of course

$$\hat{E}_{\mathrm{av}} = \frac{2U}{\hat{L}}.$$

Unless the degree of asymmetry is great,

$$\frac{\ell n \Lambda}{U}$$

will be small compared to unity and $\hat{E}_{av}$ will still be almost

$$\frac{2U}{\hat{L}}.$$

In order to present a simple quantitiative picture of the penetration of the field into the interior of the intrinsic region it is convenient to introduce the *field penetration parameter* η defined by

$$\eta = \frac{\text{minimum field intensity}}{\text{average field intensity}}.$$

For fixed intrinsic region thickness and small applied voltages, $\eta \ll 1$, indicating localization of the regions of high field intensity. For large applied voltage, $\eta \to 1$, implying substantial uniformity of field throughout the intrinsic region. Subject to the assumptions (40) relations will now be derived among the applied voltage, the intrinsic layer thickness, the asymmetry parameter Λ, and the field penetration parameter η.

The minimum field intensity occurs where $\hat{\psi} = 0$. Therefore, from (36) and (45)

$$\eta \approx \frac{[2A(B + 1)]^{1/2}}{2\hat{L}^{-1}(U - \ell n \Lambda)}. \tag{46}$$

Eliminating A from (46) by the use of (41) and $\hat{L}$ by the use of (43) leads to

$$\eta \approx \frac{2^{1/2}(B + 1)^{1/2}\Phi(B)}{U - \ell n \Lambda}$$
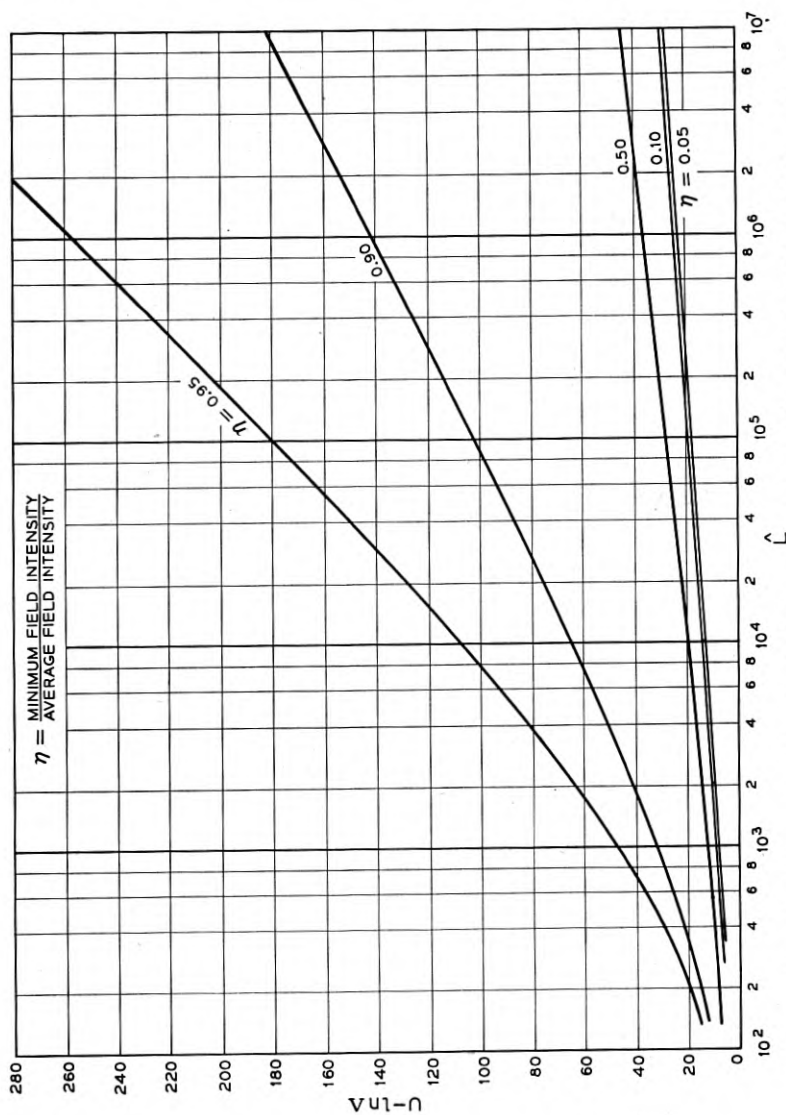
Fig. 11 — Half of voltage drop across intrinsic layer versus normalized intrinsic layer thickness for several values of the field penetration parameter η.

or

$$U - \ell n \, \Lambda \approx \frac{2}{\eta} \left(\frac{B+1}{2} \right)^{1/2} \Phi(B). \qquad (47)$$

Then substitution of (47) into (43) yields

$$\hat{L} \approx 2e^{1/2}\Phi(B) \exp \left[\frac{1}{\eta} \left(\frac{B+1}{2} \right)^{1/2} \Phi(B) \right]. \qquad (48)$$

For fixed η, (47) and (48) are parametric equations of a function $U_\Lambda(\hat{L})$. These equations were used to compute the curves of Fig. 11 in which $U - \ell n \Lambda$ is plotted against $\hat{L}$ for $\eta = 0.05, 0.1, 0.5, 0.9$ and 0.95. This figure gives a quantitative picture of the dependence of the field penetration parameter η on impressed voltage and intrinsic region thickness. [The ordinate $U - \ell n \Lambda$ is $\frac{1}{2}$ the total voltage drop across the intrinsic layer in (kT/q) units.].

The foregoing analysis clarifies the progressive elimination of the low field region near the center of the intrinsic material as the applied voltage is increased for fixed L. Now the high field regions near $y = 0$ and $y = L$ will be described. Making use of (41), together with

$$\cosh U \approx \tfrac{1}{2}eU$$

and

$$2AB \ll \Lambda, \Lambda^{-1}$$

implied by (40), in (36) leads to

$$\hat{E}(0) \approx e^{-1/2}\Lambda^{1/2} \qquad (49)$$

and to

$$\frac{d\hat{E}}{d\hat{y}}(0) \approx -\tfrac{1}{2}\, e^{-1}\Lambda. \qquad (50)$$

Hence a length characterizing the "space charge layer thickness" at $y = 0$ is

$$\frac{\hat{E}(0)}{-\dfrac{d\hat{E}}{d\hat{y}}(0)} \approx 2e^{1/2}\Lambda^{-1/2}. \qquad (51)$$

Similarly, for the p-intrinsic interface at $\hat{y} = \hat{L}$, we have

$$E(\hat{L}) \approx e^{-1/2}\Lambda^{-1/2}, \qquad (52)$$

$$\frac{d\hat{E}}{d\hat{y}}(L) \approx -\tfrac{1}{2}e^{-1}\Lambda^{-1}, \qquad (53)$$

and

$$\frac{\hat{E}(\hat{L})}{-\dfrac{d\hat{E}}{d\hat{y}}(\hat{L})} \approx 2e^{1/2}\Lambda^{1/2}. \tag{54}$$

If it is noted that the n-region and p-region Debye lengths are related to the mean Debye length $\mathcal{L}$ and the asymmetry parameter Λ by

$$\mathcal{L}_n = \Lambda^{-1/2}\mathcal{L}$$

and

$$\mathcal{L}_p = \Lambda^{1/2}\mathcal{L}$$

it is seen that (49)–(54) can be written:

$$E(0) \approx e^{-1/2}\left[\frac{kT}{q}\text{ per }\mathcal{L}_n\right], \tag{49'}$$

$$E(L) \approx e^{-1/2}\left[\frac{kT}{q}\text{ per }\mathcal{L}_p\right], \tag{52'}$$

$$\frac{dE}{dy}(0) \approx -\tfrac{1}{2}e^{-1}\left[\frac{kT}{q}\text{ per }\mathcal{L}_n\text{ per }\mathcal{L}_n\right], \tag{50'}$$

$$\frac{dE}{dy}(L) \approx -\tfrac{1}{2}e^{-1}\left[\frac{kT}{q}\text{ per }\mathcal{L}_p\text{ per }\mathcal{L}_p\right], \tag{53'}$$

$$\frac{E(0)}{-\dfrac{dE}{dy}(0)} \approx 2e^{1/2} \qquad [\text{times }\mathcal{L}_n], \tag{51'}$$

$$\frac{E(L)}{-\dfrac{dE}{dy}(L)} \approx 2e^{1/2} \qquad [\text{times }\mathcal{L}_p]. \tag{54'}$$

Equations (49')–(54') show that the maximum field intensity and the "space charge layer thickness" at either extrinsic junction is dependent only on the Debye length of the adjacent extrinsic material.

Similarly, the concentrations of the majority carrier at either interface is found to be substantially independent of U and L, and determined by the neighboring extrinsic material:

$$\frac{n(0)}{N} \approx e^{-1}, \tag{55}$$

$$\frac{p(L)}{P} \approx e^{-1}. \tag{56}$$

The minority carrier concentration at the interfaces depends too, on the neighboring extrinsic material, but is also U dependent:

$$\frac{p(0)}{N} \approx e^{-1}e^{-2U}, \tag{57}$$

$$\frac{n(L)}{P} \approx e^{-1}e^{-2U}. \tag{58}$$

At the point of minimum field intensity, $\hat{\psi} = 0$ and

$$\frac{n}{N} = \frac{p}{N} \approx e^{-1}e^{-U}. \tag{59}$$

The extremely low carrier concentrations given by (57)–(59) are not really meaningful, of course, because the analysis has neglected the carrier concentrations due to thermally generated hole-electron pairs and to saturation currents injected through the biased junctions. While these latter concentrations are negligible compared to (55) and (56), they are undoubtedly large compared to (57) and (58). Therefore, although they can be neglected in determining the electric field intensity distribution, they are of principal importance in determining the small residue carrier concentrations in the "swept" region. A computation of these concentrations can be made by regarding the fields determined in the present analysis as impressed and studying the resulting motion of the generated and injected carriers.

APPENDIX I

EVALUATION OF INTEGRATION CONSTANTS A, B, AND C.

In this section the conditions described in (7)–(12) will be used to evaluate the integration constants A, B, and C in terms of the prescribed parameters L, U, and Λ.

First a partial integration of the differential equations for the extrinsic regions will be performed to obtain from (7)–(10) relations between $\hat{E}$ and $\hat{n}$ at $\hat{y} = 0$ and $\hat{E}$ and $\hat{p}$ at $\hat{y} = \hat{L}$. Division of (25a) by (27) gives (for the n-region $\hat{y} < 0$)

$$\frac{d\hat{E}^2}{d\hat{n}} = 1 - \frac{\hat{p}}{\hat{n}} - \frac{\Lambda}{\hat{n}}. \tag{60}$$

Addition of $\hat{n}$ times (26) to $\hat{p}$ times (27) yields

$$\frac{d}{d\hat{y}}(\hat{p}\hat{n}) = 0,$$

whence

$$\hat{p} = \frac{\nu}{\hat{n}}, \tag{61}$$

where ν is a constant. Condition (7_2) then implies

$$\nu = \frac{n_i^2}{NP}. \tag{62}$$

Substituting (61) into (60) and integrating we obtain

$$\hat{E}^2 = \hat{n} - \Lambda \ln \hat{n} + \frac{\nu}{\hat{n}} + D,$$

where D is a constant of integration. Now by (7_3), for $\hat{y} = -\infty$ $E = 0$ and

$$\hat{n} = \Lambda + \hat{p} = \Lambda + \frac{\nu}{\hat{n}}$$

$$= \Lambda \left(1 + \frac{\nu}{\Lambda \hat{n}}\right).$$

Or, since

$$\frac{\nu}{\Lambda \hat{n}} \equiv \frac{n_i^2}{nN} \approx \left(\frac{n_i}{N}\right)^2 \ll 1,$$

$$\hat{n} \approx \Lambda \text{ for } \hat{E} = 0.$$

Therefore

$$D \approx -\Lambda + \Lambda \ln \Lambda - \frac{\nu}{\Lambda},$$

giving

$$\hat{E}^2 \approx (\hat{n} - \Lambda) \left(1 - \frac{\nu}{\Lambda \hat{n}}\right) + \Lambda \ln \frac{\Lambda}{\hat{n}},$$

or, for $(1 - \hat{n}/\Lambda)$ positive but not $\ll 1$,

$$\hat{E}^2 \approx \hat{n} + \Lambda \ln \frac{\Lambda}{e\hat{n}} \qquad [\text{for } \hat{y} = 0]. \tag{63}$$

A directly parallel computation for the p-region $(\hat{y} > \hat{L})$ leads to

$$\hat{E}^2 \approx \hat{p} + \Lambda^{-1} \ln \frac{\Lambda^{-1}}{e\hat{p}} \qquad [\text{for } \hat{y} = \hat{L}]. \tag{64}$$

Upon substituting (34) and (36) into (63) and setting $\hat{\psi} = U$ (12), we obtain

$$Ae^{-U} + 2AB = \Lambda \ln \frac{\Lambda}{Ae^{U+1}}, \qquad (65)$$

or, because both terms on the left are negligible under the assumptions (40)*,

$$A \approx \Lambda e^{-U-1}. \qquad (66)$$

Similarly, substitution of (35) and (36) into (63) and setting of $\hat{\psi} = -V$ yields

$$Ae^{-V} + 2AB = \Lambda^{-1} \ln \frac{\Lambda^{-1}}{Ae^{V+1}}, \qquad (67)$$

or, by virtue of (40),

$$A \approx \Lambda^{-1}e^{-V-1}. \qquad (68)$$

Combining (66) and (68) we obtain

$$V \approx U - 2 \ln \Lambda. \qquad (69)$$

This formula gives the normalized potential magnitude at the p-intrinsic interface in terms of that (U) at the n-intrinsic interface and the asymmetry parameter Λ.

Now the condition $\hat{\psi} = U$ for $\hat{y} = 0$ requires (from 38a and 39a) that

$$K\left[\left(\frac{1-B}{2}\right)^{1/2}\right] - F\left[\left(\frac{1-B}{2}\right)^{1/2} \sin^{-1}_{(1)} \operatorname{sech} \frac{U}{2}\right] = A^{1/2}C \qquad (70\mathrm{a})$$

for $-1 < B \leq 1$,

or

$$\left(\frac{2}{B+1}\right)^{1/2} F\left[\left(\frac{B-1}{B+1}\right)^{1/2}, \quad \sin^{-1}_{(1)} \tanh \frac{U}{2}\right] = A^{1/2}C \qquad (70\mathrm{b})$$

for $B \geq 1$.

In addition, the condition $\hat{\psi} = -V$ for $\hat{y} = \hat{L}$ requires

$$K\left[\left(\frac{1-B}{2}\right)^{1/2}\right] - F\left[\left(\frac{1-B}{2}\right)^{1/2}, \quad \sin^{-1}_{(2)} \operatorname{sech} \frac{V}{2}\right] = \qquad (71\mathrm{a})$$

$$A^{1/2}(C - \hat{L})$$

for $-1 < B \leq 1$,

* It will be shown later that (40) implies AB ≪ 1.

or

$$\left(\frac{2}{B+1}\right)^{1/2} F\left[\left(\frac{B-1}{B+1}\right)^{1/2}, \quad \sin_{(4)}^{-1} \tanh\left(\frac{-V}{2}\right)\right] = \quad (71b)$$

$$A^{1/2}(C - \hat{L})$$

for $B \geq 1$.

Fortunately, because of the assumptions (40a, b) the formidable relations (70a–71b) can be simplified to

$$\Phi(B) \approx A^{1/2}C + 2e^{-U/2} \qquad (72)$$

and

$$\Phi(B) \approx A^{1/2}(\hat{L} - C) + 2e^{-V/2}, \qquad (73)$$

where

$$\Phi(B) \equiv \begin{cases} K\left[\left(\frac{1-B}{2}\right)^{1/2}\right] & \text{for } -1 < B < 1 \\ \dfrac{\pi}{2} & \text{for } B = 1 \\ \left(\frac{2}{B+1}\right)^{1/2} K\left[\left(\frac{B-1}{B+1}\right)^{1/2}\right] & \text{for } B > 1. \end{cases}$$

Subtracting (73) from (72) we obtain

$$0 = 2A^{1/2}C - A^{1/2}\hat{L} + 2e^{-U/2} - 2e^{-V/2},$$

whence, substituting (66) and (69),

$$C \approx \frac{1}{2}\hat{L} + e^{1/2}(\Lambda^{1/2} - \Lambda^{-1/2}). \qquad (74)$$

Finally, substitution of (66) and (74) into (72) gives

$$\Phi(B) \approx [\frac{1}{2}\Lambda^{1/2}\hat{L} + e^{1/2}(\Lambda + 1)]e^{-U/2-1/2},$$

or, by virtue of (40c, d),

$$\Phi(B) \approx \frac{1}{2}\Lambda^{1/2}\hat{L}e^{-U/2-1/2}. \qquad (75)$$

Equations (66), (69), (74), and (75) are the desired expressions for determining A, B, C, and V when values are assigned to Λ, $\hat{L}$, and U (subject to (40)). (If necessary, some or all of the restrictions (40) could

be eliminated, but the transcendental equations to be solved for A, B, C, and V would become quite formidable.)

It should be noted that (75) permits an easy determination whether the formulae for $B < 1$ or those for $B > 1$ should be used in any particular case. Since $\Phi(1) = \pi/2$,

$$B \lessgtr 1 \text{ for } \Lambda^{1/2}\hat{L}e^{-U/2-1/2} \gtrless \pi. \tag{76}$$

Appendix II

conditions for $2AB \ll \Lambda$, Λ^{-1}

It has been stated without proof in the foregoing analysis that the conditions (40) imply $2AB \ll \Lambda$, Λ^{-1} (and hence also $AB \ll 1$). This must now be demonstrated.

For B *not* $\gg 1$, (40a–40d) are sufficient, for (41) shows that $A \ll 1$. However, for $B \gg 1$, the product AB is not necessarily small because of $A \ll 1$ and additional limitations are required. To establish suitable additional conditions we shall consider combinations of U, $\hat{L}$, and Λ for which AB *is* very small and estimate the conditions under which this smallness begins to weaken.

By eliminating U between (41) and (43) we can write

$$\Phi(B) \approx \tfrac{1}{2}\hat{L}A^{1/2},$$

or

$$B^{1/2}\Phi(B) \approx \tfrac{1}{2}\hat{L}(AB)^{1/2}. \tag{77}$$

Now for $k \approx 1$,

$$K(k) \approx \ell\text{n} \frac{4}{\sqrt{1 - k^2}}.$$

Therefore, for $B \gg 1$

$$\Phi(B) \equiv \left(\frac{2}{B + 1}\right)^{1/2} K\!\left[\left(\frac{B - 1}{B + 1}\right)^{1/2}\right] \approx \frac{\ell\text{n}\,B}{(2B)^{1/2}}. \tag{78}$$

Substitution of (78) into (77) now yields

$$\ell\text{n}\,B \approx 2^{-1/2}\hat{L}(AB)^{1/2},$$

or

$$B \approx \exp\,[2^{-1/2}\hat{L}(AB)^{1/2}]. \tag{79}$$

From (78) and (79)

$$\Phi(B) \approx \tfrac{1}{2}\hat{L}(AB)^{1/2} \exp\left[-2^{-3/2}\hat{L}(AB)\right]. \tag{80}$$

Now using the expression just obtained for $\Phi(B)$ in terms of (AB), we can eliminate $\Phi(B)$ from (43) and solve for U as a function of (AB). After some manipulations we obtain

$$\frac{2U}{\hat{L}} \approx (2AB)^{1/2} - \frac{2(1 - \ln 2)}{\hat{L}} - \frac{2}{\hat{L}} \ln\left(\frac{2AB}{\Lambda}\right). \tag{81}$$

For $2AB\Lambda$ or $2AB\Lambda^{-1}$ of the order of 10^{-2}, (40c, d) insure that the right side of (81) will not change in order of magnitude if the two terms multiplied by $\hat{L}^{-1}$ are dropped. Therefore $2AB\Lambda$ and $2AB\Lambda^{-1}$ will be of order 10^{-2} or less if

$$\frac{2U}{\hat{L}} < \frac{1}{10}\,\Lambda^{1/2},\ \frac{1}{10}\,\Lambda^{-1/2}. \tag{82}$$

Thus for a given intrinsic layer thickness there is a ceiling on the impressed voltages for which AB is negligibly small.

Transmission Properties of Laminated Clogston Type Conductors

By E. F. VAAGE

The transmission properties of ideal laminated conductors of the Clogston type are discussed by introducing the concepts of equivalent inductance, capacitance and resistance values which are analogous to their corresponding counterparts in the treatment of ordinary transmission lines. From these constants the attenuation, phase constant, and speed of propagation are obtained using conventional transmission line theory, and the results compared with those for ordinary coaxial conductors.

This paper is divided into two parts. In the first part a general discussion is given of Clogston cables and a comparison made with the conventional coaxial cable. This is illustrated with a few numerical examples, based on formulas which are developed in the second part of this paper.

INTRODUCTION

The discovery that deep penetration of the current can be obtained in laminated conductors, when the speed of propagation is made constant over the entire cross-section of the cable, is described in an earlier issue of this magazine.[1] The theoretical study of the problem was based on Maxwell's field equations dealing with a stack of parallel plates of alternate conducting and insulating layers. When applied to concentric laminated tubes, this method results in a set of extremely complex equations. S. P. Morgan has given a rigorous solution for the case when the laminated layers are of infinitesimal thickness.[2]

The present paper uses a different approach which leads to simpler approximate formulas. Available theoretical results are combined with simplifying approximations and certain somewhat arbitrary assumptions

---

[1] Clogston, A. M., Reduction of Skin Effect Losses by the use of Laminated Conductors. Bell System Tech. J., **30**, pp. 491–529, July, 1951.

[2] Morgan, S. P., Mathematical Theory of Laminated Transmission Lines. Bell System Tech. J., Part I, **31**, pp. 883–949, Sept., 1952, and Part 2, **31**, pp. 1121–1206, Nov., 1952.

in such a way that formulas for the counterparts of the usual transmission line properties of inductance, capacitance and resistance are obtained. The approximate formulas for attenuation, phase constant and speed of propagation are then derived using conventional transmission line theory.

Some computations of attenuation are presented which illustrate the interesting transmission properties of Clogston cables under ideal conditions. Exploratory work on this type of conductor is still in the early research stage, with some very difficult problems imposed by the need for a large number of thin layers with very close tolerances.

PART I — GENERAL DISCUSSION OF CLOGSTON CABLES AND A COMPARISON WITH CONVENTIONAL COAXIAL CABLE

1. SKIN EFFECT

An alternating current transmitted over a solid conductor has the tendency to crowd toward the surface of the wire. This phenomenon is known as the *skin effect* and the depth of penetration of the current is usually referred to as the *skin depth*. The skin depth is defined as the distance, measured from the surface toward the center of the wire, where the current density is reduced to $1/e = 0.367$. For a copper conductor it is given by:

$$\delta = \frac{2.61}{\sqrt{F_{mc}}}, \tag{1}$$

where

$$\delta = \text{Skin depth in mils.}$$

$$F_{mc} = \text{Frequency in megacycles.}$$

When the skin depth is a fraction of the wire radius, the ac resistance of the wire increases about as the square root of the frequency.

Laminated conductors disclosed by Clogston have the property that the ac resistance will remain very nearly equal to the dc resistance over a wide band of frequencies, if the conducting and insulating layers can be made thin enough and sufficiently uniform. The dc resistance of a Clogston conductor will be higher than the dc resistance of a solid conductor of the same over-all dimension by a factor $(w + t)/w$, where w and t are the thicknesses of the conducting and insulating layers respectively. As discovered by Clogston, the depth of penetration in a laminated conductor is much greater than in a conductor of solid copper if the

speed of propagation is constant over the entire cross-section. A coaxial cable having an inner laminated conductor and an outer laminated sheath must obey the following relation to obtain the desired effect:

$$\mu_I \epsilon_I = \mu \epsilon \left(1 + \frac{w}{t}\right),\tag{2}$$

where:

μ_I = Permeability in space between inner and outer conductor.

μ = Permeability of laminated conductors.

ϵ_I = Dielectric constant of insulation between inner and outer conductor.

ϵ = Dielectric constant of insulating layers in the laminated conductors.

w = Thickness of copper layers.

t = Thickness of insulating layers.

In (2) the expression $\epsilon(1 + w/t)$ is of course the mean dielectric constant of the laminated conductor. Since $1/\sqrt{\mu_I \epsilon_I}$ is the speed of propagation in the main dielectric, equation (2) indicates that the speed of propagation is the same over the entire cross section of the cable. Equation (2) must be satisfied to a high degree of accuracy, otherwise deep penetration is not possible.

2. DEFINITION OF CLOGSTON CABLES

Two laminated conductors arranged as a coaxial cable are shown in Fig. 1. The inner conductor consists of a solid copper wire of diameter d_1, over which a large number of alternate layers of insulation and copper are arranged as concentric thin tubes. The over-all diameter of the inner conductor is D_1. The outer conductor of the coaxial cable consists of a laminated tube of inner diameter d_2 and outer diameter D_2. The space between D_2 and d_2 is filled with thin concentric tubes of copper and insulation of the same thicknesses as for the inner conductor. The outside of the outer conductor is covered with a solid copper sheath for protection, shielding and energizing purposes. This type of cable has been named Clogston I.

By adding more layers to the outside of the inner conductor and more layers to the inside of the outer conductor, the space between them is completely filled when $d_2 = D_1$. Such a cable is shown in Fig. 2, and has been named Clogston II.

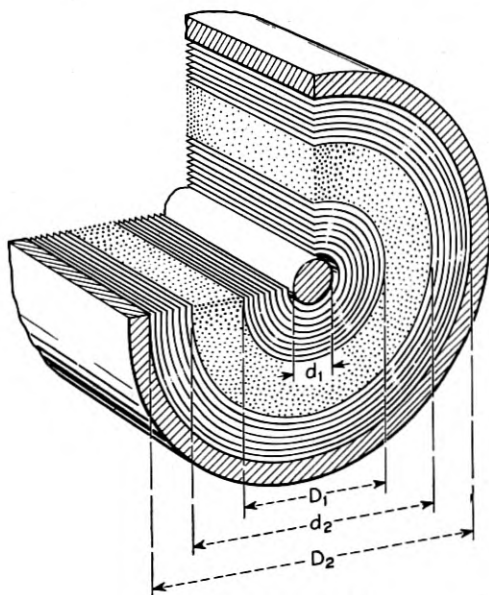Clogston I may be thought of as a physical variant of the conventional
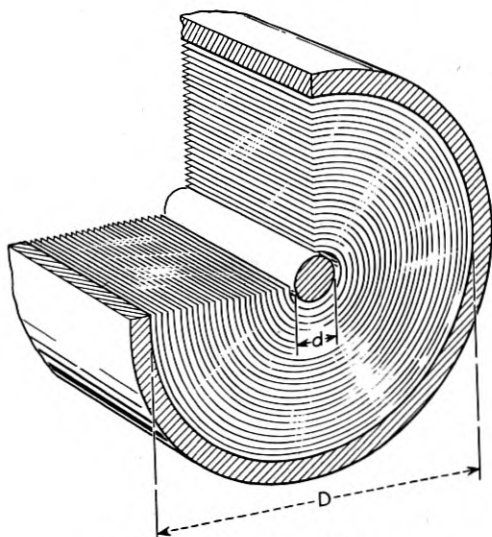
Fig. 1 — Clogston I cable.

Fig. 2 — Clogston II cable.

coaxial cable, in that when one of the conductors carriers a current in one direction, the other will carry a current in the opposite direction. In Clogston II there is also a reversal of currents somewhere between the outermost layers and the innermost. It is therefore a kind of a two conductor cable, but the point of division between the conductors is determined by the electromagnetic field configuration of the situation. This point has been worked out by S. P. Morgan and will be referred to in the second part of this paper.

3. OPTIMUM PROPORTIONING

The cross-sectional aspect of Clogston II is completely characterized by that proportion of the diameter D which is occupied by the laminations. This proportion is called the *Fill Factor* and is defined by:

$$\phi_{II} = (D - d)/D \qquad \text{Clogston II.} \tag{3}$$

The fill factor is also a useful parameter for Clogston I, though it is not sufficient to determine its geometry. It is defined by:

$$\phi_{I} = (D_1 - d_1 + D_2 - d_2)/D_2 \qquad \text{Clogston I.} \tag{4}$$

The additional parameters which, with the outer diameter D_2, will completely determine the geometry, are the ratio of the over-all thickness of the inner laminated conductor to the over-all thickness of the outer conductor, and a parameter which locates the inner diameter d_1 of the inner conductor. These parameters are defined by:

$$T = (D_1 - d_1)/(D_2 - d_2),$$
$$U = d_1/D_2. \tag{5}$$

In a conventional coaxial cable, shown in Fig. 3, the optimum value of attenuation[3] is obtained when $D/d = 3.59$. In Clogston cables no such optimum values exist.

S. P. Morgan, however, has shown that there are useful relative optimum relations in Clogston I, which direct the choice of T and U for the cables which are illustrated in this paper. For example, for a fill factor of one-half, there is a broad optimum of attenuation when $T = 1.96$ and $U = 0.0842$. Thus the over-all thickness of the inner conductor is about twice that of the outer conductor, and the diameter d_1 of the inner core is about one-twelfth of the outer diameter D_2. With

[3] Green, E. I., F. A. Leibe and H. E. Curtis, The Proportioning of Shielded Circuits for Minimum High-Frequency Attenuation, Bell System Tech. J., **15**, pp. 248–283, April, 1936.

these dimensions the cross sectional area of the outer conductor is nearly twice that of the inner conductor, so that the optimum arises from other causes than matching the conductivity of the two conductors.

The problem of determining the relative optimum values of Clogston I, in general, is complicated, but numerical studies indicate that for values of the fill factor other than one-half, the values for T and U are not greatly different.

Another factor in Clogston cables which can be optimized is the ratio of the layer thicknesses of the conducting and the insulating layers. Clogston[4] has shown that in the frequency range where the attenuation is substantially flat with frequency, this optimum ratio is equal to:

$$w/t = 2. \tag{6}$$

For this condition, the dc resistance of the laminated conductor is increased by $(w + t)/w$ or $3/2$ over a solid conductor. The dielectric constant of the main insulation, according to (2) above must equal $\epsilon_I = 3\epsilon$, which reduces the speed of propagation by $\sqrt{3}$, assuming $\mu_I = \mu$.

At frequencies where the attenuation begins to increase, other optimum values of w/t can be obtained, and the ratio will depend upon what top frequency is considered.

In a practical case a fill factor of unity will probably not be used. A little space in the center will be made available for a solid conductor for energizing or other purposes.
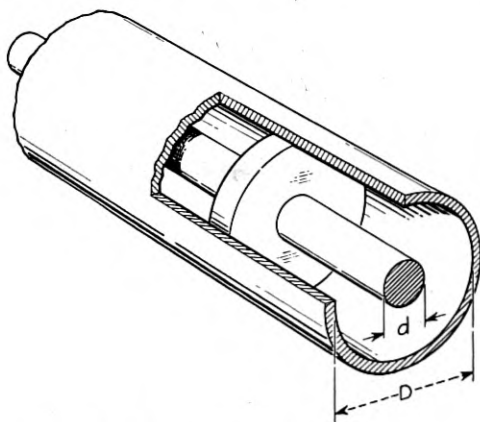

Fig. 3 — Conventional coaxial cable.

[4] Loc. cit.

4. ATTENUATION

The attenuation of a transmission circuit at high frequencies, where $\omega L \gg R$ and $\omega C \gg G$, is usually given in the following form:

$$a = \frac{R}{2} \sqrt{\frac{C}{L}} + \frac{G}{2} \sqrt{\frac{L}{C}}, \qquad (7)$$

where R is the total ac resistance of both conductors, L, C and G the inductance, capacitance and leakance of the circuit. It will be assumed that the insulation consists of polyethylene or some other material having a very low leakance. Thus as a first approximation the second term in (7) can be neglected.

In a conventional coaxial cable, in the frequency range considered, R will increase in proportion to the square root of frequency, so that neglecting leakance, (7) may be written as follows:

$$a = \frac{K_1}{D} \sqrt{F_{mc}}, \qquad (8)$$

where K_1 is a constant depending upon the dielectric constant of the insulating material and the resistivity of the conductors. D is the inside diameter of the sheath, and F_{mc} the frequency in megacycles.

In the second part of this paper it is shown that the attenuation of Clogston I or II cables can be written in the following form:

$$a_c = \frac{K_2}{D_2} + K_3 w^2 F^2_{mc}, \qquad (9)$$

where D = Over-all diameter of laminated cable.

w = Copper layer thickness.

K_2 and K_3 are constants, different for Clogston I and Clogston II, which depend upon the geometry of the cables, the dielectric constant of the insulating material and the resistivity of the conducting layers.

The first term in (9) gives a constant loss independent of frequency. The second term contributes little to the attenuation provided w is small enough. In fact, the attenuation will remain constant within $p\%$ of the first term provided:

$$w \leqq \frac{1}{10} \sqrt{\frac{pK_2}{K_3}} \frac{1}{DF'_{mc}}. \qquad (10)$$

This equation (10) determines the copper layer thickness, which will result in a "flat" attenuation within p per cent up to a frequency F'_{mc}.

By neglecting the second term in (9) and comparing the result with (8) it can be seen that the attenuation of a conventional coaxial cable and a Clogston cable are equal at a frequency given by:

$$F_{mc}'' = \left[\frac{K_2}{K_1 D} \right]^2 \tag{11}$$

The numerical examples given later in this paper indicate that a Clogston cable will have higher attenuation than a conventional coaxial cable at frequencies below F_{mc}'', and less attenuation at higher frequencies. At frequencies sufficiently higher than F_{mc}'', the attenuation of a Clogston cable will increase rapidly which is evident from the second term in (9). It is in the region between F_{mc}'' and frequencies where the second term in (9) becomes important that Clogston cables can theoretically provide less attenuation than a conventional coaxial cable.

5. IMPEDANCE, PHASE CONSTANT AND SPEED OF PROPAGATION

The equivalent impedances of Clogston I and Clogston II cables are developed in the second part of this paper and are equal to:

$$Z_I = \left[1 + \frac{L_{in}}{L_{ex}} \right] \sqrt{\frac{L_{ex}}{C_{ex}}} \qquad \text{Clogston I,}$$

$$Z_{II} = \frac{L_{in}}{\sqrt{L_{ex} C_{ex}}} \qquad \text{Clogston II.} \tag{12}$$

In these expressions L_{in}, L_{ex} and C_{ex} are the internal and external inductances and capacitances respectively. For conventional coaxial cables they are discussed by S. A. Schelkunoff in "Electromagnetic Waves" (Van Nostrand, 1943). For the Clogston analogy, Part II of the present paper gives the reasoning adopted in defining them. In a Clogston II cable the external inductance goes to zero and the external capacitance to infinity, but the product $L_{ex}C_{ex}$ nevertheless remains constant.

The impedance of a conventional coaxial cable, in the frequency range considered, is given by:

$$Z = \sqrt{\frac{L}{C}} \tag{13}$$

where L and C are the external inductance and capacitance of the conventional coaxial cable.

The equivalent impedance of a Clogston cable is lower than that of conventional coaxial cable of the same outer diameter. Numerical eval-

uations using the above formulas indicate an impedance of about 23 ohms for a half-filled Clogston I cable, and about 11½ ohms for a completely filled Clogston II cable, assuming polyethylene insulation with ϵ equal to 2.3 and copper conducting layers having a thickness twice that of the insulating layers ($w/t = 2$). These values compare with about 76 ohms for a conventional coaxial cable having air dielectric and the same outer diameter and 51 ohms for a corresponding coaxial with solid polyethylene dielectric.

The phase constants of both Clogston cables in the frequency range where the attenuation is nearly flat, are equal and independent of the geometry of the cables. They are given by:

$$\beta_I = \beta_{II} = \omega\sqrt{L_{ex}C_{ex}},\tag{14}$$

assuming uniform layer thicknesses.

For a conventional coaxial cable the phase constant, neglecting leakance, is given approximately by:

$$\beta = \omega\sqrt{LC}\left[1 - \frac{1}{2}\left(\frac{R}{2\omega L}\right)^2\right].\tag{15}$$

The computed speed of propagation, which is equal to ω/β, is about 71,000 mi/sec for Clogston cables with polyethylene insulating layers. This compares with 123,000 mi/sec for a conventional coaxial cable with polyethylene insulation and 186,000 mi/sec for a coaxial with pure air dielectric.

6. COMPARISON WITH A CONVENTIONAL COAXIAL CABLE

To illustrate the effect of the various parameters involved in the attenuation of a Clogston cable, and to compare the result with a convential coaxial cable, a few numerical examples have been evaluated.

A one-half filled Clogston I and a completely filled Clogston II have been selected arbitrarily for comparison purposes. Fig. 4 shows the attenuation characteristics of these cables for several values of copper layer thicknesses. In each case, polyethylene insulation ($\epsilon = 2.3$) is assumed, with $w/t = 2$, i.e., the insulating layers have one-half the thickness of the conducting layers. In the same figure, the attenuation characteristics of two conventional coaxial cables of the same outer diameter, one with air dielectric and one with polyethylene insulation, are shown also. The regions where Clogston cables have in theory less attenuation than conventional coaxial cables of the same outer diameter can be seen in this figure.

Fig. 4 — Comparison of attenuation — conventional coaxial, Clogston I and Clogston II.

The dotted curve shown on Fig. 4 gives the skin depth in solid copper. It will be noted that the copper layer thicknesses become smaller and smaller fractions of the skin depth as the frequency increases. This is also evident from (1) and (10) above, which show that the copper layer thicknesses are inversely proportional to the frequency, while the skin depth is inversely proportional to the square root of the frequency.

The effect of fill is illustrated in Fig. 5 for a 375-mil Clogston I cable, for fill factors of one-eighth, one-quarter, and one-half. It will be noted that the attenuation increases rapidly with decrease in fill, accompanied by an increase in the frequency band over which the attenuation is flat. The attenuation of a completely filled Clogston II cable is also shown for comparison.

The above estimates are based on ideal conditions; that is, it is assumed that the laminated structures are perfectly uniform. The effects of departures from ideal are not shown by the approximate methods used in this paper, but it has been shown by Morgan[5] that even small departures from ideal conditions will result in increases and irregularities in attenuation, and a decrease in the band over which the attenuation is approximately uniform.
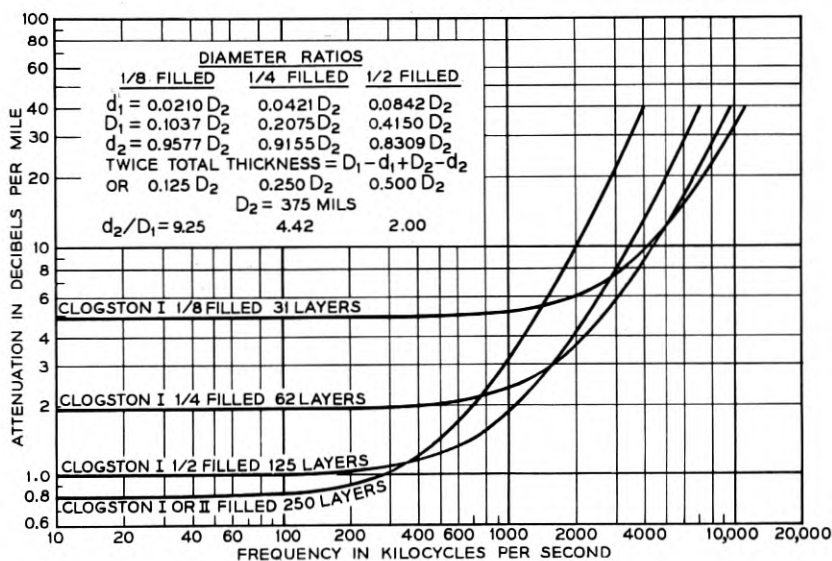

Fig. 5 — Attenuation of Clogston I.

[5] Loc. cit. Part II, pages 1161–1201.

PART II — DERIVATION AND EVALUATION OF FORMULAS

In this part of the paper, approximate formulas for the counterparts of the usual primary constants used in transmission line computations are derived for Clogston I and II type cables. From these formulas the various constants entering into the expressions for attenuation, impedance and phase as given in Part I are evaluated in terms of the dimensions of the cables and the frequency.

In deriving the formulas, certain simplifying assumptions have been made as explained in the test. The effects of these assumptions are examined by comparing the results with those obtained by more rigorous methods.

1. RESISTANCE IN GENERAL

A Clogston I cable consists of a laminated inner conductor and a laminated outer conductor as shown in Fig. 1. Each of these may be represented by the laminated conductor shown in Fig. 6. At present, we have no exact formula for the ac resistance of such conductors over a wide range of frequencies. However, S. P. Morgan has shown that for a stack of parallel plates, the resistance per unit cross-sectional area is equal to:

$$R_{ac} = R_{dc} \frac{w + t}{w} \left[1 + \frac{(n^2 - 1/5)w^4}{412} F_{mc}^2 + \dots \right], \qquad (16)$$

where R_{dc} is the dc resistance of the stack, when completely filled with
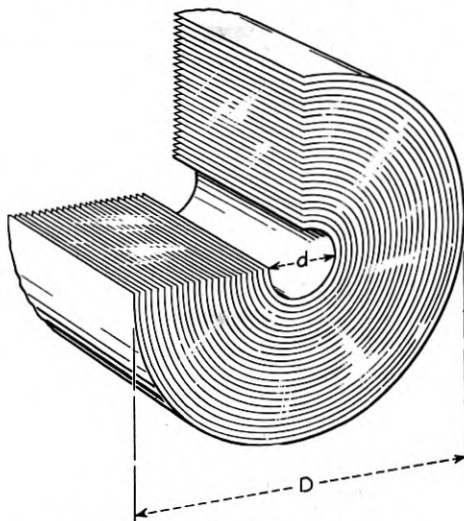

Fig. 6 — Laminated conductor.

copper. The other parameters are:

w = Thickness of copper layers in mils.
t = Thickness of insulating layers in mils.
n = Total number of layers.
F_{mc} = Frequency in megacycles.

With curvature disregarded, equation (16) also gives the ac resistance of the laminated conductor shown in Fig. 6. For only one copper layer, $n = 1$, and $t = 0$, (16) reduces to:

$$R_{ac} = R_{dc}\left[1 + \frac{w^4 F_{mc}^2}{515} + \dots\right],\tag{17}$$

which is exactly the expression for a copper tube at frequencies where the ac resistance begins to depart from the dc resistance and $\frac{1}{2}(D - d) = w$. For a single layer, the effect of curvature is very small[6] and can be disregarded. It will be assumed that (16) also holds to a fair degree of accuracy for a laminated conductor made up of a large number of layers. Since n is large, the small fraction one-fifth can be neglected. For the optimum condition (minimum attenuation) Clogston[7] has shown that:

$$w = 2t.\tag{18}$$

The total number of layers can be obtained from the following expression:

$$n = \frac{D - d}{2(w + t)} = \frac{D - d}{3w} \text{ for } w = 2t.\tag{19}$$

With (18) and (19) substituted in (16) the ac resistance is given by:

$$R_{ac} = \frac{82080}{D^2 - d^2}\left[1 + \frac{w^2(D - d)^2 F_{mc}^2}{3710} + \dots\right],\tag{20}$$

where the diameters and the copper layer thickness are given in mils and the frequency in megacycles. The resistivity of copper is taken to be 1.724×10^{-6} ohms/cm^3.

2. CLOGSTON I CABLE

2.1 Resistance

The ac resistances of the inner and outer laminated conductors of Clogston I cable, shown in Fig. 1, can be obtained from (20) above by substituting D_1 and d_1 for the inner conductor and D_2 and d_2 for the

[6] Schelkunoff, S. A., The Electromagnetic Theory of Coaxial Transmission Lines and Cylindrical Shields. Bell System Tech. J., 13, pp. 532–579, Oct., 1934. BSTJ, October, 1934, by S. A. Schelkunoff.
[7] Loc. cit.

outer conductor and adding the result. Thus:

$$R_{ac1} + R_{ac2} = \frac{82080A_1}{D_2^2} \left[1 + \frac{B_1 w^2 D_2^2 F_{mc}^2}{3710} + \ldots \right] \text{ohms/mi}, \quad (21)$$

where:

$$A_1 = \frac{D_2^2}{D_1^2 - d_1^2} + \frac{D_2^2}{D_2^2 - d_2^2}, \quad (22)$$

$$B_1 = \frac{2(D_1 D_2 - d_1 d_2)(D_1 - d_1)(D_2 - d_2)}{(D_1^2 - d_1^2 + D_2^2 - d_2^2) D_2^2}. \quad (23)$$

From (4) and (5) in the first part of this paper, it is possible to express A_1 and B_1 wholly in terms of ϕ_I, T and U. Thus, A_1 and B_1 are independent of D_2, and it follows that the first term in (21) is inversely proportional to D_2^2, while the second term, when multiplied out, is independent of D_2, assuming fixed values of ϕ_I, T and U.

2.2 Impedance, Inductance and Capacitance

In a coaxial cable the flux in the space between the two conductors gives the external inductance. The internal inductance is obtained from considering the flux within the walls of the conductors themselves but not that in the space between them. The effective inductance is then the sum of the two. Analogous considerations apply to the external and internal capacitance and the effective over-all capacitance is the value of the two acting in series.

Similarly in the frequency range where $\omega L \gg R$ and $\omega C \gg G$, but where the ac resistance is nearly equal to the dc resistance, the internal inductance and capacitance of the laminated conductors must be taken into account. Since they are in series with the external components they tend to increase the total inductance and decrease the total capacitance. The impedance of the circuit can therefore be expressed as follows:

$$Z_I = \sqrt{\frac{L}{C}} = \sqrt{\frac{L_{in} + L_{ex}}{\left[\dfrac{C_{in} C_{ex}}{C_{in} + C_{ex}} \right]}} \text{ ohms}, \quad (24)$$

$$L_{ex} = \frac{\mu_I}{2\pi} \ell n \left(\frac{d_2}{D_1} \right) \qquad \text{henries/cm}, \quad (25)$$

$$C_{ex} = \frac{2\pi \epsilon_I}{\ell n \left(\dfrac{d_2}{D_1} \right)} \qquad \text{farads/cm}, \quad (26)$$

where: $\epsilon_I = \epsilon(1 + w/t) =$ Dielectric constant of insulation between inner and outer conductors.

At the present time no exact formulas for the internal components of either L or C are available. The internal inductance, however, must be nearly equal to that of a solid wire for the inner laminated conductor and to that of a solid sheath for the outer laminated conductor, and will be assumed so in this paper.

It should be remembered that deep penetration of the currents will be obtained when Clogston condition of constant speed of propagation over the entire cross section of the cable is satisfied. The speed of propagation over the laminated conductors is equal to $1\sqrt{L_{in}C_{in}}$ and in the main dielectric between the two conductors equal to $1\sqrt{L_{ex}C_{ex}}$. Thus to obtain the Clogston condition the following relation must hold:

$$L_{in}C_{in} = L_{ex}C_{ex} . \tag{27}$$

By solving for the unknown quantity C_{in} and substituting the result in (24), the impedance of Clogston I cable is found to be:

$$Z_I = \left[1 + \frac{L_{in}}{L_{ex}}\right] \sqrt{\frac{L_{ex}}{C_{ex}}} \text{ ohms.} \tag{28}$$

Formulas for L_{in}, L_{ex} and C_{ex} in units convenient for numerical evaluation are given later in the section summarizing the formulas.

2.3 Attenuation and Phase

The attenuation of Clogston I cable neglecting leakance is obtained from the following expression:

$$\alpha_I = \frac{R_{ac1} + R_{ac2}}{2Z_I} . \tag{29}$$

The phase constant, at the frequencies considered, and where the ac resistance of the conductors does not depart appreciably from the dc value, is equal to:

$$\beta_I = \omega \sqrt{LC} = \omega \sqrt{(L_{in} + L_{ex}) \frac{C_{in}C_{ex}}{C_{in} + C_{ex}}} , \tag{30}$$

which with (27) above reduces to:

$$\beta_I = \omega\sqrt{L_{ex}C_{ex}} . \tag{31}$$

3. CLOGSTON II CABLE

3.1 *Resistance*

A Clogston II cable may be looked upon as having an "inner" and "outer" conductor which are separated by an infinitesimal amount in which R_{ac} given by (20) above is an expression for the parallel connection of R_{ac1} and R_{ac2} of a Clogston II cable having an outer diameter of D and an inner diameter of d. That is:

$$R_{ac} = \frac{R_{ac1}R_{ac2}}{R_{ac1} + R_{ac2}}. \tag{32}$$

It will now be assumed that (1) the currents flowing in one direction through the "inner" conductor and in the opposite direction in the "outer" conductor, will separate in such a way that R_{ac1} is equal to R_{ac2}, and (2) that the currents are uniformly distributed over the cross sections of the conductors. With these assumptions the respective cross sections would then be equal, and the reversal of the current would take place in a completely filled ($d = 0$) Clogston II cable at a radius equal to:

$$r = \frac{D}{2\sqrt{2}} = 0.3535D. \tag{33}$$

By substituting (32) and (33) in (20) it can be shown that:

$$R_{ac1} + R_{ac2} = \frac{328320}{D^2 - d^2} \left[1 + \frac{B_2 w^2 D^2 F_{mc}^2}{3710} + \cdots \right] \text{ ohms/mi}, \tag{34}$$

where

$$B_2 = 1 + \frac{d^2}{D^2} - \left[1 + \frac{d}{D} \right] \sqrt{\frac{1}{2} \left[1 + \frac{d^2}{D^2} \right]}. \tag{35}$$

The above assumptions relating to division and distribution of the current are not exact. S. P. Morgan has shown that the current distribution is not uniform and that the reversal of the current takes place at a radius equal to 0.3138D. As shown in a later section of this paper, the error resulting from these simplifying assumptions is not large.

3.2 Impedance, Attenuation and Phase

In a Clogston II cable, the main dielectric insulation between "inner" and "outer" conductor has vanished. Thus the external inductance approaches zero and the external capacitance becomes infinitely large.

From (25) and (26) above it is evident, however, that the product of $L_{ex}C_{ex}$ remains constant, since it is independent of the diameter ratio d_2/D_1, which in a Clogston II cable approaches unity. With (27) inserted in (24) and with $L_{ex} = 0$ and $C_{ex} = \infty$, the impedance of a Clogston II cable may be written:

$$Z_{II} = \sqrt{\frac{L_{in}}{C_{in}}} = \frac{L_{in}}{\sqrt{L_{ex}C_{ex}}} \text{ ohms.} \qquad (36)$$

The internal inductance of a Clogston II cable is not known, but will be taken equal to 0.1609×10^{-3} Henries/mi, which is the internal inductance of a pair of wires at low frequency. Thus:

$$Z_{II} = \frac{17.35}{\sqrt{\epsilon}} \text{ ohms,} \qquad (37)$$

where ϵ is the dielectric constant of the insulating layers.

The attenuation is obtained by dividing $R_{ac1} + R_{ac2}$ from (34) by $2Z_{II}$, where leakance is disregarded.

The phase constant, in the frequency range considered, is equal to the phase constant of a Clogston I cable since $L_{ex}C_{ex}$ is a constant value.

4. *Comparison of Results with Values Obtained from Rigorous Formulas*

S. P. Morgan[8] has developed rigorous formulas for the attenuation of Clogston cables, assuming infinitesimal thickness of the layers but retaining a fixed ratio of copper to insulating layer thicknesses. A correction term gives the increase in attenuation with frequency for layers of finite thickness.

The attenuation of a one-half filled Clogston I cable computed by the approximate formulas given in the present paper was 1.1 per cent higher than the value computed using Morgan's rigorous formulas. Similar computations on a completely filled Clogston II cable gave values 8.6 per cent higher. This decrease in accuracy with increase in fill is in line with the expectation that uniform distribution of the current is more closely approximated with low percentages of fill.

5. SUMMARY OF FORMULAS

The formulas developed in the second part of this paper and those for which the derivation has been indicated are summarized below. Conducting layers of copper, and insulating layers with a dielectric constant

---

[8] Loc. cit.

of ϵ and a thickness half that of the conducting layers ($w/t = 2$) are assumed throughout.

5.1 Clogston I Cable

$$R_{\mathrm{I}} = \frac{82080A}{D_2^2}\left[1 + \frac{B_1 w^2 D_2^2 F_{\mathrm{mc}}^2}{3710} + \cdots\right] \text{ ohms/mile}, \tag{38}$$

$$L_{\mathrm{in}} = 0.741 \times 10^{-3} \frac{D_2^2}{D_2^2 - d_2^2}\left[\frac{D_2^2}{D_2^2 - d_2^2}\log_{10}\frac{D_2}{d_2}\right.$$
$$\left. - 0.2172\right] \text{ henries/mile}, \tag{39}$$

$$L_{\mathrm{ex}} = 0.741 \times 10^{-3} \log_{10}\frac{d_2}{D_1} \text{ henries/mile}, \tag{40}$$

$$C_{\mathrm{ex}} = \frac{0.0388 \times 10^{-6}\epsilon_{\mathrm{I}}}{\log_{10}\dfrac{d_2}{D_1}} \text{ farads/mile}, \tag{41}$$

$$\epsilon_{\mathrm{I}} = \epsilon(1 + w/t), \tag{42}$$

$$Z_{\mathrm{I}} = \left[1 + \frac{L_{\mathrm{in}}}{L_{\mathrm{ex}}}\right]\sqrt{\frac{L_{\mathrm{ex}}}{C_{\mathrm{ex}}}} \text{ ohms}, \tag{43}$$

$$\beta_{\mathrm{I}} = \omega\sqrt{L_{\mathrm{ex}}C_{\mathrm{ex}}} \text{ radians/mile}, \tag{44}$$

$$\alpha_{\mathrm{I}} = \frac{K_2}{D_2^2} + K_3 w^2 F_{\mathrm{mc}}^2 \text{ nepers/mile}, \tag{45}$$

where:

$$K_2 = \frac{41040}{Z_{\mathrm{I}}}\left[\frac{D_2^2}{D_1^2 - d_1^2} + \frac{D_2^2}{D_2^2 - d_2^2}\right], \tag{46}$$

$$K_3 = \frac{22.124}{Z_{\mathrm{I}}}\frac{D_1 D_2 - d_1 d_2}{(D_1 + d_1)(D_2 + d_2)}. \tag{47}$$

5.2 Clogston II Cable

$$R_{\mathrm{II}} = \frac{328320}{D^2 - d^2}\left[1 + \frac{B_2 w^2 D^2 F_{\mathrm{mc}}^2}{3710} + \cdots\right] \text{ ohms/mile}, \tag{48}$$

$$Z_{\mathrm{II}} = \frac{17.35}{\sqrt{\epsilon}} \text{ ohms}, \tag{49}$$

$$\beta_{\mathrm{II}} = \omega\sqrt{L_{\mathrm{ex}}C_{\mathrm{ex}}} \text{ radians/mile}, \tag{50}$$

$$\alpha_{II} = \frac{K_2'}{D^2} + K_3' w^2 F_{mc}^2 \text{ nepers/mile,} \tag{51}$$

$$K_2' = \frac{164160}{Z_{II}} \frac{D^2}{D^2 - d^2}, \tag{52}$$

$$K_3' = \frac{44.248}{Z_{II}} \frac{D^2 + d^2 - (D + d)\sqrt{\frac{1}{2}(D^2 + d^2)}}{D^2 - d^2}, \tag{53}$$

$$A_1 = \frac{D_2^2}{D_1^2 - d_1^2} + \frac{D_2^2}{D_2^2 - d_2^2}, \tag{54}$$

$$B_1 = \frac{2(D_1 D_2 - d_1 d_2)(D_1 - d_1)(D_2 - d_2)}{(D_1^2 - d_1^2 + D_2^2 - d_2^2)D_2^2}, \tag{55}$$

$$B_2 = 1 + \frac{d^2}{D^2} - \left(1 + \frac{d}{D}\right)\sqrt{\frac{1}{2}\left(1 + \frac{d^2}{D^2}\right)}. \tag{56}$$

The parameters in (38) to (56) are defined as follows:

D_2 = Outside diameter of outer laminated conductor in a Clogston I cable, in mils.

d_2 = Inside diameter of outer laminated conductor in a Clogston I cable, in mils.

D_1 = Outside diameter of inner laminated conductor in a Clogston I cable, in mils.

d_1 = Inside diameter of inner laminated conductor in a Clogston I cable, in mils.

D = Outside diameter of a laminated Clogston II cable, in mils.

d = Inner diameter of a laminated Clogston II cable, in mils.
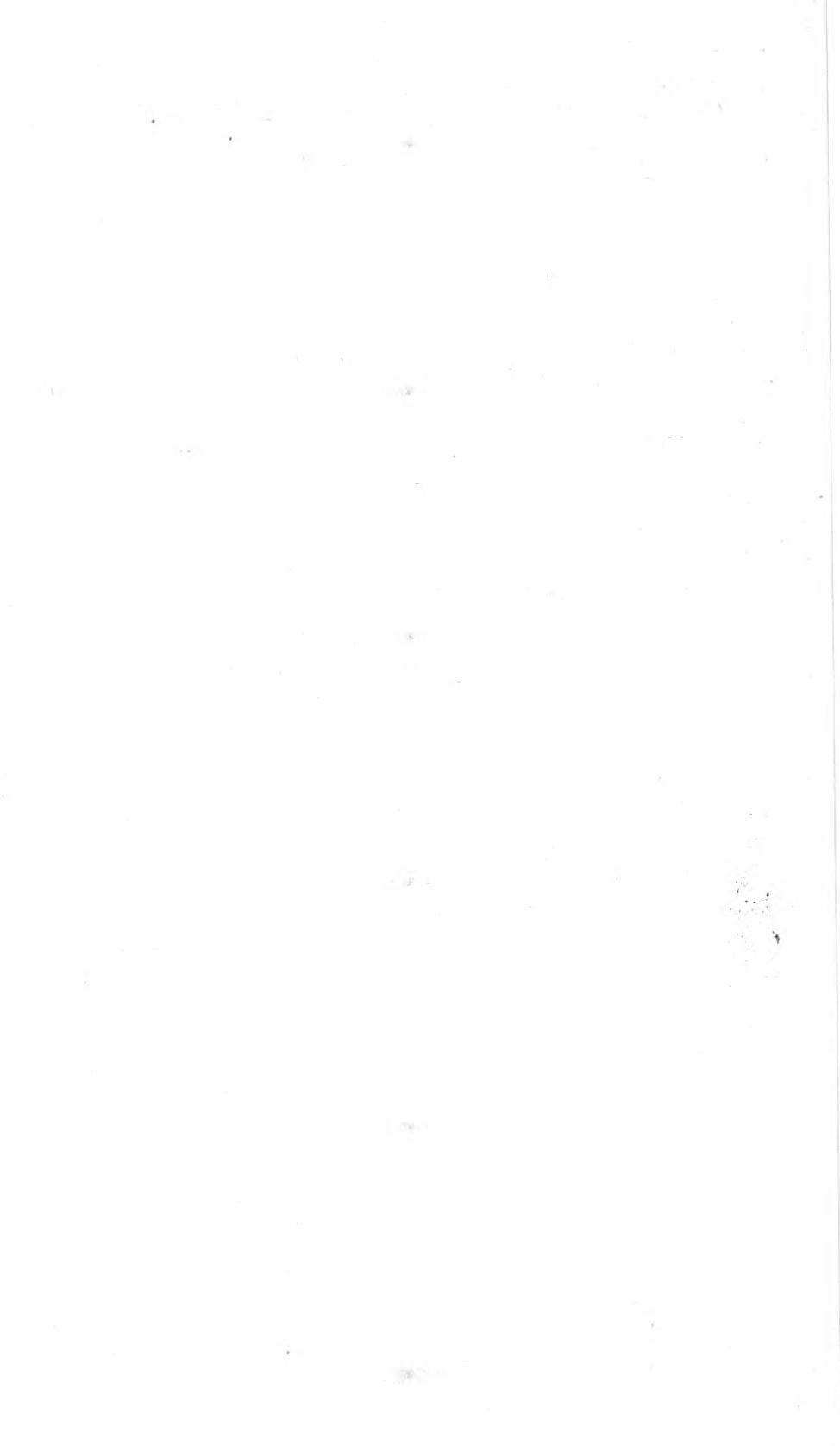
w = Thickness of copper layers in mils.

F_{mc} = Frequency in megacycles.

ϵ_I = Dielectric constant of insulation between inner and outer laminated conductors in a Clogston I cable.

ϵ = Dielectric constant of insulating layers.

ACKNOWLEDGMENT

A Coupled Resonator Reflex Klystron*

By E. D. REED

The theory of a coupled resonator reflex klystron is developed and its reduction to practice described. This tube differs from the conventional reflex klystron in that its performance characteristic is derived from the interaction between the electronic admittance due to a bunched electron stream and the input admittance of two synchronously tuned, coupled resonators. As a result: (1) power output can be made to be substantially flat over the greater part of the electronic tuning range; (2) the half-power electronic tuning range of the coupled resonator reflex klystron is more than twice that of a klystron using the same electron optical system but interacting with a single resonator; and (3) modulation linearity may be obtained over a greatly increased frequency swing.

A reduction in power output of about 3db occurs for a secondary resonator Q and coupling coefficient adjustment designed to yield a maximum flat band or maximum electronic tuning while a much smaller reduction in output power will provide a substantial improvement in modulation linearity.

TABLE OF CONTENTS

1.0 INTRODUCTION

The conventional reflex klystron derives its performance characteristics from the interaction between the electronic admittance due to a bunched electron stream and the input admittance of a resonant cavity. As is well known, this interaction results in mode shapes which are closely related to certain input properties of the passive resonant circuit. Thus, the dependence of power output upon frquency, which results from variations in repeller voltage about its mid-mode value, bears close resemblance to the input-impedance-versus-frequency plot of a parallel resonant circuit. Similarly, the curve relating frequency to repeller voltage has the same general shape as that relating frequency to the input phase angle of the resonator. Recognition of these relationships has resulted in the consideration of different and, perhaps, more useful mode shapes which might be obtained if the electronic admittance were made to interact with impedance or admittance functions of passive circuits other than that due to a single resonator.

What do we mean by "more useful" mode shapes? The answer, of course, depends on the application, although an "ideal" mode shape could probably be defined as one having a flat top, i.e., power output inde-

pendent of repeller voltage, with frequency linearly related to the latter. Moreover, these conditions should preferably obtain over the widest possible frequency range. A tube possessing such characteristics would prove exceedingly useful in a large number of applications. To list a few:

(a) Electronically swept signal generator,

(b) FM deviator,

(c) Transmitting oscillator in radio relay systems employing frequency modulation, and

(d) Local oscillator in microwave receivers with wide range AFC applied to the repeller.

Inability to realize this ideal mode shape in a practical tube might make a compromise solution appear acceptable, one consisting of a reflex klystron having a variable mode shape, i.e., a characteristic which could be adjusted to fit a particular need. Thus, application (a) requires constant power output with minor emphasis on frequency-repeller-voltage linearity, whereas applications (b) and (c) demand a high degree of modulation linearity with constancy of power output of no great importance. In application (d) the emphasis is on wide electronic tuning with both variation in power output and non-linearity in the frequency characteristic permissible.

A method to obtain this variable mode shape was achieved by the use of coupled cavities. Instead of having the bunched electron stream interact with the electric field of a single resonator, as is done in the conventional reflex klystron, we can present to it the input admittance of two coupled cavities. The resultant mode shape may then be expected to resemble the input impedance of two coupled resonant circuits just as the conventional klystron mode resembles that of a single resonator. Moreover, the input impedance of two coupled cavities can assume a large variety of contours depending on the ratio of Q's of the primary and secondary resonators and on the tightness of coupling between the two. If we were now to construct such a double cavity reflex klystron with provision to vary the secondary Q as well as the coupling coefficient continuously, we would have the means of producing a large variety of mode shapes within a single tube. Depending on the application, the characteristics could then be adjusted to give either a range of flat power output or optimum modulation linearity or wide electronic tuning.

As might be expected at this point, a price must be paid for the advantages gained in the coupled cavity approach. It consists of the power expended in supplying the losses due to the secondary resonator. This power subtracts directly from the available useful power and, there-

fore, though resulting in broadband operation, greatly improved modulation linearity and a number of other useful properties, leads to a definite reduction in output level. Whether this can be tolerated will again depend on the application. In many practical cases the performance flexibility inherent in the coupled resonator reflex klystron will more than outweigh this advantage.

It is the purpose of this paper to present the theory underlying the operation of the coupled resonator reflex klystron, as well as its experimental verification. For the sake of completeness, but also in order to emphasize methods of analysis to be used in later sections and not readily found in the literature, a review of reflex klystron theory will precede the exposition of the coupled resonator problem. In both the single and coupled resonator case, performance analysis will be based on a separate and independent study of the electronic and passive circuit admittance developed across the interaction gap and upon the graphical combination of the two in the complex admittance plane. As a by-product of this investigation, a number of driving point properties of two coupled resonant circuits will be developed which may be found of general networks interest. Following the theory of the coupled resonator reflex klystron, an experimental tube of this type will be described and a qualitative as well as quantitative verification of the theory given. Oscillograms will be presented showing the advantages of this device when used as a sweep generator, both in the microwave band and at lower frequencies. Additional applications will be indicated in the hope that others may try them.

2.0 SMALL SIGNAL REFLEX KLYSTRON THEORY

This section will be devoted to a brief review of the small signal reflex klystron theory. Emphasis will be on concepts leading to the equivalent circuit representation and to the graphical admittance-plane analysis. Both of these and particularly the graphical approach will later be used in the investigation of coupled cavity behavior.

As stated before, the operation of the reflex klystron is the result of the interaction between a bunched electron stream and the varying electric field existing inside a resonant cavity. In circuit language, this amounts to an interaction between an active and a passive element. The active one due to the electron stream is termed electronic admittance, and the passive one is the input admittance of the resonator. Derivations for the expression describing the electronic admittance may be readily found in the literature.[1] It will not be repeated here. The re-

sult of these derivations, however, and its physical significance will be discussed at some length.

2.1 *Electronic Admittance*

Consider an arrangement of four plane and parallel elements consisting of a cathode, two ideal grids and a reflector. Assume these electrodes to be of infinite extent so that all electrons will move in straight paths perpendicular to the planes of the electrodes. Let the current densities encountered be low enough so that space charge effects can be neglected. Both grids are operated at the same dc potential, V_0, positive with respect to the cathode. As shown in Fig. 1, this might be achieved by connecting them to the secondary winding of an ideal transformer having a 1:1 turns ratio. The reflector is operated at a dc potential, V_R, negative with respect to the cathode. Next, an RF voltage of amplitude V is applied to the transformer and, hence, appears across the grids. Electrons emitted from the cathode are accelerated toward the first grid and arrive at it
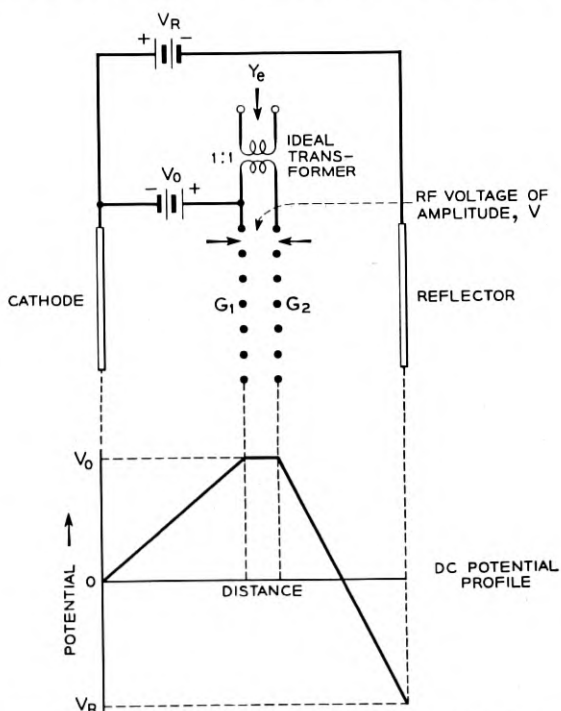

Fig. 1 — Electrode arrangement and dc potential profile giving rise to electronic admittance, Y_e, described by equation (2.1).

with a velocity corresponding to V_0 volts. During their traversal of the interaction gap, i.e., the space bounded by the two grids, this velocity is changed by an amount depending on the instantaneous direction and intensity of the electric field and on the fraction of a cycle spent in traversing the gap. They then enter the retarding dc field of the repeller drift space. Here, they are slowed down, brought to a standstill and returned to the interaction gap for a second transit. During this process, the velocity modulation acquired during the first transit is converted to intensity modulation such that the convection current returning through the gap does so in the form of sharp and well defined pulses with a repetition rate equal to the frequency of the applied RF voltage. This pulsed current may now be harmonically analyzed and its fundamental component evaluated. The ratio of this fundamental component of current to the applied RF voltage is termed electronic admittance, Y_e. Providing the amplitude, V, of the applied RF voltage is small compared to the dc accelerating voltage, V_0, the value of Y_e is given by:

$$Y_e = \frac{I_0 \beta^2 \theta}{2V_0} \frac{2J_1\left(\frac{\beta V \theta}{2V_0}\right)}{\left(\frac{\beta V \theta}{2V_0}\right)} e^{j[(\pi/2)-\theta]}, \tag{2.1}$$

where I_0 = dc beam current.

β = beam coupling coefficient.*

θ = round trip transit time in repeller drift space, in radians.

V_0 = dc beam voltage.

V = amplitude of RF gap voltage.

J_1 = Bessel function operator.

Of all these parameters affecting Y_e we shall focus our attention on two, namely V and θ. The other parameters depend on such factors as tube geometry, electron gun perveance etc. and, within the scope of this investigation, will be considered constant.

Referring to equation (2.1), it is seen that the phase angle associated with Y_e is a function of θ only, whereas the amplitude of Y_e depends on both θ and the RF gap voltage, V. Suppose now θ is held constant while

---

* β, also referred to as modulation coefficient, is given by

$$\beta = \frac{\sin \frac{\theta_g}{2}}{\frac{\theta_g}{2}}$$

where θ_g, the transit angle in the interaction gap, is expressed in radians.

V is increased from zero to some finite value, a process actually occurring in a reflex klystron during the build-up of oscillations. The amplitude of Y_e will then be of the form $2J_1(x)/x$ and will decrease according to the Bessel function plot of Fig. 2.

For an infinitesimal or zero RF gap voltage the electronic admittance, now referred to as "small-signal" electronic admittance, Y_{es}, may be derived from equation (2.1) as,

$$Y_e\big|_{V=0} \equiv Y_{es} = \frac{I_0\beta^2}{2V_0}\,\theta e^{j[(\pi/2)-\theta]}. \tag{2.2}$$

When presented in the complex admittance (i.e., g-b) plane, expression (2.2) assumes the form of a geometric spiral as shown in Fig. 4. Each point on the spiral corresponds to a particular value of θ and, since θ is a function of the repeller voltage only, to a particular value of V_R. We also note that for some values of θ the conductance component of Y_{es} is negative while for others it is positive. Thus, there is the possibility of generation of RF energy for values of Y_{es} adjusted by means of the repeller voltage to fall on the left-hand half of the admittance spiral while energy is absorbed for values of Y_{es} having a positive conductance component. As the RF gap voltage, V, builds up from zero to its final value the magnitude of the electronic admittance shrinks along a radius
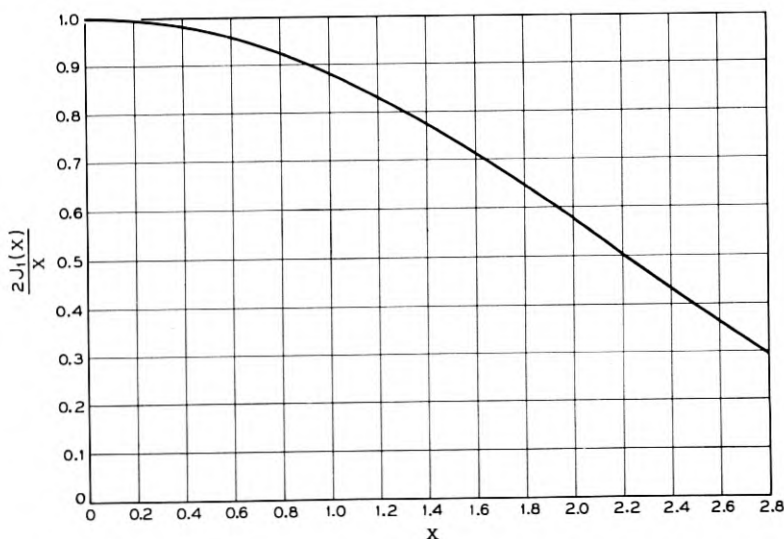

Fig. 2. Bessel function plot showing the relative variation of the amplitude of the electronic admittance (ordinate) as a function of RF gap voltage (abscissa) with repeller drift angle held constant.

vector from Y_{es} to Y_e according to the Bessel function plot of Fig. 2. For $\theta = 2\pi(N + \frac{3}{4})$ radians where $N = 0, 1, 2, 3 \ldots$ the electronic admittance becomes a pure negative conductance as may be seen by putting this relation into equations (2.1) or (2.2).

Summarizing, we have seen that the presence of an electron stream bunched in accordance with the arrangement of Fig. 1 gives rise to an admittance appearing across the grids bounding the interaction space. The phase angle of this admittance is a function only of the repeller drift angle, θ, while its magnitude depends on both θ and the RF gap voltage, V. For values of θ in the vicinity of $2\pi(N + \frac{3}{4})$ radians the conductance component is negative, a necessary condition for the production of sustained oscillations.

2.2 *Passive Circuit Admittance of Single Resonator*

Any resonant cavity may be represented by a simple parallel G-C-L-combination provided the desired resonance is sufficiently far removed from adjacent ones. In some cases such as in cylindrical or waveguide cavities, to name two, it is difficult to ascribe physical significance to the lumped elements appearing in the equivalent circuit representation. This is not so in the case of a conventional reflex klystron cavity. Since the latter always consists of a re-entrant type resonator, most of the electric field is concentrated in the interaction gap, i.e., the narrow region traversed by the outgoing and returning electrons and bounded by two parallel grids, while the major portion of the magnetic flux resides in the outer cylindrical section. Thus, the effective shunt capacitance appearing in the equivalent circuit is associated primarily with the above grid planes, a minor contribution originating in the fringing field close to the re-entrant post and the residual electric field in the outer cylindrical part of the cavity.

The input admittance, Y, of a high Q resonator when represented by C, L, and G connected in parallel is given by,

$$Y = G(1 + j2Q\delta) \tag{2.3}$$

$$= G + j2C\Delta\omega, \tag{2.4}$$

where $Q = \omega_0 C/G$ and $\delta = \Delta\omega/\omega_0 = \Delta f/f_0 = (f - f_0)/f_0$, and G represents all internal resonator losses plus the external load referred to the gap. Plotted in the complex admittance plane, the locus of the admittance vector with varying frequency is a straight line parallel to the imaginary axis and spaced a distance corresponding to G to its right. The variation in susceptance is directly proportional to the frequency deviation

from resonance. Also, the frequency deviation required to bring about a given change in susceptance is inversely proportional to both C and Q.

2.3 *Equivalent Circuit of Single Resonator Reflex Klystron*

We saw that the presence of a bunched electron stream between two closely spaced grid planes gives rise to an electronic admittance the properties of which were discussed earlier. Suppose we now let these grids become part of a re-entrant cavity so that they form the boundaries of the interaction space. This will make them elements of significance and common to both the electron stream and the passive circuit admittance. As far as the electron stream is concerned the grids become the terminals across which the electronic admittance, Y_e, is developed and in relation to the resonant circuit they constitute the major portion of the effective
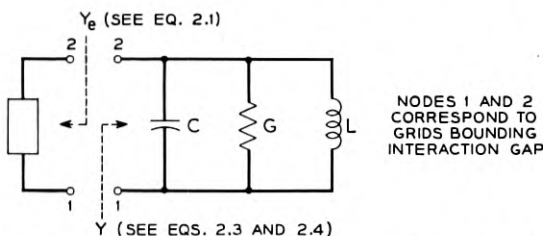
Fig. 3. — Equivalent circuit of single resonator reflex klystron.

shunt capacitance. Based on these remarks, it should be apparent that a single-resonator reflex klystron may be represented by an equivalent circuit consisting of a parallel combination of Y_e, C, G, and L as shown in Fig. 3. Furthermore, the expressions for Y_e and Y as given by equations 2.1 and 2.4, respectively, show that these two quantities are independent of each other. The electronic admittance, Y_e, is a function of the electron optics of the tube, while the passive circuit admittance, Y, is a function purely of cavity parameters, including the tightness of coupling to the external load.

2.4 *Performance Analysis Based on Complex Admittance Plane Representation*

The condition for oscillation applying to the equivalent reflex klystron circuit requires the total admittance across nodes 1 and 2 of Fig. 3 to equal zero, i.e.,

$$Y_e + Y = 0, \qquad\qquad 2.5a$$

or

$$Y_e = -Y. \qquad\qquad 2.5b$$

Attainment of this condition is preceded by a period during which the RF gap voltage increases from its initial, near-zero value to its final steady state amplitude. Concurrently, the net conductance across nodes 1 and 2 changes from its maximum negative value to zero and the electronic admittance vector shrinks from Y_{es} to Y_e.

This process of build-up of oscillations and the final steady state may be conveniently studied by the graphical representation of Fig. 4. Here, the negative of the passive circuit admittance, Y, has been superimposed on the small-signal electronic-admittance spiral. Since the net conductance across the interaction gap (nodes 1 and 2 in Fig. 3) must be negative in order for oscillations to build up, we see at once that the $(-Y)$ plot, sometimes also referred to as "load line," divides the complex admittance plane into two regions: the one to its left, in which the
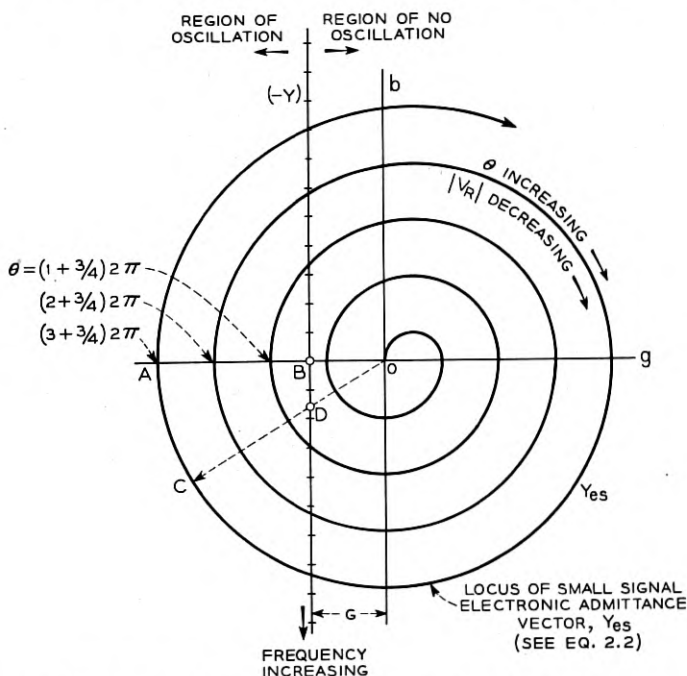

Fig. 4 — Complex admittance plane representation showing the superposition of the negative of the circuit input admittance upon the small signal electronic admittance spiral. Each point on spiral corresponds to a particular value of repeller transit angle, θ, and hence repeller voltage, V_R, and each division on the $(-Y)$ locus to an equal frequency increment.

negative electronic conductance exceeds the circuit conductance and where the build-up of oscillations is possible, and the region to its right where the condition for the build-up of oscillations is not met. Thus, for the load line in the position indicated, the tube will not oscillate in the $N = 0$ mode, but will do so in the $N = 1, 2, 3$ and higher order modes.

Still referring to Fig. 4, suppose θ has been adjusted by means of the repeller voltage to $\theta_1 = (3 + \frac{3}{4})2\pi$ radians, i.e., the center of the $3 + \frac{3}{4}$ mode. The small signal-electronic-admittance vector will then be a pure negative conductance terminating on the spiral at point A and together with the passive circuit conductance yielding a net negative conductance across the grids of value $(\overline{OA}\text{-}\overline{OB})$. Oscillations, therefore, will build up until the equilibrium condition, $Y_e = -Y$, has been satisfied. In terms of the admittance plane representation, the electronic-admittance vector will shrink without change in its phase until it terminates on the load line at point B. Since the electronic admittance for the particular value of θ chosen is a pure conductance, oscillations will occur at the resonant frequency of the cavity, f_0, which is also the frequency corresponding to the intersection of the electronic-admittance vector and the load line.

From equation 2.1 we see that,

$$\frac{\overline{OB}}{\overline{OA}} = \frac{\dfrac{2J_1\left(\dfrac{\beta V_1 \theta_1}{2V_0}\right)}{\left(\dfrac{\beta V_1 \theta_1}{2V_0}\right)}}{\left[\dfrac{2J_1\left(\dfrac{\beta V \theta}{2V_0}\right)}{\left(\dfrac{\beta V \theta}{2V_0}\right)}\right]_{V=0}} = \frac{2J_1\left(\dfrac{\beta V_1 \theta_1}{2V_0}\right)}{\left(\dfrac{\beta V_1 \theta_1}{2V_0}\right)}, \qquad (2.6a)$$

where V_1 is the steady state RF gap voltage corresponding to drift angle, θ_1. Entering the graph of Fig. 2 with $\overline{OB}/\overline{OA}$ as the ordinate, we can read off the corresponding value of $(\beta/2V_0)V_1\theta_1$. For a particular tube structure and operating conditions, $\beta/2V_0$ will be a fixed constant so that, in effect, we have determined the value of a quantity proportional to the product of $V_1\theta_1$.

Next, consider the case where θ has been changed from $\theta_1 = (3 + \frac{3}{4})2\pi$ to $\theta_2 = K(3 + \frac{3}{4})2\pi$ radians. The electronic admittance vector for $V = O$ will now terminate on the Y_{es}-plot at C. Again a build-up of oscillations will ensue and upon attainment of the steady state condition the vector will have shrunk to the value $\overline{OD}$, with the

frequency of oscillation determined by the location of point D on the load line. By an argument similar to the one used before, we have

$$\frac{\overline{OD}}{\overline{OC}} = \frac{2J_1\left(\dfrac{\beta V_2 \theta_2}{2V_0}\right)}{\left(\dfrac{\beta V_2 \theta_2}{2V_0}\right)} \tag{2.6b}$$

and once more, the plot of Fig. 2 will yield the value of

$$\left(\frac{\beta}{2V_0}\right)V_2\theta_2 \qquad \text{or} \qquad \left(\frac{\beta}{2V_0}\right)KV_2\theta_1 .$$

It is apparent that $\beta/2V_0$ is a factor common to all these determinations and may be neglected if we are only interested in the mode shape, i.e. the *relative* variation of gap voltage or output power with the repeller drift angle, θ. For modes higher than $N = 2$, the electronic admittance spiral approximates a number of semi-circles with their centers close to the origin. Hence

$$\overline{OA} \approx \overline{OC} \qquad \text{and} \qquad \frac{\overline{OD}}{\overline{OC}} > \frac{\overline{OB}}{\overline{OA}} .$$

Figure 2, then indicates that

$$V_1\theta_1 > V_2\theta_2 \qquad \text{or} \qquad V_1\theta_1 > KV_2\theta_1$$

and since K does not vary greatly from unity $V_1 > V_2$. This is as it should be since we know that a change in repeller voltage from its mid-mode value causes a decrease in power output and, hence, in gap voltage. The latter will decrease to zero when θ has been adjusted to a value such that the Y_{es} vector terminates at the intersections of the electronic admittance spiral and the load line.

Another result which becomes apparent from an inspection of Fig. 4 is this. For the condition of stable oscillation the phase angle of the electronic admittance equals that of the passive circuit except for an additive constant of 180 degrees which, however, may be disregarded in this argument. The frequency of oscillation may be determined from the input-phase-angle vs. frequency plot of the passive circuit by looking up the frequency corresponding to the particular value of phase angle to which the electronic admittance has been adjusted (by means of repeller voltage). Thus, the curve relating repeller-drift-angle to frequency is identical with the plot of input-phase-angle vs. frequency for the passive circuit. Moreover, if repeller voltage is linearly related

to the repeller-drift-angle, a condition which in most practical cases holds over a restricted repeller voltage swing about its midmode value, then the repeller voltage-frequency plot will have the same shape as the input phase angle-frequency curve of the passive circuit, differing from the latter only by a constant multiplying factor. We therefore conclude that the modulation performance of the reflex klystron, at least over the central portion of the mode, may be predicted from an examination of the driving point properties of the passive circuit.

The above method of graphical analysis is quite useful in the case of conventional reflex klystrons, although the same results may be obtained analytically by making use of the equation for electronic admittance in conjunction with that of the input admittance of a single resonator. In the case of coupled resonators the expression for input admittance becomes much more involved, as we shall see later, with the result that the graphical approach outlined above was found by far the quicker and more practical method of solution.

3.0 THEORY OF COUPLED RESONATOR REFLEX KLYSTRON

It has been shown that the performance of a reflex klystron can be analyzed by considering the electronic and passive circuit admittances separately and then combining the two graphically in the complex admittance plane. The same procedure can be adopted in the determination of mode shapes resulting from the interaction of the electronic admittance with any arbitrary circuit admittance which can be realized across the gap. Conversely, we can determine the admittance or impedance function required to produce a particular desired mode shape. In other words: given an admittance function, we can determine the resulting mode shape, and given a desired mode shape, we can predict the required admittance function. As an example, consider a mode having a flat top and vertical sides as shown in Fig. 5(a). This would be the ideal shape for a reflex oscillator to be used as an electronically swept signal source. To achieve this mode shape, we must realize an admittance function across the bunching grids yielding a constant excess of negative electronic admittance over a range of repeller voltages. Such a function is shown in the complex admittance plane along with the plot of Y_{es} in Fig. 5(b) and the corresponding input impedance in 5(c). It will result in a frequency range of constant RF gap voltage which in turn will give rise to a range of flat power provided it is developed across a constant, frequency-invariant conductance. In order to clarify this rather important consideration, let us write the admittance appearing
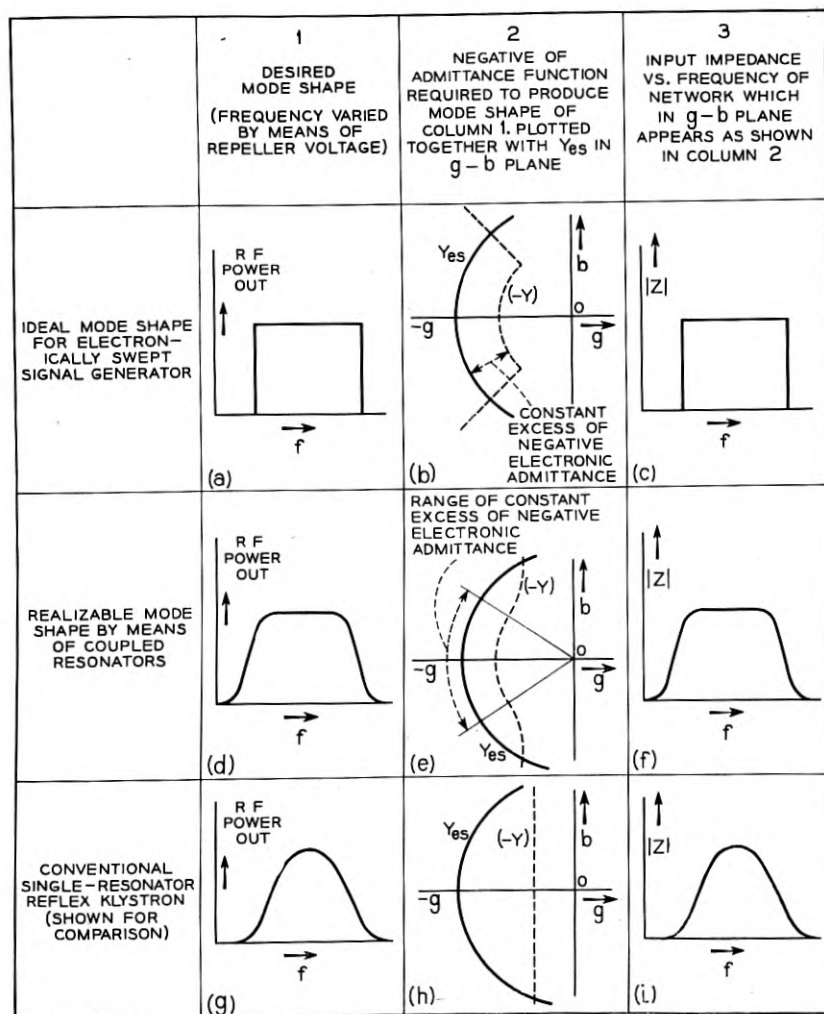
Fig. 5 — Desired mode shapes and passive circuit impedance functions required to produce them.

across the interaction gap as,

$$Y = G_T(f) + jB_T(f),$$

where $G_T(f)$ and $B_T(f)$ denote the total conductance and total susceptance both of which are functions of frequency. Over the region of constant RF gap voltage we require $| Y |$ or

$$\sqrt{G_T^2(f) + B_T^2(f)}$$

to be constant. This, however, means that the power generated (as distinct from the power delivered to the load) is not constant since it is given by,

Generated Power $= \frac{1}{2}$(RF gap voltage)^2 (Total Conductance)

$$= \frac{1}{2}V^2 G_T(f)$$

and G_T varies with frequency. Now, the term $G_T(f)$ is the sum of a number of conductances, all but one of which do not change with frequency. The frequency-invariant conductances represent the different power losses in the primary resonator plus the external load referred to the gap while the remaining frequency-sensitive conductance is due to the coupled circuit used in shaping the admittance locus. Hence the useful power output is proportional to V^2 and therefore constant if V^2 is constant, whereas the variation with frequency of the total conductance must be taken into account in evaluating generated power.

The foregoing discussion together with the illustrations of Fig. 5 should make it clear that in order to maintain a constant RF power output level, over a specified frequency range, we must have the electronic admittance interact with a circuit the input impedance of which, when referred to the gap, is also constant over the same frequency range. A circuit having such characteristics can be obtained by the use of coupled resonators as will be shown later.

Suppose the emphasis is on modulation linearity rather than flatness of power output. Attention, then, must be focused on the relation between input phase angle and frequency of the passive circuit. Here, again, we shall see that the application of coupled resonators offers advantages beyond what is possible with a single cavity.

3.1 Driving Point Properties of Two Coupled Resonators Having Equal Q's

Using the equivalent shunt representation as shown in Fig. 6, the exact expression for the input admittance, Y, of two coupled resonators
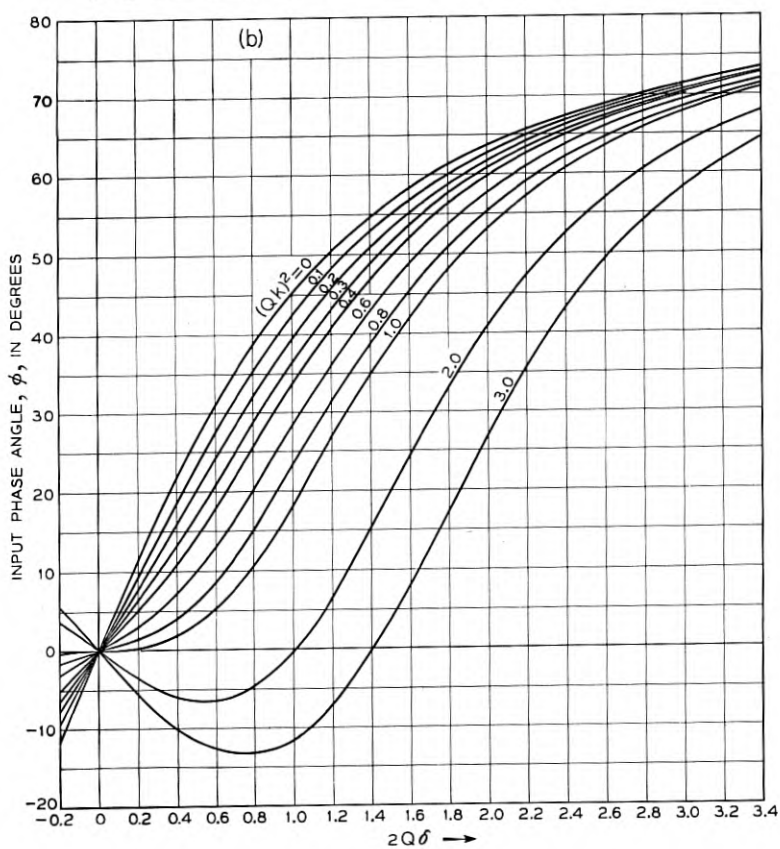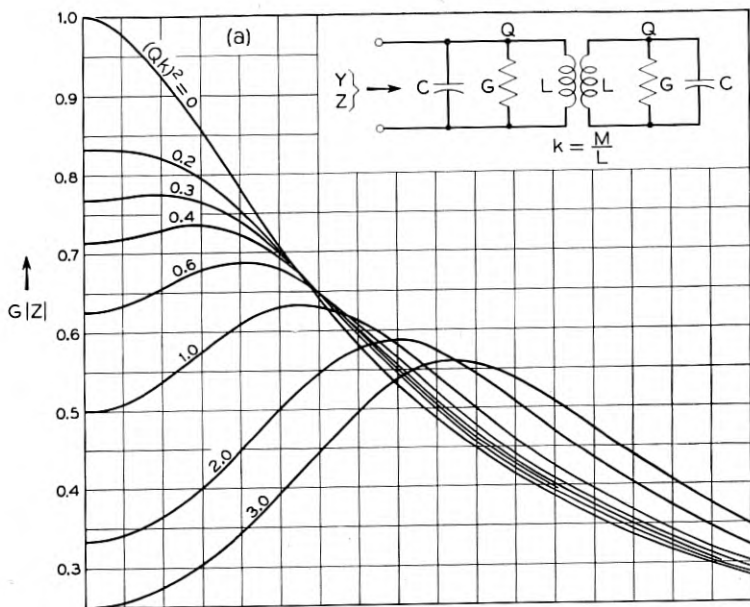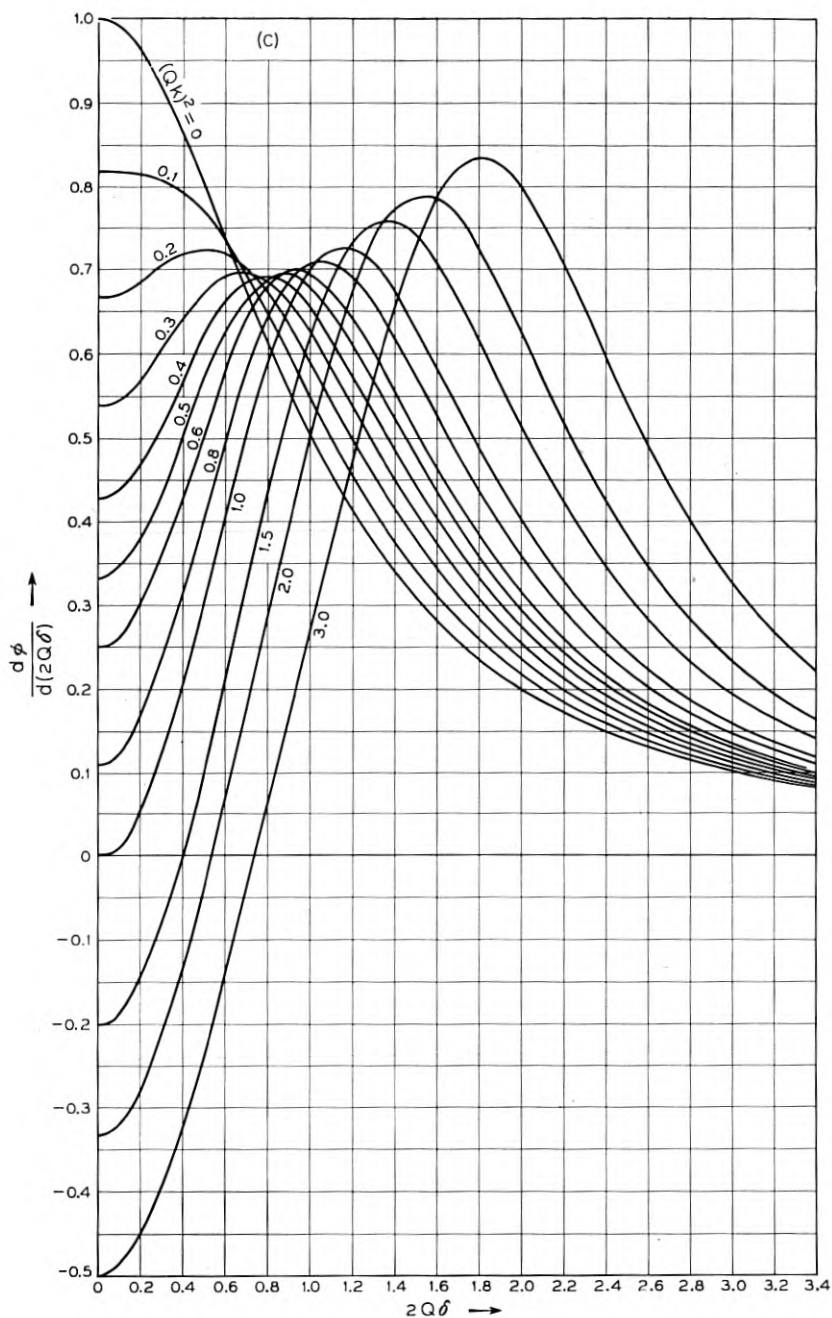
Fig. 6 — Normalized driving point properties of two coupled, synchronously tuned resonators having equal Q's. (a) Input impedance, $G|Z|$, versus frequency deviation from resonance, $2Q\delta$, for various degrees of coupling, Qk. (b) Input phase angle, ϕ, versus frequency deviation. (c) Rate of change of ϕ with $2Q\delta$ versus $2Q\delta$.

having equal Q's and equal resonant frequencies is given by,

$$Y = G(g + jb),$$

where

$$b = \frac{Q(\Omega^2 m - 1)}{\Omega m}\left(1 - \frac{\left(\dfrac{Qk}{\Omega m}\right)^2}{1 + \left[\dfrac{Q(\Omega^2 m - 1)}{\Omega m}\right]^2}\right),$$

$$g = 1 + \frac{\left(\dfrac{Qk}{\Omega m}\right)^2}{1 + \left[\dfrac{Q(\Omega^2 m - 1)}{\Omega m}\right]^2}.$$

$$\tag{3.1}$$

G = shunt conductance of primary and secondary resonators.

g, b = total input conductance and susceptance respectively normalized with respect to G,

Ω = normalized frequency = f/f_0.

f_0 = resonant frequency of both primary and secondary resonators.

k = coefficient of coupling.

$m = 1 - k^2$.

This expression for input admittance, though accurate, is rather unwieldy because of the many variables involved. A few obvious approximations, however, will change equation (3.1) into a much simpler and more meaningful expression, yet sufficiently accurate for the range of Q's and bandwidths of interest here. Let,

$$\Omega = \frac{f}{f_0} = \frac{f_0 + \Delta f}{f_0} = 1 + \frac{\Delta f}{f_0} = 1 + \delta,$$

where δ denotes the normalized frequency deviation from resonance, $\Delta f/f_0$. If we further assume that $k^2 \ll 1$ and the range of Q's is such that $(Qk)^2$ may vary from zero to about five and that the maximum value of δ is small enough so that its higher powers may be neglected, then equation (3.1) may be simplified to,

$$\frac{Y}{G} = \left[1 + \frac{(Qk)^2}{1 + (2Q\delta)^2}\right] + j2Q\delta\left[1 - \frac{(Qk)^2}{1 + (2Q\delta)^2}\right]. \tag{3.2}$$

The above equation* essentially contains three variables:

(a) The dependent variable, Y/G, i.e. input admittance normalized with respect to the shunt conductance,

(b) the independent variable, $(2Q\delta)$, which is a factor proportional to the frequency deviation from resonance, and

(c) parameter, (Qk), a measure of the tightness of coupling between the two cavities.

Compared with the input admittance for a single resonator which was given earlier, [equation (2.3),] as, $Y/G = 1 + j2Q\delta$, it is seen that the conductance component has been changed by a factor

$$\left[1 + \frac{(Qk)^2}{1 + (2Q\delta)^2} \right]$$

and the susceptance by

$$\left[1 - \frac{(Qk)^2}{1 + (2Q\delta)^2} \right].$$

Also, by setting $k = 0$, i.e., completely decoupling the secondary resonator, equation (3.2) reduces to equation (2.3) as, indeed, it should.

The information contained in equation (3.2) may be presented graphically in a number of ways. We can plot the magnitude of the normalized input impedance, $G|Z|$, as a function of $2Q\delta$ with $(Qk)^2$ as parameter as shown in Fig. 6(a).† Or we can plot the input admittance given by equation (3.2) directly in the g-b plane as in Fig. 7. Finally, we may show the variation of input phase angle with normalized frequency for different degrees of coupling as in Fig. 6(b). Each of these graphical representations has an important bearing on the performance of the coupled resonator klystron. Thus, the curves of Fig. 6(a) will have the same general shape as the RF power output vs. frequency plot of the reflex oscillator, the family of curves of Fig. 7, when superimposed on the electronic admittance spiral, can be used for a detailed analysis of mode shapes, and Fig. 6(b) determines the modulation characteristics obtainable with coupled resonators.

Having briefly touched upon the significance of the families of curves given in Figs. 6(a), 7, and 6(b), we can now proceed to discuss them in greater detail and to establish further valuable results.

---

* To check the accuracy of this equation, the curve for $(Qk)^2 = 1$ in Fig. 6(a) was replotted using the full expression given by equation 3.1 and assuming $Q = 100$. The agreement was found to be essentially perfect as far as its use in this investigation is concerned.

† For a similar presentation of the *transfer* characteristics of coupled tuned circuits see reference (6).
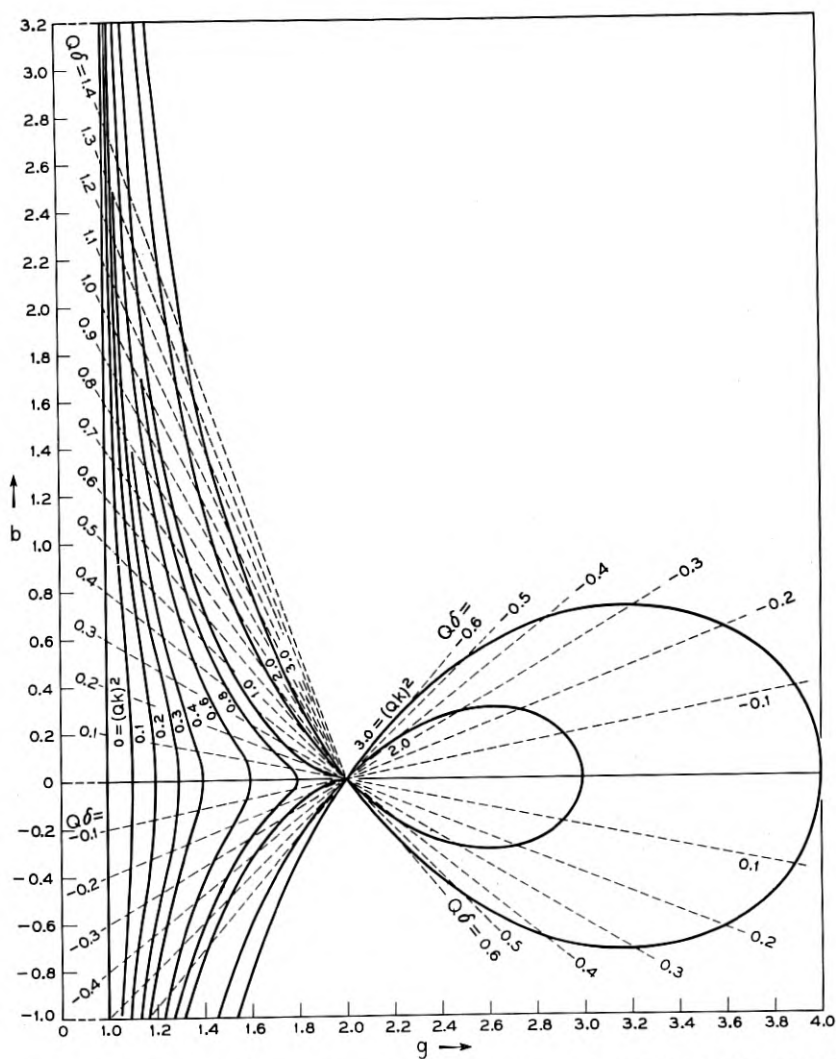
Fig. 7 — Input admittance of circuit shown in Fig. 6 plotted in g–b plane. Solid lines are loci of input admittance vectors, dashed lines connect points of equal frequencies.

3.1.1 *Variation of Input Impedance with Frequency, Fig. 6(a)*. Equation 3.2 may be written as

$$\frac{Y}{G} = g + jb,$$ (3.3)

where

$$g = \left[1 + \frac{(Qk)^2}{1 + (2Q\delta)^2} \right]$$ (3.4)

and

$$b = 2Q\delta \left[1 - \frac{(Qk)^2}{1 + (2Q\delta)^2} \right].$$ (3.5)

Hence

$$\frac{|Y|}{G} = \sqrt{g^2 + b^2}$$

and

$$G|Z| = \frac{1}{\sqrt{g^2 + b^2}}.$$ (3.6)

It is seen that g^2 and b^2 involve even powers of $(2Q\delta)$ only, so that $G|Z|$ is an even function, i.e., symmetrical about the vertical axis. For this reason, one-half the normalized impedance plot only has been given in Fig. 6(a). Inspection of this figure reveals that the effect of coupling to a secondary resonator is to broaden the frequency range over which a high impedance level can be maintained across the interaction gap. Comparing the variation of input impedance with frequency for $(Qk)^2 = 0$ (i.e., single cavity) with that for $(Qk)^2 = 0.3$, we see that for a frequency deviation of $2Q\delta = \pm 0.6$ the former shows a drop of 14 per cent while the latter only varies by ± 0.58 per cent. In terms of two coupled resonators having $Q = 100$ and operating at 4000 mc this means a variation in impedance of only ± 0.58 per cent or $\pm .052$ db over a frequency range of 24 mc.

Another result clearly brought out by the family of curves of Fig. 6(a) is that the process of broadbanding by coupling to a second resonator results in a reduction in absolute impedance level. This reduction in impedance level at midband is related to the tightness of coupling, $(Qk)^2$, by the expression,

$$[G|Z|]_{\delta=0} = \frac{1}{1 + (Qk)^2},$$ (3.7)

which may be readily derived from eq. 3.6. Note that for $(Qk)^2 = 1$, the impedance level at midband has dropped to half the value obtained with a single resonator or $(Qk)^2 = 0$. The degree of coupling corresponding to $(Qk)^2 = 1$ is noteworthy for reasons other than the one just mentioned. They will be discussed in the following section.

3.1.2 *Input Admittance Plot in g-b Plane, Fig. 7.* The graphical representation of input admittance in the complex admittance plane is of particular usefulness in the analysis of the coupled resonator reflex klystron. For the moment, however, we shall restrict the discussion to a consideration of the passive circuit only.

Each solid line in Fig. 7 is the locus of the admittance vector for a particular tightness of coupling. For $(Qk)^2 = 0$ the locus is a straight line parallel to the susceptance axis as described earlier for the case of a single resonator. For $(Qk)^2 = 0.3$ the shape of the locus approximates a circle with its center at the origin; this, of course, being true only over a restricted frequency range. As the coupling to the secondary resonator is progressively tightened, the locus is seen to bulge out in the direction of increasing conductance until, for $(Qk)^2 = 1$, it forms a cusp. This condition will henceforth be referred to as "critical coupling."* Coupling even tighter causes the formation of loops of increasing size.

For the overcoupled case, $(Qk)^2 > 1$, the admittance-vector locus crosses the conductance axis three times, with the first and third crossings coincident and independent of $(Qk)^2$ and the second crossover a function of $(Qk)^2$. The location of these intersections with the g-axis may be determined by equating the susceptance to zero, i.e., from equation (3.5),

$$2Q\delta \left[1 - \frac{(Qk)^2}{1 + (2Q\delta)^2}\right] = 0.$$

Hence the crossing to the extreme right occurs for $2Q\delta = 0$ while the first and third interesections correspond to

$$2Q\delta = \pm\sqrt{(Qk)^2 - 1}.$$

To determine the size of the loop we substitute $2Q\delta = 0$ into the expression for the conductance, i.e., equation (3.4) and obtain,

$$g\big|_{2Q\delta=0} = 1 + (Qk)^2,$$

---

* It should be noted that the term "critical coupling" as applied to the *transfer* characteristics of coupled tuned circuits, though also occuring for $(Qk)^2 = 1$, assumes a different significance in that it describes the condition of maximum flatness in response and optimum phase linearity. (See reference 6.) In the case of the coupled resonator reflex klystron, "critical coupling" forms the transition between stable performance and load hysteresis as will be shown later.

showing that the size of the loop is a sensitive function of the degree of coupling. The conductance value for the first and third crossover points is obtained by substituting

$$2Q\delta = \pm\sqrt{(Qk)^2 - 1}$$

into equation (3.4) and results in $g = 2$, i.e., a value of g independent of the coupling coefficient.

The dashed lines shown in Fig. 7 connect points of equal frequencies. It is of interest to note that these loci are straight lines crossing the conductance axis at $g = 2$. To prove this, eliminate $(Qk)^2$ between equations (3.4) and (3.5). This yields,

$$g - 1 = 1 - \frac{b}{2Q\delta},$$

whence

$$b = (-2Q\delta)(g - 2). \tag{3.8}$$

For a particular and constant value of $2Q\delta$, equation (3.8) describes a straight line of slope equal to $(-2Q\delta)$ intersecting the g-axis at $g = 2$.

3.1.3 *Variation of Input Phase Angle with Frequency, Fig. 6(b)*. The last driving point property of interest to this study is the dependence upon frequency of the input phase angle, ϕ. Referring to equation (3.2), this quantity is obtained as,

$$\phi = \tan^{-1}\left[2Q\delta \frac{1 + (2Q\delta)^2 - (Qk)^2}{1 + (2Q\delta)^2 + (Qk)^2}\right]. \tag{3.9}$$

The graphical representation of this function is given in Fig. 6(b). It shows the gradual transition from a simple S-shaped curve for $(Qk)^2 = 0$, having its only point of inflection at the origin, to the type of curve corresponding to $(Qk)^2 > 1$ which intersects the frequency axis three times. The special case of $(Qk)^2 = 1$, considered earlier and found to result in the formation of a cusp in the complex admittance plane, now gives rise to a plot of input phase which is tangent to the horizontal axis at the origin.

To investigate the condition for greatest linearity between phase angle and frequency, which, when applied to the coupled resonator reflex klystron, would be the condition for optimum modulation linearity, one could simply apply a straight edge to the curves of Fig. 6(b) and pick the best value of $(Qk)^2$ in this manner. A much more sensitive criterion of linearity, however, is the variation with frequency of the slope of these curves. An analytical expression for this slope has been

derived from equation (3.9) as,

$$\frac{d\phi}{d(2Q\delta)} = \frac{[1 - (Qk)^4] + (2Q\delta)^2[2 + 4(Qk)^2] + (2Q\delta)^4}{[1 + (Qk)^2]^2 + (2Q\delta)^2[3 + (Qk)^4] + (2Q\delta)^4[3 - 2(Qk)^2] + (2Q\delta)^6}.$$

$$(3.10)$$

The above expression, involving even powers of $2Q\delta$ only, results in a family of curves symmetrical about the vertical axis, the positive half of which is shown in Fig. 6(c). From it the value of $(Qk)^2$ for greatest linearity or most constant slope is seen to lie somewhere between 0.1 and 0.2. This is further borne out by Fig. 8, in which this region has been more fully explored. Fig. 8 constitutes a plot quite similar to that
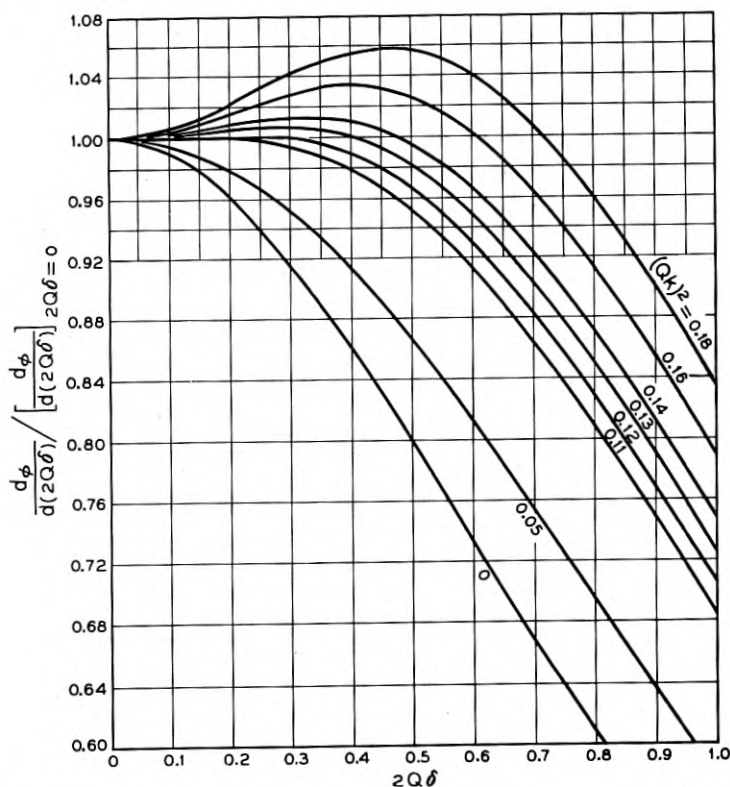
Fig. 8 — Rate of change of input phase angle with frequency normalized with respect to its center value for the circuit shown in Fig. 6. This quantity, the constancy of which is a measure of phase angle versus frequency linearity, is seen to stay absolutely constant over a frequency range corresponding to $2Q\delta = \pm0.3$ for $(Qk)^2 = 0.12$ while the single resonator case, i.e. $(Qk)^2 = 0$, shows a change of 8% over the same frequency range.

of Fig. 6(c) except that the ordinate now represents the instantaneous slope normalized with respect to its midmode value. The curves shown are for values of $(Qk)^2$ ranging from zero to $(Qk)^2 = 0.18$. Let us, for example, examine the curve corresponding to $(Qk)^2 = 0.12$; the slope is seen to stay absolutely constant up to a frequency deviation of $2Q\delta = \pm 0.3$, while the plot for the single resonator, i.e., $(Qk)^2 = 0$, included in this figure for comparison, changes by 8.4 per cent over the same frequency range.

Putting this in another way, suppose the maximum allowable deviation in slope from its mid-band value is one per cent. Fig. 8 then indicates that the permissible frequency deviation for coupled resonators having a value of $(Qk)^2 = 0.135$ is given by $2Q\delta = \pm 0.5$, whereas it must be restricted to one-fifth this value, i.e., $2Q\delta = \pm 0.1$, for the single resonator case. In terms of a midband frequency of 4000 mc and $Q = 100$, the permissible frequency excursions would be ± 10 mc and ± 2mc, respectively.

It is to be noted that the value of coupling coefficient resulting in greatest modulation linearity is considerably smaller than the value of coupling coefficient found to yield a constant impedance level.

3.2 Mode Shapes Resulting from the Interaction Between Electronic Admittance and Input Admittance of Two Coupled Resonators of Equal Q's

Having determined all relevant properties of the passive circuit as they appear across the grids of the bunching gap, we may now proceed to investigate the results arising from their interaction with the electronic admittance. The approach to be adopted is essentially the same as the one outlined earlier for the case of the conventional single-cavity reflex klystron. It involves a graphical superposition of the negative of the passive circuit admittance upon the small-signal-electronic-admittance spiral in the g-b plane, such as shown in Fig. 9(a). The location of the load lines with respect to the spiral has been chosen, somewhat arbitrarily, such that the ratio between the length of the Y_{es}-vector corresponding to $\theta = (2 + \frac{3}{4})$ cycles and that of the input admittance vector for $(Qk)^2 = 0$ and $\delta = 0$ equals two. Load lines have been drawn for five values of $(Qk)^2$ ranging from zero to unity.

The determination of mode shapes from Fig. 9(a) proceeds as illustrated in Table I. Taking the repeller drift angle as the independent variable we can obtain corresponding values of generated power (in arbitrary units) and frequency (in terms of $Q\delta$) by going through the steps indicated.

DETERMINATION OF THEORETICAL MODE SHAPES FOR 3¾ REPELLER MODE AND SECONDARY Q EQUAL TO PRIMARY Q

Fig. 9 — Graphical determination of mode shapes for a coupled resonator reflex klystron having the secondary cavity Q equal to primary Q. (a) Load lines for various degrees of coupling superimposed on small signal electronic admittance spiral. (b) Power output versus repeller drift angle. (c) Frequency versus repeller drift angle. (d) Power output versus frequency. (e) Power output, normalized with respect to its maximum value within the mode, shown as a function of frequency.

TABLE I — DETERMINATION OF MODE SHAPES FROM COMPLEX
ADMITTANCE PLANE PLOT OF FIGURE 9(a)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$K\theta_0$, Repeller Drift Angle (θ_0 = Midmode Value)	K, Ratio of Drift Angle to its Midmode Value	Small Signal Electronic Admittance	Passive Circuit Admittance	$\frac{2J_1(KV\theta_0)}{(KV\theta_0)}$ i.e. (4)/(3)	$(KV\theta_0)$ from Bessel-function plot of Fig. 2	$V\theta_0$, (6)/(2)	$(V\theta_0)^2$, Quantity Proportional to Power Output	$Q\delta$
		Measured off Fig. 9(a) in Arbitrary Units of Length						
$(Qk)^2 = 0.4$, 2¾-mode, i.e., $\theta_0 = (2 + \frac{3}{4})2\pi$ radians $= 5.5\pi$ radians								
5.20π	0.947	9.45	9.45	1.000	0.00	0.000	0.00	+0.85
5.25π	0.955	9.54	8.10	0.849	1.13	1.182	1.40	+0.67
5.30π	0.967	9.63	7.35	0.764	1.44	1.493	2.23	+0.53
5.35π	0.974	9.72	6.95	0.715	1.60	1.645	2.70	+0.42
5.40π	0.983	9.82	6.80	0.693	1.67	1.700	2.89	+0.30
5.45π	0.992	9.91	6.80	0.686	1.69	1.703	2.90	+0.17
θ_0 = 5.50π	1.000	10.00	7.00	0.700	1.65	1.645	2.71	0.00
5.55π	1.010	10.10	6.80	0.673	1.73	1.713	2.94	−0.17
5.60π	1.020	10.20	6.80	0.667	1.75	1.715	2.95	−0.30
5.65π	1.029	10.29	6.95	0.676	1.72	1.673	2.80	−0.42
5.70π	1.038	10.38	7.35	0.709	1.62	1.560	2.44	−0.53
5.75π	1.046	10.46	8.10	0.775	1.40	1.340	1.80	−0.67
5.80π	1.055	10.55	9.45	0.895	0.95	0.900	0.81	−0.85

The results of this analysis are shown in Fig. 9. This illustration, in addition to giving detailed performance characteristics for the 3 + ¾ mode, also indicates clearly the wealth of information which may be obtained from the complex admittance plane representation. Although the curves shown are self-explanatory, a few comments regarding their significance would seem to be in order. It is seen, for instance, that a coupling coefficient so adjusted that $(Qk)^2$ lies between 0.2 and 0.4 will produce a frequency range of essentially constant power. In particular, if we pick the curve for $(Qk)^2 = 0.4$ from the family of curves of Fig. 9(d), it will be seen that the variation in power over a bandwidth of $Q\delta = \pm0.4$ is ±2 per cent, while the corresponding value for the single resonator case, i.e., $(Qk)^2 = 0$, is minus 23 per cent. As the value of $(Qk)^2$ is increased beyond 0.4, the depression in the center of the mode becomes increasingly pronounced until it turns into a cusp for $(Qk)^2 = 1$. Mode shapes for $(Qk)^2 > 1$, though of no direct interest to this investigation, are indicated in Fig. 10 since they may be encountered in practice in cases of excessively tight coupling and could then be recognized as such. If the mode is traversed in the direction of increasing $|V_R|$ (or increasing frequency), such that the intersection of the electronic

admittance vector with the load line moves down along the latter as indicated by the arrows in Fig. 10(a), power changes smoothly until the Y_e-vector becomes tangent with the loop at 4. At this point there occurs a discontinuous jump in both power and frequency caused by the sudden shortening of the Y_e-vector from 4 to 7. From here on, power and frequency again become single valued functions of V_R. The mode shape corresponding to this uni-directional sweep is shown in Fig. 10(b). If the mode is traversed in the opposite sense we again encounter this discontinuity although it will now occur on the opposite side of the loop and, therefore, at a different repeller voltage and frequency. Fig.
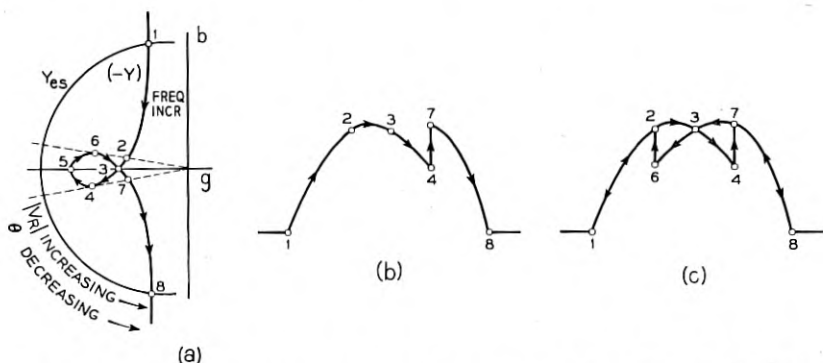

Fig. 10 — Production of load hysteresis by overcoupled cavities. (a) Load line for overcoupled cavities. (b) Oscillographic mode representation for unidirectional (sawtooth) repeller sweep. (c) Oscillographic mode representation for sinusoidal repeller sweep.

10(c) shows the load hysteresis effect as one might expect to observe it on the oscilloscope screen with a sinusoidal sweep applied to the repeller.

3.3 Driving Point Properties of Two Coupled Resonators Having Unequal Q's

The presentation of the theory of the coupled resonator reflex klystron will now be concluded by a discussion of the more general and, as we shall see, more useful case of two coupled cavities of unequal Q's. Specifically, we are considering the equivalent circuit shown in Fig. 11. By making approximations similar to those which led to equation 3.2, the input admittance for the case of unequal Q's may be derived as,[2, 4, 5]

$$\frac{Y}{G} = \left[1 + \frac{k^2 Q Q_s}{1 + (2Q_s\delta)^2}\right] + j2Q\delta\left[1 - \frac{(Q_s k)^2}{1 + (2Q_s\delta)^2}\right], \quad (3.11)$$
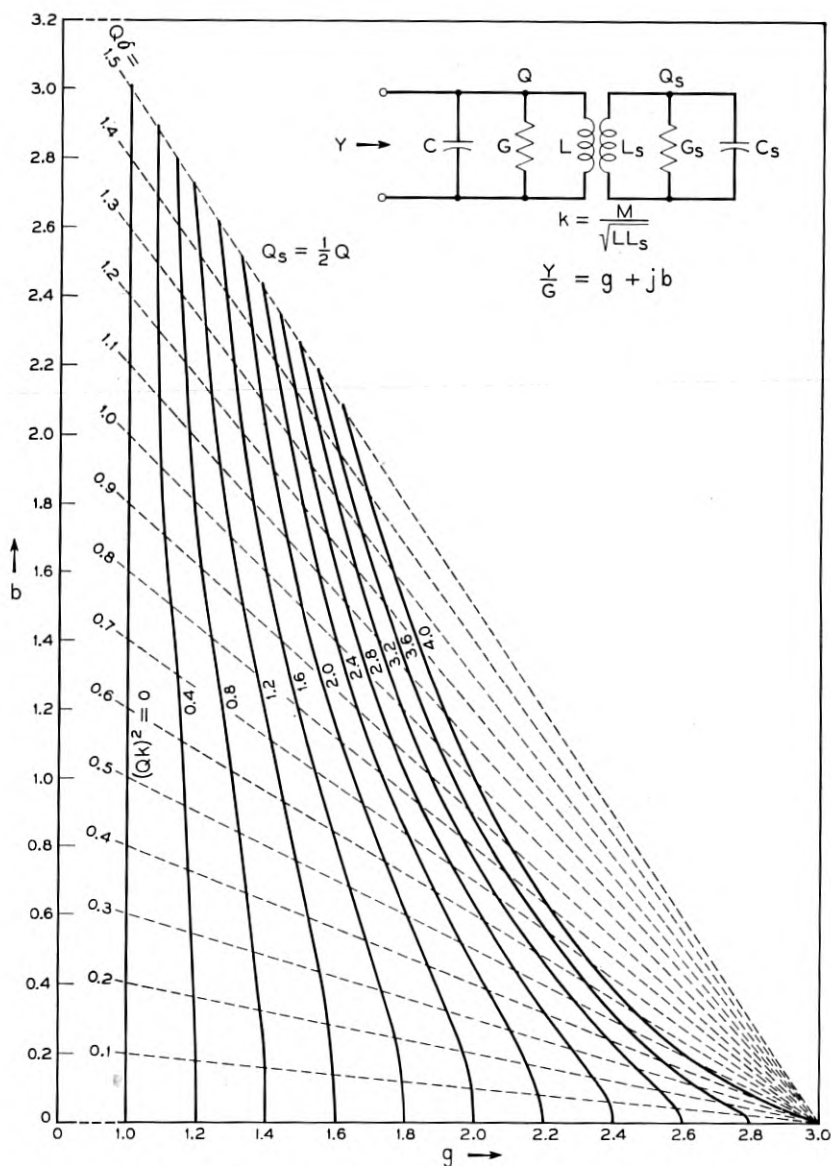
Fig. 11 — Driving point properties of two coupled, synchronously tuned resonators for the case of the secondary Q equal to half the primary Q.

or in more convenient form as,

$$\frac{Y}{G} = \left[1 + \frac{p(Qk)^2}{1 + p^2(2Q\delta)^2} \right] + j(2Q\delta) \left[1 - \frac{p^2(Qk)^2}{1 + p^2(2Q\delta)^2} \right], \quad (3.12)$$

where

$$p = \frac{Q_s}{Q}.$$

The above expression contains four variables, namely Y/G, (Qk), $(2Q\delta)$ and p, so that the complex admittance plane representation will now have to be restricted to particular values of p. One such plot, for $p = \frac{1}{2}$, appears in Fig. 11. It is similar to that of Fig. 7 except that only the positive half has been shown since equation (3.12) is symmetrical about the conductance axis; in addition, the parameter, $(Qk)^2$, has been carried to the point of critical coupling only. i.e.. the formation of a cusp, and not beyond. The frequency contours are again seen to be straight lines crossing the horizontal axis at a value of conductance equal to the input conductance for the condition of critical coupling and $2Q\delta = 0$.

Equation (3.12) shows that the susceptance term will be zero for

(a) $$2Q\delta = 0,$$

or for

(b) $$p^2(Qk)^2 = 1 + p^2(2Q\delta)^2. \quad (3.13)$$

The value of conductance corresponding to condition (a) is given by $g = 1 + p(Qk)^2$ and the value corresponding to condition (b) by $g = 1 + 1/p$. It is interesting to note that this latter value which determines the point of intersection of the frequency contours, as well as of the admittance plot for critical coupling, with the conductance axis, is independent of the actual values of Q and the degree of coupling and only dependent upon p, the ratio of Q's. Thus in Fig. 11, which constitutes a plot for $p = 0.5$, the value of conductance at which all the above named contours meet is given by $g = 1 + 1/0.5 = 3$.

From what has been said before we know that at critical coupling the admittance locus forms a cusp intersecting the conductance axis at $g = 1 + 1/p$ at the frequency, $2Q\delta = 0$. Substituting this value of $2Q\delta$ into the conductance term of equation 3.12 and equating to $1 + 1/p$ yields

$$1 + p(Qk)^2 = 1 + \frac{1}{p}.$$

Hence

$$pQk = Q_s k = 1 \text{ (for critical coupling)}. \tag{3.14}$$

The effect of reducing the secondary resonator Q is to broaden the frequency range over which a high input impedance or low admittance may be maintained. This is clearly illustrated in Fig. 12 where curves having the same value of conductance at resonance have been selected from Figs. 7 and 11 and superimposed to facilitate comparison.
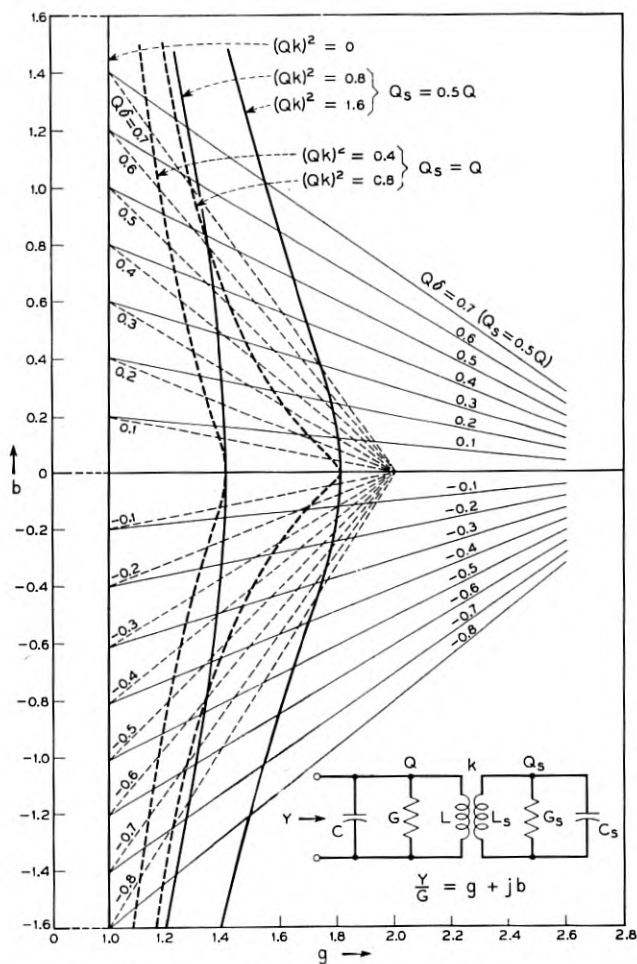

Fig. 12 — Comparison of admittance loci for coupled resonators of equal Q's with the case in which the secondary Q equals half the primary Q.

3.4 *Mode Shapes Obtainable with Two Coupled Resonators of Unequal Q's*

The complex admittance plane representation for the case of the secondary Q equal to half the primary Q is shown in Fig. 13(a) and the resulting mode shapes for the $3 + \frac{3}{4}$ repeller mode in Figs. 13(b), (c), (d), and (e). Again we notice that the coupling coefficient required for best modulation linearity is considerably smaller than that which results in a flat topped power curve. From Fig. 13(c) the most linear frequency-repeller voltage curve is associated with $(Qk)^2 \cong 0.8$ while from Fig. 13(d) a flat power mode may be obtained with a value of $(Qk)^2$ somewhat less than 1.6. In cases where neither modulation linearity nor constant power output are of importance but where the application requires a wide electronic tuning range, Fig. 13(e) shows the advantage to be gained from coupled cavities. Here, power output has been normalized with respect to its peak value within the particular mode under consideration and plotted against $Q\delta$. The ratio of half-power bandwidths for the curve corresponding to $(Qk)^2 = 2.4$ to the single resonator case, i.e., $(Qk)^2 = 0$, is seen to equal 1.73.

A phenomenon which may be encountered in the operation of the coupled resonator reflex klystron is illustrated by the $(Qk)^2 = 4.0$ curve of Fig. 13(b). Here we are dealing with a split mode in which the power, though everywhere a single valued function of repeller voltage, drops to zero over a range of repeller voltages centered about the middle of the mode. The reason for this behavior may be readily understood from an inspection of the complex admittance plane representation of Fig. 13(a). It is caused by the Y_{es}-locus for the $3 + \frac{3}{4}$ repeller mode crossing the appropriate load line and thereby resulting in a frequency band over which the condition for oscillation cannot be met.

4.0 AN EXPERIMENTAL COUPLED-RESONATOR REFLEX KLYSTRON

The reduction to practice of the theoretical results obtained in the above study raises these requirements:

(1) An arrangement must be found which allows the coupling between primary and secondary cavities as well as between primary cavity and waveguide output line to be varied continuously and independently.

(2) Either primary or secondary resonators (or both) must be tunable to allow frequency adjustments for synchronous operation.

(3) Secondary resonator Q should be continuously variable.

(4) The secondary resonator should be detachable for independent determination of Q.
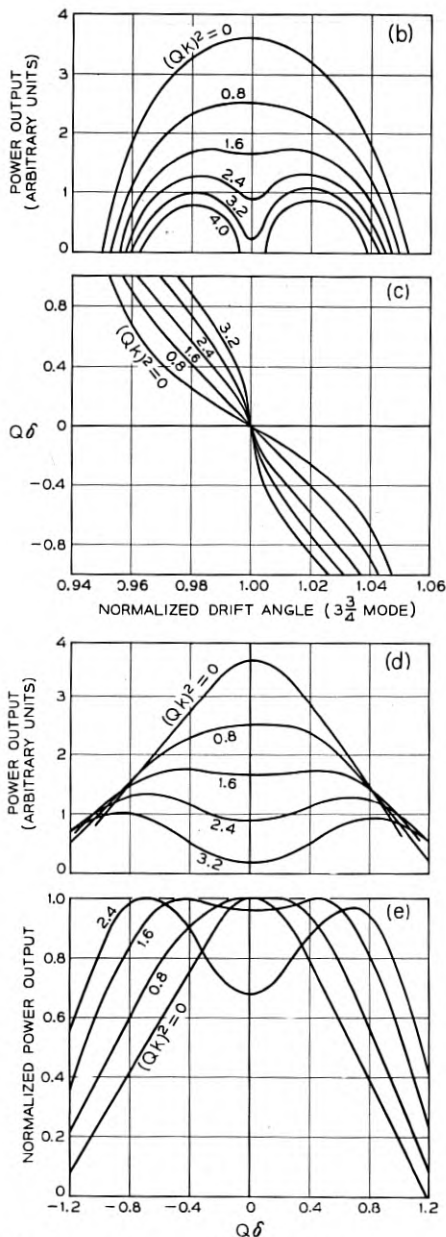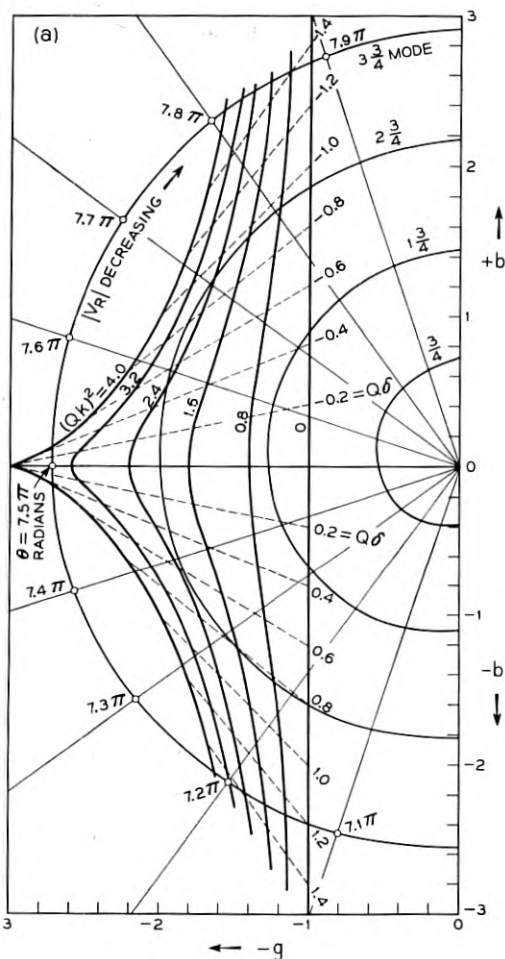
Fig. 13 — Graphical determination of mode shapes for a coupled resonator reflex klystron having the secondary cavity Q equal to half the primary Q. The effect of lowering the secondary Q may be studied by comparing the above curves with the corresponding ones shown in Fig. 9.

(5) Frequency at which these experiments are to be performed should be in a microwave band where good waveguide components and measurement techniques are available.

These considerations have led to the adoption of an external cavity type reflex klystron, the Sylvania 6BL6, as a vehicle for the experimental studies to be described, and operation in the 3700–4200 mc band.

4.1 Constructional Features of Experimental Tube and Circuit

The 6BL6 is one of a group of Sylvania low-voltage reflex klystrons[10] designed for use with external cavity resonators. Electrical connection to the interaction-gap grids is made through gold plated contact rings, formed from the disc seals which pass through the glass. The top ring is slightly smaller in diameter than the bottom ring, thereby permitting the insertion of the tube without disturbance to the associated cavity. Fig. 14 shows the external appearance of this tube and also contains a view of the major components of the passive circuit with cutouts to indicate the internal construction.
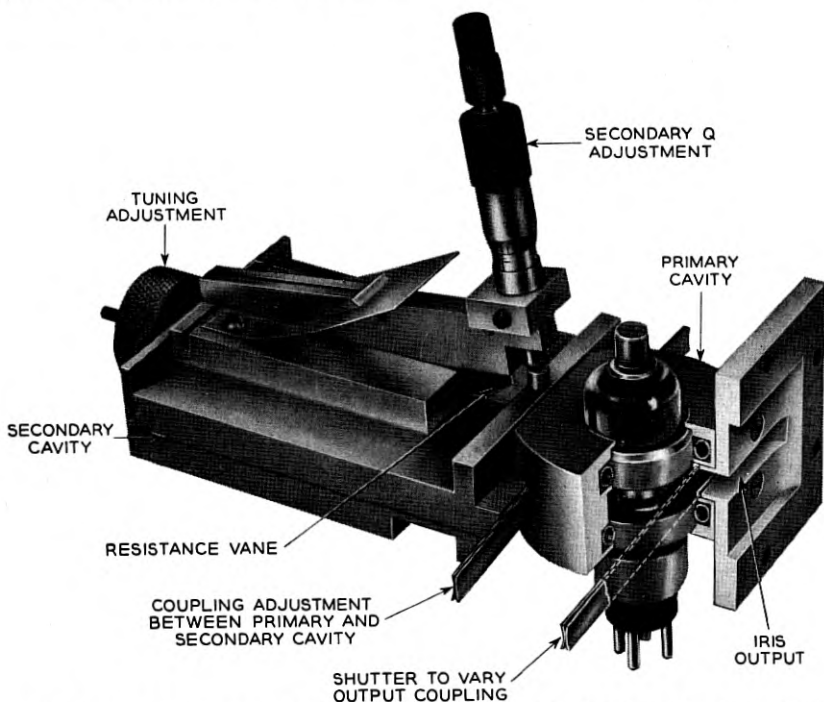
Fig. 14 — An experimental coupled resonator reflex klystron having a fixed frequency primary cavity.

Operation in the 4000-mc band was found to require a closely fitting primary cavity resonating in its principal mode. This unit, made of gold plated brass, is coupled to the output waveguide and secondary cavity respectively through two rectangular irises the sizes of which are independently variable by means of shutters. Toroidal contact springs located in circular grooves grip the inserted tube and complete the external circuit. The secondary cavity consists of a length of rectangular waveguide and uses a movable contacting plunger for tuning. A micrometer driven resistance vane may be inserted into this cavity through a longitudinal slot in its top surface thereby obtaining a wide range of continuously variable Q's. The entire unit is attached to the output waveguide by means of the adapter plate shown to the right of the primary cavity in Fig. 14.

A coupled resonator circuit having a tunable primary cavity is shown in the photograph of Fig. 15. The secondary cavity of this unit is identical to the one shown in the Fig. 14 but the primary resonator has been
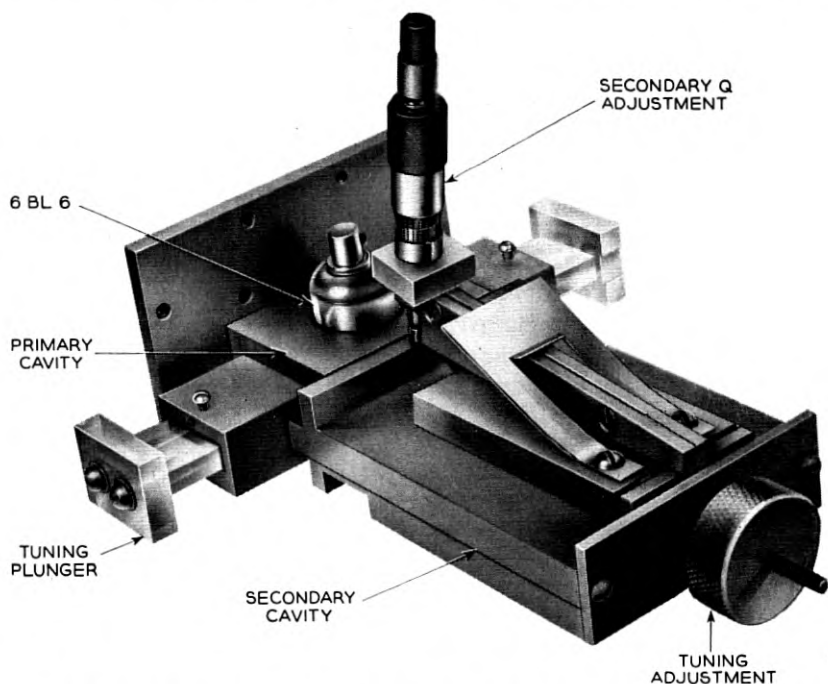

Fig. 15 — An experimental coupled resonator reflex klystron with a tunable primary cavity. The mechanical tuning range of this tube extends from 3700–4200 mc.

modified to take two non-contacting tuning plungers so dimensioned that they can be moved in very close to the glass wall of the 6BL6. In this way an operating frequency range from 3700–4200 mc was obtained.

4.2 *Qualitative Verification of Theory*

The results of a qualitative verification of the coupled-resonator reflex klystron theory are given in the form of oscillographic displays in Fig. 16. Photograph (a) shows the $(3 + \frac{3}{4})$-repeller mode of the 6BL6 with the secondary cavity completely decoupled. In photograph (b) the coupling iris and secondary Q have been adjusted for a flat topped output curve. The bandwidth of this curve for a variation in power output of ± 0.1 db was found to be 58 mc, i.e., considerably greater than the half-power bandwidth of the single resonator case above. Increasing $Q_s k$ beyond this point by either enlarging the coupling iris or withdrawing the resistance vane from the secondary resonator or a combination of both brings about the condition of critical coupling, illustrated by photograph (c). As explained earlier, this mode-shape results from the admittance plot of the passive circuit forming a cusp in the g-b plane. Comparison between this photograph and the theoretical mode of Fig. 9(b) for $(Qk)^2 = 1$ indicates good qualitative agreement. Coupling the secondary cavity still tighter, i.e., making the value of $Q_s k$ greater than unity, causes the formation of a loop in the circuit admittance plot and the consequent load hysteresis shown in photograph (d). Whereas all oscillograms discussed to this point have been obtained with a 60-cycle sinusoidal sweep applied to the repeller, the last one in this group, (e), results from a unidirectional (sawtooth) repeller sweep. The nature of this hysteresis effect has been explained earlier, and oscillograms one might expect to observe with overcoupled resonators for both sinusoidal and unidirectional sweeps were shown in Fig. 10. Again we note good qualitative agreement.

4.3 *Quantitative Verification of Theory*

As further strengthening of the theory underlying the operation of the coupled resonator reflex klystron a quantitative, experimental verification was undertaken. The methods used in connection with this work and the results obtained will form the subject of the sections to follow. To anticipate some of the conclusions which will be presented in the course of this description, let it be said here that the quantitative agreement between theory and experiment was found to be of an order high enough to justify amply the approximations involved in the ex-

pressions for both electronic admittance and passive circuit admittance, and to establish confidence in the method of analysis proposed to predict tube performance under a variety of conditions.
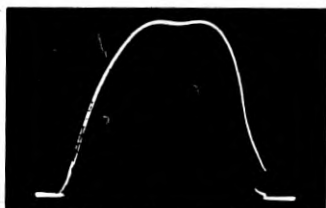
Since theoretical results had been worked out for the cases of the secondary Q equal to the primary Q and for the secondary Q equal to
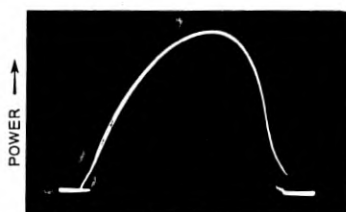
$|V_R|$ AND f INCREASING $\longrightarrow$

(a)

SECONDARY RESONATOR COM-
PLETELY DECOUPLED. HALF POWER
ELECTRONIC TUNING RANGE EQUAL
TO 49 MC.
(PEAK POWER = 100 MW)

(b)

SECONDARY RESONATOR COUPLING
AND Q ADJUSTED FOR FLAT TOPPED
MODE SHAPE. BANDWIDTH FOR
POWER VARIATION OF ± 0.1 DB IS
59 MC.

(c)

MODE SHAPE FOR CONDITION OF
CRITICAL COUPLING, i.e., $Q_S K = 1$
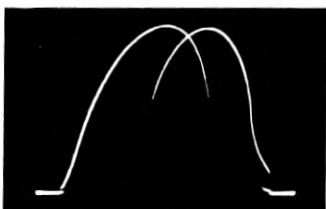
(d)

LOAD HYSTERESIS DUE TO OVER-
COUPLING, i.e., $Q_S K > 1$

(e)

SAME AS (d) EXCEPT FOR
UNIDIRECTIONAL (SAWTOOTH)
REPELLER SWEEP

Fig. 16 — Qualitative verification of theory. These oscillograms were obtained with a 6BL6 in the circuit of Fig. 14 and the following operating conditions: $V_0 = 325V$, $I_k = 28$ ma, $f_0 = 3800$ mc, $3 + \frac{3}{4}$ cycle repeller mode.

half the primary one, (see Figs. 7 and 11, respectively) it was decided to establish the same conditions in the experimental tube and to compare the results thus obtained with those predicted by theory.

Prerequisites for the execution of these experiments were:

(a) Knowledge of the primary Q.

(b) Complete calibration of the secondary cavity with respect to both the variation of Q_s with penetration of resistance vane and variation of coupling coefficient with size of coupling iris.

(c) Knowledge of the location of the load lines with respect to the small signal electronic admittance plot in the complex admittance plane.

4.3.1 *Determination of Primary Q.* The above parameter is understood to denote the "operating" or loaded Q of the tube with the output iris adjusted to its final and permanent size and the secondary cavity completely decoupled. For the case of an inductively tuned (fixed gap) reflex klystron, it may be determined experimentally using the expression,[7]

$$Q = \pi f_0^2 \left| \frac{\left|\frac{d\tau}{dV_R}\right|}{\left|\frac{df}{dV_R}\right|}\right|_{f\tau=N+3/4} - \pi(N + \tfrac{3}{4}), \tag{4.1}$$

where τ denotes the repeller space transit time. Both $d\tau/dV_R$ and df/dV_R are to be evaluated at the center of the mode.

The functional relationship between τ and V_R was obtained by placing the 6BL6 in the tunable primary cavity of Fig. 15 and determining the repeller voltage for maximum power output over the mechanical tuning range of the cavity. Since the same repeller mode was used throughout this test, τ must also have been the same at each of these frequencies, namely equal to $(N + \tfrac{3}{4})/f$ seconds. The experimentally determined plot of repeller voltage vs. frequency is given in Fig. 17. It is seen to be a straight line described by the equation,

$$f = (13.5V_R + 2495) \text{ mc} \tag{4.2}$$

and since, for the $3 + \tfrac{3}{4}$-repeller mode,

$$\tau = \frac{3.75}{f_{(mc)}} 10^{-6} \text{ sec.,} \tag{4.3}$$

the desired relation between τ and V_R is obtained from the above two equations as

$$\tau = (3.6V_R + 665)^{-1} \text{ sec.,} \tag{4.4}$$

whence the numerator of equation 4.1 follows as

$$\left| \frac{d\tau}{dV_R} \right| = \frac{3.6}{(3.6V_R + 665)^2} \, 10^{-6} \text{ sec/volt.} \qquad (4.5)$$

It remained only to replace the 6BL6 in the fixed tuned primary cavity, in which all subsequent tests were performed, and to determine the midmode repeller voltage and frequency. These values were found to be 96 volts and 3800 mc respectively, thus yielding by substitution into equation (4.5),

$$\left| \frac{d\tau}{dV_R} \right| = \frac{3.6 \times 10^{-6}}{(3.6 \times 96 + 665)^2} = 3.52 \times 10^{-12} \text{ sec/volt.}$$

The demoninator of equation (4.1) simply denotes the "modulation sensitivity" at the center of the mode. A simple measurement estab-
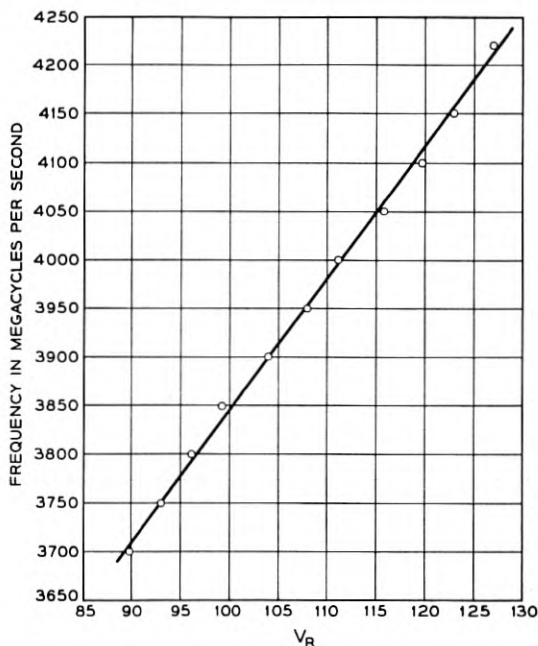
Fig. 17 — Experimental determination of the relation between frequency and repeller voltage. Points shown were obtained with a 6BL6 in tunable primary cavity of Fig. 14 with secondary cavity completely decoupled and operation in the $3 + \frac{3}{4}$ repeller mode at a beam voltage, $V_0 = 325V$. Frequency was varied by means of tuning plungers and repeller voltage recorded for maximum power output. Relation is given by, $f = (13.5V_R + 2495)$ mc.

lished its value as,

$$\left| \frac{df}{dV_R} \right| = 1.12 \text{ mc/volt.}$$

Substitution into equation (4.2) then gave the value of primary Q as,

$$Q = \pi(3.8)^2 10^{18} \frac{(3.52)10^{-12}}{(1.12)10^6} - \pi(3.75) = 130.$$

The entire Q-determination, incidentally, as outlined above, was repeated for the $2 + \frac{3}{4}$ mode and resulted in $Q = 128.8$ thus affording an excellent and independent double check.

4.3.2 *Calibration of Secondary Resonator.* To facilitate the establishment of controlled and reproducible conditions of secondary Q and coupling coefficient, two calibration curves had to be obtained. One, relating the values of secondary Q with the readings of the micrometer controlling the depth of insertion of the resistance vane and the other, relating the coefficient of coupling, k, with the coupling iris width. The latter could be varied by means of gold plated spring shutters as shown in Figs. 14 and 15.

This calibration of the secondary cavity was carried out in three distinct phases. The first phase involved the determination of the variation of Q_s with micrometer setting for values of Q_s ranging from 500 to 2000. The second phase, which was based on the results obtained in phase one, yielded the complete calibration of the coupling iris and the third and last phase, in turn dependent on results of phase two, yielded values of Q_s down to 65. The reasons for this particular sequence of measurements will become apparent in the following more detailed description.

By means of a Q-measurement technique[8] based on an oscillographic display of reflected power, points on the calibration curve were obtained as indicated by the circles in Fig. 18. It is seen that the lowest value of Q_s which could be determined by this method was 550. For lower values of Q, the sweep range of the signal generator became insufficient to display the required fraction of the resonance curve; in addition, the cavity proved excessively undercoupled to permit reliable measurements of bandwidths.

Using the values of Q_s thus determined and a particular property of coupled resonators covered earlier in this text and further elaborated below, the relation between coupling coefficient and iris width was established. It was shown earlier that an overcoupled secondary cavity gives rise to load hysteresis manifesting itself in mode shapes as sketched
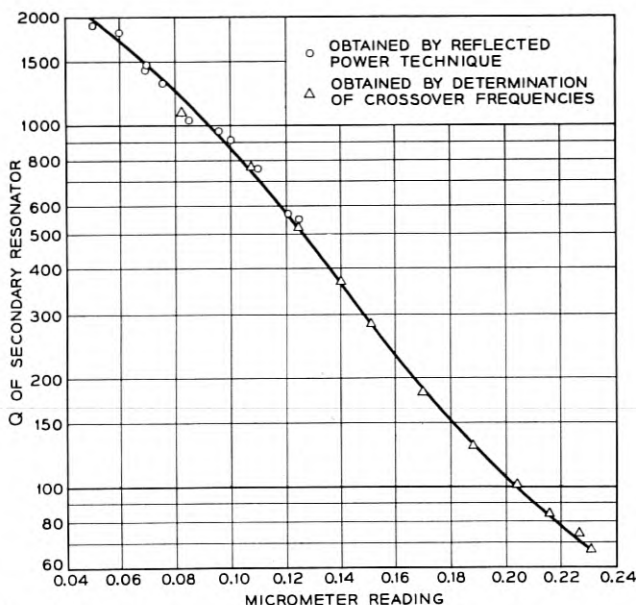
Fig. 18 — Calibration of secondary cavity. Variation of secondary Q with reading of micrometer controlling depth of insertion of resistance vane.

in Fig. 10 and shown in the form of an oscilloscope pattern in Fig. 16(d). Conditions pertaining at the crossover point were described by equation (3.13), which stated:

$$p^2(Qk)^2 = 1 + p^2(2Q\delta)^2,$$

or, since $pQ = Q_s$

$$(Q_s k)^2 = 1 + (2Q_s \delta)^2.$$

Hence the two frequencies corresponding to the crossover point are given by

$$\delta_{1,2} = \pm \frac{1}{2} \sqrt{k^2 - \left(\frac{1}{Q_s}\right)^2} \tag{4.6}$$

and their difference, $\Delta\delta$, by

$$\Delta\delta = \sqrt{k^2 - \left(\frac{1}{Q_s}\right)^2}. \tag{4.7}$$

Equation (4.7)[4, 5] provides an experimental method of determining k since $\Delta\delta$ can be accurately measured and Q_s is known from previous tests.

The actual procedure adopted in calibrating the coupling iris was this: The secondary cavity was set to a fairly high and known value of Q_s, say 1500, and the coupling iris adjusted to the smallest width which still resulted in overcoupling, i.e. in $Q_s k > 1$, and thus displayed the desired load hysteresis on the oscilloscope screen. By means of a high-Q absorption wavemeter the crossover frequencies were determined and, along with the known value of Q, substituted into equation (4.7) which then yielded the value of k. As a double check the above procedure was repeated for several values of Q_s and in accordance with theory the coupling coefficient, k, found to be a function only of the iris coupling width. A typical set of readings is given in Table II. They show that though Q_s was varied by a ratio of nearly 3:1, the resulting values of k only differed by about one-half per cent. Measurements as shown in the table were repeated for various iris widths. They resulted in the curve of Fig. 19.

To complete the calibration of the secondary cavity it was necessary to extend the relation between Q_s and micrometer readings to values lower than could be obtained by the reflected power technique. This was carried out with the aid of Fig. 19 in the following manner: With the coupling iris adjusted to its maximum width of 0.5″, the secondary Q was varied until the condition of critical coupling, as observed by the formation of a cusp on the oscilloscope pattern (see Fig. 16(c)), was reached. Since this condition is characterized by $Q_s k = 1$ and since the value of k corresponding to the iris opening could be read off the curve of Fig. 19, the value of Q_s then followed simply as the reciprocal of k. This measurement was repeated for successively smaller iris openings and the points shown as triangles in Fig. 18 were obtained. Inspection of this figure also reveals that these values of Q_s, ranging from 65 to 1100, overlap and coincide with values of Q_s determined

TABLE II

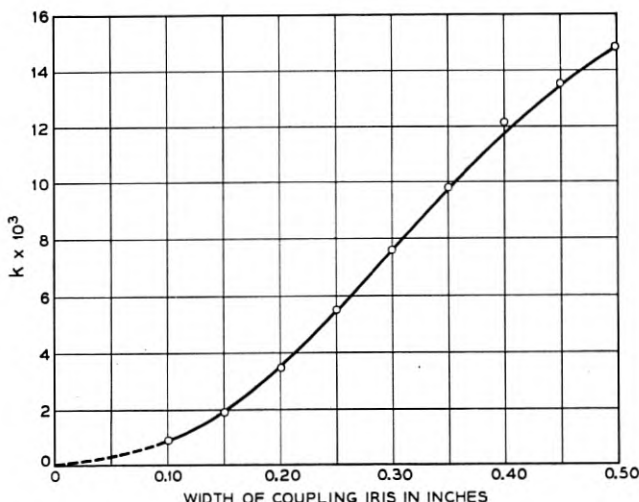| Coupling Iris Width (inches)	Secondary Cavity Q, Q_s	Crossover Frequencies f_1 and f_2 (mc)	$\Delta f = (f_1 - f_2)$ (mc)	$\Delta \delta = \dfrac{2\Delta f}{f_1 + f_2}$	From Equation (4.7), $k = \sqrt{(\Delta \delta)^2 + (1/Q_s)^2}$
0.500	1430	3749.9 3694.4	55.5	14.92×10^{-3}	14.93×10^{-3}
0.500	975	3748.4 3693.1	55.3	14.90×10^{-3}	14.90×10^{-3}
0.500	573	3745.2 3690.4	54.8	14.74×10^{-3}	14.85×10^{-3}

Fig. 19 — Calibration of coupling iris. Variation of coupling coefficient with width of coupling iris.

independently by the reflected power method. The significance of this experimental agreement is twofold. It serves to prove the correctness of the calibration curves of Figs. 18 and 19 and beyond that may be regarded as verifying much of the theory of the coupled resonator reflex klystron.

4.3.3 *Comparison of Experimental and Theoretical Mode Shapes.* Knowing the primary (operating) Q and the calibration of the secondary resonator, controlled operating conditions were established and the experimental mode shapes shown in the left hand column of Fig. 20 obtained. The first three families of curves in this column are the results of point by point measurements in which repeller voltage, taken as the independent variable, was varied by known amounts about its midmode value and the corresponding values of RF power output and frequency were determined by a thermistor-bridge-wattmeter and high-Q wavemeter respectively. The last family of curves in this group, (d), is derived from (c); in it power has been normalized with respect to its maximum value within the mode thus permitting the convenient determination of midmode percentage reduction of power and electronic tuning range to any desired power level for various values of $(Qk)^2$.

Additional details of the test conditions which gave rise to the ex-

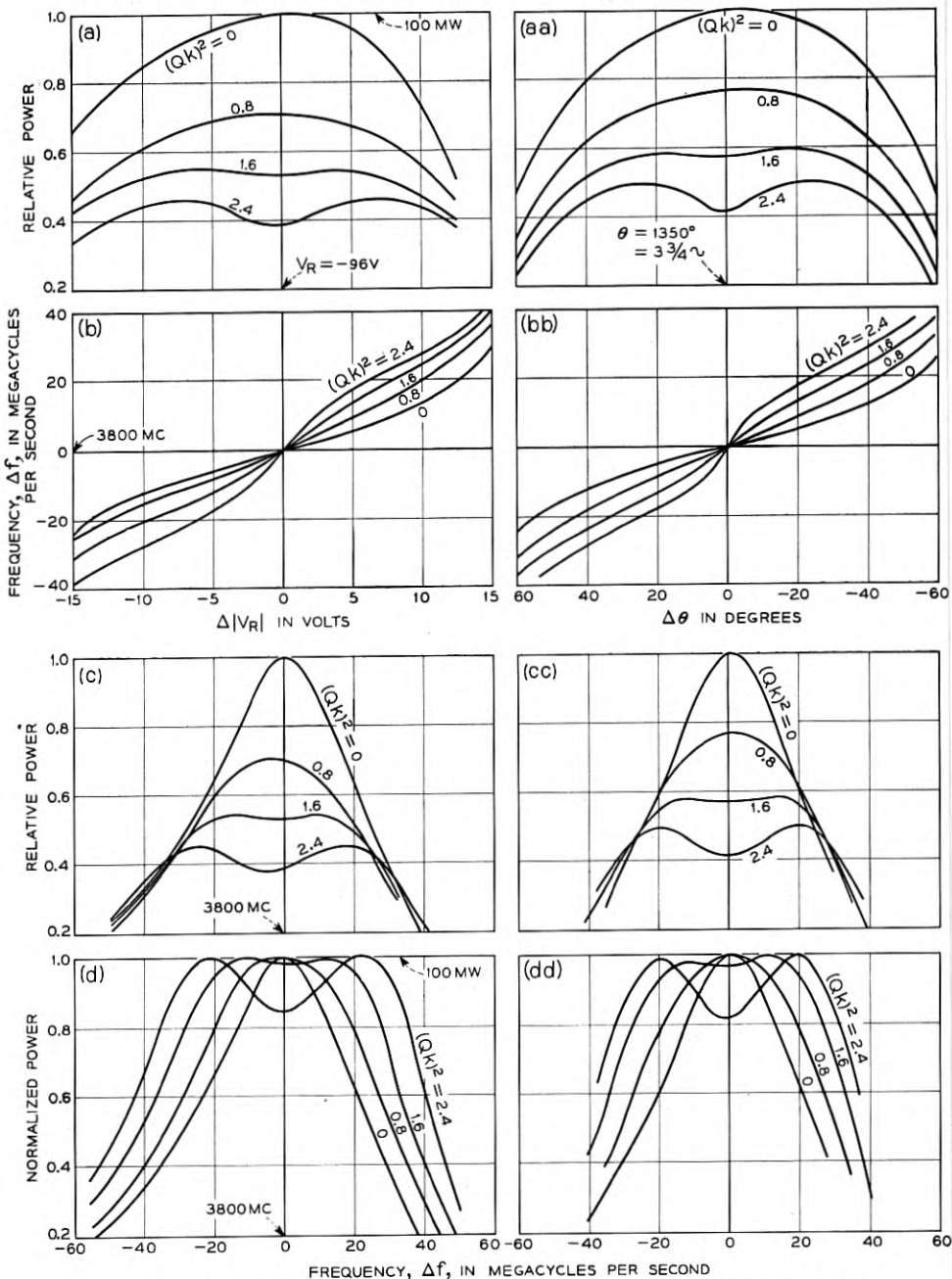Fig. 20 — Quantitative verification of coupled resonator reflex klystron theory. Curves (a) and (b) were obtained experimentally by changing the repeller voltage about its midmode value and noting the corresponding values of power and frequency. Curves (c) and (d) were deduced from (a) and (b). The theoretical curves were obtained graphically from a complex admittance plane plot similar to those of Figs. 9 and 13.

perimental curves shown are as follows:

6BL6 repeller mode $= 3 + \frac{3}{4}$ cycles

Resonator voltage, $V_0 = 325$ volts

Cathode current, $I_k = 28$ ma

Midmode repeller voltage, $V_R = -96$ volts

Midmode frequency, $f_0 = 3800$ mc

Secondary Q, $Q_s = \frac{1}{2}$ primary $Q = 65$, obtained by micrometer reading of 0.234 (see Fig. 18)

The parameter, $(Qk)^2$, was varied by adjusting the iris width, and hence k, according to Table III.

A few words, next, about the method by which the theoretical curves of Fig. 20 were derived. Input admittance plots for two coupled resonators having $Q_s = \frac{1}{2}Q$ were given earlier in this report (see Fig. 11).

TABLE III

$(Qk)^2$	k (for $Q = 130$)	Iris width from Fig. 19 in inches
0.0	0.00	0.000
0.8	6.88×10^{-3}	0.285
1.6	9.73×10^{-3}	0.350
2.4	11.90×10^{-3}	0.405

These included graphs for the values of $(Qk)^2$ listed in Table III and thus of direct use in a graphical admittance plane analysis such as described in Section 3.2. To establish correlation, however, between theory and experiment, it was necessary to determine the location of the load lines in relation to the small-signal electronic admittance plot for the $3 + \frac{3}{4}$ repeller mode. This was performed graphically by trial and error as follows: a number of electronic admittance loci were drawn and their interactions with the load line for $(Qk)^2 = 2.4$ studied. In particular, the ratio of minimum to maximum power was calculated for each trial until one was found agreeing well with the experimental value of 0.845 (see Fig. 20d). This latter trial, then, was taken as representing the proper relation between electronic and circuit admittance. Mode plots for other values of $(Qk)^2$ were determined without further reference to the experimental results and allowed to fall where they may.

In comparing the theoretical and experimental curves, then, the following points should be kept in mind: the frequency scales of the theoretical curves are normally expressed in terms of $Q\delta$. This was converted

to megacycles in Fig. 20 by using the relation,

$$\Delta f = (Q\delta)\frac{f_0}{Q} \text{ mc},$$

where Q was found to equal 130 by the independent measurement described in section 4.3.1 and f_0 taken as 3800 mc. The vertical power scales of the theoretical curves are not entirely independent of the experimental plots in as much as one of the latter, namely the plot for $(Qk)^2 = 2.4$, was used in determining the location of the load lines within the spiral diagram.

The curves presented in Fig. 20 are largely self-explaining. They indicate good agreement between theory and experiment. What disagreement there is, may be traced primarily to the assumptions involved in the small signal klystron theory not being fully met in practice. Thus, the slight discrepancy between Figs. 20(a) and (aa) may be due to the drift angle not being linearly related to repeller voltage or possibly due to the phase angle of the electronic admittance being affected by the magnitude of the gap voltage.[11] These factors, however, are eliminated in the mode plots (c) and (d), which for this reason exhibit better mode symmetry and very close agreement with theory.

4.4 Performance Data

The experimental curves of Fig. 20 were obtained with particular values of $(Qk)^2$ and Q_s/Q for which the input admittance of two coupled resonators had been computed and plotted earlier. As pointed out before, these values were chosen merely because they facilitated comparison between theory and experiment; they were not to be regarded, however, as representing optimum conditions. Whereas these curves had been obtained for a fixed secondary Q (namely equal to half the primary one), with changes in coupling-iris width producing the desired variations in $(Qk)^2$, the demonstration of optimum performance which follows was pursued along different lines. Here, the coupling iris was opened to its maximum width (0.500″) and the secondary Q adjusted until the conditions shown by the oscillograms of Fig. 21 were obtained.* These oscillograms indicate a number of mode shapes useful in applications requiring power output to be essentially independent of frequency. It is

---

* In adjusting the circuit parameters for a flat topped mode it should be borne in mind that the variation in power with frequency is a function of $Q_s k$ whereas the actual bandwidth varies inversely with Q_s. For a maximum flat band, therefore, Q_s should be chosen as small as possible consistent with a value of $(Q_s k)^2$ of about 0.35. In practice the lowest value of Q_s which can be used will be determined by the highest value of k obtainable with a given coupling iris.

seen that the useable bandwidth depends on the degree of flatness. Thus, if the application requires power to be absolutely constant, the useable bandwidth equals 39 mc; it increases to about 70 mc for a fluctuation in power of ±0.2 db. Fig. 21 also lists the values of half power electronic tuning for the oscillograms shown. They are seen to range between 107 and 113 mc.

By way of comparison, the mode shape for the case of a completely decoupled secondary resonator was shown in Fig. 16(a). Its peak power was found to equal about 100 mw, or about twice the power of the flat topped modes of Fig. 21, and its half power electronic range equaled 50
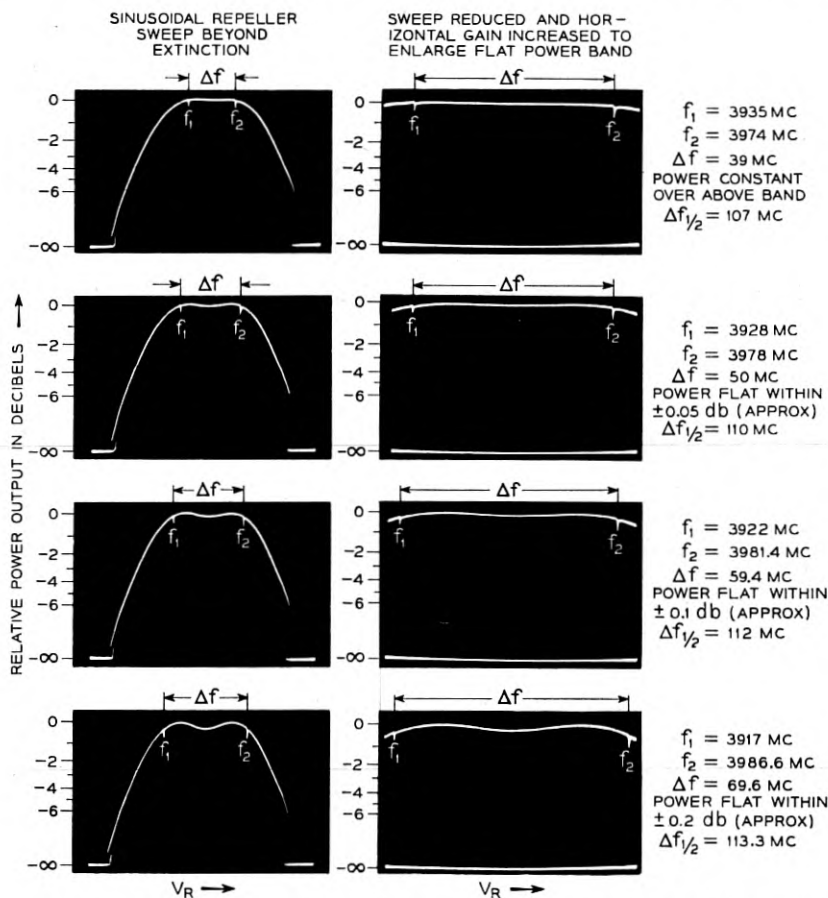

Fig. 21 — Oscillograms showing flat power bands obtainable with coupled resonator reflex klystron. 6BL6 operating in $3 + \frac{3}{4}$ mode, $V_0 = 325$ V, $I_k = 28$ ma, zero DB power level corresponds to 50 milliwatts.

mc. The price, then, we must pay for a better than two fold increase in electronic tuning and for a frequency range of constant power is a 3 db loss in the available power output.

Using the tunable primary cavity shown in Fig. 15, the mode performance just described could be obtained about any center frequency between 3700 and 4200 mc.

## 5.0 APPLICATIONS OF THE COUPLED-RESONATOR REFLEX KLYSTRON

A numer of applications making use of the variety of mode shapes which may be obtained with the coupled resonator reflex klystron were indicated in the introductory section of this paper. Some of these applications were tried experimentally and are described below, others,
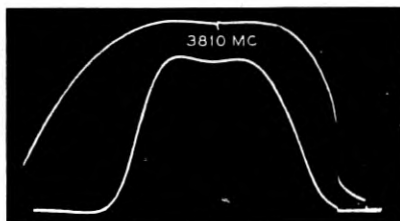
(a)

USE OF COUPLED-RESONATOR REFLEX KLYSTRON IN GAIN COMPARATOR TO DISPLAY BANDPASS CHARACTERISTIC OF MICROWAVE CHANNEL FILTER. TOP TRACE REPRESENTS INCIDENT POWER, LOWER TRACE TRANS- MITTED POWER.

CENTER FREQUENCY OF FILTER = 3810 MC
PASS BAND = 28 MC
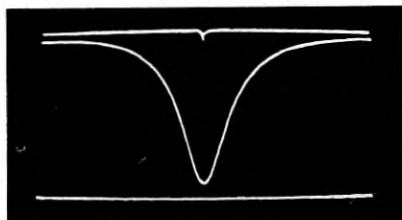
(b)

USE OF COUPLED-RESONATOR REFLEX KLYSTRON AS SWEPT FREQUENCY SOURCE IN DETERMINATION OF Q OF UNKNOWN RESONATOR BY REFLECTED POWER METHOD. TOP TRACE REPRESENTS POWER INCIDENT UPON CAVITY UNDER TEST, MIDDLE TRACE REFLECTED POWER AND LOWER TRACE ZERO POWER LEVEL. (NOTE FLATNESS OF INCIDENT POWER TRACE)
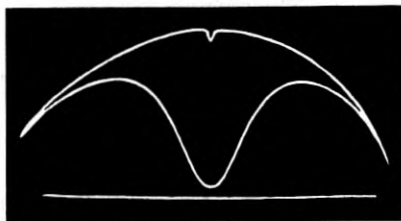
(c)

OSCILLOSCOPE DISPLAY FOR SAME RESONATOR RESULTING FROM USE OF CONVENTIONAL KLYSTRON AS SWEPT FREQUENCY SOURCE

Fig. 22 — Examples of application of coupled resonator reflex klystron as microwave sweeper.

and in particular those making use of the improved modulation linearity resulting from the use of two resonators, could not be tried since the required test apparatus[9] was unavailable. The close correlation between theory and experiment, however, as demonstrated in Section 4.3.3 leaves little doubt as to the feasibility of such applications.

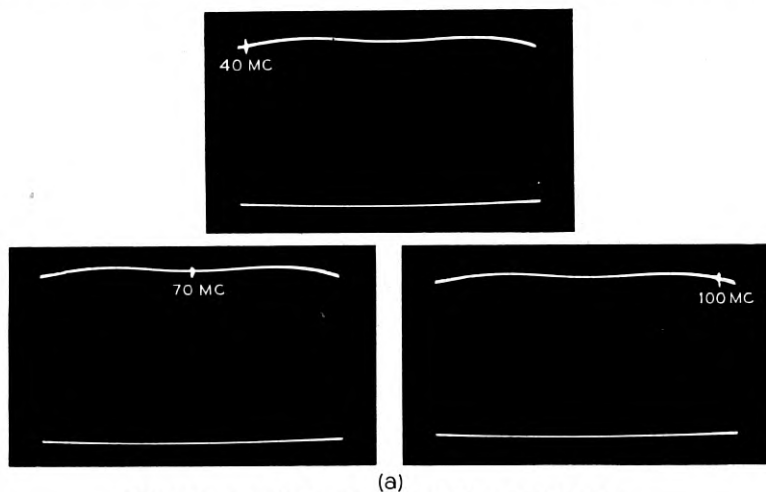Two examples of the use of the coupled resonator reflex klystron as a microwave sweeper are given in Fig. 22. The first example shows the band pass characteristic of a channel filter, 28 mc wide. The second example demonstrates the use of this tube in a Q-measurement scheme[8]

(a)

SUCCESSIVE PHOTOGRAPHS OF OSCILLOSCOPE DISPLAY SHOWING IF-POWER AS A FUNCTION OF FREQUENCY WITH SIGNAL GENERATOR TUNED TO DIFFERENT FREQUENCIES
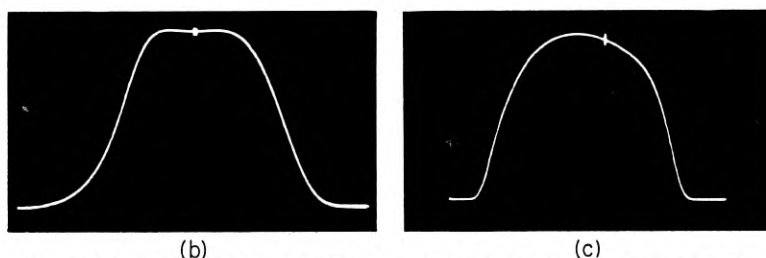
(b)

GAIN-FREQUENCY CHARACTERISTIC OF IF AMPLIFIER USING COUPLED-RESONATOR REFLEX KLYSTRON AS SWEEP SOURCE

CENTER FREQUENCY OF IF-AMPL = 70MC
FLAT BANDWIDTH OF IF-AMPL = 22 MC

(c)

GAIN-FREQUENCY CHARACTERISTIC OF SAME IF AMPLIFIER AS OBTAINED BY CONVENTIONAL (SINGLE RESONATOR) REFLEX KLYSTRON. DISTORTION IN GAIN ENVELOPE DUE TO MODE SHAPE OF KLYSTRON.

Fig. 23 — Example of application of coupled resonator reflex klystron as intermediate frequency sweeper.

based on the oscillographic display of reflected power. The advantage of a flat-topped mode shape for this measurement may be seen by comparison with the oscillogram of Fig. 22(c) which was obtained for the same resonator under test using a conventional reflex klystron as swept frequency source.

The center frequency of the flat power band obtainable with the coupled resonator reflex klystron may be shifted into the intermediate frequency range by mixing its output with a suitable local oscillator. This is shown in Fig. 23 where the first three oscillograms represent a power band flat to within ± 0.1 db and 60 mc wide centered around an IF frequency of 70 mc. The frequency markers are obtained from a signal generator tuned to the frequencies indicated and coupled lightly to the output of the mixer. Fig. 23(b) shows the gain-frequency characteristic of a 70 mc-IF strip having a useful bandwidth of 22 mc. The last oscillogram shows the gain-frequency characteristic of the same IF amplifier with the coupled resonator reflex klystron replaced by one of conventional design. Due to the inadequate electronic tuning range of this klystron and due to the asymmetry of its mode shape, the gain envelope of the IF amplifier appears distorted.

6.0 CONCLUSIONS

The theory and reduction to practice of a coupled resonator reflex klystron have been presented. This device differs from the conventional reflex klystron in that the electronic admittance interacts with the input admittance of two coupled resonators. Significant and advantageous changes in performance resulting therefrom are:

(1) Power output can be made to be substantially flat over part of the electronic tuning range. The frequency range of flat power is greater than the half power electronic tuning range of a klystron having the same electron-optical system but interacting with a single resonator.

(2) The half-power electronic tuning range of the coupled resonator klystron is more than twice that of the equivalent single resonator klystron.

(3) Modulation linearity may be obtained over a greatly increased frequency swing.

A reduction in power output of about 3 db occurs for a secondary resonator Q and coupling coefficient adjustment designed to yield a maximum flat band or maximum electronic tuning, a much smaller reduction in output power, perhaps 10 per cent, will provide a substantial improvement in modulation linearity.

Qualitative and quantitative experimental verifications of the theory were undertaken and good agreement obtained. Oscillograms of mode shapes having a range of constant power output and wide half-power electronic tuning were presented together with a number of applications of the coupled resonator reflex klystron both as a microwave and as an intermediate frequency sweep source. Other applications were indicated and additional ones may suggest themselves to the reader.

It is believed that the addition of the second resonator makes the reflex klystron a more useful device in many of its more important applications.

REFERENCES

1. Pierce, J. R., and W. G. Shepherd, Reflex Oscillators. Bell System Tech. J., **26,** pp. 460–690, July, 1947. This reference constitutes a very complete treatment of reflex klystron theory and practice. The derivation of the expression for electronic admittance is given in Appendix III, pages 639–643; of particular interest to the coupled resonator reflex klystron is Pierce and Shepherd's treatment of Frequency Sensitive Loads and Long Lines Effects, pp. 523, where it is shown that certain load conditions encountered in the operation of reflex klystrons may result in load hysteresis similar to the one described in the preceding paper as due to an overcoupled secondary resonator.
2. Martin, R. A., and R. D. Teasdale, Input Admittance Characteristics of a Tuned Coupled Circuit. I. R. E., Proc., p. 57, January, 1952, and correction to this paper published in I. R. E., Proc., p. 459, April, 1952. This paper presents a precise expression for the input admittance of two coupled resonant circuits and gives plots of specific numerical examples. The expression for input admittance given in the above named correction reduces to equation (3.2) and (3.12).
3. Bleaney, B., Electronic Tuning of Reflection Klystrons. Wireless Engineer, p. 6, Jan., 1948. A valuable background paper on the mechanics and maximization of electronic tuning.
4. Very High-Frequency Techniques, Vol. II, Radio Research Laboratory, Harvard University. (McGraw-Hill). Chapter 31, p. 849 contains a concise and valuable treatment of reflex klystron theory. It leads up to the application of external-cavity reflex klystrons in wide tuning range coaxial resonators. In the course of explaining some load hysteresis phenomena which are sometimes caused in these circuits by parasitic resonances or excitation of undesired modes, a simplified expression for the input admittance of two coupled resonators is given. This expression is similar to equation (3.12). An experimental method is also suggested (p. 869) to measure the coupling coefficient between the main resonator and the undesired resonance. Use has been made of this method in the calibration of the secondary resonator in Section 4.3.2.

5. Spangenberg, K. R., Vacuum Tubes, p. 601. (McGraw-Hill). This reference covers essentially the same material as Reference 4.
6. Reference Data for Radio Engineers, Third Edition, p. 121, Federal Telephone and Telegraph Corp. Here a universal chart of the selectivity (transfer characteristic) of two coupled tuned circuits is presented. The format and choice of variables of this presentation has inspired the selection of the equivalent quantities applying to the driving point properties of coupled tuned circuits.
7. Harmen, W. W., and J. H. Tillotson, Beam-Loading Effects in Small Klystrons. I. R. E., Proc., p. 1419, Dec., 1949. Equation (3) of this reference is identical with equation (4.1) and has been used in the experimental determination of the primary operating Q.
8. Reed, E. D., A Sweep Frequency Method of Q Measurement for Single-Ended Resonators. Proc. National Electronics Conference, Vol. VII, p. 162. Also reprinted as Bell Telephone System Monograph 1953. This method of Q-measurement was used in the calibration of the secondary resonator described in Section 4.3.2. The coupled resonator reflex klystron provides a useful tool for this type of measurement as may be seen from Fig. 22(b).
9. Albersheim, W. J., Measurement Techniques for Broad-Band Long-Distance Radio Relay Systems. I.R.E., Proc., p. 548, May, 1952. Figs. 4(a) and 4(b) of this reference give the block diagram of a "Linearity-Test-Set" developed at Bell Telephone Laboratories to determine the modulation linearity of a reflex klystron used as FM-deviator.
10. Boehlke, P. G., and F. C. Breedan, External Cavity Klystron. Electronics, p. 114, July, 1947. This paper gives an account of the design considerations which led to the development of the 6BL6 reflex klystron. It also contains useful information on its performance characteristics.
11. Garrison, J. B., A Qualitative Analysis of Hysteresis in Reflex Oscillators. Radiation Laboratory Report No. 650, dated Feb. 4, 1946. In this treatment on the causes of electronic hysteresis Garrison shows (on page 5) that a dependence of the phase angle of the electronic admittance upon the magnitude of RF gap voltage will give rise to asymmetry in the mode plot of power output vs. repeller voltage although the power vs. frequency relation will be symmetrical. (Note that the experimental mode plots for $Qk = 0$ of Figs. 20(a) and (c) show evidence of the same phenomenon.)

Abstracts of Bell System Technical Papers*
Not Published in This Journal

ANDERSON, P. W.[1]
Concept of Spin-Lattice Relaxation in Ferromagnetic Materials, Letter to the Editor, Phys. Rev., **88,** pp. 1214, Dec. 1, 1952.

BALASHEK, S., see K. H. Davis.

BAND, W., see R. A. Nelson.

BEACH, A. L., see J. J. Lander.

BELL TELEPHONE LABORATORIES Transistor Teachers Summer School.
Experimental Verification of the Relationship Between Diffusion Constant and Mobility of Electrons and Holes, Phys. Rev., **88,** pp. 1368–1369, Dec. 15, 1952.
The relationship between diffusion constant and mobility, called the Einstein relationship has been experimentally verified for electrons and holes in germanium. This has been accomplished by measuring the rate of increase in half concentration width of a pulse of minority carriers moving in an electric field.

BIDDULPH, R., see K. H. Davis.

BRIDGMAN, D. C.[2]
College Graduates and the Country's Telephone Industry, Jl. College Placement, **13,** pp. 19–27, Oct., 1952.

BULLINGTON, K.[1]
Radio Transmission Beyond the Horizon in the 40- to 4,000-Mc Band, I.R.E., Proc., **41,** pp. 132–135, Jan., 1953. (Monograph 2060).
Reliable signals have been received at distances of several hundred miles at frequencies of 500 and 3,700 mc. The median signal levels are 50 to 90

---

* Certain of these papers are available as Bell System Monographs and may be obtained on request to the Publication Department, Bell Telephone Laboratories, Inc., 463 West Street, New York 14, N. Y. For papers available in this form, the monograph number is given in parentheses following the date of publication, and this number should be given in all requests.
[1] Bell Telephone Laboratories, Inc.
[2] American Telephone and Telegraph Company.

db below the free-space field, but are hundreds of db (in one case 700 db) stronger than the value predicted by the classical theory based on a smooth spherical earth with a standard atmosphere. Antenna gains and beam widths are maintained to a first approximation and no long delayed echoes have been found.

CLARK, A. B.[1]

Development of Telephony in the United States, A.I.E.E., Trans., Commun. and Electronics Sect., **71,** pp. 348–364, Nov., 1952 (Monograph 2045).

The telephone was invented twenty-four years after the founding of the American Society of Civil Engineers and Architects, which society commemorated its Centennial in September, 1952. The telephone was eight years old when the American Institute of Electrical Engineers was founded. Since Bell's invention, the telephone business has grown tremendously and this growth has been greatly dependent on developments in science and engineering. This paper traces, and endeavors to give the significance of, the major developments. Brief mention is made of some developments still in the making and of some ideas as to the future potentialities of the business.

CONWELL, E. M.[1]

Mobility in High Electric Fields, Phys. Rev., **88,** pp. 1379–1380, Dec. 15, 1952.

An extension of conductivity theory to high fields, subject to the usual simplifying assumptions, is carried out for the cases in which the change of energy of an electron in a collision can be neglected. This yields a relationship between mobility and relaxation time which is valid over a wide range of fields.

DAVIS, K. H.,[1] R. BIDDULPH[1] AND S. BALASHEK[1]

Automatic Recognition of Spoken Digits, J. Acoust. Soc. Am., **24,** pp. 637–642, Nov., 1952.

The recognizer discussed will automatically recognize telephone-quality digits spoken at normal speech rates by a single individual, with an accuracy varying between 97 and 99 per cent. After some preliminary analysis of the speech of any individual, the circuit can be adjusted to deliver a similar accuracy on the speech of that individual. The circuit is not, however, in its present configuration, capable of performing equally well on the speech of a series of talkers without recourse to such adjustment. Circuitry involves division of the speech spectrum into two frequency bands, one below and the other above 900 cps. Axis-crossing counts are then individually made of both band energies to determine the frequency of the maximum syllabic rate energy within each band. Simultaneous two-dimensional frequency portrayal is found to possess recognition significance. Standards are then determined, one for each digit of the ten-digit series, and are built into the recognizer as a form of elemental memory. By means of a series of calculations performed

---

[1] Bell Telephone Laboratories, Inc.

on the spoken input digit, a best match type comparison is made with each of the ten standard digit patterns and the digit of best match selected.

DICKTEN, E., see R. L. Wallace, Jr.

FELKER, J. H.[1]
Typical Block Diagrams for a Transistor Digital Computer, Elec. Eng., **71,** pp. 1103–1108, Dec., 1952 and A.I.E.E. Trans., **71,** pp. 175–182, 1952 (Monograph 2046).
The superior speed capabilities of vacuum tubes have led to their use in computer designs to replace relays. Because of their small size, low power consumption, and long life expectancy, it now appears that transistors will replace tubes as computer elements. Here is a study of binary computer functions in which transistors are employed.

FRAYNE, J. G.,[6] AND J. P. LIVADARY[7]
Dual Photomagnetic Intermediate Studio Recording, S.M.P.T.E., Jl., **59,** pp. 388–397, Nov., 1952.
Selected production magnetic tracks are transferred to a recorder which lays down collinear 200-mil push-pull direct-positive variable-area and magnetic tracks. Magnetic stripe is on base of photosensitive emulsion on the opposite edge of film from photo track. The photo track may be used for reviewing, cutting, etc. Re-recording is done from assembled magnetic tracks. This method combines advantages of photo track for editing and provides superior quality of magnetic track.

HAGSTRUM, H. D.[1]
Electron Ejection from Mo by He^+, He^{++}, and He_2^+, Phys. Rev., **89,** pp. 244–255, Jan. 1, 1953.
Total yield and kinetic energy distribution have been measured for electrons ejected from atomically clean and gas covered molybdenum by the ions He^+, He^{++}, and He_2^+, in the kinetic energy range 10 to 1000 ev. Evidence is presented that one electron is excited into the kinetic energy continuum for each incident. He^+ ion and that the electrons so excited are partially internally reflected at the potential barrier of the metal. The slowest ions observed were found to eject 0.25, 0.72, and 0.13 electron per ion for He^+, He^{++}, and He_2^+, respectively. Total electron yield is found to be nearly independent of ion kinetic energy up to 1000 ev. This observation and that of the kinetic energy maximum for slow ions indicate that the electrons are released in an Auger type process for which the energy is supplied by the potential and not the kinetic energy of the ion (potential ejection). Electrons of kinetic energy greater than the upper limit predicted by present theory are observed for faster ions and are accounted for by the shift of the energy levels of the bombarding particle when it is near the metal surface. Some conclusions con-

---

[1] Bell Telephone Laboratories, Inc.
[6] Westrex Corporation.
[7] Columbia Pictures Corporation.

cerning reflection processes at a metal surface and the nature of electron ejection by the alpha-particle (He^{++}) and the molecular helium ion (He_2^+) come out of this work.

HEIDENREICH, R. D.[1]

Methods in Electron Microscopy of Solids, Rev. Sci. Instr., **23,** pp. 583–594, Nov., 1952 (Monograph 2047).

Methods of replicating solid surfaces for electron microscopy are reviewed and compared. Preparation of metal surfaces for electron microscopy is discussed, and the advantages of employing electron diffraction techniques in evaluating prepared surfaces are pointed out. Examples of the application of replicas include steel, precipitation in alloys, such as Alnico 5, and studies of slip in aluminum. Growth spirals on the surfaces of crystals of n-paraffins are demonstrated. The use of thin metal sections and of emission electron microscopy in studying metallic structures is discussed, and examples are given.

HUTCHINSON, A. R.[3]

How to Conceal Telephone Wires, Keep Desks Neat, Standardization, **23,** p. 407, Dec., 1952.

JAKES, W. C., JR.[1]

Theoretical Study of an Antenna-Reflector Problem, I.R.E., Proc., **41,** pp. 272–274, Feb., 1953.

This paper gives the results of a theoretical investigation of an antenna used with a plane reflector. This finds application in microwave relay stations, where the antenna is placed at ground level facing up and the reflector is located some distance above it. The results given show that there are certain values of λ, separation distance, reflector and antenna size for which the received power is greater than for the same antenna alone at the elevated location.

KAPLAN, E. L.[1]

Tensor Notation and the Sampling Cumulants of k-Statistics, Biometrika, **39,** pp. 319–323, Dec., 1952.

Now and then in the literature one finds results relating to multivariate distributions which are derived virtually independently of, or with considerable effort from, the corresponding univariate relations, whereas they are in fact only very mild generalizations of the latter, as will be shown. Only the familiar concepts of moments, characteristic functions, cumulants, and k-statistics and their sampling cumulants will be discussed here. It should be emphasized that these concepts are identical with those ordinarily used in multivariate situations; the only novelty lies in the concise manner of representing and handling them.

---

[1] Bell Telephone Laboratories, Inc.
[3] Western Electric Company, Inc.

KELLER, A. C.[1]

Economics of High-Speed Photography, S.M.P.T.E., Jl., **59,** pp. 365–368, Nov., 1952 (Monograph 2052).

The economics of the use of high-speed photography in research and development work are discussed. High-speed photography is a relatively new tool for engineers which can be used to measure mechanical or electrical effects or both at the same time. Examples are given which illustrate the savings in engineering manpower as well as in materials, devices and systems.

KERN, H. E., see J. J. Lander.

KOCK, W. E.,[1] AND R. L. MILLER.[1]

Dynamic Spectrograms of Speech, Letter to the Editor, J. Acoust. Soc. Am., **24,** pp. 783–784, Nov., 1952.

KOCK, W. E.[1]

Problem of Selective Voice Control, J. Acoust. Soc. Am., **24,** pp. 625–628, Nov., 1952 (Monograph 2048).

The development of devices which can be operated automatically from the phonetic content of speech may be viewed in terms of the more general problem of the reduction of channel capacity in communications systems. Significance has been observed in formant positions and movements as regard the identification of speech sound. The basic problems in the derivation of phonemes from the formant patterns are reviewed.

LANDER, J. J.,[1] H. E. KERN[1] AND A. L. BEACH[1]

Solubility and Diffusion Coefficient of Carbon in Nickel: Reaction Rates of Nickel-Carbon Alloys with Barium Oxide, J. Appl. Phys., **23,** pp. 1305–1309, Dec., 1952.

Experimental values for the solubility of carbon in nickel in the range 700°C to 1300°C yield the equation $\ln S = 2.480 - 4{,}880/T$, where S is the solubility in grams of carbon per 100 grams of nickel. Values obtained for the diffusion coefficient in the same range fit the equation $\ln D = 0.909 - 20{,}200/T$, where D is in cm² per second. These results are of some interest in the problem of the activation of thermionic oxide coated cathodes, and the experimental method used to measure the diffusion coefficients is related to phenomena occurring in vacuum tubes. To extend the usefulness of the results in this direction, rates of reaction between diffused carbon and barium oxide coatings on nickel have been measured. It was found that the rates are diffusion limited over a wide range of conditions of interest.

LEWIS, H. W.[1]

Multiple Meson Production in Nucleon-Nucleon Collisions. Revs. Modern Phys. **24,** pp. 241–248, Oct. 1952 (Monograph 2049).

LIVADARY, J. P., see J. G. Frayne

---

[1] Bell Telephone Laboratories, Inc.

LUKE, C. L.[1]
Photometric Determination of Silicon in Ferrous, Ferromagnetic, Nickel, and Copper Alloys — A Molybdenum Blue Method, Anal. Chem., **25,** pp. 148–151, Jan., 1952.

A simple, rapid photometric method for the determination of silicon in ferrous, ferromagnetic, nickel, and copper alloys has been developed. Wide applicability is its most unique and important feature. Interfering elements are removed by a carbamate-chloroform extraction and silicon is then determined by the photometric molybdenum blue method. Confirmatory data on a wide variety of samples of known silicon content are presented.

LUMSDEN, G. Q.[1]
A Quarter Century of Evaluating Pole Preservatives. Amer. Wood Preservers' Assoc., Proc., **48,** pp. 27–47, 1952 (Monograph 1999).

MACWILLIAMS, W. H., JR.[1]
Computers — Past, Present, and Future, Elec. Eng., **72,** pp. 116–121, Feb., 1953.

This article deals with the historical development of computers. It also discusses current problems and indicates future structural and functional computer trends which will help to free man from burdensome calculations and increase his material wealth while permitting him more time for pursuits not directly concerned with earning a living.

MILLER, R. L., see W. E. Koch.

MUMFORD, W. W.[1]
Optimum Piston Position for Wide-Band Coaxial-to-Waveguide Transducers[2] I.R.E., Proc., **41,** pp. 256–261, Feb., 1953.

A coaxial line can be matched to a waveguide by means of a probe antenna located ahead of a short-circuiting plunger. An impedance match can usually be achieved by varying any two of the following three dimensions: (a) the off-center position of the probe, (b) the probe length, (c) the piston position. This paper points out that there is, theoretically, an optimum piston position for greatest bandwidth, and presents some evidence corroborating this theory. Bandwidths greater than ±10 per cent to the 1 db swr points have been realized by fixing the piston at its optimum position and varying (a) and (b) above to obtain a match.

NELSON, R. A.,[1] AND W. BAND[4]
Vapor Pressure of He^3 = He^4 **Mixtures,** Letter to the Editor, Phys. Rev., **88**, p. 1431, Dec. 15, 1952.

---

[1] Bell Telephone Laboratories, Inc.
[2] American Telephone and Telegraph Company.
[4] State College of Washington.

OSBORNE, H. S.[2]

A Rose by Any Other Name, Report on Work of the Anglo-American Committee on Technical Terminology, Standardization, **24,** pp. 19–20, Jan., 1953.

PECK, D. S.[1]

Ten-Stage Cold-Cathode Stepping Tube, Elec. Eng., **71,** pp. 1136–1139, Dec., 1952 (Monograph 2054).

Developments in the art of transferring a gas discharge from one point to another in a multi-electrode tube have led to the design of a 10-stage counting tube operating up to about 2,000 pulses per second. Such a tube can be used for pulse counting, frequency division, time measurements, and similar functions.

PETERSON, G. E.[1]

Information-Bearing Elements of Speech, J. Acoust. Soc. Am., **24,** pp. 629–637, Nov., 1952.

This study deals with those aspects of speech which are phonetically significant. A technique has been developed with which phonetically equivalent speech samples may be obtained in different phonetic contexts and from different speakers. Data on two front vowels by different types of speakers are presented. The technique has also been applied to the evaluation of words containing these two vowels.

PIERCE, J. R.[1]

New Method of Calculating Microwave Noise in Electron Streams, I.R.E., Proc., **40,** pp. 1675–1680, Dec., 1952.

The noise in a temperature-limited electron beam in a drift space is calculated by a new means. Noise maxima and minima are found. The results agree with calculations made by the Rack-Llewellyn-Peterson method.

RICE, S. O.[1]

Statistical Fluctuations of Radio Field Strength Far Beyond the Horizon, I.R.E., Proc., **41,** pp. 274–281, Feb., 1953.

When a sinusoidal radio wave of extremely high frequency is sent out by a transmitter, the wave received far beyond the horizon is often observed to fluctuate. Here some of the statistical properties of this fluctuation are derived on the Booker-Gordon assumption; namely, that the received wave is the sum of many little waves produced when the transmitter beam strikes "scatterers" distributed in the troposphere. Expressions are obtained for the periods of the fluctuations in time, in space, and in frequency. These expressions extend closely related results obtained by Booker, Ratcliffe and others.

SCHIMPF, L. G., see R. L. Wallace, Jr.

---

[1] Bell Telephone Laboratories, Inc.
[2] American Telephone and Telegraph Company.

SMITH, K. D.[1]
Properties of Junction Transistors, Tele-Tech, **12,** pp. 76–78, Jan., 1953.

TOWNSEND, J. R.[5]
What We Have Learned in 1952, Standardization, **24,** pp. 16–18, Jan., 1953.

TOWNSEND, J. R.[5]
What We Have Learned in 1952, A Report to the Joint Meeting of Standards Council and Board of Directors of ASA, A.S.T.M. Bull., No. 187, pp. 22–23, Jan., 1953.

VAN ROOSBROECK, W.[1]
Large Current Amplifications in Filamentary Transistors, Letter to the Editor, J. Appl. Phys., **23,** pp. 1411–1412, Dec., 1952.

WALLACE, R. L., JR.,[1] L. G. SCHIMPF[1] AND E. DICKTEN[1]
High-Frequency Transistor Tetrode, Electronics, **26,** pp. 112–113, Jan., 1953.
Sine-wave oscillators at frequencies up to 130 mc and tuned amplifiers with substantial gain at frequencies of 50 mc or higher are obtained by using junction transistors with an added connection to the base electrode biased negative at six volts.

---

[1] Bell Telephone Laboratories, Inc.
[5] Sandia Corporation.

Contributors to this Issue

A. F. Bennett, Western Electric Company, 1914–25; Bell Telephone Laboratories, 1925–. Mr. Bennett, Director of Station Apparatus Development since 1948, is in charge of the development, design, field trials, and studies of telephone instruments and sets, coin collectors, telephone booths, and station systems. During World War II, Mr. Bennett supervised the development and engineering work on a number of ordnance items. For this he was awarded a Presidential Certificate of Merit in 1946. He was a representative of the office of Scientific Research and Development and in that capacity served in the United Kingdom in 1943. He is a member of the A.I.E.E., the Acoustical Society of America, and the Physical Society.

R. W. Burns, B.A., Indiana University, 1916; B.S. in E.E., Purdue University, 1918. U. S. Army, 1918; American Telephone and Telegraph Company, 1919–34; Bell Telephone Laboratories, 1934–. Mr. Burns has been engaged in the formulation of requirements for central office maintenance equipment, principally that used for testing exchange lines and trunks. During World War II he was in charge of a group preparing instruction books for teletype communication equipment for the armed forces. Professional Engineer, New York State. Member of Sigma XI and Eta Kappa Nu.

J. W. Dehn, E.E., Polytechnic Institute of Brooklyn, 1932. Western Electric Company, 1919–25; Bell Telephone Laboratories, 1925–. Switching Systems Development Engineer, 1952. Mr. Dehn has been principally concerned with the design and development of manual and dial telephone switching systems since joining the Bell System. During World War II he developed communication systems for the Signal Corps and trained military personnel in the operation and maintenance of this equipment. Since 1945 he has been engaged in No. 5 crossbar design. Professional Engineer, New York State.

R. F. Mallina, M.E., Vienna Technical College, 1912. London Institute of City and Guilds, 1914. Wurlitzer Piano Company, Acoustical Engineer, 1925. RCA Victor Company, Head of Apparatus Development Department, 1929. Bell Telephone Laboratories, 1929–. At the

Laboratories Mr. Mallina worked initially in acoustical research where he developed the first magnetic telephone message recorder and the five-reed telephone set. From 1936 on he was engaged in fundamental development on machine switching apparatus, first on AMA, later with the wire spring relay project. In connection with the latter, he developed the solderless wrapped connection. Mr. Mallina is also a research associate at New York University, Department of Education.

W. P. MASON, B.S. in E.E., University of Kansas, 1921; M.A., Ph.D., Columbia, 1928. Bell Telephone Laboratories, 1921–. Dr. Mason has been engated principally in investigating the properties and applications of piezoelectric crystals, in the study of ultrasonics, and in mechanics. Fellow of the American Physical Society, Acoustical Society of America and Institute of Radio Engineers and member of Sigma Xi and Tau Beta Pi.

J. W. McRAE, B.S. in E.E., University of British Columbia, 1933; M.S., California Institute of Technology, 1934; Ph.D., California Institute of Technology, 1937. Bell Telephone Laboratories, 1937–42, 1945–. U. S. Army 1942–45, where he attained the rank of Colonel and served as Deputy Director of the Engineering Division of the Signal Corps Engineering Laboratories. Returning to Bell Telephone Laboratories in 1945, Dr. McRae became Director of Radio Projects and Television Research in 1946; Director of Electronic and Television Research, 1947; Assistant Director of Apparatus Development I, 1949; Director of Apparatus Development I, 1949; Director of Transmission Development, 1949; Vice President, 1951. Legion of Merit, 1945. President of the Institute of Radio Engineers, 1953. Member of the A.I.E.E. and Sigma Xi.

T. F. OSMER, E.E, Polytechnic Institute of Brooklyn, 1935. Bell Telephone Laboratories, 1920–. As a member of the Physical Research Department until 1952, Mr. Osmer was primarily concerned with transducers, including transmitters, receivers, loudspeakers, and high quality microphones. During World War II he worked on military contracts, and since the war he has been occupied with carbon contact studies, and more recently with studies of the solderless wrapped connection, using photoelastic techniques.

R. C. PRIM, received a B.S.E.E. degree from the University of Texas in 1941, and M.A. and Ph.D. degrees from Princeton University in 1949. Following graduation from college, he was employed by General Electric Company in Schenectady until 1944, and then, as an ensign in the

Reserve, he joined the Naval Ordnance Laboratory at White Oak, Maryland. Here he conducted research on torpedo motion and control theory. He joined Bell Telephone Laboratories in 1949, and has conducted mathematical research on non-linear partial differential equations and served as a consultant on military projects. Dr. Prim is a member of the American Mathematical Society, the American Physical Society, Sigma Xi, and Tau Beta Pi.

E. D. REED, B.Sc., University of London, 1941; M.S., Columbia University, 1947. Ardente, Ltd., 1941–43; U. S. Army, 1944–46; Bell Telephone Laboratories, 1947–. Mr. Reed is engaged in the development and design of klystrons. Associate member of the Institute of Radio Engineers and member of Sigma Xi.

HARRY SUHL, B.Sc., University of Wales, 1943; Ph. D., Oriel College, University of Oxford, 1948. Admiralty Signal Establishment, 1943–46; Bell Telelphone Laboratories, 1948–. Dr. Suhl conducted research on the properties of germanium until 1950, when he became concerned with electron dynamics and solid state physics research. Member of the American Institute of Physics and the American Physical Society.

R. H. VAN HORN, B.S. in E.E., Pennsylvania State College, 1937; M.A., Columbia University, 1947. Bell Telephone Laboratories, 1937–. Mr. Van Horn is a member of the Switching Apparatus Development Department and is in charge of the machine switching apparatus laboratory. He has previously been engaged in the development of underwater sound devices and the vibrating reed selector for mobile radio applications. Member of A.I.E.E. and Acoustical Society of America.

E. F. VAAGE, E.E., Technical University of Darmstadt, 1926; M.E.E., Brooklyn Polytechnic Institute, 1932; Royal Norwegian Air Force, 1918–19; Elektrisk Bureau, 1926–27; A. T. & T. Co., 1927–34; Bell Telephone Laboratories, 1934–. Mr. Vaage is a member of a group engaged in systems studies, an outgrowth of his previous work of evaluating transmission systems. Member of American Mathematical Society and A.I.E.E.

A. S. WINDELER, B.S., Rutgers University, 1930; Bell Telephone Laboratories, 1930–. Mr. Windeler has been engaged in the design and development of toll cable, including coaxial, video pair, and microwave, types. He is currently in charge of a group concerned with the development of expanded polyethylene insulated conductors for multipair cable.

6505-1-22mw