

Additional Copies

Additional copies of 1954 Convention Record Parts may be purchased from The Institute of Radio Engineers, 1 East 79 Street, New York 21, N. Y., at the prices listed below.

<i>Part</i>	<i>Title</i>	<i>Sponsoring Groups</i>	<i>Prices for Members (M), Libraries (L), & Nonmembers (NM)</i>		
			<i>M</i>	<i>L</i>	<i>NM</i>
1	Antennas & Propagation	Antennas & Propagation	1.25	3.00	3.75
2	Circuit Theory	Circuit Theory	1.25	3.00	3.75
3	Electron Devices & Component Parts	Electron Devices Component Parts	1.50	3.60	4.50
4	Electronic Computers & Information Theory	Electronic Computers Information Theory	1.50	3.60	4.50
5	Aeronautical Electronics & Telemetry	Aeronautical & Navigational Electronics	1.50	3.60	4.50
6	Audio & Ultrasonics	Radio Telemetry & Remote Control Audio	1.50	3.60	4.50
7	Broadcasting & Television	Ultrasonics Engineering Broadcast Transmission Systems Broadcast & Television Receivers	1.50	3.60	4.50
8	Communications & Microwave	Communication Systems Microwave Theory & Techniques Vehicular Communications	1.50	3.60	4.50
9	Medical & Nuclear Electronics	Medical Electronics Nuclear Science	1.50	3.60	4.50
10	Instrumentation & Industrial Electronics	Instrumentation Industrial Electronics	1.25	3.00	3.75
11	Engineering Management & Quality Control	Engineering Management Quality Control	1.00	2.40	3.00
			15.25	36.60	45.75

Responsibility for the contents of papers published in the Convention Record of the I. R. E. rests solely upon the authors, and not upon the IRE or its members.

CONVENTION RECORD OF THE I.R.E.

1954 NATIONAL CONVENTION

LITVATY

JUL 7 1954

PART 3 - ELECTRON DEVICES AND COMPONENT PARTS

U. S. PATENT OFFICE

TABLE OF CONTENTS

Session 6: Electronic Components I - Techniques

(Sponsored by the Professional Group on Electronic Components.)

The Effect of Maintenance on Reliability of Complex Military Electronic Equipment	J.B. Arnold	3
Miniaturized Computer Applications of the Hughes Diode	S.G. Lutz	8
Subminiaturization Techniques for UHF Communication Equipment (Abstract)	Gustave Shapiro	16
Synthetic Quartz Crystals for the Electronic Industry	Danforth R. Hale and Wilfred H. Charbonnet	17
Application of Precise Components in Permeability Tuned Oscillators	David Hodgkin	22

Session 14: Electronic Components II - Application

(Sponsored by the Professional Group on Electronic Components.)

Magnetic-Core Delay Cables	Dimitri R. Stein	30
Improvements in the Field of Electrolytic Capacitors	Dietrich Altenpohl	35
An Investigation of Lowest Resonant Frequency in Commercially Available Bypass Capacitors	David T. Geiser	43
Resolution in Precision Wire-Wound Potentiometers	Robert J. Sullivan	48
Evaluation of Core Materials for Magnetic Amplifier Applications	R.D. Teasdale and H.R. Brownell	56

Session 25: Electron Devices I - Electron Tubes

(Sponsored by the Professional Group on Electron Devices.)

The Hollow Cathode in Cylindrical Geometry	B.D. Kumpfer and Herbert Brett	66
The Machining of Tungsten and its Application in the Fabrication of Philips Dispenser Cathodes	Roberto Levi	70
The GE Post Acceleration Color Tube (Abstract)	C.G. Lob	73
Amperex Type ELT Decade Counter Tube	Irwin Rudich	74
A Developmental Thyatron Capable of Current Interruption by Grid Action (Abstract)	E.O. Johnson, J.A. Olmstead, and W.M. Webster	79

Session 32: Electron Devices II - Transistors

(Sponsored by the Professional Group on Electron Devices.)

Transistors for High Power Application	John S. Saby	80
A New High Temperature Silicon Diode	C.G. Thornton and L.D. Hanley	84
Small-Signal Parameters of Grown-Junction Transistors at High Frequencies	R.L. Pritchard and W.N. Coffey	89
The Study and Design of Alloyed-Junction Transistors	L.J. Giacoletto	99
An Analytical Study of α , γ , and h Parameter Accuracies in Transistor Sweep Measurements	H.G. Follingstad	104

Session 40: Electron Devices III - Storage Tubes

(Sponsored by the Professional Group on Electron Devices.)

The Metrechon - A New Half Tone Picture Storage Tube	L. Pensak	117
Characteristics of Viewing Storage Tubes with Halftone Display (Abstract)	M. Knoll, H.O. Hook, and R.P. Stone	120
A High Writing Speed Dark Trace Tube	S. Nozick, N.H. Burton, and S. Newman	121
A Large Capacity Storage Tube for Digital Computer Applications	R.B. DeLano, Jr.	125
Noise Limitations on Storage Tube Operation	Stanley Winkler and Seymour Nozick	131

Session 47: Electron Devices IV - Microwave Tubes
(Sponsored by the Professional Group on Electron Devices.)

A Voltage-Tunable Magnetron for Operation in the Frequency Range 1500 to 3000 Megacycles. . .	Joseph A. Boyd	139
Control of Electron-Beam Spread by Positive Ion Traps	E.L. Ginzton and B.H. Wadia	145
The Multipactor Effect in Klystrons	Kees Bol	151
Backward-Wave Oscillator Characteristics (Abstract)	H.R. Johnson	156
The Propagation Properties of Cross-Wound Twin Helices Suitable for Traveling-Wave Tubes (Abstract)	M. Chodorow, E.L. Chu, and J.R. Nevins, Jr.	156

THE EFFECT OF MAINTENANCE
ON RELIABILITY OF COMPLEX MILITARY ELECTRONIC EQUIPMENT

J. B. Arnold
Aeronautical Radio, Inc.
Washington, D.C.

LIBRARY

JUL 7 1954

U. S. AIR FORCE OFFICE

Summary

From data collected by Aeronautical Radio, Inc. through field surveillance of electron tubes in military use, it can be shown that maintenance as generally performed today cannot cope with the extreme complexity of many present-day military equipments. Not only are some individual circuits difficult to analyze, but interdependency of associated circuits presents problems so intricate that diagnosis of equipment trouble cannot be readily accomplished. Wholesale replacement of tubes is often resorted to as an alternative remedy. Maintenance is further complicated by the precision adjustments necessary for successful operation of the equipment.

It can be demonstrated that these conditions not only cause unnecessary tube usage but also increase the probability of catastrophic tube failures. This, in turn, decreases equipment reliability. Typical examples are shown, and methods of alleviating the problems are suggested.

- - - - -

For the past 2½ years, Aeronautical Radio, Inc. (ARINC) has been studying the performance of electron tubes in military applications. During this time, data have been collected on more than 80,000 tubes removed from many types of equipment. One result of this investigation is the conclusion that maintenance affects both the quantities of tubes used and the reliability of electronic equipments.

The tubes under surveillance are installed in equipments operating in several types of military aircraft, in ships, and in field force and stationary installations. These equipments include communications sets of almost every type and frequency, radars of all types, navigational aids, and many special kinds of gear.

ARINC representatives are stationed at all bases where surveillance is in progress. Their job is to collect the tubes removed from sockets and to report

as much information on each removal as it is possible to obtain. These data are coded and tabulated, after which the tubes and the tabulation sheets are sent to ARINC headquarters in Washington. Here the information is summarized and analyzed, and laboratory tests are performed on certain tubes to determine the reasons for failure. However, most of the receiving type tubes are reshipped to Cornell University, where exhaustive tests are conducted to discover the nature of the failures. This work is directed by Professor Walter Jones of the Cornell School of Electrical Engineering.

Tubes such as magnetrons, klystrons, and transmitting types, which cannot be tested on normally available equipment, are returned to the respective manufacturers, who have the proper facilities for testing such types. Their assistance in the program is especially appreciated, for this cooperation represents a voluntary expenditure of their own time, money, and resources.

The surveillance program presents a good general view of military usage of electron tubes and the environments in which they operate. Analysis of the data concerning tube removals has resulted in several findings and conclusions, only one of which will be treated in this paper. This is the conclusion that maintenance of electronic equipments, as currently practiced by military technicians, affects not only the quantities of tubes required for replacement purposes but also the reliability of the equipments.

Like any other repairman, the military electronics technician must perform routine preventive maintenance as well as emergency repairs. Preventive maintenance of electronic equipments may be performed as a part of the general overhauling of the vehicle in which the equipment is installed, or it may be performed periodically as a separate operation. For example, maintenance of electronic equipment installed in aircraft occurs usually when planes are undergoing engine, propeller, or airframe checks. However, at fixed installations, such maintenance is usually carried out on a periodic basis

which may be hourly, daily, weekly, monthly, or even yearly -- or any combination of these. In any case, routine preventive maintenance may consist only of superficial testing of the fidelity of an equipment, or it may include a complete check of all tubes and other removable components.

Emergency Maintenance

Emergency repairs are made whenever the equipment fails during normal operations. In this connection, a failure is considered to be any reduction of performance below the specifications established for a particular equipment. These specifications may be either fixed or arbitrarily determined. Failures which require emergency repairs are of two general types. The first is the dangerous catastrophic failure which makes an equipment completely useless; the second is a general degradation of fidelity to a point where the equipment, although still operable, performs inadequately.

Catastrophic failures in some environments are generally more prevalent early in tube life, and consist of envelope, mechanical, and heater defects. On the other hand, the gradual depreciation of I_p , G_m and other electrical characteristics tends to increase as tubes age in operation. Data indicate that, in certain environments, catastrophic failures are much higher in the first 10 hours of tube life than during later periods of equal duration. Such a situation is shown in Fig. 1. In this illustration, the vertical axis represents the total percentage of tubes which can be expected to fail at a given time. The time is shown by the horizontal axis. Thus, it may be seen that 1% of the tubes in the test failed during the first 10 hours. Thereafter, throughout the first 600 hours of the test, the rate of failure, as demonstrated by the slope of the curve, was approximately 1% for each succeeding 50-hour period, with one exception. After 330 hours, 49 of the tubes were withdrawn from the test. The additional handling of these tubes raised the failure rate 1.4% at that time.

It should be noted that some of the tubes reported as removed during the first 10 hours actually failed at the time of installation. If such failures are not detected before the equipment is put back into service, they may cause delay by inducing additional trouble in the circuit. These subsequent failures are confusing to the technician as well as costly to the military services. On the basis of what has been said concerning failures which occur during the first

10 hours, the inference is clear: if a tube survives throughout that period, the probability of its continued performance in the same socket is greater than that of a new tube.

MORTALITY CURVE OF ELECTRON TUBES IN ONE ARINC CONTROLLED TEST

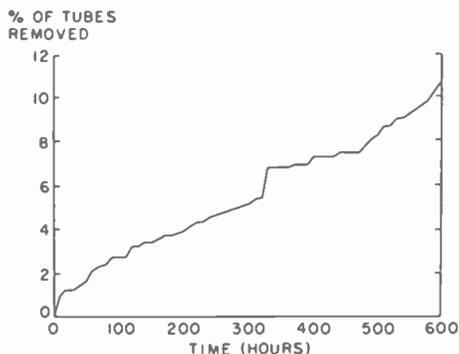


Figure 1

A third group of tubes removed by military technicians, in addition to those returned as catastrophic or as electrical failures, is made up of tubes which show no defect upon subsequent retest. As a matter of fact, these "no-defect" removals account for approximately 30% of all tubes returned to ARINC by the military. However, the percentages of such returns vary greatly from environment to environment. Fig. 2 demonstrates that 35% of the tubes removed from aircraft show no defect when retested. For ships, this figure is 30%; for mobile ground installations, it is 20%. These percentages should be contrasted with the 5% of no-defect returns from stationary installations. After retesting, many of these tubes have been reinstalled in the same sockets from which they were originally removed. Over 90% of the reinstalled tubes operated satisfactorily, and many have continued in service for several thousands of hours.

Accessibility of Equipment

The percentages of no-defect removals apparently depend in large part upon the accessibility and ease of maintenance of equipments. In aircraft, which furnish the highest proportion of no-defect returns, electronic equipment is sometimes almost inaccessible. Therefore, when making repairs, the technician quite naturally seeks ways to avoid future trouble in the particular equipment. In doing so, he often discards perfectly good tubes which would have operated

satisfactorily for many more hours. Such replacement is undesirable because it may reduce the reliability of the equipment. Furthermore, attempts to remove inaccessible equipments frequently subject them to greater shock than is ever experienced in operation. This shock definitely results in reduced tube life. To reduce the high percentage of no-defect returns from aircraft and to diminish the incidence of tube destruction through shock, equipment and airframe designers should cooperate to make electronic equipments as accessible as possible. The same problem exists to a somewhat lesser extent in shipborne and mobile equipments. In stationary installations, the greater accessibility of equipments is reflected in the low percentage of no-defect returns.

less reliable tubes in service, there are other and more important reasons why this practice is undesirable. When tubes are continually being removed and reinserted in sockets, the sockets become worn or broken and may lead to equipment failure during operation, thereby reducing equipment reliability. In addition, it has been found that the number of damaged tube envelopes is directly related to handling.

Masking of Basic Trouble

To sum up these points -- wholesale tube replacement may not only introduce a new trouble into the equipment but may increase the time required to find the original one. There is another consequence, however, that may be more serious: That is the masking of the true nature of the trouble by mass tube replacement or even by the indiscriminate replacement of a single tube. In cases such as these, the equipment soon will be back in the shop for repair. A few examples, as provided by the findings in the ARINC tube surveillance program, may be of interest.

A certain equipment of relatively new design has its electron tube filaments in series, and these in turn are in series with a thermal voltage regulator. The purpose of the regulator is to compensate for "above-normal" filament supply voltage. Several of these units were reported bad. The technicians removed the tubes, tested them, and found that several had low G_m while others had open filaments. The defective tubes were replaced and the equipments returned to operation. In a short time, the same units were back in the shop for repair. Again all the tubes were tested, the bad ones were replaced, and the equipments were released for operation. But once more they found their way back to the repair shop. When the tubes were tested for the third time it was discovered that the same trouble existed as before. Further, it was found that the actual cause of the equipment malfunction was a bad thermal relay which was not compensating for over-voltage. Since the equipment operated on the test bench after tube replacement, the technician had no reason to suspect that it would not work properly in the field for an indefinite period. Here was a situation where a system of adequate checks and test procedures would have avoided future trouble by immediately locating the defective part. In this case, the safety device -- the thermal relay -- actually reduced the equipment reliability because it was not "fail safe".

Aside from the fact that replacing or testing large quantities of tubes may decrease equipment reliability by placing

EFFECT OF ENVIRONMENT ON NO-DEFECT TUBE REMOVALS

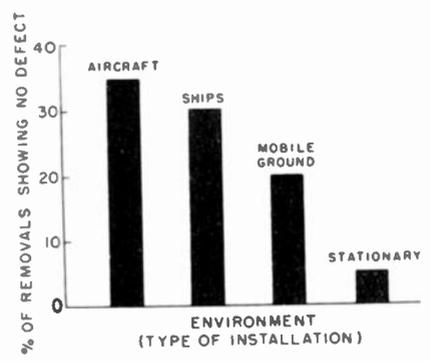


Figure 2

Another factor contributing to the removal of no-defect tubes is equipment complexity. Often a technician who does not understand the operation of an equipment will resort to replacement of all tubes in an entire chassis or equipment in order to remove one bad tube. It should be pointed out, however, that complexity is only a relative term. What is complex to a military technician may not be complex to a circuit design engineer, especially if he designed the particular circuit. It is not unusual to find that technical personnel are confused by simple circuits like cathode followers and multivibrators. Yet the same technicians are expected to service equipments involving phantastrons and microwave circuitry, which are complicated even for experienced engineers. The average period of "in-school" training for a military technician is about one year.

In another case, a tube was removed from a radar equipment because of an open cathode tab, and a new tube installed.

The equipment was returned to service. Two days later, the unit was again in need of repair, the same socket having produced the same trouble. The tube was replaced. The next day the equipment failed again for the same reason. This time 8 tubes failed successively in the same socket for the same reason before it was discovered that the tube had lost its bias due to a fault in the associated circuitry. This was a simple trouble which should have been found quite easily with a voltmeter. Yet, because the technician did not understand the function of the circuit, he relied on the simple expedient of tube replacement in an effort to solve his problems.

There are many more cases similar to these on record. They come from all branches of the military, and include such instances as three tube replacements in the final stage of a transmitter within one week because of a faulty antenna. In another case, all the tubes of one radar IF strip were replaced because of a faulty mixer crystal. In all of these situations, tube replacement temporarily remedied the trouble, and there was no outward indication of basic trouble that would lead an average technician to look further; nevertheless, equipment reliability was considerably reduced as a result. If maintenance personnel take time to make proper use of adequate test points, the basic trouble can be found and future failures avoided.

Poor alignment also causes good tubes to be removed. In some instances, equipment alignment is so complex an operation that rather than go through a procedure involving upwards of 50 critical adjustments, a technician will hand-select tubes which will allow a poorly aligned equipment to meet specifications. This is particularly likely to happen when an aircraft is being held on the ground awaiting the equipment. Unfortunately, the unit will be back in the shop for repairs as soon as the tubes drift slightly. Thus, the simple use of tubes as a cure-all is dangerous. It is useful to point out here that components other than tubes are also needlessly replaced. In one case, two transformers failed successively because of improper termination. Unnecessary tube replacement is more common, however, because tubes are easier to remove.

Effects on Reliability

Improper maintenance can have several detrimental effects on equipment reliability, in addition to causing excessive tube usage. Reliability can be reduced by damage to sockets and wiring,

which means that new trouble is added to the old. It can also be affected if maintenance, instead of curing the basic trouble, does nothing but temporarily alleviate its symptoms. This should not be interpreted as an effort to discourage the testing of tubes or proper replacement of malfunctioning tubes. Checking tubes in tube testers can eliminate such undesirable as shorts, opens, defective heaters and the like. Nevertheless, the most effective device for testing tubes is the equipment itself.

Complexity and Education

It is generally agreed that there is too large a gap between the educational level of the technician and the complexity of the equipment. It should be emphasized that this situation can be attributed only in small part to the shortcomings of the technicians. Further, it is agreed that complexity is here to stay. As human reflexes in more and more cases are incapable of coping with the speed of such devices as jets, rockets, and 250-horsepower autos, additional electronic circuits are required. Increased safety measures also increase complexity; however, it seems only reasonable that safety measures should not be of such a nature as to decrease the reliability of the equipment.

One method of reducing the gap between equipment complexity and the educational level of the technician is by lengthening the time he spends in the electronics school, and -- as a corollary -- by being increasingly selective in the choice of prospective technicians. However, since the average first enlistment of a military technician is for a period of four years, as much time as possible must be spent in performance of his primary mission, which is maintaining equipment. Further, with the shortage of civilian maintenance personnel, good military technicians are sometimes pirated away from the services. Thus, the amount of increased education which is permissible is in part dependent upon an economic factor. Whatever the length of the schooling period, however, it seems beyond question that the training program should make reliability of equipment a primary goal, rather than stressing speed in repair.

It might be suggested that one possible way of closing this gap between education and equipment complexity is to make the equipments more simple. As we have previously mentioned, however, this is almost impossible. It may be simpler to decrease the amount of technical knowledge required by the technicians. I should

like to outline some suggested methods by which this could be accomplished.

Removable Sub-chassis

If complex circuits were placed in easily removed sub-chassis, their removal and installation could be left to relatively unskilled personnel, with actual repairs being performed only by a relatively few, highly-trained technicians. There are many factors that will determine the proper number of separate circuits to be included in one sub-chassis. It should be convenient, for example, to combine all circuits which perform one function, such as an entire IF strip or a timing chain in a radar. Of course, unless the number of tubes is relatively small or the reliability of each tube is very high, the reliability of each sub-chassis will be relatively low. On the other hand, a unit which will absolutely not fail within two years would be uneconomical if it became obsolete before that time. The optimum number of circuits which can be combined into one subchassis depends on these factors, as well as practicability and environment.

The use of sub-chassis could be improved by a system of indicators which would show a bad sub-chassis. This might take the form of a light indicator which would light up when the sub-chassis failed to perform its desired function, or perhaps an oscilloscope built in so that by going through a series of switch positions the faulty assembly could be easily identified. Admittedly, this would increase the size and weight of equipments which, in some instances, are already excessively bulky. A more practical solution might be to bring out a series of easily accessible test points so that by means of an external and port-

able test set, the faulty sub-chassis could be identified.

These are methods of determining where the trouble is after it happens. It would be far preferable if failures could be predicted and prevented before they actually occur. It has been found in our surveillance program that well over one-half, and perhaps as many as three-fourths, of all tube failures are predictable. Thus, a system of marginal checks should be installed so that at intervals the equipment could be tested and probable future tube failures removed. To a degree, this is what is being attempted when a maintenance routine calls for 100% test of all tubes on a military tube tester. In reality, an equipment is its own best tester of tubes. Marginal checks have been used with a considerable degree of success by manufacturers of computers. Marginal checks combined with individual sub-chassis failure indicators would perhaps constitute the best method for improving reliability of military equipment. However, none of the various methods suggested would be very effective unless the manufacturer of the equipment were aware of the problems confronting the military maintenance man. On that point I would like to call attention to a proposal made by Mr. L. M. Clement in a paper presented at the Radio Fall meeting of the RETMA in Toronto, Canada. His suggestion was that the manufacturer's project engineer assigned to a particular equipment be sent out to the field so that he can educate himself on the use and environment of the equipment, thereby decreasing the burden of the maintenance man and the operator. This would be a forward step offering much hope of greater reliability in military electronics.

MINIATURIZED COMPUTER APPLICATIONS OF THE HUGHES DIODE

S. G. Lutz

Hughes Aircraft Company
Culver City, California

Introduction

Extremely small components, such as the grain-of-wheat-size Hughes diode, present a challenge to the engineer who wishes to take full advantage of their small size. Reducing the size of components does not generally lead to a corresponding size reduction of the equipment in which they are employed. There is a natural tendency to mount and wire such components in the same manner as large components and this results in the mounting hardware and the wiring consuming a disproportionate share of the volume. One reason for the wiring occupying so much of the volume is that connections cannot safely be soldered closer to small components than to large ones. Soldering closer than about 1/4-inch from the diodes introduces the risk of effecting a permanent change in diode characteristics if the crystal is heated above about 150°C.

Realizing good size-reduction from the use of miniature components requires breaking away from conventional practices to the extent necessary to miniaturize the wiring along with the components. Etched wiring was a step in this direction, but it yet involves soldering and its 2-dimensional inflexibility often results in waste space.

Spot-welding offers an attractive means of making connections very near the components, because capacitor-discharge welds are made in milliseconds, with much less heat than required for a soldered joint. Mounting axial-lead components on end, rather than laying them flat on etched-wiring boards or on terminal strips, also aids in miniaturizing the wiring by shortening the interconnections.

The Hughes Diode

Figure 1 shows the general construction and nominal dimensions of the Hughes germanium diode. Its glass envelope provides perfect humidity protection and is so small that, without leads, over 400 of them could be packed into a cubic inch. The leads are of dumet (copper-clad nickel), chosen because its coefficient of expansion and that of the glass are essentially identical, and these leads are tinned to facilitate soldering. As an extra advantage, the nickel core in this wire makes it much easier to spot-weld than would be the case with a solid copper wire.

If we follow the conservative directions of the manufacturer by leaving 1/4-inch between the

solder joint and the body of the diode, the mounting centers must be spaced about three times the length of the diode body as a minimum. Closer spacing can be used at the risk of diode damage unless very rapid soldering is possible on each joint. Contrasted to this, welded connections can be made as close to the seal as desired, provided that a capacitor-discharge welder is used properly.

Diode Gate Matrices using Etched Wiring

Large numbers of these diodes, along with a lesser number of resistors, are employed in the "and" and "or" gates of digital computers and the matrices composed of these gate circuits account for much of the wiring complexity of such a computer and often account for an appreciable portion of its volume. Figure 2 shows a typical matrix employing etched wiring. Some space was wasted deliberately by employing a "universal" etched-wiring pattern, so that the same basic pattern could be used for many different matrices, but its density of 13 components per cubic inch is not much below the maximum generally obtainable with etched-wiring. For the benefit of those who still think in terms of larger volumes, this is 22,464 components per cubic foot. From a more pessimistic viewpoint, though, three hundred times as many components the size of these diodes could be packed into the same volume, if their leads could be clipped off. This forms a good example of the deterioration in volumetric efficiency, or ratio of component volume to total volume, which generally is encountered as the size of components is reduced without miniaturizing the wiring along with the components.

Cemented Matrix with Spot-Welded Wiring

One good approach toward increasing the volumetric efficiency of a unit always is that of fitting the required components together as closely as possible, then seeing how little additional volume can be used for the wiring. With computer matrices, one naturally tends to pack the diodes and their associated resistors into a rectangular block, with the leads extending in opposite directions. In figuring how best to wire such a structure, one recognizes that a characteristic of gate circuits is that they have a common connection between the resistor and all its diodes, so it is logical to make these common connections on top of the block, to make ties between gates having the same inputs at the bottom of the block and to employ the remaining bottom leads for the

external connections.

Figure 3 is a matrix built in the manner just indicated, as a test to establish the maximum component density, to investigate the feasibility of spot-welded wiring in such cramped quarters and to demonstrate the short leads required when such a unit is wired from both ends in this manner. This matrix contains 32 diodes, 9 Global resistors and 4 vacant positions where diodes could have been placed if required. The unit occupies only 1/4 cubic inch, measured over the wiring but excluding the input and output leads which normally would connect to a plug. Thus, the maximum component density with this construction is 180 per cubic inch, which is believed to set a record for units composed entirely of standard components.

The cemented construction of this matrix admittedly is impractical, if for no other reason than that the entire unit would have to be discarded if one component failed, and it contains more diodes than one would wish to discard. This unit has served a useful purpose, however, in stimulating further work and in measuring subsequent designs in terms of their advantages versus their cost in decreased component density.

Stick-Capsuled Matrix

Figure 4 shows a subsequent matrix design directed toward achieving ease of production. The diodes are molded in long sticks, notched between each pair of diodes so that the desired number of diodes can be broken off, much like tearing off postage stamps. This feature permits the assemblers to handle the diodes in convenient groups of five or less. In the matrix shown, the diodes were molded in epoxy resin but tests have established that they can withstand the temperature and pressure required for alkyd molding; the latter being far more rapid.

With this construction, etched wiring is used for the gate interconnections and for the input and output connections, while the common top connections can be made by a multiple-welding operation prior to inserting the diode sticks into the etched-wiring card. Top connections to the resistors have been soldered, because of the uncertainty of welding to their solid copper leads. These connections also could be welded if resistors with weldable leads were available. All connections to the bottom board are dip-soldered simultaneously. Though there is but an eighth of an inch between the glass seal and this soldered joint, the thermal shunting action of the plastic and the low temperature of the soldering bath prevent damage to the diodes.

In the event of a diode failure, the stick of diodes containing the defective one is unsoldered and replaced. As many as four good diodes may be discarded with the bad one, but their cost generally will be less than the labor cost of this replacement. If desired, the defective diode may be broken out of its stick and be replaced. A special soldering tool has been developed for removing and replacing sticks of components and consists of a 200 watt iron with a long grooved tip.

With this construction, the spacing between components is 1/4 inch; this spacing being determined by the minimum clearances required for the etched wiring. The height of the unit is 9/16 inch, so its peak component density is 30 per cubic inch. Thus, these units are twice as compact as the flat etched board units of Figure 1, as well as being much cheaper to assemble. Compared with the cemented matrix of Figure 2, this repairability and ease of assembly have been purchased at the cost of a six-time reduction in component density.

Honeycomb Technique

In searching for a way to more of the high component-density of the cemented matrix, while retaining repairability and ease of assembly, it was recognized that the diodes and associated components could be cemented into a honeycomb of holes in a supporting board and that individual components could be replaced by softening the adhesive with a suitable solvent. Preliminary tests showed that the diodes could be spaced 1/8 inch on centers, or 64 per square inch of mounting board, and that the wiring could be spot-welded easily despite this close spacing. No difficulty was experienced in removing and replacing components other than that it was necessary to clip the leads obstructing component removal and later to weld around the clipped-out section. The diode leads were cut to a length of 1/8 inch from the seal, thus holding their over-all length just under 1/2 inch, so that the peak component density was 128 per cubic inch, or 71% of the density obtained with the cemented matrix construction.

At this stage, it was decided that the honeycomb technique should be tried on certain large and typical sections of the computer gate circuits. The sections selected were the reading gate circuits for the magnetic memory drum and the gate circuits for a conversion control unit. Fully-wired but non-operating units were decided upon, since these units were to be built as a packaging study only, and since this saved several thousand dollars worth of good diodes.

Reading Gate Matrix Unit

The memory reading amplifier gates were ideally suited for this honeycomb technique because of their nearly repetitive nature and simple binary coding under the control of a small number of flip-flops, leading to simple and systematic wiring.

The Resistor Problem

One problem encountered was the need for a resistor suitable for use with this honeycomb mounting technique. Axial leads and dimensions comparable to those of the Hughes diode were the chief requirements, with weldable leads being desirable. The Global resistors used with the earlier cemented matrix were ruled out by these wrapped-around radial leads, which prevented their insertion in the mounting holes. Figure 5 shows a comparison between the sizes of the diode and of several available resistors and further indicates why none of them were suitable. Axial leads could have been attached to Global resistor units, if a working unit had been required, but this seemed more tedious and expensive than justifiable for a packaging study; consequently, dummy resistors were fabricated from 3/32-inch wood dowel rod, using 0.020-inch nickel wire for the axial leads. There is no apparent reason why resistors of this size cannot be produced commercially whenever the demand for them justifies their development and there seems to be a steadily increasing need for such resistors for transistor circuits and other such applications in which a tenth-watt dissipation would be adequate.

Microfilmed Wiring Information

A feature which has contributed greatly to the ease of wiring these very small honeycomb units has consisted of microfilming the essential wiring information on the dural mounting plates. These plates were coated with a white lacquer, over which a photographic emulsion was applied. This emulsion was printed in contact with a negative obtained by photographing a large ink drawing to the desired scale. Incidentally, the use of a 1/8-inch dural mounting plate provides excellent heat equalization and dissipation and leads to extremely rugged units.

Figure 6 is an enlarged view of a small section of the wiring and drilling pattern for the reading matrix unit. The regularly spaced crosses and white centers of black squares mark the centers of the component holes to be drilled. The heavy lines between crosses denote buss connections. The rows of black and white squares denote rows of diodes to be connected to either the Q or \bar{Q} outputs of the same flip-flop; the particular flip-flop

being designated by the number beside each such row. Similar black and white squares also are used to indicate the connection of resistors to positive or negative bias voltages. Q leads and positive voltage leads are insulated with white vinyl sleeves, while the alternate \bar{Q} and negative voltage leads have black sleeves. Thus, most of the wiring consists of running black wires to black squares and white wires to white squares, and of bussing common connections as indicated. The remaining wiring, cabled connections to the plugs, is designated by correspondence of numbers or letters at the plug and at the circuit. This use of black and white squares and leads simplifies inspection, since a white lead to a black square is conspicuous. Figure 7 shows top and bottom wiring patterns for the complete unit, reproduced actual size.

Spot-Welded Wiring

A Unimatic capacitor-discharge welder with extension tweezer electrodes was employed for the wiring, as shown in Figure 8. These tweezers are unusual in that their points are brought together by a spring, so that they serve as a clamp to hold the wires being welded, and so that the pressure always is the same. The operator squeezes the tweezer handles to open the points and merely releases pressure on the handles to let the points close on the work. These tweezer electrodes were found more convenient and produced more consistent welds than the normally-open type of tweezers with which the welding pressure depends upon the operator's grip.

Much of the welder's energy is dissipated in the long flexible leads to the tweezer electrodes but this is inconsequential because so little energy is required to weld these fine dumet and nickel wires. Lower heat settings would be used to weld the same wires between electrodes attached to the welder. Nickel wire of 0.010 inch diameter was used for all connections. One need not worry about nickel being a critical material because all the wire weighed about as much as a 5-cent piece, and iron or stainless steel wire could have been used nearly as easily if it had been available. One generally need not worry about the resistance of these fine nickel wires either, for it amounts to about 0.050 ohm per inch and even an extra ohm in a gate circuit would be inconsequential.

Welded connections can be made without danger as close to the glass seals as desired. Even though the metal is fused at the weld, the capacitor discharge is so rapid that the quantity of heat liberated is negligible. One can hold the wires and the tips of the welding tweezers between their fingers while making a weld, and not be able

to feel any heat. Under certain conditions it is possible for "sneak" welding currents to damage the diodes but this cannot occur if the welds are made in the proper sequence, if the welding voltage is low and if the wires being welded make good enough initial contact to prevent the development of open-circuit voltages.

Welded wiring of this type has been done by three engineers and by two technicians, none of whom had any prior welding experience. Following brief instruction, none of them experienced significant trouble in producing good welds consistently. Occasional temporary difficulties arose from accidentally or unintelligently changing the welder adjustments, letting the electrodes get dirty or pitted and not keeping the work clean. One string of bad welds was traced to having brushed the work with acetone to remove excess cement and thereby depositing a film of the cement on the leads. The worst handicap to widespread adoption of spot-welded wiring, other than the critical adjustments required to weld the copper leads of most present components, is the fact that unfamiliarity breeds distrust. An engineer who had grown up with spot-welded wiring would have even greater distrust of soldered wiring, and rightly so. He would point to the danger of cold-joints, burned components and insulation, tediousness, lack of strength, danger from corrosive fluxes and all the other pitfalls of soldering which we have learned to live with.

Statistics Concerning the Reading Gate Matrix Unit

Top and bottom views of the wired unit are shown in Figure 9. The simplicity and relative "cleanliness" of the wiring is evident. One would not notice anything remarkable about the wiring of these units if one did not recognize the small size of the components and thought that the photographs were reductions of switchboard-size equipment with heavy wires and soldered connections. Such an impression is not surprising because it results from the wiring having been scaled down in proportion to the components.

The dimensions of the unit, exclusive of the plugs and frame, are $5\text{-}1/4 \times 3\text{-}1/8 \times 1/2$ inches. The unit contains 504 diodes and 209 resistors, so its average density is 94.5 components per cubic inch. Schematic diagrams for the unit were on five sheets, totalling 26 square feet. The complete unit weighs only 160 grams, or just over a quarter-pound, plugs and frame included. This, incidentally, is appreciably less than the weight of its five schematics!

The unit is amazingly rugged and may be dropped without damage because the plugs and the

aluminum frame which surround the dural plate serve to protect the components and leads. Individual components have been removed and replaced many times.

"Cribbage-Board" Conversion Control Unit

The conversion control matrices afforded a better test of the applicability and versatility of this honeycomb packaging technique because their circuits were more diversified, involved connections to many more flip-flops and cathode followers and were more typical of the arithmetic and control matrices of a computer. The dimensions of this unit are $8\text{-}3/8 \times 2\text{-}1/4 \times 1/2$ inches and the appearance of its perforated plate prior to inserting the components caused it to be dubbed a "cribbage-board". The unit contains 587 diodes, simulated resistors and ceramic capacitors, so its average density is 62.3 components per cubic inch. Top and bottom view of the wired unit, in Figure 10, show that this reduced density resulted from separating the individual matrices and from leaving space down the center for the large interconnecting cable. Though its wiring looks complicated, no more time need be required to wire such a unit than to wire a standard-size unit of equal complexity. Spot-welded wiring actually should be faster and cheaper than soldered wiring.

Conclusions

High component densities and good volumetric efficiencies have been obtained in computer gate matrices employing Hughes diodes and other correspondingly small axial-lead components by mounting the components on end to provide two wiring planes instead of one, and by spot-welding the connections to permit miniaturizing the wiring in proportion to the components. Spot-welding is an ideal method for making connections close to components which can be damaged by heat, and becomes a simple and non-critical technique when the component leads and wire are nickel, or other readily weldable material. Copper-to-copper welding is possible but the welding adjustments are apt to be objectionably critical.

Many variations of the above technique have been proposed and some have been tried. In general, variations which tend to simplify the assembly and wiring also tend to reduce the component density correspondingly, so the technique to be chosen for any specific application will depend upon what one is willing to pay in terms of size and weight for the simplification desired.

As long as vacuum tubes, rather than transistors, are used as the active circuit elements in digital computers, there is little to be gained from miniaturizing the diode matrices to the extent made possible by these techniques. Honeycomb component mounting, spotwelded wiring and micro-filmed wiring information should become useful techniques as soon as it becomes feasible to miniaturize other parts of digital computers by replacing tubes with transistors and by replacing magnetic drums with more compact memory devices. When component densities of the order of 100 per cubic inch become applicable through digital computers, it should become possible to reduce their volume and weight to that of present mechanical desk calculators.

Summary

Reducing the size of components has not generally led to corresponding size reduction of equipment, through failure to miniaturize the wiring proportionately. "3-dimensional" packaging, spot-welded wiring and similar techniques make it possible to maintain as high a volumetric efficiency in subminiature equipment as in standard-size equipment. Component densities as great as 180 per cubic inch have been attained in computer gate matrices using Hughes diodes and average densities approaching 100 components per cubic inch are practicable. Cementing components into holes in a metal plate provides excellent heat distribution, great ruggedness, and permits removal and replacement of individual components. Spot-welding with a capacitor-discharge welder and tweezer electrodes prevents thermal damage to miniature components, despite extremely close connections. Smaller resistors and other components with weldable axial leads are needed. Micro-filming wiring information on both faces of this plate simplifies wiring and inspection. These techniques should assume great usefulness in future transistorized computers.

Acknowledgements

Many members of the Hughes Miniaturization Group contributed to this work but mention should be made of Carmen Livesay, shown in Figure 8, who built and wired the first reading gate unit.

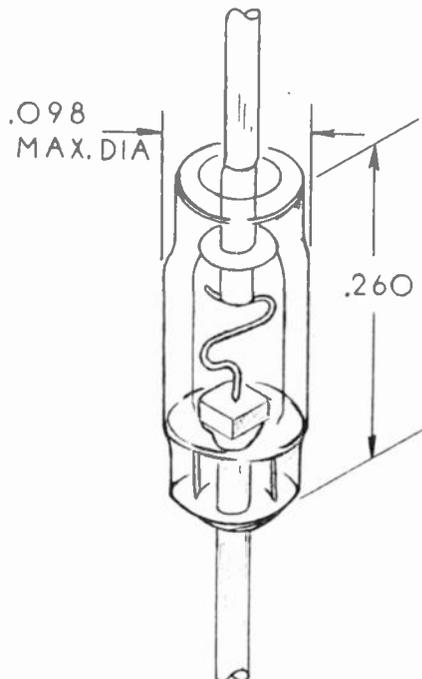


Fig. 1 - The Hughes germanium diode.

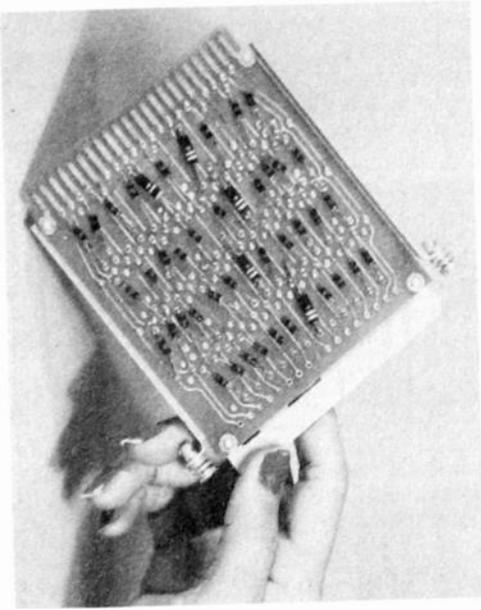


Fig. 2 - Diode matrix on etched wiring board.

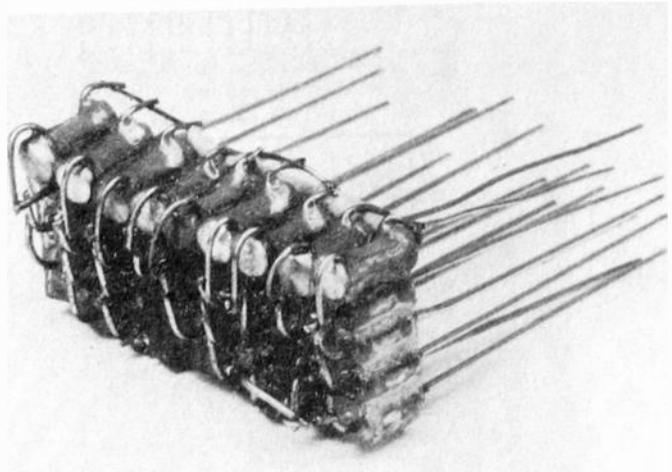


Fig. 3 - Cemented spot-welded matrix.

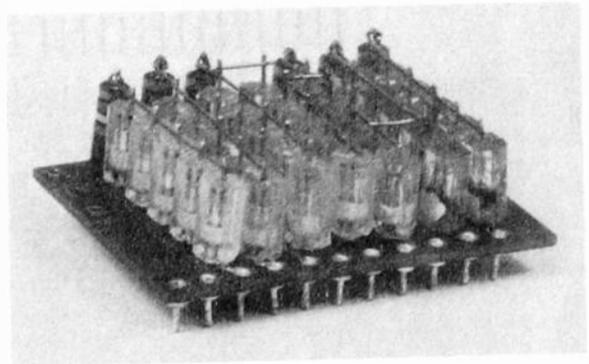


Fig. 4 - Stick-capsuled diode matrix.



Fig. 5
Size comparison between resistors and Hughes diodes.

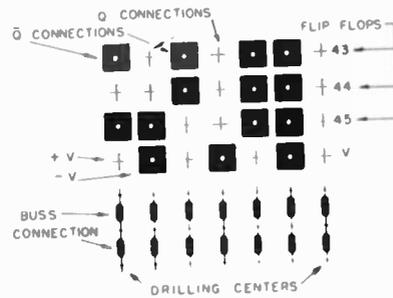


Fig. 6
Methods used for recording wiring information.

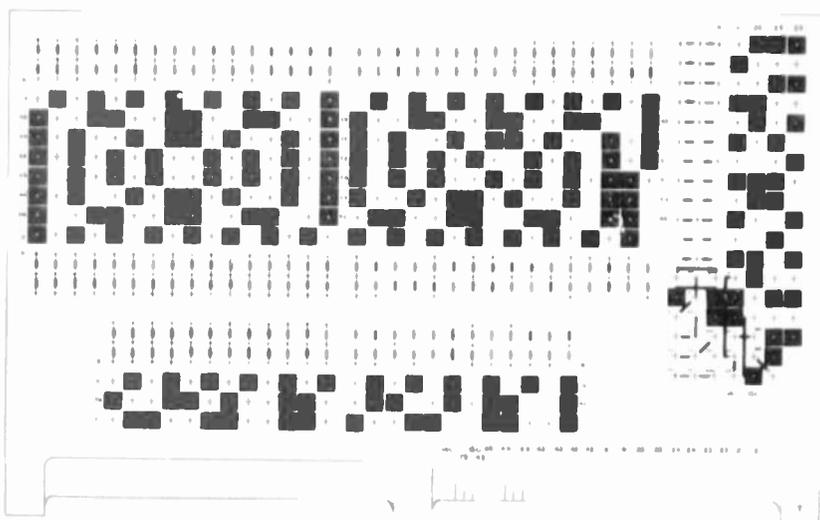
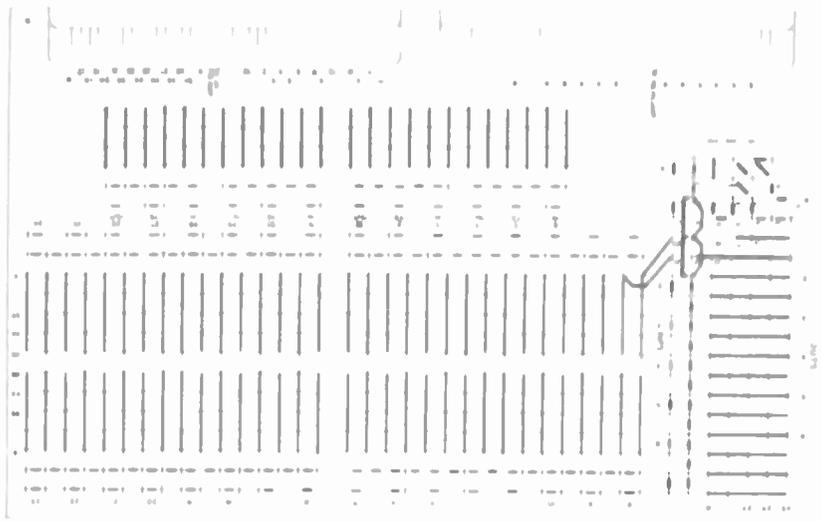


Fig. 7 - Full scale top and bottom patterns for reading head matrix board.



Fig. 8
Tweezer spot-welding electrodes in use.

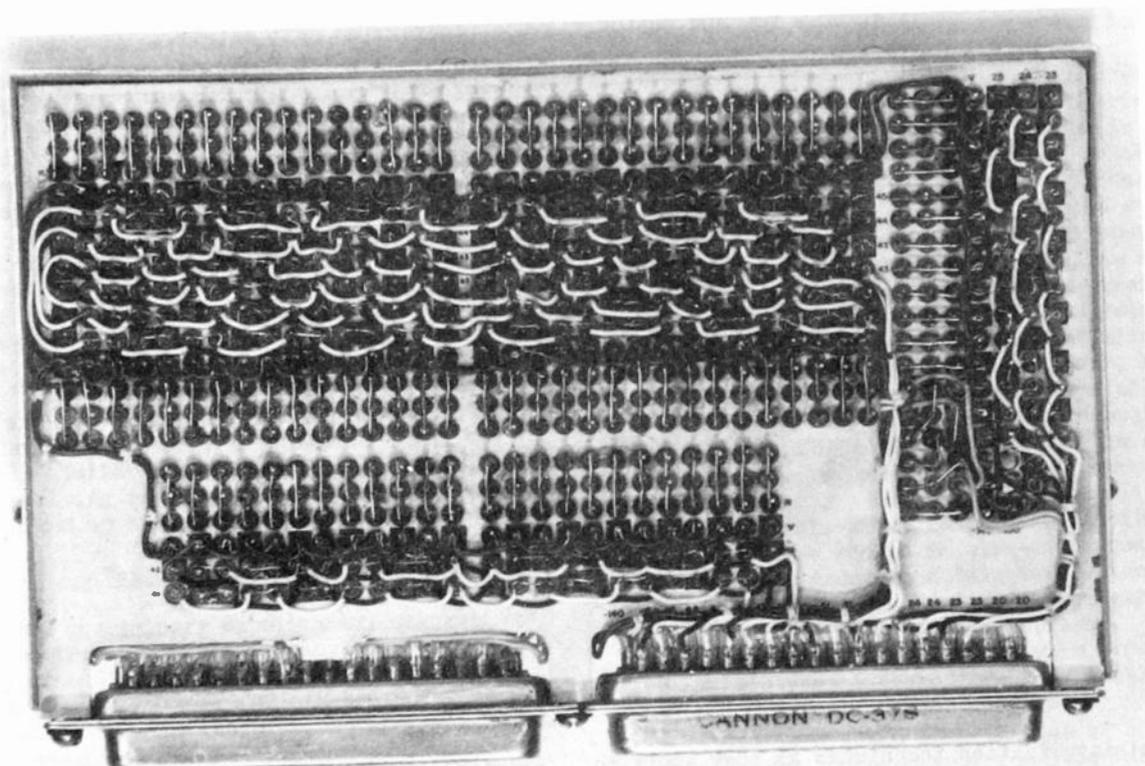
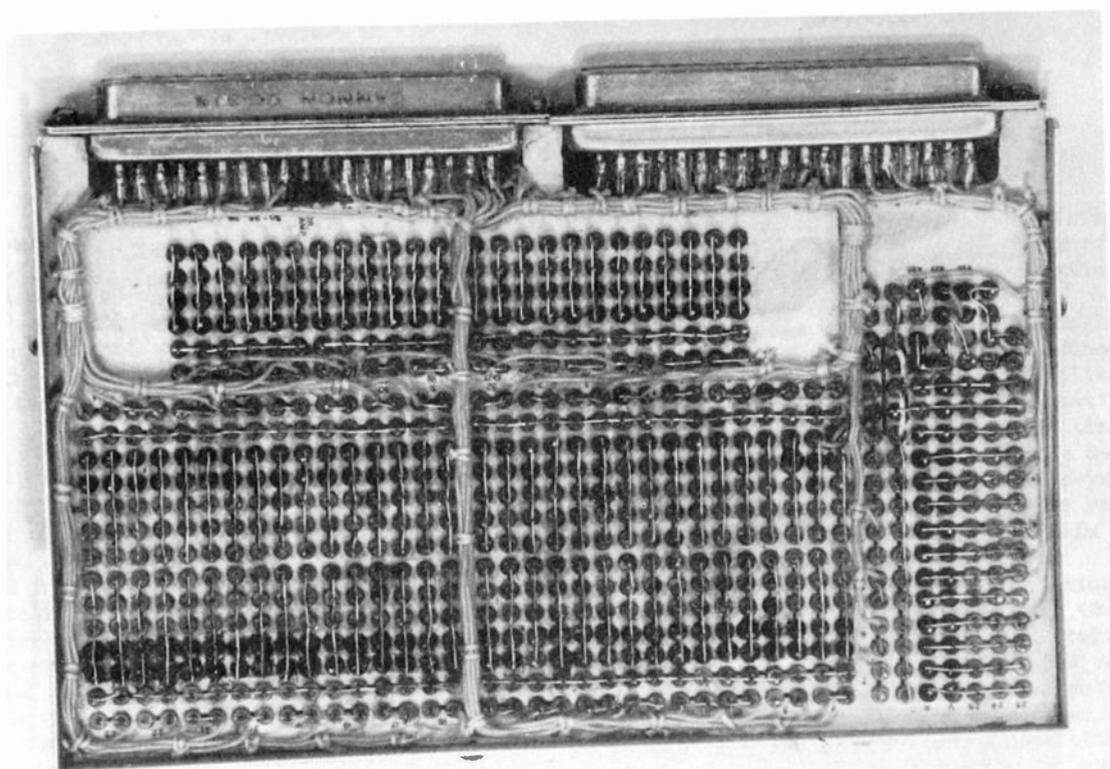


Fig. 9 - Top and bottom views of finished reading-head matrix unit.

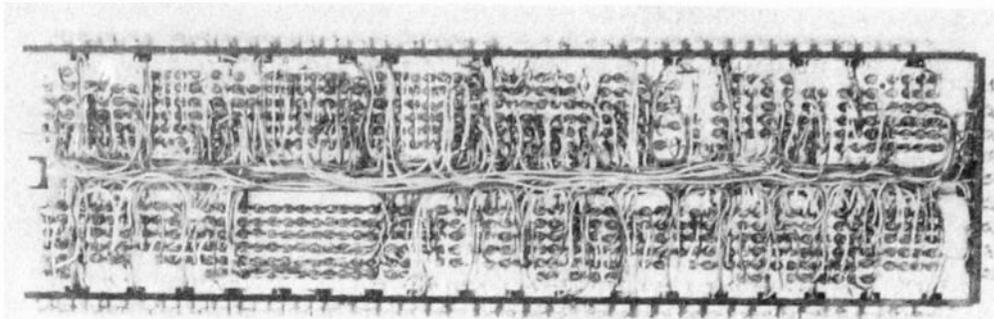
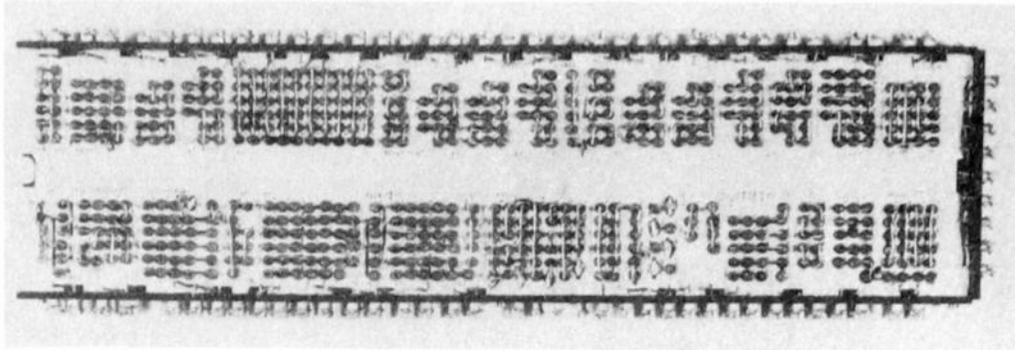


Fig. 10 - Top and bottom views of conversion unit matrix.

SUBMINIATURIZATION TECHNIQUES FOR UHF COMMUNICATION EQUIPMENT

Gustave Shapiro
National Bureau of Standards
Washington, D.C.

ABSTRACT

Subminiaturization techniques as they apply to uhf transmitting and receiving equipment are described. A combination chassis-plug system permits convenient air cooling and interconnection (dc, ac, and mechanical) of its plug-in assemblies. Where their application permits a production advantage,

printed circuits are used. Novel time-saving design and fabrication methods are disclosed. Special miniature components of wide utility are used in the small assemblies. Many of the techniques described can be applied to other subminiature electronic assemblies.

SYNTHETIC QUARTZ CRYSTALS FOR THE ELECTRONIC INDUSTRY

Danforth R. Hale and Wilfred H. Charbonnet
The Brush Laboratories Company
Division of Clevite Corporation
Cleveland 14, Ohio

Summary

Synthetic crystals of radioquartz quality, showing their natural faces, and well-sized and well-shaped for use by oscillator unit manufacturers, are being grown on an experimental scale under contract with the U.S. Signal Corps. Study of the piezoelectric properties shows that the material is the equal of natural radiograde quartz.

Introduction

Millions of quartz oscillator plates are manufactured yearly for radio frequency stabilization and for frequency selection in military and commercial radio communication equipment. These plates, known to the trade by the informal name of "quartz crystals", are cut from crystalline masses of quartz, usually imported from Brazil. In this paper we are discussing the growth of these crystalline masses themselves, which are quartz crystals in the mineralogical sense. (Fig. 1)

Frequency control by quartz crystal plates had become so important for communications that the Germans in World War II studied various substitutes for quartz, as well as the synthesis of this crystal¹ for their access to radiograde quartz (almost synonymous with Brazilian quartz) was very limited. Our own Signal Corps made some study of substitute materials² during the War and after its close, let five contracts for the investigation of the synthesis of piezoelectric crystals. Two of the contracts were directed at quartz growing, and one of these latter went to the Brush Development Company. The work to be described is the result of consecutive contracts with the U.S. Signal Corps over the past seven years.

The Bell Telephone Laboratories also are engaged in exploring the field and recently have been under contract with the Signal Corps; the process is similar in many respects to ours except involving pressures about double those we have regarded as the most satisfactory.³

Early Experiments

In the preliminary experiments, suggestions were drawn from various mineralogists, and especially from Giorgio Spezia,⁴ who in 1905-1908 heated a pressure-vessel for six-month periods with fluctuating city gas pressure in Turin, Italy. He added several grams to a number of crystals and thus proved definitely that large crystals could be grown.

During the war years in Germany, Richard Nacken, another mineralogist, attempted to grow quartz crystals, and came up with a process involving vitreous silica as raw material which

is more soluble than crystalline quartz in an aqueous solution held at constant temperature. He obtained some fundamental information and grew a number of small crystals.

His process was also tried in England and in this country, but has now been abandoned. It was found that the supersaturation of silica increases to so great a value that a large quantity of spontaneous nucleation occurs, resulting in crusts of fine quartz crystals being deposited on the walls of the vessel, and on crystal supports. Further, the vitreous silica reverts to the crystalline form, in place, within a few hours and crystal growth stops.

Two methods using cooling techniques for growing crystals from solution are: the bulk cooling method, in which a hot saturated solution provided with seeds is slowly cooled, and the continuous flow method, in which a solution is brought to a relatively hot saturated condition in one tank and continuously circulated between this and a crystallizing tank at a lower temperature.

Quartz does not have sufficient solubility, even at elevated temperature and pressure and in an alkaline solution, for the application of the direct bulk cooling method.

The continuous circulation method was therefore developed. Quartz chunks for nutrient material and seed crystals are placed in the same pressure-vessel and conditions so arranged that the nutrient quartz dissolves in one region and crystallizes out in another region over a period of days or weeks. Successful operating conditions were found to be an alkaline solution at temperatures of 570 to 750°F and pressures of 5000 to 10,000 psi. Such conditions require the use of heavy wall vessels of special alloy steel. Recent experiments have shown, however, that even at 480°F and 1200 psi radiograde crystals can be grown although the rate of growth is small.

Our early experiments were made in steel pressure-vessels (or autoclaves) of about 70 cc volume. (Fig. 2) For studying solubilities in various solvents under isothermal conditions the autoclaves were rocked in wire-wound heating furnaces. In order to obtain temperature gradients, as used in many quartz-growing experiments, they were stood upright on hot plates. One was heated by a coil of resistance wire around the top; others were operated lying on their sides with heating units around one end.

It was soon found that the use of a temperature gradient produced good results. Another type of autoclave was then designed in order to obtain as complete a separation as possible of the growing region and the dissolving region, that is to say, of the cooler and hotter regions. Two heavy-wall tubes are joined at the ends by

small cross pipes. The chambers are provided with electrical resistance heaters each with its own temperature control. One chamber is filled with the raw material - broken chunks of crystal - and the other is provided with a metal rack holding seed plates. (Fig. 3).

Such a vessel was first operated with the tube-chambers vertical. The circulation was found difficult to control and much better results were obtained by mounting the autoclave horizontal and slowly rocking it. (Fig. 4)

In current operation the vessel, charged with nutrient quartz and seeds, is filled about 2/3 full with 18% aqueous sodium carbonate solution, and closed with special pressure seals. It is made to oscillate about a horizontal axis through a displacement of about 20°, three times a minute. It is heated to operating temperature, the crystallizing chamber at about 660°F and the dissolving chamber at about 50°F higher. At operating temperature and pressure the autoclave contains no liquid-vapor interface, and this rocking motion produces a small reciprocating flow of solution from one chamber to the other because of the density difference, which in turn is due to the temperature difference.

By late 1950 the two-chamber rocking autoclave was designed, and two were in operation. We were getting a growth rate of 0.4 mm/day measured perpendicular to the minor r* face, or about 3 lbs. in a 5-week run - in more basic terms, about 1/4 lb. per week per cu. ft. of autoclave. The same year a third, larger autoclave was acquired. Each chamber of this one has an inside diameter of 6". The volume is 3.4 cu. ft. and the unit has a nominal growing capacity of 50 lbs. synthetic quartz per run.

Pilot-Plant Facility

The Signal Corps was sufficiently well satisfied with the results that we were awarded a contract for growing an increased amount of crystals, and a new phase was entered.

In August 1951 work was started on the plans for this new facility. Three autoclaves similar to the one mentioned above, were acquired, and two smaller ones having capacities of 0.7 cu. ft. Each chamber of the large autoclaves is a forged steel cylinder having a six-inch bore. To produce a vessel of this capacity for use at high pressure and temperature requires a considerable quantity of special steel: 4 1/2 tons per autoclave. Thus we ran into problems of procurement and priorities. (Fig. 5)

Since these autoclaves were to be maintained at high temperatures and pressures for long periods of time, a steel having a high creep strength at the operating temperature was chosen. For the large autoclaves the steel is ASME SA-213 Grade T-22, containing 2 1/2% of chromium and 1% of molybdenum. Its creep strength is expressed

as 22,000 pounds per square inch stress required for 1% elongation in 100,000 hours at 900°F.

The smaller autoclaves are of SAE 4620 steel produced as hot-drawn tubing with a cap welded on one end.

Next we had to house these autoclaves. For the protection of the operators in the unlikely event of bursting, the autoclaves are placed in covered pits. The general housing for the pits, work area, and office is built as an addition to an existing building at our Bedford, Ohio plant. (Figure 6) Here again we ran into priorities and procurement problems in getting steel for the building framework.

The new facility with its 5 autoclaves produced 4,136 crystals, weighing 841 lbs. during the twenty-seven months of the contract. It was in full operation 15 months.

Quartz-growing runs last upwards of three weeks. Earlier we had no positive assurance that our conditions were suitable for growth until the run was completed. We now radiograph two or three of the crystals from time to time with cobalt 60.

The product has averaged Grade 18 (using the military specification for natural quartz) which means that the quartz is 80% to 100% usable flawless radiograde material. In isolated instances, where the conditions of experimental runs were not standard, a few small losses were sustained.

In 1953 the equipment was transferred to pilot-plant status under the Signal Corps Supply Agency, for which it is making a 12-month Industrial Preparedness Study. Most of the product is being shipped to the Signal Corps Engineering Laboratories and is being largely used in their program of having synthetic quartz cut in routine production and tested for comparison with natural quartz. The output of this pilot plant amounts to approximately 100 lbs. per month.

Synthetic Crystal Forms

The typical synthetic crystal, grown from a seed cut parallel to the minor r face, and held by the edges perpendicular to the X-axis, develops into a modified X-block. (Fig. 7) The flat surface on each side of the crystal, while not guaranteed to be an X-cut exactly (at this stage of the project) varies little from it. This surface is thus ready for cementing to the flat mounting plate for cutting into AT or BT wafers.

Brush synthetic crystals can be made of such size, shape, and orientation, that wafers cut from them will yield a maximum of oscillator blanks with a minimum of waste. The current main interest is in AT-cut plates, 1/2 x 1/2 in. square and 1/2 in. diameter discs, for frequencies from 1 megacycle up. The seed for this crystal is therefore an AT plate or CT plate chosen so that multiples of this size blank can be cut, for example 6 blanks, with enough area between and around them for the saw cuts. As the seed grows, this rectangular cross section is not preserved, but decreases as the major R faces build up. When the size of the minor r face has become so small that a 1/2 x 1/2 inch AT cut cannot be diced from

*For clarity in referring to the rhombohedral faces of quartz the term and letter case "major R" will mean the major rhombohedron (11.1) and "minor r" will mean the minor rhombohedron (10.1).

the corresponding wafer, further growth is wasteful; hence we attempt to end the growing run at this point. The grown crystal weighs about 90 grams and requires about 60 days to grow.

With standard cutting techniques between 200 and 300 half-inch square AT blanks can be cut per lb. of synthetic quartz. With better techniques, especially the use of thinner saw blades and slightly closer cuts (permitted by crystals requiring less than 2 inches of cutting depth), the yield per pound will be increased.

These crystals have major R faces which provide at each end a well-shaped three sided pyramid. An important advantage here is that these faces can be used directly for assisting in orienting the crystals preparatory to cutting. For average accuracy, orientation may thus be obtained without the use of X-rays.

For other cuts and blank sizes, the seed can be chosen and growth controlled so that similarly optimum crystals can be produced.

A crystal grown on a major R seed again closely approximates an X-bar appropriate for slicing into AT or BT plates. (Fig. 8) However, the growth rate on a major R face is relatively low - about 1/5 to 1/3 of the rate on a minor r face - hence there seems to be no advantage in using this cut for seeds. The crystal on the Z-cut seed is appropriate for cutting into a variety of plates.

Piezoelectric Properties

The Signal Corps has currently an extensive program for testing synthetic quartz as to its piezoelectric behavior through supplying oscillator plate manufacturers with generous samples of synthetic crystals on a no-cost basis in exchange for reports on ease of handling, yield per lb., and tests of the finished units.

The ratio of reactance to resistance, or Q , for synthetic quartz has been measured by us and other laboratories. Values in the hundreds of thousands have been readily obtained, showing the material is satisfactory for all frequency-control uses. Measurements must still be made, however, to show whether the ultimate value of Q for synthetic quartz is lower, equal to, or higher than the Q obtainable with natural quartz. The Q of course varies with the mode of vibration and with the frequency, and is notoriously dependent on the mounting of the oscillator plate.

For the least influence of temperature on frequency, the AT plate of natural quartz must be cut at $35^{\circ} 18'$ to $21'$ from the Z-axis of the crystal this angle varying with the temperature range over which the oscillator is to be used.

It has been found tentatively that for synthetic quartz the optimum angle is about 7 minutes greater, and that the shape of the curve expressing this dependence is not the same.*

The reason for this is puzzling. The correct angle for natural quartz obtained anywhere is the angle just given: $35^{\circ} 18'$ to $21'$. Now this quartz has been grown most probably under a variety of temperatures and pressures and in the presence of a variety of other substances. Yet for synthetic quartz the required angle is different. This is one of the unsolved minor problems of the quartz investigation.

Is synthetic quartz as good as natural? Compared with Grade 1 natural radio quartz, synthetic quartz is just as good in the following particulars:

It is free of such flaws as milkiness, noticeable liquid-vapor inclusions; it is clear and water-white; it has very closely the same optical and chemical properties. It has substantially the same piezoelectric properties, with minor differences.

And in addition, synthetic quartz is better in a number of ways:

It can be guaranteed free of, or with a minimum percentage of optical and electrical twinning. It can be supplied in uniform sizes and shapes, thus it is adaptable to mass production techniques. It is completely faced. Through its pyramid ends it lends itself to lining up in jigs for cutting without X-ray measurements for use in less critical applications. It can be grown in a number of shapes the best to serve the various ultimate needs.

References

1. Swinnerton, A.C., "Investigations in European Theater" PB 28897 (1946)
2. The Brush Development Co., PB 88027 (1941)
3. Walker, A.C., "Growing Quartz Crystals for Military Needs" Electronics 24 96-99 (1951)
4. Spezia, Giorgio, Accad. Sci. Torino, Atti, 40 254 (1905); 44 95-107 (1908)
5. Brown, C.S. et. al. "Growth and Properties of Large Crystals of Synthetic Quartz", Mineralogical Mag. (London) 29 858-874 (1952)
6. Hale, D.R., "Properties of Synthetic Quartz Crystals and their Growing Technique" Brush Strokes, Dec. 1952 Brush Dev.Co. Cleveland, O.

*Reported previously by Bell Telephone Laboratories

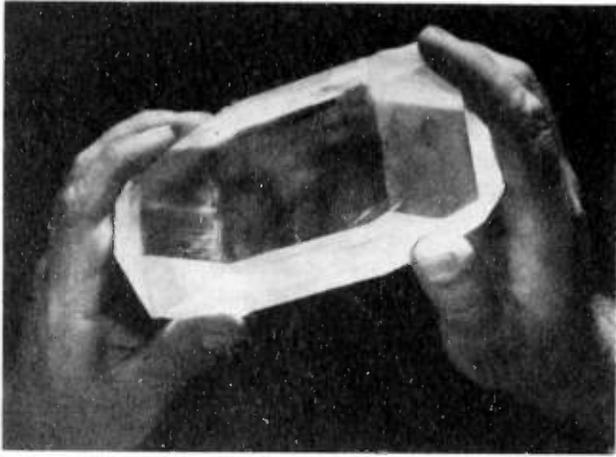


Fig. 1
Synthetic quartz crystal weighing 2.2 lbs.

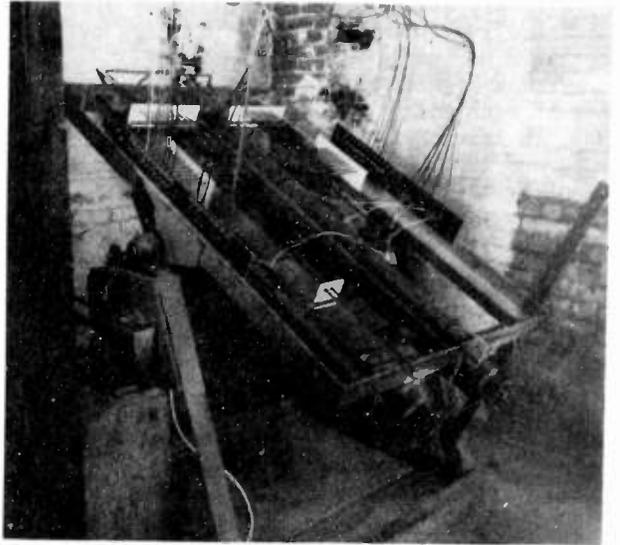


Fig. 4
Two-chamber autoclave of 0.7 cu. ft. volume.

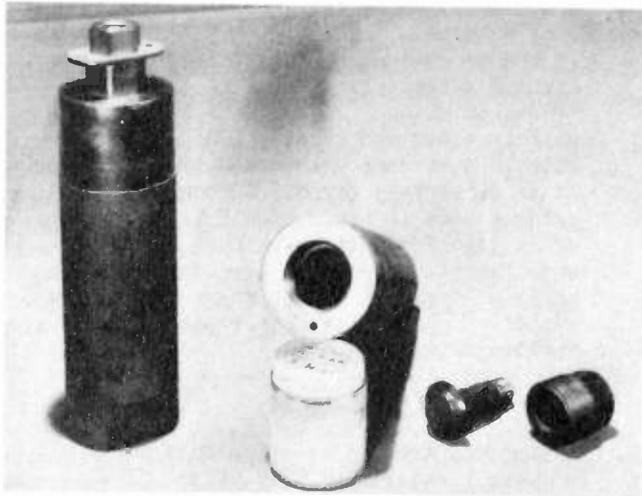


Fig. 2
A pressure-vessel with volume of 4.3 cu. in.

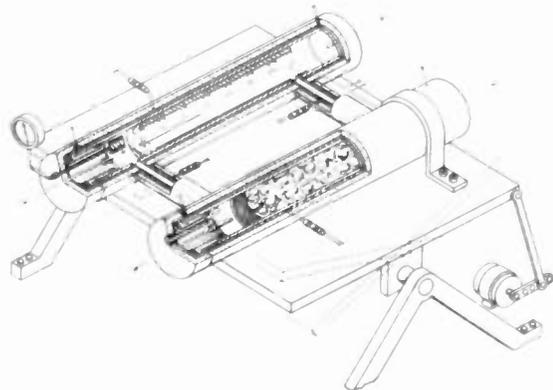


Fig. 3
Perspecting view of two-chamber rocking autoclave.

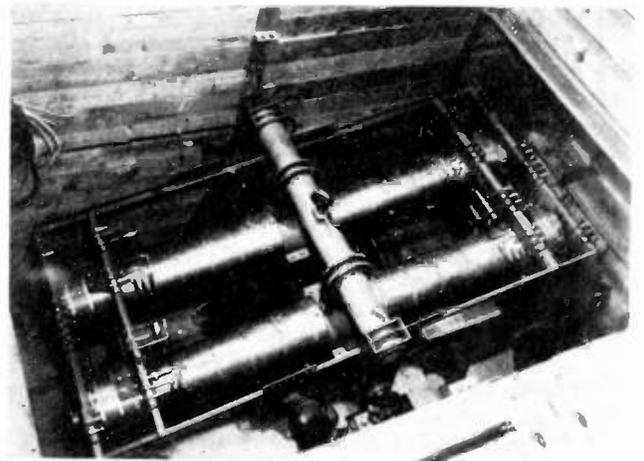


Fig. 5
Large unit mounted in protection pit.

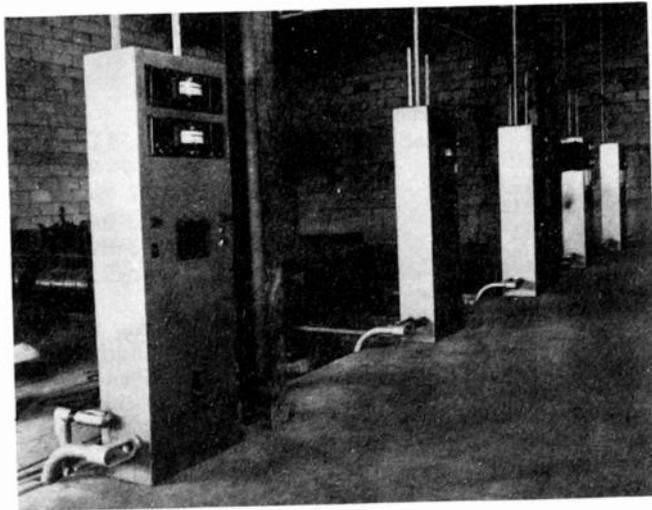


Fig. 6
General view of pilot plant.

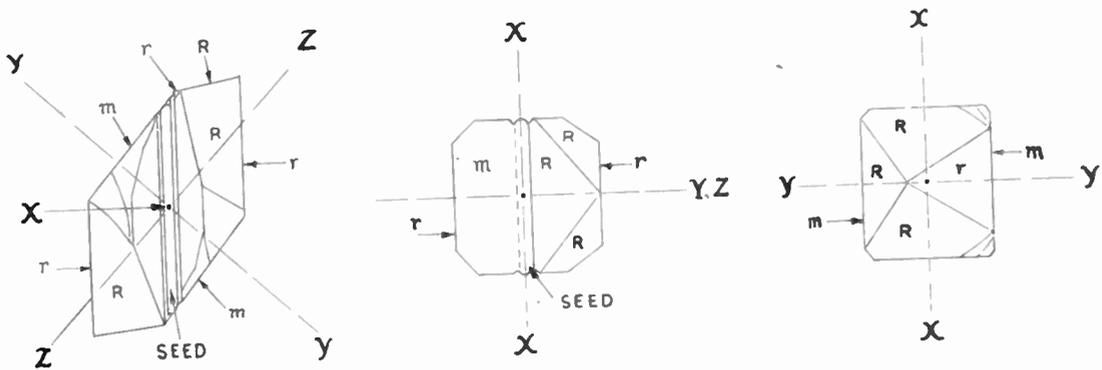


Fig. 7
Drawing of crystal grown on minor r seed.

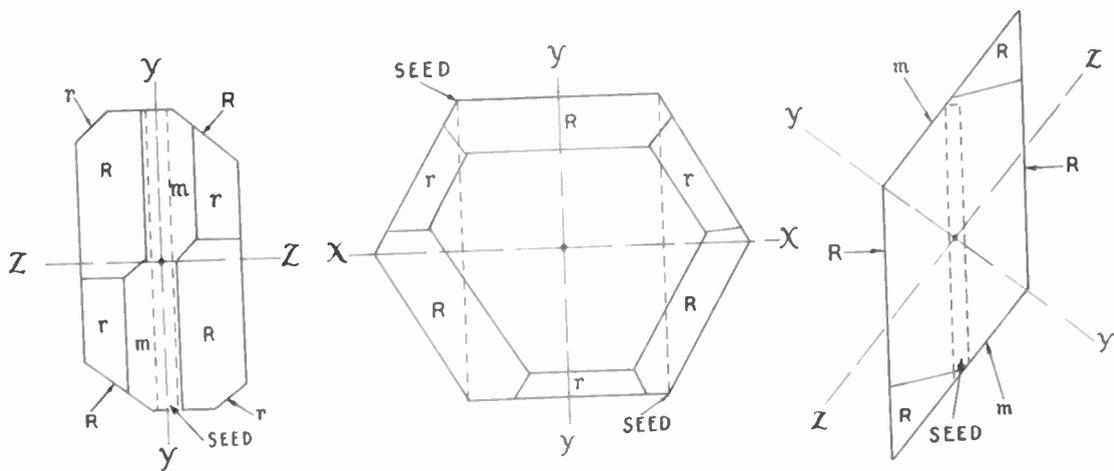


Fig. 8
Drawing of crystals grown on Z-cut
seed and major R seed.

APPLICATION OF PRECISE COMPONENTS IN PERMEABILITY TUNED OSCILLATORS

David Hodgkin
Collins Radio Company
Cedar Rapids, Iowa

Summary

In order that the frequency output of an L/C tuned oscillator be made nearly independent of environment it is necessary that the best combination of capacitors, inductors, and tuning cores be used. These components must have repeatable, linear, compensating temperature coefficients with small absolute drift during and after environmental cycling. A specially developed ceramic plate, hermetically sealed capacitor and its environmental parameters will be discussed. The properties of the inductor will be shown. The temperature coefficient of the tuning core will be explained with reasons for the requirement that the core have other than zero temperature coefficient. The final performances of the combination will be shown and discussed.

Introduction

The increasing importance of frequency stability in every frequency control application has made it necessary to obtain the best possible components for frequency determining applications. When inductance-capacity is used in frequency determining circuits even differential changes become measurable and important.

Component parameters of primary importance in permeability tuned oscillator application are:

1. Tolerance of nominal value
2. Temperature coefficient value tolerance
3. Linearity of temperature coefficient
4. Aging with environmental cycling
5. Freedom from discontinuity of value during either differential changes of temperature, voltage, or shock and vibration
6. Small frequency dependence of nominal value and differential characteristics.

Measurement of differential characteristics becomes a problem also, and the solution to this problem has had an important bearing upon successful development of the components.

It is noteworthy that the majority of the measuring schemes for measurement of temperature coefficient have used a frequency variation method for sensing change rather than a direct reading of change in absolute value by bridge measurements^{1,2,3}

The mathematical relations of differential component variations are best kept in chart or graphical form since the exact relationships are somewhat nonlinear. The relationship of component value to temperature, for example, could be represented by a simple power series.

Graphical and tabular presentation will be

more vivid and descriptive and will be used along with numerical differentiation and integration where necessary.

Variations in value due to temperature change, aging, and other factors are best referred to the nominal value and expressed as a per cent of nominal or in parts per million (PPM) of nominal.

The Capacitor^{4,5}

Figure 1 shows a cross-sectional view of the multiple-plate ceramic capacitor. Simplified, the unit is a parallel combination of individual ceramic disc capacitors. The individual plates are selected for not only capacitance value, but are selected for temperature coefficient characteristics.

To a certain extent also, the linearity of temperature coefficient is controllable, although this is a characteristic that is provided by ceramic compounding techniques. It is necessary in order to achieve maximum stability that the capacitor be sealed to moisture; otherwise the capacitance value will not remain constant with environmental changes. Humidity affects the temperature coefficient of unprotected ceramic, and the Q is also affected. The cross-sectional view (Figure 1) shows how the unit is solder-sealed.

Capacitors can be built up with as many plates as necessary to make large capacity values. While this paper will be confined to the near zero temperature coefficient types, it is, of course, possible to use materials with high dielectric constants. Thus the manufacturer can obtain up to several thousand micromicrofarads in a reasonable case size.

The Q factor at 1 megacycle for a NO50 is quite typical of good quality ceramic bodies and has a value of 1000 or more.

The capacitance value itself is adjusted to $\pm 1\%$ in production quantities.

The method of construction is designed to give inherently low inductance, and it is generally possible to use this type of unit at frequencies well above the frequency that JAN C20A types of equivalent size become resonant.

The change of capacitance with applied voltage is almost insignificant for voltages within the working rating of 500 volts. A value of .05 mmf for 500 volts applied has been measured for a 580 mmf unit. In the precision oscillator it is not a desirable circuit design practice to apply voltage to the capacitor other than,

Table I

(1) Temp. Degrees Cent.	(2) Measured Capacitance mmf	(3) Capacitance Change per Temperature Increment mmf	(4) Temperature Coefficient PPM/mm ² /C ^o	(5) Measured Capacitance Deviation (mmf) from 26°C	(6) Calculated Cap. Dev. from 26°C @.03062 mmf/ C ^o /mmf	(7) Nonlinear Deviation Meas. from 26°C minus calc. in mmf
-44	546.127	.760	-58	+2.127	+2.145	-.018
-20	545.367	.675	-59	+1.367	+1.410	-.043
+ 1	544.692	.692	-50.8	+ .692	+ .766	-.074
+26	544.	.748	-65.6	0	0	0
+47	543.252	.506	-49	- .748	- .644	-.104
+66	542.746	.691	-55.4	-1.254	-1.226	-.028
+89	542.055			-1.945	1.930	-.015
		Total	4.072 mmf			
		Average	.03062 mmf/C ^o			

Table II

(1) Temp. Degrees Cent.	(2) Measured Capacitance mmf	(3) Capacitance Change per Temperature Increment	(4) Temperature Coefficient PPM	(5) Capacitance Deviation (mmf) from 25°C	(6) Calc. Dev. if change were .0293 mmf/C ^o linear	(7) Measured Capacitance Nonlinear Dev. mmf
-42	52.077	-.699	-635	+2.077	+1.965	+.112
-20	51.378	-.752	-602	+1.378	+1.319	+.059
+ 5	50.626	-.626	-626	+ .626	+ .586	+.040
+25	50.	-.654	-523	0	0	0
+50	49.345	-.590	-590	- .654	- .733	+.079
+70	48.756	-.546	-546	-1.244	-1.319	+.075
+90	48.210			-1.790	-1.905	+.115
		Total	3.867 mmf			
		Average	.0293 mmf/C ^o			

perhaps, grid bias. Therefore, the actual capacitance voltage characteristic is negligible in the application.

The capacitor of this type shows remarkable characteristics in producing smooth non-erratic changes in capacity over a wide temperature range. Occasionally a "jumpy" unit is found which is apparently due to faulty assembly and soldering methods. When used in an oscillator it is possible to detect capacity erraticism in the order of .003 mmf.

The temperature coefficient (TC) of capacity

is one of the most important parameters in the oscillator application. This characteristic must compensate the coil and core so that frequency as a function of temperature remains nearly constant over very wide thermal cycles. These characteristics will also repeat again and again with little or no change in value or linearity.

Table I shows the measured characteristics of a typical 540 mmf unit with specified temperature coefficient of negative 50 ppm.

The temperature coefficient is not perfectly linear and column No. 7 shows the actual amount

of nonlinearity. This characteristic of linearity is shown in Table II, column 7, for a nonsealed temperature compensating ceramic capacitor of 50 mmf N600 specified value. This value was chosen because the actual capacitance change per temperature increment is about the same as for the 540 mmf N50 sealed capacitor shown in Table I.

A comparison of linearity results for both units is given in Table III. This shows that the nonlinearity of capacitance deviation at the extremes of temperature for the nonsealed capacitor is approximately seven times worse than the sealed type. (See column 2a and 5a for comparative data)

The parts per million nonlinear deviation is very much worse. However, the contribution of

this capacitor in temperature compensation would be proportional in application to the actual nonlinear capacitance deviation. The important factor to note is that for minimum nonlinearity of frequency deviation over a wide range of temperature, the major temperature compensation should be built into the sealed capacitor. If the sealed capacitor were the only contributing factor to frequency nonlinearity, column (3) of Table III would show the deviation thus produced.

Type test data from twelve (12) production capacitors of 370 mmf appears in Table IV.

Capacitance drift after several wide range thermal cycles seldom exceeds 150 ppm. Capacitance drift due to humidity cycling produces +500

Table III

(1) Temp. Degrees Cent.	(2)a. 544 mmf N50 Sealed Cap. Nonlinear Deviation mmf	(2)b. Nonlinear Deviation PPM/mm ²	(3) Frequency Nonlinear Deviation PPM	(4) Temp. Degrees Cent.	(5)a. 50N600 Non- Sealed TC Capacitor Nonlinear Dev. mmf	(5)b. Capacitance Nonlinear Deviation PPM
-44	-.018	-33	+16.5	-42	+.112	+2240
-20	-.043	-80	+40	-20	+.059	+1180
+ 1	-.074	-137	+68.5	+ 5	+.040	+ 800
26	0	0	0	25	0	0
47	-.104	-192	+96	50	+.079	+1580
66	-.028	-52	+26	70	+.075	+1500
89	-.015	-28	+14	90	+.115	+2300

Table IV

Mean TC PPM/mm ² /C ^o -50 to 90°C			Capacitance Nonlinear Deviation PPM		
Max.	Min.	Spec.	Max.	Min.	Spec.
-56.8	-47.1	-50 ±10	-243	-108	±400

to +700 ppm change⁶. This is due to the surface film of moisture present. Subsequent baking for short periods causes capacitance to return to normal.

From the foregoing data, it can be seen that the sealed ceramic plate capacitor does have, to a major extent, all the characteristics necessary to satisfy the basic requirements for precision oscillator applications.

The Mica Capacitor⁷

A hermetically sealed silvered mica of the button type construction has been developed

recently. This capacitor has shown remarkable freedom from erraticism, drift of nominal value after environmental cycling, and stability of temperature coefficient. It is possible, also, to achieve more capacitance in a smaller size than is possible in a low temperature coefficient ceramic.

The temperature coefficient of the hermetically sealed mica varies from zero to positive 25 ppm/C^o. The linearity of the T.C. is within ±150 ppm from -50 to +90°C. Capacitance value can be made to ±1% tolerance.

For precision tuned oscillator applications,

the major disadvantage in the use of this type unit is its positive temperature coefficient. To achieve a desired compensation in a composite circuit, additional parallel negative compensation of a ceramic type is necessary. This usually will increase the nonlinearity to a larger value than if a sealed type negative TC ceramic were used.

The Inductor

In the precision oscillator application where permeability tuning is employed, the inductor has many requirements not demanded of a fixed inductor. Primarily, the winding must be placed upon the form with a variable spacing so that, as the core moves into the coil from one end, the variation of inductance produces linear frequency increments. Except for fringing effects at the ends of the coil, this would make the variation of turn spacing an exponential function of coil length. Because of fringing effects, additional turns must be wound on the ends. Figure 2 shows a typical coil.

The form material that has been found best for the application is a phenolic rolled-paper-base laminated type. At first consideration this would not seem the best choice of material. However, with a temperature coefficient of expansion approximately that of copper, a stable, small dielectric constant (about 3.6), readily machineable, with a good secular characteristic provided that moisture is excluded and that preaging is employed, the material choice is the best available.

The length/diameter ratio of the coil necessary for achieving the required tuning ratio is in the order of from 2.5 to 3 to 1. This is not optimum for Q, although the Q of the coil is nearly 100 in air. When mounted and shielded the Q is normally from 35 to 50.

The exact turns location must be maintained through high and low temperature cycling or calibration accuracy will be impaired. This requires an initial care in winding the coil. Wire is under constant tension while being wound and is terminated before removal of tension. A good high temperature impregnation is then coated in less than three mil thickness which helps bond the wire to the form. A thin coating is important. Otherwise where the windings are closely spaced the impregnation may not successfully get to the form and may only act as a plastic matrix between wires. This can cause poor cyclic behavior and unpredictable temperature coefficient.

The temperature coefficient of the coil when measured in open air is about positive 25 PPM/C° and is quite linear and cyclic, providing it is not subject to variations in humidity or high dielectric constant gasses. As mounted in the oscillator, with no core in the coil, the temperature coefficient is about positive 46.5 PPM. See column 1, Table VII, for the coil TC deviation curve. This increase in TC is due to the effects

of shielding, most of which is because of the expansion effects of the extruded aluminum cover which also provides one of the hermetic seals in the oscillator.

The Core

Powdered Iron^{8,9}

Powdered iron cores have been used for many years as devices to increase effective inductance for a given size coil, to enhance the Q, for tuning, and for trimming application.

Probably the most exacting application for a powdered iron core has been in the Collins Permeability Tuned Oscillator. In this application the effective permeability must be about 2.7 with a tolerance of 0.5% or less. The homogeneity of the core permeability through its length must be alike from one core to the next. The temperature coefficient of effective permeability must be positive 6 to 7 parts per million over a wide range of temperatures, for if it is not, the temperature compensation at high and low frequency core positions is not the same. Therefore, it is impossible to obtain frequency independence of temperature. The Q is also an important factor, although the most important demand upon Q is its constancy with temperature.

Cores compounded from many grades of powdered iron have been tested, but the material most suited to P.T.O. applications is Carbonyl E. This material has relatively high Q in the 500 kc/s to 4 mc/s region, has a toroidal permeability of from 10 to 25, and can be molded to give the correct temperature coefficient.

Aging of molded powdered iron cores has been a problem that still is not completely solved, but the effects have been largely eliminated by treatment similar to the annealing of metal and ceramic products. Additional resistance to aging has been produced by coating with impervious resins such as the phenolic and epoxy materials.

Temperature coefficient data of a typical powdered iron core as used in the P.T.O. is given in Table V. The measurement was obtained by cutting the center section out of the sleeve thus eliminating the metal inserts. This provided a toroidal shape upon which wire was wound to make approximately 25 microhenries inductance.

The temperature coefficient thus measured is, of course, the toroidal TC and to find the effective temperature coefficient in the oscillator itself we must do the following. Measure the resonant frequency of the oscillator coils and fixed capacitors with no core present, and then, since we know the lowest frequency to which the oscillator is to be tuned, we can compute the frequency ratio which, squared, becomes the effective permeability. In the example at hand:

$$f_H = \text{Frequency no core inserted} = 3.63 \text{ mc/s}$$

f_L = Lowest frequency tuned = 2.5 mc/s

$$\frac{f_H}{f_L} = \frac{3.63}{2.5} = 1.452 = \sqrt{\mu}$$

$$\mu = (1.452)^2 = 2.11$$

Toroidal permeability is measured at 21.05.

Now, we can find the factor by which the toroidal TC may be multiplied to get the effective TC of the tuning core at the low frequency end of the oscillator.

$$\frac{2.11}{21.05} = 0.1002 \text{ Effective Factor}$$

Table V

Temperature Coefficient of Carbonyl E. PTO Core

(1) Temp. C°	(2) Toroidal* Temp. Coef. PPM/uh/C°	(3) Effective* Temp. Coef. PPM/uh/C°
-44		
+ 5	+51.3	+5.44
+30	+56.9	+5.70
+52	+73.4	+7.35
+76	+80.3	+8.05
+97	+140.0	+14.02

* Note the linearity of temperature coefficient below 76°C

Ferrite Core*

The same characteristics necessary in a powdered iron core are needed in a ferrite core for use in a P.T.O. Of course, the reason for using a ferrite core is to increase the effective permeability or tuning ratio. The toroidal permeability must be very much larger than the powdered iron permeability to achieve an increase in a linear frequency coverage from 1.5/1 to 2.0/1. The reason for this is nicely shown by W. J. Polydoroff and A. J. Klapperich¹⁰.

In actual usage the effective permeability for tuning a linear range of 1.5/1 is approximately 2.7 while tuning a 2:1 range requires an effective permeability of 4.53. The toroidal permeability of ferrite then must be at least 130 for safety in production.

Ferrite has certain advantages, other than

* Wherever ferrite is mentioned it is the Stackpole Carbon Company Ceramag 1848. This material has shown the best overall stability characteristics of any ferrite tested for the application.

having high permeability, over powdered iron. It is more homogeneous throughout its length. The change of permeability with time is much less than powdered iron since there is no further oxidation at operating temperatures. Disadvantages are present also. They include magnetic instability in the presence of even moderate magnetic fields, less inherent temperature stability, and somewhat less predictability in manufacture. Ranking on approximately an equal basis is the Q, at least over the range from .500 mc/s to 5.0 mc/s when used in the sleeve form.

Table VI shows the characteristic of Ceramag 1848 over a wide temperature range for both the toroidal characteristics and the effective characteristics as used in an oscillator. Column (4) shows the per cent change of toroidal permeability from nominal value over the temperature range of minus 44°C to positive 97°C. The effective value is obtained by the identical methods used for powdered core values.

$$\left(\frac{f_H}{f_L}\right)^2 = \left(\frac{3.19}{1.5}\right)^2 = 4.53, \text{ Toroidal } \mu = 160$$

$$\frac{4.53}{160} = .0283 \text{ Effective Factor}$$

Comparison of the effective temperature coefficient of ferrite and powdered iron shows that from room temperature to 52 degrees Centigrade the temperature coefficient is nearly identical, while at lower and higher temperatures the deviation increases. Work is now in progress to eliminate this nonlinear characteristic and to straighten out the excessive drop in permeability at cold temperatures.

Combined Characteristics

The individual characteristics of the components used in a precision permeability tuned oscillator have been shown. The next step is to show how these components perform when combined. The columns in Table VII are compiled to show the combination effects. Use of the ferrite core data will be made, since the ferrite has less inherent stability and will show deviations more readily. In these measurements 40 degrees Centigrade was the point nearest room temperature where measurements were taken, so throughout the discussion 40 degrees Centigrade will be used for a zero deviation point.

Column (1) shows the coil IC in PPM frequency deviation from the reference temperature.

Column (2) shows the combination of the tank capacitor and small temperature compensators, and also is dimensioned in PPM frequency deviation from reference temperature.

Column (3) shows the measured value of the coil and tank capacitors in the oscillator with no core in the unit.

Column (4) requires some further explanation

Table VI

(1) Temp. C°	(2) Toroidal Temp. Coef. PPM/uh/C°	(3) Effective Temp. Coef. PPM/uh/C°	(4) Toroidal Inductance Deviation from Nominal Inductance, Per Cent
-44	+683	+19.3	-.1115
+ 5	+432	+12.2	-.0432
30	+236	+ 6.7	0
52	+103	+ 2.9	+ .0103
76	+ 93.6	+ 2.65	+ .0197
97			

Table VII

(1) Temperature Centigrade Degrees	(2) Measured Coil T.C. Frequency Deviation PPM	(3) Measured Capacitor Frequency Deviation PPM	(4) Coil & Captr. Freq. Dev. Col. (1) plus Col. (2)	(5) High Freq. End. Mechanical TC Freq. Dev. PPM	(6) Calculated Oscillator Freq. Dev. at High Freq. Endpoint PPM
-46	+2000	-2350	-350	+600	+250
-19	+1350	-1575	-225	+415	+190
+11	+ 700	- 825	-125	+205	+ 80
+40	0	0	0	0	0
+70	- 670	+ 785	+115	-210	- 95
+98	-1300	+1445	+145	-405	-260

(7) Temperature Centigrade Degrees	(8) Measured Oscillator Freq. Dev. at High Freq. Endpoint PPM	(9) Low Freq. End Mechanical TC Freq. Dev. PPM	(10) Measured Effective Core T.C. Freq. Dev. PPM	(11) Sum of Col. (7) and (8) PPM	(12) Calculated Oscillator Freq. Dev. at Low Frequency Endpoint PPM
-46	+185	+355	+675	+1010	+660
-19	+165	+230	+415	+ 645	+420
+11	+100	+115	+150	+265	+140
+40	0	0	0	0	0
+70	-120	-115	- 55	- 170	- 55
+98	-330	-225	- 85	- 310	-165

because it is the mechanical temperature coefficient element of the oscillator. In the P.T.O. the core is mechanically loaded to a leadscrew which is also in turn loaded to a precision machined casting, called the head. To this head is mounted the coil. However, the plane of the

coil reference does not contain the point of the leadscrew reference¹¹. Thus, differential expansions make up the mechanical temperature coefficients of the oscillator. These can be calculated by means of thermal expansion characteristics combined with tuning rates. However, the

best and most accurate way to come by the factor is by measurement of frequency variation.

This measurement is made by subtracting the frequency deviations due to temperature in PPM of the oscillator with coil and tank present but with no core as in column (3) from the recorded frequency deviation in PPM at the high frequency endpoint of the oscillator linear tuning range. This high frequency endpoint contains but very little of the TC of the core since only 1/16th inch or so of the core is in the coil at this point. By repeated measurements of the above nature the value has been established as minus 7 PPM/C° frequency deviation. Column (4) shows this mechanical TC as frequency deviation in PPM.

Column (5) shows the calculated PPM frequency deviation at the high frequency endpoint. This figure is obtained by algebraically adding column (3) and column (4).

Column (6) is the actual measured value of the high frequency endpoint deviation. The error between column (5) and (6) is well within reasonable value and the small discrepancy is probably due to a fractional variation in the value of the calculated mechanical TC.

Column (7) shows the low frequency endpoint mechanical temperature coefficient deviation which is obtained by multiplying the high frequency endpoint mechanical factor by the ratio of length of expansion paths at the low frequency and the high frequency endpoints. In our case this is $\frac{1.562}{2.812} \times (-7) = -3.19$ PPM/C°.

Column (8) is the frequency deviation due to the effective core temperature coefficient.

Column (9) is the algebraic sum of column (7) and column (8).

Column (10) is the algebraic sum of the coil-tank deviation column (3) and column (9). This is the calculated low frequency endpoint. It was not possible to obtain data for the measured temperature coefficient characteristic for the actual core used for measurement. However, the correlation with other low frequency endpoint temperature coefficient measurements, where similar cores were used, is very good.

Thus, it can be seen how each component characteristic is matched against the others and in the final result the whole is better than the individual units. Quality control in manufacturing components, quality in assembly, and test is very important throughout the entire P.T.O. program. A continuing program of development is underway to find and utilize components which are even better than those shown herein.

References

1. Connor, J. A. and Troxel, D. I., "Investigation of Capacitors, Temperature Compensation Types", Naval Research Laboratory Letter 3920-139A/49 dated 6 October 1949.
- 2.(a) Bady, Isidore, "Equipment to Measure the Temperature Coefficient of Capacity", Squier Signal Laboratory Technical Memorandum M-1165, 15 January 1949.
- 2.(b) Bady, Isidore, "Precise Measurement of the Temperature Coefficient of Capacitance", Squier Signal Laboratory Technical Memorandum M-1382, 11 June 1951.
3. Smith, J. S., "Final Report - Design Study Toward Development of a Test Set to Measure Temperature Characteristics of Ceramic Capacitors", New York University, Research Division, College of Engineering, 14 October 1951.
4. Centralab, Division of Globe Union Inc., Milwaukee, Wisconsin.
5. Herlec Corporation, Grafton, Wisconsin.
6. Standard Cycle for Moisture Resistance of Component Parts. SC-D-16286.
7. Sangamo Electric Company, Springfield, Ill.
8. Aladdin Radio Industries, Nashville, Tenn.
9. Radio Cores Inc., Oak Lawn, Illinois.
10. Polydoroff, W. J. and Klapperich, A. J., "Effective Permeability of High Frequency Iron Cores", Radio November 1945.
11. Meier, J. C., "Precision Tuned Oscillators", Signal, Spring Issue, Collins Radio Company.

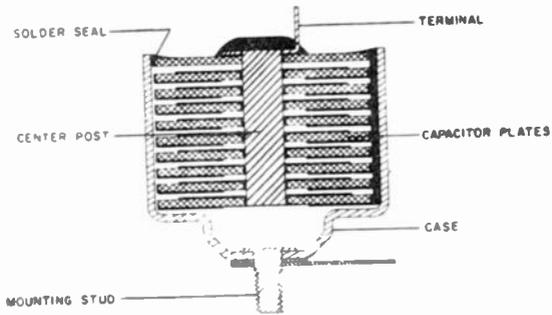


Fig. 1
Cross-sectional view of multiplate
ceramic capacitor.

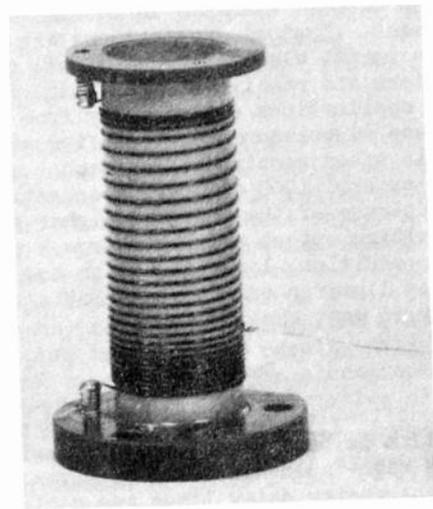


Fig. 2
Typical variable pitch oscillator inductor.

MAGNETIC-CORE DELAY CABLES

Dimitri R. Stein
Columbia Technical Corporation
New York, N.Y.

Summary

Delay cables with impedances between 1000 and 3000 ohms and time delays up to 0.6 microsecond per foot are described. These cables resemble conventional RG-type coaxial cables except that the inner conductor is a single layer coil continuously wound on a ferromagnetic core. This construction results in delay elements with high characteristic impedance, high time delay per unit length and bandwidths up to 16 mc for a delay of 1.0 microsecond.

Introduction

Delay lines are becoming increasingly important as circuit elements in modern electronic equipment. Typical applications are found in radar equipment, electronic computers, color TV transmitters and receivers, and allied fields. In these applications a time delay from a fraction of one microsecond to about five microseconds is often required. Impedances of the order of several 1000 ohms are frequently desirable because of the possible higher gain and higher voltage output available from a given tube. In addition, low attenuation and substantially linear phase characteristic are necessary in most applications.

Design Data

A modification of the distributed parameter construction is best suited for the design of efficient high-impedance delay elements. Distributed-parameter delay lines are derived from coaxial cables in which the inner conductor takes the form of a continuously wound coil. The delay and the impedance are determined by the distributed capacitance and inductance of the conductor configuration. The coiled inner conductor results in a considerable increase in the distributed inductance which in turn raises the delay and the impedance of the cable. A typical distributed-parameter delay line is shown in Fig. 1. This illustration represents a high-impedance cable which was developed by Dr. Kallmann¹ and which is known as Type RG-65/U. The electrical characteristics of a distributed-parameter delay line can be expressed by the following equations:

$$L = \mu \pi^2 d^2 n^2 10^{-11} \quad (\text{h/m}) \quad (1)$$

$$C = \frac{2l \times 10^{-12} k}{\log 10 a/b} \quad (\text{f/m}) \quad (2)$$

$$Z = \sqrt{\frac{L}{C}} = \frac{\pi n d \sqrt{\mu \log_{10} a/b}}{\sqrt{2.4 k}} \quad (\Omega) \quad (3)$$

$$T = \sqrt{LC} = \frac{10^{-5} \pi n d \sqrt{2.4 k \mu}}{\sqrt{\log_{10} a/b}} \quad (\text{usec/m}) \quad (4)$$

An examination of equations (3) and (4) shows that the impedance and the delay of a delay cable is determined by the following design constants:

a = inside diameter of outer conductors, in cm

b = outside diameter of coil, in cm

n = number of coil turns per meter

d = diameter of coil, between wire centers, in cm

k = effective dielectric constant of spacer

μ = effective permeability of core

It is obvious that a, b and d are not independent variables; rather they vary with the core diameter, the wire diameter and the thickness of the dielectric between inner and outer conductors.

Fig. 2 is a design chart showing the impedance and the time delay of a conventional distributed-parameter delay cable with a No. 38 AWG wire as inner conductor. A number of interesting conclusions can be drawn from this chart. The delay increases sharply with increasing core diameter and decreases with increasing thickness of the dielectric. The impedance displays a different tendency, i. e. it rises both with increasing core diameter and increasing thickness of the dielectric. It should be noted, however, that the impedance increases only slowly with the core diameter. Also, it is well to remember that with increasing core diameter the length of the inner conductor and its resistance become greater which adversely affects the electrical losses.

The design chart yields cables with impedance values up to 2000 ohms and delays up to about 2.0 usec per meter. In order to raise the impedance further without sacrificing delay, it is necessary to increase the inductance of the inner conductor assembly. This can be achieved by selecting a smaller wire diameter, e. g. No. 40 AWG. However, No. 40 AWG wire is already rather difficult to handle and presents a considerable problem both to the manufacturer and the user. Furthermore, it should be borne in mind that a reduction in wire size by two gauges (No. 38 to 40) almost doubles the resistance of the inner conductor assembly per unit of length and greatly increases the attenuation.

Fig. 3 shows a comparison of impedance and delay values for wire sizes No. 36, 38 and 40 AWG. The heavy lines indicate impedances of 1000 ohms and delays of 1.0 usec per meter, respectively. The broken lines designate 2000 ohms and 2.0 usec per meter, and the dotted lines refer to 2500 ohms and 3.0 usec per meter. It is apparent that with the conventional delay cable construction high impedance and long delays can only be obtained with an extremely thin inner conductor, large core diameters and consequently high losses.

A solution to this dilemma can be found by examining equations (3) and (4) which reveal that the impedance and the time delay vary with $\sqrt{\mu}$. In conventional delay cables non-magnetic materials are used for the coil core; i. e. $\mu = 1$. Yet, it is well known that the inductance of a coil can be raised by winding it on a ferromagnetic core. The application of this principle to delay cables appeared promising although it was recognized that the magnetic core had to meet a number of unusual mechanical and electrical requirements, namely:

- 1) Adequate tensile and compression strength,
- 2) Flexibility,
- 3) Uniform distribution of magnetic particles,
- 4) Adequate permeability,
- 5) Low electric losses in applicable frequency range,
- 6) Stability of electric properties in the applicable frequency and temperature range.

The development of delay cables with magnetic core was initiated several years ago by Columbia Technical Corporation, New York City. As a result of a very extensive research program, a method was developed which made it possible to produce, in a continuous process, a flexible core loaded with ferromagnetic materials. This magnetic core satisfied all the requirements mentioned above to a large extent and, above all, it afforded a practical manufacturing method which, although somewhat intricate, made possible the production of the magnetic core in virtually unlimited continuous lengths with great accuracy and uniformity.

Types of Delay Cables

With the magnetic core construction it was possible to develop a series of highly efficient high-impedance delay cables. One of these cables, known as Type HH-2500, is shown in Fig. 4. It has an impedance of 2800 ohms and a delay of about 2.0 usec per meter. The cable is built around a magnetic core of 0.4 cm diameter, over which a

coil of No. 38 AWG wire is wound. A polyethylene or teflon spacer of 0.035 cm thickness separates the inner and outer conductors. The latter are individually insulated and applied in one direction only with a long pitch. The cable is protected by a tough PVC jacket. It has an overall diameter of about 0.25". A cable with the same design constants (see Fig. 2) but with a non-magnetic core would have an impedance of only about 1300 ohms and a delay of 1.0 usec per meter. It follows that the effective permeability of the core is about 4.

The attenuation of HH-2500 for delays of 1.0, 0.5 and 0.25 usec is shown in Fig. 5. The resulting bandwidth is 8 mc for a delay of 1.0 usec. The phase characteristic of HH-2500 is slightly curved downward, i. e. the delay decreased with frequency. This characteristic is desirable in such applications where the delay element is used in circuits with low-pass filters or peaking coils which display an opposite phase characteristic. As a result, a phase compensation takes place which tends to restore pulse shapes. Fig. 6 shows an oscillogram of a 300 kc square wave after passing through 1.0 micro-second of HH-2500 Delay Cable.

In addition to Type HH-2500, two other delay cable types were developed which are now in commercial production, HH-1500 and HH-2000. The physical dimensions of these cables are identical with Type RG-65/U, which was mentioned earlier; however, the plastic core is replaced by a magnetic core and in HH-2000 a serving of insulated outer conductors is applied instead of the bare copper braid. As a result, delay and impedance are increased and the electrical losses greatly decreased. The characteristics of these cable types are listed in Table 1.

An examination of Table 1 clearly reveals the effects of the magnetic-core on the conventional delay cable construction. In HH-1500 this design feature resulted in an increase in impedance from 950 to 1600 ohms and in a delay increase from 0.042 to 0.075 usec per foot, as compared to RG-65/U. HH-2000 has an impedance and a delay of 2200 ohms, and 0.11 usec per foot, respectively. Column 4 refers to an experimental RG-65/U cable with a magnetic core and insulated outer conductors. It can be seen that it is possible to obtain the same impedance and delay as in RG-65/U by using a No. 29 inner conductor instead of No. 32, with a corresponding reduction in losses.

Fig. 7 shows the attenuation in the discussed cable types, expressed in db per usec. Up to 5 mc, HH-1500 displays less losses than RG-65/U; however, at higher frequencies the attenuation increases more rapidly than in RG-65/U. The experimental RG-65/U is considerably superior to the standard type. The incorporation of the magnetic core and of the insulated outer conductors made possible a reduction in losses by almost 75% in the frequency range up to 10 mc. Of particular in-

terest is Type HH-2000 which has a bandwidth (3 db increase) of 16 mc for a delay of 1.0 usec, resulting in extremely fast rise times of the order of 0.03 usec. Fig. 8 shows that the delay is virtually constant up to 16 mc; i. e. the phase characteristic is linear.

Electrical and Mechanical Terminations

The high impedance level of the discussed magnetic-core delay cables results in a number of important advantages; i. e. higher possible gain from a given tube, low electrical losses and large delay per lineal foot of cable. However, in order to take full advantage of the bandwidth of the cables it is of course necessary to terminate the cable correctly, preferably at both ends.

In Fig. 9 a simple termination is shown for HH-2500 Delay Cable which is satisfactory for most applications. The required inductance values for various capacitive loads are indicated in the graph. More complex networks, such as m-derived low-pass filters may be found preferable where a good match over a wide frequency range is required. The indicated series inductances may be provided by "self-peaking"; i. e. a length of inner conductor assembly is left exposed at both ends of the cable. A length of $\frac{1}{2}$ " corresponds to about 30 uh.

The mechanical assembly of the described magnetic-core delay cables is simple. HH-1500 can be terminated in the usual manner using conventional coaxial cable connectors and jacks. However, since standard connectors have low characteristic impedances, direct solder connections between the inner conductor and chassis terminal,

and the outer conductors and ground are preferable. On HH-2000 and HH-2500 the outer conductors are insulated with a solderable compound. In making ground connections it is merely necessary to bunch the outer conductors together at the ends and to heat them with a soldering iron which will melt the wire insulation. Special end-caps have been designed for HH-2500 which greatly simplify the assembly of the cable.

All three cable types are protected by a moisture-resistant tough PVC jacket. The operating temperature range is from -20° to $+80^{\circ}$ C; however, it is possible, by substituting suitable materials, to extend the upper and lower temperature limits. All described cable types are very flexible; the recommended minimum bending radius is 3". Where space is critical, it is possible to cut the cable into short lengths and to lay the individual sections side-by-side, connecting the inner conductors in series. Such assemblies may be encased in a metal can and hermetically sealed.

Acknowledgment

The author is grateful to Mr. M. B. Kline, Instrument Division, A.B. DuMont Laboratories, Inc. and his co-workers, and Mr. L. R. Kirkwood and A. J. Torre, RCA Victor Division, Radio Corporation of America, for their valuable suggestions and their active interest in the described development work.

References

1. H. E. Kallmann, "High-Impedance Cable", Proc. I. R. E., Vol. 34, pp. 248-351, June 1946.

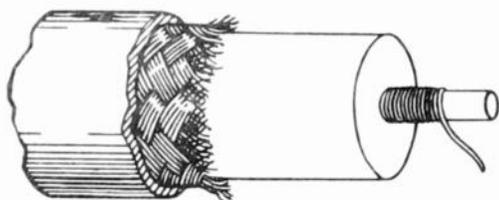


Fig. 1
Type RG-65/U high-impedance cable.

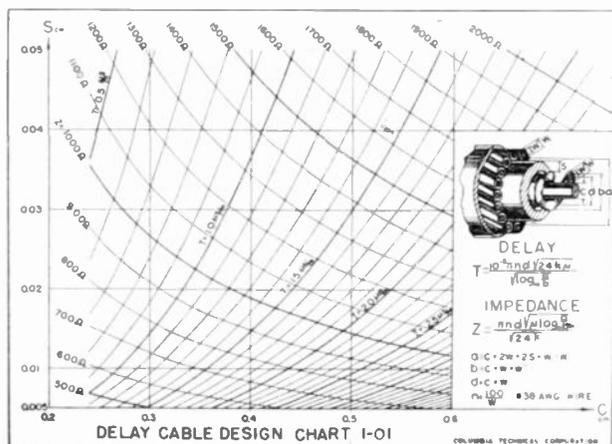


Fig. 2
Delay cable design chart for No. 38 AWG inner conductor.

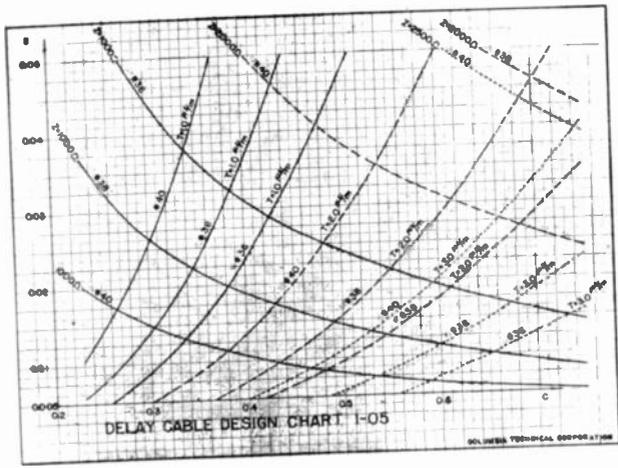


Fig. 3
 Delay cable design chart for No. 36, 38 and 40
 AWG inner conductors.

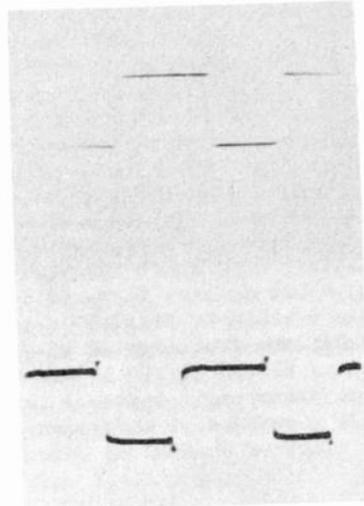


Fig. 6
 Square wave response of HH-2500 delay cable
 (top: input, bottom: output, frequency: 300 kc,
 delay: 1.0 usec).

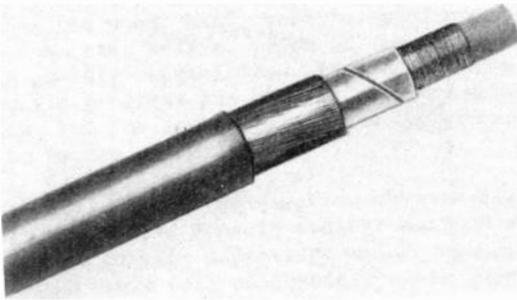


Fig. 4
 Type HH-2500 delay cable.

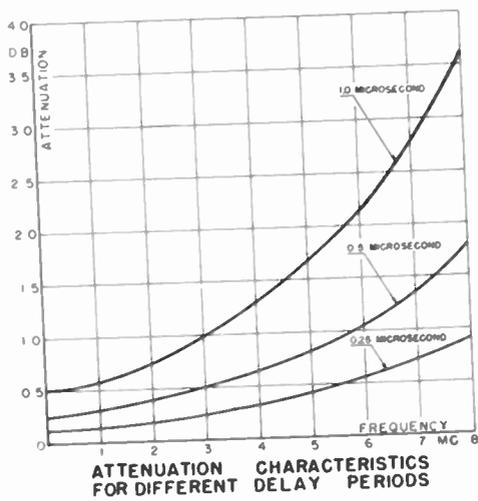


Fig. 5
 Attenuation of HH-2500 for different delay periods.

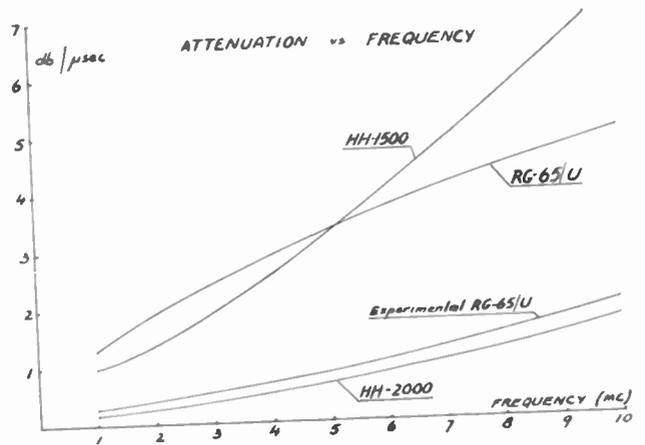


Fig. 7
 Attenuation characteristics of HH-1500, HH-2000
 and RG-65/U.

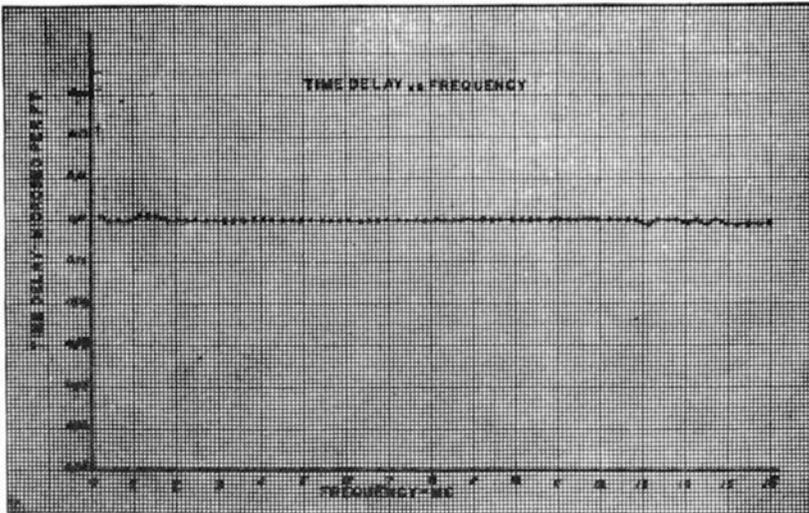


Fig. 8
Time delay vs. frequency of HH-2000 delay cable.

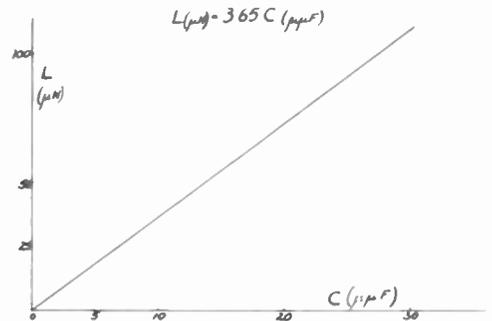
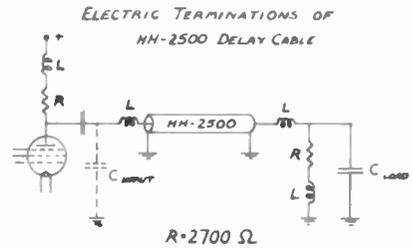


Fig. 9
Electrical termination of HH-2500 delay cable.

Table I
Comparison of delay cable characteristics.

	RG-65/U	HH-1500	HH-2000	Experimental RG-65/U	HH-2500
Impedance (Ohms)	950	1600	2200	950	2800
Delay (μsec/ft)	0.042	0.075	0.11	0.042	0.6
Unit length (ft/μsec)	23.8	13.3	9.1	23.8	1.7
Inner conductor (AWG)	#32	#32	#32	#29	#38
Dielectric O.D. (")	0.285	0.285	0.285	0.285	0.185
Cable O.D. (")	0.405	0.405	0.405	0.405	0.28
Attenuation (db/μsec)					
at 1 mc	1.3	1.0	0.2	0.3	0.5
at 4 mc	2.9	2.3	0.5	0.7	1.3
at 10 mc	5.0	7.0	1.7	2.0	4.0

IMPROVEMENTS IN THE FIELD OF ELECTROLYTIC CAPACITORS

Dietrich Altenpohl
Aluminum-Walzwerke Singen
Singen/Hohentwiel, Western Germany

I. Introduction

Electrolytic Capacitors are today produced in quantities hitherto unheard of. Nevertheless, the average electrolytic capacitor offered on the market today is, in spite of constant attempts at improvement, a rather imperfect structure. This is due to its lack of stability, and the extent of its dielectric losses. Its limited shelf life is a particularly disturbing factor. When an average electrolytic capacitor of, say 500 v is kept in unconnected storage over a period of several months or years, the decomposition of the dielectric can progress to the point where the capacitor is destroyed by excessive heating when suddenly exposed to voltage.

On the other hand, today's requirements, more than ever before, point in the direction of "high quality" capacitors. In recent years considerable progress has been made with tantalum anodes, but for general use these capacitors turn out to be too expensive¹.

This leads to the question whether everything possible has been done to achieve maximum stability in electrolytic capacitors which contain etched aluminum foil electrodes. This question must be answered in the negative, for the following reasons:

1. By using aluminum anodes of an especially high purity a decisive improvement in the stability of electrolytic capacitors can be obtained. However, up to the present only a relatively small percentage of the total output of capacitors contains electrodes made of "super pure" aluminum.

2. The vital component of the electrolytic capacitor - its dielectric (consisting of the formation layer) - has never been examined with sufficient thoroughness. In the formation process itself there is still plenty of room for improvements which can greatly affect the stability of the capacitor, or which can otherwise exert a beneficial influence on the electric characteristics.

II. Effects of Metallic Contamination

The following remarks refer to Point 1 mentioned above. The majority of anodes in electrolytic capacitors (or both electrodes in the case of AC capacitors) is made of etched aluminum foil of a purity of 99.85%. Actually the degree of purity fluctuates between 99.80 and 99.88%. This is the highest degree of purity normally obtained by electrolysis of smelted

aluminum, which can ordinarily be made available for capacitor purposes. However, during the last two decades a "super pure" aluminum, known as "Raffinal" has become available in commercial quantities². This type of aluminum has a purity of 99.99% or better³. For many years the etching of "raffinal" foil has caused considerable difficulties, which however have been largely overcome in recent years. Today it is possible actually to obtain a greater surface gain with "raffinal" than with ordinary 99.85% material. A six-fold gain at 600 v, and a foil thickness of .0037" are considered normal for "raffinal". By using an anode of 99.99% purity in place of one of 99.85% purity but similar surface gain, a marked improvement in the stability of the capacitor is achieved, particularly due to reduction of leakage currents and deformation (Fig. 1).

Fig. 2 shows that the iron content exerts by far the most detrimental influence on leakage currents, whereas the two other natural contaminants, silicon and copper, are relatively harmless. In comparing typical analyses of 99.85% and 99.99% aluminum, it will be seen that the iron content of the latter amounts to only 3 to 8% of that of the former:

Degree of Purity:	99.85%	99.99%
Iron	.05 to .10%	.001 to .004%
Silicon	.03 to .12%	.002 to .007%
Copper	.003 to .02%	.0005 to .003%
Zinc	.01 to .04%	traces

The reason for the damaging influence of iron, even in the small concentrations shown above, is its virtual insolubility in aluminum (Fig. 3).

Figure 3 shows how iron, in the form of heterogeneous aluminide crystals, is clearly recognizable in the aluminum grid. Since the solubility of iron in aluminum, under technically applied thermal conditions, is less than .01%, it can be readily seen that even in aluminum of 99.85% purity quite some ferric heterogeneities are in evidence, whereas in "raffinal" they are practically eliminated.

When a 99.85% pure aluminum foil, smooth or etched, is formed, the iron heterogeneities exert their damaging influence in two ways:

1. The diameter of iron aluminide crystals is of the order of 1μ , while the dielectric

film has a maximum thickness of about $.7\mu$. Wherever an iron crystal is reached during the forming process, the forming layer develops an abnormal composition throughout its entire thickness. The resulting mixed oxide, which contains a large percentage of iron, is a semi-conductor and allows the passage of electrons. This of course is tantamount to a deterioration of the leakage current.

2. At the same time these iron heterogeneities constitute minute electrolytic cells which generate currents giving rise to corrosion. They are less active than pure aluminum by about .03 v in the electrochemical series⁴. In this manner the electrochemical decomposition of the forming layer and the aluminum, known as deformation, is accelerated. This in turn results in a marked reduction of the expected shelf life.

The above shows clearly that switching from 99.85% aluminum to "raffinal" will result in a considerable reduction of leakage currents, and an extension of shelf life^{5,6}. It is true that such a switch results in higher costs, but in general these amount to no more than about 1 to 4% of the total cost of the finished capacitor. In Europe a good percentage of all electrolytic capacitors is today equipped with etched raffinal anodes⁷.

III. Results of the Composition of the Oxide Layer

This refers to the second point mentioned in the Introduction.

Even where the anodes of an electrolytic capacitor (respectively both electrodes in the case of AC capacitors) are made of 99.99% pure aluminum, there is no guarantee that the production methods used will in every case result in a capacitor of optimal attributes. Certain rules leading to good quality are already universally observed in the production of electrolytic capacitors - such as avoidance of chloride ions, forming with the greatest possible current density, etc.^{1,8,9}.

The dielectric constitutes the physically decisive component of any capacitor. The structure of the dielectric films on formed aluminum has been repeatedly investigated^{10,15}. Generally speaking, these investigations showed that this film, at forming potentials above about 100 v, consists of γ Al_2O_3 . Nevertheless some questions of paramount importance concerning the dielectric on formed aluminum remain unanswered to this day - especially concerning the behavior of the dielectric film during "deformation" and the composition of forming layers after the preforming process. In our investigations an attempt was made to clear up these problems. The structure of the forming layers was examined, especially by means of successive removals of thin films. This was done with a special acid blend which has the property of dissolving loosely constructed aluminum oxides without at the same time attacking metallic aluminum or stable forms of oxide, such as γ Al_2O_3 or corundum. Among others, a mixture of

chromic and phosphoric acid was used. Time and space does not permit going into further details concerning this method.

When treating the formed electrodes of commercial electrolytic capacitors with this acid mixture it becomes apparent that the forming layer consists of two strata - (1) an upper, readily soluble layer, by us designated as $\gamma_1 Al_2O_3$, and (2) a lower insoluble layer, consisting of γ Al_2O_3 . The thickness of the insoluble γ layer is easily determined by dissolving the metallic aluminum in methanol bromide after removing the γ_1 layer. This leaves the stable γ layer which is practically insoluble, even in hydrochloric acid.

In this manner a number of different forming layers was examined, and the following was found (Figs. 5 and 6):

The ratio $\gamma : \gamma_1$ increases as the forming potential is increased. However, even in capacitors of equal potentials it can fluctuate greatly. On the right hand side of Fig. 5 the methods for determining this ratio are indicated. About 25 film removals were made under controlled conditions, and the weight loss was determined in each case. As soon as no further weight loss occurs with the chromic-phosphoric acid treatment, the metallic aluminum is dissolved in methanol bromide. There remains only the γ layer, the thickness of which can then be determined. The thickness of the γ_1 layer is arrived at by adding the weight losses brought about by the successive removal processes. All results are immediately converted into terms of oxide film thickness. Fig. 5 shows that the percentage proportion of $\gamma : \gamma_1$ can vary considerably, even at equal forming potentials. It will be shown later that this is due to variations in forming methods.

Fig. 6 illustrates how the fractional composition of the forming layer depends upon the forming potential. It will be seen that the stable γ stratum increases rapidly as the forming potential is increased. This corroborates the findings in¹¹. X-Ray films however do not show any pronounced difference between the γ and the γ_1 oxide, even though the difference with respect to solubility is very great. In the following remarks we will confine ourselves exclusively to high voltage layers, where the forming potential amounts to 540 v. In cases where an amorphous or a hydrated oxide layer was present prior to the forming process, our findings were particularly revealing. During formation the dielectric is produced underneath these pre-forming layers, which latter are partly consumed in the process (Fig. 7).

The following types of preforming layers were tested:

1. Boehmite layers (Boehmite is a crystalline hydrated aluminum oxide, $Al_2O_3 \cdot x H_2O$ and can

be applied without electric current, by means of other suitable methods)^{16,17}.

2. Amorphous oxide layers, largely anhydrous - obtained by preformation in oxalic acid.

Fig. 7 shows that the reactor of crystalline Boehmite is entirely different from that of amorphous preforming layers. It will be seen that in forming the capacitor (designated as "re-formed" in the illustration) the share of stable oxide is further increased. This relative increase is particularly pronounced where an amorphous preforming layer is used. The "re-forming" time amounted to thirty hours, at room temperature.

In a further series of experiments, the re-formed capacitors were exposed to induced deformation, by unconnected storage at 65°C over a period of two weeks. Strangely enough, the stable part of the oxide layer reverts to approximately the value originally obtained after forming in hot watery boric acid solution (Figs. 8 and 9).

This tends to indicate that there are actually two modifications of the practically insoluble χ oxide. The χ oxide developed during the extended forming time during re-formation appears to be less stable than that developed during formation, and while deforming, can easily snap back into the readily soluble χ_1 modification. In how far this process is a reversible one that can be repeated at will, is not as yet established. Figs. 8 and 9 also show the distinct difference between the forming layers developed underneath the Boehmite and underneath the amorphous preforming layers. The former have a larger share of stable χ oxide, even after deformation. By far the most stable layers are obtained where an oxide-free aluminum foil is formed at high temperatures and with high current densities.

In addition to all this, the capacity was determined after each successive film removal (Fig. 10).

The following results were obtained:

When using an amorphous oxide film as preforming layer, there appears a porous stratum, dielectrically inactive, on the surface exposed to the electrolyte. We designate this as χ_2 . This is probably the original amorphous oxide layer, also known as α Al_2O_3 in its unchanged form. On the other hand, where Boehmite is used as the preforming layer, the successive removals occur more rapidly at the expense of the dielectric, as is shown by the immediate capacity increase. Thus the χ_2 layer appears to be absent, or at least very much reduced, in this case.

It seems that the χ_2 film has a very useful function in the electrolytic capacitor. It protects the dielectric, which it covers, against

solution but at the same time, due to capillary action, supports thorough wetting, so that this film structure produces a reduction of the deforming process, and also a low power factor. Of course the layer reacts differently, according to the viscosity and aggressiveness of the electrolyte used, but in general its function is a beneficial one. A Boehmite preforming layer likewise suppresses the deformation, but generally to a lesser extent, since the dielectric is practically in direct contact with the electrolyte which causes the deformation.

All of the above shows that it cannot be stated categorically that all forming layers ought to consist of as small a percentage of soluble oxide layer as possible. Theoretically, of course, a forming layer consisting entirely of χ oxide is preferred. Up to the present however it has not been possible to produce such a film. All dielectrics examined had a χ_1 stratum of varying thickness. This readily soluble stratum should, so far as possible, not be exposed directly to the electrolyte. An intervening layer, to protect it against solution, appears desirable.

Thus Fig. 10 shows that the capacity increase proceeds much more rapidly with successive removals of the dielectric, than would be expected with a presumably uniform removal of the dielectric substance. This could be explained by assuming that the dielectric is attacked more deeply in certain weak spots (Fig. 11). The nature of these spots, which have a large localized share of χ_1 oxide, is not yet known with certainty.

It can be shown that the preferential attack on these weak spots, as the dielectric film is successively removed, increases sharply as the degree of purity of the aluminum decreases (Fig. 12). Thus locally the iron heterogeneities cause an increase in the percentage of the more easily soluble oxide within the dielectric film. Furthermore there is a systematic difference between the reactions of smooth and strongly etched aluminum anodes. It appears certain that deeply etched pockets contain a high percentage of the readily soluble oxide, since the forming field intensity is reduced there, as the narrow pores act in the manner of a Faraday cage (Fig. 13). Our investigations have shown that there exist favorable as well as unfavorable geometric etching patterns with respect to the stability of the capacitor. A saw-tooth type of pattern is preferable to one having deep pores.

Another point to be made is that the presence of a χ_1 layer is probably closely related to the mechanics of building up the forming layer, since the χ_1 stratum represents the actual growth zone, into which the anions which participate in the construction of the layer can penetrate up to a certain depth (Fig. 14). The actual χ layer is undoubtedly traversed only by the aluminum ions, since due to their small size only they are capable of penetrating it.

Theoretically this loss of aluminum ought to result in the formation of empty space underneath the χ layer. However this closes up automatically due to the extraordinarily high pressures caused by the prevailing high field intensity, since the oxide film is firmly pressed against the aluminum grid. Actually there is no well-defined line of demarcation between the aluminum grid and the χ grid. It is much more likely that the lower parts of the χ layer contain an excess of aluminum atoms, and thus retain a certain amount of conductivity¹⁸.

The fact that the χ_1 and the χ layers are indistinguishable in an X-Ray diagram appears readily understandable. Apparently the two structures are nearly identical, since they can be easily converted from one to the other under the influence of an electric field, or in deformation processes. Only the χ part developed during high temperature preforming seems to be immune to this conversion into χ_1 .

We have already seen that in the mechanism of building up the forming layer the amorphous or hydrated preforming layers are partly consumed. At the prevailing high field intensities and temperatures a recrystallization into the χ_1 modification is entirely plausible. In the case of an amorphous oxide film it is true that a residual part remains on top, but this results in generally beneficial effects, such as suppression of the deformation process (Fig. 15).

It is particularly when the Boehmite layers present prior to formation are too thick that the constitution of the forming layer and the suitability of the dielectric become problematical. For one thing, it is not clear how the water content, which theoretically amounts to 15%, can escape from the lower strata of the Boehmite layer which is to be reconstructed. As the Boehmite film increases in thickness, the capacity becomes greater by 5 to 15%. This would indicate that the water, which has a dielectric constant of 80, becomes incorporated into the dielectric. Incidentally, on smooth foils the hydrated or amorphous oxide film remain virtually unconsumed; essentially they are left above the forming layer. However in etched foils the consuming effect can be clearly observed.

While it is true that Boehmite layers reduce the deformation in DC capacitors, they can, under unfavorable circumstances, increase losses. Normally, the crystalline Boehmite layer is not nearly as porous as an amorphous preforming layer, and the unconsumed upper stratum of the Boehmite layer is therefore likely to increase losses, acting as a noticeable series resistance. To what extent this loss-increasing upper layer merges into the χ_1 film is not easily determined, particularly since the χ_1 film may also, perhaps even primarily, be the center of the dielectric losses.

The search for a suitable equivalent circuit for a dielectric, led J. E. Lilienfeld and C.

Miller¹⁹ to the assumption, based entirely on electrical measurements, that it is the upper stratum in the dielectric which brings about losses, and which causes the dependence of the capacity upon frequency. This assumption is largely borne out by our own findings.

Hydrated aluminum oxides of widely differing structures can be produced in a continuous series¹⁶.

Our investigations have shown that for instance the crystalline Boehmite layers are less readily soluble in acids than the amorphous preforming layers, even where the latter were sealed off. For a number of capacitor types crystalline preforming layers are entirely suitable, provided their structure, hydration, and thickness are properly chosen.

IV. Conclusion

In order to widen the scope systematically along the lines indicated, further investigations are of course necessary. However it has already become possible to learn more about the structure and the behavior of forming layers, and thus to influence them systematically to provide whatever physical attributes are desired in the capacitor.

The surface condition, which for best results should be decided on prior to the forming process, of course represents only one link in a whole chain of factors where all links have to be properly attuned to one another in order to construct an electrolytic capacitor of optimal attributes. At the same time, the impression prevails that the structure of the dielectric has often unnecessarily been considered the weakest of these links.

To conclude with, here is a practical example: The photo flash capacitor has always been a problematic type of condenser. It requires a series of characteristics which in part are mutually contradictory:

1. Extremely small leakage currents, in order not to discharge the battery prematurely.
2. Good shelf life, so that the capacitors can be stored for months on end.
3. Low power factor, since this has a marked influence on the flash intensity.
4. Constant capacity. There is always the danger that during a sudden discharge the cathode becomes formed.
5. Very small dimensions, in order to produce handy portable equipment.

At first it was considered extremely difficult to fulfill all these requirements simultaneously. For example, in order to obtain a low power

factor it becomes necessary to employ an electrolyte of low viscosity, which however results in a deterioration of leakage currents as long as ordinary aluminum of 99.85% purity is used. Thus the ordinary types of electrolytic capacitors are entirely inadequate for photo-flash equipment²⁰. It is not surprising that early models of photo flash electrolytic capacitors resulted in many rejections. However with a 99.99% pure etched foil of high surface gain as the anode, and lightly etched foil, as thin as possible of ordinary aluminum foil as the cathode, all requirements can be met.

In using etched foil made of "raffinal," a preforming layer can ordinarily be dispensed with altogether. However for capacitors exposed to unfavorable conditions, such as high temperatures, extended unconnected storage, etc., a preforming layer may be advisable in order to reduce deformation.

Acknowledgements

The writer wishes to thank Dr. W. Herrmann, Dr. W. Ilge, and Dr. G. Dieckmann for suggestions and inspiration. He also wishes to express appreciation to Miss B. Eytel and Mr. W. Post for their painstaking care in carrying on experiments.

References

1. Alexander M. Georgiev, "The Electrolytic Capacitor," Murray Hill Books, Inc., New York.
2. "Refined Aluminum, Its Production, Properties and Applications," Light Metals I, pp. 207-211 and 250-253.
3. H. Bettler, "Production and Application of Raffinal," Aluminium Suisse, Vol. 3 (1953), p. 51.
4. R. B. Mears and R. H. Brown, "Causes of Corrosion Currents," Industrial and Engineering Chemistry, Vol. 53, Aug. 1941, p. 1001.
5. D. Altenpohl, "Aufgerauhte Anoden fuer Elektrolytkondensatoren," Elektrozentralblatt, Vol. 3, March 1952.

6. D. Altenpohl, "Improvements in the Field of Electrolytic Capacitors," 1954, published by Aluminium Walzwerke GmbH, Singen.
7. Aluminum News, September 1951.
8. P. R. Coursey, "Electrolytic Condensers, their Properties, Design and Practical Uses," London 1937, Chapman & Hall, Ltd.
9. A. Guntherschulze and H. Betz, "Elektrolytkondensatoren," Berlin 1937, published by M. Krayn.
10. W. Herrmann in Wissenschaftliche Veroeffentlichungen, Siemens & Halske, Werkstoff-Sonderheft 1940, pp. 188-212.
11. C. S. Taylor, C. M. Tucker and J. D. Edwards, "Anodic Coatings with Crystalline Structure on Aluminum," Tr. Electrochem. Soc., Vol. 88 (1945), pp. 325-332.
12. J. D. Edwards and F. Keller, "Formation of Anodic Coatings on Aluminum," Tr. Electrochem. Soc., Vol. 79 (1941), p. 135.
13. Robinson & Burnham, Tr. Electrochem. Soc., Vol. 83 (1948), p. 125.
14. Max Schenck, "Werkstoff Aluminium und seine anodische Oxydation," Bern 1948, Verlag A. Francke A.G.
15. E. I. W. Verwey, "The Structure of the Electrolytical Oxide Layer on Aluminum," Z. anorg. Chemie, Vol. 91 (1935), p. 317.
16. R. Fricke and O. Eberspaecher, Z. anorg. Chemie, Vol. 21 (1951), p. 265.
17. D. Altenpohl, Z. Aluminium, Vol. 29 (1953), pp. 361-370.
18. W. Ch. Van Geel and J. W. A. Scholte, Philips Research Rep., Vol. 6 (1951), pp. 54-74.
19. J. E. Lilienfeld and C. Miller, J. Electrochem. Soc., Vol. 100 (1953), pp. 222-226.
20. G. Dieckmann, Photo-Technik und-Wirtschaft, Vol. 5, 1953.

Theoretically this loss of aluminum ought to result in the formation of empty space underneath the γ layer. However this closes up automatically due to the extraordinarily high pressures caused by the prevailing high field intensity, since the oxide film is firmly pressed against the aluminum grid. Actually there is no well-defined line of demarcation between the aluminum grid and the γ grid. It is much more likely that the lower parts of the γ layer contain an excess of aluminum atoms, and thus retain a certain amount of conductivity¹⁸.

The fact that the γ_1 and the γ layers are indistinguishable in an X-Ray diagram appears readily understandable. Apparently the two structures are nearly identical, since they can be easily converted from one to the other under the influence of an electric field, or in deformation processes. Only the γ part developed during high temperature preforming seems to be immune to this conversion into γ_1 .

We have already seen that in the mechanism of building up the forming layer the amorphous or hydrated preforming layers are partly consumed. At the prevailing high field intensities and temperatures a recrystallization into the γ_1 modification is entirely plausible. In the case of an amorphous oxide film it is true that a residual part remains on top, but this results in generally beneficial effects, such as suppression of the deformation process (Fig. 15).

It is particularly when the Boehmite layers present prior to formation are too thick that the constitution of the forming layer and the suitability of the dielectric become problematical. For one thing, it is not clear how the water content, which theoretically amounts to 15%, can escape from the lower strata of the Boehmite layer which is to be reconstructed. As the Boehmite film increases in thickness, the capacity becomes greater by 5 to 15%. This would indicate that the water, which has a dielectric constant of 80, becomes incorporated into the dielectric. Incidentally, on smooth foils the hydrated or amorphous oxide film remain virtually unconsumed; essentially they are left above the forming layer. However in etched foils the consuming effect can be clearly observed.

While it is true that Boehmite layers reduce the deformation in DC capacitors, they can, under unfavorable circumstances, increase losses. Normally, the crystalline Boehmite layer is not nearly as porous as an amorphous preforming layer, and the unconsumed upper stratum of the Boehmite layer is therefore likely to increase losses, acting as a noticeable series resistance. To what extent this loss-increasing upper layer merges into the γ_1 film is not easily determined, particularly since the γ_1 film may also, perhaps even primarily, be the center of the dielectric losses.

The search for a suitable equivalent circuit for a dielectric, led J. E. Lilienfeld and C.

Miller¹⁹ to the assumption, based entirely on electrical measurements, that it is the upper stratum in the dielectric which brings about losses, and which causes the dependence of the capacity upon frequency. This assumption is largely borne out by our own findings.

Hydrated aluminum oxides of widely differing structures can be produced in a continuous series¹⁶.

Our investigations have shown that for instance the crystalline Boehmite layers are less readily soluble in acids than the amorphous preforming layers, even where the latter were sealed off. For a number of capacitor types crystalline preforming layers are entirely suitable, provided their structure, hydration, and thickness are properly chosen.

IV. Conclusion

In order to widen the scope systematically along the lines indicated, further investigations are of course necessary. However it has already become possible to learn more about the structure and the behavior of forming layers, and thus to influence them systematically to provide whatever physical attributes are desired in the capacitor.

The surface condition, which for best results should be decided on prior to the forming process, of course represents only one link in a whole chain of factors where all links have to be properly attuned to one another in order to construct an electrolytic capacitor of optimal attributes. At the same time, the impression prevails that the structure of the dielectric has often unnecessarily been considered the weakest of these links.

To conclude with, here is a practical example: The photo flash capacitor has always been a problematic type of condenser. It requires a series of characteristics which in part are mutually contradictory:

1. Extremely small leakage currents, in order not to discharge the battery prematurely.
2. Good shelf life, so that the capacitors can be stored for months on end.
3. Low power factor, since this has a marked influence on the flash intensity.
4. Constant capacity. There is always the danger that during a sudden discharge the cathode becomes formed.
5. Very small dimensions, in order to produce handy portable equipment.

At first it was considered extremely difficult to fulfill all these requirements simultaneously. For example, in order to obtain a low power

factor it becomes necessary to employ an electrolyte of low viscosity, which however results in a deterioration of leakage currents as long as ordinary aluminum of 99.85% purity is used. Thus the ordinary types of electrolytic capacitors are entirely inadequate for photo-flash equipment²⁰. It is not surprising that early models of photo flash electrolytic capacitors resulted in many rejections. However with a 99.99% pure etched foil of high surface gain as the anode, and lightly etched foil, as thin as possible of ordinary aluminum foil as the cathode, all requirements can be met.

In using etched foil made of "raffinal," a preforming layer can ordinarily be dispensed with altogether. However for capacitors exposed to unfavorable conditions, such as high temperatures, extended unconnected storage, etc., a preforming layer may be advisable in order to reduce deformation.

Acknowledgements

The writer wishes to thank Dr. W. Herrmann, Dr. W. Ilge, and Dr. G. Diekmann for suggestions and inspiration. He also wishes to express appreciation to Miss B. Eytel and Mr. W. Post for their painstaking care in carrying on experiments.

References

1. Alexander M. Georgiev, "The Electrolytic Capacitor," Murray Hill Books, Inc., New York.
2. "Refined Aluminum, Its Production, Properties and Applications," Light Metals I, pp. 207-211 and 250-253.
3. H. Bettler, "Production and Application of Raffinal," Aluminium Suisse, Vol. 3 (1953), p. 51.
4. R. B. Mears and R. H. Brown, "Causes of Corrosion Currents," Industrial and Engineering Chemistry, Vol. 53, Aug. 1941, p. 1001.
5. D. Altenpohl, "Aufgerauhte Anoden fuer Elektrolytkondensatoren," Elektrozentralblatt, Vol. 3, March 1952.

6. D. Altenpohl, "Improvements in the Field of Electrolytic Capacitors," 1954, published by Aluminium Walzwerke GmbH, Singen.
7. Aluminum News, September 1951.
8. P. R. Coursey, "Electrolytic Condensers, their Properties, Design and Practical Uses," London 1937, Chapman & Hall, Ltd.
9. A. Guntherschulze and H. Betz, "Elektrolytkondensatoren," Berlin 1937, published by M. Krayn.
10. W. Herrmann in Wissenschaftliche Veroeffentlichungen, Siemens & Halske, Werkstoff-Sonderheft 1940, pp. 188-212.
11. C. S. Taylor, C. M. Tucker and J. D. Edwards, "Anodic Coatings with Crystalline Structure on Aluminum," Tr. Electrochem. Soc., Vol. 88 (1945), pp. 325-332.
12. J. D. Edwards and F. Keller, "Formation of Anodic Coatings on Aluminum," Tr. Electrochem. Soc., Vol. 79 (1941), p. 135.
13. Robinson & Burnham, Tr. Electrochem. Soc., Vol. 83 (1948), p. 125.
14. Max Schenck, "Werkstoff Aluminium und seine anodische Oxydation," Bern 1948, Verlag A. Francke A.G.
15. E. I. W. Verwey, "The Structure of the Electrolytical Oxide Layer on Aluminum," Z. anorg. Chemie, Vol. 91 (1935), p. 317.
16. R. Fricke and O. Eberspaecher, Z. anorg. Chemie, Vol. 21 (1951), p. 265.
17. D. Altenpohl, Z. Aluminium, Vol. 29 (1953), pp. 361-370.
18. W. Ch. Van Geel and J. W. A. Scholte, Philips Research Rep., Vol. 6 (1951), pp. 54-74.
19. J. E. Lilienfeld and C. Miller, J. Electrochem. Soc., Vol. 100 (1953), pp. 222-226.
20. G. Dieckmann, Photo-Technik und-Wirtschaft, Vol. 5, 1953.

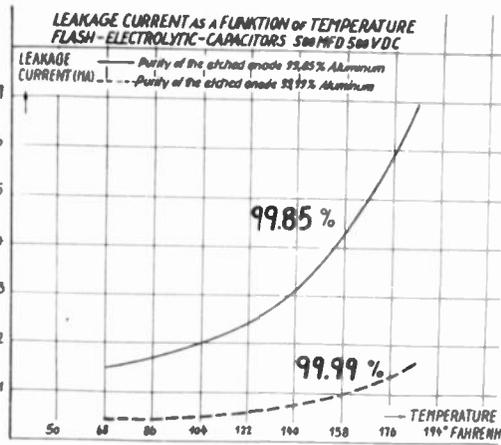


Fig. 1(a)
Decreasing purity causes increasing leakage.

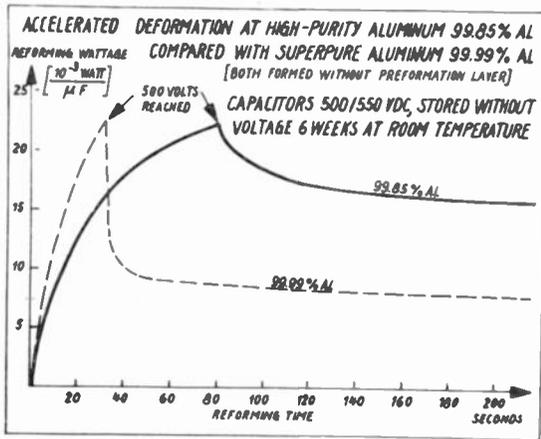


Fig. 1(b)
Decreasing purity causes increasing deformation.

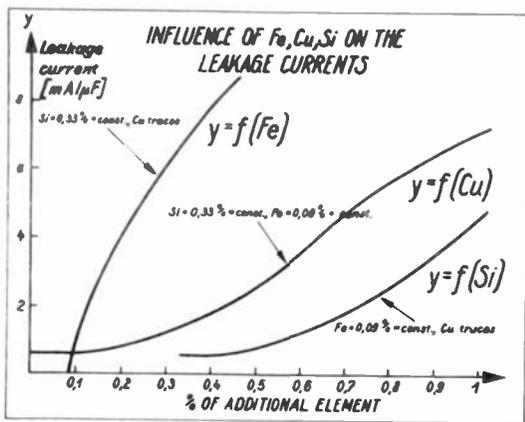


Fig. 2
Iron, the most detrimental alloying element.
(Fig. from A. Petrowski.)

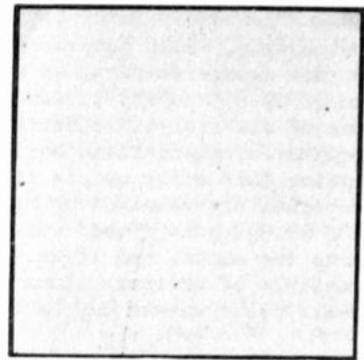


Fig. 3(a) - Superpure (99.99% aluminum).



Fig. 3(b) - High purity (99.85% aluminum).

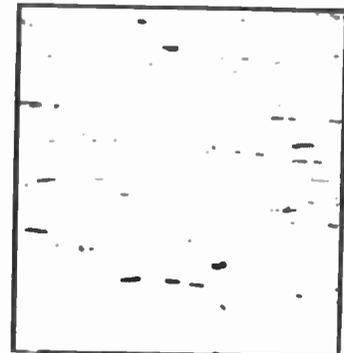


Fig. 3(c) - Normal purity (99.5% aluminum).

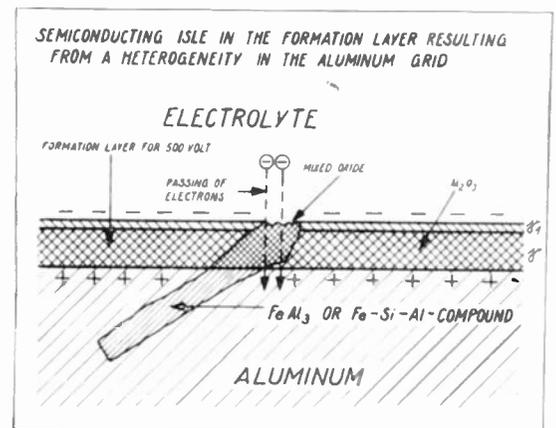


Fig. 4
Weak point in a dielectric film, resulting from an iron-heterogeneity.

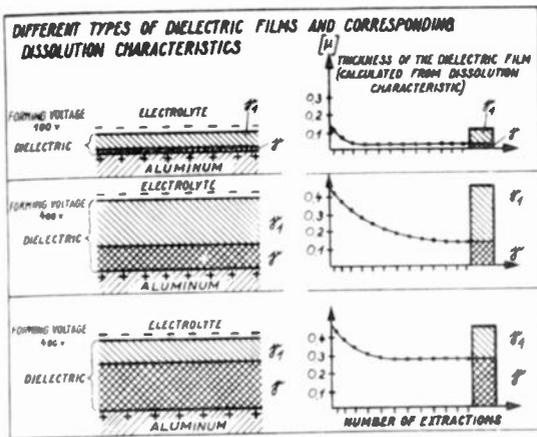


Fig. 5

The ratio, that bears the easy soluble γ - Al_2O_3 to the insoluble δ - Al_2O_3 varies widely (also at the same type of capacitor).

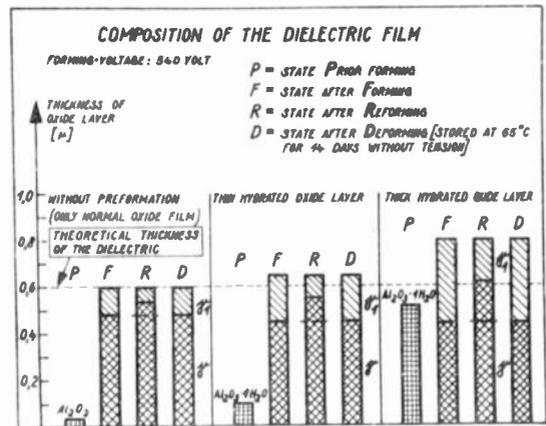


Fig. 8

Crystalline hydrated oxide as preformation layer.

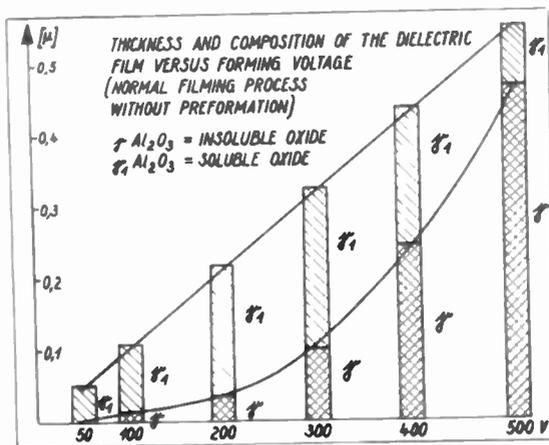


Fig. 6

The stable δ oxide increases with increasing forming voltage.

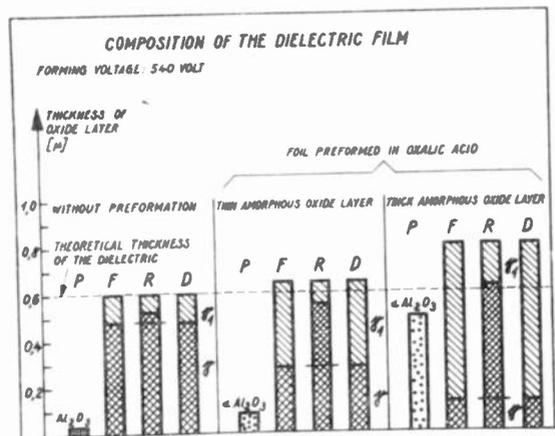


Fig. 9

Amorphous oxide as preformation layer.

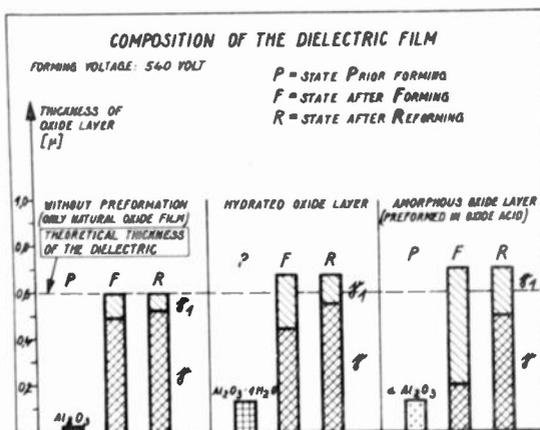


Fig. 7

Influence of different preformation layers.

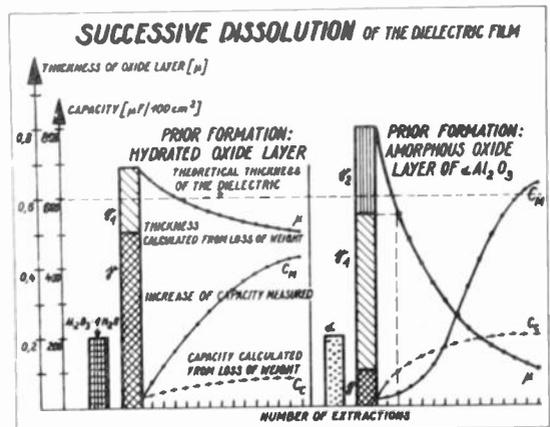


Fig. 10

Comparison of (crystalline) hydrated and amorphous preformation layers.

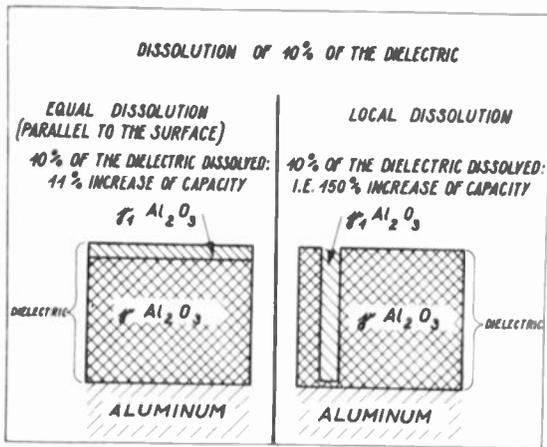


Fig. 11
Equal distribution or local accumulation of the δ_1 oxide results in great differences in the capacity-increase.

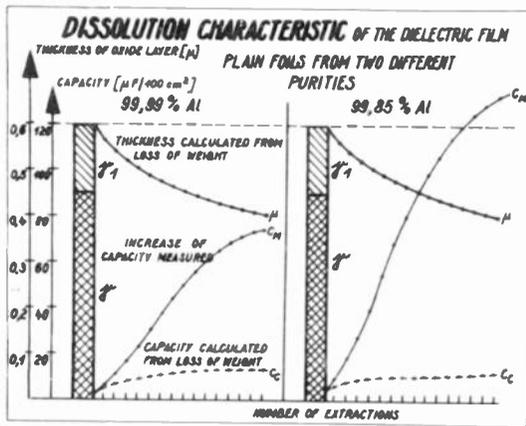


Fig. 12
By diminution of aluminum purity the local dissolution of the dielectric increases.

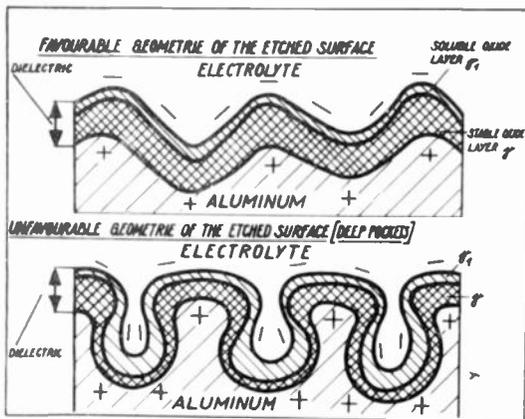


Fig. 13
The cage-effect causes in deep etching pores more easy-soluble δ_1 oxide.

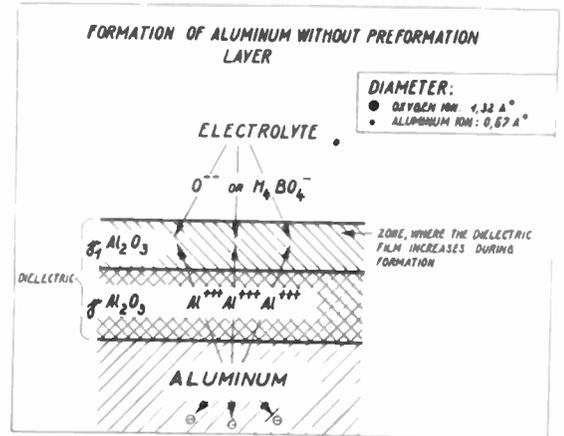


Fig. 14
Construction of the dielectric film by passing of Al^{+++} through the δ stratum.

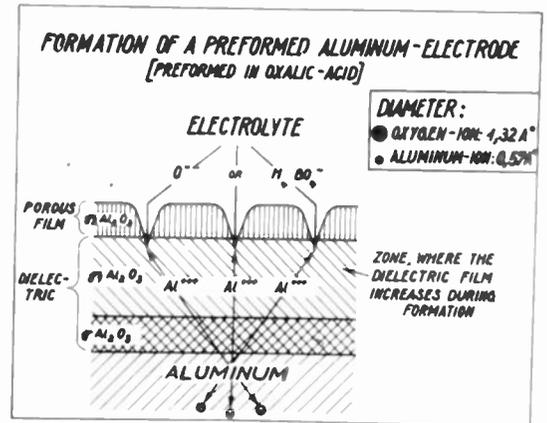


Fig. 15
Construction of the dielectric under an amorphous preformation layer.

AN INVESTIGATION OF LOWEST RESONANT
FREQUENCY IN COMMERCIALY-AVAILABLE BYPASS CAPACITORS

David T. Geiser
Sprague Electric Company
North Adams, Massachusetts

Examination of paper, mica, and ceramic capacitors for lowest resonant frequency was made of samples intended for the mass home-equipment market. The investigation was limited to the ranges 0.001 to 1.0 microfarads and zero to 100 megacycles. Effort was made to minimize external circuit effects, and suggestions are made for standard methods and jigs. Application information is included.

Introduction

The logical place to begin a discussion of bypass capacitor characteristics is with definition of the term "bypass capacitor". Literature search^{4,5,23} showed that two application definitions covered the common use of the term.

In Figure I, the capacitor is represented by the conventional 4-pole matrix "Black Box".

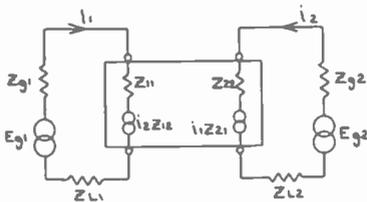


Figure I - Capacitor Application As A 4-Pole

In one desired "bypass" (coupling) application, $Z_{21} = Z_{g1}$, $Z_{12} = Z_{L2}$, and $Z_{11} = Z_{22} = Z_{L1} = Z_{g2} = 0$. In the second application, $Z_{21} = 0$, $Z_{12} = 0$, giving perfect isolation of the input and the output circuit^{18,19}. Frequently, the designer combines applications in relatively low impedance circuits and encounters serious trouble. Many times this difficulty may be traced to the fact that Z_{11} and Z_{22} are not zero or that Z_{21} and Z_{12} are likewise not zero. Inherent inductance in capacitor manufacture or application is frequently to blame.

The Two-Terminal Capacitor

The capacitor is often considered to be the two-terminal element of Figure II⁶. Strictly,



Figure II - Single Frequency Equivalent

this equivalence (for specified R, L, and C) applies only at a single frequency and element current. Assuming that element currents and resistances are negligible, a more accurate equivalence can be drawn in the form of Figure III¹⁰.

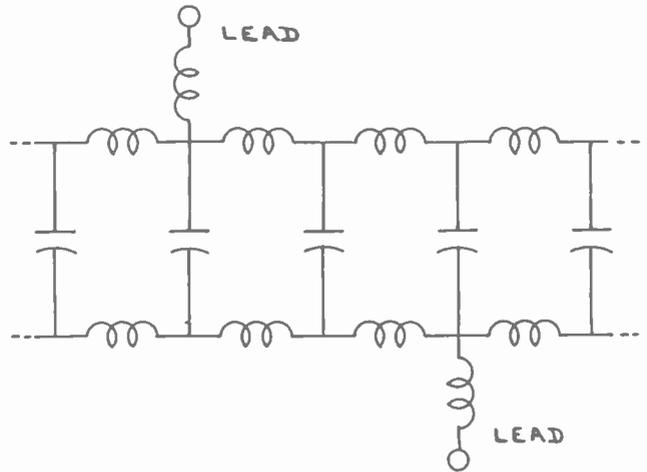


Figure III - Lossless Equivalent

In the commonly-used capacitors the leads exhibit more inductance than the capacitor proper. Several excellent charts have appeared in the literature^{11,16,22,23} to show the effect of lead inductance, and, to a certain extent, the charts of this paper display this effect.

A second effect should be noted: The nearer opposite the lead connections, the lower the effective series inductance becomes¹⁰. This is demonstrated by an increase in the two-terminal resonant frequency. Moving the lead connections farther apart lowers the effective resonant frequency, a trick that when applied to broadcast receiver bypass capacitors allows almost perfect isolation of circuits. Likewise, amateurs and others sometimes increase lead inductance²² to improve bypass action.

Resonance in capacitors can be useful, but most of the time this effect is only a liability. How can the resonant frequency be raised?^{10,16,23} First: Shorten the leads to minimum length. Second: Widen the leads. While considering lead shortening, the engineer should remember that the capacitor body is a lead representing the absolute minimum inductance available for those particular body dimensions. Thus, otherwise identical capacitors of different shapes will possess different resonant frequencies. The ultimate example of lead widening is to extend each capacitor plate and use these protruding ends for leads. This

"extended foil" construction is the basis for the highest resonant frequency rolled capacitors available.

Naturally, inductance in the external circuit adds to lead and structure inductance to cause further lowering of resonance. The capacitor structure of Figure IV shows a method used by the manufacturer to reduce mandatory inductance in the external loop. Here, for example, can be an extended-foil rolled capacitor, with connections BB'

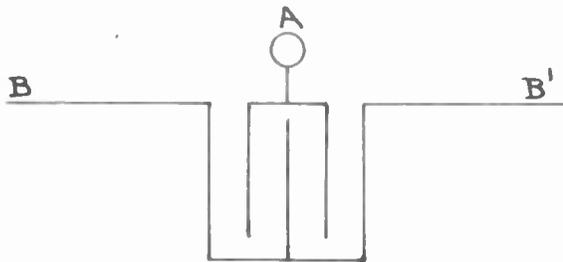


Figure IV - Shielded Capacitor Construction

the edge view of a disc. For absolute minimum inductance, this plate must be continuously connected to ground around its circumference and the connection of lead A to this plane must be short and wide. The limit is reached when lead A is also a disc. (This condition has been closely approximated in application by the writer.)

The Investigation

The portion of the investigation upon which this report is based concerned bypass capacitors intended for the mass home equipment market. Thus, the preceding discussion has been mostly a behind-the-scenes glimpse of some of the pertinent variables that should be considered when less expensive capacitors are not satisfactory.

Resonance is defined in this work as that frequency, progressing from zero, where the first minimum is to be found in the two-terminal impedance. Unless power absorption or transmission methods of measurement are used, making the measurement can be very tedious if the minimum impedance frequency does not correspond to the zero reactance frequency.

Practical considerations compel a decision of the part external leads are to play in the gathering of data. Authors differ in approach and with some exceptions either describe unknown or experimental capacitors or do not consider the proximity of other objects. It was decided to conduct three "minimum lead" tests, one with a standard reasonable wire return path (Figure VA), one with a standard composite wire and sheet return path (Figure VB), and one with total enclosure sheet return path (Figure VC).

The plate opening shown in Figure VC is to allow capacitor insertion and was closed by a covering metal sheet during test, yielding a zero (or negligible) inductance return path external to the capacitor case. In early tests, the capaci-

tor enclosure was a mercury-filled well. Subsequent use of a tight wrap of aluminum foil instead of the mercury bath permitted test non-destructive to both capacitors and jig while introducing much less than two percent error in average resonant frequency determination.

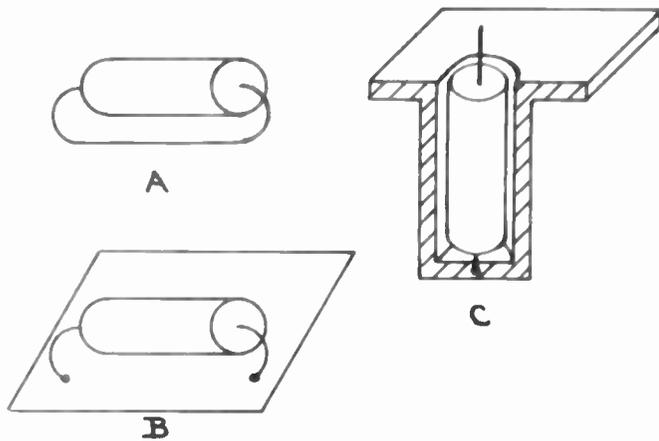


Figure V - Test Jigs

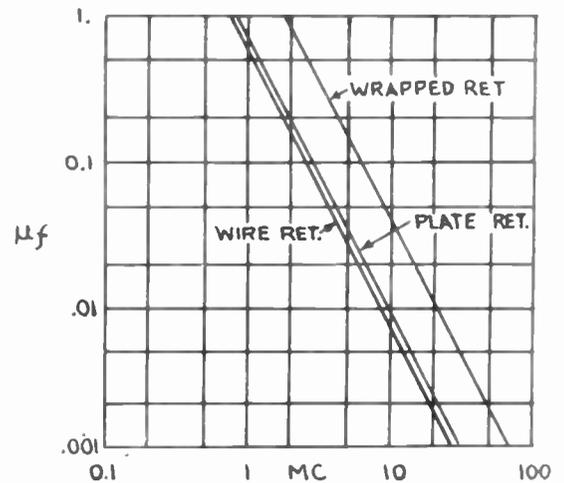


Figure VI-Resonance, Inserted Tab Paper Capacitors

Figure VI shows the results obtained by the three test methods for inserted-tab wound capacitors. As with the following graphs, capacities plotted are actual capacitance measured to better than one percent accuracy at one kilocycle. Frequency measurement was made to much greater accuracy. It should be noted that these curves each show several brands and that almost perfect agreement exists. Where some disagreement occurred, the curves are loaded to indicate approximate mean values of resonant frequency and proportional production. Maximum deviation from the curves was eight percent of frequency; the five-sigma point was close to five percent of frequency. Should these curves be used for design work, please remember actual, not nominal or marked capacitance is plotted.

Figure VII shows similar curves for metallized paper capacitors. These least expensive metallized units usually showed higher impedance at resonance than paper units generally. Note the higher resonant frequencies caused by smaller case dimens-

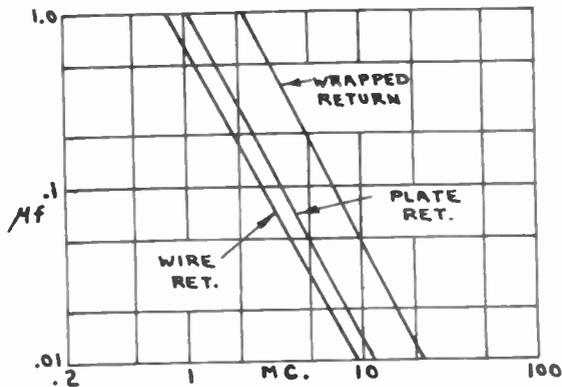


Figure VII - Metallized-Capacitor Resonance

ions. An interesting point about this graph is that the only other published resonance curves for metallized capacitors show much higher resonant frequencies, leading this writer to believe the other investigator's work²³ reflected results of interference-filter quality capacitor study, rather than the common metallized bypass.

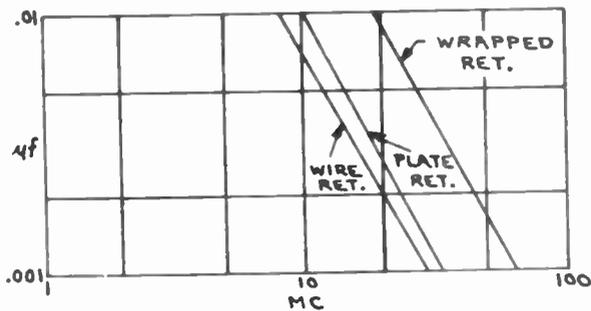


Figure VIII - Mica Capacitor Resonance

Figure VIII shows curves for least expensive mica capacitor types. Note that these curves agree almost perfectly with inserted tab rolled paper capacitors. This would indicate that the mica type should offer improvement only when higher Q is required, but in this cheap form offers practically no increase in resonant frequency.

Figure IX offers composite curves of disc and (above 0.02 microfarads) rectangular ceramic capacitors. A fourth curve has been added, showing the effect of mounting the capacitor parallel to the return sheet at a distance of one body-thickness. There are two general points of interest in these types of ceramic capacitors: they have the highest resonant frequency for a

given capacity and the rectangular types have the highest external inductive field.

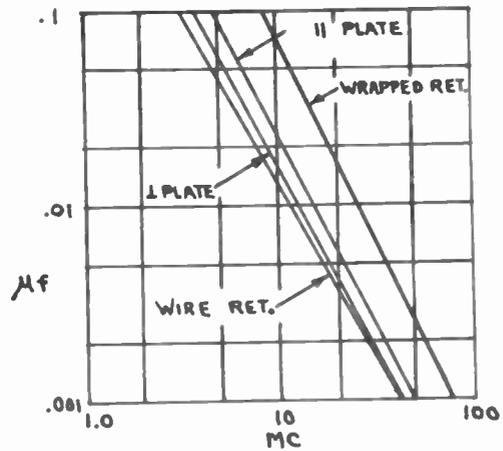


Figure IX - Ceramic Capacitor Resonance

If the user is choosing by nominal capacity, this higher resonant frequency may be more than balanced by the common practice of manufacturing capacitors with generous excess above nominal capacitance. Some particular sizes and brands may show as much as one hundred percent excess capacity.

Test Methods

The combinations of capacitance, inductance, and resistance in each capacitor makes bridge determination of resonant frequency very tedious. Time can be saved by locating approximate resonance with a grid-dip meter and making a detailed bridge examination in that area. For most engineering uses, however, careful grid-dip measurement is sufficiently accurate. Wire and sheet return measurements were made in this fashion. The writer's cylindrical-return measurements were made by the four-pole method of Figure X.

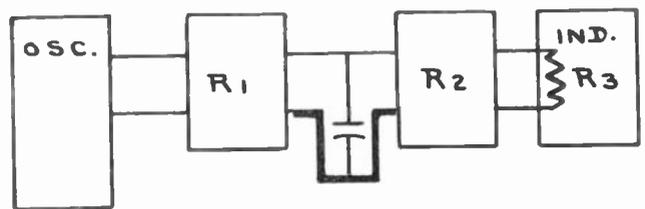


Figure X - Four-Pole Transfer Impedance Test

Here the transfer impedance of all effective branches were maintained at 50 ohms excepting frequencies where the capacitor exhibited relatively low impedance. Equally valid results have been obtained by other workers^{11,16} using the circuit of Figure XI, but those published results are carried only to 0.1 inch minimum lead length and cannot be directly correlated to these zero lead inductance measurements.

It should be noted that the electrical circuit length from the oscillator to the jig should be a very small fraction of wavelength to avoid standing wave effects. These last methods in their many possible variations offer means of checking resonance for virtually any lead length and in addition give accurate and quick approximations of "short-circuit" and "isolation" by-pass effect.

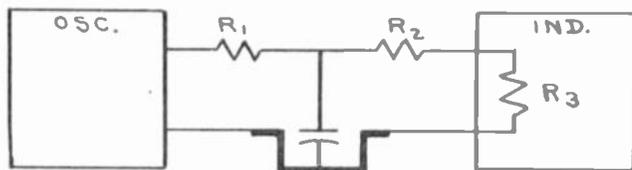


Figure XI - "T" Transfer Impedance Test

These last methods of test are easily adapted to use of a sweeping signal generator and oscilloscope to visually indicate resonance. Figure XII is a photograph of the response curve of a capacitor undergoing the four-pole test of Figure X. The dip to the left shows the network response approaching zero frequency, the deep dip shows capacitor resonance, and the shallow dip to the right shows one of the reasons for this paper: a capacitor resonance in the oscilloscope input circuit.

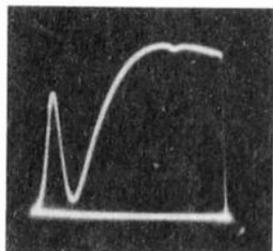


Figure XII - Resonance, Single Capacitor

The photograph of Figure XIII shows the effect of adding a smaller parallel capacitor in a simulated attempt to obtain better high-frequency by-passing. As may be seen, that method is of limited usefulness.

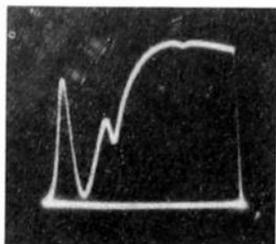


Figure XIII - Resonance, Paralleled Unequal Capacitors

One last comment about test methods used in preparation of this paper: all lead bends were made on one-eighth inch radius starting one-eighth inch from the capacitor.

Conclusions

Several general conclusions may be drawn from the literature and discussion here. These may be simply summarized:

1. Capacitors have inductance and resonance.
2. Avoid resonant frequencies unless commercially available resonant capacitors are used or unless the job can be hand-tailored.
3. Do not parallel unlike capacitors unless an anti-resonance is allowable. Instead use decoupling networks or special feed-through capacitors properly installed.
4. Do not operate a capacitor above its resonant frequency if the circuit is sensitive to the presence of inductance.
5. When in doubt about any capacitor application, ask the manufacturer.

Acknowledgement

These results include tests on more than eight hundred capacitors of 1951 production and two hundred of 1953-1954 production. The early work was financed by a sponsor who has asked to remain anonymous. In addition, the writer is indebted to his present employer, Sprague Electric Company, for facilities to enable the later work and assistance in this publication.

Chronological Bibliography

1. H. J. McLeod, "Variation with Frequency of the Power Loss in Dielectrics", Phys. Rev., Vol. 21, Jan. 1923, p. 56.
2. S. L. Brown and M. Y. Colby, "Electrical Measurements at Radio Frequencies", Phys. Rev., (Ser. 2) Vol. 29, May 1927, p. 717.
3. E. B. Moullin, "Theory and Practice of Radio Frequency Measurements", Chas. Griffin and Co. Ltd., London, 1931, pp. 259-285.
4. I. G. Maloff, "Mica Condensers in H. F. Circuits", Proc. I.R.E., Vol. 20, April 1932, pp. 647-656.
5. A. L. M. Sowerby, "Non-Inductive Condensers", Wireless World, Vol. 31, July 15, 1932, pp. 33-34.
6. R. F. Field and D. B. Sinclair, "Inductance and Resistance of Variable Air Condensers", Proc. I.R.E., Vol. 24, pp. 255-274.
7. W. H. Ward, "Self-Inductance of Variable Air Capacitors", J. Sci. Instr., Vol. 13, August 1936, pp. 251-260.
8. L. Hartshorn and W. H. Ward, "H. F. Chokes,

- Condensers, and Resistors", J. Sci. Instr., Vol. 14, April 1937, pp. 132-135.
9. v. L. Linder and J. Schniederermann, "Behavior of Paper Condensers at High Frequencies", Zeits. f. Fernmeldetechnik, Vol. 18, May 1937, pp. 73-77 (in German).
 10. v. L. Linder and J. Schniederermann, "Effect of Self Induction of Wound Condensers on their Impedance", F. T. Z., Vol. 60, July 6, 1939, pp. 793-798, (in German).
 11. R. L. Haskins, "Resonance Chart, Capacitor and Leads", Electronics, Aug. 1939, p. 35 (and reverse).
 12. E. Leider, "Inductance and Damping of Wound Condensers", Hochfrequenztechn. u. Electro-Akustik, Vol. 54, Nov. 1939, pp. 153-156 (in German).
 13. Fomenko, "Self-Inductance of Capacitors", Izv. Electroprom. Slab. Toka, No. 9, 1940, pp. 47-53 (in Russian). English abstract in Wireless Engr., Vol. 20, Dec 1943, p. 627
 14. W. H. Cazaly, "Instruments: Test and Measuring Gear and its Uses, V", Wireless World, Vol. 48, Oct. 1942, pp. 232-236.
 15. Deeley, "Impedance of Electrolytic Capacitors at R.F.", Electronics, Apr. 1943, p. 209D.
 16. A. P. Green and C. T. McComb, "Resonance in Mica Capacitors", Electronics, March 1944, p. 119.
 17. F. I. Hall, "V.H.F. Behavior of Capacitors", Electronics, March 1944, p. 114.
 18. W. M. Allison and N. F. Beverly, "A New High-Frequency Capacitor", Electrical Engineering (Transactions) Vol. 63, Dec. 1944, pp. 915-916.
 19. A. Watton, Jr., "The Duct Capacitor", Proc. I.R.F., Vol. 36, April 1948, p. 550.
 20. H. W. Kline, "Increasing the Self-Resonant Frequency of Variable Capacitors", Radio-Electronic Ed., Radio and TV News, Vol. 11, July 1948, p. 13, 30.
 21. P. S. Rand, "Don't Pamper Your Harmonics", QST, Feb. 1951.
 22. G. Grammer, "Bypassing for Harmonic Reduction", QST, April 1951.
 23. R. Davidsoh, "R.F. Characteristics of Capacitors", Wireless World, Vol. 58, Aug. 1952, pp. 301-304.

RESOLUTION IN PRECISION WIRE-WOUND POTENTIOMETERS

Robert J. Sullivan, Senior Member IRE
Potentiometer Division, Fairchild Camera and Instrument Corp.
Hicksville, N. Y.

Summary-----This paper describes and defines the various types of potentiometer resolution for which some standardization has been accomplished. Several methods of measuring and analyzing these forms are discussed. Shorting resolution, or the "shorting effect" of the brush is described and illustrated and a distinction between normal, shorting and unshorting resolution is made. Effects of contact width, high wires, broad spacing of resistance wire turns, and speed of rotation on potentiometer resolution is discussed and illustrated. Resolution as an important factor in determining potentiometer linearity and conformity is discussed and a number of design criteria are established. Limits of resolution for various types of potentiometers are given and several potentiometer design limitations imposing restrictions on resolution are cited. Infinite resolution potentiometers of the deposited metal film type are briefly discussed.

INTRODUCTION

In the majority of precision potentiometers the electrical characteristics of greatest importance to the designer and user of the potentiometer are total resistance, linearity or conformity, noise and resolution. Linearity and conformity are basic terms used to define the accuracy of the output voltage with respect to mechanical input for linear potentiometers in the case of linearity and functional or non-linear potentiometers in the case of conformity. Both of these characteristics are "built-in" characteristics of the potentiometer and, as will be illustrated later in this paper, their values and limitations are primarily related to resolution. Noise,¹ which has been defined as the equivalent parasitic, transient, contact resistance, in ohms, appearing between the sliding contact of the potentiometer and the winding surface, to a very slight extent is related to resolution, but it is more dependent upon choice of contact materials, cleanliness of the winding and contact and other characteristics of the potentiometer which are more definitive of the quality of the unit rather than any "built in" mechanical or electrical feature, such as resolution.

Resolution or "granularity" of the potentiometer was recognized early in the development of the precision potentiometer as we know it today by the Potentiometer Group of the Radiation Laboratories M.I.T. as an important factor contributing to the accuracy of potentiometers developed and produced during World War II.² Rad Lab defined voltage resolution as the change in voltage from one step to the next for unit voltage.³ Basic contact design studies which led to the theory of rod contacts and their advantage in obtaining and maintaining the utmost of accuracy from a potentiometer of given resolution were

carried out by the M.I.T. group, and the results successfully applied to improve by a factor of nearly 100% the average linearity characteristics of so-called semi-precision linear potentiometers, the only type in large volume production during the early and middle stages of World War II.⁴

Terms defining other forms of resolution, such as ohmic resolution and angular resolution, were also defined by Rad Lab but with the tremendous changes in potentiometer design and performance resulting from post war development and production, most of these terms have subsequently proved to be inadequate to define the variety of forms of resolution existing today. Two standardization groups, the RETMA potentiometer sub-committee G2.1 and a committee of the AIA (Aircraft Industries Association) are currently working on the standards of definition, measurement and interpretation of all electrical and mechanical terms related to precision potentiometers, and have established tentative definitions for resolution terms as follows:

AIA Electrical Resolution The electrical resolution is a measure of the accuracy to which a potentiometer may be set. (The potentiometer wiper is moved through the smallest, finite angular displacement which will produce a change in resistance or voltage.) The maximum incremental change in resistance or voltage output observed anywhere in the total mechanical rotation of the potentiometer shaft or specified portion thereof, when multiplied by 100 and divided by the total variable resistance or voltage, or specified portion thereof, is the percentage resolution of the potentiometer.

AIA Angular Resolution The angular resolution of the potentiometer is the angular rotation of the shaft required to produce the increment output change as defined under electrical resolution.

The RETMA committee, which generally has attempted to obtain more complete coverage in its standardization work, defines five different terms related to resolution as follows:

Resolution Figure The resolution figure is the maximum value of any form of resolution over the entire resistance element or any specified portion thereof. (Note: For linear windings, specify a resolution figure. For non-linear windings, specify one or more resolution figures and the zone or zones over which they apply.)

Theoretical Voltage Resolution The design change in voltage per turn in a particular portion of the shaft rotation with unit voltage applied between the end terminals.

Theoretical Angular Resolution The design number of degrees shaft rotation per turn in a particular portion of the rotation.

Actual Angular Resolution The actual number of degrees shaft rotation per voltage step in a particular portion of the rotation.

Actual Voltage Resolution The actual change in voltage per angular step in a particular portion of the shaft rotation with unit voltage applied between the end terminals.

As yet the RETMA committee has not attempted to define a form of ohmic resolution, as did the Radiation Laboratory, since their definitions have been confined for the moment principally to potentiometers rather than precision rheostats. In the rheostat application, this is, of course, of decided interest and actual ohmic resolution might be appropriately defined as the actual change in resistance per angular step in a particular portion of the shaft rotation.

You will note that the RETMA committee defines both theoretical and actual forms of resolution, which implies that there may be considerable difference between the resolution designed into the potentiometer and the amplitude of or the angular separation between the resolution pulses appearing in the voltage output of the potentiometer. The magnitudes of these differences and a number of the controlling influences will be apparent as we proceed.

ANALYSIS OF RESOLUTION

For a clearer picture of what resolution actually is, let us examine the typical output curve of a linear wire-wound potentiometer. Consider the theoretical voltage output of a potentiometer of 1,000 turns of resistance wire, having a voltage E applied across the terminals and assume that the shaft is rotated at a constant speed of 1 RPM over a function angle of almost 360 degrees, Fig. 1(a). At time = 0, the wiper has a potential difference of 0 volts to ground, as seen in Fig. 1(b). At $t = 30$ seconds, the wiper potential has changed to 50 volts, and at $t = 60$ seconds, the wiper potential has reached 100 volts. Superficially it appears as though the varying potential from wiper to ground assumes the form of a continuous sawtooth wave, but closer examination of the wave form shows it to be made up of a discreet number of voltage steps, and very close examination shows these steps to be variable in both amplitude and duration but apparently following a definite pattern, as seen in Fig. 1(c). These steps are what is identified as resolution in the potentiometer.

For a further analysis of the characteristic wave form, let us consider a very simple wire-wound potentiometer having a total of only ten turns of resistance wire, as illustrated in Fig. 2. Assume that the wires are equally spaced and that each turn of wire has a resistance of one ohm and that the sliding contact, C, alternately shorts two turns together for 25% of the time as it moves

uniformly from terminal A to terminal B and remains on a single turn of wire for the other 75% of the time. You will note that in the initial position the contact C is shown shorting the terminal position to turn 1, and in position 2 it is shown only contacting turn #4. With a potential of 10 volts across the winding, the potential drop across each turn is, of course, one volt. If we were to place a vacuum tube voltmeter between terminal A and the wiper C and consecutively record the output voltage for each of the possible positions as the wiper moves uniformly from terminal A to terminal B, we would obtain a total of 19 different readings and 18 increments of change starting with 0 volts and ending with 10 volts. The amplitude of each voltage variation recorded with the VTVM would be as shown in table I. At a position on the terminating point A, the output voltage is, of course, 0 and as the wiper moves towards terminal B, in its first position, turn #1 is shorted to the terminal and the output voltage remains at 0. As the short is broken and the wiper moves to a position only on turn #1, a distinct step increase in the output is observed. Since there are ten equal resistors in series across the 10 volts input, the potential drop at this point is one volt. As the wiper moves into the position where turns 1 and 2 are shorted together, one ohm of resistance is shorted out leaving only 9 ohms across the 10 volts, and the potential drop is now $2/9$ ths of 10 volts, or 1.1111 volts. As the wiper moves on to turn #2 in the unshorted position, the voltage is again distributed across ten turns and the output voltage of the potentiometer increases to 2.0000 volts, an increase of .8889 volts, etc.

If we were to plot the output voltage as the wiper is moved uniformly from one end of the winding to the other, it would follow a theoretical curve as illustrated in Fig. 3. Note that as we progress from turn 1 to turn 10, the amplitude of each voltage step varies depending upon the position of the wiper on the winding. This can perhaps more clearly be seen in the pictorial representation shown in Fig. 4. Note that the amplitude of the longer duration pulses resulting from unshorting of turns gradually decreases as the amplitude of the shorter duration pulses resulting from the shorting position of the wiper gradually increases, and that as the center of the winding is approached, the amplitude of the pulses becomes nearly equal.

To differentiate between the two, the pulse which occurs as two adjacent turns become shorted has been identified as the minor resolution pulse, and the pulse which occurs as the turns unshort, the major resolution pulse. From table I and Fig. 3 it can be seen that the sum of amplitudes of any major pulse and the minor pulse it follows is always theoretically equal to the voltage drop between turns, or normal resolution. The number of normal resolution pulses is of course equal to n , the number of active turns of the winding, and in linear units its theoretical voltage amplitude, Δ , may be determined from $\Delta = V/n$ where V is the total voltage drop across the active portion

of the winding. For a 1000 active turn linear potentiometer with 100 volts applied across the active portion of the winding the normal resolution, for example, would be 0.1000 volts or as usually expressed, 0.10%. Whereas the total number of major resolution pulses resulting from the contact moving to the unshorted position is also theoretically equal to n , the total number of minor resolution pulses is equal to $n-2$, since as the first two and last two active turns become shorted, no change in the output voltage occurs. The total number of theoretical resolution pulses therefore is equal to $2n-2$, and for our 1000 turn potentiometer, 1998 increments, or nearly twice the number of active turns, will be theoretically observed in the output voltage.

To determine the amplitude of any minor resolution pulse, Δ_{mi} , as we short turn X to turn $X+1$ in a linear potentiometer the following may be used:

$$(1) \quad \Delta_{mi} = VX \left(\frac{1}{n-1} - \frac{1}{n} \right)$$

The amplitude of any major resolution pulse, Δ_{ma} , as we go from the shorted condition turn X to turn $X+1$ to the condition where a single turn of wire $X+1$ is contacted, can be determined for a linear unit from the following:

$$(2) \quad \Delta_{ma} = \Delta - \Delta_{mi}$$

or:

$$(3) \quad \Delta_{ma} = \frac{V}{n} - VX \left(\frac{1}{n-1} - \frac{1}{n} \right)$$

As noted previously, resolution is also oftentimes expressed in terms of angular degrees. The theoretical normal angular resolution of a linear potentiometer can be determined by dividing the active electrical angle by the total active number of turns. The increase in angular resolution resulting from the shorting effect between turns, Fig. 3, unlike the increase in voltage resolution is theoretically a constant, but to compute this increase involves detailed knowledge of the resistance wire diameter and spacing as well as contact surface configuration, discussion of which is beyond the scope of this paper.

At Fairchild, a dynamic study of the resolution steps in the output of a great number of linear and non-linear potentiometers has shown that the increments due to the shorting effect between turns have a relatively much higher frequency than the pulses from normal resolution. This

suggests that in dynamic applications of potentiometers of high resolution, the shorting resolution may be of very little practical advantage. However, in potentiometers of inherent low resolution, such as sector potentiometers, very low resistance units, some special types of non-linears, and in high resolution applications involving static voltage balance or division, etc., consideration might be given by the systems designer to the added resolution due to the shorting effect as a means of improving system performance or accuracy. Several means are occasionally employed by the circuit engineer to "improve" the resolution of a potentiometer. One of these is most commonly used when the circuit values are such that a very low resistance potentiometer is required and their inherent lower resolution is not desired. A potentiometer of much higher resistance, with its attendant higher resolution is specified, and a fixed or variable shunt resistor is tied between its input terminals to reduce the resistance to the desired value. By using a variable resistor, the paralleled resistance of the two may be more exactly established and so long as little or no current is drawn through the potentiometer wiper, the accuracy of the output function will be unaltered.

SOME PRACTICAL CONSIDERATIONS OF RESOLUTION

From some of the previous discussion, it may be apparent that the theoretical design limits of resolution may be restricted by a number of other specified potentiometer characteristics. The total number of turns in the active portion of the potentiometer winding is oftentimes predetermined by:

1. The physical size of the potentiometer
2. The specified or desired electrical or function angle
3. The total resistance requirement of the potentiometer
4. The choice of winding mandrel material or shape, as predetermined by operating or environmental specifications
5. The choice of the alloy of the resistance wire, which in turn may be limited by other performance characteristics required such as long life, temperature coefficient of resistance, or noise limitations.

Linearity or conformity, as will be seen a little later, have a significant influence on as well as are greatly influenced by resolution. Generally it can be said that better resolution is inherent in the larger diameter potentiometers, longer electrical angles and the higher resistance values. The highest resolution theoretically might be obtained by the winding of the finest size of lowest resistivity wire on a mandrel of minimum cross sectional area, but other practical limitations as well as performance requirements seldom make it possible to achieve this ultimate. The lower resistivity wires, for example, are high in copper content and as a result are both difficult to draw and wind in the

fine sizes. They also have rather poor life expectancies, inferior corrosion resistance, high temperature coefficients and poor noise characteristics. To obtain the desired low ohms per turn, specially drawn and insulated copper mandrel wire which is both strong and tough, and in small diameters easily wound and formed, is used as a mandrel in the manufacture of the majority of the high resolution high accuracy potentiometers of today. Single turn potentiometers of 0.02% resolution (5,000 turns) and lower are not uncommon, and multi-turn potentiometers with resolution less than 0.004% (25,000 turns) for the ten-turn types are readily available. For even lower values multi-turn potentiometers of up to forty turns have been produced.

So far, resolution has been discussed principally in terms of the theoretical considerations and values. Referring back to Fig. 2, it is readily apparent that the shape and size of the sliding contact must be given very important consideration if high and uniform actual resolution in the potentiometer output voltage is to be achieved. Other considerations such as uniformity of wire spacing, uniformity of the length and cross sectional area of each turn, uniformity and smoothness on the contact surface of the turns, absence of shorted turns, etc. are all important factors contributing towards good "effective" or actual resolution. Generally, in the design of the contact shape every attempt is made to avoid shorting more than two turns of wire as the contact moves from one end of the winding to the other, and to preserve this condition for as great a number of operational cycles as possible by proper choice of contact material and deflection properties of its supporting spring. Fig. 5(a) illustrates the type of wiper designed by Radiation Laboratories group for their RL270 series of precision linear potentiometers.² The contact was cut from a length of Paliney #7⁵ wire and had the shape of a small cylinder or rod. This type of contact became identified as a rod type of contact and is still in widespread use in the same or similar form. Figs. 5(c) and 5(d) illustrate several current variations of the rod contact now in use. Note that they all have one common characteristic, the cylindrical shape in the area where the resistance wire is contacted. It is this shape in combination with a properly designed spring which results in the optimum resolution and linearity characteristic commensurate with the other performance requirements of the potentiometer. With an improperly designed wiper the contact wears very rapidly and a flat is developed on the bottom of the contact. The result is a greater tendency for the wiper contact to skip turns, short out a variable number of turns, "reach back" and re-touch turns already passed, and generally produce an erratic resolution effect in the potentiometer output. Fig. 5(b) illustrates a most recent Fairchild low deflection sensitivity wiper design developed to reduce the contact mass concentrated on the end of the contact assembly with a resultant increase in the natural frequency of the spring. It permits the low noise higher speed operation of the potenti-

ometer for many hundreds of thousands of cycles without any appreciable degradation in potentiometer resolution at reduced wiper pressure.

Perhaps the most important consideration in the design of the contact for optimum resolution characteristics is the choice of the contact diameter in the area where the resistance wires are contacted. If the contact diameter is too large with respect to the wire diameter and spacing of the turns, a flat develops very rapidly on the contact with resultant erratic resolution. If the contact diameter is too small, catching and tearing of the turns of wire will occur with a resultant rapid deterioration over a relatively short period of life. At Fairchild as many as twelve different basic contact variations of size and physical characteristics are standardized for a single type of linear potentiometer and "contact matching" is employed to the fullest extent to preserve the optimum resolution, linearity and noise characteristics of the potentiometer throughout life. Actual voltage resolution to some extent varies with several other factors such as wiper current, noise in combination with wiper current, etc. The tendency to skip turns and for the contact to resonate and produce erratic resolution, for example, increases as the speed of the potentiometer shaft is increased. For extremely high speed potentiometers, i.e. 300 to 1800 RPM, specially shaped and lightened contacts are employed to minimize these tendencies.⁶

RELATIONSHIP OF RESOLUTION AND LINEARITY OR CONFORMITY

The potentiometer designer must give extremely important consideration to resolution with respect to the linearity or conformity requirement of the potentiometer. Since resolution produces a stepped effect in the output of the potentiometer, it can be seen that it automatically imposes a limitation upon the theoretical accuracy of the potentiometer function. Referring to Fig. 3, it can be seen that the normal resolution jumps are one volt and since linearity tolerance is expressed as a plus and minus limitation, i.e. $\pm 5\%$, or a 1% spread, it follows that the best possible linearity obtainable from the potentiometer illustrated would be $\pm 1/2$ volt or $\pm 5\%$. For a potentiometer with 1000 active turns of resistance wire (resolution 0.10%), the best possible linearity obtainable would be $\pm 0.05\%$, or a total tolerance spread equal to the resolution. Because of the many other factors contributing to error in the average precision potentiometer such as mechanical eccentricity, variations in mandrel dimension, lack of uniformity in the resistance wire, etc. a much greater number of turns must be designed into the winding to achieve a desired linearity. Several "rules of thumb" have been evolved, as follows — For potentiometers of the single turn type employing mandrels of circular cross sectional area, at least twice as many turns of wire must be employed to meet a given linearity tolerance, i.e. a potentiometer having a linearity requirement of $\pm 1.0\%$ must have a

resolution of at least .05% or a minimum of 2000 turns of resistance wire. For mandrels of other than circular cross sectional area, such as the phenolic card type of winding, this factor is more nearly three to one, and in the multi-turn potentiometers approaches four to five to one. In addition, it can be generally stated that with the very fine sizes of resistance wire (.0007 to .0015) these factors must be increased slightly, and with the very heavy resistance wires (.0030 to .0060) they may be reduced slightly. It is assumed, of course, in applying these rules proper tolerance consideration has been given particularly to the factors controlling eccentricity and that a correct selection of contact diameter ratio to resistance wire diameter and spacing has been made. Fig. 6, for example, illustrates the poor linearity obtained from a winding of approximately 2500 turns (.04% resolution) in which a contact of improper diameter has been utilized to contact the winding. The erratic behavior of the contact and its effect on the resolution may be seen by the variable band width of the pattern and the linearity is $\pm 10\%$. Fig. 7 shows the linearity pattern resulting from the same potentiometer in which only the contact has been changed to the proper diametric ratio with respect to winding wire size and spacing. Note the uniform resolution band width and appreciable improvement in linearity to $\pm 0.06\%$, a gain equal to the resolution.

MEASUREMENT OF RESOLUTION

While most potentiometer resolution requirements are generally stated or designed in terms of the theoretical resolution, others require that the actual resolution shall not exceed certain specified voltage, percentage or angular limits, entailing a measurement of these values. As seen in Figs. 6 and 7, a fair idea of the quality and average amplitude of the voltage resolution can be obtained by observation of the varying band width of the linearity pattern. By employing an accurate protractor, a bridge circuit and observing the angular and voltage increments from various nulled positions over the potentiometer winding, a more accurate measurement of conformance to the specification can be obtained. This method is both arduous and time consuming, but it is currently used by many inspection departments.

Fig. 8 illustrates a method successfully employed at Fairchild to obtain qualitative and to some extent quantitative resolution data on both linear and non-linear types of precision potentiometers. The voltage appearing at the input to the condenser will consist of the characteristic sawtooth wave (for a linear potentiometer) upon which is superimposed the voltage resolution steps. The wave form appearing at the input of the D.C. amplifier depends, however, on the action of the RC network. This RC network will act as a differentiator if its time constant is very much smaller than the period of the applied wave. With a potentiometer of 1000 turns spread linearly over approximately 360° and being turned at 1 RPM, the sawtooth wave has a period of 60 seconds. If the RC network is adjusted to have a time con-

stant of for example 1 second, the sawtooth wave will be completely differentiated while the resolution pulses, having a period of from 0.06 seconds or less will come through with practically no change in wave shape. As a result, the recorder will show the turn to turn changes arising from the resolution of the potentiometer. The use of a pen recorder is particularly advantageous since a permanent record of the resolution trace is obtained for reference or study purposes, and a voltage calibration marker can be introduced on the tape prior to recording. Care must be taken not to exceed the high frequency response characteristics of the recorder. For the measurement of resolution at higher rotational speeds, an oscilloscope may be substituted and if desired, a permanent record made with an oscilloscope camera.

Fig. 9 illustrates a typical section of a resolution trace of a linear potentiometer, and Fig. 10 that of a non-linear unit. Note the pips due to the shorting effect between adjacent turns and uniform turn spacing in the linear unit. Fig. 11 shows the characteristic curve obtained with a poorly designed or excessively worn contact. Note the very erratic amplitude of the resolution pulses and the apparent wide spacing between turns as a result of the contact skipping turns.

In most well designed linear potentiometers no particular problems will be encountered in obtaining smooth and uniform resolution characteristics within the usual limits specified and these characteristics can be retained within reasonable tolerance throughout many hundreds of thousands of cycles of operating life. With some non-linear potentiometers, particularly those of high slope ratio, obtaining the specified actual resolution or the theoretical resolution required to obtain the desired functional accuracy (the latter rather than the former is almost invariably specified), some design problems are occasionally encountered. For the majority of windings a much greater number of turns than needed are designed into the winding and the checking of actual potentiometer resolution becomes more a matter of assuring the quality of the unit by observing the resolution trace for absence of turns skipped by the contact, shorted turns, and wide spaces between turns. Each of these defects have a characteristic appearance on the resolution trace, as illustrated in Fig. 12. Widely spaced turns show up as an abnormal separation between resolution pulses, as seen in Fig. 12(a). Shorted turns show up in the same manner, and while it is difficult to differentiate between a wide space and a single pair of shorted turns, shorts generally occur in groups, and their appearance is usually as indicated in Fig. 12(b). Skipped turns are readily identified by the characteristic double or triple amplitude pulse as shown in Fig. 13(a). As a rule, these large pulses are preceded by a wide space. Fig. 13(b) illustrates the appearance of a low frequency noise pulse and the resolution pulse resulting from the contact "reaching back" to touch a turn it had already passed over. In

most cases the noise pulse will cause the pen to deflect off scale whereas the amplitude of the pulse caused by the contact reaching back will be proportional to the number of turns skipped. Both types of pulses will oftentimes be followed by erratic amplitude pulses and spaces, as seen in Fig. 13(b).

Figures for typical maximum resolutions found in present wire-wound potentiometers were cited earlier, and they generally can be supplied in large volume to meet the average current systems requirements. With the present emphasis on component miniaturization and the need for continued high accuracy, the physical limits are becoming such that high resolution potentiometers can only be obtained in the higher resistance values. Manufacturers of miniature sector or segment type potentiometer units in some instances have found it necessary to wind wire as fine as 0.00040 inches in diameter in order to meet existing specifications. To meet the increasing need for miniature potentiometers of high resolution as well as higher operating temperatures, Fairchild recently announced the development of a 3/4 inch diameter Filmpot having an infinite resolution characteristic and capable of operating in ambient temperatures exceeding 450°F. The "infinite" or "zero" resolution feature of the unit stems from the use of a resistance element consisting of a high temperature non-conductive inorganic disk upon which has been vacuum deposited a thin film of highly non-corrosive metals of the desired resistance. The output of the Filmpot is a continuous stepless curve whose functional accuracy is almost entirely dependent upon other mechanical factors in the potentiometer. While they are presently unavailable in large volume production, they show promise of becoming an important factor in contributing to the improved resolution and accuracy of tomorrow's electronic equipment.

ACKNOWLEDGEMENTS

The author wishes to acknowledge his indebtedness to all the members of the Fairchild Wire Wound Lab who assisted in gathering data referenced for portions of this paper. Particular acknowledgement is made to engineers Stanley Friedman and Aaron Blaustein for their work in deriving and proving out satisfactory circuits and mechanical details involved in the continuous measurement of resolution.

Acknowledgement is also made to the Air Force Cambridge Research Laboratories, Cambridge, Mass., for permission to use the photographs shown in Figs. 6 and 7.

CITED REFERENCES

1. J. R. Altieri, "Causes and Measurement of Residual Potentiometer Noise", Instruments, page 1712, Volume 26, Nov. 1953
2. Robert J. Sullivan, "The RL270 Series of Precision Potentiometers", Report 864, M.I.T. Radiation Laboratories, Cambridge, Mass., Mar. 25, 1946
3. J. F. Blackburn, "Components Handbook", M.I.T. Radiation Laboratory Series, Volume 17, Chapter 8, McGraw-Hill Book Co., New York (1949)
4. Fred E. Dole and Robert J. Sullivan, "Life Tests of Contact Material on Standard Linear Wire-Wound Potentiometers", Report 617, M.I.T. Radiation Laboratories, Cambridge, Mass., Mar. 12, 1946
5. J. M. Ney Co., Hartford, Conn., Bulletin R-12
6. "Electro-Mechanical Report to the Steering Committee" #2, 1946, Air Force Cambridge Research Center, Cambridge, Mass.

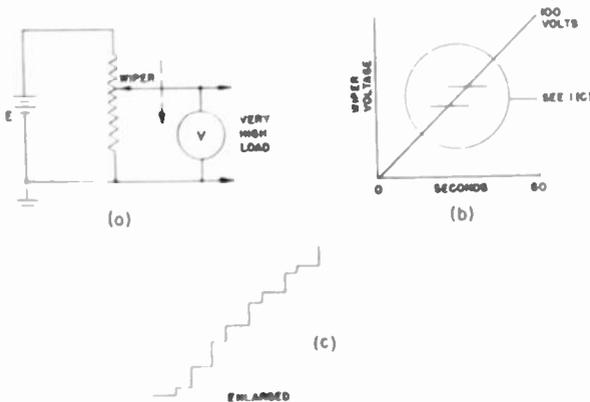


Fig. 1

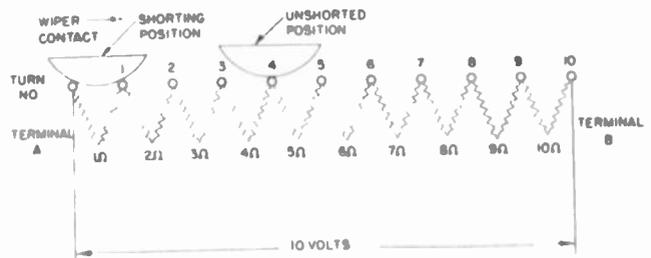


Fig. 2

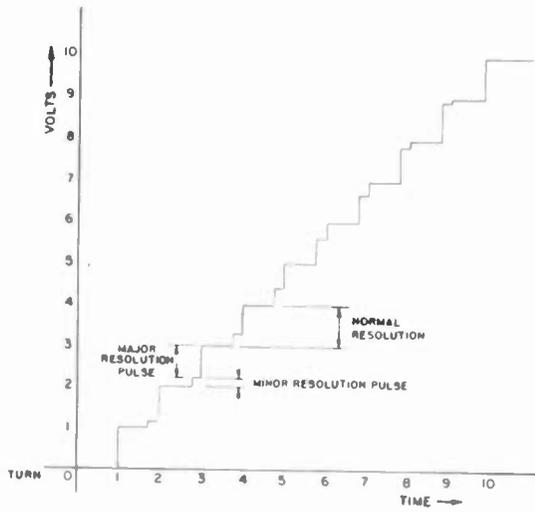


Fig. 3

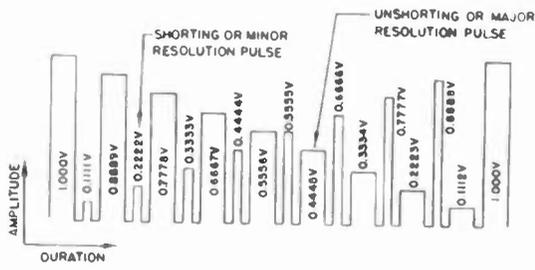


Fig. 4

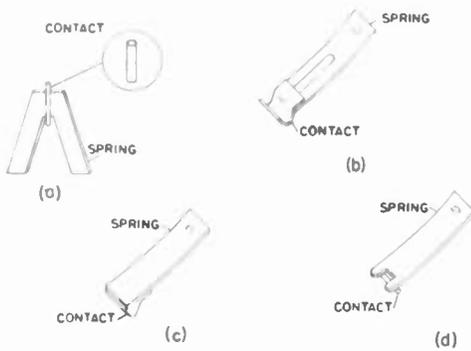


Fig. 5

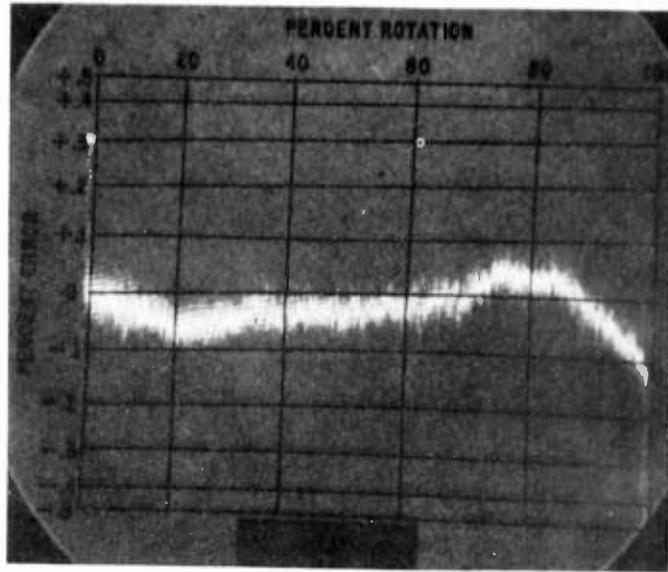


Fig. 6

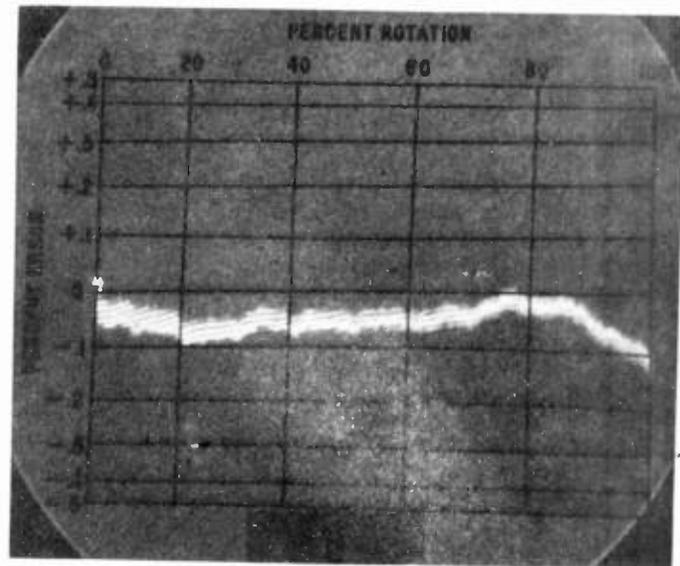
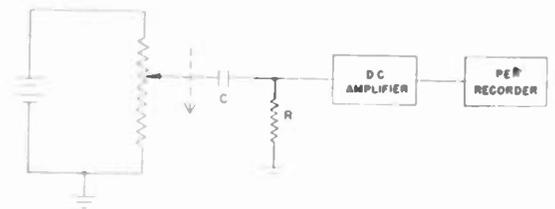


Fig. 7



RESOLUTION CHECKER

Fig. 8

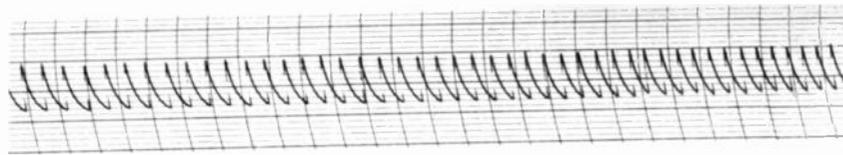


Fig. 9

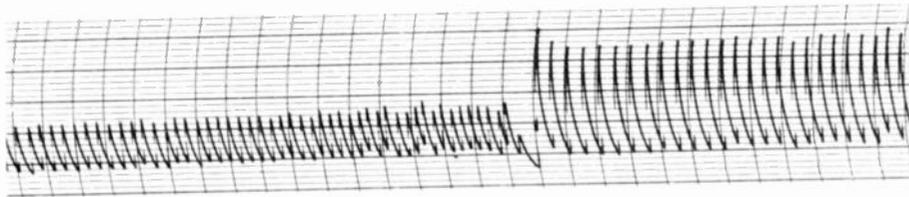


Fig. 10

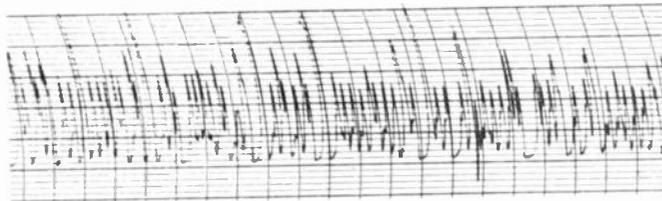
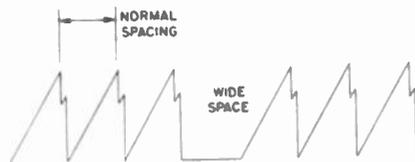


Fig. 11



(a)



(b)

Fig. 12

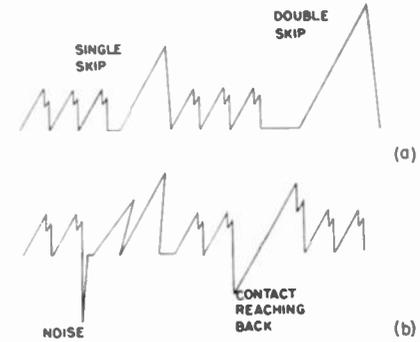


Fig. 13

VOLTAGE PULSE NO.	OHMIC PROPORTION	WIPER VOLTAGE	RESOLUTION CHANGE
Start	0/10	0.0000	----- 1.0000V
#1	1/10	1.0000	----- 0.1111V
#2	1/9	1.1111	----- 0.8889V
#3	2/10	2.0000	----- 0.2222V
#4	2/9	2.2222	----- 0.7778V
#5	3/10	3.0000	----- 0.3333V
#6	3/9	3.3333	----- 0.6667V
#7	4/10	4.0000	----- 0.4444V
#8	4/9	4.4444	----- 0.5556V
#9	5/10	5.0000	----- 0.5555V
#10	5/9	5.5555	----- 0.4445V
#11	6/10	6.0000	----- 0.6666V
#12	6/9	6.6666	----- 0.3334V
#13	7/10	7.0000	----- 0.7777V
#14	7/9	7.7777	----- 0.2223V
#15	8/10	8.0000	----- 0.8888V
#16	8/9	8.8888	----- 0.1112V
#17	9/10	9.0000	----- 1.0000V
#18	10/10	10.0000	----- 1.0000V

TABLE I

EVALUATION OF CORE MATERIALS FOR MAGNETIC AMPLIFIER APPLICATIONS

R. D. Teasdale

H. R. Brownell

Magnetic Metals Company
Camden, New Jersey

Introduction

In order to produce useful and economical magnetic amplifiers, the engineer must tailor his design to the characteristics of commercially available ferromagnetic materials. Unfortunately, at the present time, there is little standardization as to what information is most important and as to how the pertinent data can best be presented to aid the designer in selecting suitable core materials. Further, the data which is available is restricted almost entirely to wrapped cores and neglects the characteristics of other shapes, such as stamped rings, DU laminations and EI laminations which are attractive economically and which are quite adequate for many applications.

It is the object of this paper to present data on commercially available alloys and shapes in the form of useful normalized control characteristic curves and tables. These data are obtained from 400 cycle measurements on a half-sine wave self-saturating magnetic amplifier circuit employing the so-called constant current flux reset type of measurement. These curves and tables are used:

1. to compare various magnetic alloys (79% nickel, 50% nickel, 3% silicon-iron) with and without strong grain orientation,
2. to compare the performance of these magnetic materials in various shapes, such as wrapped cores, stamped rings, DU laminations, and EI laminations,
3. to compare the performance of these magnetic materials in various thicknesses, from 2 mils through 14 mils,
4. to determine the effect of a suggested modification in the dimensions of the DU lamination for silicon-iron alloys.

It is felt that designers can use these curves to make reasonable estimates of performance for a wide range of specific applications.

Measurement Method, Materials Tested

The criteria for a suitable method of

measurement are numerous. The method selected should give information which is useful both to those who design magnetic amplifiers and to those who produce magnetic alloys. It should give results in a form which permits calculation or estimation of the response of magnetic amplifiers using the material being tested. The test circuits should be reproducible and inexpensive enough to permit widespread use of the method. The method chosen should provide for periodic checks on accuracy and should be simple and fast enough to be useful for production testing.

Further, it is axiomatic that the best analogue of a device is the device itself. Failing exact correspondence, the test circuit should be as close to a typical operating magnetic amplifier circuit as time and economy permits. In addition, the results, insofar as possible, should depend only upon the core material of the sample being tested and not upon the associated circuitry.

The test method which was selected as best satisfying these criteria follows quite closely a method first described by Conrath¹ and discussed in considerable detail by Roberts in an excellent conference paper². As illustrated in Figure 1, in this method the core is tested in a simple self-saturating magnetic amplifier circuit which employs an excitation current of half sine waves in conjunction with a variable DC supply for control. Thus, the excitation is gated on and off at the repetition rate of the test frequency here 400 cps. This method has been referred to as a constant current flux-reset method, in the sense that the DC control current strives to reestablish a flux level after each excitation pulse. For a fixed excitation amplitude A , high enough to insure that the particular core is driven well into saturation, values of the average rectified output voltage are recorded as the DC control current is varied through zero to a value large enough to produce saturation. The recorded values are normalized and plotted to yield a curve similar to that shown in Figure 2.

Neither the test method nor the type of presentation used here has been adopted definitely as a standard by the industry; however, this test method involving a self-saturating magnetic

amplifier circuit, or a variation of it, is now in use by so many companies who work with magnetic amplifiers, that it has been selected as the most useful for obtaining test data on magnetic core materials.

Tests were made on the following magnetic materials:

- Carpenter Hymu 80 alloy, non-oriented
(Molybdenum Permalloy)
(4% Mo, 79% Ni, 17% Fe)
- Carpenter 49 alloy, non-oriented
(49% Ni, 51% Fe)
- Orthonic alloy, rectangular loop oriented
(50% Ni, 50% Fe)
- Microsil oriented silicon-iron alloy
(3% Si, 97% Fe)
- Crystalligned oriented silicon-iron alloy
(3% Si, 97% Fe)

For these different alloys, tests were made on the following shapes in various thicknesses:

- Wrapped Cores 1" x 1.375" x 0.25 toroid,
path length 3.72", area
.0469 in²
- Punched Rings 1.125" x 1.5" x 0.1875"
toroid, path length 4.15",
area .0338 in²
- 1DU Laminations rectangular shape shown on
Fig. 6. Outside dimensions
0.875" x 2.5", path length
5.25", area .045 in²
- 37DU Laminations rectangular shape shown on
Fig. 8. Outside dimensions
1.5" x 3.0", path length
6.75", area .141 in²
- 37DU Modified
Laminations rectangular shape shown on
Fig. 8. Outside dimensions
1.25" x 3.0", path length
6.5", area .0625 in²
- 75EI Laminations rectangular shape shown on
Fig. 6. Outside dimensions
1.875" x 2.25", path length
4.5", area .090 in²

These area values are nominal. Before actual measurements, each sample used was weighed; the actual area of each sample was computed from the weight and recorded for use in normalizing.

Figure 2 shows a general layout of the test equipment. Three separate windings were used, oriented to minimize coupling. A voltage compensation method was used in calibrating and checking the amplifier and output equipment, i.e., what is actually plotted is that (normalized) average value of voltage into the amplifier which produces the same output reading as was recorded with the peaked waveform from the core.

Because of the high peaks in the waveform of induced voltage from the core, the output circuit requires some care. It is necessary to use a wide band audio amplifier and to guard against peak clipping by observing patterns on an oscilloscope. The amplified voltage is fed through a bridge rectifier into a load resistance, and the output voltage read with a DC millivoltmeter.

Method of Presenting Data

For convenience in design, the test results are presented in the form of normalized control characteristics as illustrated in Figure 2. The average value of the output voltage per turn per square inch of core cross-section is plotted against the DC control ampere-turns per inch to yield a curve which is characteristic of the core material in this particular circuit for the particular amplitude of excitation employed.

For purposes of numerical tabulation, quality control, and production testing, it is convenient to associate numerical values with each curve. With reference to Figure 2, these normalized values are:

- V_0 , the zero intercept in millivolts per turn per square inch
- S , the maximum slope of the linear region
in volts per turn per square inch
ampere turns per inch
- V_s , the saturation value of average output voltage corresponding to a specified H_s in millivolts per turn per square inch
- H_A, V_A coordinates of point A, the lower limit of the linear range
- H_B, V_B coordinates of point B, the upper limit of the linear range
- $D = 100 \frac{V_B - V_A}{V_s}$, the length of the linear range as a percentage of the saturation value

$V_C = \frac{V_A + V_B}{2}$, the output voltage value corresponding to the center of the linear range

$H_C = \frac{H_A + H_B}{2}$, the coercive force at 400 cps in ampere turns per inch

It is evident that the knowledge of the six numbers V_0 , S , V_S , D , H_C and V_C (in conjunction with the specified values of H_S and of the permissible deviation from linearity) is sufficient information to plot the characteristic curve. Tables 2, 3, 4 and 5 give these six normalized values for high-nickel, low-nickel and silicon-iron alloys. The values listed in these tables would permit one to draw the corresponding core characteristic curves, if desired, but this is not necessary, as the tabulated values give sufficient information to permit design of a large class of magnetic amplifiers directly.

For production testing, it is possible as a result of experience to specify two voltage levels V_1 and V_2 , equally spaced about V_C , which are sure to fall on the linear region for the material and shape being tested. It is then not necessary to draw the complete curve but merely to record values:

V_0 , the zero intercept
 H_1 , the DC control level corresponding to a specified V_1
 H_2 , the DC control level corresponding to a specified V_2
 V_S , the saturation value;

then two additional values can be computed

$S = \frac{V_2 - V_1}{H_2 - H_1}$, the slope of the linear region

$H_C = \frac{H_1 + H_2}{2}$, the coercive force at 400 cps.

If the testing is confined to toroidal cores it is possible to normalize the results still further by multiplying the slope S by the ratio of the outer to the inner diameter of the toroid.

Because these numerical values depend upon the excitation level A , and upon H_S , V_1 and V_2 , the selection of proper values for these parameters is important. Curves such as those shown in Figure 3 were plotted to determine suitable values of excitation level A . The other three

parameters were selected as a result of test experience with many cores. The final values selected are given in Table I.

A certain degree of arbitrariness is inevitable in determining points A and B , the limits of the linear range. Here points A and B are selected as those points at which the characteristic curve deviates by a specified amount from the straight line through the linear range, as shown in Figure 2. The deviation is specified as 0.02 AT/in for the high-nickel and low-nickel alloys and as 0.1 AT/in for the silicon-iron alloy. In every case choices were made to yield conservative results for the numerical data.

Interpretation of Curves

From Figure 2, it is seen that the characteristic curve is similar to the right hand side of a dynamic hysteresis loop. For data on magnetic materials:

V_0 corresponds to $B_m - B_r$. Low values of V_0 correspond to rectangular loop materials.
 (O, V_C) are the coordinates which correspond to the origin of the dynamic hysteresis loop
 V_S corresponds to $2 B_m$
 S corresponds to the slope of the linear portion of the dynamic hysteresis loop, or to the maximum value of μ_d .
 D corresponds to the limits of the linear range
 H_C corresponds to the coercive force of the material at the test frequency.

It is also possible to relate the design of a large class of magnetic amplifiers to the information given by the characteristic curves and the tables. Roberts² has obtained close agreement between calculated and measured curves for a half-wave magnetic amplifier whose load impedance is high with respect to the inductive effect of the gate winding when the core is saturated. With reference to magnetic amplifiers, it is evident that V_0 corresponds to the level of output voltage at zero control current and V_S corresponds to the maximum output voltage or the power handling capacity of the magnetic amplifier. S corresponds to the gain of the magnetic amplifier using this core material. The greater the slope the higher the gain. For magnetic amplifiers

with high impedance control windings the slope should be directly proportional to the gain.

Comparison of Magnetic Alloys

Detailed comparisons of different alloys are not justified, because of differences in processing and in temperatures of anneal. However, several general comments can be made.

Some major differences among the various magnetic alloys are shown in the curves of Figure 3. Thus the Hymu 80 alloy saturates at low values of DC control (low H_C), with a high gain but a low value of saturation voltage. The Orthonic alloy saturates at intermediate values of DC control with a lower gain but at a higher saturation voltage. The Crystalligned silicon-iron alloy saturates at high values of DC control with a high value of saturation voltage, but the gain is quite low.

Reference to tables II, III and V substantiates these conclusions. One can use Hymu 80 and get increased sensitivity (gain) at the expense of power handling capacity. Or one can use silicon-iron alloy and sacrifice gain to obtain high power handling capacity. (Note that the modified 37DU lamination of silicon-iron is particularly favorable in power handling capacity.) Orthonic combines the better features of both alloys at the expense of higher H_C and lower gain than the Hymu 80 alloy.

The Hymu 80 and Carpenter 49 alloys give the lowest values of H_C and also the highest gains in the 2 mil thickness. The Orthonic alloy gives the longest length of linear range D and, in the wrapped cores, the lowest value of V_0 . The silicon-iron alloy exhibits the highest saturation voltage V_S , particularly for the modified 37DU laminations.

Tables III and IV permit a comparison of the effect of orientation on 50% nickel alloys. Orthonic alloy (Table III) shows strong grain orientation while Carpenter 49 alloy (Table IV) does not. The principal effect of the orientation seems to be in higher value of saturation voltage and in the higher gains for the Carpenter 49 alloy. The other differences in the tabulated values are largely explained by the lower temperatures used in annealing the Orthonic material.

Comparison of Wrapped Cores, Stamped Rings, and Laminations

Figures 4 through 7 illustrate the effect of

various shapes upon the normalized characteristic curves of different alloys, and tables II through V give more extensive data. [It should be emphasized that all these curves represent results on material now available in commercial production and are not values obtainable only from materials specially selected in the laboratory. As such, these curves do not, of course, represent the ultimate results obtainable.] To permit valid comparisons, all the samples listed in each table and represented on each curve sheet were given the same type anneal at the same top temperature.

In general, the curves and tables show that the wrapped cores have the highest gain S , the highest saturation voltage V_S , the lowest intercept V_0 , and, for Orthonic only, the longest linear range D. The rings have the longest linear range D, (except for Orthonic), the lowest H_C , and gains almost as high as the wrapped cores.

For the DU laminations, the corresponding values are higher than was anticipated -- an indication that such laminations are feasible wherever high sensitivity is not a primary requirement. The DU laminations permit easy adjustment of stack height and are less expensive. Figure 6 shows that the DU laminations are superior to the EI laminations; further, the variations after restacking are much greater for the EI laminations than for the DU laminations.

Effect of Material Thickness

It is evident that as magnetic material becomes thinner, it becomes more subject to surface contamination and to mechanical distortion. For thicknesses below 4 mils, in particular, it is almost impossible to isolate the mechanical and magnetic properties. This fact tends to limit the use of laminations, which must be stacked, to thicknesses of 2 mils and greater. Conversely, when it is necessary to use material thinner than 2 mils, as for high frequency applications, the use of wrapped cores is indicated.

As a matter of practice, the optimum anneals for the same alloy in different thicknesses are somewhat different.

The curves and tables indicate several important trends in the design variables with changes in thickness. In general, as the material becomes thinner, the gain S increases, the saturation voltage V_S increases, and H_C decreases -- all desirable trends, counter-

balanced in part by the increased cost and the mechanical weakness of thinner materials. (The few exceptions to this general trend are for the 2 mil laminations, and, therefore, are probably due to mechanical distortion.)

For thinner samples of the Hymu 80 and Orthonic alloys, the length of linear range D does increase slightly, but no corresponding general increase is evident for the Carpenter 49 and silicon-iron alloys.

Effect of Modifying DU Dimensions

It has been suggested that a more desirable flux distribution and better saturation characteristics can be achieved by modifying the dimensions of the DU lamination. To test this suggestion, several samples of silicon-iron alloy were modified as shown in Figure 8 to change the leg cross-sectional area considerably while reducing the path length only slightly.

Many samples were prepared and tested. Typical results, as given in Table V and plotted in Figure 8, show that this modification does indeed produce superior performance for silicon-iron alloy. For example, as compared to the regular DU lamination, Table V shows that the modified DU laminations exhibit higher gain S, higher saturation voltage V_s , and a longer linear range D, the only disadvantage being an increase in H_c . These conclusions apply for both thicknesses investigated.

Conclusions

1. The constant-current flux reset test method is described and used to obtain data for comparison of commercially available magnetic core materials in various shapes and thicknesses.
2. Test results are normalized for greater usefulness in the design of magnetic amplifiers, and typical characteristic control curves are plotted.
3. For purposes of tabulation, quality control, and production testing, the essential features of the characteristic curves are compressed into six normalized numbers: V_0 , S, V_s , D, H_c , and V_c , as illustrated in Figure 2. These basic numbers are tabulated for Hymu 80, Orthonic, Carpenter 49, and silicon-iron alloys.
4. Three different alloys are compared in Figure 3. The Hymu 80 and Carpenter 49 alloys have the lowest H_c values and, in the 2 mil thickness, also exhibit the highest

gains S. Orthonic alloy gives the longest length of linear range D, and, in wrapped cores, the lowest values of V_0 . The silicon-iron alloy has the highest saturation voltage V_s (particularly for the modified DU laminations) but the lowest value of gain. In general, it is necessary to sacrifice power handling capacity to get increased sensitivity, and conversely.

5. In selecting shapes, the following general statements apply: All else being equal, wrapped cores have the highest gain S, the highest saturation voltage V_s , the lowest intercept V_0 , and, for Orthonic only, the longest linear range D. Punched rings have the longest linear range D (except for Orthonic), the lowest H_c and gains almost as high as the wrapped cores. Values for the DU laminations, while lower, are competitive for many applications.
6. In general, as a magnetic material becomes thinner, the gain S increases, the saturation voltage V_s increases, and H_c decreases; however, cost increases, and surface contamination and mechanical distortion become serious problems.
7. It is possible to obtain superior performance from regular 37DU laminations of silicon-iron alloy by modifying the dimensions as shown in Figure 8. The modified DU laminations have higher gain S, higher saturation voltage V_s , and a longer linear range D, at the expense of an increase in H_c . Comparison curves are plotted in Figure 8, and numerical values are given in Table V for both thicknesses investigated.

Acknowledgments

This paper has benefited greatly from the helpful suggestions of H. F. Porter, of Magnetic Metals Company, W. C. Johnson of Princeton University, R. W. Roberts of the Westinghouse Electric Corporation, and C. A. Belsterling of the Franklin Institute. Much of the measurement work was done by B. R. Huddell and D. W. Stanton.

References

1. J. R. Conrath - "Magnetic Amplifier Gapless-Core Tests", Electronics, November 1952, McGraw Hill, New York, N. Y., p. 119.
2. R. W. Roberts - "Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers", AIEE Conference Paper, New York, N. Y. January, 1954.
3. Huhta, H. - "Flux Resetting Characteristics of Several Magnetic Materials", AIEE Technical Paper 54-161, New York, N. Y. January, 1954.

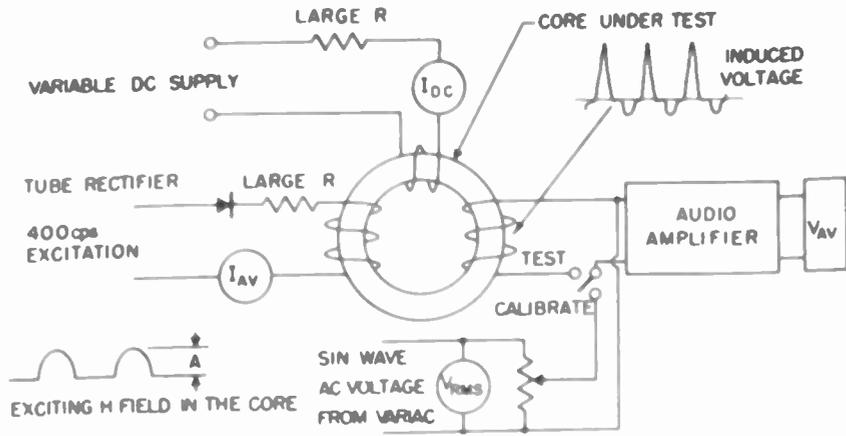


Fig. 1 - Test method and wave forms.

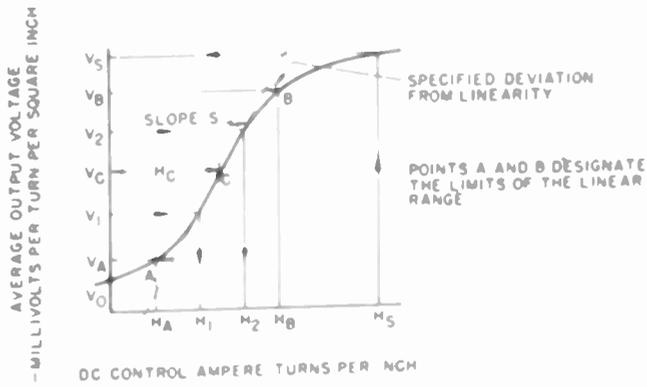


Fig. 2 - Typical normalized characteristic curve.

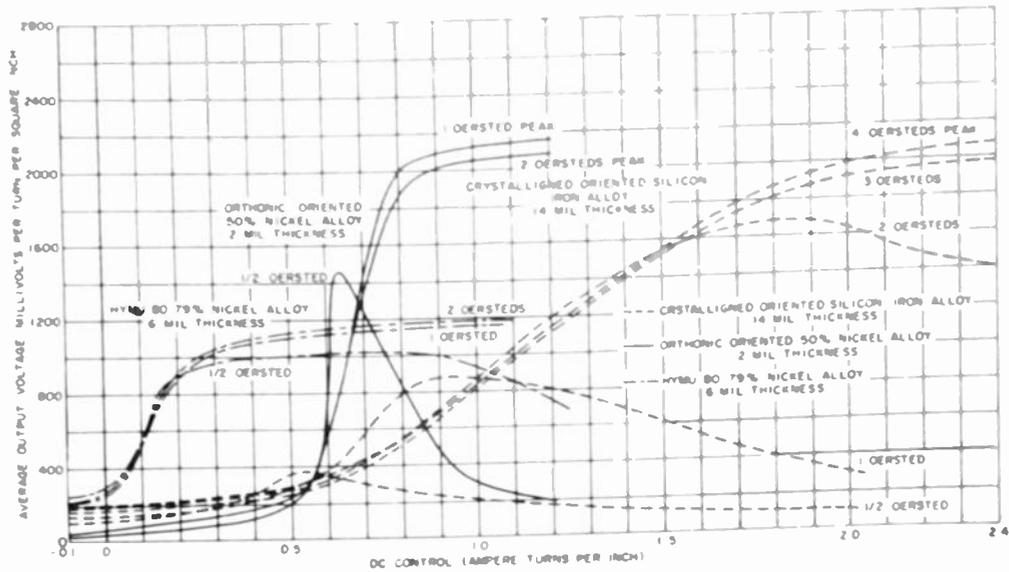


Fig. 3 - Effect of peak excitation upon characteristic curves; wrapped cores 400 cps.

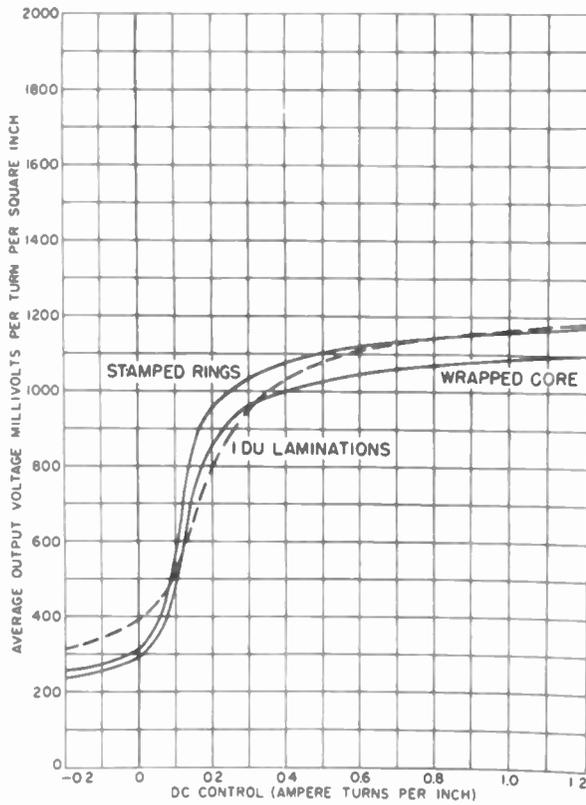


Fig. 4
Carpenter Hymu 80 79% nickel alloy in various shapes; 6 mil thickness 400 cps.

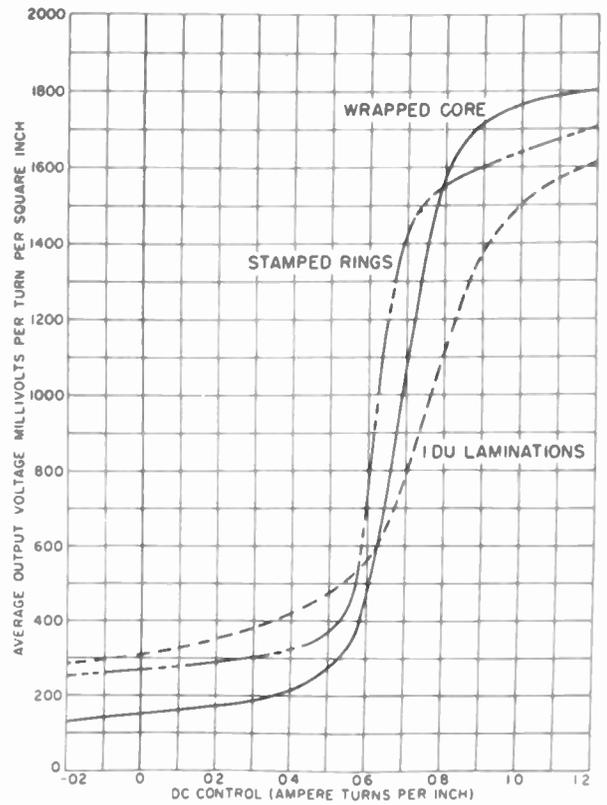


Fig. 5
Orthonic oriented 50% nickel alloy in various shapes; 4 mil thickness 400 cps.

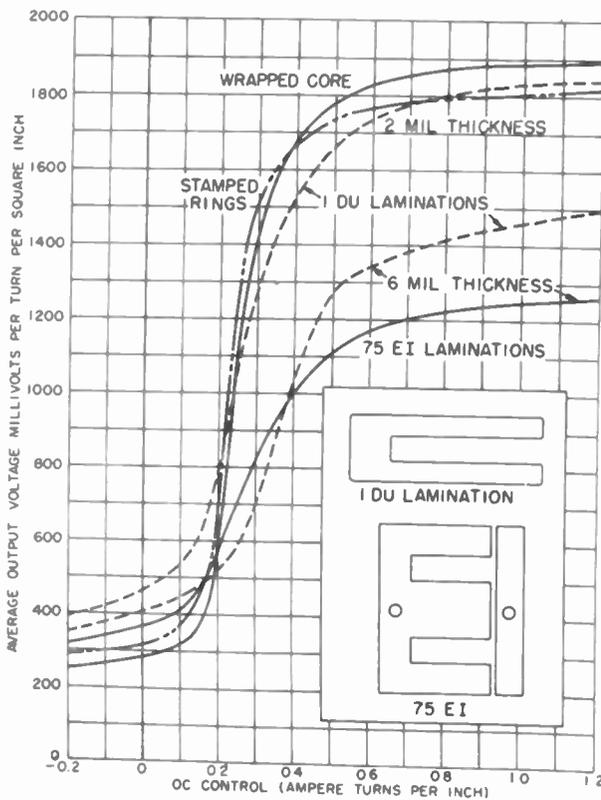


Fig. 6
Carpenter 49 alloy in various shapes; 400 cps.

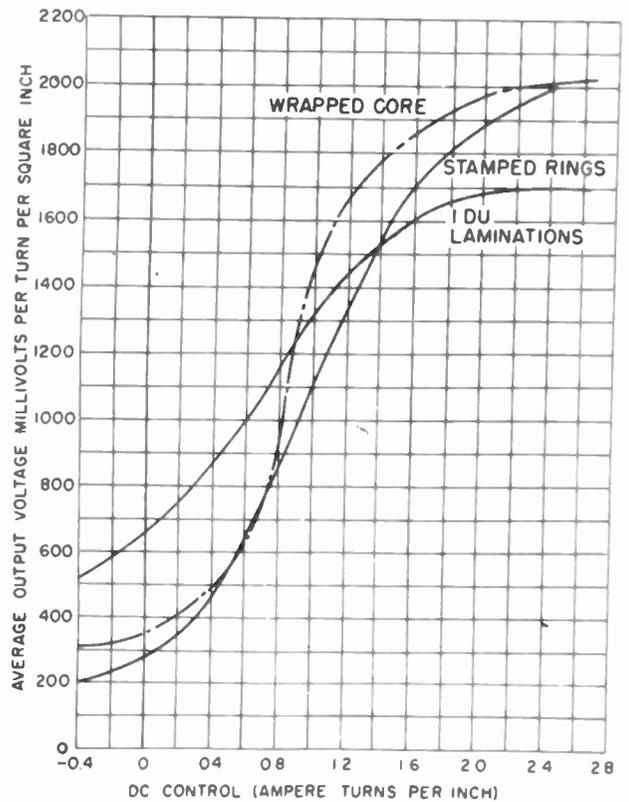


Fig. 7
Microsil oriented silicon-iron alloy in various shapes; 4 mil thickness 400 cps.

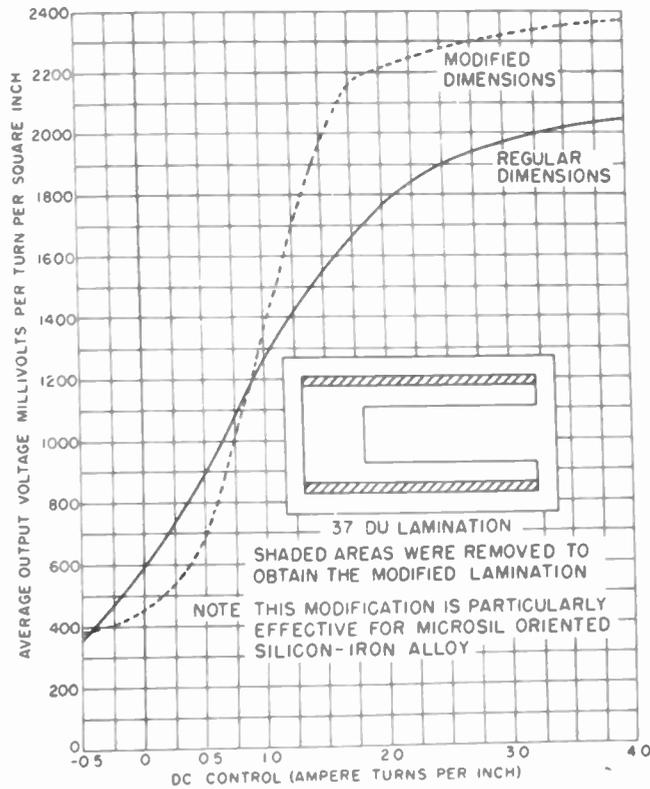


Fig. 8
Effect of modifying lamination dimensions.
Microsil oriented silicon-iron alloy;
4 mil thickness 400 cps.

TABLE I
SPECIFIED PARAMETERS FOR VARIOUS ALLOYS

400 Cycle Tests	Ampere Turns Per Inch			Millivolts per turn per in ²	
	Peak Excitation A	DC Control for Saturation H _s	Departure from Linearity	Production Test Values	
				Lower Voltage V ₁	Upper Voltage V ₂
79% Nickel	4.0 (2 oersteds)	1.2	0.02	500	800
50% Nickel	4.0 (2 oersteds)	1.2	0.02	700	1100
3% Silicon-Iron	8.0 (4 oersteds)	3.0	0.1	700	1300

TABLE II

NORMALIZED DATA FOR HYMU 80 ALLOY IN
VARIOUS SHAPES AND THICKNESSES

<u>Shape</u>	<u>Thickness</u> (mils)	<u>V_o</u>	<u>S</u>	<u>V_s</u>	<u>D (%)</u>	<u>H_c</u>	<u>V_c</u>
Wrapped Cores	2	280	32.0	1200	51	0.11	600
	6	290	6.0	1100	48	0.12	600
Rings	2	290	40.0	1280	59	0.10	700
	6	310	5.8	1230	55	0.12	700
1DU Lamination	2	410	4.4	1130	39	0.15	700
	6	360	3.6	1150	41	0.15	650

TABLE III

NORMALIZED DATA FOR ORTHONIC ALLOY IN
VARIOUS SHAPES AND THICKNESSES

<u>Shape</u>	<u>Thickness</u> (mils)	<u>V_o</u>	<u>S</u>	<u>V_s</u>	<u>D (%)</u>	<u>H_c</u>	<u>V_c</u>
Wrapped Cores	2	135	7.1	1850	73	0.66	950
	4	143	5.6	1800	71	0.67	990
	6	130	4.9	1750	75	0.82	950
Rings	2	300	6.6	1750	61	0.58	1000
	4	275	8.2	1660	60	0.55	920
	6	260	4.8	1630	58	0.62	880
1DU Laminations	2	320	2.3	1450	63	0.77	950
	4	330	3.1	1600	49	0.73	970
	6	225	3.3	1800	54	0.70	1050
37DU Lamination Reg.	4	325	4.8	2100	45	0.55	1200
37DU Lamination Mod.	4	345	4.6	2120	45	0.55	1200

TABLE IV

NORMALIZED DATA FOR CARPENTER 49 ALLOY
IN VARIOUS SHAPES AND THICKNESSES

<u>Shape</u>	<u>Thickness</u> (mils)	<u>V_o</u>	<u>S</u>	<u>V_s</u>	<u>D (%)</u>	<u>H_c</u>	<u>V_c</u>
Wrapped Cores	2	265	13.0	1920	57	0.27	910
	6	245	5.6	1510	60	0.34	815
Rings	2	320	10.5	1800	57	0.23	950
	6	260	6.0	1820	59	0.34	945
1DU Laminations	2	460	5.0	1850	45	0.24	1015
	6	405	4.3	1510	35	0.40	980
75EI	6	360	2.6	1300	43	0.26	740

TABLE V

NORMALIZED DATA FOR SILICON IRON ALLOY
IN VARIOUS SHAPES AND THICKNESSES

<u>Shape</u>	<u>Thickness</u> (mil)	<u>V_o</u>	<u>S</u>	<u>V_s</u>	<u>D (%)</u>	<u>H_c</u>	<u>V_c</u>
Wrapped Cores	4	390	3.4	2030	44	.82	1100
	14	200	2.1	2090	51	1.02	900
Rings	4	270	1.4	1880	64	.96	950
	14	300	.80	1640	60	1.10	920
1DU Laminations	4	690	.88	1700	35	.72	1100
	14	510	.68	1690	61	.92	1050
37DU Laminations Reg.	4	590	.73	1970	48	.78	1085
	14	590	.40	1780	42	1.42	1270
37DU Laminations Mod.	4	450	1.4	2300	67	.99	1340
	14	660	.53	2160	53	1.55	1530

THE HOLLOW CATHODE IN CYLINDRICAL GEOMETRY

B. D. Kumpfer and Herbert Brett
Signal Corps Engineering Laboratories, Belmar, N. J.

Summary--A form of the hollow cathode is described which provides a radial electron beam of small axial extent. Experimental investigation of such cathodes indicates that emission densities of several amperes per square centimeter are obtainable and that such emission is stable because the shielded position of the coating renders it immune to ion bombardment and electron back bombardment. Evidence is presented which indicates that all parts of the hollow cathode contribute to the total emission observed from hollow cathodes.

The so-called "hollow cathode", in which electronic emission is produced from a small aperture in a heated metal enclosure coated on its inner surface with a thermionic emitting material, has been investigated both theoretically and experimentally by several workers. Pencil beams having current densities of tens of amperes per square centimeter, when calculated on the basis of the aperture area, have been produced. The behavior of emitters of this type differs fundamentally from that of conventional space-charge limited emitters as given by the well known Child-Langmuir three-halves power emission law in that the current varies linearly with voltage over the major portion of the characteristic, and true temperature-limited saturation is not reached in practical cases. Another peculiarity of hollow cathodes is that the beams so produced conform to the shape of the orifice in cross section and are hollow under some conditions. Typical geometries investigated previously are indicated in Fig. 1. These include the hollow sphere and the hollow cylinder. Analytically, the hollow sphere has been subjected to the closest scrutiny because of its mathematical simplicity.

To date, the mechanism of thermionic emission from such cathodes is even less perfectly understood than conventional emission phenomena. The cross-section of the beams suggest that emission occurs principally from the periphery of the hole and is enhanced by the increased electric field gradient at the edges. According to this simple concept, the role of the structure behind the aperture is merely to act as a dispenser to replenish the emitting material as it is depleted in the region of the edges. In this case, the geometry of the enclosure is relatively unimportant.

A more sophisticated theory to explain the behavior of the hollow cathode has resulted from an analysis by H. S. Wu, Von Foerster and others at the University of Illinois. This approach postulates the generation of a stable electron gas cloud within the isothermal sphere having a maximum electron density adjacent to the surface, a potential maximum at the center of the enclosure, and an electric field intensity of several hundred volts per square centimeter at the inner surface of the sphere. Wu computes the total pressure acting on the walls of the sphere as the sum of the electrostatic pressure and the pressure due to the kinetic energies of the electrons and shows that if a hole is made

through the surface of this "thermionic teakettle", a copious stream of electrons will flow from the aperture. It is also shown that the apparent depth of the electron shell within the sphere is relatively small; being only 3×10^{-4} centimeter in an example considered.

The hollow cathode of the pencil beam type just described has possible applications in cathode ray devices such as klystrons, traveling wave tubes, and display tubes. Its advent suggested the possibility of other types for other purposes. Workers at the Signal Corps Engineering Laboratories at Belmar, N. J., recently became interested in a form of the hollow cathode to produce a radial electron beam. This type is of interest because of its potential application to magnetrons, resonant, radial beam klystrons, and radial beam display tubes. An experimental investigation of the characteristics of hollow cathodes for producing such beams is the subject of this paper.

The geometry of the radial beam hollow cathode is shown in Fig. 2. As indicated, this shape may be thought of as being generated by revolving a spherical pencil beam hollow cathode about the axis shown. The result is a toroidal enclosure with a circumferential slot around the outer surface. When an electron gun of this kind is used in conjunction with a cylindrical anode, a radial stream of electrons which is thin in the axial direction results. Actually, in the experimental embodiment used, a half-torus with a slot in the outer flat surface was employed for ease of fabrication. A practical tube incorporating this form is shown in Fig. 3. The cathode is divided in the plane of the slot to facilitate spray-coating the inner surface. The emitter is heated indirectly by a conventional alumina insulated tungsten wire heater in the base. A heat shield was employed to conserve heating power.

A cylindrical anode of 80 x 80 nickel screen was used in these experimental diodes and the inner wall of the glass envelope was coated with willmitite to permit investigation of the beam cross-section. The magnified image of the screen mesh made observation of the beam dimensions possible.

In order to study the effects of aperture geometry, experimental diodes were constructed with gaps of .005, .010 and .015 inch. Also, since the electric field distribution near the aperture was assumed to be important, some tubes were built with outwardly flared edges to enhance the gradient. Gaps of .010 and .015 inch were tried.

The tubes were processed in accordance with modern vacuum techniques. Rigorous oven bake-out at 450°C. followed by the outgassing of internal metal parts by radio-frequency heating was employed. Getters were incorporated in all tubes. The activation of the hollow cathodes was normal except that there was a marked absence of poisoning effects

usually encountered during activation at higher voltages. Stable emission was rapidly reached.

Fig. 4 shows the thermionic performance of a typical radial beam hollow cathode. The slot width is .005 inch and the cathode diameter is .240 inch as indicated in the figure, making the aperture area .025 square centimeter. Emission current density as a function of anode voltage is shown for several cathode temperatures. It will first be noted that the characteristic is essentially linear over most of the range investigated here rather than following the three-halves power law dictated by space-charge limitation. Secondly, the current density is considerably greater than that predicted by the Child-Langmuir law for an equivalent diode, as shown dotted in the figure. It will also be noted that a change in emitter temperature changes the slope of the emission line, rather than extending the current value at which saturation is reached, as in the conventional space-charge limited case. In this respect, the emission of the hollow cathode resembles the temperature-limited portion of the normal diode curve.

Within the limits of experimental error and discounting the variation from tube to tube due to differences in thermionic activity, the total current obtained was nearly independent of slot area and shape. Therefore, the greatest current densities were observed in tubes with the smallest gaps. There was no measurable improvement in emission in tubes with flared apertures.

The electron beam produced by the cylindrical hollow cathode, as viewed on the willemitte screen, appeared double at low current densities and low cathode temperatures. Small mechanical irregularities in the aperture edges were visible in the image. At higher current densities, the hollowness of the electron stream could not be verified due to smearing of the image by stray secondaries. The width of the beam at the anode could be measured since the image of the screen mesh was clearly visible on the willemitte. The beam for the .005 inch slot tube was approximately twelve times that width. It was not determined if this divergence is due to the inherent ballistics of the cathode or the divergence of the electric field. The beam width was essentially independent of current density, indicating that external space charge does not play an important part in producing the observed divergence.

Since the two most prevalent concepts as to the mechanism of emission from hollow cathodes differ principally in the relative roles of the emission from the aperture edges as compared to that from the remaining parts of the structure, an experimental effort was made to isolate these regions. First, several tubes were constructed in which the coating was omitted from the flat surfaces adjacent to the apertures. Such tubes behave poorly when processed in the manner described above. Activation was slow and there was a marked tendency toward emission poisoning during activation. Current densities observed were considerably below the theoretical performance of an equivalent diode.

Another experimental tube was then constructed

in which it was possible to observe the emission from the slot alone, without electron contribution from the back structure. The construction of this tube is illustrated in Fig. 5. The emitter is a directly heated planar strip filament of thin nickel foil with a slot aperture of the same area used in the cylindrical tubes. Emitting material was applied only on one side in a thin strip on either edge of the slot. A planar anode is disposed on the uncoated side of the heater element and a concave element intended to simulate the back part of the hollow cathode was placed on the coated side. Current could then be drawn through the slot to the anode or directly to the unheated "reflector" electrode for comparison purposes.

Fig. 6 compares the current drawn directly to the back electrode with that through the slot on a relative current density basis, assuming that the area of the coated portion is approximately ten times that of the slot. These data were taken at low cathode temperature to produce reflector current saturation before the dissipation of that electrode was exceeded. At this temperature, the current to the anode behaves in a manner similar to the direct emission to the reflector. The most striking difference observed was the absence of poisoning during activation when current was drawn through the slot to the anode. Poisoning was experienced when current was drawn directly to the back electrode. Evidently, the shielded position of the coating renders it immune to ion bombardment from the cathode-anode region.

The effect of impressing an electric field between the reflector and the emitter is shown in Fig. 7. Holding the anode voltage constant, it was found that the anode current increased when the reflector was made positive with respect to the emitter, and decreased when it was made negative. This effect is shown in another form in Fig. 8, where anode current is shown as a function of anode voltage for various reflector conditions. Here it is seen that the characteristic becomes more linear when current is drawn to the reflector. This has nearly the same effect as raising the cathode temperature as shown by the straight line above. A three-halves power curve is also shown dotted for comparison and the theoretical curve for an equivalent planar diode is shown below. Note that the emission through the slot is about ten times the theoretical current based on the aperture area, indicating that the entire coated area is contributing to the total emission.

From these experiments one can draw the following conclusions:

- a. That the radial beam hollow cathode provides a source of electrons of moderate current density in a radial beam of small axial extent.
- b. That the activation of cathodes of this type is simplified by the fact that the emitting surface is essentially shielded from ion bombardment and that this feature should also make it useful in magnetron tubes where back bombardment of the emitting surface by electrons is a problem.

c. That, although emission from the edges of the aperture plays a dominant part in causing the emission curve to depart from the three-halves power law, it does not account for the high current densities observed.

d. That the contribution of electrons from the back structure is important in providing high current densities.

e. That emission from the hollow cathode resembles temperature limited emission more nearly than space charge limited emission.

¹H. S. Wu and H. M. Von Foerster, "Thermodynamics and Statistics of The Electron Gas", Part I, University of Illinois Technical Report No. 3-1; Oct. 1950

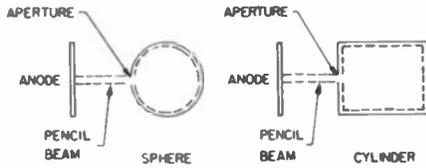


Fig. 1

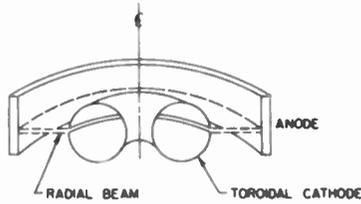


Fig. 2

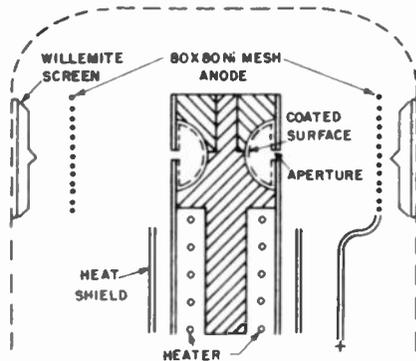


Fig. 3 - Construction of radial hollow cathode.

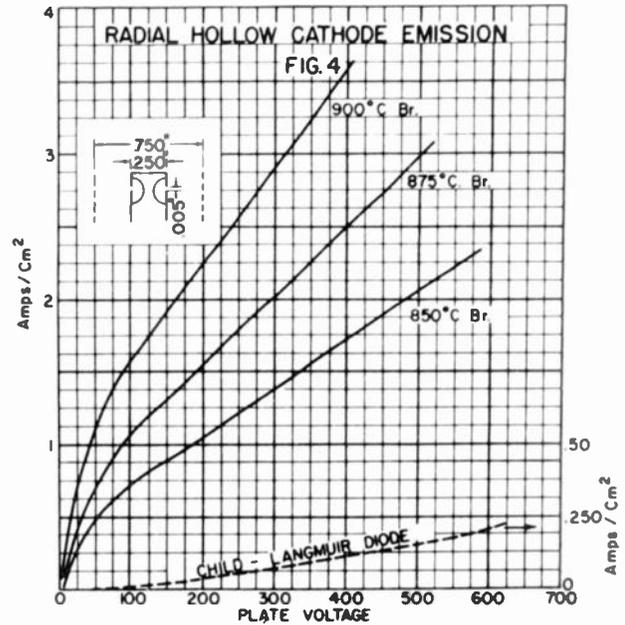


Fig. 4 - Radial hollow cathode emission.

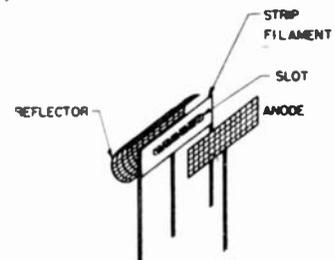


Fig. 5 - Construction of planar slot tube.

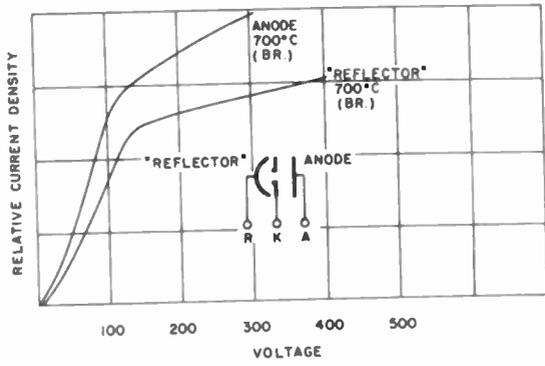


Fig. 6

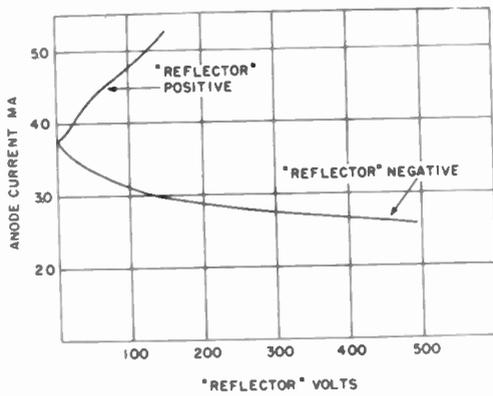


Fig. 7

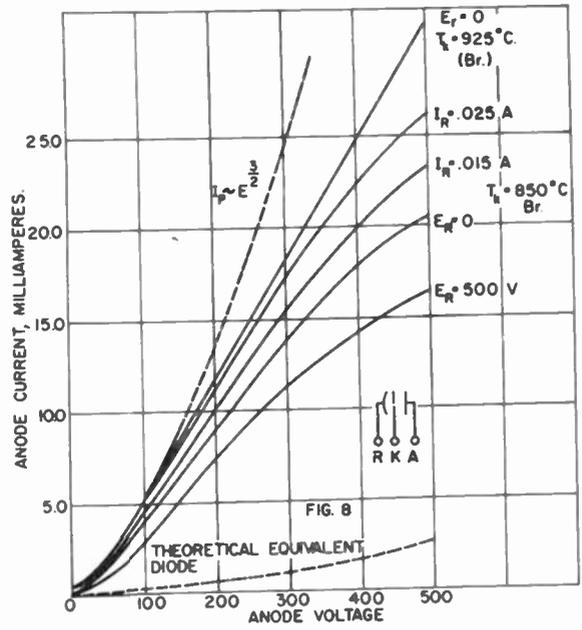


Fig. 8

THE MACHINING OF TUNGSTEN AND ITS APPLICATION IN THE
FABRICATION OF PHILIPS DISPENSER CATHODES

Roberto Levi
Philips Laboratories, Inc.
Irvington-on-Hudson, New York

In a paper¹⁾ presented by this author at last year's National Convention of the Institute of Radio Engineers, some of the advantages of the new "impregnated" cathode over the "L" cathode were discussed. A comparison of the two cathodes, in both their planar and cylindrical versions, is shown in Figures 1 and 2. The most important component of both the "L" and impregnated cathodes is a tungsten part which has been sintered to a high density and has accurately controlled dimensions, porosity and gas permeability; its outer surface is the emitting area.

In the past, tungsten sintered to the required high density, that is in the neighborhood of 80 - 85% of the theoretical value, could not be machined; the combination of hardness and brittleness of the material at ordinary temperature plus the difficulties inherent to the machining of most sintered porous bodies made this impossible. Therefore the parts, depending on their size and shape, were preformed either by pressing tungsten powder in a die, or by machining them from a partly sintered and very porous bar having a density of the order of 55% of the theoretical value. In the latter case, because of the very weak bond between the particles, these were torn out in clusters during the machining operation rather than cut and therefore smooth surfaces could not be obtained. Whether preformed by pressing tungsten powder in a die or by machining from a presintered ingot, the parts were then sintered to the required degree. During this final sintering operation, considerable shrinking and warping occurred making it exceedingly difficult in most cases to attain the required shape and dimensions.

The author has developed a technique²⁾ for machining tungsten ingots sintered to a high density by which parts of specified porosity, can be made in shapes and with tolerances comparable to those possible with steel or brass. Since the tungsten has been sintered to the required degree before machining, the problem of shrinking and warping is eliminated.

Basically, the technique consists in first infiltrating the fully sintered porous ingot with a metal which complies with the following requirements.

- a) at a temperature somewhat higher than its melting point it must be capable of wetting tungsten and of penetrating the porous body by capillary action;
- b) both the infiltrant and tungsten should be mutually insoluble not only below but

also above the melting point of the infiltrant and should form no alloys in the metallographic sense;

- c) it should act as a lubricant during the subsequent machining step.

Gold, copper and alloys of the two in all proportions appear to conform best to the above requirements.

The second step in the process is the actual machining of the impregnated body. Since the infiltrant acts as a filler and as a lubricant, it will prevent the tearing out of particles, burnishing and extremely high tool wear which would otherwise occur.

Finally, the infiltrant is removed by volatilization, thus restoring the density and porosity of the machined part to the original values established for the tungsten during the sintering operation.

While the infiltration step is similar to that employed in the manufacture of certain types of electrical contact materials, its purpose, as we have seen, is quite different. Moreover, the successful practice of the entire technique imposes a number of additional requirements.

For example, the impregnation must be carried out with the greatest care, for if a small portion of the body is not properly infiltrated, breakage of the tungsten or of the tool may result. Furthermore, in order for the machined part to retain dimensional stability and proper porosity the tungsten ingot must first be sintered for a sufficient length of time at a temperature considerably higher than the highest temperature to which it will be subjected during infiltration, evaporation of the infiltrant or other required steps such as cathode activation and operation. This is achieved by the proper combination of factors such as characteristics of the tungsten powder, pressure used in forming the bars and dew point of the sintering atmosphere, so that the required sintering temperature will be within a convenient range.

It should also be noted that not all the metals used as infiltrants in contact materials are best suited in connection with the technique described in this paper and vice versa. For instance, while the high electrical conductivity of silver makes this metal highly desirable in contact materials, silver is not as satisfactory as copper or gold in our case, since at temperatures above its melting point it shows a slight solubility for tungsten. The

reprecipitation of tungsten onto the larger grains, which takes place upon cooling, changes somewhat the desired porosity attained during the sintering operation. On the other hand, the use of gold as the infiltrant in contact materials is normally prohibitive in view of its cost, but for our application it is probably the most suitable metal; since the gold is subsequently removed by volatilization and can be recovered, its cost factor becomes negligible.

Details of the technique are illustrated by the following typical example:

- 1) Tungsten powder having the characteristics shown in Table 1 is pressed into bars at 2,000 Kg/cm² and presintered for twenty minutes at 1100°C Br to permit subsequent handling; its density at this stage is about 55% of the theoretical value.
- 2) The ingots are then sintered for twenty minutes at 2400°C Br in an atmosphere of cracked anhydrous ammonia. During this stage considerable shrinkage and warpage occurs and the density of the ingot reaches a value of 83% - 84%.
- 3) The sintered ingots are next impregnated with copper at 1350°C Br for a period of not less than 10 minutes in the case of ingots 3/8" x 1/8"; larger ingots require a longer impregnating time. This is carried out in a hydrogen atmosphere by placing the ingot horizontally on top of a weighed amount of OFHC copper, slightly in excess of the amount which can infiltrate the pores; in this particular case 8% to 10% of the weight of the tungsten. The temperature is slowly raised to a point slightly below the melting point of the impregnant and held there a few minutes to permit the interior of the ingot to attain the same temperature as the surface. When the temperature is finally raised to 1350°C Br the molten copper will penetrate the tungsten body from the bottom by capillary action, a process facilitated by the fluxing action of the hydrogen.
- 4) After cooling the impregnated bar is machined with carbide or high speed steel tools and all normal machining operations are possible.

No tungsten grains can be detected under microscopic examination of a freshly machined surface since a thin copper film has smeared over the entire area. If, however, the copper film is chemically removed from the surface of the machined grains the smoothness and flatness of the grains indicate that they have actually been cut out by the tool and not merely torn out.

- 5) Finally the copper is volatilized by heating the machined parts in a vacuum

furnace at 1800 - 1900°C Br. The resulting parts under spectroscopic examination show only an extremely faint trace of copper and no shrinkage has taken place since the tungsten frame had previously been sintered at a much higher temperature, namely, 2400°C Br.

These various stages are illustrated in Figure 3.

Referring now to Figure 4, there is illustrated therein a photomicrograph of a polished section of the tungsten surface, after removal of the copper. As can be seen, the average pore size is of the order of a few microns and when the material is used in Philips Dispenser Cathodes all of the surface should contribute to the emission since the maximum distance between pores is very small compared to the migration length of barium on tungsten.

The technique described allows the fabrication of tungsten parts having a density up to slightly less than 90% of the theoretical value. While all normal machining operations can be carried out with relative ease in the case of material of density up to about 83%, further increases in density will make the machining progressively more difficult, because of a rapid increase in the percentage of non-connecting pores which cannot be infiltrated. However, since in cathode applications a tungsten part must have a large percentage of interconnecting pores, the technique is adequate for the fabrication of cathodes having the highest densities which may be required.

It is also interesting to note that this process allows the assembly of magnetron cathodes of the "impregnated" type by screwing directly the tungsten emitting sleeve to the molybdenum end pieces as can be seen in Figure 5. Furthermore, the emitting surface of planar type cathodes can be lapped to a flatness within a few fringes of sodium light which appears to be quite advantageous in a number of applications. Since the lapping operation is carried out before volatilization of the filler metal, the pores of the tungsten body have no tendency to close up. In addition the new technique allows fabrication of extremely fine parts which could not previously be made either by pressing in a die or by machining the presintered ingot because of the inherent weakness of the material.

While the main application of the technique described has so far been in the fabrication of Philips Dispenser Cathodes, it is expected that new uses will develop in other fields.

- 1) R. Levi, "New Dispenser Type Thermionic Cathode" 1953 I.R.E. Convention Record, Part 6, pp.40-42.

2) R. Levi, U.S. Patent No. 2,669,008 Feb. 16, 1954.

3) E.S. Rittner and R.H. Ahlert, Private communication.

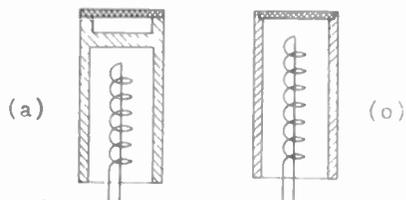


Fig. 1
 (a) planar type L-cathode;
 (b) planar type impregnated cathode.

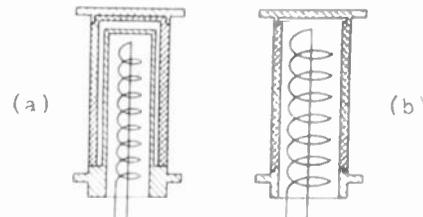


Fig. 2
 (a) cylindrical type L-cathode;
 (b) cylindrical type impregnated cathode.

ELUTRIATION ANALYSIS

FRACTION NO.	PER CENT	EQUIV. PARTICLE RADIUS
1	29	$>6\mu$
2	25	
3	11	
4	13	
5	22	$<2\mu$

SCOTT DENSITY 68.4 GRAMS PER CUBIC INCH

Table I - Characteristics of tungsten powder.

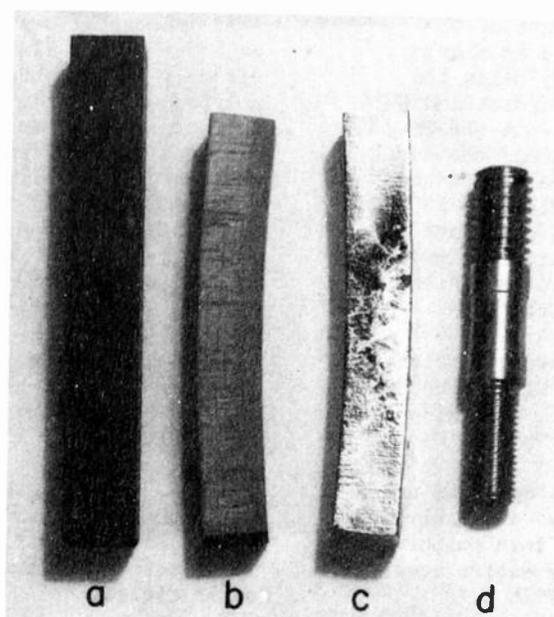


Fig. 3
 (a) pressed and presintered tungsten ingot;
 (b) similar ingot after high temperature sintering;
 (c) copper impregnated sintered ingot;
 (d) pure tungsten machined part made by new method shows turning, drilling and threading.

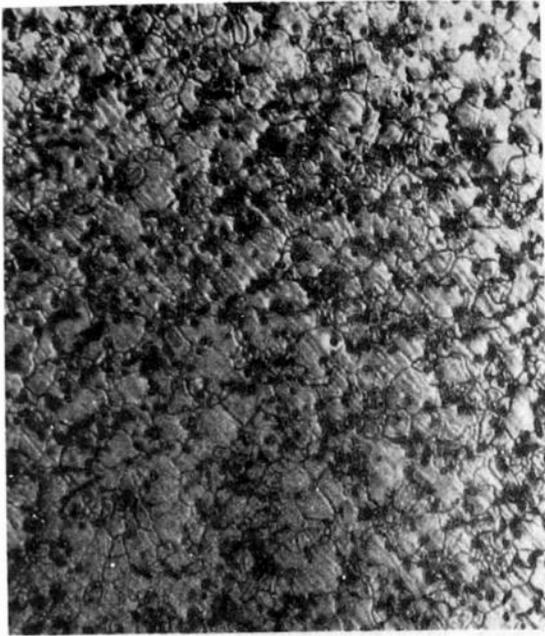


Fig. 4
Photomicrograph of tungsten surface after
removal of filler metal.

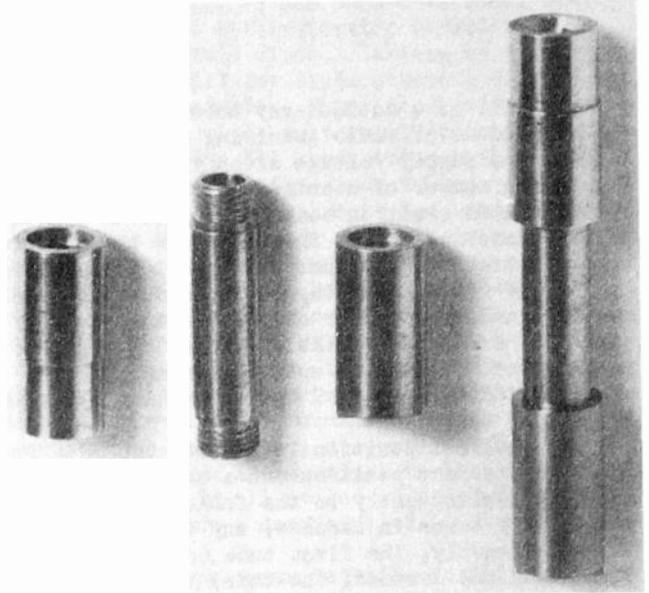


Fig. 5
Threaded magnetron cathode parts and their
assembly by screwing.

THE GE POST ACCELERATION COLOR TUBE

C.G. Lob
General Electric Company
Syracuse, New York

ABSTRACT

The tube to be described is a high brightness, high-definition color tube using three electron guns and a viewing screen consisting of vertical phosphor stripes alternating red-green-blue, red-

green-blue, etc.

Performance data obtained from developmental models of this tube will be presented and many of the advantages of such a device pointed out.

AMPEREX TYPE ELT DECADE COUNTER TUBE

Irwin Rudich
Amperex Electronic Corporation
Hicksville, L. I., N. Y.

The ELT is a cathode-ray tube having roughly the dimensions of radio receiving tube and operation with a supply voltage of only 300 v. According to the number of counting pulses applied, its ribbon-shaped electron beam is shifted in a horizontal plane and passes in succession through the ten apertures of a cylindrical anode, thus impinging on the fluorescent layer with which the envelope is lined. The number of pulses can thus be read on the outside of the envelope, by means of a rectangular luminescent spot appearing opposite one of the figures 0 to 9 indicated on the circumference of the bulb.

As the last position is passed, the beam is reset to its zero position and a counting pulse is applied simultaneously to the following tube. With several ELT tubes in cascade, any number can thus be read directly, the first tube counting the units, the second the decades, the third the hundreds, etc.

In this paper the basic operation of the ELT is described, as well as the principles of 30 kc, 100 kc and 2 megacycle counting circuitry. A note on predetermined counting and resolving time for random pulses is also included.

This tube was developed at the Philips Research Laboratories in Eindhoven, Holland, by Dr. J. L. M. Jonker and co-workers. Amperex Electronic Corporation is distributing and furnishing technical service on this tube in the United States.

The electrode system used in this tube and its symbolic representation, may be seen in Fig. 1. We see it has a cathode, control grid, beam forming electrodes, accelerating electrode, deflection plates, a slotted screen, anode and fluorescent coating on the inside of the glass wall. We also notice 2 screen grids, used to suppress secondary emission as in an ordinary pentode. The cathode, control grid, beam forming electrodes, accelerating electrode, and deflectors constitute an electron gun which forms a ribbon-shaped beam focused on the slotted screen. The deflection plates serve to position this beam on the slotted screen. The screen has 10 slots from one edge of the screen to the other. The current which passes the screen then is adjusted by the slot area to vary according to which slot the beam is focused on. This current is received by the anode and thus a dependence of anode current on the deflector plate voltages is established. The anode itself has small apertures which allow a portion of the beam striking the plate to pass on to the fluorescent screen and give a visual indication of the electron beam position.

In Fig. 2 we can see a representation of the mechanism of the formation of the anode current

characteristic. Fig. 2(a) is obtained when the slots in the screen are as shown in the sketch, while Fig. 2(b) shows what will occur when a horizontal slot is placed in the slotted screen under the vertical slots on the lower left side.

Now let us connect the tube as shown in Fig. 3 and measure the anode current I_a as a function of the common voltage applied to the anode and deflector plate D' while deflector D is held at a fixed voltage. The curve, shown in Fig. 4 is obtained when these results are plotted. It may be seen that the current increases with an undulatory motion as the deflector voltage is decreased. If in addition we insert a resistor of appropriate size in the anode lead, supply it from the 300-volt line and draw its load line as in Fig. 4(a) all the undulations are intercepted. There are ten stable and nine unstable intersections. It may easily be seen that such a position as a is stable, for if the beam moves slightly to the left, the current decreases, and the voltage on the deflector anode increases. This change in voltage is of the right nature to return the beam to its original position. Just the opposite happens in an unstable position, b .

The current to the auxiliary anode A' as the deflector plate voltage varies is also shown at the left of the figure.

One other point may be noted on this curve. That is the differences in voltage between two stable points on the curve. This is approximately 14 volts and is the voltage which is necessary to move the beam from one slot to the next slot, that is to count one.

In Fig. 4(b) we see that by increasing the right deflection plate voltage 14 volts, the beam is in the stable position C' corresponding to number 1 while the anode and right deflection plate voltage are still at the value corresponding to point a or number 0 (230 volts).

If the voltage on left deflection plate D is slowly increased, the beam will move in such a direction as to decrease the anode current. In turn this will cause the voltage on the anode and right deflection plate D' to increase and try to maintain the voltage difference between D and D' constant. This keeps the beam in the slot. This is simply a restatement of the stability criterion. Turn now to Fig. 5. Because there is a shunt capacitance associated with the anode, a rapid rise in the potential of D cannot be followed by D' and the beam will jump to the next slot. In order to maintain the beam in this position, the pulse must not decay faster than the anode voltage can change. As the pulse on the left deflection plate decays

the anode current increases causing the voltage on the anode and D' to decrease. As D returns to its rest potential of 156 volts, D' decays to a new value roughly 14 volts lower than its value before the pulse. Thus a permanent change in deflector voltages has been accomplished and the beam is stably located in the next slot.

The value of the anode interelectrode capacity, the load resistance and the peak of the available current "hump" determine the requirements of the pulse shown in Fig. 5(b). For the average pulse of 13.6 volts the rise time should be less than .7 microseconds and the fall should take longer than 7 microseconds.

We now turn to Fig. 6 which shows one ELT counting stage and its reset multi-vibrator.

On the tenth pulse the beam falls on the auxiliary reset anode and causes a voltage drop in R5 which is coupled to the reset one shot multi-vibrator. The components of this circuit are so proportioned as to generate a pulse at point D properly shaped to operate the next decade stage. From the plate of the right half of the twin triode a negative going pulse is available for the grid of the ELT. This negative pulse cuts off the beam current long enough to allow the anode A2 and D' to return to the voltage which corresponds to zero count. These wave forms are shown in Fig. 7.

The reset time in this circuit is determined by the anode load resistor of 1 megohm and the anode and deflection plate capacity of 16.5 maximum. There is an anode voltage difference of approximately 145 volts between the zero and 9 position. A time of around 21 microseconds is required for the capacity to discharge this voltage through the 1 megohm load.

This time is the main limit on the counting rate using this circuit - which is limited to 30 kc, because of necessary tolerances on associated circuitry.

We can see in Fig. 8(a) the "Staircase Waves" on the anode of the ELT tube when pulses of a uniform frequency of 3,000 cps are counted. At 30 kc, as in Fig. 8(b) the reset time takes up most of the time that the voltage would be at zero position.

It is interesting to note that the resolving time of the tube proper is very low during the count from 0 to 9. However, in the 30-kc circuit the dead time of the pulse shaping circuit that generates the pulse shown in Fig. 5(b), is necessarily greater than 7.7 microseconds.

It can be shown that the resolving time for random pulses in the simple circuit shown in Fig. 6 is shorter than the 30-kc continuous counting rate would indicate. Since the dead time is a maximum of for 33 microseconds for only one out of ten counts, the average contribution is only 3.3 microseconds. This in many instances will contribute little error to counting, especially when used with customary geiger counters that have a

dead time of more than 50 microseconds.

If there is the necessity to count continuously up to 100 kilocycles, a modified first stage circuit (Fig. 9) has been developed. To the ELT and the reset multi-vibrator we add two twin triodes and two twin diodes. Instead of cutting off the grid of the ELT for flyback we cut off the normally heavily conducting paralleled twin triode. The fast rise in plate voltage of this tube is communicated to the ELT anode through the diode and rapidly increases the anode voltage and beam position from the nine value to the zero value. In other words, the ELT anode capacity is rapidly discharged through a low impedance path in parallel with the 1 megohm normal load. Because of the greater complexity of this 100 kc circuit it has been found necessary to select ELT tubes for this first stage use. However, it can be said that a good percentage will work reliably so this is a practical circuit that requires no more tubes than an equivalent performance stage using all twin triodes. It offers an advantage in that ELT tubes can now be used in all stages of a decade scaler that operates with a dead time of 10 microseconds.

Extending the above technique using such high Ω_m , low capacity tubes as the EFP60 (a secondary emission pentode) it has been possible in the laboratory to obtain a maximum counting rate of 2.2 megacycles.

In Fig. 10 can be seen basic additional circuitry that permits predetermined counting. The remainder of the circuit is the same as in the normal 30-kc circuit of four decades.

Suppose we wanted to count 7236 pulses. We first determine the complimentary number for a four decade counter that is 10,000 minus 7236 equals 2764.

If we close switch S and adjust the potentiometers, P1, P2, P3, P4, the anode and right deflection plate voltages of each can be set at any DC values. For example, we can set stage one so that the beam is opposite 4 on the fluorescent screen of that ELT, stage two to read 6, stage three to read 7 and stage four to read 2. On the face of the counting unit we then read 2764 since the actual number is in inverse order to the progression of signal through the unit.

We can now open switch S and the 2764 still will remain on the counters since the beams of the tubes are in stable positions and the diodes will not conduct. The cathode of the diodes are more positive than the preset voltages when the switch is opened.

If we now introduce 10,000 minus 2764 equals 7236 pulses into the first stage, all the stages will switch to zero. When this occurs the last stage will deliver a pulse which can trigger a relay to actuate a sorter or packaging device or stop a timer. The pulse can also close switch S (which is actually a high conduction triode) and start the preset cycle over again without the need

for further adjustment of the potentiometers.

Due to the additional capacity introduced by the diodes the maximum counting rate of the practical circuit is 12,500 counts per second.

Designed and manufactured using long life computer tube technique the ELT has low heater power, using 6.3 volts and .3 amperes. The drain on the 300-volt B supply (which incidentally does not have to be regulated) is less than one milli-ampere for the tube proper. The unusual advantages can be seen by the fact that a scaler capable of counting and indicating up to 10 million with a maximum counting rate of 30,000 counts per second

can be built in a cabinet less than 17" wide by 6-5/8" deep by 6-1/8" high with a total power consumption of only 77 volt-amperes.

The features of the ELT render this decade counter tube particularly suitable for those applications where small dimensions, high counting speed, direct indication and reliability are required.

Typical applications include radio-active radiation counters such as decade scalars, business machines, electronic computers, totalizers, differential analysers, batch counters, production counters, timers and frequency meters and checkers.

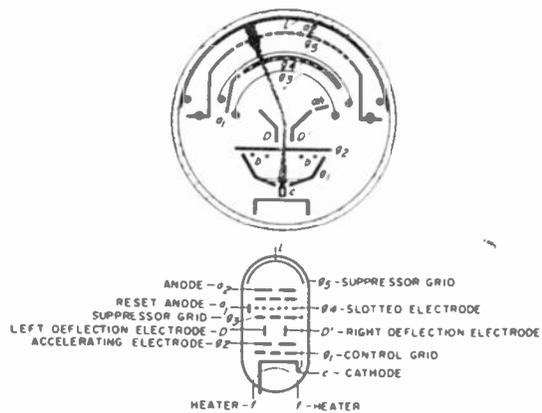


Fig. 1
Cross section and diagrammatic representation.

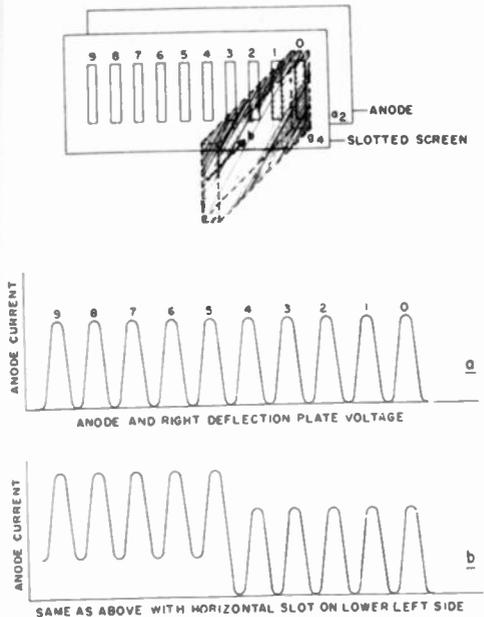


Fig. 2 - Slotted screen produces undulatory anode current characteristics.

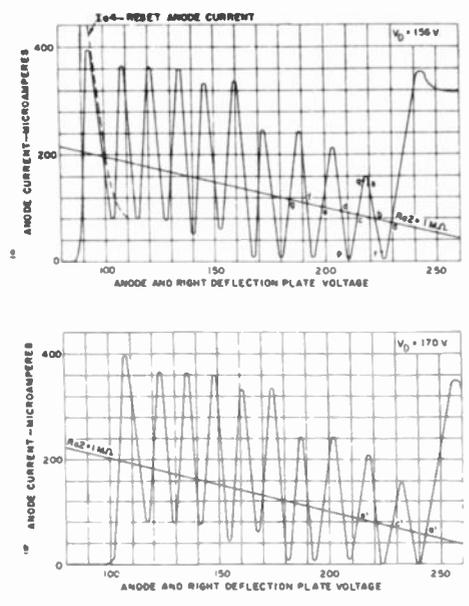


Fig. 4 - Measured characteristic and effect of changing left deflection plate (input signal) voltage.

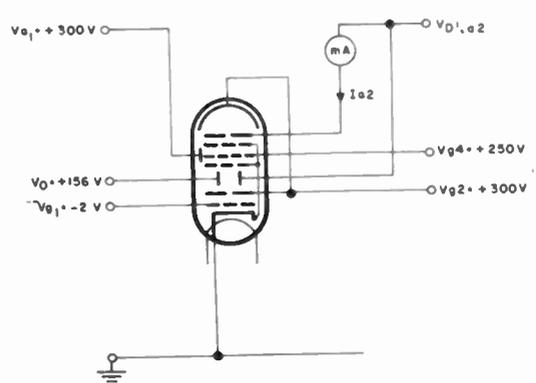


Fig. 3 - Measuring circuit for Fig. 4.

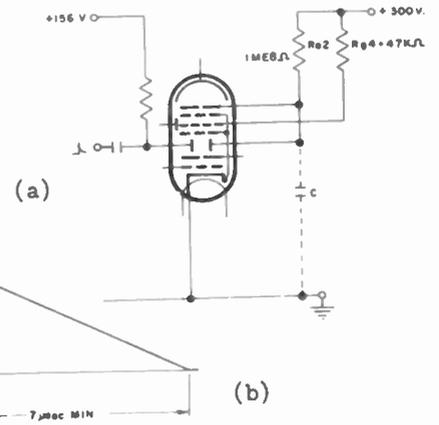


Fig. 5 (a) EIT counting circuit; (b) limits on counting pulse on left deflection plate.

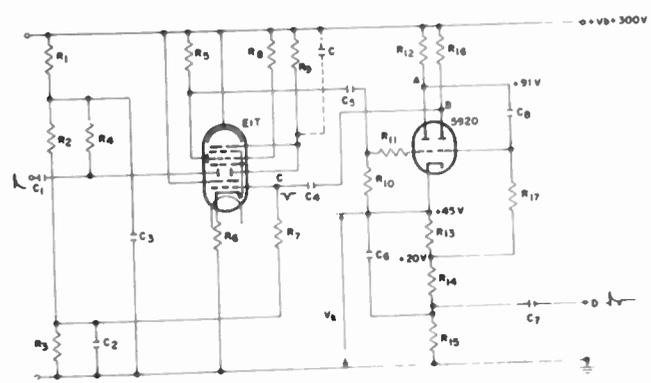


Fig. 6 One EIT counting stage and reset multivibrator.

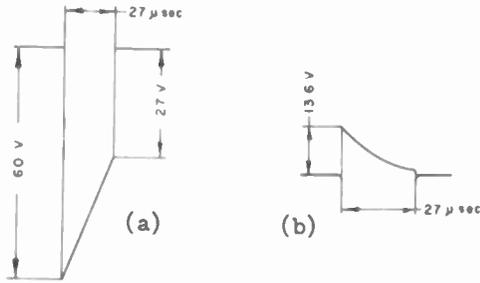


Fig. 7

Waveforms generated by reset multivibrator:
 (a) reset pulse to cut off grid of previous EIT.
 Point C of Fig. 6; (b) pulse to be applied
 to left deflection plate of next EIT. Point D
 of Fig. 6.

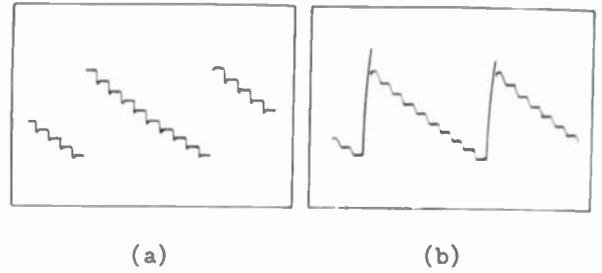


Fig. 8

Anode and right deflection plate voltage:
 (a) oscillogram at 3,000 cps;
 (b) oscillogram at 30,000 cps.

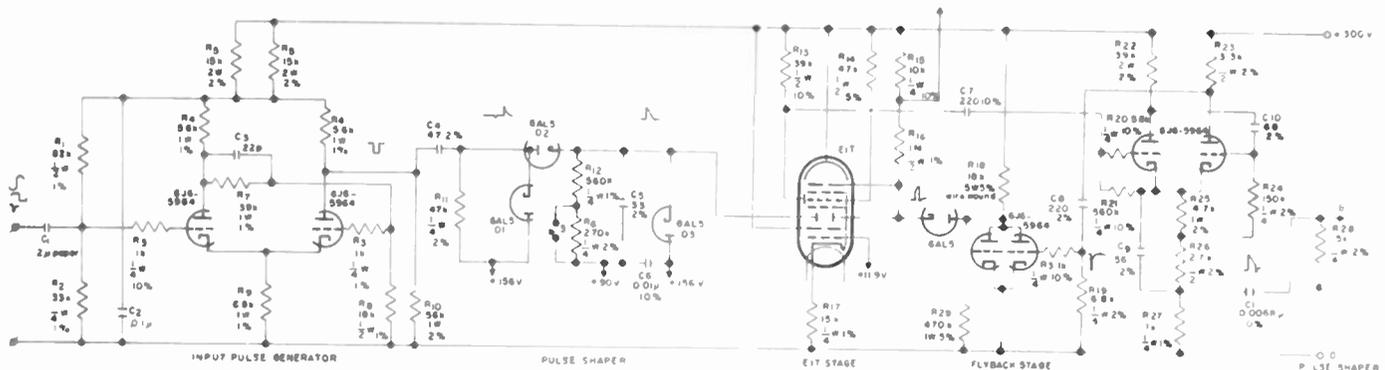


Fig. 9

Modification of input circuit to allow 100-kc continuous counting.

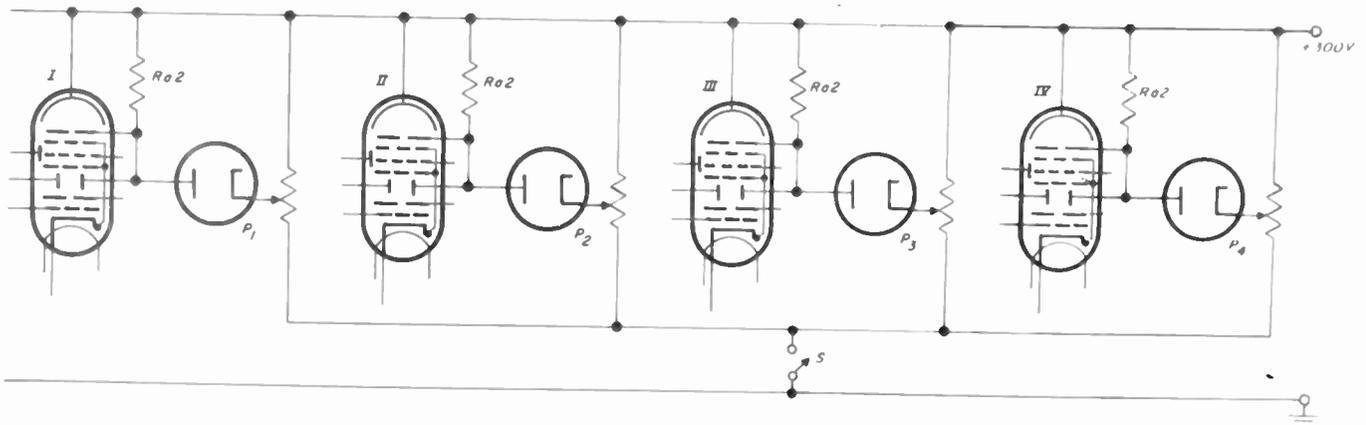


Fig. 10

Basic circuit of a four-decade predetermined counter. In order to obtain a cycle of 7236 counts (i.e. the complementary of $10,000 - 7236 = 2764$), the potentiometers $P_1 - P_4$ must be so adjusted that the counter tubes I - IV are preset to Figs. 4, 6, 7 and 2, respectively.

A DEVELOPMENTAL THYRATRON CAPABLE OF
CURRENT INTERRUPTION BY GRID ACTION

E.O. Johnson
RCA Laboratories Division
Princeton, New Jersey

J.A. Olmstead
RCA Victor Division
Camden, New Jersey

W.M. Webster
RCA Victor Division
Princeton, New Jersey

ABSTRACT

This paper describes a new type of thyatron whose grid can interrupt as well as initiate conduction. In addition to having an essentially zero recovery time this developmental tube has the added feature of being free of noise and oscillations commonly observed with conventional hot cathode gas

tubes. With respect to initiation of current flow, arc drop during conduction, and other salient characteristics, the tube is similar to conventional thyratrons. Design aspects, performance, and principles of operation of this tube will be presented.

TRANSISTORS FOR HIGH POWER APPLICATION

John S. Saby
Electronics Laboratory
General Electric Company
Syracuse, New York

Introduction and Summary

The advent of junction transistors has promised reliability for complex systems and great economy of space, weight, and power - all of vital interest to modern electronics. But transistors designed for milliwatts cannot deliver the watts which are needed somewhere in nearly every system. It has been found possible to use power transistors for outputs of watts, tens of watts, even more than 100 watts; but to do this with stability at high ambient temperatures, yet with a minimum of heavy fins or other bulky cooling equipment, requires transistors designed specially for high power and applied with full regard for thermal as well as electrical characteristics.

Power transistors must be designed to meet three nearly independent conditions:

1. The maximum permissible output current must be high enough to provide the required output without exceeding the maximum voltage limits imposed by power supply or transistor.

2. Since maximum power output is usually determined by maximum voltage and currents, and these in turn specify load impedance, (usually low) the power transistor must be designed to operate with reasonable gain into low impedance loads.

3. The transistor must be designed to safely dissipate the required power under ambient conditions imposed by the user. Developmental transistors have been constructed which have been used in amplifiers delivering good quality audio output above 20 watts, and d.c. control output above 100 watts.

Maximum Output Requirements

For transistor circuitry in which $V_{c \max}$ is the maximum collector voltage (whether this limit is imposed by power supply or transistor specifications), and $I_{c \max}$ is the maximum collector current per transistor, minimum values being negligible, the relations between these quantities and ideal maximum output dissipation and load resistance R_L are summarized in Table I, where the quantity $M = (V_{c \max} I_{c \max})$, representing "maximum output volt amperes", constitutes a figure of merit for transistor output.

Where minimum values of I_c and V_c are not negligible with respect to maximum values, maximum values minus minimum values must be substituted wherever maximum values appear in the table. In any case, the values in the table represent absolute upper limits on output, as they are based on extreme possible variations of voltage and current

across load lines, as illustrated in Fig. 1 and Fig. 2. Note that the load lines and, hence, R_L for maximum output are determined entirely by $V_{c \max}$ and $I_{c \max}$. Hence, R_L is often far below the optimum value for maximum gain. "Impedance matching" for maximum output has no direct relation to transistor parameters when the limiting factors are maximum values of current and/or voltage for a transistor.

It will be noted that Class B push-pull operation provides the same maximum output per transistor as Class A single ended operation, but with less distortion, due to even harmonic cancellation. However, R_L for maximum output is only half as great for push-pull, resulting in lower gain. Even under the most conservative assumptions, Class B push-pull operation requires 3 - 5 times less dissipation per transistor than Class A.

The output limitations of transistors have been discussed here primarily to emphasize the amount of current-voltage range required for reasonable output. For example, a type of transistor for which $V_{c \max} = 50$ volts, $I_{c \max} = 1$ ampere has $M = 50$ volt amperes. Yet the absolute maximum sine wave output per transistor for this type would be $50/8$ or 6.25 watts regardless of use - single ended, push-pull, or paralleled units. The maximum dissipation required (i.e., Class A) for the output current-voltage range of 50 volt amperes would be 12.5 watts. However, as will be seen below, it may be desirable to design for more than 50 watts safe dissipation at room temperature in order to permit operation at a useful dissipation at higher ambient temperatures.

Gain Requirements

Table II illustrates how the usual expressions for power gain are affected as r_e becomes negligible at high I_e .

The small signal power gains at high level operating points as given there are not precisely equivalent to the large signal power gain; however, these expressions are very useful for approximate calculations. Observe that at high levels, the factor $1/(1-\alpha)$ appears even in the grounded base expression. The grounded emitter circuit has more gain than the grounded base circuit, but generally causes much more distortion in large signal use, unless alpha decreases only negligibly at high currents - conditions which normally would require specifications of a smaller maximum current and, hence, would result

TABLE II

Circuit Condition	Low Level Power Gain	High Level Power Gain
Grounded base	$\frac{R_L \alpha^2}{r_e + r_b' (1-\alpha)}$	$\left(\frac{R_L}{r_b'}\right) \alpha^2 \left(\frac{1}{1-\alpha}\right)$
Grounded emitter	$\frac{R_L \alpha^2}{(1-\alpha)(r_e + r_b' (1-\alpha))}$	$\left(\frac{R_L}{r_b'}\right) \alpha^2 \left(\frac{1}{1-\alpha}\right)$
Grounded Collector	$\frac{1}{1-\alpha}$	$\frac{1}{1-\alpha}$

in less maximum output than grounded base operation with the same distortion. The grounded collector circuit appears interesting principally in cases where the load resistance must be much less than r_b' . The grounded base and grounded emitter circuits both require r_b' as low as possible. It has been possible to successfully construct power transistors with r_b' of the order of one ohm and less.

Dissipation Requirements

The basic dissipation limit for all junction devices is set by thermal runaway arising from their characteristic increase of current with junction temperature, the resulting increase in dissipation, which leads to still higher junction temperatures, etc. The condition for thermal stability for junction transistors (namely, that the increase in dissipation due to temperature increase must be less than the corresponding increase in heat flow from the junction) can be written⁽¹⁾

$$V_c(dI_c/dT_j) < 1/\theta \quad \text{Eq. (1)}$$

where T_j is the collector junction temperature, and θ the total thermal resistance, i.e., the junction temperature rise above ambient per watt dissipation. If the thermal increase in collector current is assumed to be independent of total current, and if the rate of increase of junction current with temperature is assumed to be approximately 7% per degree, the stability condition for low impedance circuits can be written:

$$\text{for grounded base operation } V_c < \frac{15}{\theta I_{co}} \quad \text{Eq. (2)}$$

and if the current multiplication for low impedance grounded emitter operation is relatively independent of current:

for grounded emitter operation

$$V_c < \left(\frac{1-\alpha}{\alpha}\right) \left(\frac{15}{\theta I_{co}}\right) \quad \text{Eq. (3)}$$

High impedance circuits are more stable to the extent that increased collector current re-

sults in decreased collector voltage. The grounded emitter circuit is seen to be much less stable than grounded base for simple low impedance circuits. By special bias circuits, however, it is possible to improve the grounded emitter stability to some extent, particularly for small signal amplification. However, suitable stabilization circuitry presents major problems in practical applications of large signal grounded emitter circuits.

The dotted lines in Fig. 2(a) show the locus of thermal instability in a transistor. Note that at higher voltages the dissipation must be smaller (so that T_j and, hence, I_{co} will be lower) for stability, i.e., for a given collector dissipation, a high voltage-low current operating point is closer to thermal runaway than a high current-low voltage point. The emitter dissipation is larger for high currents, resulting in an effect, usually small, in the opposite direction. This is not in strict agreement with the widely accepted belief that thermal breakdown in transistors is a simple function of dissipation only⁽²⁾; however, this discrepancy is ordinarily well within the bounds of the safety factors which must be applied in practice. The specification of a conservative maximum junction temperature, as shown in the figure, corresponds to the specification of maximum dissipation at any one ambient temperature and follows the actual shape of the breakdown characteristic fairly well. Such specification of maximum permissible junction temperature maintains this relationship independent of ambient temperature, as illustrated in Fig. 2(b) and 2(c). There is less safety factor at the higher voltages, but this, of course, is the fundamental reason for the existence of practical independent specifications for maximum collector voltage. Whenever closer ratings are desired, it may be desirable to establish higher values of V_c for lower ambient temperatures. This rating system (first applied to the General Electric 2N43-44-45 transistors by A. F. Perkins⁽³⁾) follows the actual thermal behavior of transistors, and is simple to apply, using the relation:

$$T_j = T_{amb} + \theta W \quad \text{Eq. (4)}$$

The transistor manufacturer specifies value of $T_{j \max}$ and the thermal resistance of the transistor. The user substitutes a required ambient temperature and obtains the maximum permissible dissipation for his conditions. The

(1) The author acknowledges valuable discussion on this point with Dr. Johannes S. Schaffner, who has independently obtained a relation similar to Eq. (1).

(2) On the other hand, Eq. (1) applied to the inverse-biased junction diode yields the simple dissipation condition $V_b I_b < 15/\theta$.

(3) C.H. Zierdt, "Hermetically Sealed Medium Power Fused Junction Transistors", PG&E TRANSACTIONS of the IRE, Vol II, Feb. 1954.

total thermal resistance, of course, is composite but usually can be roughly divided into two parts:

$$\theta = \theta_i + \theta_o \quad \text{Eq. (5)}$$

The internal thermal resistance θ_i is the temperature rise from transistor case to collector junction per unit dissipation. This is built into the transistor. The external thermal resistance θ_o is the case temperature rise above ambient per unit dissipation. The manufacturer can evaluate this quantity approximately for free convection, for example, but the user may bury the transistor in a poorly ventilated location, resulting in higher θ_o , or at the other extreme, might locate the transistor in a circulated oil bath, or other "infinite" sink, and reduce θ_o practically to zero. The external thermal resistance, therefore, depends almost entirely upon the user. For the General Electric 2N43-44-45 transistor, for example, $T_j \text{ max} = 100^\circ \text{ C}$, $\theta_i = 0.2^\circ \text{ C/mw}$, $\theta_o = 0.3^\circ \text{ C/mw}$ (free air convection). The values of θ_i and θ_o are extremely important to the user. The rating system expressed by Eqs. (3) and (4) can be expressed graphically as shown in Fig. 3.

The maximum safe junction temperature is often in the range $100^\circ - 150^\circ \text{ C}$. It is, therefore, apparent that dissipation ratings at 2 watts and above require total thermal resistance below 50 degrees per watt, even for use at moderately low ambient temperatures. The region of smallest thermally conducting area, which often accounts for most of the thermal resistance, is that closest to the collector junction. For this reason, the internal thermal resistance is one important criterion of the thermal design of a transistor - the lower its value, the less external cooling must be provided by the user. Low total thermal resistance can be achieved only by careful design since, for example, even a heavy copper strip can easily account for 10 degrees rise per watt, and the heaviest gauge aluminum chassis is far from being an infinite heat sink. In all cases, the total thermal resistance must be included in the transistor rating, just as the total weight of conductors, fins, and any other cooling equipment must be included in the total weight of the system.

Experimental Results

Developmental high power p-n-p alloy diffusion junction transistors have been constructed using a large construction designed for minimum decrease in alpha at high currents and for low base resistance. Characteristics obtained by V. P. Mathis are shown in Figs. 4, 5, and 6. Units of this type have operated continuously above 75 watts in cases having very low thermal resistance. The permissible dissipation in practice depends, of course, upon the thermal resistance of the case. For example, in an ordinary medium size (6AG7) metal tube bulb, mounted on an open chassis ($\theta_o \approx 12^\circ \text{ C/w}$) the room temperature continuous rating is approximately 5 watts. The internal sealed-in liquid-vapor cooled construction (having θ_i less than 5° C/w) provides sufficient heat

capacity, however, that more than 50 watts dissipation can be maintained safely for intervals of a half minute or so, greatly increasing the control switching capacity. Correspondingly higher ratings are obtained using larger cases or forced ventilation. An external heat sink clamped to any portion of the external case is effective in raising the dissipation rating still further.

The available gain at high currents into matched impedance loads is more than 30 db, corresponding to $r_c > 5,000 \text{ ohms}$ and $r_b < 2 \text{ ohms}$. However, as pointed out above, maximum output requires a load resistance usually of the order of 100 ohms, much lower than the match for maximum gain, so that practical gains at maximum output are more commonly in the range 10 - 20 db, and are almost independent of collector resistance.

The operating characteristics of a practical push-pull amplifier using these units are listed in Table III:

TABLE III

MEASURED PERFORMANCE DATA	
CLASS B PUSH-PULL GROUNDED BASE POWER AMPLIFIER	
Power Output	: 23 watts
Power Supply Voltage	: 45 volts
Load Impedance (seen by each transistor)	: 40 ohms
Power Input from Driver	: 2.3 watts
Collector Efficiency	: 60%
Collector Dissipation at Full Output (per transistor)	: 7 watts
Total Distortion at Full Output	: 3%

Although the dissipation at full sine wave output is more than 7 watts, the 6AG7 cased transistors are ample for voice or music application since the dissipation averaged over any half minute period is usually less than 2 watts even in fortissimo passages. The amplifier chassis and its current are depicted in Figs. 7 and 8.

The maximum d.c. switching output is over 100 watts per transistor, based upon $V_c \text{ max} = 75\text{V}$, $I_c \text{ max} = 2 \text{ A}$, as implied by Table I. The peak dissipation required for 100 watts peak d.c. output is approximately 25 watts, well within the intermittent ratings noted above.

These results do not represent ultimate obtainable performance - merely an example of progress to date. Although the kilowatt transistor is not as yet a reality, these early results justify our confidence that transistors will eventually be designed to handle the entire range of reliable high power operation required by electronic communication and control.

ACKNOWLEDGEMENT

This work has been supported by the U.S.A.F. Air Research and Development Command, U.S.A.F. Air

Material Command, Army Signal Corps, and Navy Bureau of Ships, under Contract AF33(600)-17793.

SUMMARY OF ABSOLUTE MAXIMUM OUTPUT RELATIONS

	D C "SWITCH" CONTROL	CLASS A SINE WAVE-SINGLE	CLASS B SINE WAVE PUSH-PULL
SUPPLY VOLTAGE REQUIRED	$V_c \text{ max}$	$1/2 V_c \text{ max}$	$1/2 V_c \text{ max}$
IDEAL MAX. POWER OUTPUT (TOTAL)	M	$1/8 M$	$1/4 M$
POWER OUTPUT (PER TRANSISTOR)	M	$1/8 M$	$1/8 M$
COLLECTOR DISSIPATION PER TRANSISTOR	$1/4 M$ (MAXIMUM)	$1/4 M$ MAX (QUESCENT) OR $.035 M$ IDEAL AVG (76% EFF) OR $1/8 M$ IDEAL PEAK OR $.08 M$ PRACTICAL AVG. (60% EFF)	
LOAD RESISTANCE R_L FOR MAX OUTPUT	$V_c \text{ max} / I_c \text{ max}$	$V_c \text{ max} / I_c \text{ max}$	$1/2 (V_c \text{ max} / I_c \text{ max})$

$M = (V_c \text{ max}) (I_c \text{ max})$ VOLT AMPERES

Table I

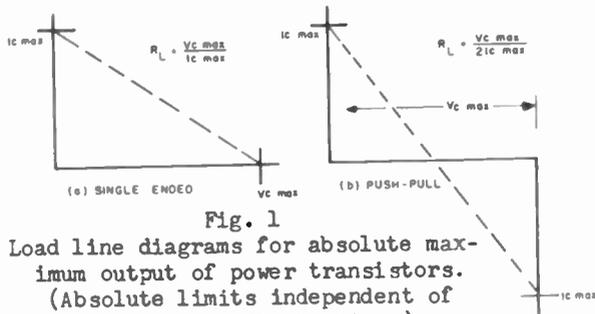


Fig. 1

Load line diagrams for absolute maximum output of power transistors. (Absolute limits independent of circuit or dissipation.)

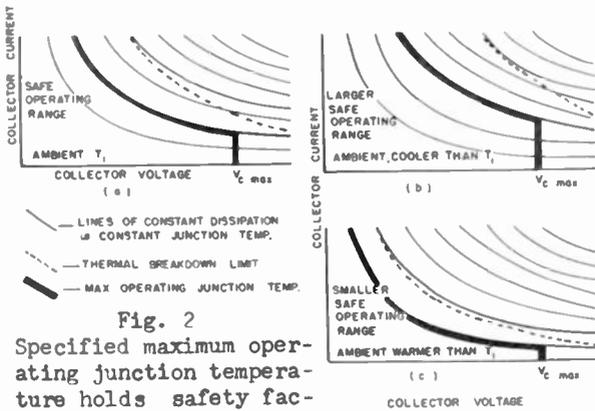


Fig. 2

Specified maximum operating junction temperature holds safety factor approximately constant over wide ambient temperature range.

GRAPHICAL TRANSISTOR POWER RATINGS

EQUIVALENT TO $T_j = T + \theta D$

WHERE

$T_j \text{ max} = 100^\circ\text{C}$

$\theta = 0.2^\circ\text{C}/\text{mas}$

$\theta_0 = 0.3^\circ\text{C}/\text{mas}$

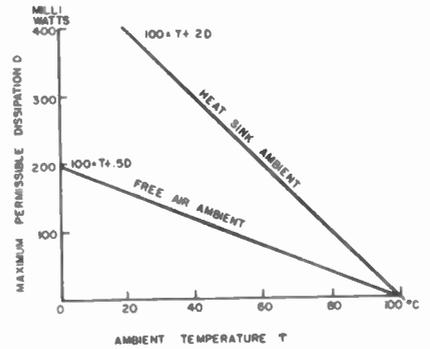


Fig. 3

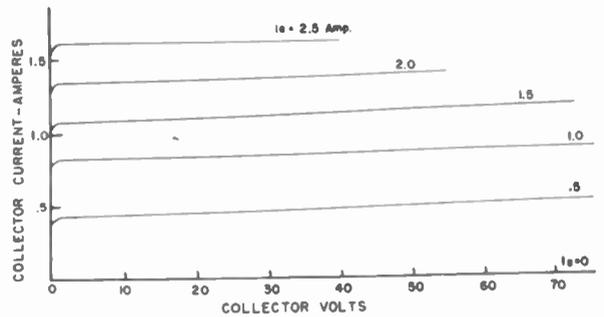


Fig. 4

Typical collector characteristics - large area power transistor (pnp).

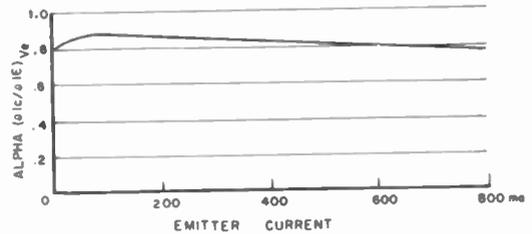


Fig. 5

Alpha vs. emitter current - large area power transistor (pnp).

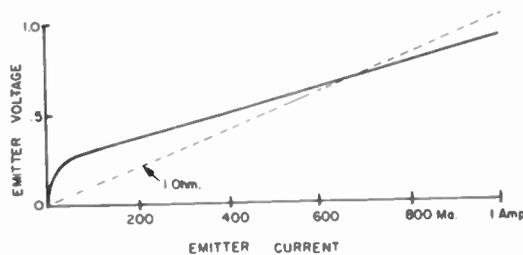


Fig. 6

Emitter characteristic - large area power transistor (pnp). High current base resistance < 1 Ohm.

A NEW HIGH TEMPERATURE SILICON DIODE

by
C. G. Thornton and L. D. Hanley
Sylvania Electric Products, Inc.
Electronics Division
Ipswich, Massachusetts

Summary

A method of preparing high back voltage silicon by bombarding the surface with oxygen has been devised and a new type of point contact diode prepared. The diode is characterized by: (1) very low saturation currents of the order of 1 microampere, (2) high inverse operating voltages (70-200v), (3) very low barrier capacitance ($<.3\mu\text{mf}$) and rapid recovery times, and (4) operation at elevated temperatures up to 200°C.

Introduction

Microwave mixer crystals for radar and video applications have been satisfactorily produced from heavily doped silicon for several years, however, most attempts to produce high back voltage point contact silicon rectifiers by similar techniques have resulted in units which were far inferior to germanium point contact rectifiers.

This paper describes the preparation and properties of a new type of high back voltage point contact diode. The diode is characterized by: (1) very low saturation currents of the order of 1 microampere, (2) high inverse operating voltages (70-200v), (3) very low barrier capacitance ($<.3\mu\text{mf}$) and rapid recovery times, and (4) operation at elevated temperatures up to 200°C. The diode is produced by a surface bombardment technique which produces the desired rectification properties and leaves the silicon surface with a thin protective layer.

The use of silicon in high back voltage rectifiers is predicated on the fact that it will withstand a considerably higher temperature rise than germanium without losing its rectification properties. Whereas the upper limit of satisfactory operation for germanium rectifiers is reached in the range of 75 to 100°C, silicon junction rectifiers have been constructed which operate at temperatures up to 300°C without suffering a serious loss of rectification efficiency.

Preparation

The technique to be reported here was developed, after an investigation of the ionic bombardment procedure described by R. Ohl¹ of Bell Telephone Laboratories in 1952. The apparatus used is shown schematically in Figure 1. The filament, grid, and heated support for the silicon are all enclosed in a suitable high vacuum system to which a given gas is introduced to a predetermined pressure through a controlled

leak. When the system is energized, electrons from the heated filament will pass through the grid structure into the space above the silicon sample where collisions with molecules of the ambient gas will take place resulting in ion formation. These ions are accelerated through a potential field of several kilovolts and uniformly bombard the silicon surface.

Several different gases were used in the initial investigation including helium, argon, nitrogen, and hydrogen. Suitable steps were taken to obtain a good ($<10^{-5}$ mm.) vacuum in the system, and the purity of the incoming gas was carefully assured. Under these conditions, only a very slight change could be detected with bombardment. It was decided to try oxygen as a bombarding gas in contrast to the inert or reducing gases mentioned above. The results are shown in Figure 2 as obtained first with helium and then oxygen on the same starting material. The surprising result was obtained that the peak inverse voltage could be increased more than an order of magnitude by this treatment when oxygen was used as compared with helium with reverse resistances as high as 10 megohms up to -40 v. It was discovered that the process was readily reproducible and that similar results could be obtained with both high and low resistivity silicon. The chief difference between the two resistivities is the time required to complete the bombardment, and with lower resistivity silicon a lower forward resistance is obtained. On the basis of this observation, it was concluded that the reverse characteristic is relatively independent of the carrier concentration in the silicon substrate.

In view of the extremely large increase in the peak inverse voltage, it is believed that the effect of the bombardment is to reduce the effective carrier concentration at the surface through the introduction of new trapping levels in the forbidden region. Possible mechanisms include: (1) actual bombardment damage as described by Lark-Horovitz² for bulk semiconductors tending to shift the Fermi level toward the center of the forbidden region, and (2) an effective change in the chemical composition (formation of an oxide or an adsorbed oxygen layer) at the surface which could function as a trapping site. Additional experiments are presently being carried out which should shed further light on this aspect of the problem.

A typical point contact rectification curve for oxygen bombarded silicon is shown in Figure 3. The initial material was about .5 ohm cm. p-type silicon with a peak inverse voltage of about 10v. Although the forward resistance of the diode is

increased slightly by the treatment, the increase in resistance in the reverse direction is considerably greater causing an increase in rectification ratio at 5v from 60 to 10^5 .

An additional advantage accrues to the fact that after oxygen bombardment the surface is found to be extremely inert with respect to the effects of humidity and other surface contaminants. This is attributed to the possible formation of a thin oxide protective layer which appears to passivate the silicon surface. It was found early in the investigation that the effects of the bombardment could be annealed out at temperatures above 300°C and, therefore, the time during which effective bombardment can be carried on in this temperature range is limited to the time required for equilibrium to be established between the annealing and bombarding processes. At temperatures below 250°C the bombarded surfaces are found to be completely stable. For this reason lower temperatures are presently used and the bombardment can be carried on to advantage for several hours at the rate of 5 microamperes per square centimeter.

In the actual process of preparing the diode, single crystals of boron doped silicon having the appropriate resistivity are used. These ingots such as shown in Figure 4 are sliced to obtain wafers about 20 mils in thickness which are then brought to a high metallographic polish on the sides to be bombarded. The reverse sides of the slices are plated to obtain a good ohmic contact. These slices are then diced and assembled into diodes as shown in Figure 5. Both glass and ceramic type packages are shown. The latter is useful where it is desirable to avoid photoelectric effects.

Applications

The rectification characteristics of the finished diode and the effect of temperature up to 150°C is shown in Figure 6. The changes in diode characteristics with temperature should not be interpreted as resulting entirely from changes in the silicon surface as there is unavoidable expansion in the package at elevated temperatures.

It is to be noted that the diode adequately covers a range of operating temperatures where germanium has lost its useful semiconductor properties. Such diodes have been operated at temperatures of 200° or higher without serious changes in the rectification properties.

In view of the recent development of several types of silicon junction rectifiers, it is useful to compare their electrical properties and applications. The capacitance of silicon junction rectifiers prepared in this laboratory are the order of 10^{-10}farad compared with measured values of less than $.3\mu\text{farad}$ for diodes of the point contact variety. The frequency response of the two types of rectifiers are shown in Figure 7. The frequency cutoff for the bombarded point contact diodes is approximately 400 megacycles and could probably be extended further by a special package design of the type used in microwave circuits. This may be compared with a frequency cutoff of from .1 to 15 megacycles for junction units.

In many computer type circuits the recovery time of a switching diode is of considerable importance. The recovery time may be specified as that time required to regain 90% of the initial reverse resistance after switching from the heavy forward current obtained with a forward voltage to a voltage in the reverse direction. Characteristics for germanium and silicon point contact rectifiers are shown in Figure 8. The recovery time obtained for germanium point contacts varies from .5 microseconds in a very favorable case to as long as 5 microseconds. For the silicon diodes described in this paper the measured recovery time averages less than .1 microsecond. The diodes also have very marked photoelectric effects and may find considerable application in the field of photoelectric devices.

References

1. R. Ohl, Properties of Ionic Bombarded Silicon, Bell Technical Journal, January 1952, pp. 104-21.
2. K. Lark-Horovitz, Semiconducting Materials, Butterworths Scientific Publications, Ltd., London 1951.

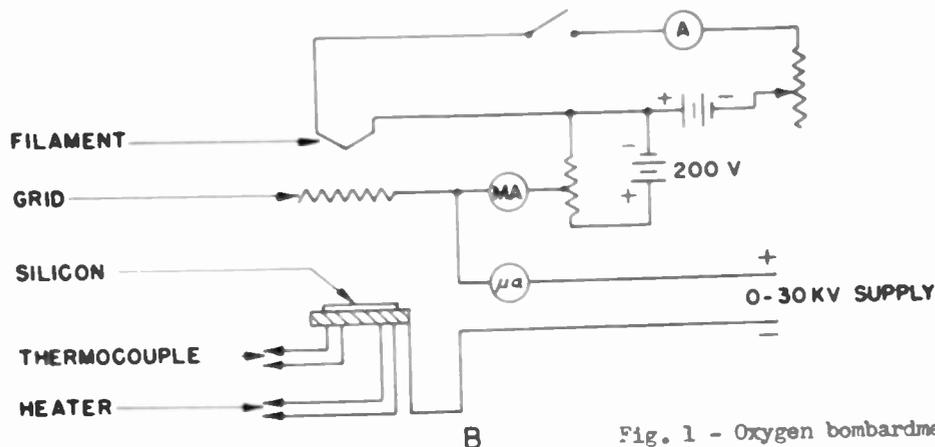


Fig. 1 - Oxygen bombardment apparatus.

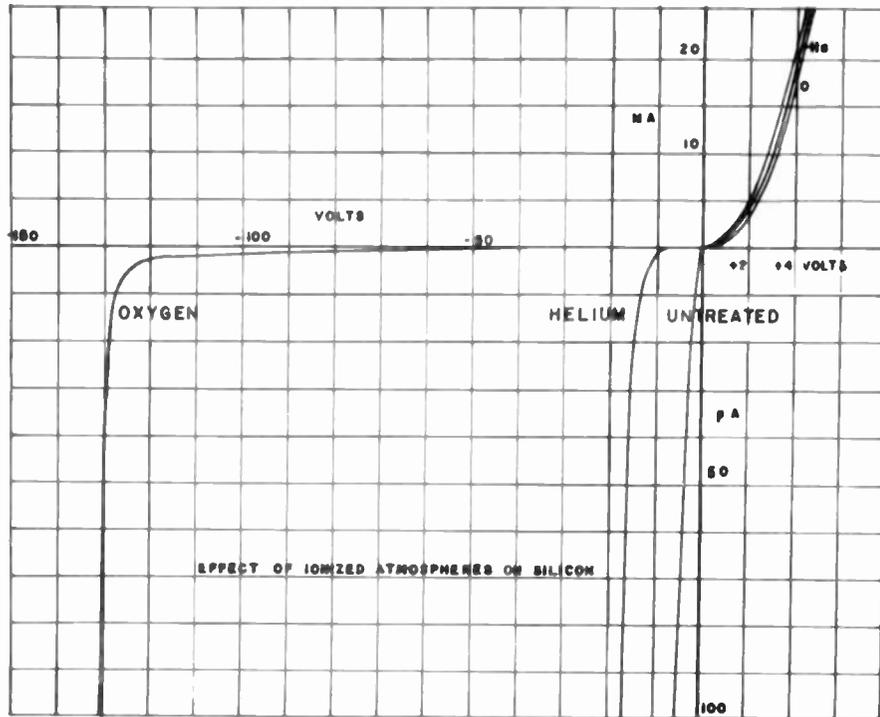


Fig. 2

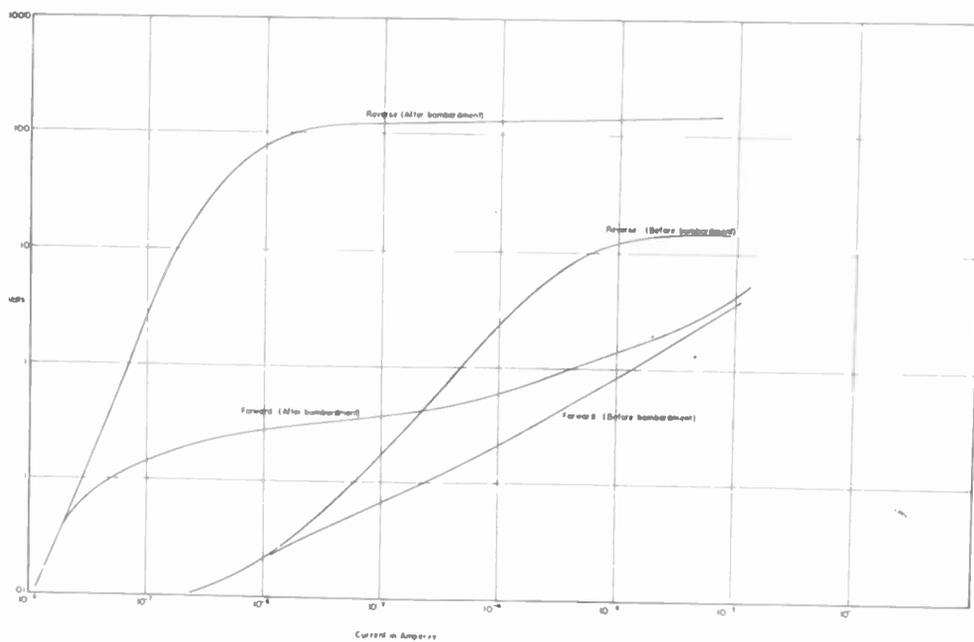


Fig. 3



Fig. 4

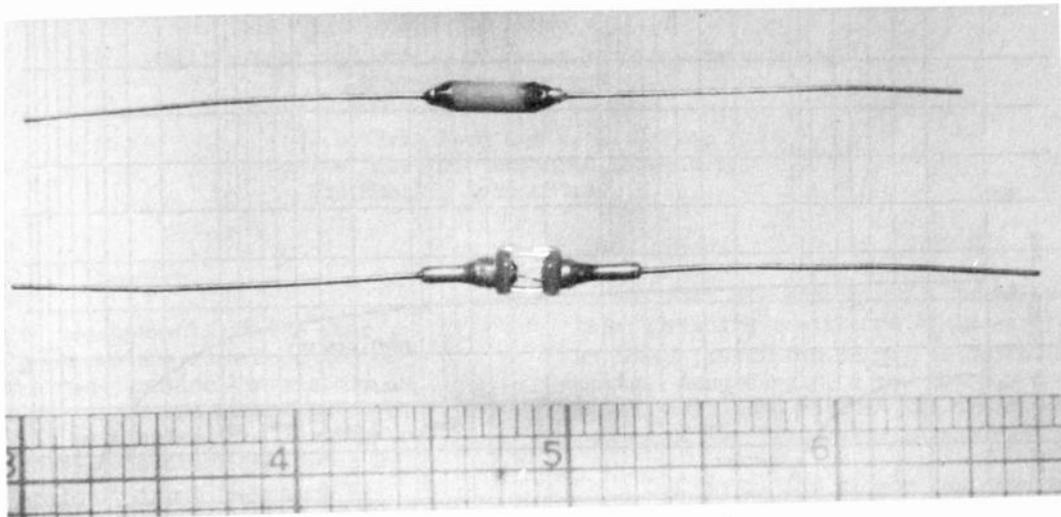


Fig. 5

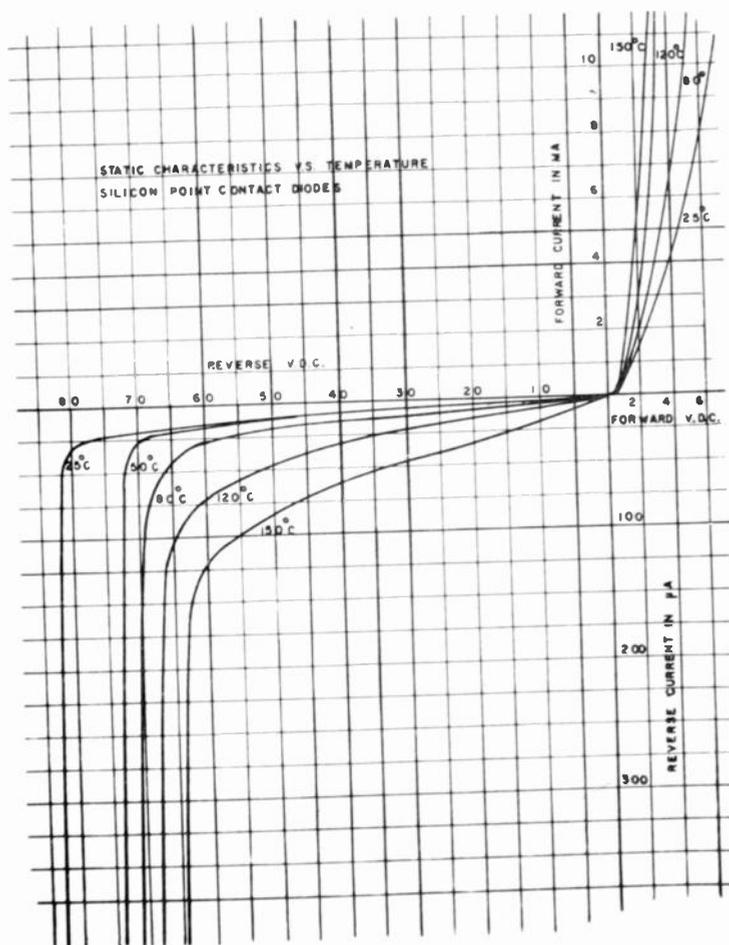


Fig. 6

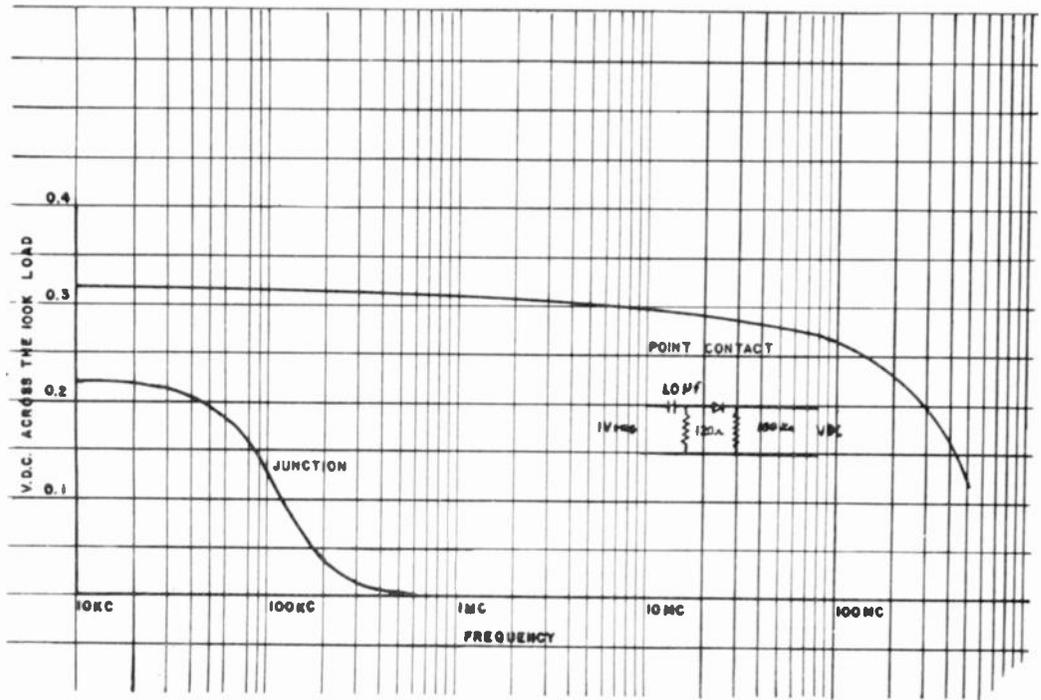
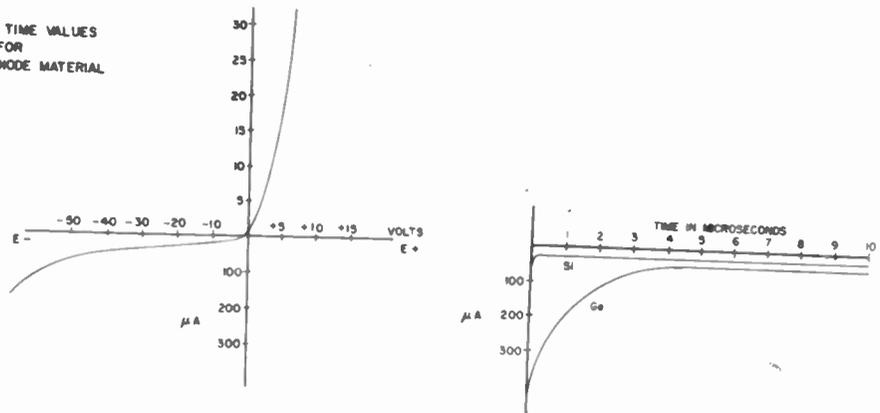


Fig. 7

RECOVERY TIME VALUES FOR DIFFERENT DIODE MATERIAL



DIODE TYPE	POINT CONTACT	SWITCHING POINT CONTACT	HIGH BACK VOLTAGE POINT CONTACT
MATERIAL	Ge	Ge	Si
RECOVERY TIME RANGE	1-5 μ sec	3-2 μ sec	0.5-5 μ sec

Fig. 8

SMALL-SIGNAL PARAMETERS OF GROWN-JUNCTION TRANSISTORS AT HIGH FREQUENCIES*

R. L. Pritchard and W. N. Coffey
General Electric Research Laboratory
The Knolls, Schenectady, N.Y.

Summary

Results of measurements of the four small-signal h parameters as a function of frequency for grown-junction transistors in grounded-base operation are described in some detail. In particular, the results for open-circuit voltage-feedback parameter h_{12} and for short-circuit input impedance h_{11} at high frequencies indicate that the grown-junction transistor can not be represented by the usual theoretical model of an ideal one-dimensional transistor in series with a lumped constant base spreading resistance r_b' . A theoretical analysis of a new model is described in which account is taken of the distributed nature of the transistor parameters and of the spreading resistance of the base region. Under simplifying assumptions, results of the analysis are in fairly good agreement with experimental results for the grown-junction transistor. A new approximate equivalent circuit based on this analysis also is described, in which the lumped constant base spreading resistance is replaced by a complex base impedance whose magnitude decreases with increasing frequency.

Introduction

During the past year, there has been a great deal of interest in the subject of the frequency variation of junction-transistor parameters. Several writers^{1,2,3} have calculated independently a theoretical variation for a model consisting of an ideal one-dimensional transistor with an added external base spreading resistance. Several equivalent circuits based on these analyses have been described.^{1-3,4,5} Furthermore, experimental measurements of transistor parameters for fused-junction (or alloy) transistors appear to be in reasonably good agreement with results predicted from this theoretical model.⁶

The purpose of the present paper is to describe experimental results of the measurement of small-signal parameters for a number of grown-junction transistors. These results

showed that the concept of a lumped constant base spreading resistance r_b' in series with an ideal transistor is not valid for these units. Accordingly, a new model was devised which takes into account the distributed nature of the transistor parameters.

The parameters chosen for measurement were the series-parallel or h parameters, defined by the equations

$$\left. \begin{aligned} e_1 &= h_{11}i_1 + h_{12}e_2 \\ i_2 &= h_{21}i_1 + h_{22}e_2 \end{aligned} \right\} \quad (1)$$

A grounded-base configuration of the transistor was employed. In general, these parameters are easily measured for a junction transistor with its low impedance and high output impedance. For convenience at the higher frequencies, the short-circuit output admittance $y_{22} = h_{22} - h_{12}h_{21}/h_{11}$ was measured in place of h_{12} .

Experimental Measurements and Results

Method of Measurement

The short-circuit input impedance h_{11} was measured by means of a General-Radio type 916-A R.F. impedance bridge. On the other hand, open- and short-circuit output admittances h_{22} and y_{22} were measured on the General-Radio type 821-A Twin-T impedance measuring circuit. For the measurement of h_{11} , d-c bias to the emitter was supplied through a high-impedance shunt feed system, and the measured values of impedance then were corrected for the a-c shunting effect of the d-c feed system as necessary. The short-circuit current amplification factor h_{21} ($= -a$) was measured in amplitude and phase with a heterodyne system devised and built by one of the authors (WNC). Amplitude and phase measurements were made at a fixed frequency of 10 kc/s with a variable carrier frequency to correspond to the desired signal frequency. Finally, at low frequencies the magnitude of h_{12} was

*This work was supported by the Air Material Command, Signal Corps and Bureau of Ships under contract AF33(600)17793.

measured with a simple transmission system employing a high impedance tuned circuit to effectively open circuit the emitter-base circuit of the transistor.

Results Expected for Old Theoretical Model

Before describing the results of the measurements, it might be well to review briefly the results that might be expected for the ideal one-dimensional theoretical model plus a constant external base spreading resistance, as shown in Fig. 1a.

The frequency variation of the current-amplification factor α is well known⁷ and needs little comment. In general, α decreases in magnitude very much like the response of a simple low-pass RC circuit. However, the phase shift associated with α generally is somewhat greater than that obtained with the simple RC circuit, e.g., the phase shift of α may exceed 90° . The frequency for which $|\alpha|$ has decreased to 0.707 times its low-frequency value α_0 is defined as the α -cutoff frequency.

If the collector-base diffusion capacity due to space-charge layer widening⁴ is negligible, as it generally is in a grown-junction transistor at moderate values of d-c emitter current, then the collector-base circuit for an open-circuit emitter consists essentially of the collector-barrier capacity C_c in series with the base spreading resistance r_b' . Consequently, the output capacity C_{22} at low frequencies is constant and is equal to C_c . However, C_{22} decreases slowly at high frequencies when r_b' becomes comparable with $(1/\omega C_c)$. On the other hand, the output conductance g_{22} increases as the square of the frequency from a constant low-frequency value to an ultimate high-frequency value of $1/r_b'$.

The open-circuit voltage feedback parameter h_{12} also can be determined from this simple picture of C_c and r_b' in series as the voltage drop across r_b' for a constant unit a-c voltage applied between collector and base. However, at low frequencies, h_{12} has a limiting value determined by an inherent voltage-feedback mechanism⁸ and by collector conductance, neither of which is indicated in the simple circuit of Fig. 1a. As frequency is increased, the open-circuit emitter voltage, and hence h_{12} , increases linearly with frequency, as shown in Fig. 1b.

Ultimately, at high frequencies, h_{12} reaches a limiting value of unity.

Finally, the short-circuit input impedance h_{11} consists of the impedance of the forward-biased emitter-base junction in series with the impedance $r_b'(1-\alpha)$. At moderate d-c emitter currents, the former impedance normally is negligible with respect to the latter. Since the phase shift associated with α is negative, i.e., lagging, up to approximately five times the α -cutoff frequency, the term $(1-\alpha)$ has a positive reactive part. Hence, the impedance $r_b'(1-\alpha)$ should behave as an inductance in parallel with a resistance for frequencies well beyond the α -cutoff frequency, as shown in Fig. 1b.

Results of Measurements on Grown-Junction Transistors

Measurements of the four h parameters were made on a large number of grown-junction transistors having many different types of internal properties. Consequently, there was no one set of results that could be called truly typical. However, certain types of behavior were found to be common—at least qualitatively—to most transistors. Accordingly, the experimentally observed variation for each h parameter is shown below for one particular transistor. The relationships between results for this transistor and corresponding results for other transistors are discussed briefly. It should be emphasized that whereas there was considerable variation among transistors in the magnitudes of the parameters, the relative variation with frequency was common to many transistors.

The behavior of $-h_{21} = \alpha$ is shown in magnitude and phase in Fig. 2 for two values of d-c emitter current I_c . For either value of I_c , the shape of both amplitude and phase curves is in excellent agreement with theoretical curves for α of a one-dimensional transistor. The dependence of the α -frequency curves upon I_c is relatively small and can be considered a second-order effect. Many of the other transistors that were measured also displayed this type of frequency variation. However, an equally large number of transistors displayed a distinctly different type of α -frequency behavior. In these cases, the phase shift of α was considerably

less than the theoretical value and exhibited a maximum with respect to frequency. Furthermore, in such cases, the α -frequency variation generally was quite dependent upon d-c bias. This type of variation will be discussed below in somewhat more detail.

In Fig. 3 is shown the frequency variation of the open-circuit output conductance g_{22} and capacitance C_{22} . The capacity C_{22} is essentially constant with frequency, although a slight decrease is noticeable at higher frequencies. Note that C_{22} is somewhat dependent on d-c emitter current I_E ; this indicates that collector-base diffusion admittance is not negligible for this transistor. Consequently, the open-circuit output conductance also is quite dependent upon I_E . An equally strong dependence of g_{22} upon d-c collector voltage E_C for a fixed value of $I_E = -1.0$ ma also is shown by the curves in Fig. 3. This is to be expected as a result of the well-known dependence of C_{22} upon E_C and since g_{22} and C_{22} are directly related for a collector-base circuit that is approximated by a collector capacitance in series with a base spreading resistance r'_b . Over the lower range of frequencies shown, g_{22} varies approximately as the 1.6 power of the frequency. This rate of frequency variation is in fair agreement with that predicted for the simple C_c - r'_b series representation for the collector-base circuit. However, in general, quantitative agreement between measured g_{22} and g_{22} calculated in terms of r'_b is not very good.

Also shown in Fig. 3 is the short-circuit output conductance g_{22}^s . In general, in the frequency range near the α -cutoff frequency g_{22}^s is more or less constant as shown. Note that g_{22} approaches g_{22}^s at the higher frequencies. On the other hand, at lower frequencies, g_{22} must decrease to a considerably smaller value, e.g., the order of 10μ mho.

All of the other transistors measured showed a similar variation of h_{22} with frequency. In some cases, C_{22} and g_{22} were quite independent of d-c emitter current so that collector-base diffusion capacity could

be assumed truly negligible. For many transistors with relatively high values of C_c , a fairly appreciable decrease of C_{22} was observed at increasing frequencies. Again, this would be in agreement with the simple series C_c - r'_b model.

The variation of the magnitude of the open-circuit voltage-feedback parameter h_{12} over a wide range of frequencies and the variation of the phase of h_{12} over a smaller range is shown in Fig. 4. A linear variation with frequency that would be associated with constant base spreading resistance r'_b is shown by the dotted curves. Note that the departure of $|h_{12}|$ from a linear frequency variation occurs at values of $|h_{12}|$ that are very much less than unity.

All of the other grown-junction transistors that were measured showed a similar departure of $|h_{12}|$ from a linear frequency variation at frequencies of the order of the α -cutoff frequency. In general at these higher frequencies, $|h_{12}|$ varies as the 0.4 to 0.7 power of the frequency. Also, the phase shift of h_{12} at these frequencies generally is of the order of 30° to 50° .

It should be quite obvious that these results can not be explained by the simple picture of a constant base spreading resistance r'_b in series with collector capacity. To explain the experimental results, it would be necessary to replace r'_b by a complex base impedance having a magnitude that decreases with frequency.

Finally, the frequency variation of the short-circuit input resistance and reactance r_{11} and x_{11} respectively is shown in Fig. 5 for two values of d-c emitter current I_E . For both values of I_E , the resistance r_{11} increases with increasing frequency until a maximum is attained at about 4-5 Mc/s. On the other hand, the reactance x_{11} attains a maximum at a lower frequency, decreases through 0 at approximately the frequency for which r_{11} attains its maximum, and becomes distinctly capacitive reactive at higher frequencies. The similarity between the curves of Fig. 5 and corresponding curves for the driving-point impedance of a parallel RLC circuit is quite striking and suggests a sort of resonance in the input circuit of the transistor.

This general behavior for h_{11} was typical of all grown-junction transistors measured. Furthermore, with transistors having a fairly wide range of "resonance" frequencies, the resonance frequency always occurred at approximately 0.8 to 1.0 times the α -cutoff frequency. It should be emphasized that the observed capacitive reactance x_{11} can not be explained with the model of the ideal one-dimensional transistor plus constant r_b^i . In particular, for a d-c emitter current $I_e = -1.0$ ma, the impedance of the forward-biased emitter-base junction is of the order of 20 ohms at the α -cutoff frequency. Hence, in that model, h_{11} is determined solely by $r_b^i(1-\alpha)$, which should be inductive reactive for frequencies well beyond the α -cutoff frequency. Consequently, it should be obvious from these results for h_{11} , as well as those for h_{12} , that the concept of a constant base spreading resistance r_b^i can not be applied to grown-junction transistors.

New Theoretical Distributed Model

Description of Model

In an attempt to explain the behavior of the experimental results obtained, an analysis was made of a new theoretical model which takes into account the distributed nature of the transistor parameters and of the base spreading resistance. A sketch of this distributed model is shown in Fig. 6a. Electrical contact is assumed to be made uniformly over the upper and lower surfaces of the base region and uniformly over the cross-section areas of emitter and collector regions, i.e., these surfaces are assumed to be equipotential.

For purposes of analysis, this two-dimensional problem is separated into two one-dimensional problems. Thus, diffusion current is assumed to flow only in the direction (y) perpendicular to emitter and collector barriers. Alternatively, the model can be considered as comprising an infinite number of elementary transistors having a common emitter and common collector terminal but with successive base terminals connected by elementary resistances r as shown in Fig. 6b.

At any value of transverse distance x the a-c voltage-current relations of the appropriate elementary transistor can be written simply as

$$\begin{aligned} \Delta i_e &= \left[(y_{11}^i + j\omega c_e^i) e_e + y_{12}^i e_c \right] \Delta x \\ \Delta i_c &= \left[y_{21}^i e_e + (y_{22}^i + j\omega c_c^i) \right] \Delta x \end{aligned} \quad (2)$$

The parameters y_{ij}^i are diffusion admittances per unit length for the elementary transistor, while c_e^i and c_c^i are, respectively, emitter-base and collector-base barrier capacitances per unit length. Each of the diffusion admittances can be expressed in terms of physical parameters such as base thickness, carrier concentration, etc. In general, each y_{ij}^i is complex and frequency dependent.

By solving the differential equations (2) subject to appropriate boundary conditions, as in the case of the conventional transmission line, a-c terminal voltage-current relations can be calculated for the over-all distributed model. From these relations, the small-signal parameters of the distributed model can be determined readily. Unfortunately, the diffusion-admittance parameters y_{ij}^i in general are complicated functions of transverse distance x , for the following reason. The d-c base current, which exists in any practical transistor, flowing through the distributed resistance r of the base region produces a transverse voltage drop across the base. Hence, with a constant voltage applied between emitter and base terminals, the voltage appearing between emitter and base of any elementary transistor depends upon the position (x) of that transistor. Therefore, the d-c current per unit area for an elementary transistor, which depends in an exponential manner upon emitter-base voltage, may vary considerably with distance x . Consequently, the value of each diffusion admittance parameter y_{ij}^i , which is directly proportional to d-c current per unit area, may be quite dependent upon position x .

Note that as terminal d-c emitter-base voltage is increased, the d-c base current will increase, and hence the d-c current distribution through the base becomes increasingly non-uniform. In particular, an increasing amount of d-c current between emitter and collector tends to flow through the region of the base near the base terminal contact. Thus, a junction-transistor triode effectively behaves as an internally biased tetrode transistor⁹ as d-c emitter current is increased. Experimental verification of this effect is

given below.

If the emitter-base voltage distribution is non-uniform, the diffusion-admittance parameters y'_{ij} are functions of transverse distance, and a simple solution for the terminal voltage-current relationships is not possible. However, if it is assumed that the emitter-base voltage distribution is essentially uniform, the parameters y'_{ij} may be considered constant, and a relatively simple solution may be obtained for the small-signal parameters of the distributed model. Due to space limitations, only a brief description of the results will be given here.

As might be expected from consideration of the related RC transmission line problem, the small-signal parameters of the over-all theoretical distributed model each involve hyperbolic functions of a complex argument Γh , where h is the base-base distance and Γ is a sort of propagation constant for the three-terminal transmission line. In particular,

$$\Gamma h = \left\{ R_b \left[y'_{11} + y'_{12} + y'_{21} + y'_{22} + j\omega(c'_e + c'_o) \right] h \right\}^{1/2} \quad (3)$$

where $R_b = rh$ is the total d-c base-base resistance. The argument Γh is complex and increases in magnitude with increasing frequency.

Low-Frequency Results

In general, the solutions for the small-signal parameters of the distributed model can not be described in simple terms. However, at fairly low frequencies when the argument Γh is not too large, series expansions of the hyperbolic functions may be employed. In this case, approximate solutions for the terminal small-signal parameters are obtained for the distributed model which are of exactly the same form as the equations for the simple one-dimensional theoretical model of a transistor plus an external base spreading resistance r'_b .

Furthermore, in this case the base spreading resistance r'_b can be simply related to the d-c base-base resistance R_b ; viz.,

$$r'_b = R_b/3. \quad (4)$$

Hence, it turns out that the concept of a constant base spreading resistance r'_b is

valid at low frequencies for a grown-junction transistor. (However, it should be emphasized that this low-frequency r'_b is not the same as the low-frequency value of the equivalent-Tee base resistance r_b ; see reference 8; J. M. Early.)

Two other items of interest may be calculated from the low-frequency expansions. When the emitter terminal is open circuited and a voltage is applied to the collector terminal, the emitter voltage e_1 at medium frequencies is

$$e_1 = h_{12} e_2 = j(\omega r'_b C_{22}) e_2 = j(\omega R_b C_{22}/3) e_2.$$

At the same time, the open-circuit voltage e_{b2} appearing at the upper (floating) base contact is 3/2 times as large, i.e.,

$$e_{b2} = (3/2) e_1. \quad (5)$$

On the other hand, if both base contacts are tied together, then the voltage available at the emitter terminal is reduced by a factor of 4, i.e., in this case

$$r'_b = R_b/12. \quad (6)$$

Experimental verification of these results at low frequencies is not especially easy since the nature of the base contact in a practical transistor can modify considerably the value of measured base spreading resistance r'_b . Nevertheless, agreement for a number of transistors is not especially bad as shown by the comparison in Table I.

Transistor	$(R_b/3)$ (ohms)	Table I r'_b			(e_b/e_e)	
		b-1	b-2	both*	b-1	b-2
172-2	270	250	220	35	1.5	2.0
199-4	830	350	1100	110	1.65	1.75
36-5	900	1000	1100	300	1.55	1.45
31-6	1300	1300	1300	300	1.6	1.5
35-5	2400	3000	3800	700	1.7	1.5

*Denotes Base that is grounded.

Note that agreement between experimental results and calculated values is quite poor for transistor 199-4, but is very good for transistor 31-6. In general, agreement is poorest in the case of r'_b for both bases grounded.

High Frequency Results

At moderate frequencies, the series expansions of the hyperbolic functions are no longer valid, and the hyperbolic functions must be considered in some detail. However, at higher frequencies, e.g., of the order of 0.2 - 0.5 times the α -cutoff frequency, it is possible in general to replace the hyperbolic tangent by unity and to approximate $|h|$ by the relatively simple expression

$$|h| \approx \sqrt{jaR_b C},$$

where C denotes the sum of the emitter and collector diffusion capacities plus emitter and collector barrier capacities. In this case, fairly simple approximate expressions can be obtained for the parameters of the distributed model of the transistor.

If collector-base capacity is moderately small, e.g., a few $\mu\mu\text{f}$, and if the collector diffusion capacity due to space-charge layer widening can be neglected, the expressions for the parameters of the distributed model assume their simplest form. In this case the current-amplification factor $-h_{21} = \alpha$ for the distributed model is substantially the same as that of the ideal one-dimensional model.⁷ Similarly, the open-circuit output admittance $h_{22} = g_{22} + jaC_{22}$ for the distributed model is not greatly different from that of the ideal model plus r'_b . The capacity C_{22} is essentially constant and equal to collector-barrier capacity C_c . However, the conductance g_{22} for the distributed model varies as the $3/2$ power of the frequency rather than as the square of the frequency.

On the other hand, the behavior of h_{12} and h_{11} is quite different from that predicted from the ideal model plus r'_b . The theoretical variation of h_{11} with frequency is shown by the normalized curves of Fig. 7. Short-circuit input resistance r_{11} reaches a maximum value of approximately $0.8 \sqrt{R_b r'_e}$ at about the α -cutoff frequency; here r'_e is the emitter resistance calculated by Shockley, et al.,¹⁰ as $r'_e = kT/qI_e$. On the other hand, the reactance x_{11} passes through 0 at about 1.2 times the α -cutoff frequency and becomes capacitive reactive at higher frequencies. It might be noted that the cross-over frequency of x_{11} may be decreased by 20-40% when the effect of

collector capacity is taken into account in calculating h_{11} .

A simple physical explanation for the behavior of h_{11} is as follows: The a-c base current flowing in the distributed resistance of the base region leads the emitter current because of the phase shift of α . Hence, the distributed base resistance effectively behaves as a distributed inductance. This in turn effectively resonates with the distributed diffusion capacity of the forward-biased emitter-base junction.

Finally, the open-circuit voltage feedback parameter h_{12} for the distributed model increases approximately as the square root of the frequency and has a phase shift of approximately 45° for frequencies of the order of 0.2-2.0 times the α -cutoff frequency. Alternatively, the feedback mechanism may be described approximately by the equivalent circuit shown in Fig. 8. This circuit is quite similar to the conventional equivalent circuit except that r'_b has been replaced by a complex base impedance z'_b whose frequency behavior is shown also in Fig. 8. Note also that the usual emitter impedance is not included in this circuit. A simple physical explanation for the behavior of h_{12} is that as frequency is increased, less of the base region is utilized for a-c because of the bypassing of the distributed base resistance by the distributed capacity.

The theoretical results presented here are in fairly good qualitative agreement with the experimental results described above in Figs. 2-5 for transistor 199-4, which was selected as having a low collector capacity. For example, the shapes of the curves of h_{11} in Figs. 5 and 7 are in fairly good qualitative agreement. Also, the nature of the observed frequency variation of $|h_{12}|$ and of g_{22} is in good agreement with the calculated variation. Similarly, for other transistors in general, qualitative agreement between experimental results and calculations based on the distributed model is quite satisfactory, although quantitative agreement is not always especially good.

Effect of Large Barrier Capacities

If emitter and collector barrier capacities are not negligible, the theoretical, and experimental, results described above may be modified considerably. One particularly striking effect of large barrier capacity is in the modification of the frequency variation of the current-amplification factor α . This is shown by curves in Fig. 9. for an experimental transistor having large emitter and collector capacitance (e.g., $C_{22} = 30 \mu\text{mf}$). Results of theoretical calculations based on the distributed model for the two values of d-c emitter current are shown by the dotted curves. Note that the α -cutoff frequency now is considerably dependent upon d-c emitter current (decreasing with decreasing I_e). Also note the relatively low value of phase shift ϕ of α , and the fact that ϕ attains a maximum value and then decreases with increasing frequency.

Internal Tetrode Action for Triode

The results described above for the new distributed model have been based upon constant diffusion admittance parameters, i.e., upon a uniform d-c current distribution across the transverse direction of the base. In practice, at moderate d-c emitter currents the d-c current distribution is not uniform, and the grown-junction transistor triode tends to behave as an internally biased tetrode,⁹ as described above. One measure of this tetrode action is the value of the low-frequency a-c base spreading resistance r_b^l . A second measure of the non-uniform d-c current distribution is the value of the open-circuit floating base potential E_{b2} appearing at the upper base terminal of the transistor when d-c bias is applied to the other three terminals. When this floating base potential exceeds the thermal voltage $kT/q = 25 \text{ mv}$, the d-c current distribution should be quite non-uniform. Furthermore, these two sets of results should be related. In general, they are, as shown by Fig. 10, in which normalized low-frequency r_b^l is plotted as a function of d-c emitter current for two different transistors. Also shown for these two transistors is the variation of the floating base potential E_{b2} with emitter current. For transistor 31-6, r_b^l decreases only slowly with increasing emitter current, and E_{b2} is relatively small for I_e up to

approximately 1-2 ma. On the other hand, for transistor 199-4, the d-c floating base potential is considerably higher at a given d-c emitter current, and r_b^l decreases much more rapidly with increasing emitter current.

Conclusion

Results of measurements of the frequency variation of the small-signal h parameters for grown-junction transistors which have been described clearly indicate that this type of transistor can not be represented by the usual model of an ideal theoretical transistor in series with a lumped constant base spreading resistance. A theoretical analysis of a new distributed model indicates that the concept of a lumped constant base spreading resistance is not valid for this type of transistor because of the distributed nature of the transistor parameters and of the spreading resistance of the base. From the results of the analysis, a new approximate equivalent circuit has been devised, in which base-spreading resistance is replaced by a complex base impedance whose magnitude decreases with increasing frequency.

Acknowledgements

The writers are grateful to R. N. Hall and C. B. Collins for supplying all of the grown-junction transistors used in this investigation and for helpful discussions concerning the theoretical analysis of the distributed model of the transistor. A number of people assisted with the measurements including J. L. Lawrence and D. J. Locke, and in particular, R. Johnston.

References

1. R. L. Pritchard, "Frequency Variations of Junction-Transistor Parameters," Presented at Electrochemical Society Meeting, New York, 15 April 1953; scheduled for publication in Proc. I.R.E.
2. H. O. Johnson, "Diffusion Reactances

junction transistors has been given by J. R. Nelson, "Transistor I-F Amplifiers," Tele-Tech 12, p. 69, Fig. 3; Dec. 1953.

7. a) W. Shockley, M. Sparks, and G. K. Teal, "p-n Junction Transistors," Phys. Rev. 83, p. 161; 1 July 1951.
- b) R. L. Pritchard, "Frequency Variations of Current-Amplification Factor for Junction Transistors," Proc. I.R.E. 40, 1480-1481; Nov. 1952.
8. J. M. Early, "Effects of Space-Charge Layer Widening in Junction Transistors," Proc. I.R.E. 40, 1404-1405; Nov. 1952.
9. R. L. Wallace, Jr., L. G. Schimpf, and E. Dickten, "A Junction Transistor Tetrode for High-Frequency Use," Proc. I.R.E. 40, 1395-1397; Nov. 1952.
10. W. Shockley, et.al., Footnote 7a, p. 156.
3. J. M. Early, "Design Theory of Junction Transistors," Bell Sys. Tech. Jour. 32, 1271-1312; Nov. 1953.
4. R. L. Pritchard, "Collector-Base Impedance of a Junction Transistor," Proc. I.R.E. 41, 1060; Aug. 1953.
5. C. W. Mueller and J. I. Pankove, "A P-N-P Triode Alloy Junction Transistor for Radio-Frequency Amplification," RCA Rev. 14, p. 594, Fig. 4; Dec. 1953, also Proc. I.R.E. 42, p. 389; Feb. 1954.
6. See, for example, in the literature:
 - a) Mueller and Pankove p. 594-595.
 - b) Experimental evidence that the base spreading resistance r_b' is constant with frequency for fused-

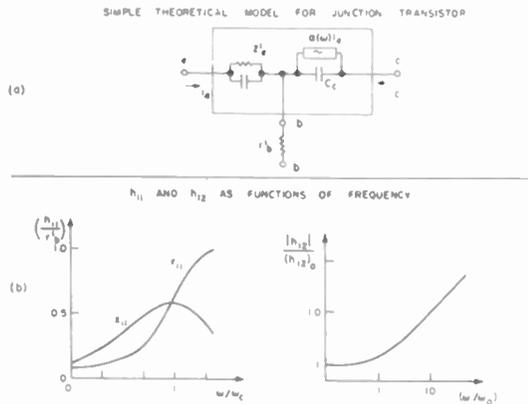


Fig. 1
Equivalent circuit for ideal one-dimensional theoretical model of junction transistor with external base spreading resistance; relative frequency variation of h_{11} and h_{12} for this model.

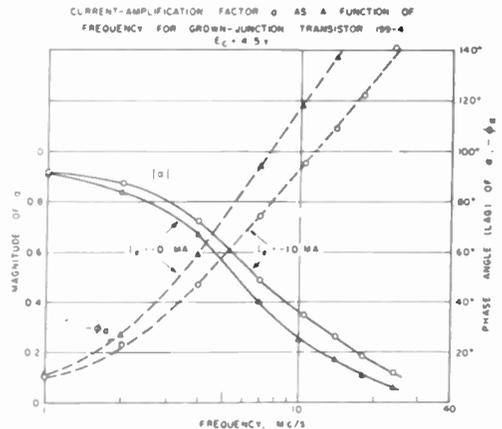


Fig. 2
Frequency variation of short-circuit current-amplification factor $-h_{21} = \alpha$ for grown-junction transistor.

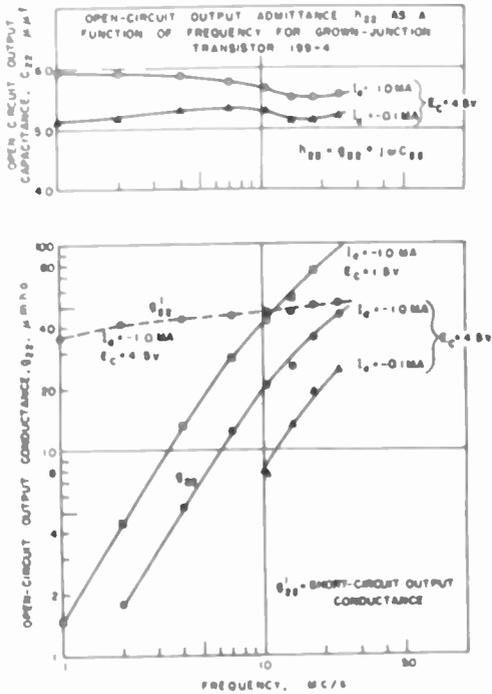


Fig. 3
Frequency variation of open-circuit output admittance h_{22} for grown-junction transistor.

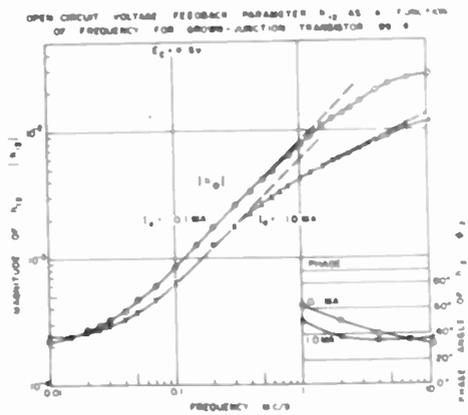


Fig. 4
Frequency variation of open-circuit voltage-feedback parameter h_{12} for grown-junction transistor.

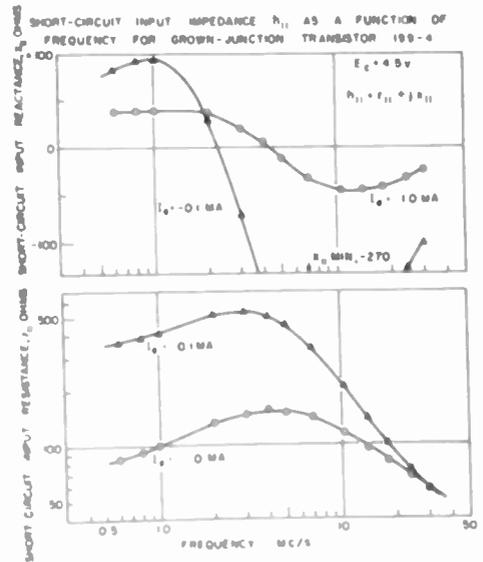


Fig. 5
Frequency variation of short-circuit input impedance h_{11} for grown-junction transistor.

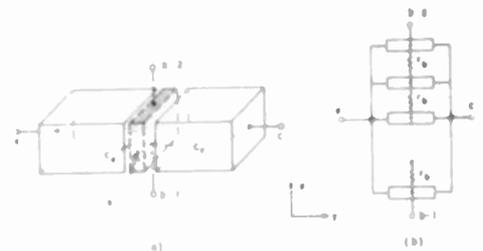


Fig. 6
Distributed theoretical model for grown-junction transistor.

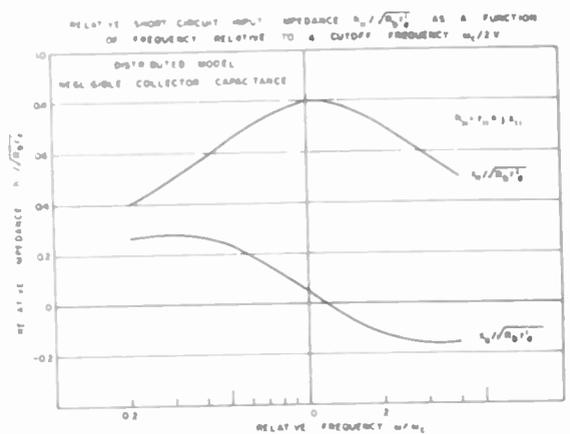


Fig. 7
Theoretical variation of short-circuit input impedance with relative frequency for distributed model of grown-junction transistor having negligible collector capacitance.

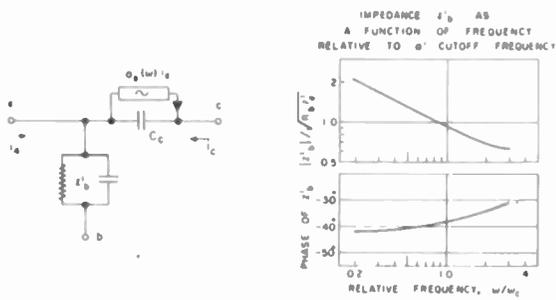


Fig. 8

Approximate equivalent circuit for theoretical distributed model of grown-junction transistor at high frequencies, including theoretical frequency variation of complex base impedance z'_b in the circuit.

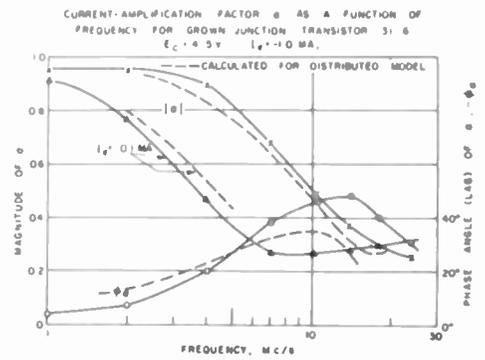


Fig. 9

Frequency variation of short-circuit current-amplification factor $-h_{21} = \alpha$ for grown-junction transistor having large barrier capacitance.

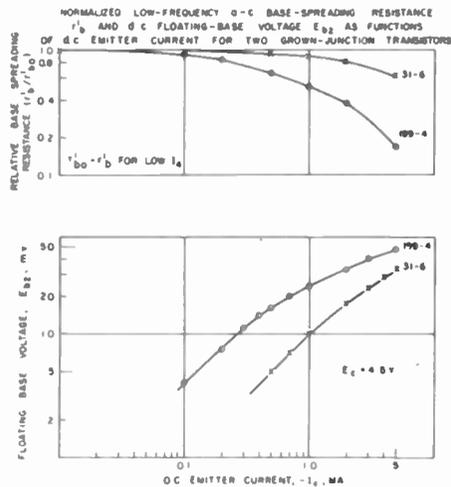


Fig. 10

Illustration of internally biased tetrode effect for two grown-junction transistor triodes; variation of low-frequency base spreading resistance and of dc floating base voltage with dc emitter current.

THE STUDY AND DESIGN OF ALLOYED-JUNCTION TRANSISTORS

by

L.J. Giacoletto
 RCA Laboratories Division
 Radio Corporation of America
 Princeton, N.J.

Introduction

This paper will present some measurements and consider the design of a junction transistor. A germanium p-n-p alloyed-junction transistor is considered herein, but the results are generally applicable to both n-p-n and p-n-p junction transistors, to junction transistors made by other techniques, and to junction transistors made with other semiconductors. Due to space limitations only a small part of the complete study can be presented here.

General Considerations

Direct-current measurements of the junction transistor indicate that the usual transistor theory¹ is applicable provided suitable allowances are made for certain "extraneous" or extrinsic elements. The most important of these extrinsic elements is a base-lead resistance, $r_{bb'}$. The next in importance is a leakage conductance, g_L . These extrinsic elements are shown in Fig. 1 together with the "basic" or intrinsic transistor.

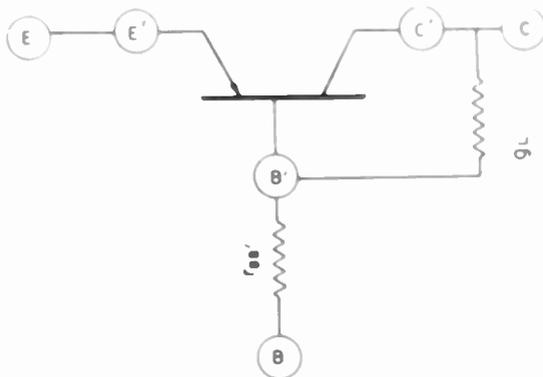


Fig. 1

Composite Transistor Showing Placement of Extrinsic Elements

These same extrinsic elements, now determined by ac measurements, are shown in Fig. 2 appended to a π equivalent circuit of the intrinsic transistor. The resulting equivalent circuit is a hybrid π for a common-emitter circuit. The leakage conductance can be conveniently lumped in with the intrinsic feedback admittance, $y_{b,c}$, but there is not much that can be done with the base-lead resistance to simplify the circuit. The full importance of the base-lead resistance will become apparent shortly.

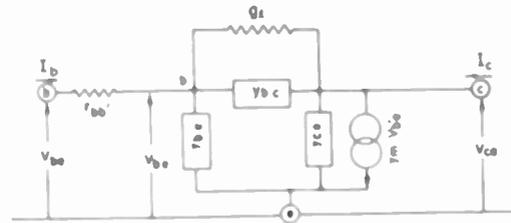
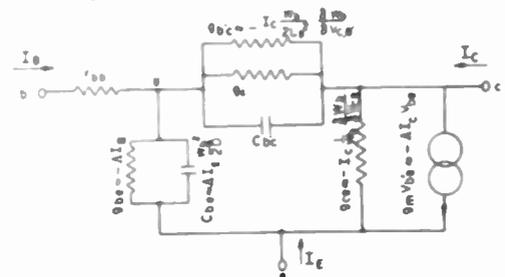


Fig. 2

Common-Emitter Hybrid π Equivalent Circuit

All the parameters in this hybrid equivalent circuit have been measured as a function of dc current, dc voltage, and frequency. In addition, these measured results have been compared with analytic formulations of the same quantities. Comparison of measurement with theory is considerably complicated by the fact that one cannot get to the intrinsic transistor for direct measurements. As a consequence, the intrinsic parameters have to be deduced from terminal measurements. The net result of this comparison work has been quite gratifying as it is found that most of the basic aspects of the transistor can be formulated analytically. Thus, it has been found that for most practical purposes the junction transistor can be represented as shown in Fig. 3. Here, the parameters of the intrinsic transistor that are of importance have been formulated analytically in terms of the operating point and the material constants and dimensions used. In this figure, the Greek capital letter, Λ , is used in place of $\frac{c}{\tau}$, the other symbols used have their usual meaning.



$$C_{bc} = \frac{\kappa_b \epsilon_0 \Lambda c}{2 \kappa_0 \epsilon_0 \mu V_{ce}} \frac{1}{\sigma_b} - I_c \frac{W_D}{2D} \frac{W_D}{V_{ce}}$$

$$\frac{\beta W_D}{\beta V_{ce}} = \frac{1}{2} \left[\frac{2 \kappa_0 \epsilon_0 \mu}{\sigma_b V_{ce}} \right] \frac{1}{2}$$

Fig. 3

Analytic Formulation of Hybrid π Equivalent Circuit

Satisfactory analytic formulations of the extrinsic elements, the base-lead resistance and the leakage conductance, are not available so these elements must be determined either by measurement or by estimation from previous experience. The base-lead resistance will be directly proportional to the resistivity of the semiconductor material and will depend upon the geometry of the base connection. For some configurations analytic formulation of the base-lead resistance may be possible. The leakage conductance will be greatly influenced by surface treatment and may undergo changes with time.

The circuit shown in Fig. 3 is an approximation to more exact but also more cumbersome circuits. The important approximations involved in obtaining Fig. 3 are: (1) that the collector voltage is sufficient for current saturation; (2) that the base width, W_b , is smaller than the diffusion length of minority carriers in the base, L_b ; and (3) that the radian frequency is less than $\frac{3D}{W_b^2}$. There are also other approximations of lesser importance.

One of the purposes for using a hybrid n equivalent circuit as shown in Fig. 3 is to obtain a representation that is valid over a wide range of frequencies. This is of great importance to both device and circuit designers. The equivalent circuit parameters can be determined by low-frequency measurements and yet be employed for high-frequency calculations. This is a situation similar to electron tubes where interelectrode capacitances are determined by measurements at 1 kc/s and used in calculations up to a few hundred megacycles/second.

Input Self-Admittance

Consideration will now be given to the input self-admittance together with its constituent components. Variation with dc emitter current is shown in Fig. 4. Measurement conditions are as shown in the figure. The base-lead resistance, r_{bb} , is seen to be essentially independent of current except for a rather sharp increase at very small currents. The intrinsic transistor input self-conductance, $g_{b'e}$, increases linearly with emitter current. The computed variation of this quantity is shown by the dashed line. The terminal measurement of the input self-conductance is given by the curve labelled g_{bb} . Likewise, the terminal measurement of the input self-capacitance is given by the curve labelled C_{bb} . The intrinsic transistor input self-capacitance is given by the curve labelled $C_{b'e}$. In accordance with theory this capacitance is linearly proportional with current - at least for small currents. From the initial slope of this linear dependency the value of the base thickness, W_b , can be computed. This forms a relatively simple and accurate method for determining this important transistor constant. By suitable interpretation of input self-admittance parameters as a function of collector-to-emitter dc voltage three other independent evaluations of the base

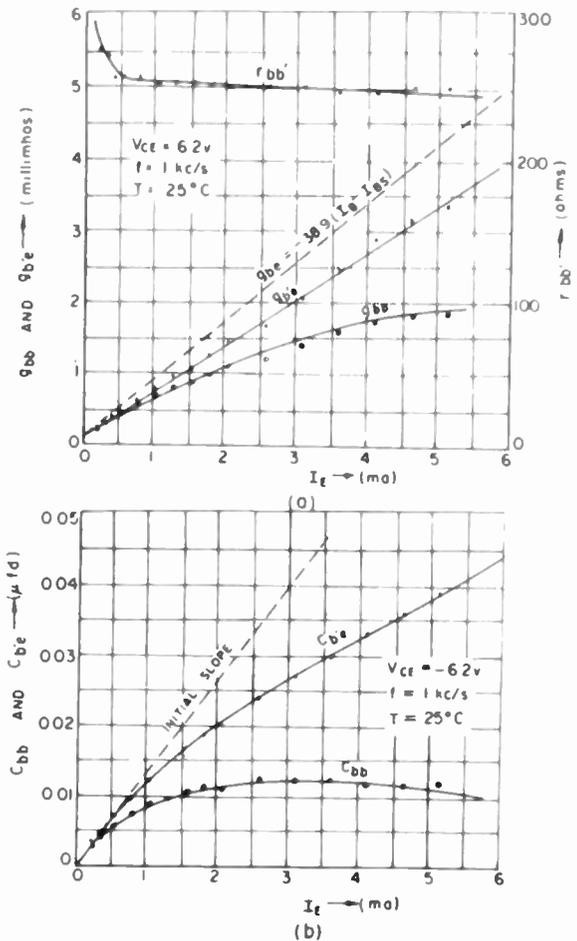


Fig. 4
Variation of Input Self-Admittance and Its Constituent Components with Emitter Current

thickness can be obtained. All four values were found to be in good agreement. Returning to the functional dependency of $C_{b'e}$ with emitter current, a linear variation is again found at high currents but with a slope variation that is very nearly half of the initial slope. Some words of explanation may be of interest in connection with this complex behavior of this parameter. A basic assumption of most theory is that the injected minority carrier density is small in comparison with the normal base material majority carrier density. This basic assumption is satisfied for only relatively small currents. Thus, for the transistor under consideration here the injected hole density at the emitter becomes equal to the base material electron density at an emitter current of 0.4 milliamperes (emitter current density of 0.5 amperes/cm²). As noted in Fig. 4, this is approximately the current at which $C_{b'e}$ begins to depart from the initial linear variation. When the operating current density is increased an electric field through the base becomes significant and more and more of the current is carried by the electric field. In the limit, half of the current is carried by the electric field and half by the diffusion gradient. This accounts for

the two-to-one relationship between the initial and final slope of $C_{b'e}$. One might ask whether the breakdown in the basic assumption does not have serious consequences for the other parameters. Fortunately this does not appear to be the case insofar as low-level small-signal operation is concerned. However for large-signal operation⁴ and for large-current operation⁵ suitable allowances must be made.

The variation of the input self-admittance parameters as a function of collector-to-emitter voltage is shown in Fig. 5. The various quantities

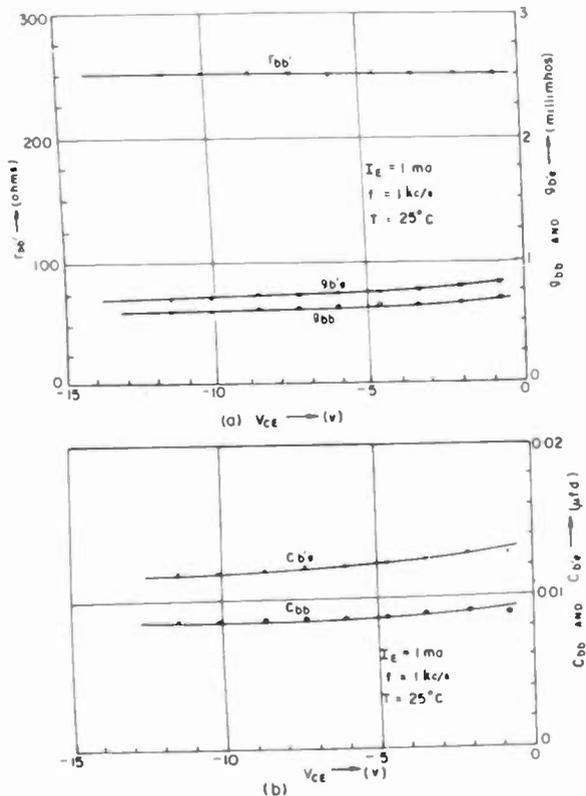


Fig. 5
Variation of Input Self-Admittance and Its Constituent Components with Collector-to Emitter Voltage

do not vary greatly with voltage. However, as mentioned before, the variation present can be used to determine the base thickness.

The frequency dependency of the input self-admittance parameters is shown in Fig. 6. The terminal self-conductance and self-capacitance have a pronounced frequency dependency. This is to be expected in view of its composition. In contrast, the constituent parameters are essentially independent of frequency. As was indicated before, theory states that the intrinsic transistor parameters should be independent of frequency as long as the period is large in comparison with the

time of transit through the base. For the transistor under study here, the frequency correspond-

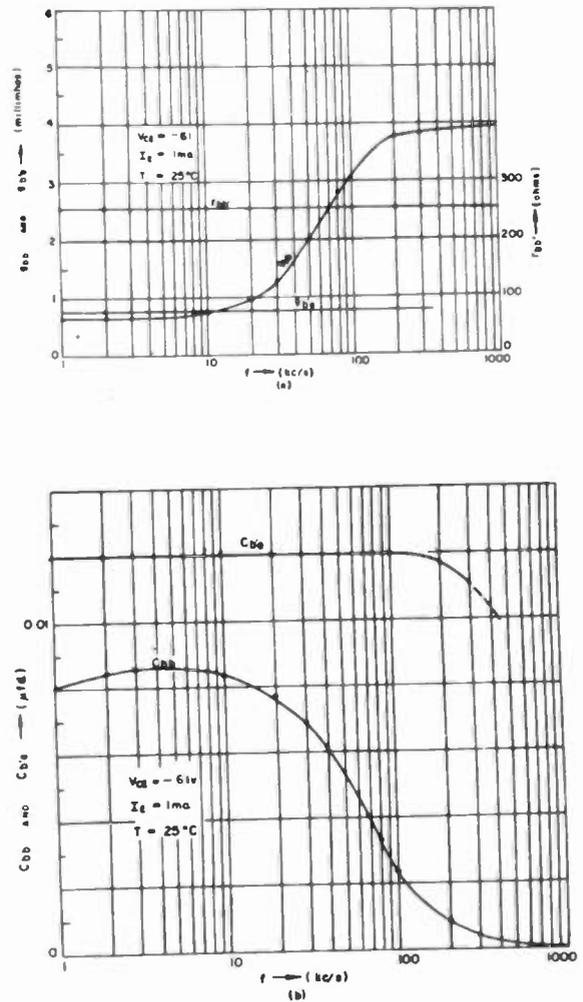


Fig. 6
Variation of Input Self-Admittance and Its Constituent Components with Frequency

ing to $\frac{3D}{2\pi W_b}$ is 685 kc/s. Therefore it is not surprising that the intrinsic self-capacitance, $C_{b'e}$, changes with frequency above 100 kc/s.

In the complete study of the junction transistor, the other three parameters are examined in much the same way as was done for the input self-admittance parameter. Space does not permit the presentation of these data. One of the results of this complete study is that the hybrid π equivalent circuit for this transistor is as shown in Fig. 7. The operating conditions for which the transistor parameters are valid are shown on this figure. It is perhaps worth repeating again that this equivalent circuit is independent of frequency for most practical purposes.

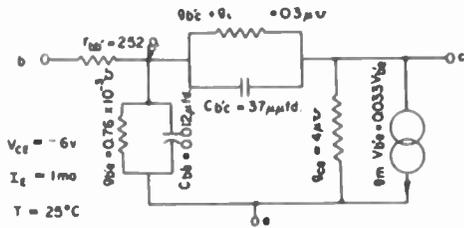


Fig. 7

Hybrid π Equivalent Circuit for a p-n-p Alloyed Junction Transistor

Maximum Power Amplification

One of the circuit considerations that is of considerable interest is the common-emitter single-frequency power amplification. For maximum power amplification, the input and output admittances must be conjugately-matched. Before presenting the exact results for this operating condition, it is of interest to consider a formulation which is approximately valid at higher frequencies. At these higher frequencies the capacitor, $C_{b'e}$ is approximately a short-circuit, and the generator resistance for maximum input power is therefore $r_{bb'}$. On the output side it is found that the output impedance is also approximately a pure resistance at higher frequencies. This comes about because an output voltage causes the g_m generator to be activated by an in-phase voltage arising from a voltage division between $C_{b'e}$ and $C_{b'c}$. That is, at higher frequencies,

$$g_o \approx - \frac{C_{b'c}}{C_{b'e}} \wedge I_C$$

After formulating the input and output match analytically, it is found that the approximate maximum power amplification at higher frequencies is

$$P.A._{HF} = \frac{g_m}{4\omega^2 r_{bb'} C_{b'c} C_{b'e}} = \frac{1}{4\omega^2 r_{bb'} C_{b'c} \frac{W_b^2}{2D}}$$

With the aid of this formula, a very useful figure of merit can be obtained. This figure of merit,

$$\frac{1}{4\pi} \left[\frac{g_m}{r_{bb'} C_{b'c} C_{b'e}} \right]^{1/2}$$

is the frequency at which the power amplification is unity; it is also approximately the maximum oscillation frequency for the transistor. Both the maximum power amplification and the figure of merit are essentially independent of operating dc current but will increase somewhat with larger collector dc voltages.

The approximate power amplification at higher frequencies for the transistor in Fig. 7 is shown by the dashed line in Fig. 8. The

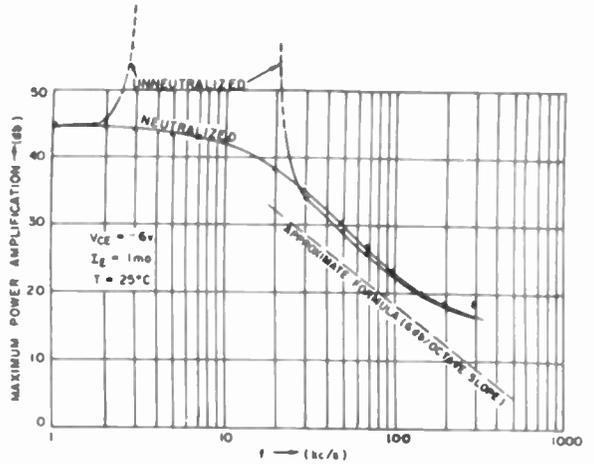


Fig. 8

Variation of Maximum Power Amplification with Frequency

6 decibels per octave slope is closely in accord with the exact power amplification. The figure of merit in this case is seen to be 900 kc/s. The exact maximum power amplification is shown in Fig. 8 for both the unneutralized and neutralized case. If the transistor is conjugately-matched at both the input and output, oscillations are obtained for a range of frequencies - in this case from about 3 to 20 kc/s. To a large measure, these oscillations are due to the feedback capacitor, $C_{b'e}$. Neutralization, as for instance with a collector-to-base inductor can be used to eliminate the oscillations. However, since the internal base point is not accessible, neutralization is of small value at higher frequencies as is seen in Fig. 8.

It is readily apparent that in order to obtain power amplification at higher frequencies, $r_{bb'}$ and $C_{b'e}$ must be reduced. The former can be done by using lower resistivity germanium, and the latter by making the transistor with a smaller base thickness. Also, the feedback capacitor, $C_{b'c}$, must be reduced. This can be done by reducing the area of the collector. The aforementioned methods of improvement have been employed by C.W. Mueller and J.I. Pankove⁵ to make a radio-frequency transistor with an oscillation frequency of 75 mc/s.

The input admittance corresponding to maximum power amplification is shown in Fig. 9 for both the unneutralized and neutralized operating condition. It is seen that neutralization serves to greatly reduce the magnitude of the input capacitance and to make both input conductance and capacitance more nearly constant with frequency. At lower frequencies the input capacitance for the unneutralized circuit is considerably larger than $C_{b'e}$. This condition is brought about by Miller effect in the same manner as in an electron tube. At higher frequencies it is seen that the input admittance approaches a pure conductance corresponding to the base-lead resistance.

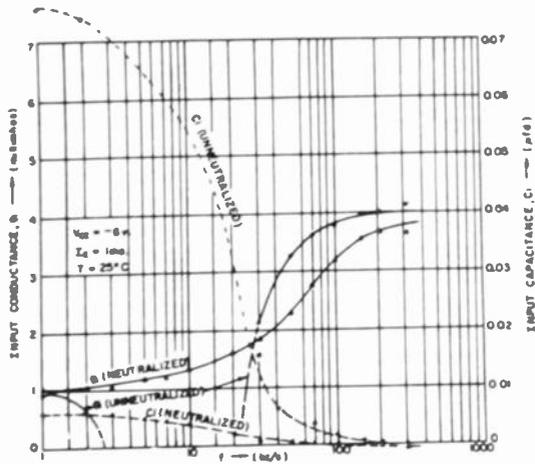


Fig. 9

Variation of Input Admittance for Maximum Power Amplification with Frequency

The output admittance corresponding to maximum power amplification is shown in Fig. 10.

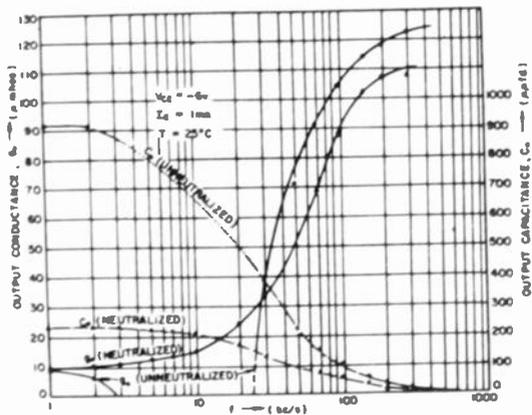


Fig. 10

Variation of Output Admittance for Maximum Power Amplification with Frequency

Here again, neutralization serves to reduce the output capacitance and to make both the output conductance and capacitance more nearly constant

with frequency. At higher frequencies the output admittance approaches a pure conductance as was mentioned before.

Conclusions

The hybrid π equivalent circuit shown in Fig. 3 should be valuable to both device and circuit designers since the parameters can be formulated analytically in a reasonably simple manner and since the circuit is independent of frequency throughout the range of practical importance. The manner in which the transistor operation is dependent upon operating voltage, current, frequency, and temperature can be predicted so that calculations can be carried out to obtain the answer desired. In the hybrid π equivalent circuit it is important to distinguish between the parameters of the intrinsic transistor and the extrinsic elements.

Although the material presented herein is of a p-n-p alloyed junction transistor made with germanium, the results have general applicability to other junction transistors including n-p-n transistors, transistors made from other materials, and transistors made by different assembly techniques.

References

1. W. Shockley, M. Sparks, and G.K. Teal, "p-n Junction Transistors," *The Physical Review*, vol. 83, pp. 161-162; July 1, 1951.
2. J.M. Early, "Design Theory of Junction Transistors," *B.S.T.J.*, vol. 32, pp. 1271-1312; Nov. 1953.
3. W.M. Webster, "On the Variation of Junction Transistor Current Amplification Factor with Emitter Current," to be published.
4. L.J. Giacoletto, "Power Transistors for Audio Output Circuits," *Electronics*, vol. 27, pp. 144-148; Jan. 1954.
5. C.W. Mueller and J.I. Pankove, "A p-n-p Triode Alloy Junction Transistor for Radio-Frequency Amplification," *RCA Review*, pp. 586-598; December, 1953 and *Proc. I.R.E.* pp. 386-391, February 1954.

AN ANALYTICAL STUDY OF z , y , and h PARAMETER ACCURACIES IN TRANSISTOR SWEEP MEASUREMENT

H. G. Follingstad
Bell Telephone Laboratories, Inc.
Murray Hill, N. J.

Summary

A comparison is made of measurement accuracies which can be realized for low frequency z , y , and h parameters using practical sweep techniques. The comparison includes all regions of the transistor characteristic and involves both point-contact and junction transistors in the common-base and common-emitter connections. Formulas are developed which determine error in terms of the transistor parameters and actual finite terminations. Errors are calculated assuming practical terminating impedance values from simple, well-established sweep design techniques, and assuming typical transistor parameter values in each region. The analysis shows that at low frequencies the most generally useful and accurate sweep measurements are obtained with hybrid (h) parameters. Sweep bandwidth requirements impose the most severe restrictions on the open-circuit and short-circuit measuring technique. Therefore, the accuracy of all of the parameters can be improved by point-by-point current-voltage or bridge methods to such a degree as to make acceptable some parameters which are unsuitable for sweep measurement. However, the increasing accuracy in point-by-point measurements applies as well to the hybrid parameters. Therefore it may be concluded that the hybrid (h) parameters are the only set so far proposed which will yield acceptable accuracies in both sweep and point-by-point types of low frequency measurement.

1.0 Introduction and Purpose

Open-circuit and short-circuit four-pole parameters¹ of a network are defined by assuming infinite or zero impedance terminations. Therefore inherent measurement error always exists with finite terminations. As the frequency of measurement increases to a value where parasitic reactances at the terminals become dominant or unpredictable, measurements of open and short-circuit parameters become impractical and other methods must be used. At low frequencies, however, terminations may usually be made sufficiently higher or lower than the network impedances that an acceptable approximation can be made to at least one of the six² possible sets of four-pole parameters.

In the present transistor art three of the six sets have been frequently measured with reasonable accuracy on one or more existing transistor types in various modes of connection. They are the open circuit impedance (z), short circuit admittance (y), and hybrid (h) parameters. The equivalent circuit equations which define them are shown in Fig. 1. It can be seen that the z and y parameters have impedance and admittance dimensions, respectively, whereas the h parameters are a mixed set having the following dimensions:

- h_{11} - input IMPEDANCE with a-c short-circuited output.
- h_{12} - feedback VOLTAGE RATIO with a-c open-circuited input.
- h_{21} - forward CURRENT RATIO with a-c short-circuited output.
- h_{22} - output ADMITTANCE with a-c open-circuited input.

The relations between the z , y , and h parameters are given in the Appendix, Fig. A1.

In this paper, a comparison is made of measurement accuracies which can be realized for z , y , and h parameters at low frequencies using terminating impedance values from well-established sweep design techniques. The comparison is also based on typical transistor parameter values in all regions of the transistor characteristic and involves both point contact and junction transistors in the common-base and common-emitter connections.

The primary objective of the comparison is to determine which set of parameters are most accurately measured on a sweep basis so as to obtain a single set which is equally useful in sweep and point-by-point measurements.

2.0 Source and Termination Error Formulas

In transistor measurement popular conceptions of estimated accuracies are often wrong due to neglect of active feedback effects in the various regions of the

transistor characteristic. An analogous wrong conclusion in the vacuum tube art is that the output impedance of a cathode follower, for example, is equal to the cathode resistor value. To accurately determine measurement error, formulas have been developed for the z , y , and h parameters which show the exact relation between measured and defined values in terms of two error factors. The latter conveniently relate the Thevenin source and terminating impedances to the transistor parameters.

The error factor derivations follow the sample z_{11} case in Fig. 2. In Fig. 2(a) the input impedance, $\frac{v_1}{I_1}$, deviates from open circuit impedance, z_{11} , as I_2 deviates from zero. The additive error term, $z_{12} \frac{I_2}{I_1}$, is conveniently translated into a "termination error factor" multiplying z_{11} . The error factor is a function of the termination, z_T , and the z parameters of the transistor. A further deviation from z_{11} occurs as the transistor input current, I_1 , in Fig. 2(b) deviates from a known calibrated current, $I_1 \text{ CAL}$. The ratio of these currents is a "source error factor" multiplying the input impedance measurement, $\frac{v_1}{I_1}$. The source error factor is conveniently expressed as a partial function of the generator source impedance, z_S . Fig. 2(c) illustrates that the measured input impedance is z_{11} modified by the two aforementioned error factors. The error factors for the z_{11} case are derived in the Appendix.

Complete error formulas for all of the z , y , and h parameters are given in the Appendix, Figs. A2-A4. The formulas are a convenient tool at any frequency for analysis of measurement error. In this paper they form the basis for low frequency analysis of error in sweep measurement.

3.0 Low Frequency Sweep Measurement Objectives

In the low frequency range two basic sweeper types are postulated, one a slow, precise recorder type with $\pm 2\%$ or better accuracy, and the other a fast oscilloscopic type with the order of $\pm 10\%$ accuracy. The proposed display coordinates are shown in Fig. 3. The z , y , or h parameter is displayed as an ordinate versus swept input current as abscissa with collector voltage as family parameter. The latter independent variables have been most consistently useful in the transistor design art.

The bandwidth objectives for these sweepers are indicated in Fig. 3. The z , y , or h information is obtained by a small derivative or probe current with a frequency

one decade above the display bandwidth of the swept waveform to permit simple filtering and detection. The display bandwidth, Δf , includes 400 harmonics of the sweep frequency to provide freedom from harmonic distortion of the display waveforms. The frequency data for the oscilloscope type sweeper are shown in brackets with a probe frequency of 100 KC, a display band of 10 KC, and a sweep frequency of 25 cps. The frequency data for the recorder type are scaled down by a factor of 100. Both sweeper types have a maximum probe current of 10 μa and maximum swept input current range of 50 ma.

4.0 Practical Source and Terminating Impedance Values in Low Frequency Sweep Measurement

Assuming the previous frequency, bandwidth, and current objectives, the basic plan for determining practical, realizable impedance values was to assume simple but essential building-blocks and arrange them in an optimum way for sweeping each of the twelve z , y , and h , parameters. Additional complexity of circuit functions can conceivably produce somewhat better open and short circuits but this is unnecessary if acceptable accuracies can be obtained in at least one set of parameters using the simpler building blocks.

The four essential sweeper functions are: (1) Probe current or voltage generator (2) Low frequency input sweep current (3) Probe measuring circuit (4) d-c family parameter supply voltage. Arrangement of these building blocks for sweeping r_{11} is indicated in Fig. 4. Sweep measurement of r_{11} is based on constant probe current and an open circuit termination. Thus all of the branches shown must have high impedances relative to the transistor. The probe source is shown as a current generator, I_G , with shunt generator resistance, r_G . The sweep circuit has a necessarily low impedance to the low frequency sweep current, I_1 , but has high impedance to the probe frequency by means of parallel resonance. The open circuit input voltage, v_1 , is monitored by a high impedance, wide band measuring circuit which filters, amplifies, and detects the r_{11} information and presents it to the Y axis of the display mechanism. Input current, I_1 , is calibrated by the voltage drop across a calibrated resistance, $r_{11} \text{ CAL}$. The output termination necessarily has a low impedance to the d-c family parameter, collector voltage, but a high impedance to the probe frequency and its side bands containing the r_{11} waveform components. The impedance levels shown are examples for the 1 KC probe frequency case. The probe source impedance can be quite high

and is limited to 20 megohms only by a practical limit on probe voltage for the given probe current of 10 μ a. Similarly the probe measuring circuit impedance of 10 megohms is limited by the wide measurement bandwidth ($Q = 5$). The limiting impedance levels are those in the sweep and collector voltage paths. Here both circuits must present a high impedance to the r_{11} sidebands, limiting the coil Q 's to 5. The coils must also be able to pass high d-c currents of the order of 50 ma or more. These factors limit the resonant resistance to about 500 kilohms. Thus each of the composite source and terminating impedances, r_S and r_T , are 500 kilohms.

The four essential sweeper functions take on various forms and values in each of the sweepers as summarized in Figs. 5 and 6. The indicated values are based on practical design limitations. In Fig. 5(a), for example, with 0.2 μ pf capacitance across r_G , probe current source resistance, r_G , at 100 KC is limited to 1.5 megohms. At 1 KC the limit of 20 megohms is due to probe voltage limits for a 10 μ a current. As a voltage source, an r_G of 1 ohm is practical at both 1 KC and 100 KC. In Fig. 5(b), the voltage monitor resistance, r_M , is limited by the wide measurement bandwidth required ($Q = 5$). The current monitor resistance, r_M , is a minimum of 10 ohms because of transistor and tube noise limitations when measuring small currents (10 μ a) in a relatively wide band. In Fig. 6 the sweep and family parameter circuits, when functioning also as open circuit terminations, must present a high impedance to the information side bands, limiting the coil Q to 5. To prevent detuning from resonance by transistor or stray capacities a minimum of 80 μ pf tuning capacitance is required. These factors limit r_{SW} and r_F at 100 KC to about 100 kilohms. The limitation at 1 KC is not stray capacity but is the iron core inductance required to maintain the impedance while passing large d-c currents. Thus r_{SW} and r_F at 1 KC are limited to about 500 kilohms. Identical limitations exist when the sweep or family parameter circuits are in parallel with the probe source and measuring circuit. If the measuring circuit is absent these circuit impedances can be raised by a factor of 10 ($Q = 50$) as shown. As a short circuit termination, a resistance of 1 ohm is practical for both circuits at 1 KC and 100 KC.

Arrangements of the twelve z , y , and h sweepers are shown in the Appendix, Figs. A5-A10. The composite source and terminating impedances are also indicated for the 1 KC and 100 KC cases. These impedance values are the basis for error comparison of the z , y , and h parameters in this paper.

5.0 Parameter Values of Typical Transistor Models in all Regions of Operation

In sweep measurements errors due to finite terminations should be determined in all regions of the transistor characteristic. A general regional division has been proposed for switching type transistors⁴ which of course can include transmission types as well. An adaptation of this regional scheme is shown in Fig. 4. The regions may be defined as follows⁵:

- Region (1) Collector Voltage Saturation.
- Region (1-2) Transition.
- Region (2) Active.
- Region (3) Collector Voltage Cutoff.

The measurement accuracy required in each region depends upon the importance of that region in a particular application. In switching applications greatest accuracy is required in the transition (1-2) and active (2) regions. Passive regions (1) and (3) often may be satisfactorily specified by d-c measurements alone or with reduced accuracy in swept a-c measurements. In low level transmission applications accuracy in active region 2 is of major concern, and reduced accuracy is usually permitted in the other regions.

Evaluation of measurement error in each region is based on the relation of the transistor parameters in that region to the source and terminating impedances. Models of both point contact and junction transistors have been assumed in all of the regions. Use of these models permits a realistic estimate of measurement errors likely to be met in the great majority of transistor types. The basic internal parameters of the models are shown in Fig. 7. A six-point approximation is assumed. Regions 1 and 3 are each represented by a single limiting point whereas the transition and active regions are approximated by four points as shown at the bottom of Fig. 7. The assumed values of r_e , r_b , r_c , and "a" in each region are probably not more than a factor of 2 away from most present types. Results of conversion of r_e , r_b , r_c , and "a" values to the corresponding r , g , and h parameters are shown in the Appendix, Figs. A11-A17. The r , g , and h values may be directly inserted into the error formulas as well as the source and terminating impedance data of Sec. 4.0 to obtain measurement error in all regions of the characteristic.

6.0 Comparison of Errors in Low Frequency z , y , and h Sweep Measurement

Insertion of the termination and transistor parameter data of secs. 4.0 and 5.0 into

the error formulas of sec. 2.0 permit direct error comparisons of the low frequency z , y , and h parameters in all regions of the characteristic. The comparisons from region to region are shown in Figs. 9-13, inclusive. They include both point-contact and junction transistors in both common-base and common-emitter connections and both at 1 KC and 100 KC, with the following exceptions: (1) Comparisons of point contacts in the common-emitter case and the y data in the common-base case is omitted because sweeping causes either instability or large errors with the required negative impedance-stabilizing terminations. (2) A 100 KC comparison is omitted for the common-emitter junction because the transmission cutoff frequencies are often lower than 100 KC so that the parameter magnitudes may deviate radically from the low frequency real values.

In Figs. 9-13 the error data are presented with percentage error plotted vertically and the regional positions horizontally. Shaded areas indicate unacceptably large errors. In the 1 KC case a maximum error of 2% is consistent with the order of accuracy of recorder type display mechanisms. In the 100 KC case a maximum error of 10% is compatible with average oscilloscopic display accuracies. However, a maximum error of up to 30% is reasonable in Regions 1 and 3 because they are passive regions and may often be specified as well by d-c measurements.

In Fig. 9 the familiar common-base point contact at 1 KC, which historically has been measured with r parameters, indicates that the h parameters are preferred, although acceptable accuracies are obtained with some of the r 's. The advent of the junction transistor caused a general departure from r parameters to h and g parameters. The superior accuracy of the h parameters over either the r or g sets is evident in the accuracy trends of Fig. 10 which show the common-base junction at 1 KC. In Fig. 11 the common-emitter junction comparisons at 1 KC reflect superiority of the g parameters over the r or h set. However, the h parameters yield acceptable accuracies even though not quite as good as the g set. Figs. 12 and 13 show that even at 100 KC the h parameters as a set remain more accurate than the r or g sets in both the common-base point-contact and junction cases. The general superiority of the h parameter set is evident when the five comparisons are taken as a whole.

If the curves of Figs. 9-13 are analyzed individually on the basis of acceptability in each region, the parameter selector chart in Fig. 14 is obtained. The shaded areas indicate that for the condition specified the parameters have entered

the forbidden error region in the previous figures. Fig. 14 shows that in the regions of greatest importance (1-2 and 2) the h parameters are clearly the most generally accurate set. However, it is also apparent that the g parameters for the common-emitter junction case alone are superior to the others. The r_{12} parameter also has some merit. It should be noted that the r and g parameters, except for the above cases, tend to be most accurate in the regions of least importance (1 and 3) whereas the h parameters are most accurate in the regions of greatest importance. The one weak point of the h set is in Region 1 where h_{11} and h_{12} are marginal. These cases may be supplemented by g_{11} and g_{12} measurements or by d-c measurements in Region 1.

It should be emphasized that the numerical results are based on sweep measuring techniques and cannot be extrapolated directly to point-by-point measurements. Because the sweep technique imposes the most severe restrictions on terminating impedances, a point-by-point parameter selector chart would have less unacceptable parameters. However, the same trend toward greater accuracy applies also to the h parameters in the point-by-point case.

7.0 Conclusions

Comparisons have been made of typical z , y , and h parameter accuracies in all regions of the transistor characteristic and have involved both point-contact and junction transistor types in both the common-base and common-emitter connections. Error calculations were based on practical terminating impedance values from simple, well-established sweep design techniques and assumption of typical transistor parameter values in each region. The results of the analysis afford many possibilities for comparison. A particularly useful result is obtained when maximum allowable percentage errors are assumed, consistent with the relative importance of each region and consistent with the inherent accuracy of the display mechanism used. Under these conditions it may be concluded that at low frequencies the most generally useful and accurate sweep measurements are obtained with hybrid (h) parameters. The sweep type measurement imposes the most severe restrictions on the open-circuit and short-circuit measuring techniques. Therefore all parameter accuracies can be improved by point-by-point current-voltage or bridge methods to such a degree as to make acceptable some parameters which are useless in sweep measurement. However the increasing accuracies in those methods

apply as well to the hybrid (h) parameters. Therefore it may be concluded that the hybrid (h) parameters are the only set so far proposed which will yield acceptable accuracies in both sweep and point-by-point types of low frequency measurement.

8.0 Acknowledgment

This evaluation study was originally made as part of a program on Transistor Test Methods sponsored by the U.S. Armed Forces Joint Services under Contract No. DA36-039 sc-5589, administered by the Signal Corps. Acknowledgment is gratefully made to the author's colleagues for helpful criticisms and suggestions as to assumptions and methods of presentation.

9.0 Bibliography

- 1,2
E. A. Guillemin, Communication Networks, Vol. II, Chap. 4, John Wiley & Sons, Inc., New York, N.Y.; 1935.
- 1,2
L. C. Peterson, "Equivalent Circuits of Linear Active Four-Terminal Networks", Bell System Technical Journal, Vol. 27, pp. 593-622; October, 1948.
- 3
H. G. Follingstad, "A Transistor Alpha Sweeper", Convention Record of the I.R.E., 1953, Part 9.
- 4,5
A. E. Anderson, "Transistors in Switching Circuits", Proc. I.R.E., Vol. 40, pp. 1541-1558; November 1952.

Appendix

Al.0 Derivation of Source and Termination Error-Factor Formulas for z_{11}

Al.1 Measured value of z_{11}

The measured value of z_{11} may be expressed as in Fig. 2(c):

$$\frac{\text{MEASURED INPUT IMPEDANCE}}{i_{1\text{CAL}}} = \frac{v_1}{i_{1\text{CAL}}}$$

$$= (\text{S.E.F.})(\text{T.E.F.})z_{11} \quad (\text{A1})$$

Using the terms of Fig. A2,

$$\frac{v_1}{i_{1\text{CAL}}} = \left(\frac{i_1}{i_{1\text{CAL}}} \right) (A)z_{11} \quad (\text{A2})$$

where

$$\text{S.E.F.} = \frac{i_1}{i_{1\text{CAL}}}, \quad \text{T.E.F.} = A, \quad \text{Az}_{11} = \frac{v_1}{i_1} \quad (\text{A3})$$

Al.2 Derivation of Termination Error Factor for z_{11} (T.E.F. = A)

From the circuit equations of Fig. 1,

$$\frac{v_1}{i_1} = z_{11} + z_{12} \frac{i_2}{i_1} \quad (\text{A4})$$

and

$$\frac{i_2}{i_1} = h_{21} + h_{22} \frac{v_2}{i_1} \quad (\text{A5})$$

From Fig. 2(a),

$$i_2 = -\frac{v_2}{z_T} \quad \text{or} \quad v_2 = -i_2 z_T \quad (\text{A6})$$

Substituting equation (A6) into equation (A5) gives

$$\frac{i_2}{i_1} = \frac{h_{21}}{1+h_{22}z_T} \quad (\text{A7})$$

To obtain equation (A7) in terms of z's use the following relations from Fig. A1:

$$h_{21} = -\frac{z_{21}}{z_{22}} \quad (\text{A8})$$

$$h_{22} = \frac{1}{z_{22}} \quad (\text{A9})$$

Substituting equations (A8) and (A9) into equation (A7) gives

$$\frac{i_2}{i_1} = \frac{-z_{21}}{z_{22}+z_T} \quad (\text{A10})$$

Substituting equation (A10) into equation (A4) gives

$$\frac{v_1}{i_1} = z_{11} + z_{12} \left[\frac{-z_{21}}{z_{22}+z_T} \right] \quad (\text{A11})$$

Factoring out z_{11} gives

$$\frac{v_1}{i_1} = z_{11} \left[1 - \frac{z_{21}z_{12}}{z_{11}z_{22} + z_T z_{22}} \right] \quad (\text{A12})$$

Defining τ as in Fig. A1,

$$\tau = \frac{z_{21}z_{12}}{z_{11}z_{22}} \quad (A13)$$

Substituting equation (A13) into equation (A12) gives

$$\frac{v_1}{i_1} = \left[1 - \frac{\tau}{1 + \frac{z_T}{z_2}} \right] z_{11} = Az_{11} \quad (A14)$$

The termination error factor is

$$\text{T.E.F.} = A$$

$$= 1 - \frac{\tau}{1 + \frac{z_T}{z_{22}}} \quad (\text{Compare Fig. A2}) \quad (A15)$$

A1.3 Derivation of Source Error Factor for z_{11}

$$\left(\text{S.E.F.} = \frac{i_1}{i_{1\text{CAL}}} \right)$$

From Fig. 2(b),

$$i_{1\text{CAL}} = \frac{z_s}{z_s + z_{11\text{CAL}}} i_s \quad (A16)$$

and

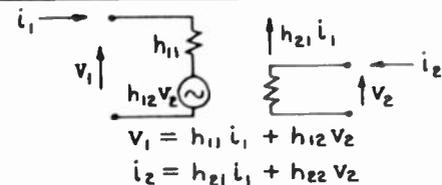
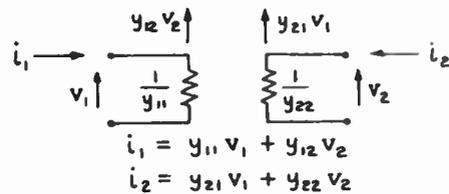
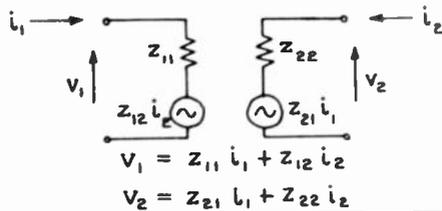


Fig. 1

Four-pole z , y , and h equivalent circuits and equations.

$$i_1 = \frac{z_s}{z_s + \frac{v_1}{i_1}} i_s \quad (A17)$$

Therefore

$$\frac{i_1}{i_{1\text{CAL}}} = \frac{z_s + z_{11\text{CAL}}}{z_s + \frac{v_1}{i_1}} \quad (A18)$$

Dividing by z_s ,

$$\frac{i_1}{i_{1\text{CAL}}} = \frac{1 + \frac{z_{11\text{CAL}}}{z_s}}{1 + \frac{\frac{v_1}{i_1}}{z_s}} \quad (A19)$$

Substituting equations (A3) into equation (A19) gives the source error factor:

$$\begin{aligned} \text{S.E.F.} &= \frac{i_1}{i_{1\text{CAL}}} \\ &= \frac{1 + \frac{z_{11\text{CAL}}}{z_s}}{1 + \frac{Az_{11}}{z_s}} \quad (\text{Compare Fig. A2}) \end{aligned} \quad (A20)$$

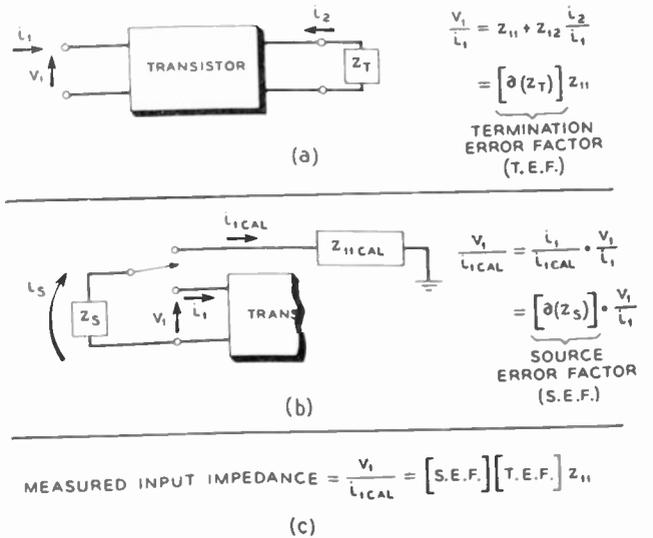


Fig. 2

Errors in measurement of z_{11} due to finite terminations.

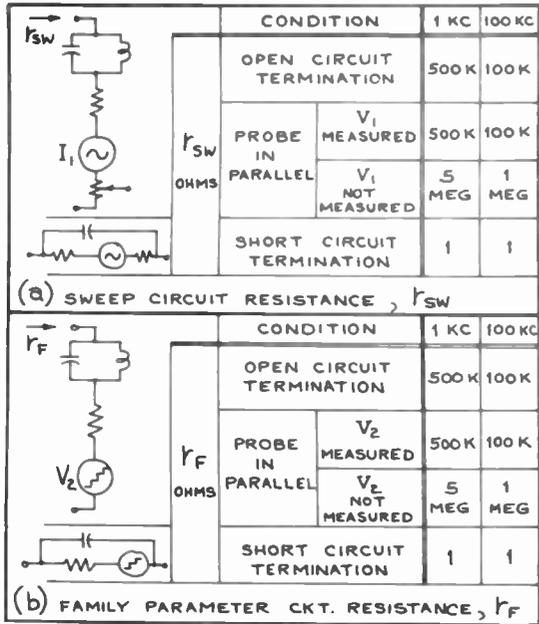


Fig. 6

Building-block impedance values for low frequency z, y, and h sweepers.

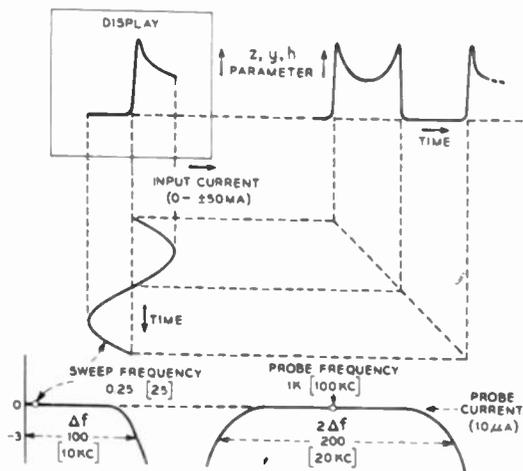


Fig. 3

Sweep measurement objectives for low frequency z, y, and h parameters.

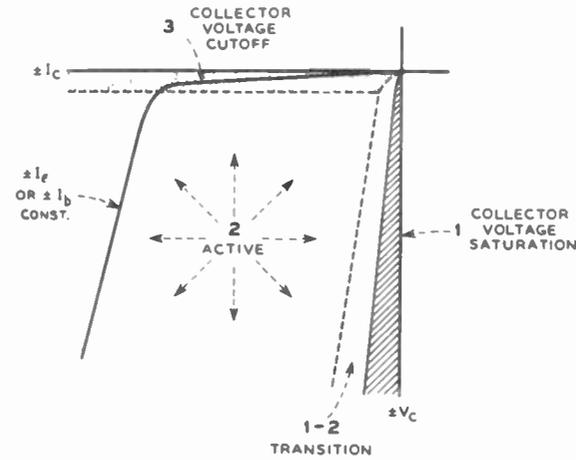


Fig. 7

A regional division of the transistor characteristic.

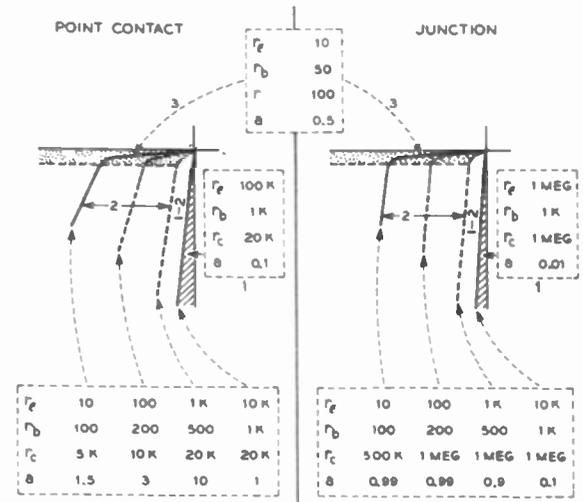


Fig. 8

Assumed point-contact and junction-transistor models.

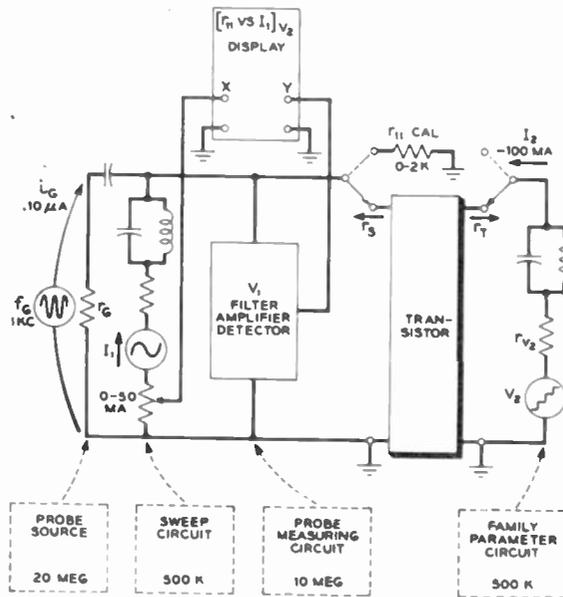


Fig. 4

Arrangement of building blocks for an r_{11} vs. I_1 v_2 sweeper (with indicated impedance levels at 1 kc).

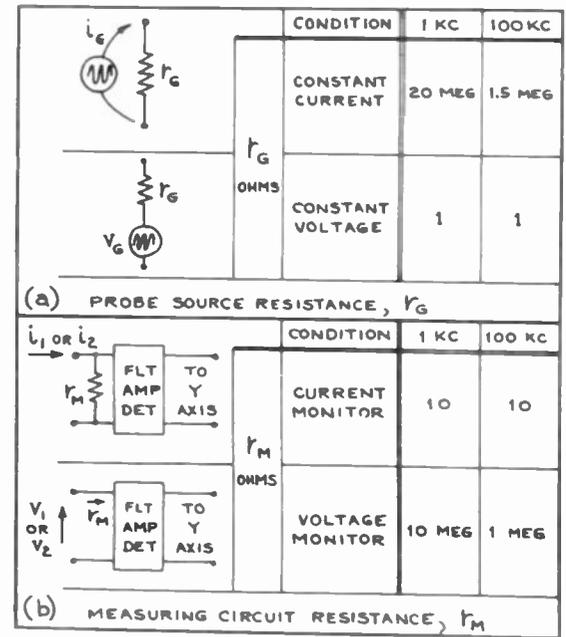


Fig. 5

Building-block impedance values for low frequency z, y, and h sweepers.

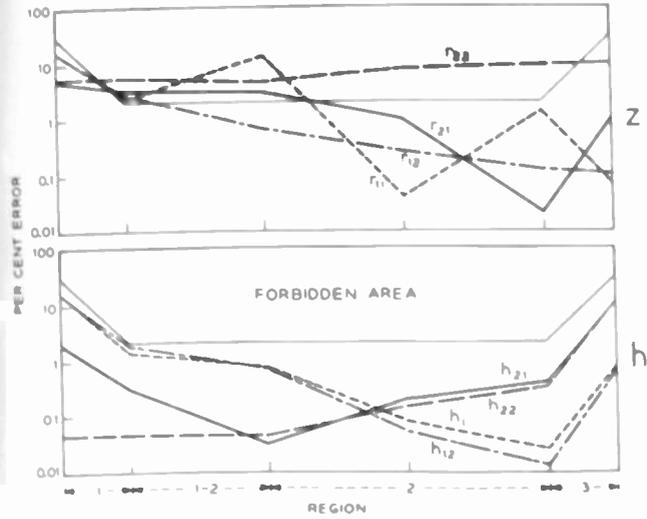


Fig. 9
z, y, and h error comparisons for the common-base point-contact transistor at 1 kc.

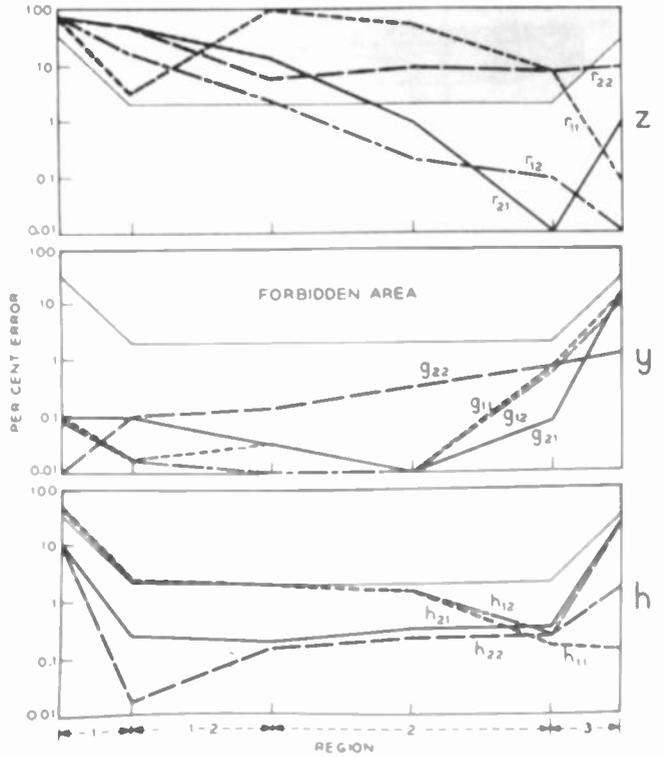


Fig. 11
z, y, and h error comparisons for the common-emitter junction transistor at 1 kc.

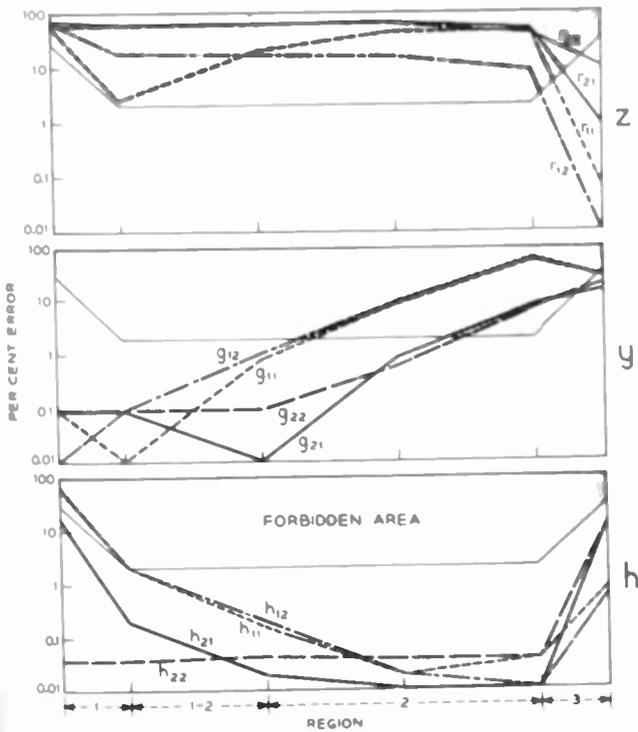


Fig. 10
z, y, and h error comparisons for the common-base junction transistor at 1 kc.

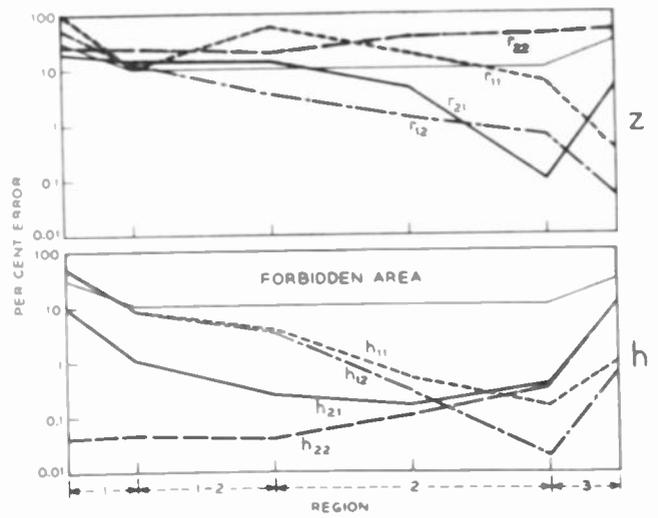


Fig. 12
z, y, and h error comparisons for the common-base point-contact transistor at 100 kc.

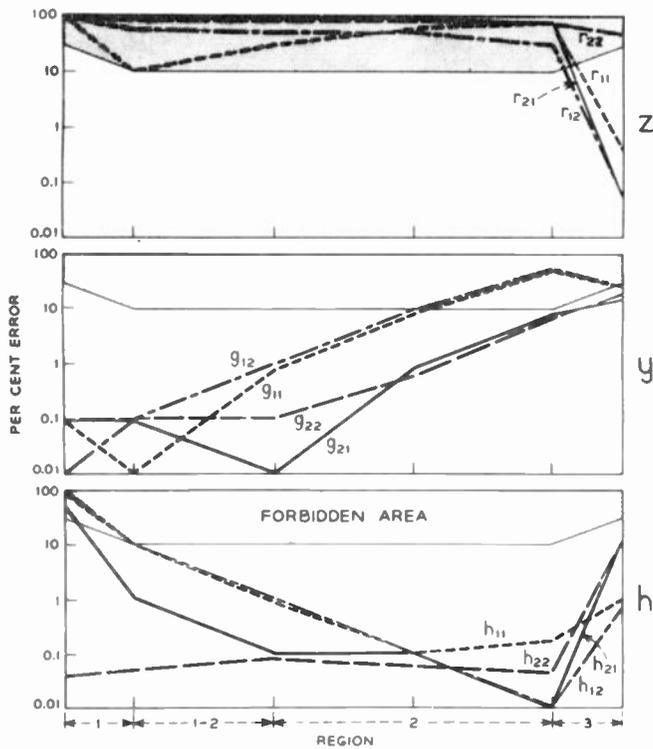


Fig. 13
z, y, and h error comparisons for the common-base junction transistor at 100 kc.

REG.	f _{KC}	CONN.	TYPE	r ₁₁	r ₁₂	r ₂₁	r ₂₂	g ₁₁	g ₁₂	g ₂₁	g ₂₂	h ₁₁	h ₁₂	h ₂₁	h ₂₂
1-2 AND 2	1	CB	PC												
			JCN												
		CE	JCN												
3	100	CB	PC												
			JCN												
1	100	CB	PC												
			JCN												
		CE	JCN												

Fig. 14
Transistor parameter selection chart for low frequency sweep measurements.

Z ₁₁	1/y ₁₁ (1-τ)	h ₁₁ /1-τ
Z ₁₂	-y ₁₂ /y ₁₁ y ₂₂ (1-τ)	h ₁₂ /h ₂₂
Z ₂₁	-y ₂₁ /y ₁₁ y ₂₂ (1-τ)	-h ₂₁ /h ₂₂
Z ₂₂	1/y ₂₂ (1-τ)	-1/h ₂₂
1/z ₁₁ (1-τ)	y ₁₁	1/h ₁₁
-z ₁₂ /z ₁₁ z ₂₂ (1-τ)	y ₁₂	-h ₁₂ /h ₁₁
-z ₂₁ /z ₁₁ z ₂₂ (1-τ)	y ₂₁	h ₂₁ /h ₁₁
1/z ₂₂ (1-τ)	y ₂₂	h ₂₂ /1-τ
z ₁₁ /1-τ	1/y ₁₁	h ₁₁
z ₁₂ /z ₂₂	-y ₁₂ /y ₁₁	h ₁₂
-z ₂₁ /z ₂₂	y ₂₁ /y ₁₁	h ₂₁
1/z ₂₂	y ₂₂ /1-τ	h ₂₂
$\tau = z_{e1}z_{e2}/z_{11}z_{22} = y_{21}y_{12}/y_{11}y_{22} = 1/1 - \frac{h_{11}h_{22}}{h_{21}h_{12}}$		

Fig. A1
Relations between z, y, and h parameters.

MEAS-URED VALUE	SOURCE ERROR FACTOR (S.E.F.)	TERMINATION ERROR FACTOR (T.E.F.)	PARAMETER
$\frac{V_1}{I_{ICAL}}$	$\frac{I_1}{I_{ICAL}} = \frac{1 + \frac{Z_{1ICAL}}{Z_S}}{1 + \frac{AZ_{11}}{Z_S}}$	$A = 1 - \frac{\tau}{1 + \frac{Z_T}{Z_{22}}}$	Z ₁₁
$\frac{V_1}{I_{2CAL}}$	$\frac{I_2}{I_{2CAL}} = \frac{1 + \frac{Z_{2ICAL}}{Z_S}}{1 + \frac{DZ_{22}}{Z_S}}$	$B = 1 - \frac{1}{1 + \frac{Z_T}{Z_{11}}}$	Z ₁₂
$\frac{V_2}{I_{ICAL}}$	$\frac{I_1}{I_{ICAL}} = \frac{1 + \frac{Z_{2ICAL}}{Z_S}}{1 + \frac{AZ_{11}}{Z_S}}$	$C = 1 - \frac{1}{1 + \frac{Z_T}{Z_{22}}}$	Z ₂₁
$\frac{V_2}{I_{2CAL}}$	$\frac{I_2}{I_{2CAL}} = \frac{1 + \frac{Z_{2ICAL}}{Z_S}}{1 + \frac{DZ_{22}}{Z_S}}$	$D = 1 - \frac{\tau}{1 + \frac{Z_T}{Z_{11}}}$	Z ₂₂

Fig. A2
Error factor formulas for open circuit impedance (z) parameters.

MEASURED VALUE	SOURCE ERROR FACTOR (S.E.F.)	TERMINATION ERROR FACTOR (T.E.F.)	PARAMETER
$\frac{i_1}{V_{ICAL}}$	$\frac{V_1}{V_{ICAL}} = \frac{1 + y_{11CAL} Z_S}{1 + E y_{11} Z_S}$	$E = 1 - \frac{\tau}{1 + \frac{1}{y_{ee} Z_T}}$	y_{11}
$\frac{i_1}{V_{ECAL}}$	$\frac{V_2}{V_{ECAL}} = \frac{1 + y_{12CAL} Z_S}{1 + H y_{22} Z_S}$	$F = 1 - \frac{1}{1 + \frac{1}{y_{11} Z_T}}$	y_{12}
$\frac{i_2}{V_{ICAL}}$	$\frac{V_1}{V_{ICAL}} = \frac{1 + y_{e1CAL} Z_S}{1 + E y_{11} Z_S}$	$G = 1 - \frac{1}{1 + \frac{1}{y_{ee} Z_T}}$	y_{21}
$\frac{i_2}{V_{ECAL}}$	$\frac{V_2}{V_{ECAL}} = \frac{1 + y_{22CAL} Z_S}{1 + H y_{22} Z_S}$	$H = 1 - \frac{\tau}{1 + \frac{1}{y_{11} Z_T}}$	y_{22}

Fig. A3
Error factor formulas for short circuit admittance (y) parameters.

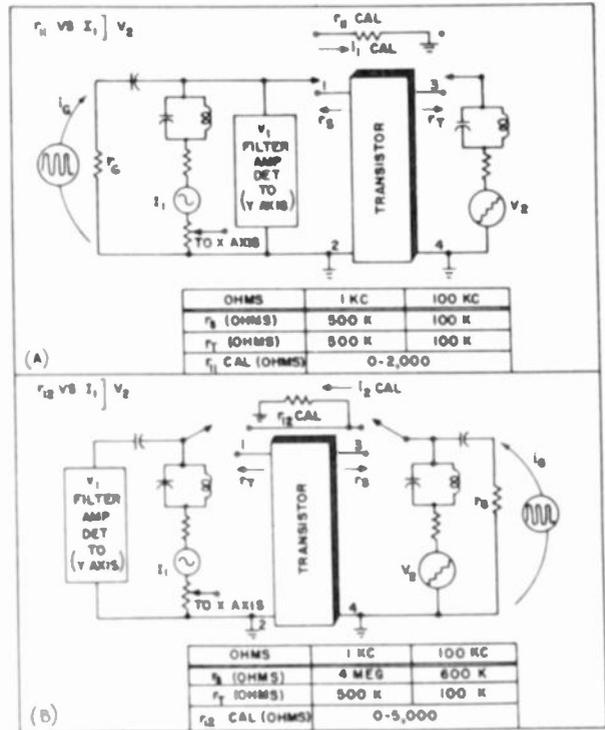


Fig. A5
Building-block arrangements for r_{11} and r_{12} sweepers.

MEASURED VALUE	SOURCE ERROR FACTOR (S.E.F.)	TERMINATION ERROR FACTOR (T.E.F.)	PARAMETER
$\frac{V_1}{i_{ICAL}}$	$\frac{i_1}{i_{ICAL}} = \frac{1 + \frac{h_{11CAL} Z_S}{Z_S}}{1 + \frac{J h_{11}}{Z_S}}$	$J = 1 - \frac{\tau - 1}{1 + \frac{1}{h_{22} Z_T}}$	h_{11}
$\frac{V_1}{V_{ECAL}}$	$\frac{V_2}{V_{ECAL}} = \frac{1 + \frac{Z_S}{Z_T}}{1 + M h_{22} Z_S}$	$K = 1 - \frac{1}{1 + \frac{Z_T}{h_{11}}}$	h_{12}
$\frac{i_2}{i_{ICAL}}$	$\frac{i_1}{i_{ICAL}} = \frac{1 + \frac{Z_T}{Z_S}}{1 + \frac{J h_{11}}{Z_S}}$	$L = 1 - \frac{1}{1 + \frac{1}{h_{22} Z_T}}$	h_{21}
$\frac{i_2}{V_{ECAL}}$	$\frac{V_2}{V_{ECAL}} = \frac{1 + h_{22CAL} Z_S}{1 + M h_{22} Z_S}$	$M = 1 - \frac{\tau - 1}{1 + \frac{Z_T}{h_{11}}}$	h_{22}

Fig. A4
Error factor formulas for hybrid (h) parameters.

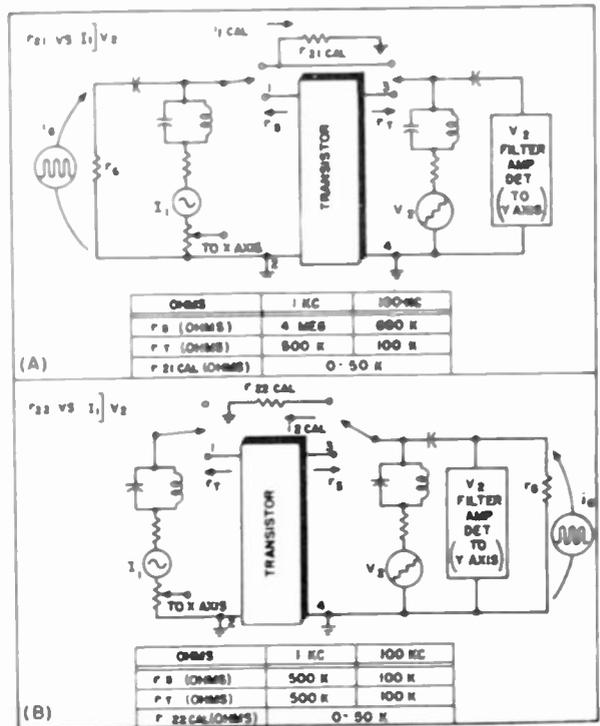


Fig. A6
Building-block arrangements for r_{21} and r_{22} sweepers.

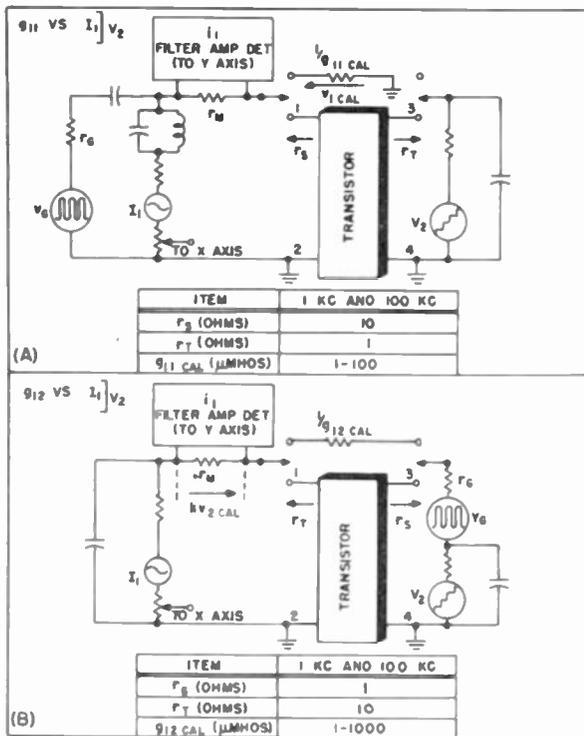


Fig. A7
Building-block arrangements
for g_{11} and g_{12} sweepers.

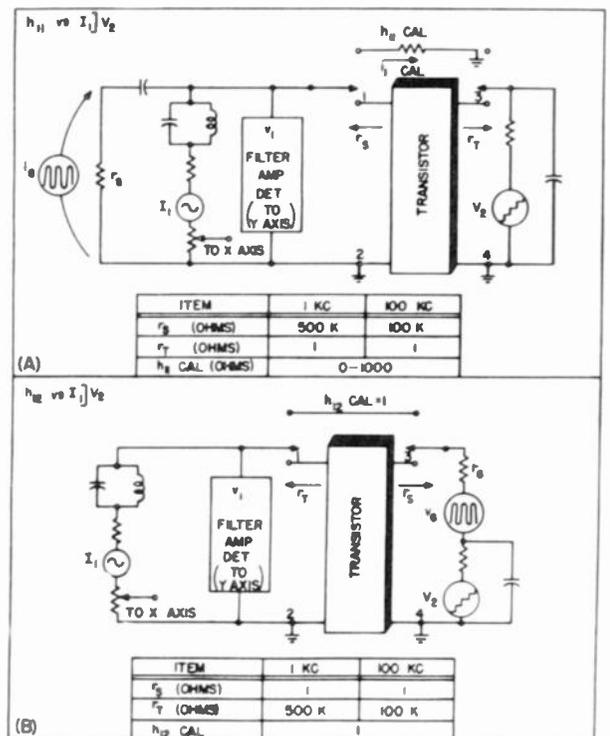


Fig. A9
Building-block arrangements
for h_{11} and h_{12} sweepers.

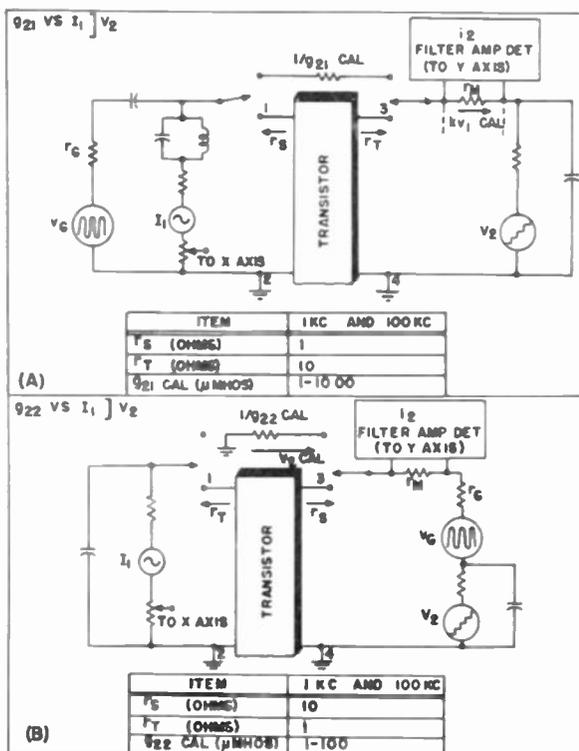


Fig. A8
Building-block arrangements
for g_{21} and g_{22} sweepers.

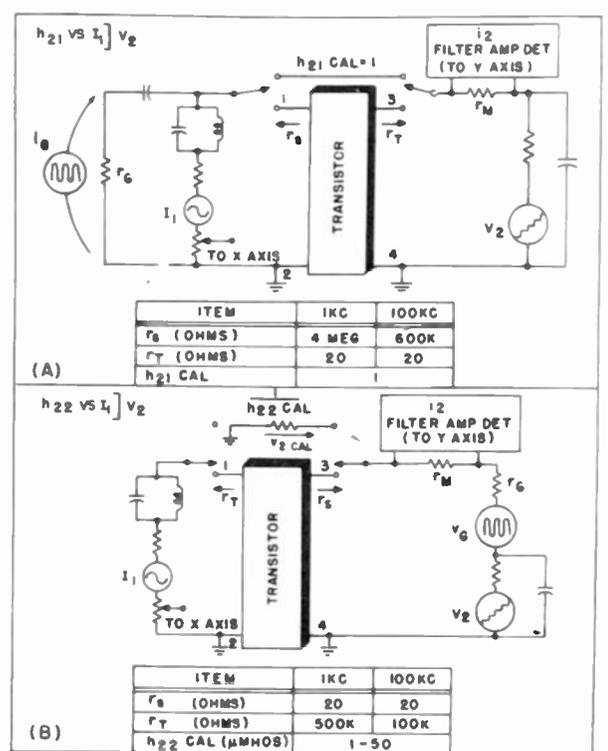


Fig. A10
Building-block arrangements
for h_{21} and h_{22} sweepers.

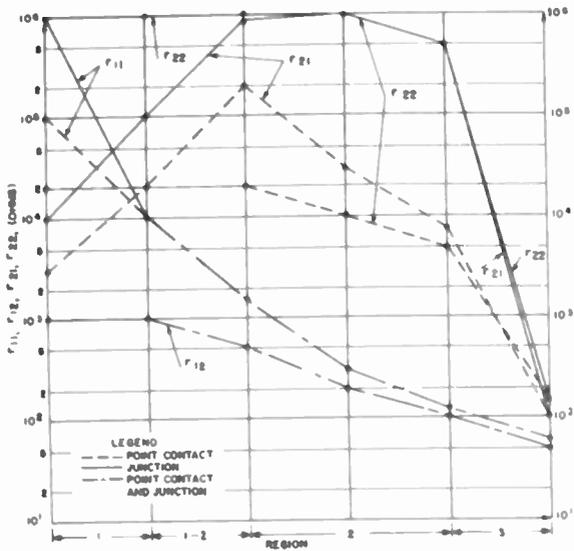


Fig. A11
Grounded base r parameters of the transistor models.

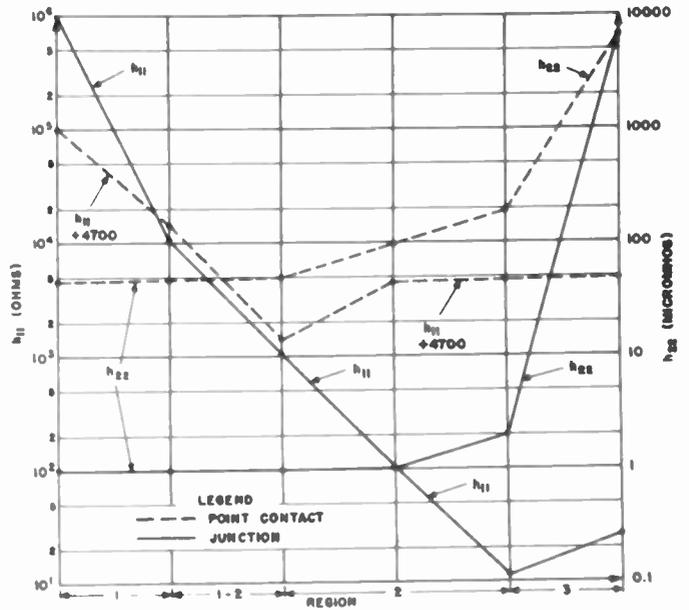


Fig. A13
Grounded base h_{11} and h_{22} parameters of the transistor models.

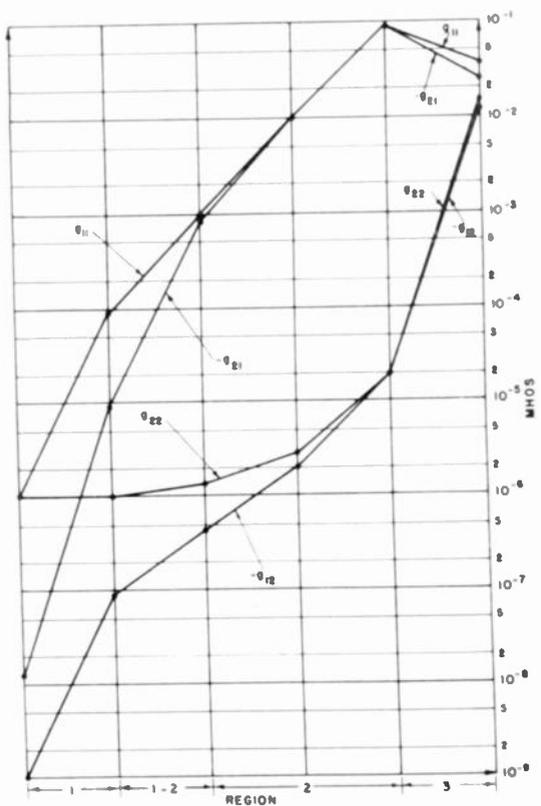


Fig. A12
Grounded base g parameters of the junction transistor model.

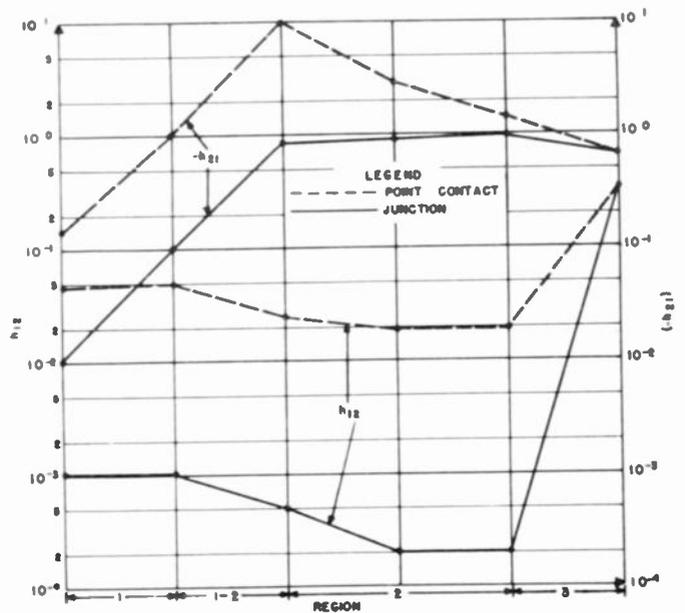


Fig. A14
Grounded base h_{12} and h_{21} parameters of the transistor models.

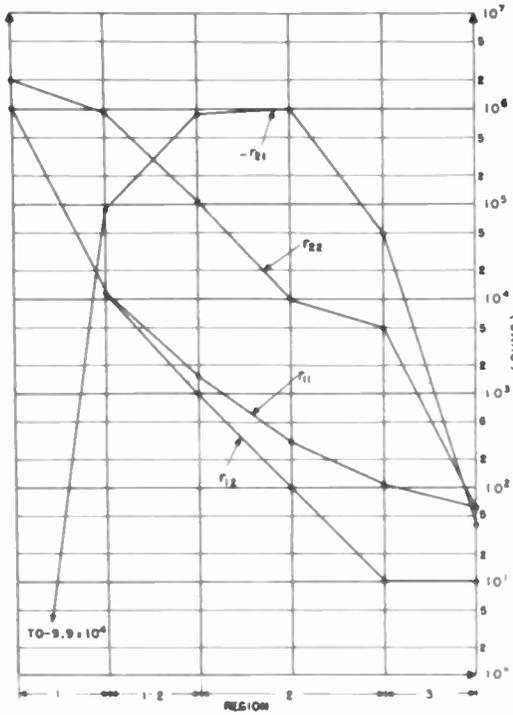


Fig. A15
Grounded emitter r parameters of the
junction transistor model.

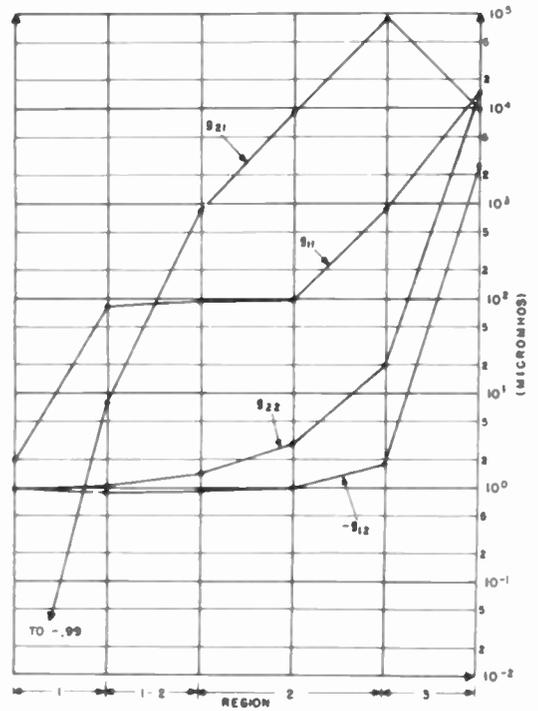


Fig. A16
Grounded emitter g parameters of the
junction transistor model.

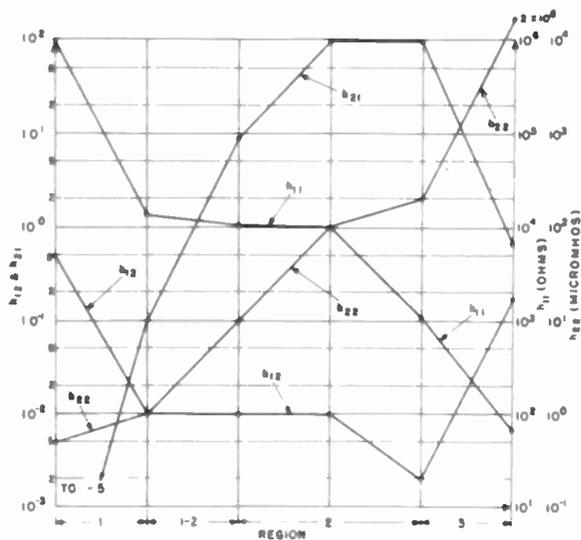


Fig. A17
Grounded emitter h parameters of the
junction transistor model.

THE METRECHON - A NEW HALF TONE PICTURE STORAGE TUBE

by

L. Pensak

RCA Laboratories Division
Radio Corporation of America
Princeton, N.J.

The Metrechon is new because, although it has been under development for some time and operating samples have been built, it is not yet in production and this is the first public statement describing it. It is a half-tone picture storage tube because it can store a radar or television picture with a continuous range of signal values (half tones) and reproduce them in the output. It is a storage tube because, according to a proposed IRE definition, it does this by means of static electric charges. The name is also new and was chosen by the convention of a descriptive name based on Greek words. They are, 'metre' which means 'to measure', and 'echo' which means 'to keep or to hold'. The 'n' is the residue of an attempt to tack on the old suffix 'tron'.

A schematic diagram of the tube is shown in Fig. 1. It consists of a low velocity gun similar in principle to the type used in the image orthicon, a storage target and a conventional cathode ray gun. The low velocity gun is used for the reading operation and its cathode is tied to ground. The cathode ray gun is used for the writing operation and its cathode is operated at about 2000 volts negative with respect to ground. The target is shown enlarged below the tube outline. It consists of a thin sheet of insulator with a metal mesh on one surface. There are several ways of building such a target. One way is to metalize a mesh on a sheet of mica. Another is to seal a thin film of high resistivity glass to a sheet of electrolytic mesh. Because the fineness of the mesh determines the maximum number of storage elements on a target of limited diameter, a 700 wire per inch mesh was used. The metal side of the target is faced toward the reading gun and is called the front surface. Within a few thousandths of an inch from the back, a second fine mesh screen of metal is located and is called the "writing screen". The bulb wall back of the writing screen is coated with an electrical conducting film which serves to collect the secondary electrons that emerge thru the writing screen and also to act as an anode to the writing gun. The remaining electrodes are conventional. The magnetic focussing and deflecting coils around the reading gun have been described in the literature and so are not shown. Fig. 5 is a photograph of a sample tube and its mechanical dimensions are 23 inches long and 2 inches maximum diameter.

Fig. 2 is a drawing of the target very much enlarged and shows the zero equipotential lines under different conditions when the mesh is held at about 2 volts positive with respect to the reading gun cathode. There is some electron

landing on the mesh. However those electrons that land on the insulator, because of their low velocity, do not produce any secondary electrons and therefore drive it negative till no further landing can take place and are subsequently totally reflected. The reflected electrons go back along the line of deflection to a secondary emission multiplier section which amplifies the current. The output of the multiplier produces a signal which is proportional to the magnitude of the current reflected from the target.

The writing operation will be such as to drive the insulator more negative than zero. In this case the zero equipotential lines emerge from the surface and start to cut down the area thru which the electrons can approach the metal. This increases the fraction of the beam current that is reflected. When the insulator is driven still more negative the lines emerge further and close over to prevent any landing on the mesh. This causes full reflection of the reading beam into the multiplier and produces a peak or saturated signal. Fig. 3 also shows how the beam current varies with insulator potential. A typical transfer characteristic is also shown where the input signal is shown in terms of charge density on the target. Because the usual cathode ray gun does not have a linear beam current with grid voltage, correction must be made for it. Approximately 70 percent of the curve is linear when plotted as shown.

It should be noted that at no time, in the presence or absence of a stored charge is there any electron landing on the insulator. Therefore the reading action is one by which the stored charge is measured without disturbing it. This is the reason for the name containing the word "metre". It is also the reason why the target can be scanned repeatedly to reproduce many copies of output signals with the full range of half tones.

The writing operation can be performed by applying the input signals to the grid of the writing gun and thereby modulating its beam current from zero to any value up to its maximum. During the time of writing the writing screen at the back of the target is held at 100 volts negative with respect to ground. Because of the close spacing there is created a retarding field of sufficient strength to effectively prevent the escape of any secondary electrons from the point of landing of the beam. The target is thus driven negative by an accumulated charge which is the product of the beam current and the time the beam takes to pass a given spot. The negative potential thus produced on the

back of the insulator couples by the high capacity of the very thin insulator film to the front surface to affect the reading as described earlier. The potential required to cut off reading beam landing is no more than five volts. This is a small fraction of the retarding voltage on the writing screen and therefore provides a linear build up of charge with either time, or current, or number of passes. The tube will therefore integrate linearly any number of accumulated signals until the reading action is saturated.

The writing speed is very high, being of the order of 10 million spot diameters per second. This is due to the low capacity of the storage elements which are essentially supported at their edges instead of being backed up by metal as is the case with most previous storage tubes. This permits an effective reduction in capacitance per unit area to one tenth that of a corresponding metal backed insulator of the same thickness. The voltage excursion is also small so that a saturated writing of a television picture can be obtained in a single scan with a peak beam current of no more than three microamperes. This speed is ample for most television or radar applications and the writing gun was chosen for high resolution and a peak beam current of 10 microamperes. For special applications where higher speeds are required it is possible to use different writing guns with 50 times the peak beam current.

The erasing operation can be performed by switching the writing screen to zero volts and hitting any spot with the writing beam to give point for point erasure. The beam voltage was chosen to produce an excess of secondary electrons for the material of the storage target. If the bombarded spot was negative due to writing, then the secondaries are drawn to the writing screen until the area is driven just enough positive to pull back the excess secondaries and no further potential changes occur. This is the normal erased state of the back of the target. It is obvious that the writing and erasing operations cannot be done simultaneously. Therefore erasing must be timed to occur when no new writing signals are occurring. In a radar system this can be the recovery time of the transmitter. This is also the interval during which the deflection returns to the starting point. Therefore the writing beam can be turned on to a fixed value and deflected along a line where the new signals will be written and clear it for them. Thus the whole radar pattern will be visible in the reading output except where it is about to be replaced. Because the writing and reading operations can be performed without interfering with the reading, the output is continuously available from the instant a signal is written in to the time it is erased.

The resolution of the tube is shown by Fig. 4. It was obtained by writing down a bar pattern generated by an oscillator synchronized with a television scan and written in a single frame. Each frequency shows the corresponding number of lines per target diameter. We see that

100 percent modulation of the reading output signal is obtainable for almost 300 lines per target diameter and that 500 lines still gives almost 10 percent modulation. The corresponding resolution number of a television picture is lower because not all of the target area is utilized and is about 370 lines. There is reason to believe that the resolution is limited by the fineness of the target mesh. Finer mesh has been made and will be used in future tubes when and if necessary.

The storage time of the tube can have several possible meanings. The most useful is the time during which the reading operation can produce useful copies. With a television type reading it was found that the storage time was about 10 to 30 minutes before a fully written signal would decay to half value. This limit was set by the ionization of the traces of gas remaining in the tube by the reading beam current and causing positive ions to land on the target and to erase the negative areas. The effect was stronger in the center of the target than around the edges. If the reading operation can be interrupted then the storage time can be measured in terms of the number of copies. For this purpose the number is 36,000 to 100,000 copies. In terms of the insulator's ability to hold charge it was found that appreciable signal remained days after it was written.

As for applications of the tube, it was primarily developed for the purpose of storing radar PPI patterns and displaying them in television scan on a kinescope. Tests of this type of operation are under way at the present time and it is hoped that there will be found a better detection of signal in noise due both to the constancy of visible signals in the display and to the fuller utilization of the effects of integration. Other applications will be investigated as tubes become more available.

Although the tests already made indicate the likelihood that the tube could be an elegant solution to many systems problems it would be well to recognize that much work remains to be done in learning how to operate it under the necessary conditions. It is a complex device and will require a good bit of circuitry to develop for its specific operation. In any case it will be merely a link between a complex writing circuit and a television type or other reading system. Thus its use is expected to be restricted to those applications where its special properties will justify the expense and complexity of its installation. It is not expected that it will be of use to the computer field because of the problem of obtaining sufficiently accurate scanning circuits to make full use of its storage capabilities. A more detailed description of its operation in various modes will appear in publications for those who wish to understand the tube better. As for the availability of samples for testing, the tube is not yet in production and is not scheduled for the time being. We know of some design changes that must be made and expect to find more in our tests before the tube will be ready for production.

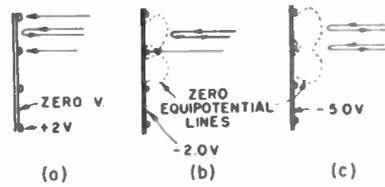
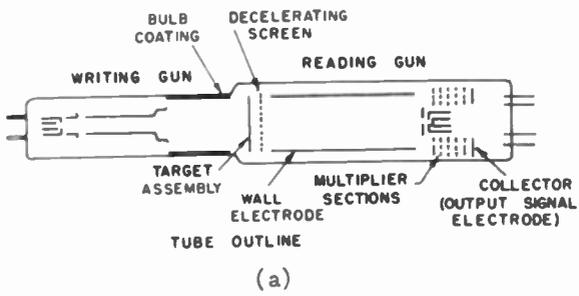


Fig. 2 - Equipotential diagrams.

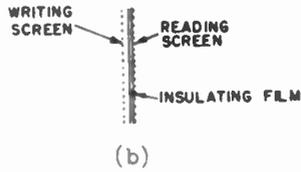


Fig. 1

(a) tube outline; (b) target cross section.

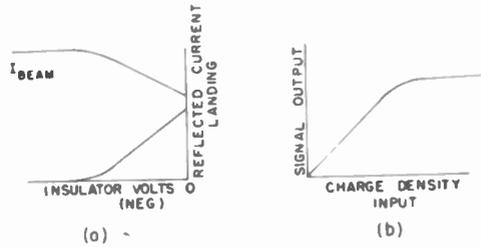


Fig. 3

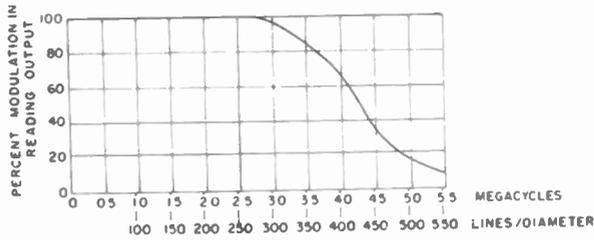


Fig. 4 - Resolution curve.

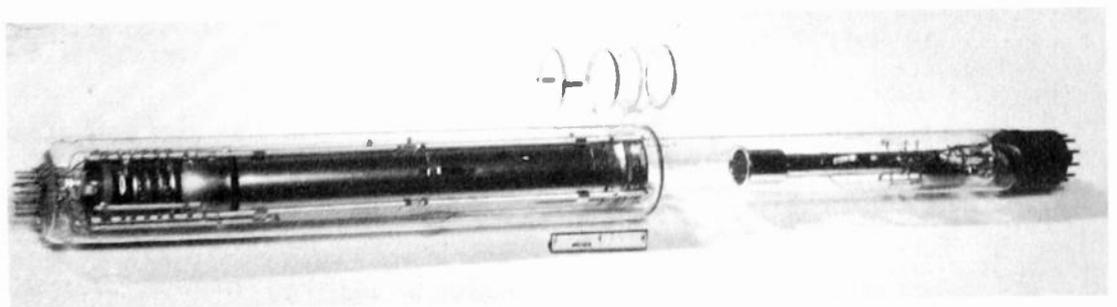


Fig. 5

CHARACTERISTICS OF VIEWING STORAGE TUBES WITH HALFTONE DISPLAY

M. Knoll, H.O. Hook, and R.P. Stone
RCA Laboratories Division
Princeton, New Jersey

ABSTRACT

Typical characteristics of transmission control type viewing storage tubes are discussed, using as an example a new version of halftone storage tube with primary current modulation for writing, and grid-control reading. This tube has 400 line resolution, a highlight brightness of 400 feet-Lamberts at 7 kv anode voltage, and is able to store single television halftone frames up to 15 seconds without holding.

Calculated and experimental static viewing-current characteristics are compared, showing a

considerable decrease in slope with increasing gradient across the storage layer. Dynamic characteristics of the visual output as a function of the electrical input are approximately linear. To keep the amplification factor sufficiently constant variations in storage, grid-wire diameter must be $< \pm 1$ per cent for accurate halftone display. Viewing durations without holding are mainly limited by the residual gas pressure. Using pulse-holding techniques, black-and-white pictures have been viewed for 27 hours.

A HIGH WRITING SPEED DARK TRACE TUBE

S. Nozick, N. H. Burton, and S. Newman

Storage Tube Group
U.S. Naval Material Laboratory
Brooklyn, New York

Summary.

As an outgrowth of several years of dark trace tube study, an improved skiatron is presented with a writing speed better than three times that previously obtainable. The improvement was obtained by special electron optical design rather than by raising the anode voltage. The analysis of skiatron writing speed qualities is presented as a function of electron gun and screen characteristics, with experimental verification. The special design characteristics arrived at are presented and analyzed in view of the laboratory study. The tube and system are discussed and contrasted with previous dark trace tubes.

Introduction

Dark trace, or skiatron, tubes have certain inherent advantages over bright display tubes which make their use desirable for information displays. These advantages are: the ability to integrate visually, the ability to retain information for an extended period of time, and the ability to present optimum contrast with high ambient illumination. The last named, the daylight viewing characteristic of dark trace tubes, in conjunction with the storage characteristic, explains the interest in skiatrons. One limitation to the utilization of the skiatron tube has been the low information display rate. A tube is described herein which greatly increases the display rate. A special electron optical system is presented which reduces the deflection defocusing encountered when high beam currents are employed.

Analysis

After basic analysis of the problem of the low information display rate of dark trace tubes, it was apparent that an appreciable improvement was required. A search of the available literature revealed that a detailed study of the writing speed of skiatrons had not as yet been reported. Therefore a study of the characteristics affecting the writing speed was made. It was found that changes in the accelerating potential, the beam current, or the screen sensitivity would influence the writing speed.

The first possibility, suggested by extrapolating the work of previous researchers in the field, was to increase the accelerating potential. However, limitations to this technique were immediately apparent. The magnitude of the possible increase in the writing speed due to increased potential would be limited. In addition, the insulation and corona problems associated with potentials higher than fifteen kilovolts limit any proposed tube's application.

The results of Nottingham's work¹ indicated

that a relationship exists between contrast and the charge density deposited on the skiatron screen. Based upon this relationship and the work of Windsor² an analysis was prepared and experimental verification obtained of the relationship of writing speed, beam current, and spot size³. The formulation arrived at is:

$$W = K I/s \quad (1)$$

where: W is the writing speed
I is the beam current incident on the screen
s is the spot size
K is a constant determined by the screen sensitivity

This equation led to the information display rate, which is equal to the writing speed divided by the spot size. Thus:

$$F = W/s = K I/s^2 \quad (2)$$

where: F is the information display rate
This is a more accurate formulation for comparing different tubes since it is actually the maximum number of elements of information that can be displayed per unit time.

An increase in the writing speed could be attained by increasing the factor K in the above formula. This, however, would involve a modification in the basic screen material, entailing a long study of the crystal state of the scotophor and the effect of the variations in the composition of the scotophor on the formation of the F centers. Available information indicates that an increase in the sensitivity of the pure potassium chloride screen would be very difficult and here again, the possible increase would probably not be sufficient to warrant the effort.

The remaining factor in the writing speed equation, the ratio of beam current to spot size, showed the most promise. Our objective, therefore has been to increase the ratio of the screen current of a dark trace tube to the spot size while maintaining satisfactory resolution.

Description of Tube

To attain our desired objective a dark trace tube, ML-9211, was constructed by Dr. Holborn of the National Union Radio Corporation, containing a wide aperture tetrode gun with the low heat capacity mica supported screen of the National Union Radio Corporation. The gun is capable of approximately ten milliamperes emission through the wide aperture system.

This tube was studied to determine if it satisfied the basic requirement of the problem, an

improvement in the writing speed. The writing speed was measured by writing a single raster on the scotophor and examining it. The vertical deflection waveform was a sawtooth and the horizontal waveform a sinusoid. The frequency of the sinusoid was increased until the trace just disappeared at the vertical centerline. The writing speed was then calculated by measuring the width of the raster, noting the frequency of the sinusoid, and substituting these values into the formula:

$$W = 2 Kfa \quad (3)$$

where: W is the writing speed in centimeters per second
 f is the frequency of the sinusoid in cycles per second
 a is the amplitude in centimeters
 (The small vertical velocity component can be neglected)

The results of this study were not completely satisfactory. The writing speed, at one milliamper beam current, was 7.1 kilometers per second as compared to a previously determined value of 5 kilometers per second for the standard R-2112-G, or an increase of about 40 percent. No figures were attainable at higher beam currents due to severe defocussing.

An increase in the beam current beyond one milliamper was accompanied by increased deflection defocussing. To obtain a quantitative measure of this phenomenon the spot size was measured as the beam current was increased. It was found that a comparison of the standard spot size "A"⁴ of the shrinking raster method, the line width at the center of the raster, with spot size "D" produced a measure of the deflection defocussing. Spot size "D" is defined as the line width determined by shrinking a fifty line raster until the top two lines merged. The results of these measurements, (Graph 2), and those of the spot size of the R-2112-G, confirmed the observation that the increasing spot size of the ML-9211 restricted it to operation with beam currents of one milliamper or less. The criterion for this limit was taken to be a spot size of one millimeter, which is comparable to that of the P-7 cascade screens currently in use.

Electron Optics

The work of Schlesinger⁵ reveals that the basic cause of this type of deflection defocussing is the large diameter of the electron beam in the field of the deflection coils. With this as a guide, a lens system was designed to reduce the diameter of the beam entering the field of the deflection coils.

The lens system is composed of a primary, short focal length lens and a secondary, long focal length lens. The primary lens, placed on the neck of the tube just forward of the gun structure (Fig. 1) converges the beam at a point just before the field of the secondary lens. The secondary lens, in turn, refocusses the beam at the scoto-

phor. Comparing this system to the conventional focus lens (Fig. 2), we see that the diameter of the beam has been reduced by producing an effective point electron source much closer to the focus coil.

The primary lens field, calculated from the geometric character of the tube employed, is approximately four kilogauss, as measured in the air gap with a kilogauss meter. Since this field magnitude cannot be produced by a practical electromagnetic lens, an Alnico V magnet was magnetized to a stable condition having a variation of from three to six kilogauss, by mechanical variation of the air gap dimension.

The secondary lens is a standard 2D2 electromagnetic focus coil. This lens acts in its usual capacity to focus the reduced diameter beam at the screen.

The initial attempts to employ the system indicated a substantial reduction of the deflection defocussing found at two milliamperes beam current. The position and field strength of the primary lens were then varied to determine the optimum operating conditions. The graph of the results (Graph 1) indicates the position and field for minimum deflection defocussing to be at 12.5 centimeters from the tube reference line and a field strength of 5 kilogauss. These values were found to yield good results for all beam currents employed. The writing speed of the ML-2911 employing the two lens focus system was then determined.

Results

The results of the various spot size determinations are summarized in graphical form (Graph 2). It will be noted that, while the ML-9211 and the R-2112-G have equivalent spot sizes "A" and "D" for beam currents up to nine hundred microamperes, the ML-9211 has much more severe deflection defocussing for higher beam currents. The aperture limiting of the R-2112-G is apparent when compared to the ML-9211 at these higher beam currents. The improvements obtained with the two lens focussing system are verified by the displacement of the spot size curves of the ML-9211 employing the system. The curves are displaced appreciably in the direction of higher beam current. The spot size "A" is maintained below .5 millimeters while the spot size "D" is practical for beam currents up to two and a half milliamperes.

The results of the writing speed measurements are shown in Graph 3. The R-2112-G is capable of a usable writing speed as fast as five kilometers per second. The ML-9211, employing the conventional focus system, attains a maximum of seven kilometers per second, while the two lens focus system raises this value to fifteen kilometers per second. Thus the combination of the ML-9211 and the two lens system provides a considerable improvement as compared to the R-2112-G. The final figure attained is an improvement of more than three hundred percent.

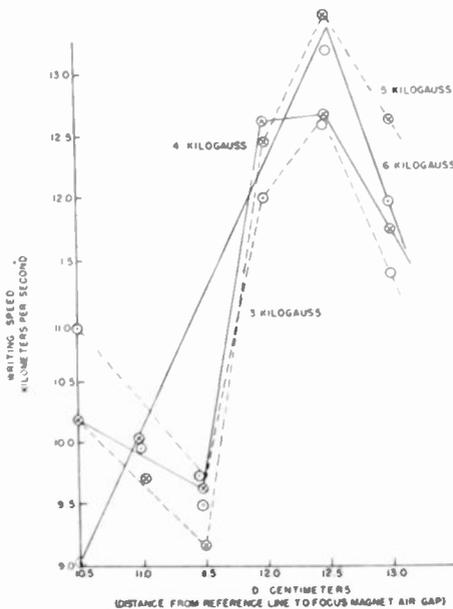
The information display rate (Graph 1), which provides a good figure of merit, varies from a maximum of 1.37×10^7 elements per second at 500 micro-amperes beam current for the R-2112-G to 2.30×10^7 elements per second at 1000 microamperes for the ML-9211, and to approximately 3.05×10^7 elements per second at 1800 microamperes for the ML-9211 with the two lens system. This is an improvement of 124 percent.

Conclusions

It has been demonstrated that it is possible to employ high beam currents in cathode ray tubes, without appreciably increasing the spot size, and in this manner, increasing the maximum writing speed. With the aid of this development the uses of the dark trace tube may now be extended to encompass a wider area in the field of radar and oscillography. In oscillography it will now be possible to obtain recordings of faster non-periodic phenomena than could previously be recorded. The two lens system insures that the trace will be sharp and clear, and enables the display of more rapid phenomena.

For radar applications the dark trace tube now comes closer to having the high writing speed characteristics of bright trace tubes, with the additional advantages of the scotophor. Thus, in many special radar applications, the dark trace tube will now be an improvement over the phosphor tubes. With its integration abilities and daylight viewing it can prove most useful in the fields of signal to noise enhancement, where retention of the signal is required.

The two lens focussing system itself is applicable to any cathode ray tube which is to be



Graph 1
Variation of writing speed with D at various values of flux density of the ML-9211 (double focus).

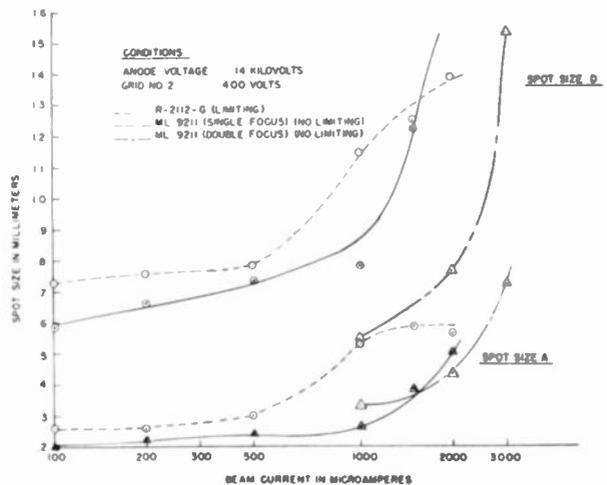
operated at high beam currents. It is useful where extremely small spot size is required and aperture limiting is not feasible. In all cases, the two lens system can be installed without radically altering the circuitry since the only addition to existing electromagnetic focus systems is the primary lens, a permanent magnet. It may enable the construction of shorter cathode ray tubes for television by reducing the deflection defocussing found when wide deflection angles are used.

Acknowledgement

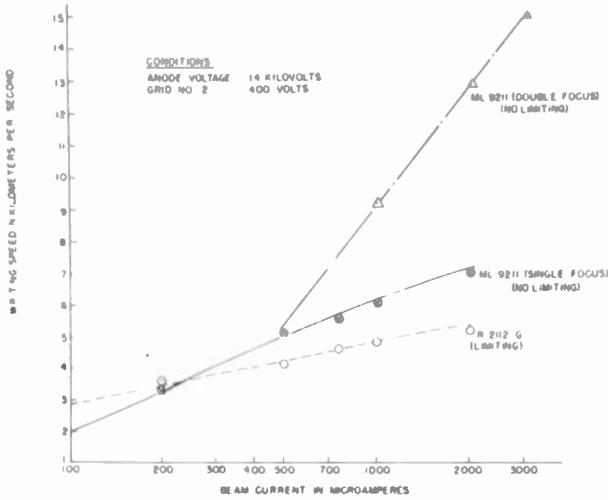
The authors would like to express their appreciation to Dr. F. Holborn for his kind cooperation and to Mr. D. H. Andrews and Mrs. F. R. Darne without whose encouragement this work would not have been possible.

References

1. W. B. Nottingham, Cathode Ray Displays. Vol. 22, Rad. Lab. Series, Ch. 18. McGraw-Hill (1948)
2. R. B. Windsor, "Production and Properties of Skiatron Dark-trace Cathode-ray Tube with Microcrystalline Alkali Halide Screens", Thesis, Rad. Lab., M.I.T., (1944)
3. S. Nozick, Bulletin of A.P.S., Vol. 29, No. 1, 41, Jan. 28, 1954
4. R. D. Rawcliffe and T. Soller, Cathode Ray Displays. Vol. 22, Rad. Lab. Series, Chap. 17. McGraw-Hill (1948)
5. K. Schlesinger, Electronics, Vol. 22, Oct., 102. (1949)



Graph 2
Spot size of R-2112-G and ML-9211.



Graph 3
Writing speed.

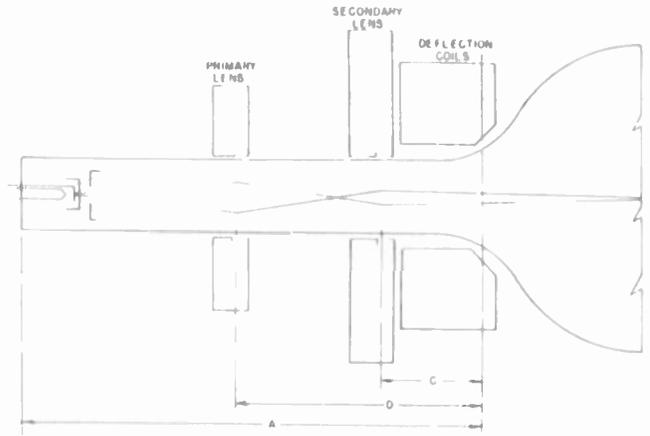
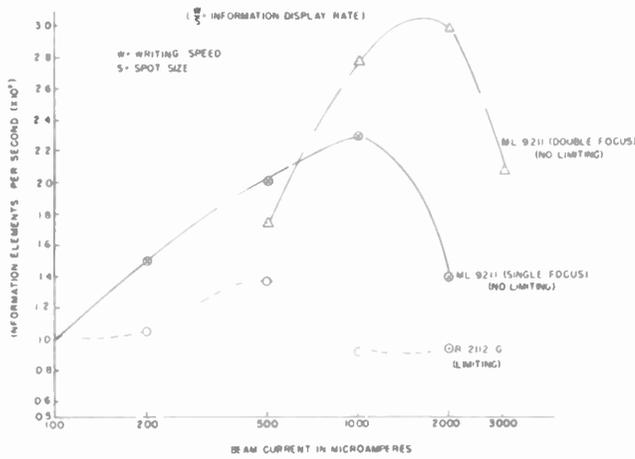


Fig. 1
Double focus system.



Graph 4
Information display rate.

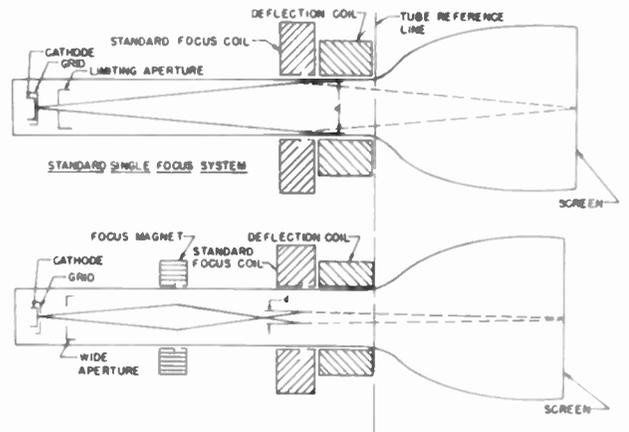


Fig. 2
(a) Beam diameter (single focus vs. double focus);
(b) double focus system.

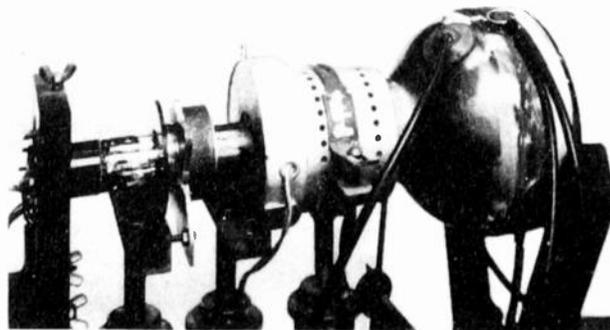


Fig. 3
High writing speed tube assembly.

A LARGE CAPACITY STORAGE TUBE FOR DIGITAL COMPUTER APPLICATIONS

R. B. DeLano, Jr.

International Business Machines Corporation
Poughkeepsie, New York

Introduction

The IBM-93 cathode-ray storage tube, which in certain applications stores 10,000 bits of information, represents a significant advance in random access memory systems. This paper considers tube operation and describes the target structure which gives the new tube improved performance over tubes used in Williams type memory systems. Also included in the discussion is a description of the special circuits required and the test results obtained on a group of these tubes.

By inhibiting the secondary redistribution of electrons, this new tube is able to overcome one of the basic limitations existing in Williams type systems. In these systems, the storage phenomenon itself and the major factor limiting storage capacity, namely spill, are a result of the redistribution of secondary electrons. These factors make it difficult to increase the storage capacity. A tube in which the secondary redistribution is reduced should have much larger storage capabilities.

The IBM-93 belongs to the class of tubes which uses a wire mesh on the storage surface to prevent the redistribution of secondary electrons and uses backplate modulation to provide the second binary state (1,2). A collector grid attracts secondary electrons as they are emitted, producing a further reduction in secondary redistribution.

Tube Construction

Fig. 1 shows a picture of the tube. The re-entrant bulb structure is used to simplify the problem of mounting the target. The gun which is used in this tube was originally developed for the IBM-85 tube used in the Williams-type memory system of the IBM type 701 Electronic Data Processing Machines (3). This gun was designed to give a small spot size and a minimum of deflection defocusing. Since the new tube has similar requirements, the same gun is used. The target located in the left end of the tube is supported by bulb spacers and three leads. The bulb is sealed at the extreme left end after the target is located properly.

The target construction is rather unique and is shown in Fig. 2. One of the ceramic frames is held on a mandrel and placed in a grid winding machine. A square piece of mica having an evaporated metal coating on one side is placed on the frame with the metallic side down. This metallic coating is the backplate. The machine winds four tenths mil diameter tungsten wire over the mica and ceramic. The winding pitch is 300 turns per inch. A low-inertia tensioning device maintains a tension of one-half breaking strength in the wire as it is wound. A second winding is placed on top of the first at right angles to it, forming a mesh. The mesh squares are smaller than the beam diameter, so that the beam covers several squares. The wires are secured to the frame with ceramic cement. The frame has the same temperature coefficient of expansion as the tungsten wire to prevent breaking of wires when the tube is heated during processing. The collector grid is wound on a second frame with wires running at an angle of 45° with respect to the frame sides. The two frames are cemented together with ceramic cement, spacers being used to keep the mesh and the collector grid ten mils apart.

Tube Operation

The output circuit for the tube is shown in Fig. 3. The gun is operated with its cathode at -2300 volts. The collector grid is the most positive electrode in the tube. The pentode is normally cut-off so that the mesh and backplate are at nearly ground potential. The mica surface has a secondary emission coefficient greater than one and a high resistivity.

The binary state called a zero is the equilibrium potential of the surface. If the beam is turned on and the backplate is at ground potential, the mica surface beneath the beam becomes charged to the equilibrium potential. At this potential enough secondary electrons are returned to the spot to make the effective secondary emission coefficient equal to one. The other state, called a plus, is obtained by (1) applying a negative pulse to the backplate while the beam is on and (2) turning the beam off before the end of the backplate pulse. The portion of the mica surface which is not in contact with the

mesh, swings negative with the backplate pulse and secondary emission gives a net electron current away from that portion of the area beneath the beam. This current charges the area positively with respect to the rest of the target. Reading is accomplished by bombarding the area containing the bit in question. If this bit is a zero, there will be no output, since the area is at equilibrium. If it is a plus, the surface will swing negative to reach equilibrium, and there will be a negative output pulse.

The negative backplate pulse is obtained by applying a positive pulse to the pentode grid. The diode conducts and a 100-volt pulse is developed across the mica between the mesh and the backplate. The diode does not conduct while a signal is being read from the tube so that the signal does not have to drive the large mesh to backplate capacitance. Since the input capacitance of the amplifier is much smaller than the capacitance between the mesh and backplate, the use of a diode makes it possible to increase the amplifier input resistance without decreasing the input circuit bandwidth. Because the storage tube may be considered to have a constant-current output, increasing the amplifier input impedance increases the signal magnitude. The diode also reduces the magnitude of the backplate pulse at the amplifier input, thereby improving amplifier recovery.

Fig. 4 shows the arrangement of grid and backplate pulses used in reading and writing. In the plus cycle, the beam is turned on to interrogate the bit during time interval "a" and the plus is rewritten during "b".

The type of a zero cycle shown here is called a reset zero. At the end of the interrogation interval "a" the bit is at the normal zero potential so that the rest of the cycle is unnecessary from this standpoint. In the case of repeated references to a zero bit, however, the fringe electrons of the primary beam would tend to change neighboring bits to zeros. Interval "b" having both the beam and backplate pulse on, has been added so that fringe electrons striking neighboring bits will reset them in the plus direction. The plus which is written during "b" is then erased during interval "c", leaving the bit at the normal zero potential.

Complete resetting of neighboring bits is not possible, but the improvement obtained may be seen in Fig. 5. These curves compare the reset and non-reset operation of the tube. Non-reset oper-

ation means that the beam is turned off at the end of time "a" in the zero cycle (See Fig. 4). The curves show the number of references to a zero bit which will reduce the plus signal of an adjacent bit to 50 per cent of its original value when the two bits are spaced by the amount shown. The improvement which occurs for a large number of references is due to the backplate pulse during the zero cycle.

There is another advantage in having a backplate pulse during each cycle. The portion of this pulse which appears across the diode is about eight volts and is also across the amplifier input. The signal magnitude is only about a millivolt and it occurs seven microseconds after the end of the backplate pulse. It is important to note that there is an amplifier recovery problem here which is considerably reduced by having the backplate pulse occur at regular time intervals. With this arrangement, the amplifier does not have to recover completely between pulses.

Amplifier Circuit

Four cathode-coupled grounded-grid stages are used in the amplifier to provide a non-inverted signal which prevents driving the grid of the second stage positive and charging the grid-coupling capacitor. Discharge of the capacitor would block the amplifier for a long period. Fig. 6 shows an elementary schematic diagram of the amplifier. With the grounded-grid stages, the backplate pulse is negative at the input to each stage and is limited by driving the cathode followers to cutoff. Each stage has a gain of about seven except for the last, which has regenerative feedback from the plate of the modified cathode follower to the inverter grid. The gain in this stage is thereby doubled.

The signal is amplified to a level of a few volts by these stages and then inverted by the high-gain pentode stage. The coupling time constant in the pentode grid circuit is one-half microsecond to differentiate and reduce the duration of overshoots produced in previous interstage coupling networks. The pentode grid serves as a diode, and clips the positive excursions produced by differentiation. A cathode follower is provided for the output.

A composite output signal is shown in Fig. 7. This is a superposition of the signals from 8192 storage locations. The upper waveform was obtained by storing a plus at each of the locations and the lower waveform by storing a complete raster of zeros.

The large pulse on the right in each waveform is the backplate pulse.

Storage Tube Test Equipment

The interaction between bits varies over the target area. Consequently, storage capacity must be determined by using a raster containing a large number of storage locations and testing each location in this raster for interaction. Fig. 8 shows the equipment which is used to perform this test automatically.

The large frame contains most of the logical circuitry. This includes use and regeneration counters and a deflection register. Timing pulses for the grid and backplate pulse generators are provided. Automatic error-indicating circuits and indicator lights for the circuits are also located in this frame.

A chassis, which is not visible in the figure, converts the output of the deflection registers into corresponding deflection voltages for the storage and monitor tubes located in the chassis on the bench. The new storage tube does not require more precise deflection circuits than those used with the IBM-85. This is because both tubes have the same spot size and the beam must strike a bit within one-tenth of a beam diameter.

Alongside the tubes are the backplate pulse generator and the signal amplifier. The grid drive, temporary storage trigger, and amplifier sampling circuits are located on this chassis.

There are two different applications for a storage tube of this type. One application, which is of the buffer storage type, requires only a few references to neighboring locations before they are regenerated. The other application is in the working memory of a computer, where a large number of references is required. In the latter application, the number of references required increases as the storage capacity is increased unless more time is allowed for regeneration.

The tester will operate with rasters containing up to 8192 storage locations. Hexagonal packing is used in the raster. In this arrangement, each bit has six nearest neighbors; whereas each bit has four in square packing. If both configurations have the same distance between nearest neighbors, hexagonal packing allows more bits to be stored in a given area. Each of the tubes was tested with 8192

bits spaced 10 mils apart, and with 1024 bits spaced 22 mils apart. Neither of these rasters cover all the target, which is $1\frac{3}{8}$ inches square. However, they are large enough to evaluate the tube.

The tester starts with a raster full of either zeros or plusses. One storage location is selected by the use address counter and a binary state of opposite sign is written at this location. This bit is referred to a number of times determined by the setting of a dial on the panel. Each time the bit is referred to, it is checked automatically so that forgetting errors will be detected. These errors may be caused by improper reset operation and result in a zero changing to a plus. After these references, the original binary state is rewritten at this location so that unless the fringe electrons have caused interaction between this bit and the others in the raster, all the storage locations contain the same binary state. The complete raster is then regenerated and the signal coming from each bit is checked automatically.

If an error is made, neon indicator lights show whether the error was caused by forgetting or by interaction between bits. In either case, the tester stops making references to the bit, but continues regenerating the information stored in the tube. This enables the operator to see where the error was made by referring to a monitor tube which displays the stored information.

If no error is made, the use address counter moves to the next location and the process is repeated. After the complete raster is checked several times in this manner, it is manually reset to the opposite sign and the raster is checked again. In this way the effects caused by pounding a plus on a field of zeros, and that caused by pounding a zero on a field of plusses are checked. The discriminating level is set so that the same number of references is obtained in both directions and this number is used as an index of performance.

By adjusting the gain of the deflection amplifier, reference numbers may be obtained for various spacings. Fig. 9 shows the reference number of three tubes, A, B, and C, as a function of the spacing between bits. The reference number at 10 and 22 mils is shown for four other tubes, D, E, F, and G. The entire lot of seven tubes was produced at the same time and the data indicates the amount of spread between tubes. It is possible to produce these tubes with consistent operating characteristics.

Storage Capacity

The IBM-85, as used in the IBM 701, stores 1024 bits spaced 52 mils apart and under worst conditions a bit may be referred to 342 times before adjacent bits are regenerated. These bits occupy an area measuring 1.6 inches on a side.

The target for the IBM-93 was made 1.375 inches square so that both tubes would have the same outside diameter. When the new tube stores three times as many bits, the raster has 56 bits along each side with a spacing between bits of 23 mils. The number of references required increases in proportion to the number of bits stored so that this application would require about 1000 references. Referring to Fig. 9, we see that the worst tube has 2000 references at 23 mils spacing. Thus, 3000 bits can be stored with a reasonable safety factor.

For buffer storage and calculator applications requiring less than 100 references, one tube can store 10,000 bits.

The IBM-93 is the result of a program to develop a tube specifically for digital computer storage. The

success achieved so far indicates that the cathode-ray tube is still a very attractive storage means.

Acknowledgment

Dr. D.R. Young did the initial development work on this project and it has been continued by a large group of people. Although it is not possible to mention all their names here, the author would like to acknowledge that their contributions made this development possible.

References

1. "Barrier Grid Storage Tube and its Operation", A.S. Jensen, J.P. Smith, M.H. Mesner, and L.E. Flory, RCA Review, March 1948, pp 112-135.
2. "An Electronic Storage System", E.W. Bivans, J.V. Harrington. IRE National Convention, March 1950.
3. "Improved Cathode-Ray Tube for Application in Williams Memory System," W.E. Mutter. Electrical Engineering, April 1952.



Fig. 1 - The IBM-93 cathode-ray storage tube.

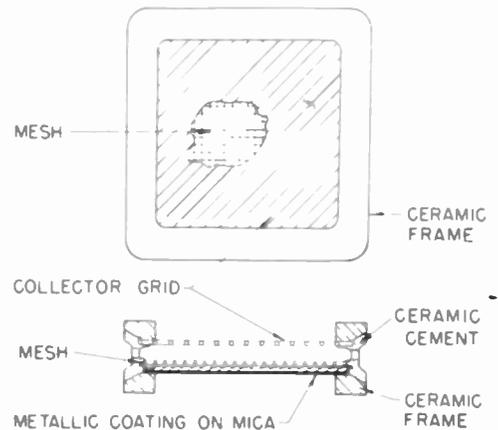


Fig. 2 - Target construction.

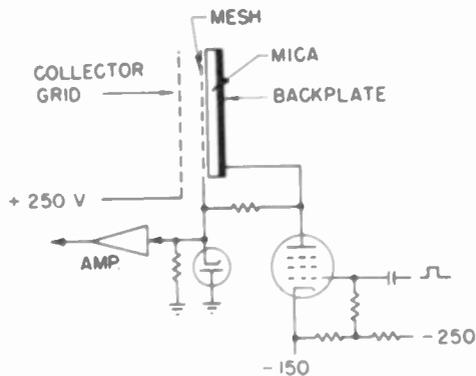


Fig. 3 - Target connections.

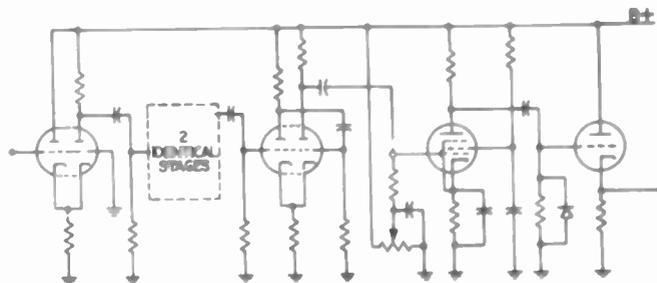


Fig. 6 - Schematic diagram of amplifier.

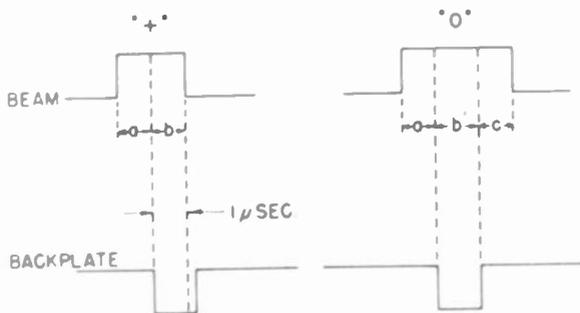


Fig. 4
The two operating cycles - plus and zero.

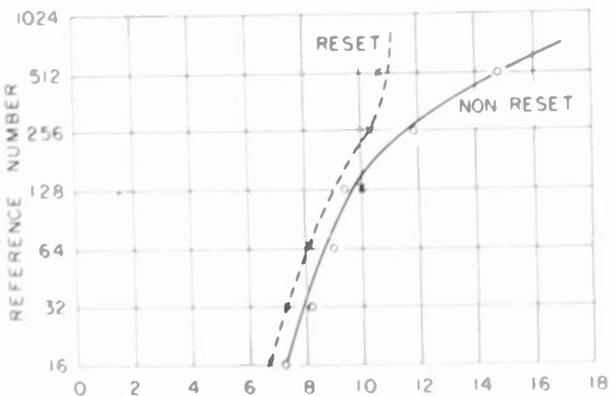


Fig. 5
Number of references to a zero bit which will reduce the plus signal of an adjacent bit to 50% of its original value.

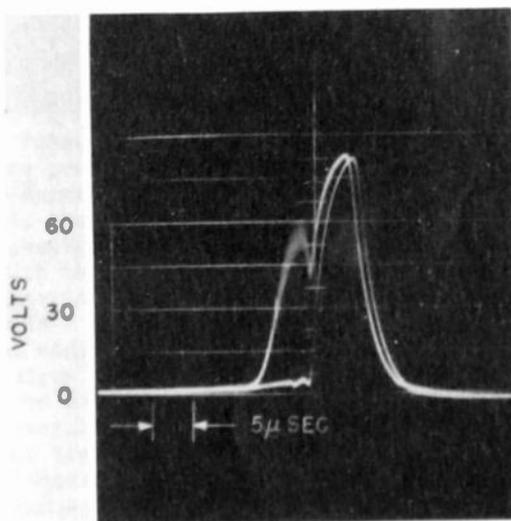


Fig. 7
Composite output signal from 8192 storage locations.

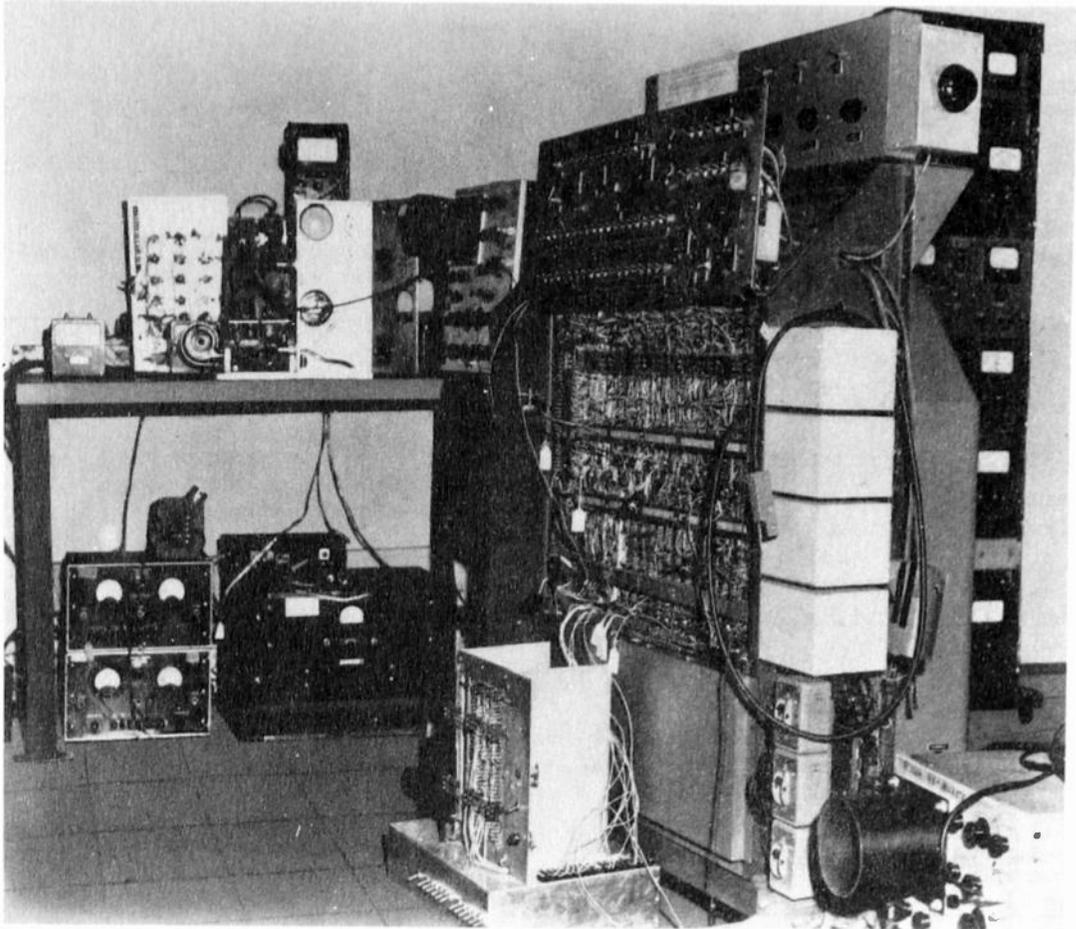


Fig. 8 - Test equipment.

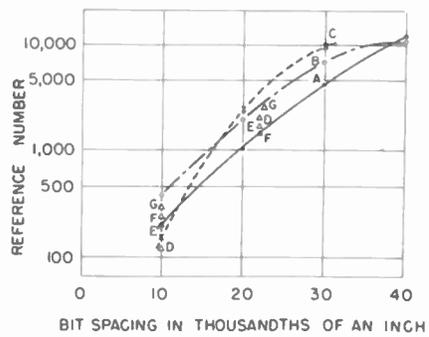


Fig. 9
Reference number vs. spacing between bits.

NOISE LIMITATIONS ON STORAGE TUBE OPERATION

Stanley Winkler and Seymour Nozick
U.S. Naval Material Laboratory
Brooklyn, N.Y.

Summary: In practice, the operation of storage tubes is limited by the presence of noise. The experience obtained in using storage tubes is presented together with the techniques and methods of optimizing the signal-to-noise ratio. The difficulties introduced by gas, lack of voltage regulation, the ripple content of supply voltages, sweep circuit jitter, the choice of beam current and other factors are discussed. The effects of inhomogeneous storage surfaces in both electrostatic and electromagnetic tubes, misalignment of the electron gun, and various distortion factors are considered. The requirements for obtaining satisfactory operational conditions are given, with emphasis placed on practical information to assist the equipment designer.

I. Introduction

The storage of information for a definite period of time has many uses and the modern development of cathode-ray storage tubes gives promise of providing large content, information storage devices. Storage tubes can be applied to signal-to-noise improvement in radar receivers¹, in moving target indicators for radar system (MTI)², as a memory bank in digital and analogue computers³, in television film transmission⁴, in television cameras⁵, and in high speed transient recorders^{6,7}. In practice the operation of storage tubes is limited by the presence of noise. The performance of storage tubes is evaluated by their ability to accurately reproduce a given input signal, the maximum writing speed, the quantity of information which can be stored, the memory span, the accessibility of stored data, and factors related to efficiency of operation. All of these characteristics are affected by noise and the importance of noise is accentuated by the low signal levels encountered at the outputs of storage tubes⁸. The aspect of noise depends upon the particular application. For example, in radar receivers where the storage tube functions as an integrating element, the minimum input signal-to-noise ratio is increased by the presence of noise, while in moving target indicators (MTI), the noise may give rise to false indications. In television cameras, noise produces loss of definition. Other applications are affected in a similar manner. The purpose of this paper is to present the noise limitations on the operation of storage tubes and to indicate the possibilities of minimizing these factors in actual applications, with emphasis in practical information to assist the equipment designer. This paper collects the answers to questions most frequently asked by visitors to our laboratory.

Definition of Noise. Noise is defined⁸ as the undesired signal constituting the difference

between the input and output of a storage tube. The sources of noise can be divided into three categories: random noise arising from the electron beam, distortion introduced by the transfer effect, and signal interference or noise external to the storage tube.

In the present state of development of storage tubes, the operation of storage tubes is difficult even under laboratory conditions. Our experience indicates that they are delicate and must be handled and mounted carefully. Vibration or jarring can cause permanent damage to the screens and storage surfaces. Each tube is an individual and the results obtained vary considerably from tube to tube. Because of these differences, systems using two or more storage tubes have the additional requirement that each tube be provided with the means for separate adjustment. The group at the Material Laboratory engaged in the evaluation and development of storage tubes has been fortunate in having access to many of the developmental storage tubes. These include the Raytheon QK245C, QK245F and QK357; the RCA Graphecon C73366; Williams System Cathode Ray Tubes; the Haeff Tube EM3GA; the Philco Dark Trace Tube; the DuMont Dark Trace Tube; the National Union Dark Trace Tubes R-2112B, R-2112C, R-2112E, R-2112G and the present model R-2112H; and several Freed Radio experimental Dark Trace tubes. A new development, the ML-9211, has been described earlier in this session⁹. Discussion with other groups working with these tubes has confirmed the findings at the Material Laboratory. A 20 db signal-to-noise ratio is a workable figure and a bandwidth of six to ten megacycles attainable. It can be expected that signal-to-noise ratios of 40 to 50 db will be achieved and the operating bandwidth extended to 20 megacycles in the near future. Although at the present time random noise is a negligible factor, the ultimate limitations will come from the electron gun performance and storage surface characteristics.

All of the present day storage tubes suffer from the defect of non-uniformities in the storage surfaces. This produces the shading effects similar to those observed in the RCA Graphecon and in the Raytheon QK245F and the spurious signals produced by blemishes which are so troublesome in Williams system tubes. The rate of insertion of information as well as the quantity of information which can be stored is affected by the electron beam current and spot size. The need exists for electron guns capable of providing a higher ratio of beam current to spot size¹⁰. Such guns would not only increase the writing speed but also extend the number of tonal ranges which can be reproduced. At the present time satisfactory operation of storage tubes in equipment can be

obtained providing that the techniques and methods of low noise and distortion free operation are carefully followed. These techniques will be described below with particular reference to the operation of the QK245F, the QK245C, the R-2112H and the ML-9211.

II. Outline of a Storage Tube

A storage tube consists of an electron gun, deflection circuits and storage elements. For the operation of these component parts, regulated voltage sources are necessary. Three functions are essential, namely, writing, reading and erasing of information. To perform these functions, tubes are built with one, two or three electron guns. The electron guns employed in storage tubes are modified conventional cathode ray tube guns. The output of a storage tube may be visual, electrical or a combination of both. For electrical outputs, the read-out amplifier must be chosen carefully. Figure 1 shows the block diagram of the complete equipment and figure 2 is a photograph of the operating equipment.

III. Electron Gun Operating Voltages

Up to the present, the use of alternating current for filament heating has been found satisfactory, but the filament voltage must be held to better than 5%, and preferably to 1%. This requirement is important when employing low beam currents such as the 1 microampere (ua) in the Williams system or the 2 to 6 ua in the reading beams of the Raytheon and Graphecon tubes. Larger beam currents would increase the writing speed, and therefore, would be desirable. However, in the Graphecon this will be accompanied by a reduction in the storage time. In the QK245 we have found that an ion spot appears in the center of the storage surface and slowly builds up as repeated readings are made. The effect is similar to a blemish except that the area grows with continued use and can be erased. To avoid this effect, very low reading currents must be used. Using a beam current of 2 ua, more than 30,000 consecutive readings were made without noticing any harmful effects or deterioration of the stored pattern.

Variation in the grid bias will also cause changes in the beam current. This is best illustrated by the following table:

Table I. QK245F - Reading Operation

Grid Bias (Volts)	Beam Current (ua)
-56.0	1
-55.0	2
-54.0	3
-53.5	4
-53.0	5
-52.5	6
-52.0	7

From these data it can be estimated that a 1% change in grid bias will cause a 25% change in beam current. The grid bias voltage must be

obtained from a well regulated power supply with 0.1% regulation or better. In addition, the ripple and noise must be kept below 3 millivolts and should be near to 1 millivolt.

It should not be inferred that greater fluctuations can be tolerated with larger beam currents, since fluctuations in 500 ua writing beams will appear as variations in the stored outputs. Where tonal ranges are expected, these variations will limit the number of tones which can be furnished.

Since spot size determines the storage capacity, the highest possible accelerating voltages should be used in order to obtain a small spot size. The equations for estimating the required anode voltage regulation are derived in Appendix I. For a magnetic tube, the regulation may be approximated by

$$R = 4L/K \quad (1)$$

where R is the voltage regulation
 L is the allowable deflection error as a fraction of spot size
 K is the resolution expressed as a number of spots obtainable per line which is line length divided by spot size

For electrostatic tubes this becomes

$$R = 2L/K \quad (2)$$

An idea of the order of magnitudes involved can be obtained by assuming a 100 mm line length with a 1 mm spot size and an allowable error of 0.05. This yields for a magnetic tube

$$R = 0.2\%$$

and half this value for an electrostatic tube, or

$$R = 0.1\%$$

The necessity for better regulation in an electrostatic tube than in a magnetic tube, for the same resolution, provides a partial explanation for the common experience that higher resolving power is obtained with magnetic tubes.

IV. Storage Element Operating Voltages

The storage element is the screen itself in a Williams tube or in a dark trace tube, while in a Graphecon it consists of a copper mesh, a signal plate and an insulator. In the Raytheon storage tube, the storage element has three components, a first screen, the storage screen and a collector-reflector electrode, all of which are supplied with voltages. In this tube the first screen should be operated at as high a value as possible to obtain the smallest spot size. About 450 volts with an anode voltage of 2200 volts was found optimum. The storage screen and collector-reflector electrode voltages vary with the function. For writing, the optimum voltages were 300 volts and -300 volts; for reading, 27 volts and

200 volts; and for erasing, 75 volts and -300 volts, respectively. Carefully regulated power supplies are essential for these voltages and additional filtering may be indicated. Ripple and other noise voltages will modulate the reading and writing beams and will appear in the output where their importance is exaggerated by the low signal levels available. Another source of trouble is any slow drift of these voltages, which will cause a deterioration of the signal quality.

V. Deflection and Focus Circuits

A basic limitation in most storage tubes is the registration or the ability of the electron beam to repeatedly strike the same point in the storage surface. The regulation of anode voltage required for a given deflection error has been computed in a previous paragraph. In Appendix I, the equations for approximating the regulation required of the deflection circuits is given as

$$R = \frac{2L}{K} \quad (3)$$

where R is the deflection voltage regulations for electrostatic tubes and the deflection current regulation for magnetic tubes. L is the allowable fractional spot size deflection error and K is the resolution as defined above. For a 0.05 deflection error with a 1 mm spot size and 100 mm line a 0.1% regulation is needed. The deflection errors caused by anode voltage and deflection voltage (or current) variation are additive and the estimates of 0.1% and 0.2% regulation should be considered as maximum.

Superimposed upon this amplitude variation is the jitter in the scanning raster. During writing, this jitter will produce smears rather than distinct lines. For this reason, the 525 line, interlaced raster is less satisfactory than a 200 line, non-interlaced raster. A stationary raster is necessary for good operation, and single frame writing should be employed wherever possible. The deflection linearity determines the scanning linearity of the electron beam. A rule of thumb value for the maximum allowable jitter is

$$\text{jitter} = \frac{0.2 \text{ sT}}{D} \quad (4)$$

where s is the spot size

T is the period of line scan

and D is the effective tube diameter or usable line length

Jitter in the system greater than this amount appears as poor read out resolution and can be mistakenly attributed to poor focus. A case in point actually occurred in a field application of a skiatron (dark trace) storage tube, where excessive deflection system jitter produced a condition diagnosed by field personnel as "poor focus in a gassy tube". In magnetic focus and deflection systems, mechanical vibration of the deflection yoke or focus coil can produce a similar effect. When attempting to use storage tubes under conditions where shock or vibration may be encountered,

careful shock mounting is recommended.

Mechanical misalignment of the electron gun or the deflection yoke causes errors in beam registration. If the electron beam does not intercept the deflection flux at right angles or is not maintained perpendicular to the target, keystoneing of the raster occurs. The problems of mechanical alignment are accentuated in multiple gun tubes such as the Graphecon. The accurate co-alignment of writing and reading is essential to avoid improper results ranging from spurious and garbled signals to the complete omission of signal if the reading beam strikes the target in between the written pattern. However, in well-constructed tubes the total misregistration should not exceed 1% as measured by the difference between written and read out patterns.

The small beam currents used during the reading operation require anastigmatic focus coils. The astigmatism found in most commercial focus coils is objectionable. The focus coil positioning is very critical and no skewing to achieve centering can be tolerated. Centering must be achieved by suitable deflection circuitry and the focus coil located at the optimum focus point.

VI. Operating the Storage Tube

Successful operation of the storage tube requires relatively noise free and distortionless amplifiers. The design of low-noise broadband amplifiers with Gaussian frequency response is described in the literature. Schematic diagrams of the input and output amplifiers used at our laboratory is shown in figure 3. The bandwidth employed depends on the application but should be no wider than necessary. In this connection it may be noted that the maximum rise time which can be utilized by a storage tube depends upon its own rise time. This can be computed from the formula

$$R.T. = \frac{\text{spot size}}{\text{scanning velocity}} \quad (5)$$

Thus, for a spot size of 1 mm and a scanning v velocity of 2.5 mm per usec (100 mm. scan in 40 usec), the rise time is 0.4 usec. Thus, an amplifier rise time better than 0.2 or 0.1 usec., corresponding to bandwidths of 2 to 4 mc., would be unnecessary. The design of the readout amplifier is complicated by the high output capacitance (50 uuf) of the storage tube. This decreases the output signal with broad bandwidths because of the necessarily low load resistance used, and the resultant sensitivity to extraneous signals reduces the signal-to-noise ratio obtainable. Several methods of reducing the effective output capacitance are available, namely: a negative capacitance amplifier, an overpeaked compensating amplifier, and, what has been found most effective in using the Raytheon tube, a cathode follower stage followed by an overcompensated video amplifier with the storage and signal meshes kept at the same AC potential as the signal output electrode. Using this system, signal-to-noise ratios of 30 db were obtained. To avoid saturation of the readout amplifier during writing, erasing or retracing,

gating should be provided during these intervals. Otherwise spurious signals from a blocked amplifier slowly recovering may occur.

Complete shielding, using mu-metal shields, is essential over the entire length of the storage tube as well as the pre-amplifier portion of the readout amplifier. Shielding of the video readout amplifier and the use of short ground returns to avoid pickup loops was found desirable. Poor resolution and poor registration of write and read beams are frequently caused by improper shielding of the tube. A good practice, which was introduced by R.C.A., is the degaussing of electron guns to remove magnetism in the gun and bulb parts. R.C.A. reports a residual magnetism of 0.03 gauss after degaussing, corresponding to 0.25% of error in beam displacement.

The choice of writing beam current depends upon the rate of information insertion, since the writing speed is proportional to the ratio of beam current to spot size. However, the writing beam current is limited by the level at which the peak signals barely saturate the storage surface. At saturation, the maximum video signal output exists.

The reading current is generally a compromise. Typical read out signals for 1.2 ua and 6.0 ua are shown in figure 4. To achieve the largest signal output, a high reading beam is required. However, ion effects tend to cloud or distort the signal, reducing its amplitude and tones. In magnetic tubes, this starts by an erasure at the center and moving out. In electrostatic tubes, a deterioration over the whole stored pattern ensues. Since the ion effects depend upon current and time, fast single read outs permit high read currents while long repetitious readouts require low beam currents. Storage tubes with tonal ranges are designed particularly for low reading currents.

An effect, which limits storage tube operation, is a blemish on the storage surface which produces spurious signals. The time occurrence of these is dependent on the stored pattern position. Little can be done except to try to reduce the effect by suitable band rejection filters or by relocating the stored pattern elsewhere on the storage surface. The storage surfaces are usually a dielectric deposit on a mesh or other metallic surface. These screens have a tendency to flake off with hard handling causing blemishes, and hence lack of storage capability at certain locations on the storage screen. One trouble which should generally be avoided is a continuation of the electron beam following a sweep failure. This concentrated beam can burn holes in the storage surface which would tend to limit the tube usefulness by creating a blemish. Simple protective circuits, which cut off the storage tube beam in the event of deflection failure, are available.

In tubes with mechanically supported meshes, these metallic meshes are free to vibrate with respect to one another. With voltages impressed, the action is that of a condenser microphone or a microphonic audio signal which can obscure the

desired video readout. This is illustrated in figure 5.

VII. Conclusions

Although additional developmental work is needed before completely satisfactory utilization of storage tubes in field applications can be achieved, it is possible at the present time to obtain adequate low level performance with existing tubes (see figures 6 and 7) provided that careful design of equipment and good engineering practice is followed. Signal-to-noise ratios of 20 to 30 db have been attained in practice and an increase in this figure should be possible with better electron guns and improved manufacturing techniques. Random noise is a negligible factor at this time, but the ultimate limitations will depend on electron gun performance and storage surface characteristics. In applications where both low noise and long access time is required, flicker noise may become important, but, since flicker noise is an inverse frequency effect, as the low frequency cut-off is raised, contribution from this source is reduced.

In utilizing storage tubes, the system distortion and noise producing elements must be minimized. Operation of the storage tube and the readout pre-amplifier should be carried on inside a shielded cage. See figure 8. Mu-metal shielding of the tube is necessary to reduce magnetic defocussing effects. Good regulation and filtering of power supplies is essential. In obtaining wide bandwidth with the desirable high load impedance, frequency over-compensation was found useful.

Measurements of square wave outputs of the Raytheon QK245F recording storage tube yielded a signal-to-noise ratio of 30 db.

Acknowledgements

It is a pleasure to acknowledge the contributions of our colleagues, W.S. Treitel and P.R. Liegey, who participated in the experimental program. We wish to thank both D.H. Andrews, Assistant Head, Electron Tube Branch, NML; and Mrs. F.R. Darne, Chairman, Sub-panel on Special Tubes of the RDB, for their encouragement and support of our work.

References

1. J.V. Harrington and T.F. Rogers, "Signal-to-noise improvement through integration in a storage tube", Proc. I.R.E., vol. 38, pp.1197-1203; October 1950.
2. R.A. McConnell, A.G. Emslie and F. Cunningham, "A moving target selector using deflection modulation on a storage mosaic", M.I.T. Rad. Lab. Report 562; June 1944 referred to in R.A. McConnell, "Video storage by secondary emission from simple mosaics", Proc. I.R.E., vol. 35, pp.1258-1264, November 1947.
3. S.H. Dodd, H. Klemperer and P. Youtz, "Elec-

trostatic storage tube", Elec. Eng., vol. 69, pp. 990-995; November 1950.

4. F. Schroeter, "Image storage in television reception", Optik, vol. 1, pp. 406-409, 1946.
5. S.V. Forgue, "The storage orthicon", R.C.A. Review, vol. 8, pp. 633-650, 1947.
6. L.E. Flory, J.E. Dilley, W.S. Fike and R.W. Smith, "A storage oscilloscope", R.C.A. Review, vol. 12, pp. 220-229, June 1951.
7. A.A. Benner and L.M. Seeberger, "Graphecon writing characteristics", *ibid*, pp. 230-250.
8. Stanley Winkler and Seymour Nozick, "Noise in storage tubes", Phys. Rev., vol. 91, p. 492, July 15, 1953.
9. S. Nozick, N.H. Burton and S. Newman, "A high writing speed dark trace tube", I.R.E. National Convention, Session on Electron Devices III - Storage Tubes, March 25, 1953.
10. S. Nozick, "Writing speed and tonal range of dark trace tubes", Bulletin of the American Physical Society, vol. 29, n.1, p.41 January 1954.

Appendix I - Regulation Required for Anode Voltage and Deflection Circuit.

A. Magnetic Tubes

For a magnetic deflection tube

$$y = CIV^{-1/2} \quad (1)$$

where y is the spot displacement from the center of the tube

I is the deflection yoke current
 V is the anode voltage

$$dy = CV^{-1/2}dI - 1/2 CIV^{-3/2}dV \quad (2)$$

$$\Delta y = CV^{-1/2}\Delta I - 1/2 CIV^{-3/2}\Delta V \quad (3)$$

$$\Delta y = y \left(\frac{\Delta I}{I} - \frac{1}{2} \frac{\Delta V}{V} \right) \quad (4)$$

$$L = \frac{\Delta y}{s} = \frac{y}{s} \left(\frac{\Delta I}{I} - \frac{1}{2} \frac{\Delta V}{V} \right) \quad (5)$$

Since V and I need not vary in the same sense and since we are examining the worst condition, choose the spot location at the periphery of the tube or $y = \frac{D}{2}$.

$$\frac{2Ls}{D} = \left| \frac{\Delta I}{I} \right| + 1/2 \left| \frac{\Delta V}{V} \right| \quad (6)$$

$$\frac{2Ls}{D} = R_I + 1/2 R_V \quad (7)$$

where R_I is the centering current regulation.
 R_V is the anode voltage regulation.

L is the allowable fractional spot size registration error.

s is the spot size.

D is the useful storage surface diameter.

The centering current and anode voltage regulations permitting a maximum allowable fractional spot size registration is shown in equation (7).

B. Electrostatic Tubes

For an electrostatic tube

$$y = CE V^{-1} \quad (8)$$

where y is the spot displacement from the center of the tube.

E is the deflection voltage.

V is the anode voltage.

$$dy = C V^{-1} dE - CE V^{-2} dV \quad (9)$$

$$\Delta y = C V^{-1} \Delta E - CE V^{-2} \Delta V \quad (10)$$

$$\Delta y = y \left(\frac{\Delta E}{E} - \frac{\Delta V}{V} \right) \quad (11)$$

$$L = \frac{\Delta y}{s} = \frac{y}{s} \left(\frac{\Delta E}{E} - \frac{\Delta V}{V} \right) \quad (12)$$

Since ΔV_d and ΔV_o need not vary in the same sense and since we are examining the worst conditions, choose the spot location at the periphery of the tube, or $y = D/2$.

$$\frac{2Ls}{D} = \left| \frac{\Delta E}{E} \right| + \left| \frac{\Delta V}{V} \right| \quad (13)$$

$$\frac{2Ls}{D} = R_d + R_v \quad (14)$$

where R_d is the centering voltage regulation.
 R_v is the anode voltage regulation.

L is the allowable fractional spot size registration.

s is the spot size.

D is the useful storage surface diameter.

The relationship between centering and anode voltage regulation permitting a maximum allowable spot size registration is shown in equation (14).

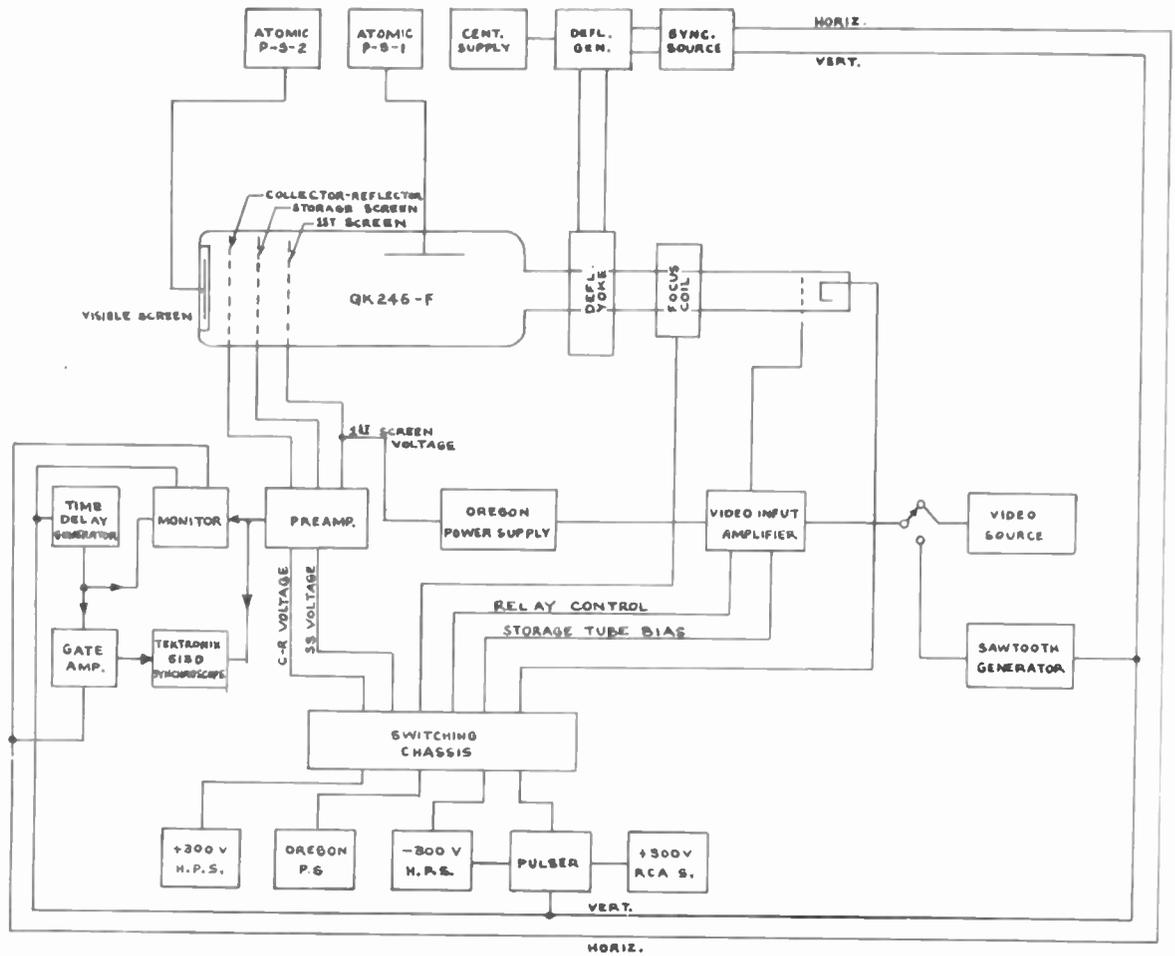


Fig. 1 - Equipment block diagram.

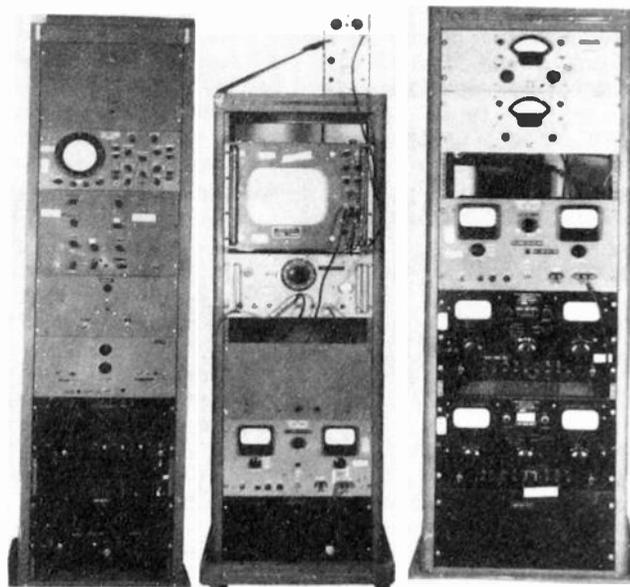
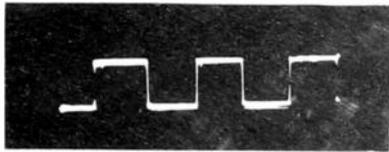
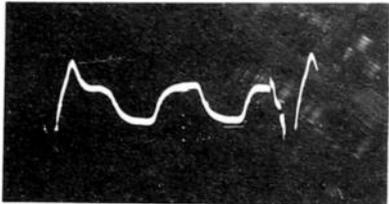


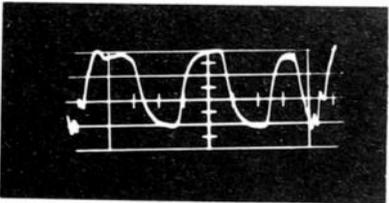
Fig. 2 - Storage tube equipment.



Storage tube input



Read out with 1.20 microamperes
12.6 millivolts per centimeter



Read out with 6.0 microamperes
35 millivolts per centimeter

Fig. 4 - Typical read out signals.

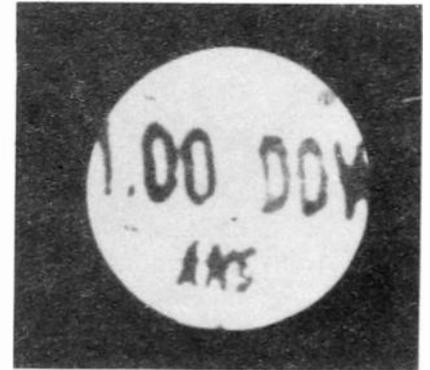
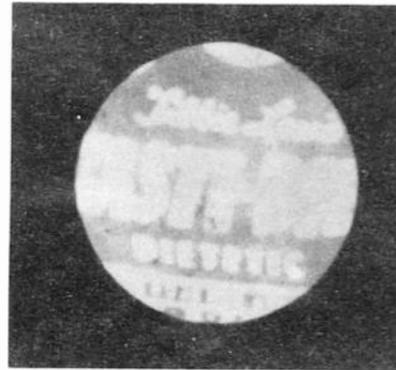


Fig. 6
Television signals read out after an hour's
storage in the QK245F.

138

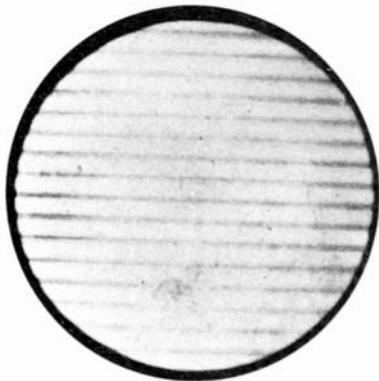


Fig. 5
Monitor screen showing
vibration during writing.

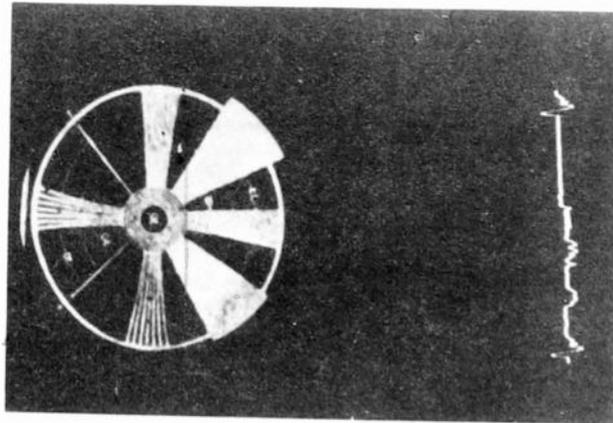


Fig. 7
QK245C monoscope pattern and "snatched" line.

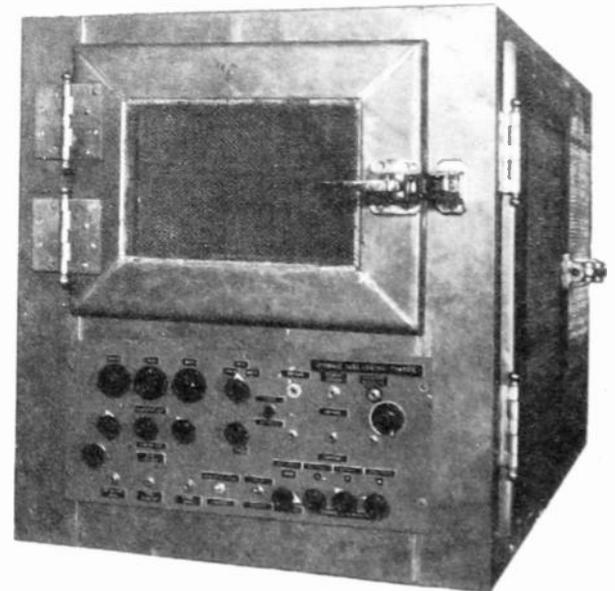


Fig. 8 - Shielded enclosure.

A VOLTAGE-TUNABLE MAGNETRON FOR OPERATION IN THE FREQUENCY RANGE 1500 TO 3000 MEGACYCLES*

Joseph A. Boyd
University of Michigan
Ann Arbor, Michigan

Introduction

This paper presents the results of work done in the University of Michigan Electron Tube Laboratory toward the development of a voltage-tunable magnetron. This magnetron can be electronically tuned in the frequency range of 1500 to 3000 mc, and delivers a useful power output in the order of 250 milliwatts.

Design of Magnetron and Cavity

In the design of this magnetron (designated the Model 11) and its cavity, two factors were considered essential if a satisfactory voltage-tunable magnetron was to be developed. These factors were:

- (a) Low anode-to-anode capacitance;
- (b) A high-impedance external cavity, the impedance being a slowly varying function of frequency. By high impedance we mean in the order of 100 ohms.

Consideration of the above circuit requirements resulted in the conclusion that a waveguide structure offered a possibility for fulfilling these requirements. The interdigital anode structure is adaptable to mounting in a waveguide structure and was selected for this reason. It also has wide range tuning possibilities and is mechanically simple.

The ridge waveguide¹ offers a possibility for increasing the circuit impedance over that of a rectangular waveguide for a given spacing between anode supports. The ridge waveguide is characterized by its lowered cutoff frequency and lowered impedance (compared with a rectangular guide of equal dimensions), and by a wide bandwidth free from higher-mode interference.

A section of standard S-band rectangular waveguide was used with a ridge .84 inch high and 1.0 inch wide. This structure has a characteristic impedance at infinite frequency (Z_{∞}) of 160 ohms, and a cutoff frequency of 1250 mc. The characteristic impedance as a function of frequency is given by the relation

$$Z_o = \frac{Z_{\infty}}{\sqrt{1 - \left(\frac{f_c}{f}\right)^2}} \quad (1)$$

where f_c = cutoff frequency of ridge waveguide

*This paper is based on work done for the Signal Corps, U.S. Army Contract No. DA-30-039-SC5423.

f = operating frequency of ridge waveguide.

Since it is desirable to have a coaxial output from the magnetron, a tapered-ridge waveguide-to-coaxial-line junction was used to terminate each end of the ridge waveguide. This arrangement is shown in Fig. 1. The junctions used are based on a design given by Cohn². These junctions have an upper cutoff frequency of approximately 4800 mc, therefore, when used in conjunction with a ridge waveguide with a lower cutoff frequency of 1250 mc, give a very wide-band structure. The impedance which this circuit presents to the magnetron is, to a very large extent, determined by the termination used on the coaxial output terminals. Theoretically, this impedance would be a pure resistance whose magnitude is given by Eq. 1 when the coaxial output terminals are terminated in their characteristic impedances.

Design of Model 11 Magnetron

In the design of the interaction space for the Model 11 magnetron, there were two factors, the anode-to-anode capacitance and the shape of the anode bars, which were considered to be critical to voltage-tunable operation. Thus in addition to the usual consideration of current, voltage, frequency, magnetic field, and interaction space dimensions, special attention was given to these two parameters.

Hull³ has suggested that the operation of a typical magnetron can be improved by redesigning the interaction space to reduce the higher-order space (or Hartree) harmonics of the RF field. His analysis was made for the case of π -mode operation and for planar geometry. It appears that the greatest improvement is obtained for low values of operating voltage and magnetic field. The shape of that portion of the anode bars nearest the cathode determines the field configuration between the anode and cathode; therefore, it is necessary to consider only this portion of the anode bars. It can be shown that a good approximation to the desired anode shape can be obtained by making the anode bars round with a diameter of $2/3d$ (where d is the center-to-center spacing of the anode bars.)

In the operation of a low-Q magnetron, the presence of the space harmonics other than the π -mode components can cause significant harmonic content to exist in the output voltage of the magnetron due to the fact that the impedance at the harmonic frequencies may be appreciable for a low-Q, wide-band circuit. In other words, the circuit will not serve as a filter to assure that only the fundamental frequency will occur in the output voltage.

On the basis of these considerations and the fact that round anode bars give less anode-to-anode capacitance, it was decided to use round anode bars in the design of the Model 11 magnetron. However, the design suggested by Hull for the complete elimination of space harmonics called for an anode diameter of $2/3d$ (where d is the center-to-center spacing of adjacent anode bars), which is inconsistent with a design for low anode-to-anode capacitance. Therefore, the spacing between adjacent anode bars was made equal to the diameter of the anode bars.

Listed below are the values of the parameters which resulted from the design of the interaction space.

- $f = 3000$ mc (chosen to suit available test equipment)
- $P_i = 50$ watts
- $I_b = .050$ amperes (required to satisfy the maximum power demand at an anode voltage of 1000 volts)
- $N = 12$ (number of anode bars, based on existing magnetrons)
- $d_a = 0.035$ inch (diameter of anode bars)
- $r_c/r_a = 0.6$ (ratio of cathode to anode radii, based on existing magnetrons)
- $r_a = .166$ inch (determined by size of molybdenum rod available, and anode radius of existing magnetrons)
- $L_c = .150$ inch (length of emitting surface on cathode required to satisfy maximum power requirements with a safety factor of two)
- $l_s = .290$ inch (overlap length of anode bars: chosen to allow for length of emitting portion of cathode and end hats)
- $E_o = 250$ volts
- $B_o = 550$ gauss
- $C_p = .9\mu f$ (anode-to-anode capacitance, no cathode present)

The Magnetic Circuit

In the design of the magnetic circuit there were two factors which were considered to be important to the successful operation of the tube: (1) that the magnetic field be as nearly uniform as possible throughout the interaction space, and (2) that relatively high fields, in the order of 3000 gauss, be obtained without unreasonable requirements for an electromagnet. Fig. 2 shows a drawing of the Model 11 magnetron with an oxide-coated cathode in place. As indicated, the iron pole pieces, part 1, serve as anode supports. The molybdenum rods are brazed onto the pole pieces. Parts 2 and 3 are made of Kovar to facilitate making the glass seals as indicated.

Voltage-Tunable Magnetrons Models 11A and 11B

Two modifications were made in the design of the Model 11 magnetron:

- (a) Square anode bars are used with an inside anode diameter of .235 inch. The anode bars are .030 inch. The bars formed by grinding .045 molybdenum rods to the desired dimension. This tube,

which is otherwise identical to the Model 11, has been designated Model 11A.

- (b) The inside anode diameter was changed to 0.205 inch, and .030 inch molybdenum rods were used as anode bars. With this arrangement, the spacing between adjacent anode bars is also .030 inch. This tube, which is otherwise identical with the Model 11, has been designated Model 11B.

The purpose of the Model 11A was to make possible a comparison of the two similar tubes, one with round anode bars, and one with square anode bars.

The design of the Model 11B was based on a consideration of the operation of a voltage-tunable magnetron. The output of the Model 11 is in the order of 15 milliwatts, which indicates a very small rf voltage across the anode bars. For small rf voltages, it is reasonable to expect that bunching and interaction efficiency will be enhanced by increasing the ratio r_c/r_a so that the rf fields will penetrate farther into the interaction space. The Model 11B was designed for the purpose of determining the validity of this assumption. Tungsten cathodes with a diameter of .150 inch were used in the Model 11B, giving an r_c/r_a ratio of .75, as compared to a nominal value of .6 for conventional high-Q magnetrons.

EXPERIMENTAL RESULTS

A total of thirty-one tubes were studied in this investigation; in this section typical results as obtained from the various forms of the tube are presented. Several types of cathodes were used, namely; oxide-coated, thoriated tungsten, tungsten, and a button cathode. Results of this investigation indicated that the pure tungsten cathode is most satisfactory for use in the Model 11 tubes. The first tube constructed used an oxide-coated cathode. This tube could not be operated cw because the back-heating was sufficient to prevent control of the cathode temperature required for coherent operation of the tube. The thoriated tungsten cathode made possible the cw operation of the tubes but was more sensitive to backheating than the pure tungsten cathode. The tube using a button cathode could only be operated under pulsed conditions because of the backheating of the cathode. Data are presented only on those tubes which contained pure tungsten cathodes.

Model 11 No. 83

Typical data from the Model 11 tubes are shown in Fig. 3. This shows the voltage tuning characteristic of the tube as cw frequency measurements were made from 1780 to 2450 mc.

The break in the tuning curve (as indicated by the dotted section) is caused by readjustment of the cathode input power (which is necessary in order to maintain a coherent signal output). It is possible to operate the tube over a considerably

wider frequency range when the anode voltage is swept.

When the tube was used as a local oscillator in a spectrum analyzer it was possible to vary the frequency from 1840 to 2730 mc with a coherent signal output throughout the frequency range. Conditions for this test were as follows:

$$\begin{array}{ll} B = 3940 \text{ gauss} & E_{dc} = 3350 \text{ volts} \\ I_{fil} = 8.95 \text{ amps} & E_{ac} = 1200 \text{ volts} \\ & \text{(peak-to-peak)} \end{array}$$

Model 11B No. 112

Tungsten rods were used as the anode bars in this tube. The heat dissipating capacity of the anode structure is increased over that of the tubes which use molybdenum anode bars. The processing technique was modified in the construction of this tube in that the cathode was vacuum-fired before mounting in the tube. This vacuum-firing of the cathode before mounting eliminated the coating of the glass envelope of the tube which had previously occurred in normal operation of the tubes. (This coating was due to molybdenum evaporated from the cathode).

This tube was tested under dynamic operating conditions, and the results are as shown in Figs. 4 and 5. The conditions of operation are indicated on the figures. A 60-cycle voltage was superimposed upon the dc anode potential, giving an anode-voltage swing of 900 volts. Filament power was supplied from a battery to eliminate the frequency modulation caused by the variation of the magnetic field which is produced by an ac heater current.

Fig. 4 shows the variation of the frequency with the anode voltage. To obtain these data the horizontal sweep on an oscilloscope was synchronized with the ac anode voltage, and therefore, could be calibrated in frequency.

In addition to the voltage-tunable mode indicated in Fig. 4, there were two other modes which were very weak, but still detectable over portions of the voltage range. One of these modes tuned from 2845 mc at 2100 volts to 3170 mc at 2300 volts; the second tuned from 4080 mc at 2050 volts to 4370 mc at 2200 volts. In the voltage range 2350 to 2850 only the main mode was present. Under certain conditions of operation, there are discontinuities in the operation of the main mode. These discontinuities in the operation are detectable in both the power output and anode current curves (shown in Fig. 5). Continuous operation can be obtained by proper adjustment of the various parameters. Frequency deviations of 400 to 600 mc can be obtained with no difficulty in maintaining continuous operation; for frequency deviations of the order of 800 to 1000 mc, the adjustment of the cathode temperature becomes more critical.

With the modulation voltage reduced to 450 volts peak-to-peak, the frequency of operation

could be shifted as much as 50 mc by varying the filament heater current and still maintain coherent operation over the voltage range. An increase in cathode temperature decreases the frequency of oscillation. The tube seemed to operate best at, or near, the lower cathode temperature.

For cw operating conditions it was possible to vary the filament current from 8.9 amps to 9.3 amps and still maintain a coherent signal output. The power output varied from 52 mw at an anode current of 5 ma to 350 mw at an anode current of 18.75 ma. The corresponding frequency variation was from 2450 mc to 2320 mc.

Figure 5 shows the variation in power output and anode current with the anode voltage. The frequency range in Fig. 5 is approximately 900 mc (as shown in Fig. 4) and the peak power is of the order of 550 milliwatts. Fig. 5 shows the variation of anode current with anode voltage, the average current being approximately 15 ma. It should be noted that there are no discontinuities in the operation and it was determined that the tube was oscillating coherently throughout the range. The decrease in the power output in the right (high frequency) side of Fig. 5a is due to the loading of the cathode circuit at that frequency.

This tube was operated with battery supplies for the filament and anode, and an electrically regulated power supply for the electromagnet. Under these conditions, the output signal (as observed on a spectrum analyzer) was stable; no jitter was present in cw operation. The output could be tuned over several hundred megacycles without readjusting the cathode temperature. Qualitative measurements using the spectrum analyzer indicated that the coherent output signal was at least 65 db above the noise present in the output signal.

Satisfactory operation was obtained with anode voltages as low as 800 volts and frequencies in the order of 2000 mc. Under these conditions the output power was reduced but the signal was stable and relatively free of noise.

Model 11A

The Model 11A magnetron is similar to the Model 11 except that square anode bars are used. The Model 11A was constructed for the purpose of comparing the operation of similar tubes with round and with square anode bars. In general, the Model 11A is more sensitive to cathode temperature than either the Model 11 or Model 11B and cannot be tuned over as wide a frequency range. The output signal contains more noise than in either the Model 11 or Model 11B. The power output is slightly greater than for the Model 11, but less than the Model 11B. The data indicates that the improvement in operating characteristics of the Model 11B over the Model 11A is due primarily to the lower value of anode-to-anode capacitance.

DISCUSSION OF EXPERIMENTAL RESULTS

The power output of the voltage-tunable magnetrons has been shown to be very closely related to the total shunt impedance of the rf circuit, the power being a minimum when the shunt impedance is a minimum. With the wide band non-resonant circuit the peak power output was in the order of 500 mw; however, when the shunt impedance of the rf circuit was increased by shorting one end of the cavity, power output of 4 watts was obtained. Under these conditions the frequency of the output signal could be varied as much as 200 mc by varying the anode voltage.

A pure tungsten cathode which operates at approximately 2200 degrees centigrade was found to be most suitable for voltage-tunable operation if coherent signal output is desired. The operation of the magnetron is sensitive to the temperature of the cathode; however, the filament input power can be adjusted to give stable, coherent operation.

FREQUENCY CHARACTERISTICS

The experimentally determined tuning characteristics of the various forms of the voltage-tunable magnetron can be compared by the use of the following equation⁴

$$\frac{f}{f_0} = \left[1 - \sqrt{1 - \frac{E_H}{E_u}} \right] \quad (2)$$

where

E_H = threshold potential
 E_u^H = Hull cutoff potential
 f_0 = is the frequency of the current induced in the anodes by an electron at the anode with a tangential energy of E_u .

$$f_0 = \left[\frac{E_u e}{m 2\pi^2} \right]^{1/2} \frac{n}{r_a} \quad (3)$$

Fig. 6 shows the theoretical tuning curve plotted for comparison with the experimental curves for the three forms of the tube. In plotting the experimental data, f is the operating frequency and E_a is the anode potential. E_u and f_0 were determined from the operating conditions, assuming π -mode of operation. It should be noted that all forms of the tube had essentially linear tuning character-

istics and all tubes operated at higher anode potentials than indicated by the Hartree equation when π -mode of operation is assumed.

It was assumed that the Model 11 magnetron would operate in the π -mode. Rotating-probe measurements on this tube and its associated cavity structure show that with an rf voltage applied to one of the output terminals of the cavity, a π -mode voltage distribution does exist on the anode bars.

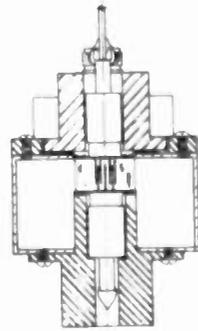
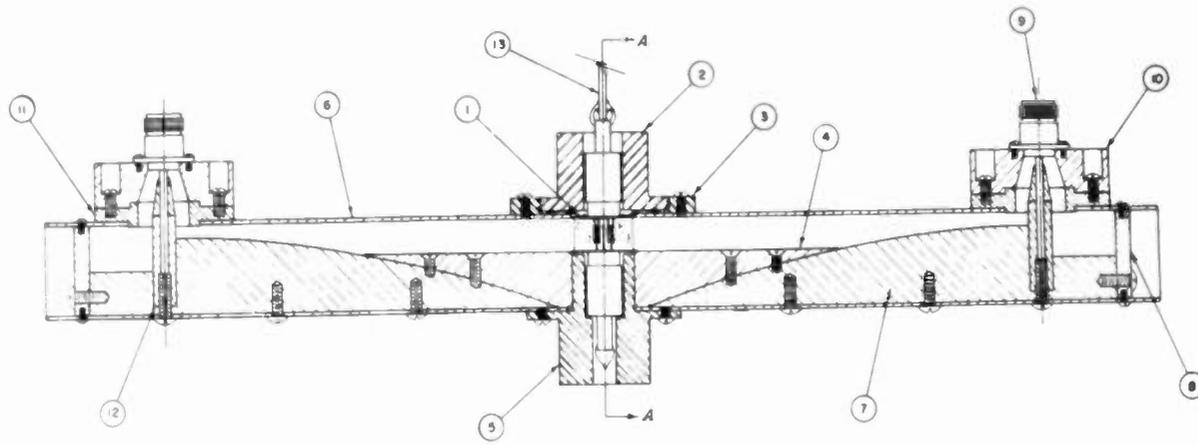
CONCLUSION

A voltage-tunable magnetron has been developed with operating characteristics that make the tube useful as a source of microwave power. It has been shown that the tube is satisfactory for use as a local oscillator spectrum analyzer, and as an rf power source for most rf measurement applications. Work is in progress to determine the feasibility of using the tube as the local oscillator in microwave receivers.

With the present circuit arrangement, the anode-to-anode capacitance appears to be the limiting factor in obtaining higher power from the tubes. A reduction in the value of this capacitance is necessary if significantly higher power is to be obtained. Some improvement in efficiency and power output may be obtained by increasing the external circuit impedance; however, the improvement to be obtained in this way is limited. It is believed that voltage-tunable operation of a magnetron can be obtained under space-charge-limited operation of the cathode if the impedance of the cavity and anode structure can be made sufficiently large.

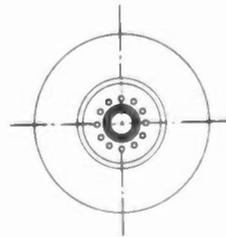
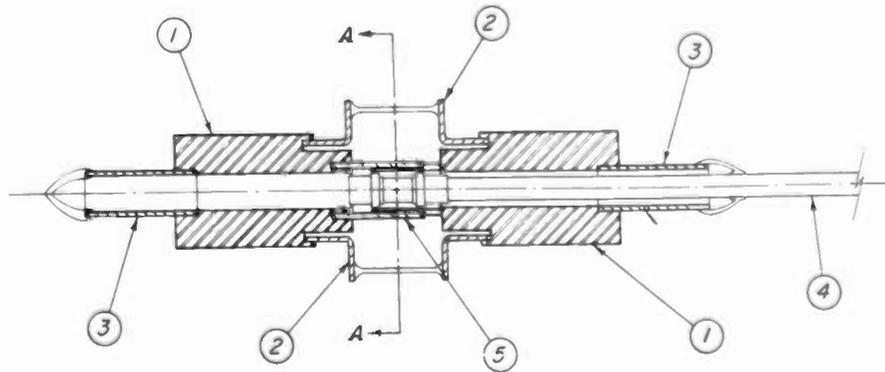
REFERENCES

1. S. B. Cohn, "Properties of Ridge Waveguide," Proc. IRE, August 1947.
2. Radio Research Laboratory Staff, "Very High Frequency Technique," vol. II, McGraw-Hill, 1947.
3. J. F. Hull, "Analysis of a Magnetron Interaction Space for Elimination of Hartree Harmonics," Memorandum Report, SCC SCL-RTB2, Project 322A, Evans Signal Laboratory, January 1951.
4. J. A. Boyd, "A Voltage-Tunable Magnetron for Operation in the Frequency Range 1500 to 3000 Megacycles", University of Michigan, Electron Tube Laboratory, Technical Report No. 15, November 1953.



SECTION A-A

Fig. 1



SECTION A-A

Fig. 2

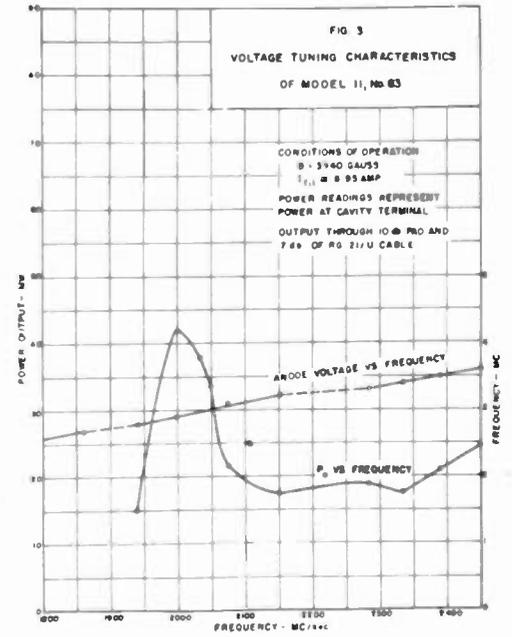


Fig. 3

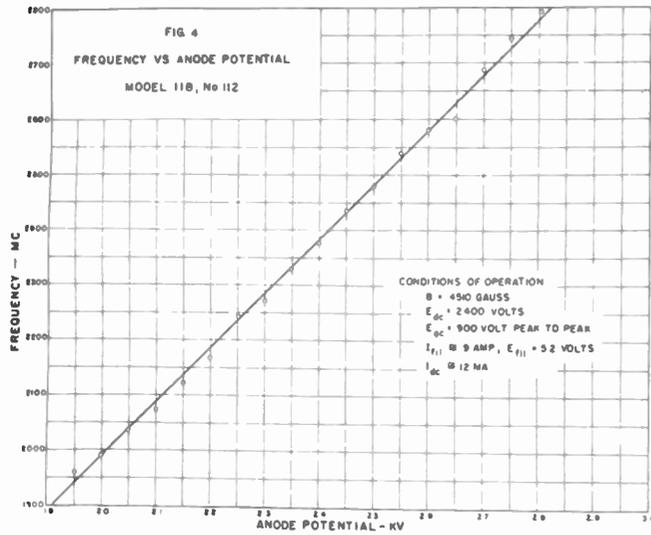


Fig. 4

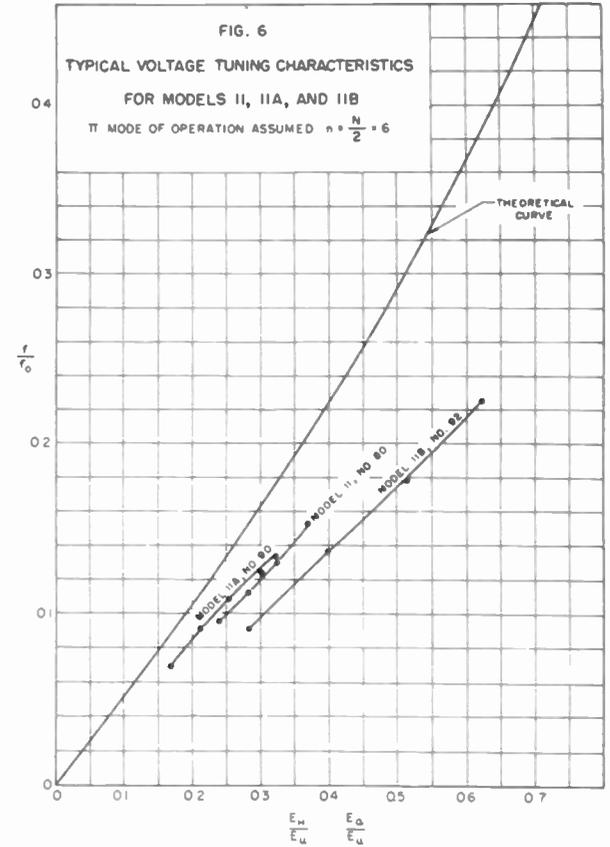


Fig. 6

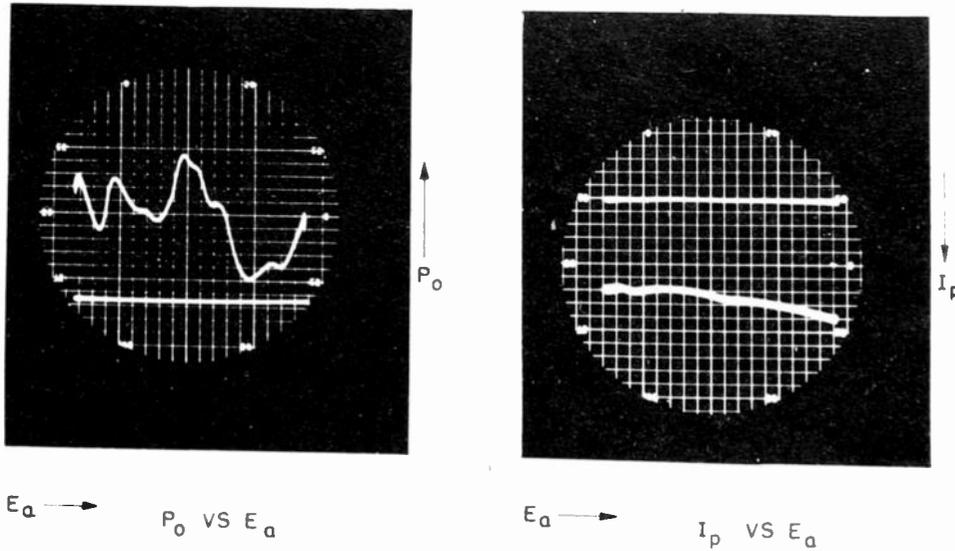


Fig. 5

Power and current characteristics of Model 11B, No. 112.
 $B = 4510$ gauss; $E_{dc} = 2400$ volts; $E_{ac} = 900$ volts peak to peak;
 $I_{fil} \approx 9$ amp; $I_{dc} \approx 15$ amp; $E_{fil} = 5.2$ volts.

111

CONTROL OF ELECTRON-BEAM SPREAD BY POSITIVE ION TRAPS

E. L. Ginzton and B. H. Wadia
Stanford University
Stanford, California

I. Introduction

The negative space charge of a stream of electrons causes the production of forces on the particles comprising the stream. These forces are directed away from the stream and hence tend to increase its cross section. In the case of electron beams required in klystrons and traveling-wave tubes this effect is, at present, countered by the use of a magnetic field directed along the axis of the beam.

A simpler method, which has been known for some time,¹ consists of confining a sufficient number of positive ions in the drift space to neutralize the negative charge of the electrons and thus to suppress the very cause of the divergent forces. Positive ions are created within the tube by collision of the beam electrons with molecules of residual gas. An appreciable fraction of these ions is continually being removed from the beam by several causes, the main one being a steady loss to the cathode. It is suggested that if this loss to the cathode is eliminated by interposing a positive hump of potential in the path of the ions, complete neutralization should be possible even at good vacuums.¹

The few attempts that have been made at positive-ion trapping have not met with much success and have yielded results which were not explainable on the basis of the simple theory then prevailing. It has been found possible to extend and modify this simple theory in order to obtain a complete understanding of the phenomenon and to correlate theory with experiment.

II. The Untrapped Positive Ions

A. The Neutralizing Effect of Ions

At the generally-accepted good vacuums, a large number of gas molecules is usually present within the tube envelope; the presence of these molecules results in the formation of positive ions by virtue of their collisions with electrons. The rate of positive-ion formation is dependent upon the nature of the gas, its pressure, the number of electrons present in the beam, and their velocity.

Since these ions are positively charged, they will move toward points of lower potential which are near the axis of the beam. The slow electrons, which are the other product of the ionizing collision, will move in the opposite direction, namely to the drift tube, and will be lost. It may then be said that the beam will soon form a core of positive ions clustered around its axis. If, for the present, we assume that there is no appreciable loss of ions, a sufficient amount of positive charge must eventually collect within the beam to neutralize the electron charge completely.

Any excess ions that are created constitute a net positive charge which results in a radial "hill" of potential down which these surplus positive charges are drained to the drift tube. Hence,

under the assumption of no loss, the number of ions in the beam must closely approach the number of electrons if a sufficient amount of gas is available for ionization.

B. Minor Causes of Ion Loss

The assumption that no ions are lost from the beam space is, of course, not justified. Several processes which cause a continual removal of ions from the drift tube may be listed; their effects can be shown to be negligible by consideration of their magnitudes:²

1. Recombination with electrons to form a neutral molecule is a process which is continually taking place within the beam. However, at pressures in the vicinity of 10^{-7} mm Hg, the rate of ion formation far exceeds the rate of recombination.

2. The thermal energy of the ions is in the vicinity of $1/40$ th of a volt and is negligible when compared to the normal potential variations across the beam which tend to hold the ions together. The loss due to thermal energy might become important as one approaches the state of complete neutralization because, under such conditions, the potential variation across the beam is very small.

3. The impact energy imparted to the ion by the colliding electron is small when compared to the potential drops in the beam. Like the loss due to thermal energy, this cause of ion removal will not assume serious proportions until a nearly neutralized condition is reached.

4. An insulator in the vicinity of the beam becomes negatively charged by picking up stray electrons and hence constitutes a sink for positive ions. However, most electron beams are surrounded partially or completely by metallic structures and are therefore shielded from the effects of the sinks described above. The effectiveness of the drift tube as a shield determines the extent to which these charges will affect positive-ion neutralization.

C. The Major Cause of Ion Loss

It would seem from the above discussion that electron beams enclosed in metallic drift spaces should reach total neutralization even at pressures as low as 10^{-7} mm Hg since at such pressures a sufficient number of gas molecules is available to form the required number of ions.

That this is not the case was shown by L. M. Field¹, who disclosed a major source of ion loss which had not been considered previously. The accelerating field from the cathode-anode region penetrates into the drift space via the anode aperture through which the beam must enter. This field is the strongest at the axis near the aperture and is of such a polarity as to remove ions

from the parts of the beam near the aperture. The removal of the ions at the head of the beam leaves a low potential region to which ions must flow from parts of the beam further down the drift space. Thus the effect of the draining field is felt for quite a distance into the beam.

The penetration of the cathode-anode field into the drift space will therefore prevent the beam from reaching complete neutralization. In fact, it can be shown³ that only a slight amount of neutralization is possible at the beginning of the beam when pressures better than 10^{-7} mm Hg are considered. Therefore, the spread of an electron beam, under the pressure conditions normally available in commercial tubes, may be taken to be practically the same as in perfect vacuum.

III. The Elementary Theory of Ion Trapping

The principle behind the elimination of axial drainage may first be introduced by considering a thin beam of electrons flowing through a comparatively large drift space as shown in Fig. 1. Let it be our intention to obtain a completely neutralized beam at all points to the right of some section AA. Some means is then necessary to prevent the positive ions from flowing across the section toward the cathode. Under such circumstances, all the ions created in the beam to the right of AA will tend to be held within the beam boundaries. It would then seem that the removal of ions can only be caused by the minor processes mentioned previously and the rate of removal due to these is so small that a sufficient number of ions should collect to neutralize the beam completely even at good commercial pressures.

In order to understand the stopping of ion flow across a given section, the potential distribution in that vicinity should be examined. On the left-hand side of AA, the positive ions are drawn rapidly away toward the cathode so as to leave that part of the beam unneutralized. The potential on the axis there will always be less than V_0 , the drift tube voltage, by some amount which depends upon the parameters of the beam. If the beam were to begin holding ions to the right of AA, the potential on the axis to the right would approach the value V_0 depending on the extent of the neutralization. For the case of full neutralization all points to the right of AA should reach the potential V_0 , since no net negative charge remains.

In order that a field tending to retain ions be obtained at section AA under all conditions of neutralization, the potential on the axis there must be raised above the value V_0 so that a hump of potential is interposed in the path of the ions.

A grid held at the drift tube potential is the simplest means of obtaining the potential hump. It will be clear subsequently that this is the only method which can give satisfactory ion trapping. Nevertheless, considerations of power dissipation usually forbid its use. Aperture plates, cylinders, and ring electrodes concentric with the beam have been suggested as alternative means of obtaining the required variation of potential along the axis. Such gridless type of trapping is the subject of discussion in this article. It is found that many complications arise when this method of trapping is used and, as a result, the design of such a system

of beam control becomes extremely complex. The simple theory outlined above needs to be extended and modified in order to explain the behavior of gridless ion traps.

IV. Experimental Data on Gridless Ion Trapping

The following experimental check was offered at the time when ion trapping was first suggested¹. An initially parallel beam of 130 ma at 7300 volts, which normally spread to twice its starting diameter in 36 cm, was prevented from spreading by the application of a trapping voltage of 15 volts at its entrance aperture. An increase in the trapping voltage showed neither an adverse nor a beneficial effect on the beam.

Attempts have since been made to use ion trapping for holding together a beam of electrons and no appreciable degree of success has been reported.⁴ It was therefore decided to perform an experiment which would test the various facets of gridless ion trapping and perhaps yield some data which might give some indication as to the nature of the phenomena taking place.

The schematic diagram in Fig. 2 shows the essential parts of the equipment used in the experiment. The beam-testing tube⁵ consists of a bombarded tantalum cathode injecting a beam of perveance 1.07×10^{-6} and an entrance angle of 2.5° with an initial diameter of 2.54 cm into a drift tube of the same diameter. Its anode is followed by ring electrodes, numbered one to six, which are to be used for trapping. The position of the trapping section may be varied by selecting one out of the six electrodes. The rest of the drift tube is broken up into six sections and a collector which are insulated from each other by glass seals.

Curves of the percentage of total current reaching the collector as a function of the trapping voltage, for different positions of the trapping section are given in Fig. 3. A glance at these curves indicates a definite pattern of behavior which cannot be predicted by the simple theory mentioned previously.

V. Detailed Consideration of Gridless Trapping

The process of ion trapping as disclosed in Section III is assumed to be a purely axial phenomenon; the movement of the ions and the insertion of the potential hump are both considered as being along the axis. While these assumptions may hold for thin beams of low perveance, a more general theory must consider trapping as a volume phenomenon. Therefore, one needs to look more closely at the fields involved in the region around the trap and the effect which these fields have on the distribution of ions.

A. Potential Distribution Around the Trapping Electrode

The potential in the drift tube is the result of contributions from three different causes. First, the positive voltage on the trapping electrode raises the potential at every point in the drift space. Second, the presence of the negative particles in the beam tends to lower it. Third, if the combined effect of these two causes results in

trapping positive ions, the presence of these positive charges will modify the potential due to the other two depending on the particular distribution of the ions.

The potential distribution due to the voltage V_t on the trapping electrode (shown in Fig. 4) can be obtained by electrolytic tank measurements while that due to the beam may be got approximately by solving the problem of a cylinder of charge coaxial with a conducting wall. Since potentials are additive, these two may be added numerically or graphically to obtain the resultant potential distribution.

It will soon be clear that the only equipotential contour relevant to our discussion is the one having the same potential as the drift tube. Such an equipotential must exist because there are points within the drift space where the increase in potential due to the positive trapping voltage just cancels the drop due to the negative space charge in the beam. The net potential at these points, in the absence of ions, will be equal to the drift tube potential.

The location of this contour line depends on the value of the trap voltage and the parameters of the electron beam being used; a separate set of calculations must be gone through for each set of conditions. Figs. 4 and 5 show contours calculated for three values of trap voltage and a 2000-v beam of perveance 1.07×10^{-6} , just filling the drift tube.

The contribution made to the potential distribution by the presence of positive ions will alter the location of the V_0 contour; but the position of this contour, as will be seen subsequently, determines the distribution of positive ions. This interdependence makes it impossible to determine the actual distribution of the ions and the final position of the V_0 equipotential. Consequently, the rest of the discussion is based on qualitative argument rather than formulas and numbers.

B. The Trapping Phenomenon

At all points in the beam, positive ions are continuously being formed with potential energies which depend on the value of the potential at the place where they are created. Ions which are formed at a place where the potential is high with respect to its surroundings, will tend to flow away from the point in much the same way as water flows away from tops of hills. The rate of formation of ions at pressures of the order of 10^{-7} mm Hg is so slow that, under such pressure conditions, the rate of ion drainage due to even the smallest fields can exceed the rate of formation, and high potential regions, such as the one mentioned above, may be taken to be virtually free of ions. Conversely, ions formed in places which are surrounded completely by a surface of higher potential cannot have sufficient energy to pass this bounding surface and hence must remain trapped therein.

The distribution of ions within a given beam space will therefore depend on the nature of the potential contours within that space. In order to get a qualitative idea of the disposition of the ions at different values of trapping voltage, let us consider the potential contours shown for the three cases in Fig. 5.

Figure 5(a) shows that contour line which should, by the simple theory of Section III, have trapped enough ions completely to neutralize the beam to the right of the trap. The portion of the diagram shown shaded is a region where the potential is higher than V_0 and hence all the positive ions formed in that region are drained away either to the walls or to regions of lower potential. The shaded region cannot therefore have any of its space charge neutralized.

The space which is left unshaded is essentially one of potential below V_0 and is also a region which is completely surrounded by the V_0 equipotential and hence isolated from the draining field of the cathode. Positive ions found in this space will remain there and begin to neutralize the electron beam. As soon as neutralization starts, the V_0 equipotential begins to shift toward the right because of the accumulation of positive charge to the right of it. This makes the allowable trapping region move further and further down the drift tube as more and more positive ions become trapped. Ions will continue to be trapped until the whole space enclosed by the new V_0 equipotential has reached the drift tube voltage. After this, the positive charges will begin to move out to the drift tube and be lost.

Figure 5(b) shows the potential distribution when the value of V_t is below the minimum prescribed by the simplified theory of Section III. The space shown cross-hatched is a region with potentials above V_0 and hence must be essentially free of ions. The area marked with dots is below V_0 in potential but has a direct connection with the unneutralized low potential part of the drift tube to the left of section AA where the cathode is continually removing ions. This region, therefore, cannot retain ions.

Trapping can only take place inside a space enclosed completely by some such potential contour as is shown dotted and dashed in Fig. 5(b). Let this potential be denoted by xV_0 where x is less than one. As before, when trapping begins, the xV_0 potential will begin to shrink because of the positive space charge enclosed within it. Some final equilibrium distribution will be established, consistent with the above considerations. Under such conditions, the neutralization cannot be very great, first, because the trapping space must needs be small, and second, because trapping takes place within a contour of potential lower than V_0 . This is analogous to the fact that a shallow bowl holds less water than a deep one.

Figure 5(c) illustrates the potential distribution for high values at V_t . A little reasoning will show that, the higher one takes V_t , the farther right does the V_0 equipotential shift, thus increasing the size of the unneutralized portion of the beam.

The treatment of the three cases above gives an idea of the disposition of the trapping zone with respect to the electrode for different values of voltage. It must not be imagined, however, that ions attain a uniform density at all points inside the trapping region. Their behavior in this space is more a matter of conjecture than of accurate mathematical calculation. The phenomenon might, perhaps, be more clearly understood if one were to liken the trapping zone to a sunken area of ground

being filled up by a slow seepage of water. The deepest portions of the hollow will, of course, be filled first and eventually will hold the largest depths of water, while the surrounding areas near the lip of the hollow will be filled last and will be in small depths of water. In our electrostatic case, matters are further complicated by the effect which the positive ions themselves have in altering the potential distribution in the surrounding domain. Despite this complication, the same general reasoning holds and the qualitative picture remains the same.

C. Effect on Beam Transmission

It has been established above that, for a given trapping voltage, there exists within the drift tube a surface which encloses a certain volume of the drift space inside which positive ions are held. The positive ions will accumulate rapidly in that part of the beam until the whole volume enclosed by the surface mentioned is practically at uniform potential, after which the excess of positive ions will migrate to the drift tube wall.

An electron entering this neutralized space will be entering an equipotential region and must therefore maintain the magnitude and direction of the velocity with which it enters. The slope of the trajectory of each electron as it crosses into the trapping region is hence a parameter of great importance in determining the subsequent behavior of the beam.

1. The Perfect Ion Trap. It would be best to start out by examining what is needed in order to obtain a truly parallel beam by means of ion trapping. In view of what has been said above, the principle of focusing by trapping positive ions may be enunciated in one sentence. The various electrons comprising the beam must be made to enter the unipotential region when their paths hold zero slope. If one can obtain this condition, all the electrons of the beam must continue to travel from there on with zero slope since no radial forces exist in a unipotential region. Perfect transmission should therefore be obtained if the minima of the various electron trajectories are made to coincide with the position of the trapping surface.

2. Realization of the Ideal Conditions. In order to understand the extent to which perfect trapping may be achieved, the following two points must be considered:

(a) The contour behind which ions are trapped is some surface of revolution (similar to the ones shown dotted in Fig. 5) whose actual shape and position cannot be determined by analytical, graphical or numerical methods.

(b) In order that the minima of the trajectories of the electrons of the beam may be made to coincide with the position of the trapping surface, the location of these minima should be accurately known. The most elaborate treatment of this problem available to date⁶ suffers from two drawbacks. First, an estimate of the position of the beam minimum obtained from this treatment is only an approximate one due to the assumptions inherent in setting up the problem. Second, the treatment only deals with the approximate behavior of the

electron at the edge of the beam; no mention is made of the behavior of the electrons within the beam envelope. The reduction of electron velocities which accompanies the drop in potential in an unneutralized beam will cause a piling up of charge in the regions where the velocity is reduced, i.e., the charge density will be higher near the axis than at the edge of the beam. The inner electrons will be subject to greater repelling forces since their trajectories enclose a larger average charge density, and hence will go through their minima quicker than the outer electrons; therefore, the minima will not all lie in the same cross sectional plane.

Figure 6 shows a sketch of the relative positions of the trapping equipotential and the beam minima as they might possibly occur at some stage during the trapping experiment. It is also shown in the figure that the different electron trajectories enter the neutralized zone at different slopes and hence behave differently subsequent to their entry. It must then be apparent that while the trapping surface is in the appropriate position for some parts of the beam it must, by the very nature of its curvature, be improperly situated for the rest of the beam. Therefore, while electrons at a particular radius will remain parallel to the axis, the rest of the electron paths are at different slopes. Cross-overs could very well occur and complete transmission is improbable.

VI. Explanation of the Experimental Results

Inasmuch as a strict mathematical treatment of the trapping phenomenon is not possible, quantitative estimates cannot be made to match with the values of transmission obtained by experiment. But the theoretical treatment of Section V can predict a definite trend of behavior which seems to be borne out by experiment.

A. Effect of Trapping Voltage

As the trapping voltage is increased from zero a quick rise in beam transmission is first observed followed by a slower fall. This is not entirely unexpected when one considers the phenomena outlined in connection with Fig. 5.

At really low trapping voltages, the draining effect of the cathode field can still leak into the drift tube beyond the trapping ring and only partial neutralization can exist as was explained in connection with Fig. 5(b). As one increases the trapping voltage, the larger value at V_c offers better opposition to the cathode drainage and transmission improves.

However, excessive trapping voltages have the effect of pushing the neutralized region farther down the drift space. The slope of the beam under such circumstances would be positive as it enters the trapping region, thus resulting in a diverging beam.

Between these two effects, there obtains some optimum condition where the trapping surface is in the best possible position and the transmission reaches its greatest value before beginning to fall. Such a peaked behavior is shown by each of the experimental curves.

B. Position of Trap

It is interesting to observe the effect of shifting the position of the trapping electrode. It is seen that, as one moves the trapping section closer to the cathode, the voltage V_t required to achieve the maximum transmission becomes higher. This is probably because as one gets closer to the cathode its draining effect becomes greater and one needs larger trapping voltages to counteract the unwanted fields. In this connection it should be remembered that in order to cause a given change in voltage near the axis, a considerably larger voltage alteration is required at the ring electrode.

Another effect seen from the family of experimental curves is that the value of maximum transmission obtained is different for different positions of the trapping ring. One merely needs to refer to Section V-C-2 in order to find the explanation for this effect. There, it was pointed out that the position of the trapping surface relative to the beam minimum has considerable effect on the resultant transmission. If the trapping surface falls in the region where the beam is diverging, the electrons will continue on with the same positive slope and only a slight improvement can be expected. When the trap voltage is applied to the electrode close to the cathode, the beam is still converging as it enters the trapping zone and should continue converging until it crosses over and then diverges. The position of the trapping ring which brings the trapping surface closest to the minimum will give the best possible results.

VII. Concluding Remarks

It has been found possible to obtain an improved understanding of the process of ion trapping by considering it as a volume phenomenon. A detailed examination reveals that this method of controlling beam spread involves more complications than are accounted for in the simple theory which is only representative of the special cases of thin beams of low perveance.

The conditions for perfect trapping are simply enunciated; however, it is also realized that it is not very easy to achieve them with gridless traps. The main difficulty lies in the fact that the designer is not able to predict the distribution of the positive charges within the beam; as a result of this, the position of the unipotential region is no longer a matter of deliberate design. Hence it is not possible to obtain a trap configuration which will assure that the various electrons in the beam will enter the trapping region

precisely at the moment when the slopes of their trajectories are zero.

There have been instances of very efficient trapping with gridded drift tubes used in some experimental klystrons. This is certainly understandable in view of what has been discussed; because, by positioning the grid at the estimated minimum of the beam, the trapping equipotential can be forced to stay close to the minima.

In the absence of a proper analytical design, several gridless electrode configurations can be tried out experimentally. Chances of success should be better in the case of those electrode configurations whose potential distributions come close to that of a grid (at drift tube potential) stretched across the beam at its point of minimum. One such example would be that of an Einzel⁷ lens around the minimum of a beam whose diameter is small compared to the drift space.

A pilot experiment, which was performed in order to check the validity of these suggestions, showed a large improvement in beam transmission due to the use of a large ratio of drift-tube diameter and a trapping configuration similar to that of an Einzel lens. A beam of perveance 2.6×10^{-6} and a minimum diameter of $5/32$ in. was injected into a drift tube of $1/2$ in. diameter and 6 in. long. For a beam voltage of 300 volts, over 80 per cent of the cathode current was received at the collector when 95 volts were applied to a trapping electrode of the Einzel type situated near the minimum of the beam.

It is not unreasonable to expect that subsequent trials with different types of trapping devices could yield further improvements in beam transmission.

References

1. L. M. Field, K. Spangenberg, and R. Helen, "Control of electron beam dispersion at high vacuum by ions," *Electrical Communication*, 24: 108-121 (1947).
2. L. M. Field, Ph.D. dissertation, Stanford University, 1944.
3. B. H. Wadia, unpublished work.
4. Private communication with workers at Varian Associates, San Carlos and Palo Alto, Calif.
5. The construction and processing of this beam-testing tube was made possible through the co-operation of Eitel-McCullough, Inc., San Bruno, Calif.
6. L. P. Smith and P. L. Hartman, "Formation and maintenance of electron and ion beams," *J. Appl. Phys.*, 11:220-224 (1940).
7. K. R. Spangenberg, *Vacuum Tubes*, p. 386, Van Nostrand Co., New York, 1949.

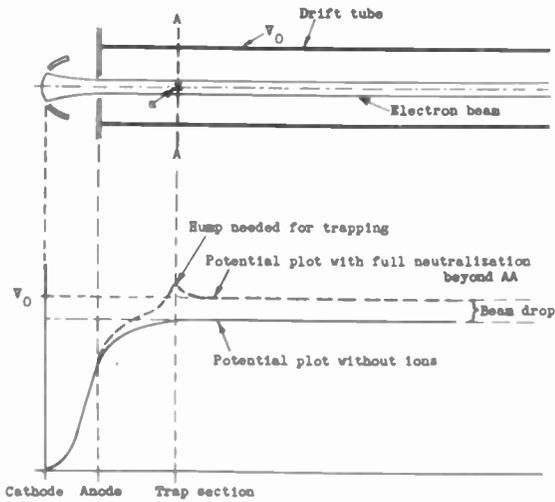


Fig. 1 - The principle of positive ion trapping.

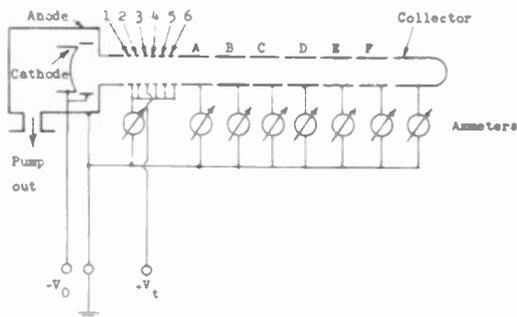


Fig. 2 - Beam-tester schematic.

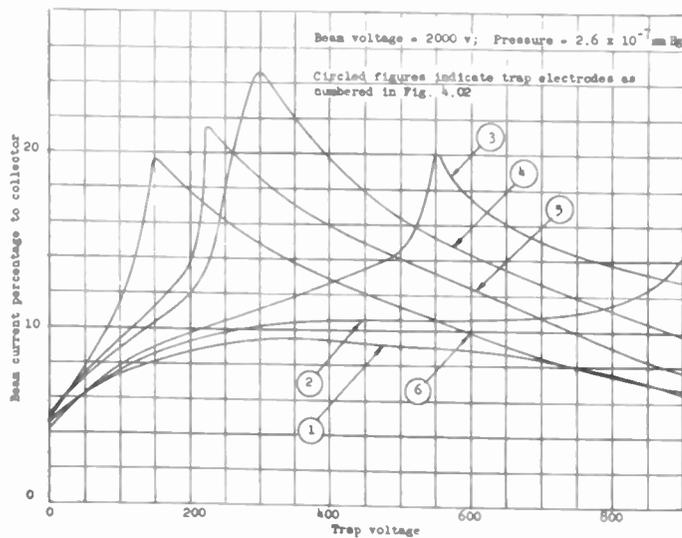


Fig. 3 - Results of the ion trapping experiment.

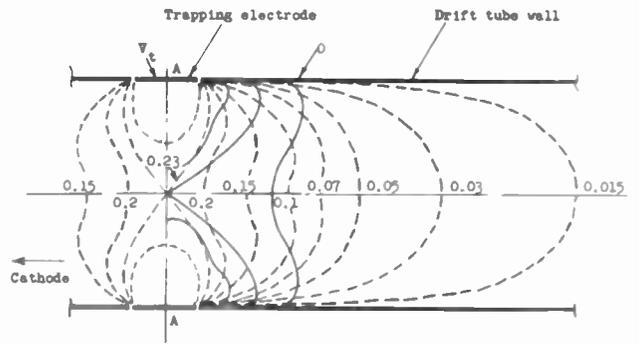
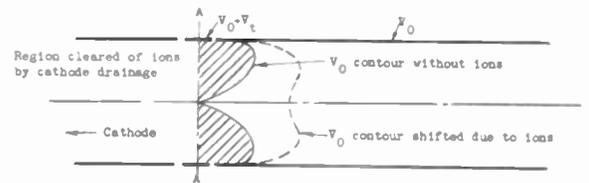
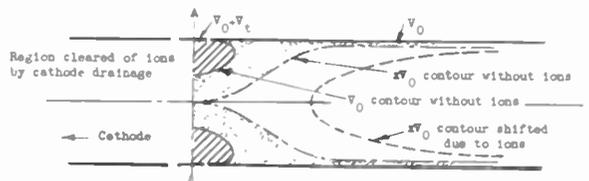


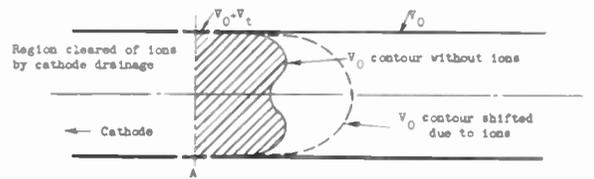
Fig. 4
Potential distribution around trapping electrode.



(a)



(b)



(c)

Fig. 5
Location of the trapping surface: (a) V_t adjusted to just start trapping; (b) low value of V_t ; (c) high value of V_t .

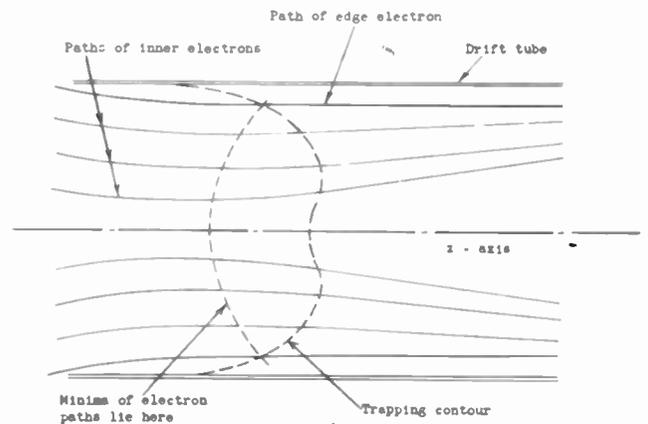


Fig. 6
Position of the trapping surface with respect to electron trajectories.

THE MULTIPACTOR EFFECT IN KLYSTRONS*

Kees Bol
Sperry Gyroscope Company
Division of The Sperry Corporation
Great Neck, New York

Summary

The multipactor effect in klystrons absorbs useful energy and is most undesirable. This paper describes a study of the problem undertaken with the ultimate aim of learning to suppress the effect.

Introduction

One of the problems of klystron design is avoiding what is known as the multipactor effect. This effect is a greatly enhanced form of the secondary loading that occurs to some extent in all klystron interaction gaps. It is characterized by its dependence on the strength of the r-f electric field in the gap and also on the strength of the magnetic focusing field where this is present.

The kinematic explanation of multipactor is that at suitable r-f field strengths secondary electrons are driven across the gap in, roughly, an odd number of half cycles. If these electrons then create further secondaries on impact, they will in turn be driven across the gap and so continue the process. Even if the multipactor ratio is somewhat less than unity it is clear that for each secondary electron created by beam interception the total power loss will correspond to many more. If the ratio exceeds unity the process becomes more or less independent of beam interception, and can in fact occur without any beam at all. Fig. 1 shows some data taken on a resonator with the beam off. The parameter is magnetic field strength, which is obviously of crucial importance.

Although the multipactor effect does occur in klystrons not having magnetically controlled beams, it is rarely serious in these cases and can usually be eliminated without difficulty. It is most serious in high power tubes with magnetic beam control and in these the effect is not only most severe but also most difficult to eliminate.

No analysis of the multipactor effect is as yet sufficiently detailed to allow prediction of power loss as a function of gap configuration, secondary emission coefficients, beam interception, r-f voltage, and magnetic field strength. Even for the simplest case of multipactor between two infinite parallel plates the present theory allows little more than prediction

of the necessary gap voltage. With a more realistic geometry and with a magnetic field present, even less quantitative statements are possible. But in making even qualitative statements about an actual case, it is of course very useful to have studied particular cases, however idealized.

To this end the following cases will be briefly discussed: (a) the parallel plane case; (b) a symmetric gap configuration producing a strong radial component of r-f field, with magnetic field; and (c) the coaxial cylinder case. Finally there will be presented some test results illustrating the basic role of the magnetic field in actual cases.

Multipactor Theory

Parallel Plane Case

The theory for this case was worked out by Abraham¹ who first noticed the effect in klystrons. An r-f voltage of amplitude V is applied across two infinite parallel planes a distance d apart. The equation of motion is

$$\frac{d^2x}{dt^2} = \eta \frac{V}{d} \sin \omega t$$

which when integrated twice yields:

$$x = \frac{\eta V}{\omega^2 d} [(\omega t - \omega t_0) \cos \omega t_0 + \sin \omega t_0 - \sin \omega t] + v_0(t - t_0)$$

Here η is the charge to mass ratio of the electron, t_0 the departure time, and v_0 the emission velocity. The multipactor condition requires that $x = d$ when $(\omega t - \omega t_0) = m\pi$ where $m = 1, 3, 5, \dots$. Hence the voltage amplitude required is

$$V = \frac{\omega^2 d^2}{m\pi\eta} \frac{\cos \omega t_0 + \frac{2}{m\pi} \sin \omega t_0}{1 + \frac{m\pi v_0}{\omega d}}$$

¹W. Abraham, "Interactions of Electrons and Fields in Cavity Resonators," Stanford University Doctoral Thesis, 1950.

*The work described in this paper was sponsored by the U.S. Navy on contract NOnr-217(01).

If $0 < \omega t_0 \ll 1$ and $m\pi V_0 \ll \omega d$ then

$$1. \quad V_m \approx \frac{\omega^2 d^2}{m\pi\eta}$$

The assumption on starting phase can generally be justified², but not the neglect of emission velocity. This is easily seen as follows: If u_0 is the velocity of the klystron beam, then the gap spacing is usually so chosen that $\omega d/u_0 \approx 1$. Hence

$$\frac{m\pi V_0}{\omega d} = m\pi \left(\frac{u_0}{\omega d} \right) \left(\frac{V_0}{u_0} \right) \approx m\pi \frac{V_0}{u_0}$$

So for a 1000 volt beam and 5 volt emission energy the above quantity is about $m/5$, which is certainly not negligible except for perhaps the lowest mode. Nevertheless, when d is the least separation of two complicated surfaces equation (1) is a useful indicator of the approximate voltage at which multipactor might set in.

It is instructive to compare the relationship demanded by equation (1), with that existing in a typical klystron gap. If V_0 is the beam voltage, u_0 the beam velocity, then

$$\frac{V_m}{V_0} \approx \frac{2}{m\pi} \left(\frac{\omega d}{u_0} \right)^2 \sim \frac{2}{m\pi}$$

Since the gap voltage for efficient operation is comparable with beam voltage the condition of equation (1) will usually be satisfied for all modes. However, if the gap is widened so that $\omega d/u_0 \approx 1.5$, then $V_m/V_0 \approx 1.5/m$ and the lowest mode can no longer be reached.

Mere compliance with equation (1) is fortunately no sign that multipactor actually will occur. A second fundamental requirement is that the energy gained by a given electron must be high enough so that it can cause emission of secondaries. An electron emitted at $t_0 = 0$ with zero velocity and having a transit angle of $m\pi$ has a velocity $x = 2 \eta V/\omega d$ or an energy in volts given by:

$$V_m' = 2 \frac{\eta V_m^2}{\omega^2 d^2} = \left(\frac{m\eta V_m}{\omega^2 d^2} \right) \frac{2}{m\pi} V_m = \frac{2}{m\pi} V_m$$

Since $V_m \sim m^{-1}$ it follows that the energy gain is proportional to m^{-2} . For example,

²ibid.

if $V_1 = 1000$ volts then $V_1' = 650$ volts, whereas $V_3' = 70$ volts. This is so low that multipactor would be most unlikely.

A more complete theory for the parallel plane case, of which the above is a short abstract, can be found elsewhere.^{3,4}

Parallel Planes With DC Bias

The effect of a bias voltage applied across a gap is to make the transit times for going and coming electrons different. By making the bias high enough it is possible to prevent multipactor altogether. The analysis is rather lengthy and will therefore not be reproduced here.⁵ The conclusion reached as to the amount of bias necessary to suppress multipactor in the various modes is as follows:

$m = 1$	$V_{DC} = .34 V_1$
$m = 3$	$V_{DC} = .12 V_3$
$m = 5$	$V_{DC} = .07 V_5$

For the less important higher modes the bias is therefore quite small, but for the lowest mode it is material.

Coaxial Cylinder Case

The interest here is the same as in the preceding paragraph: that by creating sufficient asymmetry in the gap multipactor might be prevented. Coaxial cylinders present asymmetry of the desired type in a convenient form. But since the equation of motion, while simple, is non-linear a large number of trajectories were plotted with an analog computer. Only the lowest mode was studied and the results are shown in Fig. 2. It is evident that a great deal of asymmetry is necessary to be effective - more, indeed, than can be built into a gap configuration.

Magnetic Focusing Effects⁶

Qualitatively the dependence of multipactor on magnetic field strength is easy to understand. If for example the electric field and the magnetic field are both axial, then the transverse component of the emission velocity leads to a circular path in a transverse plane while the axial motion remains unchanged. The frequency of the circular motion, on the other

- - - - -

³ibid.

⁴Multipactor study reports furnished to U.S.N. Office of Naval Research by the Sperry Gyroscope Company.

⁵ibid., Appendix B

⁶ibid., Appendix D

hand, is determined only by the magnetic field and is in fact just the cyclotron frequency. So if the magnetic field strength is such that electrons undergo an integral number of rotations while traversing the gap then they are exactly focussed.

If there is a radial electric field in the gap, then the transverse motion becomes cycloidal. The frequency is no longer exactly the cyclotron frequency unless the transverse emission velocity is zero. There is consequently no exact focussing, although optima should still occur at about the same field strengths as in the case of no radial field. The principal change caused by the radial field is the increase in the extent of the radial motion. With ordinary operation the increase amounts to one or two orders of magnitude above the radial motion caused by the emission velocity.

The results of some simplified calculations are shown in Fig. 3. It was assumed that the radial motion was slight (a very rough approximation); that the emission velocities were negligible; that the gap geometry was symmetrical; that the radial field strength varied linearly across the gap; and that the axial motion was the same as in the parallel plane case. The figures show the displacement of the electrons, on arrival at the opposite gap face, as a function of magnetic field strength. Plainly the higher the multipactor mode (characterized by transit time in r-f cycles) the lower the field strength needed to produce a focus. This follows at once from the fact that longer transit time in the gap permits a longer period of transverse motion, or a weaker field. At the other extreme, as the field becomes stronger the transverse motion ultimately vanishes so that "brute force" focussing takes place for all magnetic field strengths above some minimum.

Experimental Results

Multipactor behavior in the absence of a magnetic field is fairly well explained by the parallel plane theory.⁷ In part this is so because multipactor unrestrained by a magnetic field simply does not occur unless the actual geometry closely resembles the parallel plane case. In the presence of a magnetic field, however, the situation is drastically altered. The purpose of the data to be presented is to make clear the extent of the change.

The resonator used is shown in Fig. 4. It was designed to allow the resonance frequency to be kept constant at

- - - - -

⁷Abraham, Op. Cit.

about 5 kmc while the gap geometry was varied over a wide range, and this was achieved. The power input was varied up to about 200 watts. The maximum magnetic field strength was about 2000 gauss in all but a few cases in which it was nearly 2500 gauss. To make the origin of the multipactor visible a very thin coat of willemite was painted on the center posts, with fair results.

In Fig. 5 the data are presented as regions of multipactor in a power input-magnetic field strength plane. The results are striking in several respects. There is scarcely any dependence on gap voltage, provided the power input exceeds about 20-40 watts; there is remarkably little dependence of either starting voltage or critical magnetic field strength on gap geometry; and the apparent spacings involved in the "A" mode are fantastically large.

The relative unimportance of gap voltage on the multipactor is readily explained. The constraining influence of the magnetic field allows multipactor to occur between the sides of the center posts, and hence there is no unique spacing as in the parallel plane case. The distance traversed in a given mode simply increases as the voltage increases, while simultaneously lower order modes become possible at closer spacings. Furthermore, the axially directed electric field is no longer uniform with distance so that the parallel plane theory loses practically all validity.

According to the parallel plane theory the starting voltage for a given multipactor mode should vary as d^2 . Therefore, if the cavity impedance were independent of gap spacing the minimum power required should vary as d^4 . Thus a 20 per cent change in spacing should cause a 100 per cent change in minimum power. Instead it is found that the minimum power is determined primarily by the secondary emission coefficient of the surface involved. For the particular test results given here multipactor appeared to require a gap voltage of 400-600 volts to start, and continued for higher voltages.

The most striking feature of the results is the persistence of the magnetic modes labeled "A" and "B" in the figure, and the appearance in each case of what is labeled the "gas discharge" region. The "B" mode does not fit the theory of magnetic focussing represented by Fig. 3. This is not too surprising since that theory was based on an axial displacement similar to that for the parallel plane case. The "A" mode occurs quite close to $\omega_c/\omega = 1$, and possibly the deviation from unity is the result of an error in field calibration.

At $\omega_c/\omega = 1$ the simple theory indicates focussing for all modes. Furthermore, it is possible to have resonance with the radial component of electric field at this value of ω_c . In other words, the "A" mode may correspond to radial motion rather than axial motion of the electrons. The "gas discharge" may mark the exact point of resonance. It should be noted that the ambient pressure in the cavity was of the order of 10^{-6} mm Hg, and that the pressure necessary to account for the observed brilliant discharge could only have arisen from intense multipactor bombardment of the willemite.

Finally it must be stated that the extent of the multipactor region depended primarily on the secondary emission coefficient of the surfaces. Without the willemite multipactor always cleaned up after prolonged operation. Even with the willemite some low field modes sometimes disappeared in a few seconds. (These modes have not been shown on Fig. 5.)

Conclusion

The parallel plane theory of multipactor predicts critical voltages at which multipactor can occur. For gridded gaps this theory appears adequate. For gridless gaps in a strong axial magnetic field the situation is far different. There

then appear critical magnetic field strengths at which multipactor occurs for all sufficiently high voltages. An adequate theory for this case does not yet exist. Experimentally the critical magnetic field strengths seem to depend very little on the geometry of the gap. It also appears that self-sustained multipactor does not occur in a gridless gap when the surfaces are thoroughly outgassed. However, if beam interception provides a supply of secondaries, then at certain magnetic field strengths there is certain to be some amount of multipactor loading.

The means of minimizing multipactor implied so far can be summed up as follows:

1. Minimize beam interception.
2. Adjust the design so that the beam focussing magnetic field is not critical for the gap.
3. Increase gap spacing to raise multipactor starting voltage above operating gap voltage.
4. Apply a d-c bias voltage to the gap.
5. Reduce the secondary emission coefficient of the gap surfaces.

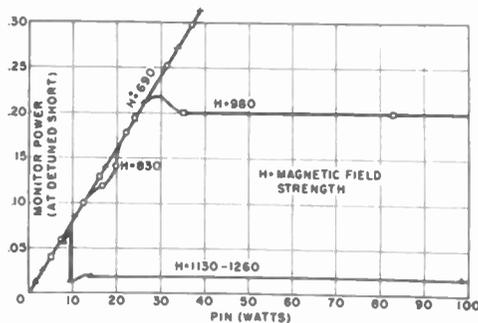


Fig. 1 - Cold multipactor data.

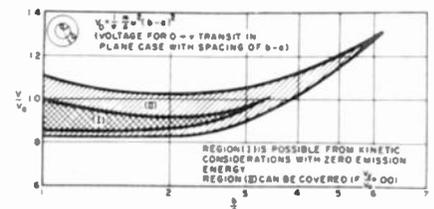


Fig. 2 - Coaxial multipactor.

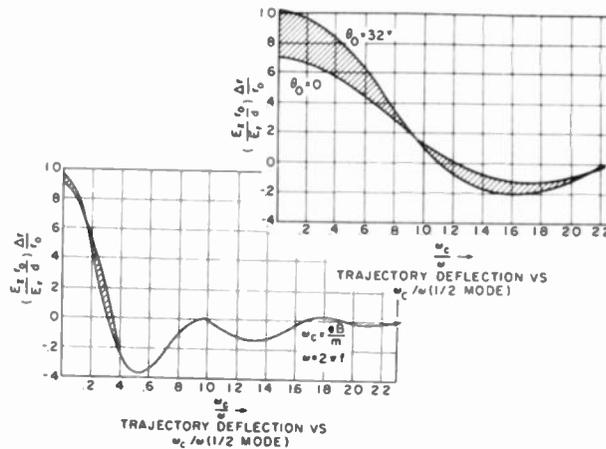


Fig. 3 - Trajectory deflection vs. ω_c/ω .

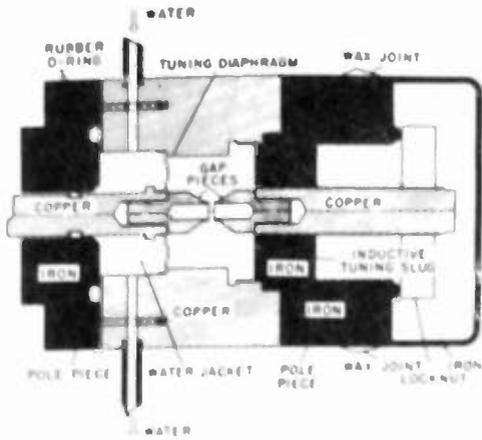


Fig. 4 - 5-kmc multipactor test cavity.

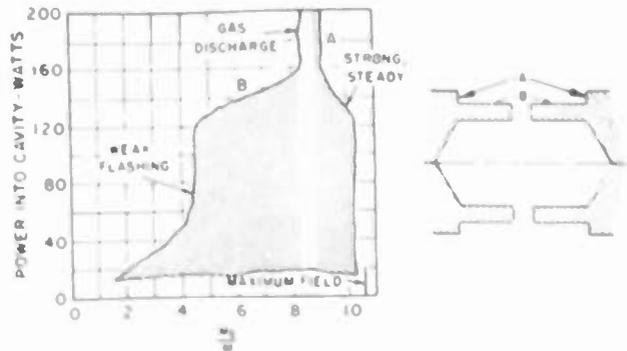


Fig. 5(b) - Results obtained with test cavity.

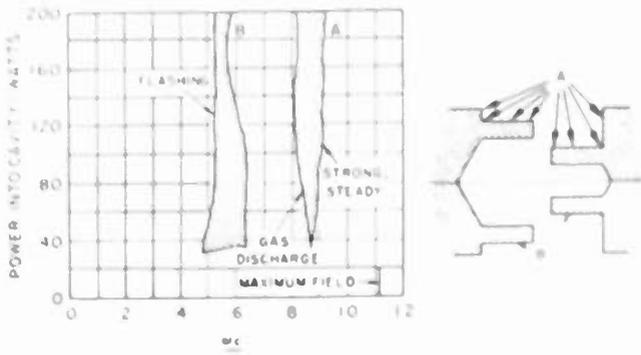


Fig. 5(a) - Results obtained with test cavity.

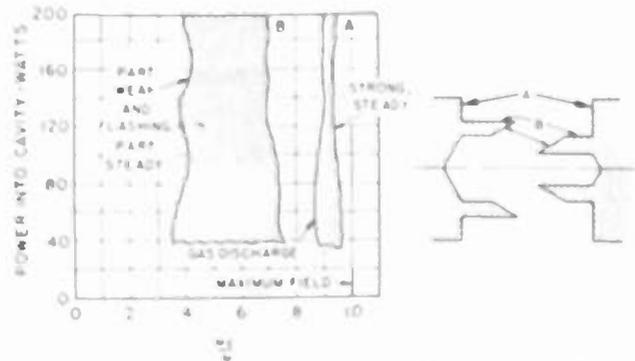


Fig. 5(c) - Results obtained with test cavity.

BACKWARD-WAVE OSCILLATOR CHARACTERISTICS

H.R. Johnson
Hughes Research and Development Laboratories
Culver City, California

ABSTRACT

This paper contains theoretical and empirical design information on backward-wave oscillators with power outputs below 100 watts. The main design parameters which are discussed are tuning range, power output, and frequency stability. All the experimental work has been carried out with the he-

lix-type tubes with hollow beams. Tuning ranges of 9:1 have been obtained, with reasonably constant output power over a 2:1 range. One tube oscillated from 500 to 4,500 mc, another from 7,000 to 14,000 mc. At 3,000 mc, 30 watts output has been obtained for 400 watts input.

THE PROPAGATION PROPERTIES OF CROSS-WOUND TWIN HELICES SUITABLE FOR TRAVELING-WAVE TUBES

M. Chodorow, E.L. Chu, and J.R. Nevins, Jr.
Microwave Laboratory
Stanford University,
California

ABSTRACT

A cross-wound twin helix used as a propagating structure for traveling-wave tubes has advantages over the conventional single helix, particularly for high-voltage operation. Because of the symmetry of the fields in a twin helix, impedance of the fundamental component is much higher than for the single helix, and impedance of the back-

ward-wave harmonics is much less. This will improve the gain in amplifier operation and make backward-wave oscillation much less apt to occur for the twin helix. Calculations of propagation constants and impedances for a range of dimensions have been made, and measurements of these properties are in good agreement with the theory.