RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

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

VOLUME XXII | MARCH 1961 | NUMBER 1

CONTENTS

RCA REVIEW is regularly abstracted and indexed by *Industrial Arts Index Science Abstracts* (I.E.E.-Brit.), *Electronic Engineering Master Index, Chemical Abstracts, Proc. I.R.E.*, and *Electronic & Radio Engineer.*

RCA REVIEW

BOARD OF EDITORS

FOREWORD

IN THIS ISSUE of the *RCA Review* there are published a number of papers on a subject of rapidly increasing interest and importance—electrical energy conversion.

For many years our civilization has depended for most electrical energy on power supplied by fossil fuel to a heat engine or, to a lesser extent, on water power. In either case, the energy has come from the sun through a fairly involved process, and the over-all efficiency of conversion from solar to electrical energy is extremely low. More fundamentally, the power received from the sun originates in nuclear energy stored in the material of which the sun is constituted.

Methods for producing electrical power have not changed materially during the past half century. The steam engine, heated by fossil fuel and driving a rotating generator, has been the backbone of our power plants. During the past two decades, however, a number of technological strides have been made which can revolutionize the source of our power, the means of converting from the primary energy to electricity, and the efficiency of utilization of this energy. As a result of these advances, we can now anticipate some rather radical innovations; among the more interesting of these are the following:

1. Nuclear energy derived directly from the materials making up the earth can become our primary source of power, rather than the nuclear energy which comes to us from the sun by radiation and which is subsequently converted to fuel materials through an inefficient and time-consuming process.

2. We can look forward to the time when nuclear energy will be converted to electricity, directly and efficiently, without the intervention of a heat engine.

3. Recent developments in thermoelectric, thermionic, and magnetohydrodynamic converters, now make possible direct electronic conversion of heat to electrical power, with improvement in simplicity, flexibility, possibly efficiency, and elimination of the mechanical generator.

4. In locations where the use of radiant energy from the sun is practical, it will be possible to convert directly from solar to electrical energy at an efficiency of greater than 15 per cent.

5. Batteries are envisioned in which a continuously renewed fuel supply converts chemical energy directly to electrical power with an efficiency of at least 80 per cent.

Each one of these power sources can have its part in our economy, and one or more of them can revolutionize the whole power industry.

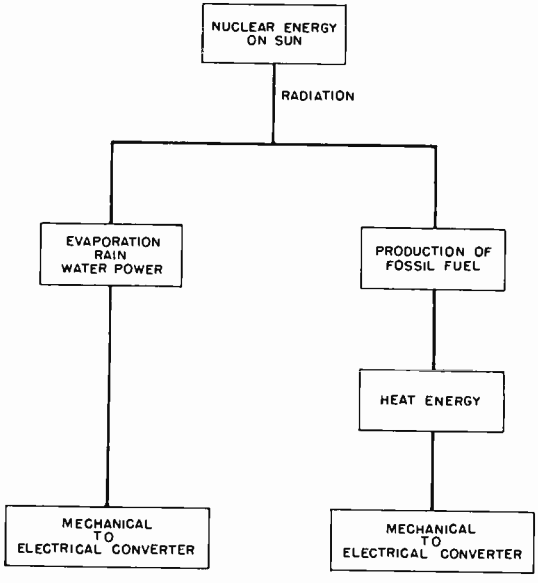


Fig. 1—Conventional methods for converting solar energy into useful power.

In each case electronics is basic to at least one stage in the conversion process.

Figure 1 shows the steps which are generally used today in converting energy from the primary source, the sun, into electrical energy. Parallel paths show alternate methods. Figures 2 and 3 show possible steps in going from the primary source of energy to electrical energy using some of the more recent developments. Any stage in which electronics is basic to the operation is enclosed with a double line. It is seen from Figure 1 that electronics is not required in the operation of conventional generators. In Figure 2, the primary source of power is still nuclear energy from the materials making up the sun;

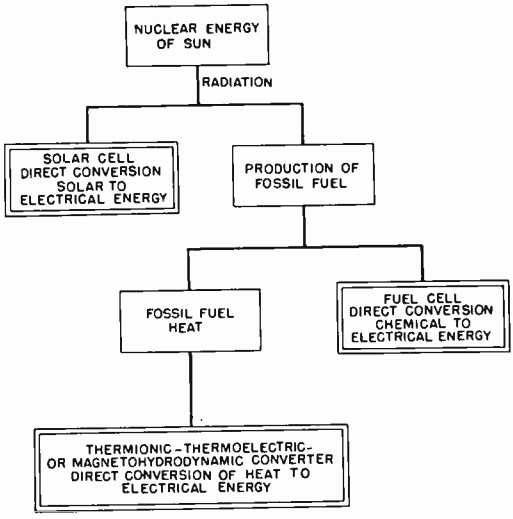


Fig. 2—New methods for harnessing solar energy.

in Figure 3, it is nuclear energy from materials on the earth. Specially to be noted is the fact that double lines enclose at least one box in each case.

Electronics grew because it was a necessity to communication, sound and sight broadcasting, and many phases of military systems. As an industry, it is growing faster than business in general, because (a) activities which were the basis for its initial growth are still

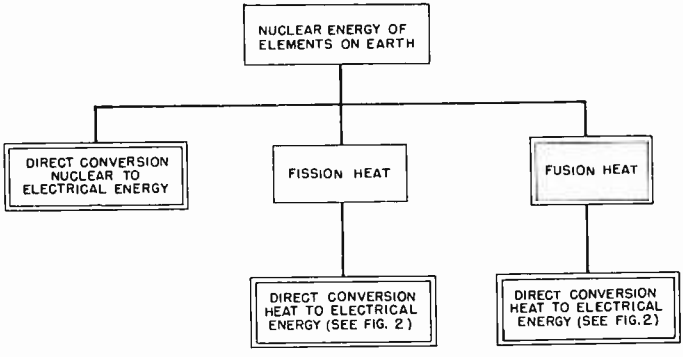


Fig. 3—Methods for utilizing the nuclear energy contained in the earth.

expanding; (b) its applications in the area of industrial controls, automation, and data processing are increasing; and (c) because of the new technology of the electronics of solids. If, in time, the industry we know reaches a plateau, the potential revolution in power conversion may well give rise to another period of extraordinary growth.

RCA Laboratories
Princeton, N. J.

IRVING WOLFF
Vice President, Research, Retired

PLASMA SYNTHESIS AND ITS APPLICATION TO THERMIONIC POWER CONVERSION*†

BY

KARL G. HERNQVIST

RCA Laboratories,
Princeton, N. J.

Summary—The work described in this paper pertains to thermionic energy converters (T.E.C.) in which the electron space charge is neutralized by positive-ion injection into the interelectrode space. A novel method of representing the potential energy diagram for the plasma is described; this method facilitates the understanding of the interaction between plasmas and solids. Based on this model, a detailed plasma energy balance for different types of cathode materials is given.

Results of a theoretical analysis of the space-charge problem in a cesium plasma type T.E.C. are described. The cases studied were for the two simplest space-charge states corresponding to a single potential maximum or a single potential minimum with an arbitrary applied potential between the electrodes. The different space-charge states that are possible when ions and electrons are injected into the interelectrode space are discussed.

Experimental studies of plasma synthesis are described where ions and electrons originate at different sources. The experimental tube operating as an energy converter had a power output of about 0.55 w/cm² at a cathode temperature of 1100°C.

INTRODUCTION

ONE OF THE major problems associated with the thermionic energy converter (T.E.C.) is that of overcoming the effect of interelectrode space charge in hindering the flow of electrons from the cathode to the anode. This problem of space-charge cancellation is not unique to the thermionic energy converter, but is common to high-current electron tube devices for rectification and switching purposes. An excellent way of overcoming space-charge effects is to introduce positive ions in a quantity sufficient to neutralize the electronic space charge. The interelectrode space then becomes occupied by electron–ion pairs, that is a plasma.

In conventional gas-discharge devices such as the thyratron, the plasma is generated by electron impact on neutral gas atoms. In general, this method of plasma generation requires electrical energy,

* This research was sponsored by Advanced Research Project Agency, Department of Defense.

† Manuscript received 18 October 1960.

and is not very efficient because energy is lost in a large number of collisions which do not lead to ionization. Although gas discharges have been used to demonstrate the thermionic converter principle,[1,2] the conversion efficiency was low because of the large fraction of the precious electrical output used for ionization. A much more efficient method of plasma generation, and one that is suitable for the thermionic energy converter, is that of plasma synthesis. Here, ions and electrons are generated separately and brought together to form a plasma, in contrast to the method of direct electron–ion pair generation.

In the type of plasma synthesis discussed in this paper, the electrons are thermionically emitted from the cathode. The ions are generated by resonance ionization of cesium vapor at a hot, high-work-function metal surface. Three cases of plasma synthesis suitable for different cathodes (and heat source for temperatures) are discussed. In one case, both ions and electrons are generated at a high-work-function surface. In another case, a multi-work-function surface is considered where electrons and ions originate at different points on the cathode surface. Thirdly, the case where electrons and ions originate at separate sources is considered. Experimental results for this last method of plasma synthesis are described.

In principle, it is possible by these methods of plasma synthesis to assure charge neutrality and thus zero electric field throughout the interelectrode space. For the case of interest for thermionic converters, the gas pressure is low and thus particle collisions occur mainly with the electrodes. In this case it can be shown that the zero-field solution is, in general, unstable. Potential distributions having a single potential minimum or a single potential maximum can exist in steady state. Different space-charge states characterized by specific potential distributions may exist in steady state depending on the ratio of electron and ion injection rates. Results obtained from an extensive space-charge analysis are described in this paper.

Detailed Plasma Energy Balance for the Thermionic Energy Converter

The principle of operation of a T.E.C. where cesium ions for space-charge neutralization are generated at the cathode surface have been

[1] L. Malter, E. O. Johnson, and W. M. Webster, "Studies of Externally Heated Hot Cathode Arcs, Part I — Modes of the Discharge," *RCA Review*, Vol. 12, p. 415, September, 1951.

[2] G. Medicus and G. Wehner, "Noble Gas Hot Cathode Diodes with Negative Arc Drop," *Jour. Appl. Phys.*, Vol. 22, p. 1389, November, 1951.

described in the literature.[3-6] For optimum performance, the ratio of ion-to-electron emission from the cathode should be such as to maintain charge neutrality (i.e., an equal number of ions and electrons per unit volume) in the interelectrode space. This requires that the ratio of ion-to-electron current density be approximately equal to the square-root of the electron-to-ion mass ratio, about 1/500 for cesium. Secondly, the cathode-to-anode spacing of the T.E.C. should be less than the mean free path for electrons in cesium vapor to minimize the internal impedance.[3] As discussed in detail in the section on space-charge stability, the desired zero-field distribution for the potential in the interelectrode space is not stable. However, potential distributions having a single potential minimum or a single potential maximum can exist in steady state. The plasma-like medium occupying the interelectrode space is then in the strictest sense of the word not a true plasma, since electric fields are likely to be present throughout most of the interelectrode space. However, in discussing the interaction with the boundaries, a very good approximation can be made by assuming that the interelectrode space is occupied by a plasma. Furthermore, in discussing such boundary interaction problems, it is convenient to consider the plasma as a medium having a bandgap energy equal to the ionization potential of the gas (3.89 volts for cesium). Thus, in analogy with a semiconductor, we will be concerned with the empty electron levels represented by the ions. This represents a deviation from conventional plasma concepts and proves to be extremely helpful in considering interaction between plasmas and solids. The energy represented by the bandgap would be lost from the system in the form of radiation if electron–ion recombination took place; therefore, this energy must be accounted for in all cases of plasma generation.

Figure 1 shows the potential-energy diagram for electrons in a T.E.C. Complete space-charge neutralization is assumed in the interelectrode space. Since the diagram shown in Figure 1 represents the potential energy as seen by the electrons, ions used for space-charge neutralization are represented by empty energy levels for electrons, as shown by the dotted line. If, for example, cesium ions are generated

[3] N. D. Morgulis and D. M. Marchuck, "Conversion of Thermal Energy Into Electrical Energy by Means of Thermionic Emission," *Ukranian Phys. J.*, Vol. 2, p. 359, 1957.

[4] K. G. Hernqvist, M. Kanefsky, and F. H. Norman, "Thermionic Energy Converter," *RCA Review*, Vol. 19, p. 244, June, 1958.

[5] G. M. Grover, D. J. Roehling, E. W. Salmi, and R. W. Pidd, "Properties of a Thermoelectric Cell," *Jour. Appl. Phys.*, Vol. 29, p. 1611, November, 1958.

[6] V. C. Wilson, "Conversion of Heat to Electricity by Termionic Emission," *Jour. Appl. Phys.*, Vol. 30, p. 475, April, 1959.

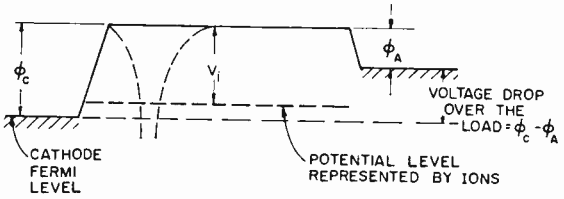


Fig. 1—Potential-energy diagram for electrons in a thermionic energy converter.

at the high-work-function cathode due to resonance ionization, this ionization or "bandgap" energy V_i originates from the heat delivered to the cathode, which raises the potential energy of the emitted electrons. Part of V_i, namely ϕ_A, is lost for each electron–ion pair when the electron is collected at the anode. The rest of the energy, $V_i - \phi_A$, is lost to the anode in an Auger de-ionization process for the ion. Thus it is seen that in the case of cesium ions formed at a high-work-function cathode surface, it is convenient to represent the ions by their empty electron levels originating on an energy scale at a point approximately $(\phi_c - V_i)$ away from the cathode Fermi level.

To improve the efficiency of the thermionic energy converter, it is desirable to lower the cathode work function from the 3.9 volts required for ionizing cesium at the cathode.[4] It would be desirable, therefore, to have a patchy multi-work-function surface where areas of high and low work functions are intimately intermixed. The high-work-function areas ($\geqq$ 3.9-volt) would serve for ion generation, and the low-work-function areas would yield improved electron emission.

The situation prevailing at a multi-work-function electron–ion emitter is best understood if reference is made to the potential-energy diagram shown in Figure 2. Here the potential distribution at two adjacent cathode areas of different work function have been superimposed. The cathode Fermi levels for these two areas are, of course, identical. Let the low-work-function areas have a work function ϕ_L and the high-work-function areas ϕ_H. As discussed above, the plasma is represented by the electron level and the ion level (dashed line in

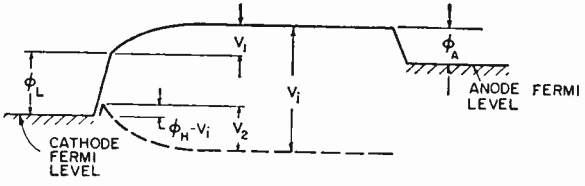


Fig. 2—Potential-energy diagram for electrons assuming a multi-work-function emitter.

Figure 2). The plasma levels will adjust themselves with respect to the cathode Fermi level, depending on the vapor pressure, in such a way as to yield a retarding field for both electrons and ions emitted from the cathode. This will yield an effective work function for the cathode equal to $\phi_L + V_1$. The ions generated due to resonance ionization at the high-work-function areas will appear outside the cathode at a potential level $\phi_H - V_i$ above the cathode Fermi level. The ions will enter the plasma against the retarding potential V_2. In this case, therefore, the energy V_i necessary for the creation of an electron–ion pair originates from the heat delivered to the cathode ($\phi_L + V_1$ for the electron and the remainder for the ions).

The electron current density may be written[4]

$$j_e = 120 T_c^2 \exp\left\{\frac{-e(\phi_L + V_1)}{kT_c}\right\}, \tag{1}$$

and the ion current density is

$$j_p = e\frac{N_g \bar{c}}{4} \exp\left\{\frac{-eV_2}{kT_c}\right\}, \tag{2}$$

where N_g is the vapor density, $\bar{c}$ is the average Maxwellian velocity of the vapor, e is the electronic charge, and k is Boltzmann's constant. For space-charge neutralization

$$\frac{j_p}{j_e} = \sqrt{\frac{m_e}{m_p}}, \tag{3}$$

where m_e = electron mass and m_p = ion mass. Furthermore, from Figure 2,

$$V_i = V_1 + \phi_L + V_2 - (\phi_H - V_i),$$

or

$$V_1 + V_2 = \phi_H - \phi_L. \tag{4}$$

The vapor density is related to the pressure p (in mm of Hg) by

$$N_g = 3.56 \times 10^{16}\, p \frac{273}{T_g}, \tag{5}$$

where T_g is the vapor temperature in °K.

Furthermore,

$$\bar{c} = 1.455 \times 10^4 \sqrt{\frac{1836\, T_g m_e}{m_p}}. \tag{6}$$

It is convenient to express Equations (1-6) in terms of the effective electron work function:

$$\phi_L + V_1 = \frac{1}{2}\left[\phi_H + \frac{kT_c}{e} \ln\left(5 \times 10^{-4} \frac{T_c^2 T_g^{\frac{1}{2}}}{p}\right)\right]. \tag{7}$$

Thus it is seen that a reduction in effective work function and a corresponding increase in the emitted current may be achieved by raising the vapor pressure, p. A practical upper limit on p exists due to the increase in the internal impedance of the converter. The requirements for a multi-work-function emitter may be summarized as follows:

(1) It must possess areas of low and high work functions intimately intermixed to minimize local space-charge effects.

(2) The high work function ϕ_H must be 3.9 volts or higher.

(3) The emitter must have a long life at the operating temperature.

It appears that a variety of electron emitters presently used may fulfill some of these requirements. For instance the L-cathode and thoriated tungsten would satisfy all except the life requirements. Carbides such as uranium[7] and thorium[8] carbide offer possibilities of satisfying all the requirements.

For heat source temperatures in the range 1000 to 1300°C, cathode work functions of the order of 2 volts are called for to obtain reasonably high efficiencies for the T.E.C. In this case, the use of a multi-work-function surface would be impractical due to the high contact fields existing between areas of low and high work function. One way to overcome the effects of these contact fields would be to separate

---

[7] R. W. Pidd, G. M. Grover, D. J. Roehling, E. W. Salmi, J. D. Farr, N. H. Krikorian, and W. G. Witteman, "Characteristics of UC, ZrC, and (ZrC) (UC) as Thermionic Emitters," *Jour. Appl. Phys.*, Vol. 30, p. 1575, October, 1959.

[8] R. Fox and W. Gust, "Thermionic Emission Characteristics of a Thorium Carbide Heat Converter," *American Physical Society Bulletin*, Vol. 4, Series II, p. 322, June 18, 1959.

out the different areas so that a bias voltage can be applied which overcomes the effect of the contact fields. Of course, in practice this implies introducing a third electrode whose sole purpose is that of generating ions in the interelectrode space. This third electrode then must be in thermal but not electrical contact with the cathode.

For the case when ions are generated due to resonance ionization at this third high-work-function electrode, the corresponding potential-energy diagram for electrons is shown in Figure 3. The energy to create an electron–ion pair for the plasma V_i is here supplied partly by the heat delivered to emit electrons from the cathode (ϕ_c) and partly ($V_i - \phi_c$) by the bias voltage. The correct bias voltage called for is $\phi_i - \phi_c$, where ϕ_i is the work function of the ion emitter.

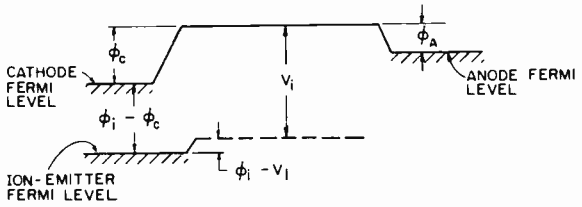


Fig. 3—Potential-energy diagram for a three-electrode thermionic energy converter.

The power delivered by the bias supply is in the form of electricity, of course, and must be taken from the electrical output of the T.E.C. To preserve converter efficiency, this bias power must be small compared to the converter output. For an ideal converter, the lifetime of an ion is determined by its transit time in going through the interelectrode space. The transit time of an ion is then approximately equal to $\sqrt{m_p/m_e}$ times the transit-time of an electron. Thus for cesium, each ion allows the passage of about 500 electrons through the interelectrode space. As an example, let the output voltage of the converter be 0.5 volt, $\phi_c = 2$ volts, and $\phi_i = 4$ volts. In this case the ratio of bias power to converter output power would be 0.016, which is a sufficiently small quantity.

Space-Charge Stability of the Thermionic Energy Converter

Consider a plane-parallel converter diode where both ions and electrons are emitted with Maxwellian velocity distribution from one electrode and are collected by the other. If charged particles are injected at such a rate as to give exact charge neutrality everywhere in the interelectrode region, a zero-field solution is, in general, un-

stable. A slight excess of electrons will depress the potential profile, thus slowing down the electrons and accelerating the ions. The result is an enhanced negative space charge. The cumulative action continues until it is limited by the competing effect of electron reflection from the potential minimum. The system may settle into a steady state characterized by a potential distribution with a potential minimum of finite depth. A similar argument holds if the positive space-charge density is initially in slight excess; the steady state is then characterized by a potential distribution with a maximum of finite height.

A comprehensive space-charge analysis has been carried out for the plasma diode.[10] The analysis yields relations among the operating parameters for the diode. These operating parameters are cathode-to-anode spacing, electron current density, cathode temperature, and the ratio β of the charge density of the injected ion current and the injected electron current. In summary, the result of this analysis shows that there are two different regimes of operating parameters. For one regime steady-state solutions exist, characterized by having a space-charge state with a potential maximum for $\beta > 1$ and a potential minimum for an excess of electrons $\beta < 1$. In the second regime no steady-state solution exists. In general, this second range of operating parameters encompasses those for which a-c oscillations have been experimentally observed in the cesium plasma diode.[8-10]

In practice, the three-electrode T.E.C. may consist of an electron emitter and an electron collector facing each other. The third electrode, the ion emitter, may be placed between or in close proximity to these electrodes. The charge flow then presents a two- or perhaps three-dimensional space-charge problem. Since this becomes impractical to handle mathematically, a simplified model must be chosen for analytical treatment. A model which should provide results which are at least in qualitative agreement with experiment is the plane parallel one, where electrons are emitted from one electrode and ions from the other. This case is illustrated in Figure 4. As discussed above, the zero-field solution is unstable. A single potential maximum (positive state) or a single potential minimum (negative state) is a possible solution to the space-charge problem. In addition, a potential distribution having a potential minimum in front of the electron emitter and a potential maximum in front of the ion emitter (Z-state)

---

[9] H. L. Garvin, W. B. Teutsch, and R. W. Pidd, "Generation of Alternating Current in the Cesium Cell," *Jour. Appl. Phys.*, Vol. 31, p. 1508, August, 1960.

[10] F. M. Johnson, "Direct Conversion of Heat to Electromagnetic Energy," *RCA Review*, Vol. 22, p. 21, March, 1961.

is a stable configuration in this case. Since, in general, the potential tends to go negative outside the electron emitter (a potential minimum will form very close to the cathode) and positive outside the ion emitter, it appears that the Z-state is the one that would be naturally established when the particles are injected. This state, however, is associated with a very deep potential minimum close to the cathode which prevents the electrons from reaching the collector (the ion emitter in this simplified model). One way to pregenerate the right kind of particle distribution in the interelectrode space is to initiate a hot cathode arc discharge. Such a discharge generates a plasma in the interelectrode space due to impact ionization of the vapor. Such a plasma would satisfy the right starting conditions to avoid the build-up of a space-charge distribution corresponding to the Z-state.

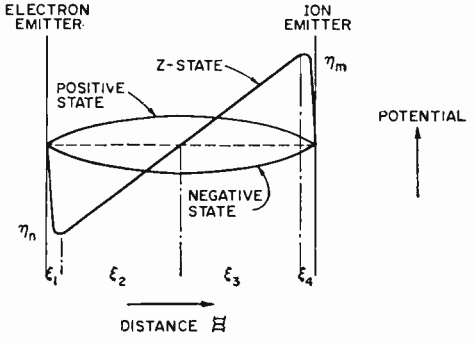


Fig. 4—Possible potential distributions for diode having electron emitter and ion emitter facing each other.

Experimental Studies of Plasma Synthesis

To test the principle of plasma synthesis in which ions and electrons are generated at separate electrodes, the tube shown in Figure 5 was built. The barium-impregnated cathode has a flat surface ¼ inch in diameter. The anode is a nickel disc, about 0.4 inch in diameter. The cathode-to-anode distance is about 0.1 inch. The ion emitter is a tungsten ribbon, 0.040 inch wide and 0.001 inch thick, suspended between the cathode and the anode, and it can be heated separately.

The test circuit is shown in Figure 6. The cathode is heated by a d-c supply. The ion emitter is heated by half-wave rectified 60-cycle current.[3] Performance measurements are made during the "off" part of the cycle when there is no voltage gradient along the ion emitter.

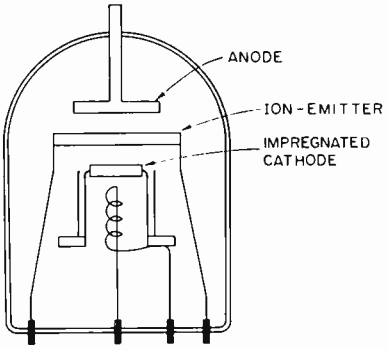


Fig. 5—Experimental three-electrode thermionic energy converter.

The variable load is connected between the cathode and the anode, and provisions are made for measuring the load current and voltage.

In order to be able to initiate a discharge in the converter, a short (~ 1 microsecond) pulse is applied repeatedly to the ion emitter from a pulse generator. The pulse generator is synchronized with the heating voltage of the ion emitter and the pulse occurs at the beginning of the "off" part of the heating cycle. A stiff bias for the ion emitter is provided by the variable d-c supply and rectifier. The rectifier avoids shorting of the pulse applied to the ion emitter.

Figure 7 shows an actual oscillogram representing electron current to the collector as a function of time. It is seen that the application of the discharge pulse takes the system out of the Z-state and initiates a space-charge distribution which is more like the positive or negative state and thus allows the passage of a much larger current. It was further found that current stability was greatly improved by introducing a longitudinal magnetic field of the order of 50-100 gausses. The data reported here were taken with this magnetic field applied.

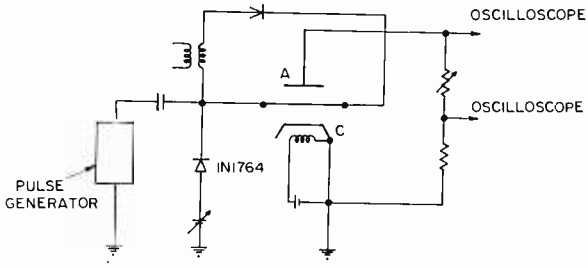


Fig. 6—Circuit diagram for thermionic energy converter showing connection of pulse generator for initiation of discharge.

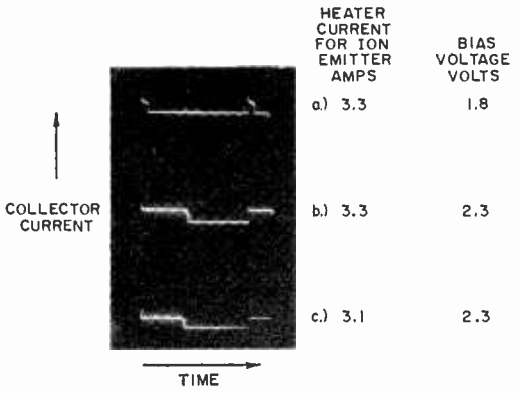


Fig. 7—Oscillograms showing collector current as a function of time: (a) inadequate bias voltage (Z-state has too short a lifetime); (b) proper bias and ion-emitter temperature; (c) too low ion-emitter temperature.

To understand the full significance of the current–time diagram shown in Figure 7, the time sequence of ion-emitter heater voltage, bias voltage, and pulse voltage as well as collector current are sketched in Figure 8. At time T_1 the ion-emitter heater voltage is on, bias voltage is about 2 volts and the collector current is low (Z-state). At time T_2 the heater voltage becomes zero, i.e., no voltage gradient exists

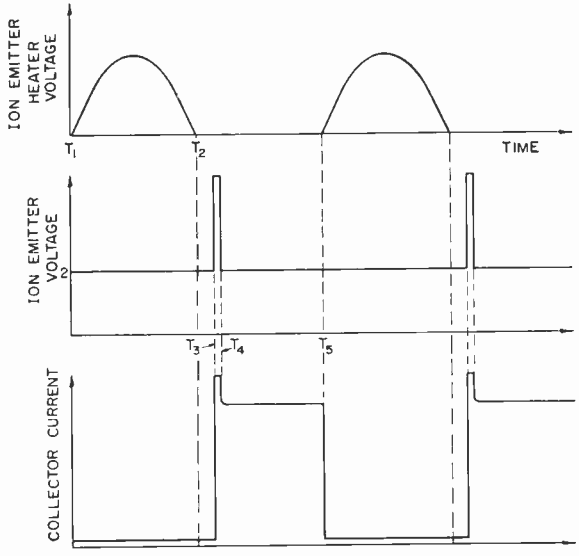


Fig. 8—Heater voltage and bias voltage for ion emitter as well as collector current as a function of time.

along the ion-emitter filament. At time T_3 the pulse voltage is applied and a discharge is initiated in the converter. At time T_4 the pulse voltage is zero and the discharge is extinguished. At this time, since the Z-state is not re-established, a large current passes through the interelectrode space. At time T_5 the ion-emitter heater voltage is again applied and a voltage gradient occurs along the ion emitter. Consequently, the proper bias conditions are no longer satisfied for some part of the ion emitter. As a result the system is forced into the Z-state and the collector current becomes small.

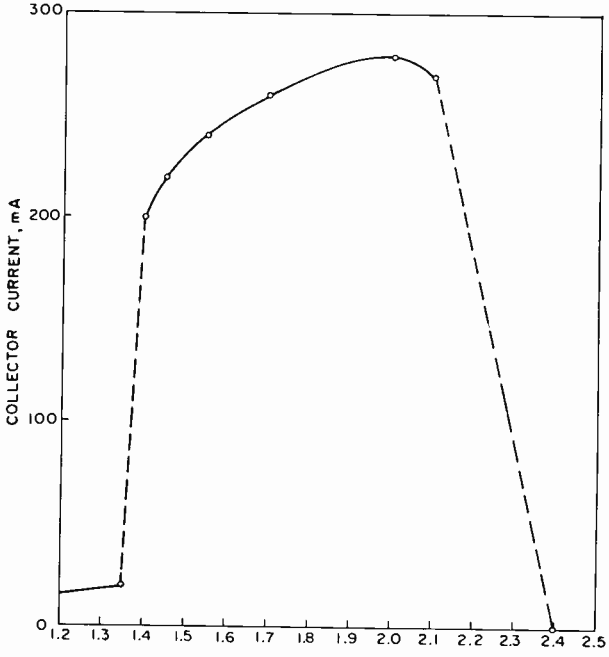


Fig. 9—Collector current versus bias voltage.

Figure 9 shows the effect of bias voltage on collector current. Too low bias voltage results in a retarding field and an insufficient supply of ions. If the bias voltage is too high, ions are accelerated and thus pick up speed before entering the plasma. Moving faster, their ability to neutralize the electronic space charge is diminished. As indicated earlier, the proper bias voltage should be equal to $\phi_i - \phi_c$. For tungsten, $\phi_i = 4.5$ volts, and for an impregnated cathode, $\phi_c = 2.2$ volts. Thus $\phi_i - \phi_c = 2.3$, which is in good agreement with the experiment.

Figure 10 shows the effect of ion-emitter temperature, T_i, on the collector current. In general, the ion current increases with T_i. Thus at a certain temperature the ion supply is sufficient to maintain a positive or negative space-charge state. Further increase of T_i does not appreciably improve the performance.

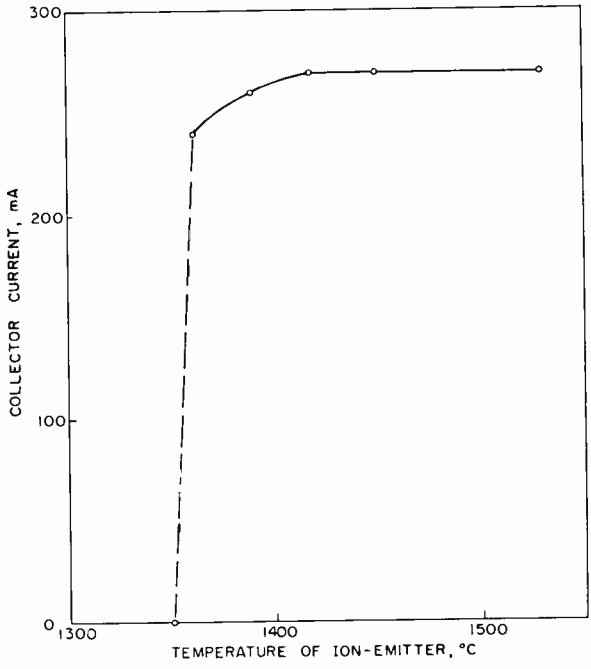


Fig. 10—Collector current versus ion-emitter temperature.

Figure 11 shows the volt–ampere characteristic as well as the corresponding relation between power output and output voltage. The bias power was about 10 per cent of the maximum power output. Although the tube was built to test the principle of plasma synthesis and no attempts were made to optimize the geometry, it is seen that the performance of the device as an energy converter is quite good. Its performance is equal or perhaps slightly better than the best results reported for the close-spaced diode in this temperature range.[11]

CONCLUSIONS

A new method of overcoming space-charge effects in a thermionic

[11] G. N. Hatsopoulos and J. Kaye, "Analysis and Experimental Results of a Diode Configuration of a Novel Thermoelectron Engine," *Proc. I.R.E.*, Vol. 46, p. 1574, September, 1958.

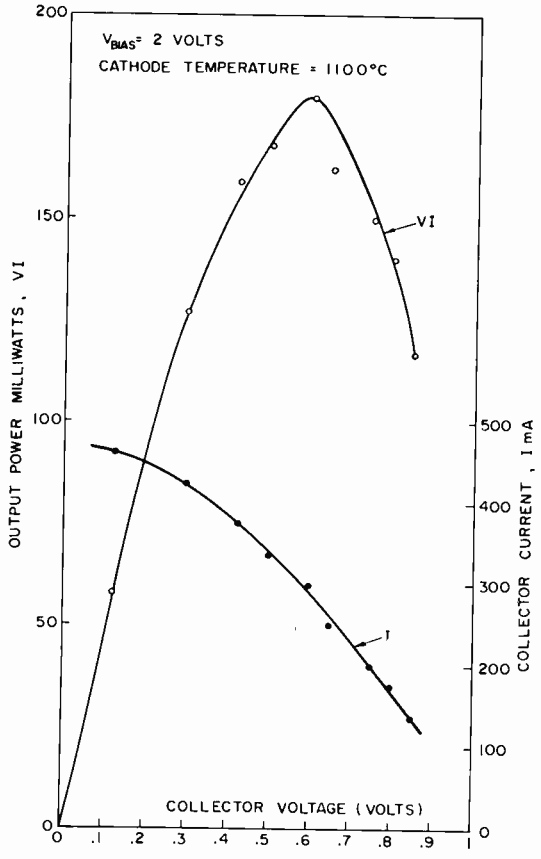


Fig. 11—Collector current and power output versus voltage.

energy converter has been demonstrated. Here, a third electrode has been introduced whose sole purpose is that of generating ions for space-charge neutralization. This plasma triode does not require any critical spacings and is suitable for the heat source temperature range from 1000 to 1500°C. This temperature range has heretofore been reserved for the close-spaced vacuum diode converter which is difficult to fabricate because of the critical tolerances. Thus, although much work still remains in geometry and materials optimization for the plasma triode, no fundamental obstacle now exists in making an efficient, rugged thermionic converter for this temperature range.

DIRECT CONVERSION OF HEAT TO ELECTROMAGNETIC ENERGY*†

By

Fred M. Johnson

RCA Laboratories
Princeton, N. J.

Summary—The conversion of heat into electromagnetic energy is achieved by utilizing the intrinsically unstable space-charge properties of a thermionic cesium plasma diode. Experimental studies of this phenomenon are described. A physical model for the observed relaxation oscillations is proposed which is in qualitative agreement with experiments.

Introduction

A novel method by means of which it has been possible to convert heat directly into electromagnetic energy is based upon operating a thermionic plasma diode in a particular mode of operation. The purpose of this paper is to describe this particular mode of operation and to present some of the experimental results.

Under ordinary conditions, the cesium-vapor-filled thermionic energy converter is a device which converts heat directly into d-c electrical energy.[1] However, during the study of the performance of these converters, it was discovered that under certain operating conditions the output current is self-modulated at a frequency which lies in the 0.5-2 megacycle range.[2-5] The amplitude of the current modulations could approach close to 100 per cent of the direct current. In

* This research was sponsored by the Advanced Research Project Agency, Department of Defense.

† Manuscript received October 18, 1960.

[1] Karl G. Hernqvist, "Plasma Synthesis and Its Application to Thermionic Power Conversion," *RCA Review*, Vol. 22, p. ——, March, 1961, and references cited therein.

[2] Fred M. Johnson, "Relaxation Oscillations in a Cesium Plasma Diode," *Report on 20th Annual Conference on Physical Electronics*, p. 88, M.I.T., Cambridge, Mass., March, 1960.

[3] H. L. Garvin, W. B. Teutsch, and R. W. Pidd, "Generation of Alternating Current in the Cesium Cell," *Jour. Appl. Phys.*, Vol. 31, p. 1508, August, 1960.

[4] R. Fox and W. Gust, "Oscillations and Current Characteristics of the Tantalum Plasma Diode," *Bull. Amer. Phy. Soc.*, Series II, Vol. 5, p. 80, January 27, 1960.

[5] M. Gottlieb and R. J. Zollweg, "Measurements of Oscillation Frequency and Saturation Emission Currents in Cesium Diodes," *Bull. Amer. Phy. Soc.*, Series II, Vol. 5, p. 383, June 15, 1960.

order to understand this relaxation oscillation phenomenon, an experimental and theoretical investigation was undertaken. Some of the results have already been presented.[2,6] Other research[7] had shown, by experimentation, that current oscillations occurred under certain conditions in a plasma extracted from an arc discharge. The conditions for oscillations were such that ion reflection took place in the plasma. In fact, it seems to be generally the case that whenever a mechanism is provided for changing the electron-to-ion space-charge ratio, the possibility of attaining an unstable space-charge configuration exists, which in turn would result in current oscillations.

Theoretical Analysis

In order to understand the a-c case, an analysis[6] was first made of the steady-state mode of operation of the thermionic energy converter. Some of the results of this analysis are outlined here. It was shown that the analysis yields two operating regimes, one where a steady-state solution to the space-charge problem exists, and one where no such solution is allowed. The operating conditions for this second mode of operation are consistent with those for which a-c operation is experimentally observed. The equations describing the space-charge distributions indicate that for the oscillatory mode of operation, quasi-steady-state solutions to the space-charge problem may exist for times short compared with the ion transit time. Based on these studies, a model for the a-c operation can be deduced. The analysis was undertaken for the case of a diode geometry where both ions and electrons are emitted with Maxwellian velocity distributions from one electrode and collected at the opposite electrode. In addition, the analysis was generalized for the case of an arbitrary potential difference between cathode and anode. It would thus encompass most cases of practical interest in the field of thermionic energy converters. The analysis was made for the two simplest space-charge states corresponding to a single potential maximum or a single potential minimum.

As a result of this analysis a mapping of the parameters governing the two regimes, namely the steady state and the unstable state, have been obtained.[6] Although a rigorous solution of the oscillatory case is exceedingly difficult to obtain, it can be shown that steady-state solutions are approached for times which are short compared to the

[6] Fred M. Johnson and Karl G. Hernqvist, "Relaxation Oscillations in Synthesized Plasmas," paper presented at the 13th Gaseous Electronics Conference, Monterey, California, October 16, 1960. A comprehensive paper on this subject will be published elsewhere.

[7] K. G. Hernqvist, "High-Voltage Electron Extraction from an Arc-Discharge Plasma," *RCA Review*, Vol. 21, p. 170, June, 1960.

transit time of the ions. Thus relaxation oscillations occur where the system switches between such quasi-steady-state configurations. The switching time, and hence the frequency, is determined by the transit time of the slowest particles (the ions).

The physical basis of the proposed model is as follows. Consider Figure 1. Both electrons and ions are emitted from the cathode. The three cases considered differ in the polarity of the anode surface potential with respect to the cathode. The operation of the d-c energy

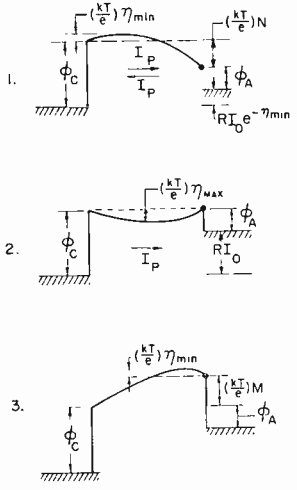


Fig. 1—Potential energy diagram for electrons for various anode potentials and corresponding plasma states. (The ion flow is designated by I_P.)

converter is indicated in No. 2 of Figure 1. Here the anode surface potential is shown to be equal to the cathode surface potential. When heat is applied to the cathode, the electrons have their potential energy raised with respect to the Fermi level. Those electrons that reach the top of the potential barrier may leave the cathode surface and will reach the anode provided interelectrode space-charge effects have been overcome. The space-charge problem is solved by introducing positive charges by means of resonance ionization of cesium atoms at the cathode surface. The contact difference of potential, $(\phi_c - \phi_A)$, is equal to the output voltage I_0R which appears across the load, R, where I_0 is the saturated electron current. If an excess of positive charge is present, the potential distribution has a maximum value as shown in No. 2. Note that for this type of potential distribution, all the positive ions that enter also leave the interelectrode space since they are accelerated towards either electrode.

This briefly describes the operation of the d-c converter. To understand the a-c converter consider the following.

If the rate of ion generation should be diminished, the interelectrode charge density will change from a predominantly positive one to a relatively negative one. When this occurs, the assumed potential distribution cannot be maintained, and a potential distribution as shown in No. 1 of Figure 1 will be established. As soon as this potential minimum is established, however, the electron current will be drastically diminished to $I_0 \exp\{-\eta_{min}\}$ where $(kT/e)\eta_{min}$ is the depth of the potential minimum, T is cathode temperature, k Boltzmann's constant, and e is the electronic charge. Consequently, the output

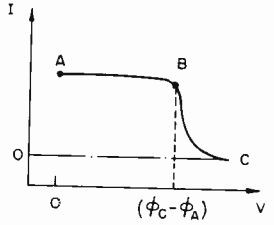


Fig. 2—Typical I–V characteristic of cesium plasma diode.

voltage will also be diminished and the anode surface potential will now be positive by an amount $(kT/e)N$ volts with respect to the cathode surface potential. Note that for this case, a space-charge sheath will be established close to the anode and some of the ions passing from the cathode to the anode will be reflected at this space-charge sheath and re-enter the interelectrode space. The number of ions reflected is a function of the potential difference between the cathode and anode which, in turn, is determined by the current flowing through the external load. Due to these reflected ions, an excess of positive charge will rapidly be established which, in turn, will cause the potential distribution to flip back from the minimum to a maximum value. From then on the cycle repeats. This ion-pumping effect is believed to be responsible for the relatively coherent nature of the observed oscillations.

If on the other hand the anode surface potential is negative with respect to the cathode (as shown in No. 3 of Figure 1), no ion reflection is possible even if the potential were to change from a minimum to a maximum value, since the corresponding potential variation is negligible with respect to $(kT/e)M$.

The I–V characteristic of a typical cesium plasma diode is shown schematically in Figure 2. In the region $A \rightarrow B$ both a-c and d-c operation are possible. It was found experimentally, however, that the

internal relaxation oscillations never occur in the region $B \rightarrow C$, which corresponds to case No. 3 of Figure 1. Thus one of the necessary conditions for oscillations to occur is that the anode surface potential be positive with respect to the cathode surface potential. These experimental observations are consistent with the theoretical model described above.

Experimental Results

The experimental arrangement used for the observation of relaxation oscillations is similar to that previously described by Hernqvist[8] *et al.* (See Figure 3). The cesium-filled diode used in this series of

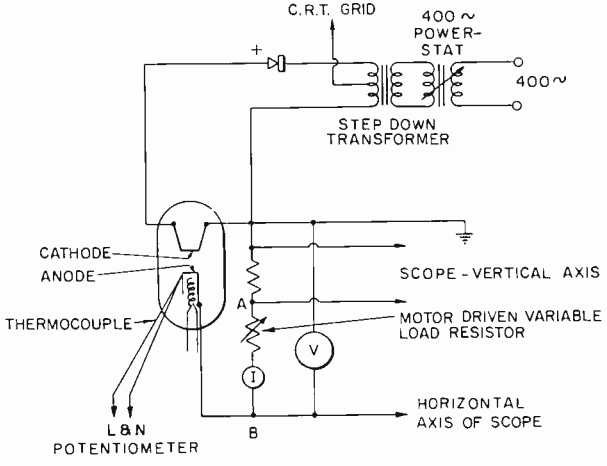


Fig. 3—Diagram of experimental arrangement for measuring I–V and a-c characteristics of cesium diode.

experiments consisted of a tantalum ribbon cathode and a tantalum anode. The temperature of the anode could be varied by means of a separate heater. The temperature was monitored by means of a thermocouple attached to the anode. The filament was heated by either 400- or 60-cycle half-wave rectified a-c, and observations of the diode characteristics were made during the off part of the heating cycle. The volt–ampere characteristics of this device are shown in Figure 4. Note that if the voltage is varied over the range $A \rightarrow B$, oscillations of the output current are observed. The frequency of these oscillations was in the 0.5 to 1 megacycle range. Since the oscillogram for the I–V characteristics was taken at a relatively slow sweep speed, the

[8] K. G. Hernqvist, M. Kanefsky, and F. H. Norman, "Thermionic Energy Converter," *RCA Review*, Vol. 19, p. 244, June, 1958.

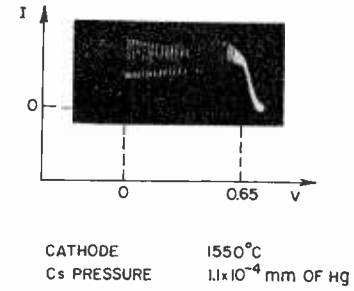


Fig. 4—Typical *I–V* characteristic of cesium diode operated under conditions for which oscillations are seen to appear.

60-cycle modulation of the relaxation oscillations is observed in Figure 4. Oscillograms of the oscillations under various conditions are shown in Figures 5, 6, and 7. The effects of varying load resistance are shown in Figure 5, whereas Figure 6 shows the effects of changing the anode temperature. For a given set of operating conditions, it has always been found possible to quench the oscillations by increasing the anode temperature to about 800°C or beyond. The frequency dependence of these oscillations as a function of cesium pressure for a particular experimental tube is shown in Table I. The effects of an external magnetic field on these relaxation oscillations were also investigated. In general, the magnetic field caused a change in the oscillation wave shapes from a jagged to a more rounded appearance.

Finally, a cesium diode having a thin wire probe in the interelectrode space was constructed. This permitted simultaneous observations of the relaxation oscillations inside the plasma as well as in the external circuit. An oscillogram of such a measurement is shown

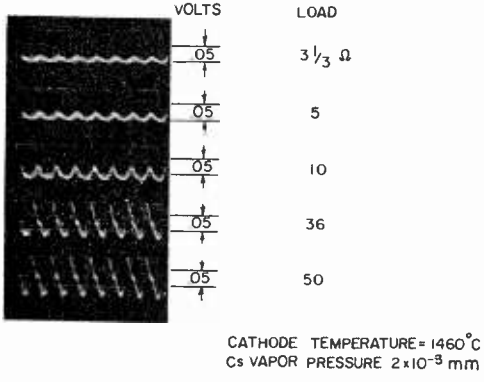


Fig. 5—Relaxation oscillation wave shapes for various load resistances.

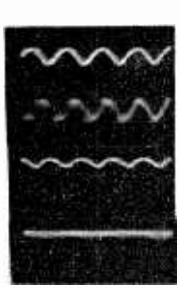


Fig. 6—Relaxation oscillation wave shapes as a function of anode temperature.

in Figure 7. The voltage swing of these oscillations, as measured by the probe, is of the order of $kT/e(\simeq 1/10$ volt), whereas the output voltage across the load may be an order of magnitude greater and is, in fact, a function of load resistance.

Table I

Cs Pressure (mm of Hg)	Frequency of Relaxation Oscillations (kilocycles)
10^{-4}	610
3×10^{-4}	655
2.5×10^{-3}	770
2×10^{-2}	790
6×10^{-2}	780

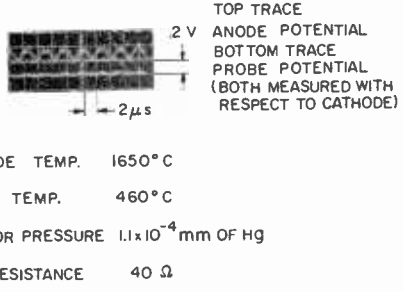


Fig. 7—Simultaneous observation of relaxation oscillations at the anode (top trace) and at the interelectrode probe (bottom trace).

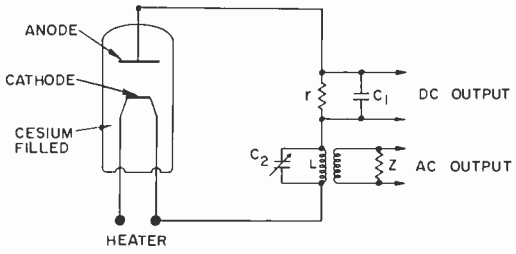


Fig. 8—Circuit diagram for a-c converter.

For practical application of the thermionic energy converter, which would provide both an a-c and a d-c output, a circuit as indicated in Figure 8 was used. Values of C_2 and L were chosen so as to have the tank circuit in resonance with the relaxation oscillations of the plasma diode. The beneficial effect of introducing a tank circuit can best be appreciated by referring to Figure 9. In this case, a plasma diode of cylindrical geometry was investigated.

Conclusions

Direct conversion of heat into electromagnetic energy can be achieved by means of the cesium plasma diode. The frequency is governed by the transit time of the ions, which, for a practical device, limits the frequency range to approximately 100 kc to 5 mc.

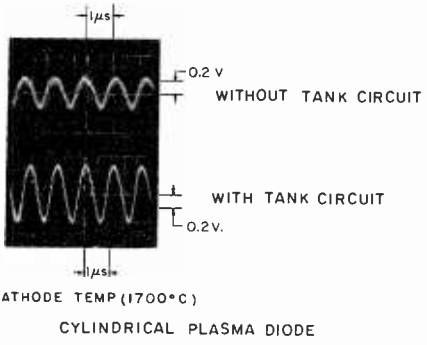


Fig. 9—Oscillogram of output voltages of a-c converter, with and without tank circuit.

LARGE-AREA THIN-FILM PHOTOVOLTAIC CELLS*

By

H. I. Moss

RCA Laboratories
Princeton, N. J.

Summary—*The production of single-crystal material for use in the fabrication of photovoltaic cells of silicon, cadmium sulfide, and other materials involves considerable time and expense. However, polycrystalline photovoltaic cells fabricated from thin films or layers of these materials would be relatively inexpensive, large-area devices. Thin-film cells of sufficiently high efficiency would provide more electrical output per unit weight than a collection of relatively thick single-crystal cells.*

Some progress in this direction has been made with cadmium sulfide. Large-area thin-film photovoltaic cells have been fabricated by vacuum depositing cadmium sulfide onto heated transparent conducting substrates. These films are approximately 2 microns in thickness. A layer of copper was applied to the exposed surface of the cadmium sulfide films to form a barrier layer. Back-wall illumination of these cells produces solar conversion efficiencies of about 1 per cent.

The spectral response and current–voltage characteristics of a typical thin-film cell are given and compared to the same characteristics obtained from a single-crystal cadmium sulfide photovoltaic cell. Also given are plots of short-circuit current and open-circuit voltage as a function of light intensity for the cadmium sulfide film cells.

Introduction

PHOTOVOLTAIC CELLS are self-contained current and voltage generators, which produce a potential difference between their terminals when exposed to light or ionizing radiation. The photovoltaic effect was first discovered by Becquerel[1] in 1839 when light was directed on one of a pair of electrodes immersed in an electrolyte solution. Following Becquerel's initial discovery, a number of "dry" photovoltaic cells were discovered and their properties investigated.

The photovoltaic effect in selenium was discovered by Adams and Day[2] and investigated more fully by Bergmann.[3] Grondahl[4] discovered

* Manuscript received October 18, 1960.

[1] M. E. Becquerel, "On Electric Effects Under the Influence of Solar Radiation," *Compt. Rend.*, Vol. 9, p. 561, 1839.

[2] W. G. Adams and R. E. Day, "The Action of Light on Selenium," *Proc. Roy. Soc.*, Vol. A25, p. 113 (1877).

[3] L. Bergmann, "On a New Selenium Barrier Layer Photocell," *Z. Physik*, Vol. 32, p. 286, 1931.

[4] L. O. Grondahl, "The Copper-Cuprous Oxide Rectifier and Photoelectric Cell," *Rev. Mod. Phys.*, Vol. 5, p. 141, April, 1933.

the effect in cuprous oxide and cells of this material were developed to a point of practical usefulness by Lange[5] and Schottky.[6] Due to the relatively low efficiencies of such cells (about 1 per cent), practical applications were restricted to laboratory measurement devices and photoelectric exposure meters. In 1954 the silicon p-n junction electronvoltaic cell was unveiled;[7,8] this converted radioactive radiation into electrical energy with practical efficiencies. Also announced in 1954 was the silicon p-n junction photovoltaic cell; this had a solar conversion efficiency of 6 per cent[9,10] (presently about 10-14 per cent for production line units and 15 per cent[11] for laboratory units). The advent of this cell made it possible to convert useful amounts of solar energy directly and efficiently into electricity. At present, silicon cells cost about 200 dollars per watt for high-efficiency cells and about 100 dollars per watt for low-efficiency cells.

A further significant development in the photovoltaic field was the cadmium sulfide cell.[12,13] These cells are fabricated from cadmium sulfide single crystals, made n-type by suitable doping. The surfaces are roughened and copper is electroplated onto one face and diffused into the interior at 350-450°C for less than one minute, forming a barrier layer between one and ten microns thick. Electrical contacts are then made to the n region of the crystal and the barrier layer. Back-wall illumination of these cells can produce solar conversion efficiencies of 5-8 per cent. These photovoltaic cells have spectral responses well matched to the solar spectrum, and their properties indicate that they may be useful in the field of solar energy conversion.

---

[5] B. Lange, "New Kind of Photoelectric Cell," *Z. Physik*, Vol. 31, p. 139, 1930.

[6] W. Schottky, "Origin of Photoelectrons in Copper-Oxide Photoelectric Cells," *Z. Physik*, Vol. 31, p. 913, 1930.

[7] P. Rappaport, "The Electron-Voltaic Effect in P-N Junctions Induced by Particle Bombardment," *Phys. Rev.*, Vol. 93, p. 246, January, 1954.

[8] P. Rappaport, J. J. Loferski, and E. G. Linder, "The Electron-Voltaic Effect in Germanium and Silicon P-N Junctions," *RCA Review*, Vol. XVII, p. 100, March, 1956.

[9] D. M. Chapin, C. S. Fuller, and G. L. Pearson, "A New Silicon P-N Junction Photocell for Converting Solar Radiation into Electrical Power," *Jour. Appl. Phys.*, Vol. 25, p. 676, May, 1954.

[10] G. L. Pearson, C. S. Fuller, and D. M. Chapin, "The Bell Solar Battery," *Bell Laboratories Record*, Vol. 32, p. 232, June, 1954.

[11] C. Ruddenberg, "High Efficiency Silicon Solar Cells," *14th Annual Power Sources Conference*, Atlantic City, N. J., May 17-19, 1960.

[12] D. C. Reynolds, G. Leies, L. L. Antes, and R. E. Marburger, "Photovoltaic Effect in Cadmium Sulfide," *Phys. Rev.*, Vol. 96, p. 533, October, 1954.

[13] L. L. Antes, "Progress in Cadmium Sulfide," *Trans. I.R.E. PGCP*, Vol. CP-4, No. 4, p. 129, December, 1957.

The large energy gap of cadmium sulfide, 2.4 e.v., indicates that the photovoltaic efficiency should be quite low. However, the photovoltaic effect is produced with high efficiency in the cadmium sulfide–copper system by light whose energy is less than that of the energy gap.

The mechanism of the photovoltaic effect in cadmium sulfide is not yet fully understood. However, recent work at the RCA Laboratories by Williams and Bube[14] indicates that efficient single-crystal cadmium-sulfide–copper photovoltaic cells can be fabricated without any heat treatment. Experiments provide evidence that the photovoltaic current produced by such cells is due to the photoemission of electrons from the copper into the cadmium sulfide. This work by Williams and Bube and the potential usefulness of the cadmium sulfide photovoltaic cell for solar energy conversion served as the impetus for the experimental work reported here.

THIN-FILM PHOTOVOLTAIC CELLS

The production of single-crystal material for use in the fabrication of photovoltaic cells of silicon, cadmium sulfide, etc., involves considerable time and expense. However, polycrystalline photovoltaic cells fabricated from thin films of these materials would result in relatively cheap, large-area devices divorced from the limitations imposed by single-crystal production and application. Thin-film cells of sufficiently high efficiency would provide more electrical output per unit weight than a collection of relatively thick single-crystal cells. Indeed, it appears that the solar cell array of the future will be in the form of a large-area polycrystalline photovoltaic device fabricated from either sintered or vacuum-evaporated materials.

W. R. Cherry[15] has described the essential properties that a thin film of a polycrystalline semiconductor must possess to be applicable to photovoltaic-cell fabrication. In the case of a p-n junction cell, the thickness of the film must be great enough to absorb an appreciable amount of the photons having an energy greater than the band gap of the film material. Also, the carrier diffusion lengths must equal or exceed the film thickness. In the case of a polycrystalline layer, the minimum grain size must be at least equivalent to the film thickness. This insures that a carrier diffusing towards the junction will not be intercepted by a grain boundary and thus given an opportunity to

---

[14] R. Williams and R. H. Bube, "Photoemission in the Photovoltaic Effect in Cadmium Sulfide Crystals," *Jour. Appl. Phys.*, Vol. 31, p. 968, June, 1960.

[15] W. R. Cherry, "Advanced Photovoltaic Devices," *14th Annual Power Sources Conference*, Atlantic City, N. J., May 17-19, 1960.

recombine. In the case of a metal–semiconductor junction photovoltaic cell, the above requirements for a thin layer may not have to be as stringent as for a p-n junction device.

There have been relatively few reports in the literature concerning the application of polycrystalline semiconductor films to photovoltaic cell fabrication. Carlson[16] attempted to fabricate photovoltaic cells from vacuum-evaporated films of cadmium sulfide in contact with layers of copper or cuprous sulfide. The measured conversion efficiencies of these cells were very small, about 0.1 per cent. The fabrication of cells from sintered layers of cadmium sulfide was reported by Lind, et al.[17] There was an optimum thickness for the cadmium sulfide layers, about 5 microns. As the layers became thicker, less light would reach the active junction. For layers thinner than 5 microns, the copper electrode, which was applied to the surface of the cadmium sulfide, would tend to short through the porous cadmium sulfide layer. The conversion efficiencies of these sintered cells were not greater than a few tenths of one per cent. An interesting type of photovoltaic cell has been described by Goldstein and Pensak.[18] They vacuum deposited films of cadmium telluride from which a high-voltage photovoltaic effect (100 volts/cm) was observed. Experimental evidence indicated that the cadmium telluride films involved an array of p-n junctions arranged in such a way that their voltages were additive. The high resistance of these films limited the conversion efficiencies to very small values. Recent work has shown that layers of powdered silicon can be fabricated into photovoltaic cells which exhibit conversion efficiencies of 1 per cent.[19] Also, cast silicon cells have been made which are polycrystalline with grain sizes of 5×10^{-2} cm and show efficiencies of about 10 per cent.[15] The use of organic semiconductor materials in the form of thin layers for photovoltaic cells has been investigated by Kearns and Calvin.[20] One organic system showing a photovoltaic effect uses magnesium phthalocyanine disks coated with a thin film of air-oxidized tetramethyl *p*-phenylenediamine.

---

[16] A. Carlson, "Research on Semiconductor Films," *WADC Technical Report* 56-52, Jan., 1956.

[17] E. L. Lind, J. J. Loferski, L. Pensak, P. Rappaport, and S. M. Thomsen, "Investigation of Materials for Photovoltaic Solar Energy Converters," 4th Interim Report, 1 Jul 56-31 Oct 56, 30 Nov 56, Contract No. DA-36-039-sc-64643.

[18] B. Goldstein and L. Pensak, "High-Voltage Photovoltaic Effect," *Jour. Appl. Phys.*, Vol. 30, p. 155, February, 1959.

[19] M. B. Prince, "Large Area Silicon Solar Cells," *14th Annual Power Sources Conference*, Atlantic City, N. J., May 17-19, 1960.

[20] D. Kearns and M. Calvin, "Photovoltaic Effect and Photoconductivity in Laminated Organic Systems," *Jour. Chem. Phys.*, Vol. 29, p. 950, October, 1958.

The maximum voltage developed by one of these disks was 200 millivolts. The maximum power output was only 3×10^{-9} milliwatts. The low power output was probably caused by a high cell resistance. The use of organic semiconductor materials for photovoltaic cells will have to wait until materials are available which exhibit greater conductivities than the materials presently available.

The following sections of this paper describe some recent work at the RCA Laboratories designed to investigate the feasibility of producing large-area, thin-film cadmium sulfide photovoltaic cells.*

Preparation of Cadmium Sulfide Cells

Cadmium sulfide layers were deposited onto heated transparent conducting (tin oxide) Pyrex† substrates. The pressure during the evaporation was slightly lower than 10^{-5} mm Hg. These layers were about 1-2 microns thick and were hard and adherent. Microscopic observation revealed the layers to be microcrystalline and free of pin

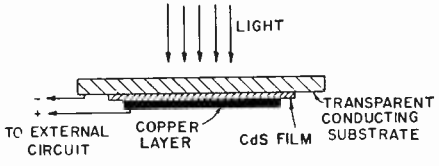


Fig. 1—Geometry of CdS evaporated-layer photovoltaic cell.

holes. Their optical properties appeared to be very similar to cadmium sulfide single crystals and they were transparent to wavelengths greater than 5200 Å. There was no visible evidence of layer deterioration even after standing for several months in the atmosphere. The evaporated layers without additional treatment are relatively conducting owing to the presence of excess cadmium. Indications are that the layers have resistivities of about 100 ohm-cm.

Fabrication of the photovoltaic cell was completed by applying an opaque layer of copper to the exposed surface of the cadmium sulfide layer. The geometry of the finished cell is illustrated in Figure 1. Electrical contacts to the cell were made on the copper layer and on the tin oxide. The transparent tin oxide coating appears to make an ohmic contact to the evaporated cadmium sulfide layer.

---

* An evaporated film cadmium sulfide photovoltaic cell exhibiting an efficiency of 3.5 per cent has been described in a paper by A. E. Middleton, D. A. Gorski and F. A. Shirland presented at the American Rocket Society Space Power Systems Conference, Santa Monica, Calif., Sept. 27-30, 1960.

† Registered trade mark.

Cell Characteristics

Back-wall illumination of these cells with light from an artificial sunlight source produced open-circuit voltages between 0.2 and 0.5 volt (copper becomes positive) and short-circuit current densities between 2.5 and 5 milliamperes per square centimeter. Cells with active areas of 2 cm^2 have been fabricated; however, the highest short-circuit currents were obtained from the smaller area cells. A 75-watt tungsten photospot lamp was used as the light source; the

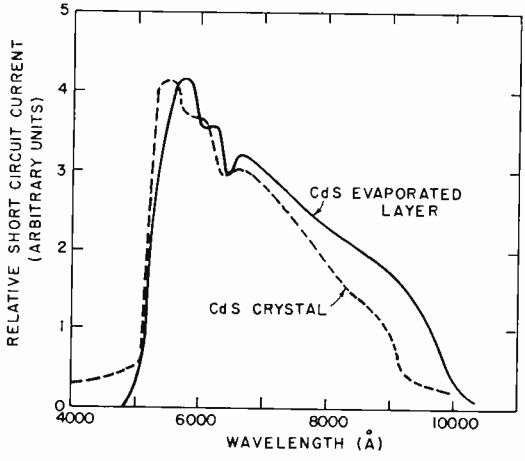


Fig. 2—Spectral response of CdS-Cu photovoltaic cells.

voltage to the lamp was 130 volts. It has been reported that operation of a tungsten photospot lamp at 15 volts over the nominal rating results in a light source more nearly resembling direct sunlight.[21] This, of course, is a compromise, since the efficiency of a photovoltaic cell depends on the spectrum used for illumination, and one cannot readily duplicate the solar spectrum with an artificial light source. For this work the lamp-to-cell distance was adjusted to reproduce the same short-circuit current in an experimental cadmium sulfide cell as produced in bright sunlight, the intensity of which was determined with an Eppley pyrheliometer. This lamp arrangement served as a convenient laboratory method to measure photovoltaic-cell characteristics.

Figure 2 shows the spectral response of a typical cell when illumi-

[21] D. A. Hammond, F. A. Shirland, and R. J. Baughman, "A Cadmium Sulfide Solar Generator," *WADC Technical Rept.* 57-770, December, 1957.

nated with a Bausch and Lomb grating monochromator containing a tungsten light source. The cell response has been normalized to take into consideration differences in the spectral intensity transmitted by the monochromator. A peak is present at about 5700 Å which drops sharply at the absorption edge (5200 Å) and more gradually at longer wavelengths. This curve is similar to curves obtained from single-crystal cadmium sulfide–copper photovoltaic cells whose barrier electrodes were fabricated in the same manner as the evaporated-layer cells. The dashed curve in Figure 2 is the spectral response of such

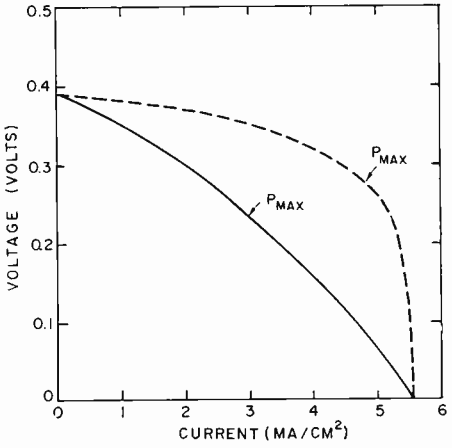


Fig. 3—Current–voltage characteristic of CdS-Cu photovoltaic cells. The solid line is for a CdS evaporated layer, and the dashed line is for a CdS single crystal.

a single-crystal cell, normalized so its maximum coincides with the maximum of the evaporated cadmium sulfide cell. The shift in the maximum of the curve for the evaporated-layer cell to slightly longer wavelengths compared to the maximum of the single-crystal curve is probably due to a decrease in the sharpness of the absorption edge of the evaporated layer. This decrease could arise because of the microcrystalline nature of the evaporated layer.

Figure 3 shows the current–voltage characteristic for a typical cell when illuminated with the artificial light source equivalent to a sunlight radiation density of 86 milliwatts/cm^2. The curve was obtained by varying the load resistance and measuring the voltage and current while the cell was illuminated. The maximum power, P_{max}, that this cell can deliver is represented by the area of the largest

rectangle that can be fitted under the curve.[22] In this case the maximum power is 0.69 milliwatt/cm² when the external load is 77 ohms; the solar conversion efficiency is 0.81 per cent. In general, the more rectangular the current–voltage characteristic, the higher the efficiency. The dashed curve in Figure 3 shows the current–voltage characteristic of a single-crystal cadmium sulfide photovoltaic cell, the barrier of which was fabricated in the same manner as the evaporated layer cell. The maximum power delivered by this single-crystal cell is

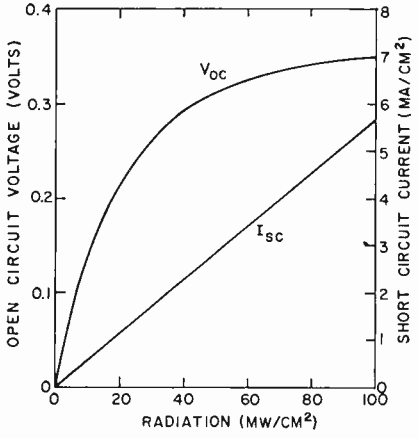


Fig. 4—Variation of current and voltage as a function of light intensity for a CdS evaporated-layer photovoltaic cell.

1.95 milliwatts/cm². This crystal had been grown* from the vapor phase in an atmosphere containing a few millimeters partial pressure of iodine and had a resistivity of about 1 ohm-cm.

In order to compare the two sets of data in Figure 3, the current–voltage curve for the single-crystal cell has been normalized by dividing it by a constant factor such that the short-circuit current coincides with the measured short-circuit current obtained for the evaporated-layer cell. Both cells had comparable open-circuit voltages. The rectangular characteristic of the single-crystal cell is quite obvious while the characteristic of the evaporated-layer cell is almost linear. A number of factors can contribute to produce the poor current–voltage characteristic observed for the evaporated-layer cell. Among

[22] K. Lehovec, "The Photo-Voltaic Effect," *Phys. Rev.*, Vol. 74, p. 463, August 15, 1948.

* Crystals grown by L. A. Barton.

the possible contributing factors are high resistance or rectifying properties at the cadmium-sulfide–tin-oxide contact as well as a high bulk resistance for the cadmium sulfide layer.

Figure 4 shows the variation of the open-circuit voltage, V_{oc}, and of the short-circuit current, I_{sc}, as a function of light intensity. Variations in light intensity were obtained with neutral density wire-mesh filters. The current is directly proportional to the intensity, while the voltage varies as the logarithm of the light intensity at low levels and tends toward a saturation value at high intensities. This is the usual behavior for photovoltaic cells.

CONCLUSION

It has been shown that it is possible to fabricate relatively inexpensive large-area thin-film cadmium sulfide photovoltaic cells by utilizing vacuum evaporation techniques. Solar conversion efficiencies approach 1 per cent. It is believed that suitable doping of the cadmium sulfide evaporated layers to increase conductivity and proper treatment of the surface of the layers prior to applying the copper electrode will result in increased efficiencies, perhaps by factors of five to ten.

ACKNOWLEDGMENTS

The author wishes to thank R. Williams, R. H. Bube, H. L. Pinch, S. M. Thomsen, L. Pensak, and P. Rappaport for many helpful discussions and suggestions.

SPECTRAL RESPONSE OF PHOTOVOLTAIC CELLS*†

BY

J. J. LOFERSKI‡ AND J. J. WYSOCKI#

Summary—*This paper presents a theoretical and experimental study of the spectral response of p-n junction photovoltaic cells.*[1-18] *The current flowing across the junction is calculated as a function of the absorption constant of the radiation for various values of the semiconductor parameters. The results of machine calculations of practically interesting cases are shown in the figures. The use of normalized spectral-response curves to determine the values of minority-carrier diffusion lengths is demonstrated. It is shown that by using certain approximations it should be possible to determine minority-carrier diffusion lengths on both sides of the junction with relative ease. These principles have been applied to experimental GaAs and silicon photovoltaic cells to determine the values of diffusion lengths in the finished junctions.*

INTRODUCTION

IF THE ABSOLUTE VALUE of absorption constant α of a semiconductor is known as a function of photon energy $h\nu$, it is possible to calculate the shape of the spectral response of a photovoltaic cell with a specified geometry, minority-carrier diffusion lengths (L_n and L_p), and surface recombination velocity. Consequently, if the position of the p-n junction is established by some independent means, L_n and L_p can be estimated from the shape of the spectral response curve of the junction. Representative theoretical curves of response versus α for various plausible combinations of semiconductor parameters have been calculated. Experimental photovoltaic spectral-response curves of diffused Si and GaAs p-n junction photocells have been used to verify the theory.

THEORETICAL

In this section the photovoltaic spectral response of a plane p-n junction located at a distance l below the surface of a semiconductor

* This work was supported by the U.S. Army Signal Research and Development Laboratory, Fort Monmouth, New Jersey under contract DA36-039-sc-78184. It is reported in contract reports commencing December 15, 1958.

† Manuscript received 16 January 1961.

‡ Formerly, RCA Laboratories, Princeton, N. J., now at Brown University, Providence, R. I.

RCA Laboratories, Princeton, N. J.

wafer of thickness b (Figure 1) is calculated. Contributions from both the p and the n sides of the junction are included.

The problem consists of determining the fraction of liberated minority carriers that arrive at the junction contributing to the photovoltaic current. It is assumed that the minority carriers move by diffusion, that they have a lifetime in the bulk symbolized by τ, and that the surface of the p-n junction can be characterized by a surface recombination velocity s. The appropriate differential equation for minority holes has the form

$$\alpha N_0(1-R)\,e^{-\alpha x} + D\frac{d^2p}{dx^2} - \frac{p}{\tau} = 0. \tag{1}$$

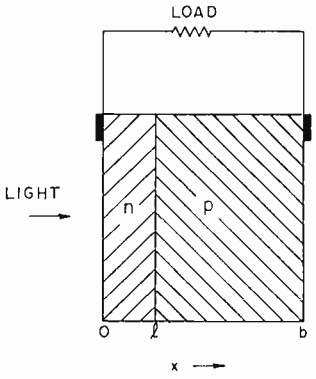


Fig. 1—Diagram of a photovoltaic cell.

This equation represents the net change in minority-hole concentration per unit volume in an illuminated semiconductor. The first term is the number of minority holes introduced into the volume at a distance x below the surface by light (or any kind of ionizing radiation) whose absorption in the semiconductor follows Lambert's Law, i.e.,

$$N(x) = N_0 e^{-\alpha x}. \tag{2}$$

Here α is the absorption constant, which is a function of λ, the wavelength of the incident light, and N_0 is the incident photon current density at $x = 0$. The fraction of light reflected at the surface is R. The second term in Equation (1) is the net increase in hole concentration by diffusion from the surrounding space. q is the electronic

charge and D is the diffusion constant. The third term is the number of minority carriers lost by recombination. Note that $p = p_0 + \Delta p$ where Δp is the minority hole concentration in excess of the equilibrium value, p_0. The general solution of Equation (1) is

$$p = Ae^{\nu x} + Be^{-\nu x} + Ce^{-\alpha x} \tag{3}$$

for $\alpha \neq \nu$. The constant C is

$$C = \frac{\alpha N_0 (1 - R)}{D(\nu^2 - \alpha^2)} . \tag{4}$$

The coefficients A and B are determined by introducing the following conditions which are appropriate to the p-n junction shown in Figure 1:

$$\left.\frac{dp}{dx}\right|_{x=0} = \frac{s}{D_p} p(0), \tag{5}$$

$$p(l) = 0, \tag{6}$$

i.e., the junction is a sink for minority carriers. The diffusion current flowing across the junction is given by

$$-qD_p \left.\frac{dp}{dx}\right|_{x=l} = I_p(l). \tag{7}$$

The same basic considerations apply for electrons generated in the base. Equation (1) represents the net change in minority electron concentration if one substitutes $n = \Delta n + n_0$ in place of p and changes the sign of the electronic charge q. The boundary condition at $x = l$ is the analog of Equation (6) and the minority electron current to the junction is

$$+qD_n \left.\frac{dn}{dx}\right|_{x=l} = I_n(l). \tag{8}$$

In this case the charge is of opposite sign to that in Equation (7), and the electrons add to the electric current when dn/dx is positive. At $x = b$, the location of the ohmic contact to the base, it is assumed that $n(b) = n_0$.

By applying these boundary conditions, the following equations

result for the hole current:

$$I_p(l) = \frac{qN_0(1-R)}{\alpha(1-\nu_p^2/\alpha^2)} (\beta_p \nu_p e^{\nu_p l} - \gamma_p \nu_p e^{-\nu_p l} - \alpha e^{-\alpha l}) \tag{9}$$

where

$$\beta_p \Delta_p = e^{-\nu_p l}(h+\alpha) - e^{-\alpha l}(h+\nu_p),$$
$$\gamma_p \Delta_p = e^{-\alpha l}(h-\nu_p) - e^{\nu_p l}(h+\alpha),$$
$$\Delta_p = e^{\nu_p l}(h+\nu_p) - e^{-\nu_p l}(h-\nu_p);$$
$$h = s/D_p; \text{ and}$$
$$\nu_p = (D_p \tau_p)^{-1/2} = L_p^{-1}.$$

Similarly the electron contribution to the current is

$$I_n(l) = \frac{qN_0(1-R)}{\alpha(1-\nu_n^2/\alpha^2)} (\beta_n \nu_n e^{\nu_n l} - \gamma_n \nu_n e^{-\nu_n l} - \alpha e^{-\alpha l}), \tag{10}$$

$$\beta_n \Delta_n = \alpha e^{-(\nu_n l + \alpha b)} - \nu_n e^{-(\nu_n b + \alpha l)}, \tag{11}$$

$$\gamma_n \Delta_n = -\alpha e^{(\nu_n l - \alpha b)} - \nu_n e^{(\nu_n b - \alpha l)}, \tag{12}$$

$$\Delta_n = 2\nu_n \cosh(\nu_n b - \nu_n l) \tag{13}$$

where

$$\nu_n = (D_n \tau_n)^{-1/2} = L_n^{-1}.$$

In order to compare theory and experiment, it is necessary to know the ratio of the short-circuit current at a given wavelength λ (and therefore at a given α) to the maximum current, I_{max}, which can be obtained from the photons of that wavelength, namely

$$I_{max} = qN_0(1-R). \tag{14}$$

This ratio is referred to as the collection efficiency, denoted by Q, and it is given by

$$Q = \frac{I_p}{I_{max}} + \frac{I_n}{I_{max}} = Q_p + Q_n. \tag{15}$$

The two terms in Equation (15) have been calculated separately with

the aid of a computer for many permutations of plausible values of α, l, b, v_n and v_p. The values were chosen on the basis of the following considerations:

1. The maximum value of τ which one can hope to find in a semiconductor is about 1 millisecond, while values as low as 10^{-10} sec are possible. For an assumed value of D of 100 cm^2/sec, the range of values of L corresponding to the indicated range in τ lies between 10^{-2} cm $> L = v^{-1} > 10^{-6}$ cm.
2. The maximum value of α likely to be encountered for the semiconductors of interest in this application is about 10^6 cm^{-1}.

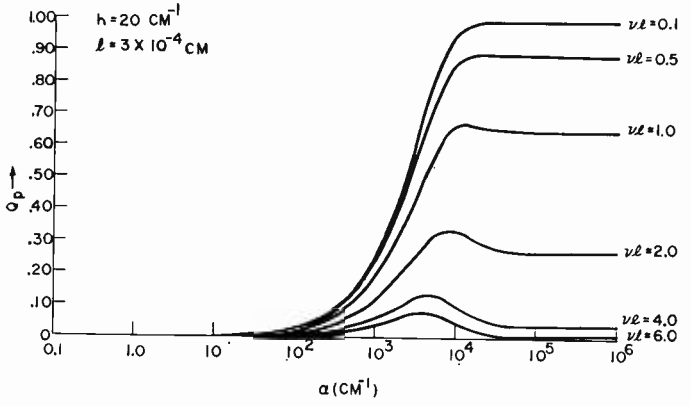


Fig. 2—Collection efficiency Q_p versus α for different values of vl.

Although the minimum value of interest would correspond to $\alpha L \approx 0.1$ (based on previous experience), values as low as 0.1 cm^{-1} have been included. Thus the range of α values used in the calculation was such that $10^6 > \alpha > 0.1$.

3. Five values of junction depths, l, were used ranging from 1 to 25 microns.
4. Four values of $h = s/D$ ranging from 20 cm^{-1} to 1000 cm^{-1} were used. Their choice was based on the fact that surface recombination velocities of a few hundred cm/sec can be readily obtained on germanium and silicon surfaces. Furthermore, values in excess of 10^4 cm/sec were not likely to be encountered in practice.

Some of the results of these calculations are shown in the figures. Figures 2 and 3 pertain to a cell with a junction depth of 3×10^{-4} cm. These figures show Q_p (collection efficiency of the diffused skin

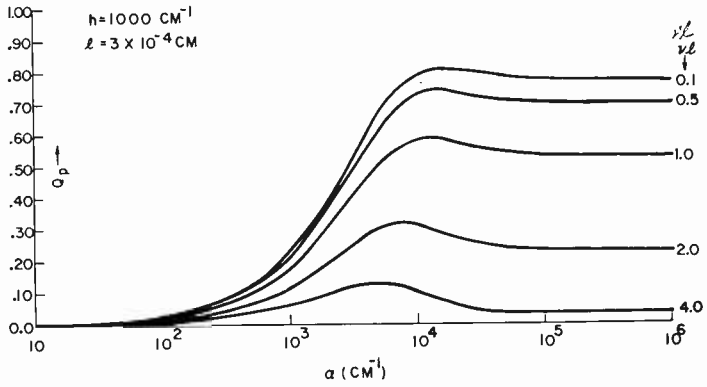


Fig. 3—Collection efficiency Q_p versus α for different values of νl.

or front part of the cell) versus α with νl as a running parameter. The value of h is different for each figure. It is evident that the value of ratio $l/L = \nu l$ seriously affects the shape of the curve at large values of α. Note that these curves can be used for any material for which α versus $h\nu$ is known by simply including an $h\nu$ scale along the abscissa.

Figure 4 shows a similar set of curves for $l = 7 \times 10^{-4}$ cm and $h = 20$ cm^{-1}. Note here that Q_p starts to rise at lower values of α than it did for the shallower junction of Figure 2. Thus infrared radiation is more efficiently used by the deep junction.

Figure 5 shows the contribution from the base or back side of the junction, i.e., from the region $l < x < b$. In efficient solar cells there is usually an appreciable contribution to Q from the back region because $\nu l \leq 0.1$. Note that this region contributes very little for

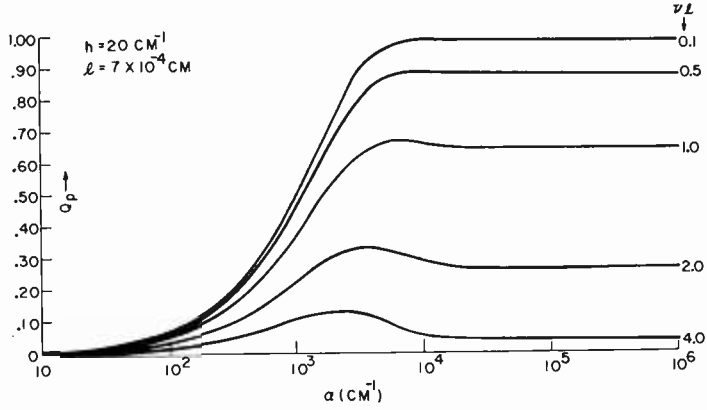


Fig. 4—Collection efficiency Q_p versus α for different values of νl.

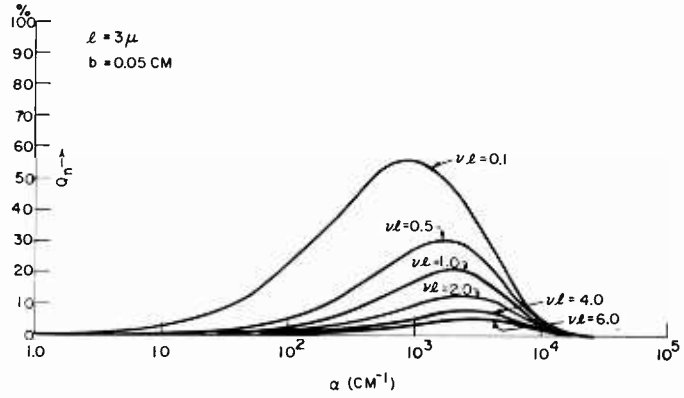


Fig. 5—Collection efficiency of the rear part of the junction versus α for different values of νl.

$\alpha > 10^4$ cm^{-1}, so that the shape of the total Q curve in that vicinity is not affected by the base response.

Figure 6 shows the total Q for $l = 3$ microns and $\nu l = 0.1$ on both sides of the junction, i.e., $L_n = L_p$. If this curve is compared with that of Figure 2, it can be seen that the base region makes a noticeable contribution to Q for $\alpha < 10^3$ cm^{-1}.

Figure 7 shows curves of Q_{total} versus $h\nu$ in silicon where the data of Dash and Newman were used to transform from an α to $h\nu$ scale on the abscissa.[19] The values of $L = 5 \times 10^{-3}$ cm in the base wafer ($\nu l = 0.02$) and $l = 10^{-4}$ cm correspond to values encountered in commercial solar cells.

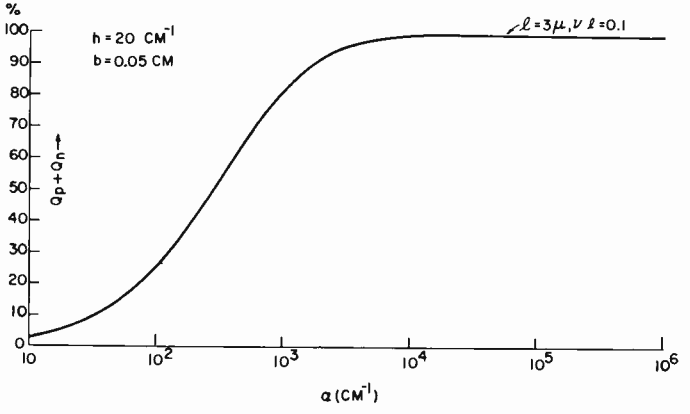


Fig. 6—Total collection efficiency versus α for $\nu l = 0.1$.

Determination of Minority Carrier Diffusion Lengths

In the preceding section, the general equations for the spectral response of an illuminated junction were presented. One can fit these equations to the absolute spectral response of a junction and, given the values of s, α and l, determine the diffusion lengths, L. The experimental measurement of such absolute Q-versus-α curves is, however, quite difficult for the kind of accuracy required to make curve fitting meaningful. The use of normalized spectral-response curves to extract information about minority-carrier diffusion lengths and, under certain circumstances, about junction depths is discussed in this section. Such relative spectral-response curves are much simpler to obtain.

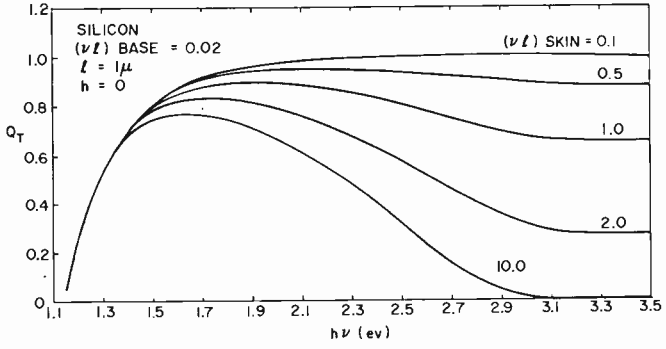


Fig. 7—Total collection efficiency versus $h\nu$ for silicon.

Consider a family of photovoltaic cells fabricated on wafers of a certain semiconductor in such a way that the junction depth l varies from one cell to the other in the series. Assume further that the surfaces of the cells are all treated in the same way so that s is constant in the family, and also that the diffusion length in the region between $0 \leq x \leq l$ is constant and has the same value in all the cells after processing. The cells in this group will then differ from each other only in the value of the ratio l/L. An examination of the equations for $Q_p/Q_{p_{\max}}$ leads to the conclusion that the shape of normalized spectral-response curves in the region of high values of α is a strong function of νl. This conclusion that the shape is primarily a function of the product νl and not of l alone is illustrated in Figure 8. The calculated normalized curves of $Q_p/Q_{p_{\max}}$ for the two different junction depths, 7×10^{-4} and 2.5×10^{-3} cm, practically coincide when superimposed. In superimposing the curves, it was necessary

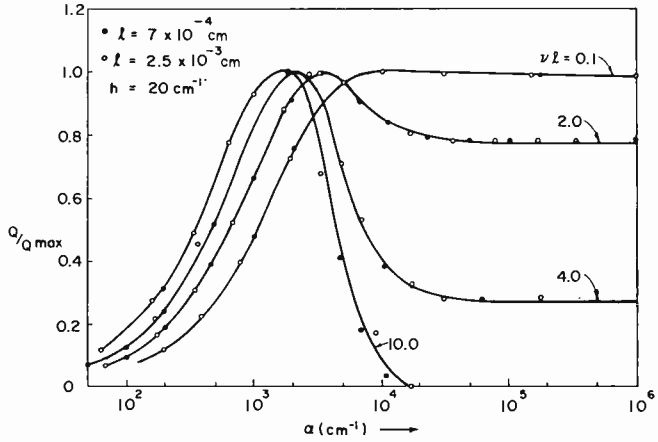


Fig. 8—$Q_p/Q_{p\max}$ versus α for different values of νl and l.

to shift the curves for $l = 2.5 \times 10^{-3}$ cm along the abscissa; the maxima of such curves occur in the region $\alpha l \sim 1$ as can be seen from a comparison of Figures 2 and 4. Furthermore, changes of surface recombination velocity (or of $h = s/D$) over a wide range do not distort the shape much at high α as is evident from Figure 9 where the curves for $s \sim 8 \times 10^2$ cm/sec are compared to those for $s \sim 10^5$ cm/sec.

In order for these considerations to apply to experimental data, there must not be any great contribution to Q_{total} from the base region at high α which could cause some distortion of the curve shape.

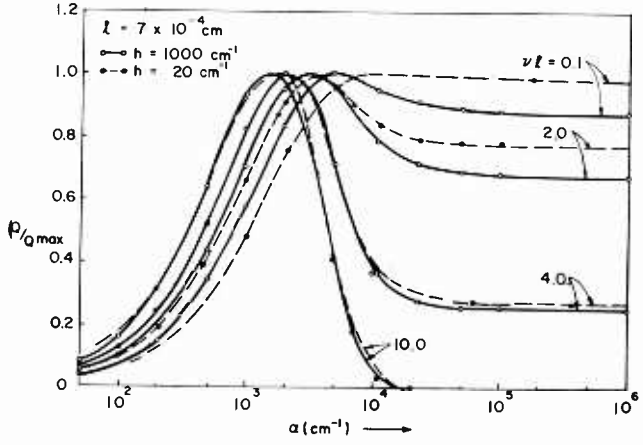


Fig. 9—$Q_p/Q_{p\max}$ versus α for different values of νl and s.

The above analysis indicates that only a value of vl can be deduced from the normalized spectral-response curve at high values of α. Therefore l must be established by independent means if the diffusion length, L, is required. The value of l can be determined from a knowledge of the junction fabrication conditions or else by subsequent sectioning of the junction. Such an analysis has been used to determine L in GaAs photovoltaic cells. Spectral-response analysis with GaAs is hampered by the lack of data on the dependence of α on $h\nu$. In spite of this handicap, it is still possible to extract the desired

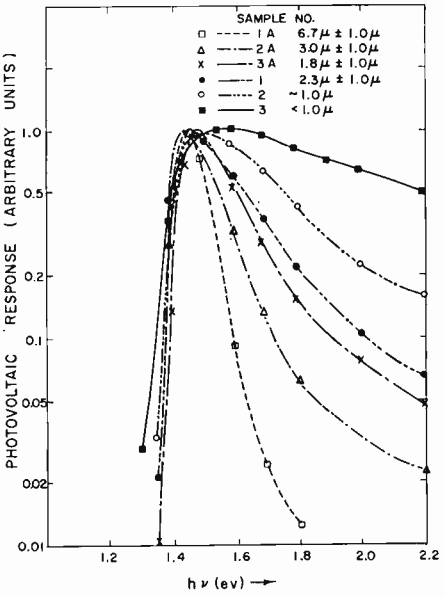


Fig. 10—Spectral response of GaAs cells.

information from such a study. A set of GaAs cells were made by diffusing Zn into wafers cut from adjacent parts of a crystal. The value of l was determined as indicated above by calculation from the diffusion time and temperature and by subsequent sectioning. The normalized spectral response versus $h\nu$ of a family of such cells is shown in Figure 10. From a comparison with the calculated curves of Figure 8 it was possible to estimate that vl is equal to about 10 in cell 1A in which $l \sim 6.7$ μ. From this value of vl, the value of L was deduced to be about 0.7 μ in the diffused skin. Assuming a value of $D_n = 25$ cm^2/sec in heavily doped p-type GaAs, τ is about 2×10^{-10} sec for electrons in the p-type region. Such a low value of τ is con-

sistent with the fact that the cells under study did not have solar energy conversion efficiencies greater than 6 per cent.

The values of L determined by such an analysis are probably correct to within a factor of two, so that this method provides a powerful and useful tool for evaluating this parameter in finished p-n junctions. If, in a given cell, the value of L is changed by some independent means, say by heat treatment or irradiation, it is possible to determine the change experienced by this parameter.

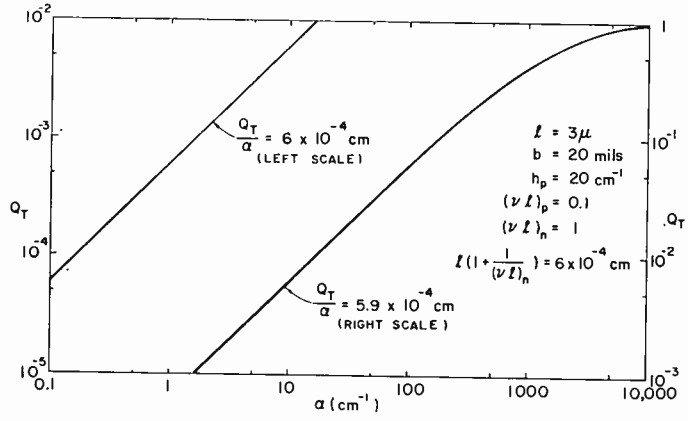


Fig. 11—Q_{Total} versus α.

Approximations of the General Solution

This section presents approximations which permit evaluation of diffusion lengths on both sides of a p-n junction.

(a) Consider the Q_{total} versus α curve in the region where α is so small that the radiation generates carriers uniformly throughout the semiconductor volume. If the following conditions, which are often encountered in practical cells, are satisfied on both sides of the junction,

$$\alpha \ll \nu, \qquad h < \nu, \tag{16}$$

then it can be shown that

$$Q_n + Q_p = Q_{\text{total}} \rightarrow \alpha\,(L_n + L_p). \tag{17}$$

A plot of Q_{total} versus α is shown in Figure 11 where a log-log plot demonstrates the linear relation between Q_{total} and α up to $\alpha = 100$

cm^{-1} for the indicated parameters. Thus, by analyzing the data at low α, it is possible to evaluate the sum $L_n + L_p$. In silicon p-on-n diffused junctions, the base diffusion length is usually much greater than the diffusion length in the skin, so that this part of the analysis yields the base diffusion length.

The analysis of the spectral response presented thus far has not included any contribution from the depletion layer of width W, and the assumption is made implicitly that this contribution is small. If the contribution to the current from the space-charge region were included, Equation (17) would change to

$$Q_{\text{total}} \rightarrow \alpha\,(L_n + L_p + W). \tag{18}$$

The value of W can be determined from junction capacitance measurements.

(b) Another interesting approximation is that for which

$$vl \ll 1,\ (v/\alpha)^2 \ll 1,\ h < v, \tag{19}$$

in which case the normalized response is

$$Q/Q_{\max} \sim (1 - e^{-\alpha l}). \tag{20}$$

In this case, an analysis of $Q/Q_{\max}$ in the vicinity of $Q = Q_{\max}$ can yield a value of junction depth in a nondestructive experiment. In order that Equation (20) be applicable, it is also necessary that the contribution from the base region be small compared to that from the diffused skin. These conditions are encountered in alloy-type p-n junctions.

(c) If the following conditions are satisfied,

$$vl \gg 1,\ \alpha \ll v,\ \alpha l \ll 1, \tag{21}$$

it is found that

$$Q_n \rightarrow \alpha l. \tag{22}$$

This result can be correlated with Equations (17) and (18). When Equation (21) applies, one finds that

$$Q_{\text{total}} = \alpha\,(l + L_n + W). \tag{23}$$

Equation (23) is often found to be applicable in shallow p-n junction photovoltaic cells.

(d) As is evident from Figures 2 through 4, Q approaches a limiting value at high values of α given by

$$Q = 2e^{-\nu l} \tag{24}$$

for $\nu l \gg 1$. A determination of this limiting value of Q combined with independent information concerning the value of l yields the value of the minority carrier diffusion length, L, in the diffused skin.

In summary, it has been shown that if the absolute value of Q is known as a function of α in the low-α and high-α regions, it is possible to evaluate diffusion lengths on both sides of the junction. Under some conditions, it is also possible to determine the junction depth, l. This method permits measurements of very small diffusion lengths, even values less than 1 micron. It also allows measurement of small changes of diffusion lengths and, therefore, of minority carrier lifetimes.

Experimental Absolute Spectral Response Curves

Measurements of absolute spectral response made on some GaAs and Si cells are presented in this section.

Gallium Arsenide

As indicated, one of the important goals of a spectral analysis of a solar cell is the possibility of obtaining the lifetimes on both sides of the junction. In one step of the analysis, a transformation of the data from a response-versus-energy scale to a response-versus-absorption-constant scale is made. The transformation is not yet possible for GaAs since the absorption constant is not known as a function of photon energy over the whole range of photon energies. It is possible, however, to obtain an estimate of the diffusion length and lifetime in the diffused skin by the following procedure. The spectral response of the cell at high photon energies saturates at a value which depends upon the value of νl in the diffused skin. When the response saturates, it is no longer a function of photon energy and absorption constant as shown earlier. Consequently, the diffusion length in the skin can be determined without knowing the absorption constant. In this case, Equation (24) is applicable if $\nu l \gg 1$.

The above procedure was used on two GaAs cells with junction depths of 2μ and 4.6μ. The absolute response measurements were made as follows. The short-circuit current of the cell was measured at a given photon energy. The number of incident photons was then measured by means of a calibrated photoemissive cell. The number

of photons actually penetrating the cell was calculated by correcting for the photons reflected from the surface. The ratio of the short-circuit current to the corrected number of incident photons is the required collection efficiency, Q.

The results of this measurement for the two GaAs junctions are shown in Figure 12. Q is plotted as a function of photon energy. The response tends to saturation at the higher photon energies. By extrapolating the data toward the apparent saturation values, the diffusion lengths in the skin are calculated to be 0.5 and 0.6μ. The

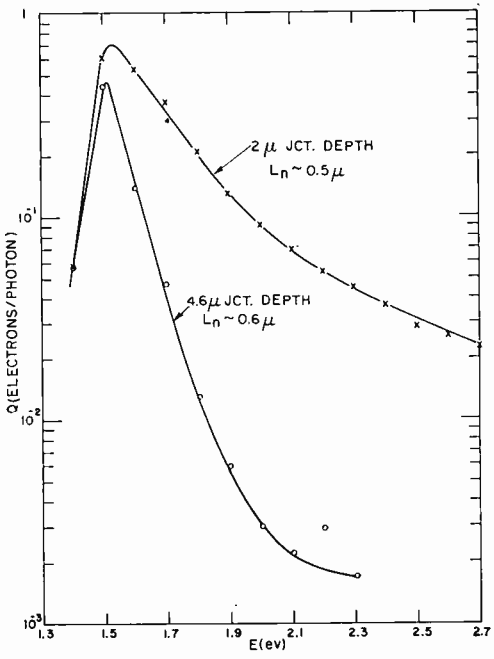


Fig. 12—Absolute response of GaAs photovoltaic cells.

lifetime is thus 10^{-10} sec, if one assumes a diffusion constant of 25 cm^2/sec. This value of lifetime is consistent with the value deduced earlier from the normalized response as well as with values deduced from other measurements involving diffused skins on GaAs.

Silicon

The literature contains examples of experimental studies of spectral response for single silicon cells. The primary objective of these experiments on a family of silicon cells is to prove the applicability of the theoretical expression for the spectral response on a material

which is better understood. Absolute measurements were made for a more direct comparison with theory. The cells consisted of phosphorus-diffused p-type silicon with a base resistivity of 0.24 ohm-cm, and boron-diffused n-type silicon with a base resistivity of 0.85 ohm-cm. The junction depths varied from 1 to 25μ.

An example of the results obtained on the #3—1μ (p-on-n) cell is shown in Figure 13. The absolute response, Q, is plotted against absorption constant, α. The data of Dash and Newman was used to convert from photon energy to absorption constant. The average data for Q from several measurements is shown. Two methods were used in the measurements. In the d-c method, the short-circuit current and

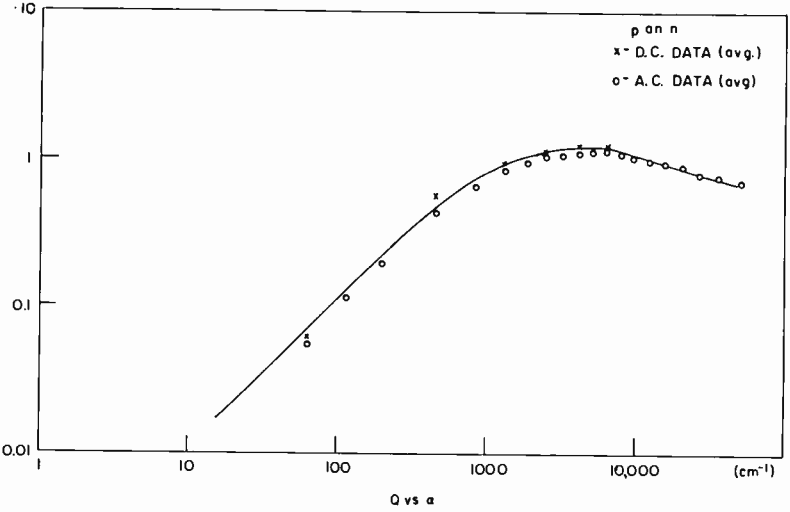


Fig. 13—Absolute response, Q, of silicon cell #3-1μ.

number of incident photons were determined at discreet photon energies by means of d-c meters. The a-c data was determined by measuring the relative response of the cell with a constant light-energy input and subsequently calibrating one point on the response. The standard deviation of the measurements for either method is ±10 per cent. The a-c and d-c data differ by 10 per cent in the region of large response and up to 30 per cent for smaller responses. The peak response in Figure 13 goes above unity by about 25 per cent. Since the peak response occurs for photon energies less than 2 ev, the magnitude of the peak is not indicative of electron multiplication, but rather of the over-all error involved in the measurement.

The response in Figure 13 is consistent with theory in the following respects. First, the shape of the curve is what one would expect for this particular junction. Secondly, the response at low α is proportional to α. A diffusion length can consequently be calculated from the low-α region. The response saturates at high α allowing the diffusion length in the diffused skin to be determined.

The absolute results for other p-on-n cells in this family are shown in Figure 14. Only the a-c data is included. Qualitatively, all the cells follow theoretical expectations except #4—3μ. The response at high α is a strong function of the junction depth. The peak response becomes greater as the junction depth is reduced. The peak response shifts

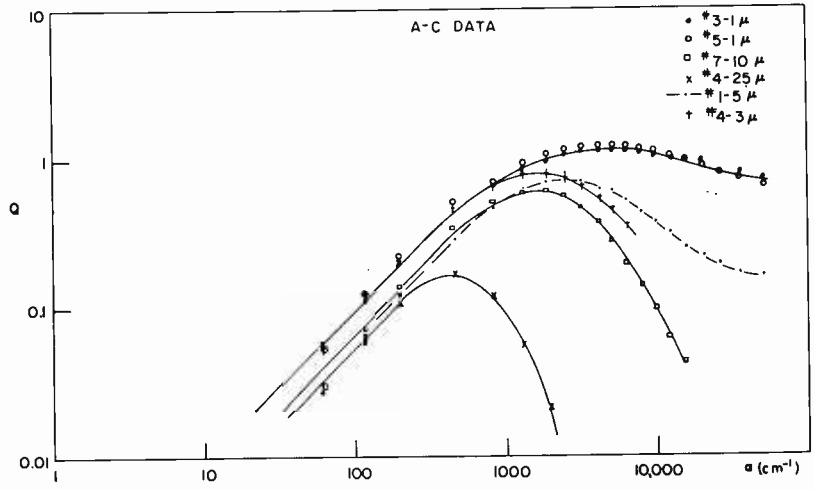


Fig. 14—Absolute response of p-on-n silicon cells.

to lower α as the junction depth is increased. Cell #4—3μ, the exception to this behavior, may actualy have a junction depth other than 3μ. Measurements of the junction depths have not been made as yet because such measurements destroy the cells. The junction-depth values quoted are those obtained from blanks diffused at the same time.

The a-c data for n-on-p cells is shown in Figure 15. Similar remarks apply to this data except that the response does not go above unity for the 1-μ junction and that the response at high α's falls rapidly even for the cell with a 1-μ junction depth indicating a very low value of diffusion length in the n-type skin.

Diffusion lengths in the various cells can be calculated from the

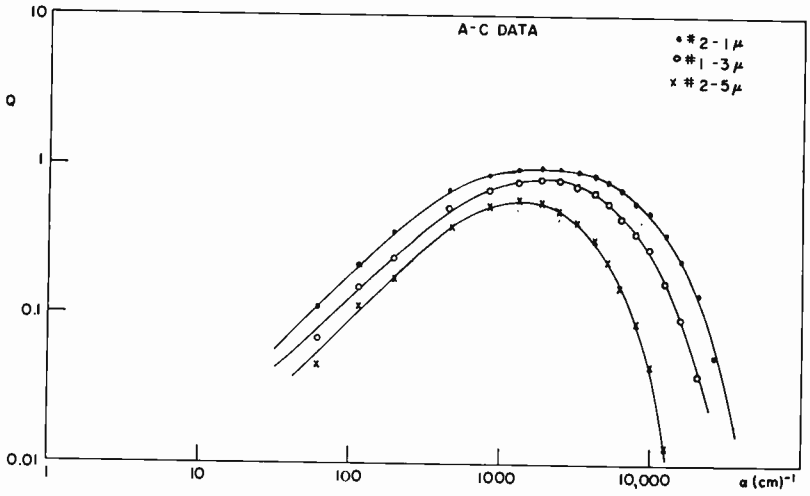


Fig. 15—Absolute response of n-on-p silicon cells.

Table I—Data on Silicon Cells

Cell	L_p (μ)	L_n (μ)	τ_n* (sec)	τ_p (sec)	$\tau_{measured}$† (sec)
p on n					
#3-1μ	11	0.91	4×10^{-9}	1.3×10^{-7}	2×10^{-7}
#5-1μ	11	0.84	3.6×10^{-9}	1.3×10^{-7}	$1\text{-}2 \times 10^{-7}$
#4-3μ	10	—	———	10^{-7}	$1\text{-}2 \times 10^{-7}$
#1-5μ	4.1	1.9	1.8×10^{-8}	1.8×10^{-8}	2×10^{-7}
#7-10μ	7	—	———	5.1×10^{-8}	———
#4-25μ	5.7	—	———	3.4×10^{-8}	2×10^{-7}
n on p					
#2-1μ	—	18	3.1×10^{-7}	———	$3\text{-}4 \times 10^{-7}$
#1-3μ	—	12	1.4×10^{-7}	———	2×10^{-7}
#2-5μ	—	8.9	7.7×10^{-8}	———	2×10^{-7}

* For p-on-n cells, $D_p = 9.5$ cm^2/sec, $D_n = 2$ cm^2/sec. For n-on-p cells, $D_p = 1$ cm^2/sec, $D_n = 10.3$ cm^2/sec. These diffusion constants were estimated from M. Prince.[21]

† In the pulse-injection technique, $\tau_{measured}$ is primarily the lifetime of the carriers injected into the base when the emitter lifetime is small and the emitter doping is high.

experimental data according to the conditions listed above. The lifetimes obtained from the diffusion lengths can be compared to those measured by other techniques such as the pulse-injection technique.[20] A comparison of the data is given in Table I. Observe that the measured lifetime in the shallower junctions agrees quite well with the lifetime in the base of the cell deduced from the spectral response. The comparison is much poorer for the deeper junctions. This discrepancy is not understood at present. Consequently, the cells will be re-examined to improve the accuracy of the absolute measurements.

CONCLUSIONS

The spectral response of p-n junctions has been examined both theoretically and experimentally. The equations governing the response of the junction to light have been derived, and the limiting forms of these equations have been examined for various interesting cases. Experimental results generally agree with the predicted behavior of the spectral response. However, several discrepancies indicate that a more careful determination of the absolute spectral response is required.

The spectral response offers a relatively simple way of determining the diffusion lengths on both sides of a p-n junction. Furthermore, even relatively small diffusion lengths can be determined, and changes in the diffusion lengths as a result of processing, radiation damage, etc., can be easily studied. The spectral response consequently deserves further attention and use.

ACKNOWLEDGMENTS

The authors wish to acknowledge the help and support of Paul Rappaport and Peter Vajk. The families of silicon solar cells were kindly furnished by John Scott-Monck. The spectral measurements were made with the help and equipment of W. Spicer. We also wish to thank F. Herzfeld for his aid in machine computations.

REFERENCES

[1] K. Lehovec, "The Photo-Voltaic Effect," *Phys. Rev.*, Vol. 74, p. 463, August 15, 1948.

[2] R. Cummerow, "Photovoltaic Effect in p-n Junctions," *Phys. Rev.*, Vol. 95, p. 16, July 1, 1954.

[3] R. Cummerow, "Use of Silicon p-n Junctions for Converting Solar Energy to Electrical Energy," *Phys. Rev.*, Vol. 95, p. 561, July 15, 1954.

[4] D. M. Chapin, C. S. Fuller, and G. L. Pearson, "A New Silicon p-n Junction Photocell for Converting Solar Radiation into Electrical Power," *Jour. Appl. Phys.*, Vol. 25, p. 676, May, 1954.

[5] W. Pfann and W. van Roosbroeck, "Radioactive and Photoelectric p-n Junction Power Sources," *Jour. Appl. Phys.*, Vol. 25, p. 1422, November, 1954.

[6] H. U. Harten and W. Schultz, "Einfluss von Diffusionslänge und Oberflächenrekombination auf den Sperrschicht-Photoeffekt an Germanium," *Zeit. Phys.*, Vol. 141, p. 319, July, 1955.

[7] D. A. Jenny, J. J. Loferski, and P. Rappaport, "Photovoltaic Effect in GaAs p-n Junctions and Solar Energy Conversion," *Phys. Rev.*, Vol. 101, p. 1208, February 1, 1956.

[8] J. J. Loferski, "Theoretical Considerations Governing the Choice of the Optimum Semiconductor for Photovoltaic Solar Energy Conversion," *Jour. Appl. Phys.*, Vol. 27, p. 777, July, 1956.

[9] P. Rappaport, J. J. Loferski, and E. G. Linder, "The Electron-Voltaic Effect in Germanium and Silicon P-N Junctions," *RCA Review*, Vol. 17, p. 100, March, 1956.

[10] M. B. Prince and M. Wolf, "New Developments in Silicon Photovoltaic Devices," *Jour. British I.R.E.*, Vol. 18, p. 583, October, 1958.

[11] J. J. Loferski, P. Rappaport, and J. J. Wysocki, *Proc. 13th Annual Power Sources Conf.*, 59 (1959).

[12] M. Wolf, "Limitations and Possibilities for Improvement of Photovoltaic Solar Energy Converters," *Proc. I.R.E.*, Vol. 48, p. 1246, July, 1960.

[13] J. J. Wysocki, J. J. Loferski, and P. Rappaport, *Proc. 14th Annual Power Sources Conf.*, 32 (1960).

[14] D. N. Nasledov and B. V. Tsarenkov, "Spectral Characteristics of GaAs Photocells," *Soviet Physics—Solid State*, Vol. 1, No. 9, p. 1346, March, 1960.

[15] V. S. Vavilov, L. S. Smirnov, and V. M. Patshevich, "The Diffusion Length of Charge Carriers in Silicon Photocells," *Soviet Physics—Solid State*, Vol. 1, No. 9, p. 1344, March, 1960.

[16] V. K. Subashiev, "Determination of Recombination Constants from Spectral Characteristic of a Photoelement with a p-n Junction," *Soviet Physics—Solid State*, Vol. 2, No. 2, p. 187, August, 1960.

[17] B. Ya. Moizhes, "On the Theory of Photocells with a p-n Junction," *Soviet Physics—Solid State*, Vol. 2, No. 2, p. 202, August, 1960.

[18] A. M. Goodman, RCA Laboratories, has proposed making such measurements by analyzing surface photovoltage spectral responses curves —private communication.

[19] W. C. Dash and R. Newman, "Intrinsic Optical Absorption in Single-Crystal Germanium and Silicon at 77°K and 300°K," *Phys. Rev.*, Vol. 99, p. 1151, August 15, 1955.

[20] S. Lederhandler and L. Giacoletto, "Measurement of Minority Carrier Lifetime and Surface Effects in Junction Devices," *Proc. I.R.E.*, Vol. 43, p. 477, April, 1955.

[21] M. B. Prince, "Drift Mobilities in Semiconductors. II. Silicon," *Phys. Rev.*, Vol. 93, p. 1204, March 15, 1954.

THE EFFECT OF SERIES RESISTANCE ON PHOTOVOLTAIC SOLAR ENERGY CONVERSION*†

By

JOSEPH J. WYSOCKI

RCA Laboratories
Princeton, N. J.

Summary—The series resistance in a photovoltaic cell is divided into two components: contact and sheet resistance. Each of the components is examined theoretically and experimentally, and qualitative agreement between theory and experiment is shown. It is concluded that contact resistance reduces the conversion efficiency more than sheet resistance.

INTRODUCTION

THIS PAPER considers the effect of series resistance on the photovoltaic process from both a theoretical and an experimental viewpoint. To facilitate the discussion, the resistance is divided into two components: contact resistance and sheet resistance. As the name implies, contact resistance arises in making contact to the junction. Sheet resistance, on the other hand, arises in the thin channel defined by the illuminated surface and the junction. The generated current must flow through this channel to reach the load. The effects of these resistances on the conversion efficiency, current–voltage characteristic, and the short-circuit current in GaAs solar cells is determined. As an indication of the importance of these effects, it may be pointed out that Prince calculates that the addition of 5 ohms to a silicon cell with an area of several cm^2 is sufficient to reduce the conversion efficiency by 70 per cent.[1]

GENERAL

It is convenient in this discussion to deal with current densities. When series resistance is included, however, the occurrence of terms involving the product of resistance and current is unavoidable. To maintain an approach consistent with the use of current densities, therefore, it is necessary to use a modified resistance consisting of the

* Manuscript received October 18, 1960.

† This work was supported by the U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N. J., under contract DA36-039-sc-78184.

[1] M. Prince, "Silicon Solar Energy Converters," *Jour. Appl. Phys.*, Vol. 26, p. 534, May, 1955.

product of the resistance and the active cell area (ohm cm^2). Fortunately, this approach also allows generalization to a cell of arbitrary area.

The behavior of any photovoltaic cell depends strongly upon the junction current density, I_j, as well as on the series resistance.[2] The variation of I_j with the material properties of the semiconductor can be quite complex.[2,3] In the following discussion, I_j will be assumed to depend upon the junction voltage, V_j;

$$I_j = I_0 \, (e^{\lambda V_j} - 1). \tag{1}$$

The constant, λ, is equal to $q/(aKT)$ where $1 \leq a \leq 2$. I_0 itself is a function of the lifetime and material constants, and is exponentially dependent upon the semiconductor bandgap, E_g;

$$I_0 \propto C \, e^{-\lambda E_g}. \tag{2}$$

The constant, a, will be taken to be 2 in the following discussion since experimental cells frequently have values close to 2. Using this value for a, I_0 for GaAs is computed to be 1.4×10^{-8} amp/cm^2 corresponding to a minority-carrier lifetime of 10^{-8} sec. The use of GaAs cells to illustrate the results, together with the above values of a and I_0 in no way restricts or limits their validity.

Contact Resistance

Contact resistance arises in making contact to the junction and can, therefore, be considered as a fixed resistance located at the output terminals as shown in Figure 1. In this figure, I_s is the generated current density; R_c is the contact resistance in ohm cm^2; and I_L and V_L are the load-current density and voltage, respectively.

By manipulating the Kirchhoff current–voltage laws for this simple circuit, the conversion efficiency, short-circuit current-density, and various other properties can be calculated. A value for I_s of 45 ma/cm^2 will be used, corresponding to a solar power input of 135 mw/cm^2 for GaAs cells.[4]

---

[2] J. J. Wysocki and P. Rappaport, "Effect of Temperature on Photovoltaic Solar Energy Conversion," *Jour. Appl. Phys.*, Vol. 31, p. 571, March, 1960.

[3] C. T. Sah, R. N. Noyce, and W. Shockley, "Carrier Generation and Recombination in P-N Junctions and P-N Junction Characteristics," *Proc. I.R.E.*, Vol. 45, p. 1228, September, 1957.

[4] J. J. Loferski, "Theoretical Considerations Governing the Choice of the Optimum Semiconductor for Photovoltaic Solar Energy Conversion," *Jour. Appl. Phys.*, Vol. 27, p. 777, July, 1956.

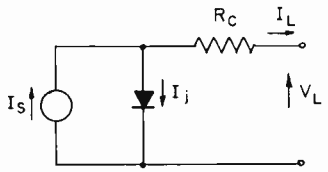


Fig. 1—Equivalent circuit of photovoltaic cell with contact resistance R_c.

The current–voltage characteristic of the solar cell is

$$V_L = \frac{1}{\lambda} \ln \left[\frac{I_s - I_L}{I_0} + 1 \right] - I_L R_c. \tag{3}$$

A plot of this equation is shown in Figure 2. The current–voltage characteristic becomes less rectangular as R_c increases, approaching a straight line for high values. The open-circuit voltage is independent of R_c, but the short-circuit current density falls as R_c is increased above 10 ohm cm^2.

It is frequently instructive to analyze the results shown in Figure 2 to obtain the junction current–voltage characteristic, λ, and I_0. Such an analysis can only be performed if R_c is taken into account. Otherwise, the situation shown in Figure 3 may result. This figure is a semilog plot of I_j versus V_L with R_c as a parameter. The slopes of these curves can only be identified with λ for the case of zero R_c. Furthermore, the intercept of the straight-line portion of these curves with the I_j axis is higher than I_0 when R_c is greater than zero. While the above plot will give misleading values of I_0 and λ when R_c is present, the plot may be used to determine R_c if the junction characteristic for zero R_c is known.

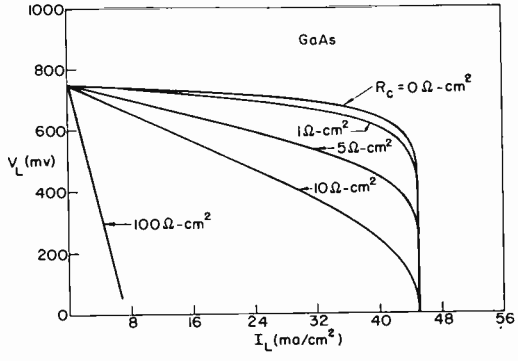


Fig. 2—Current–voltage characteristic of solar cell (Equation (3)).

The conversion efficiency, η, of a solar cell is

$$\eta = \frac{V_L I_L}{1.35}, \tag{4}$$

for a solar power input of 135 mw/cm^2. By using Equation (3), η

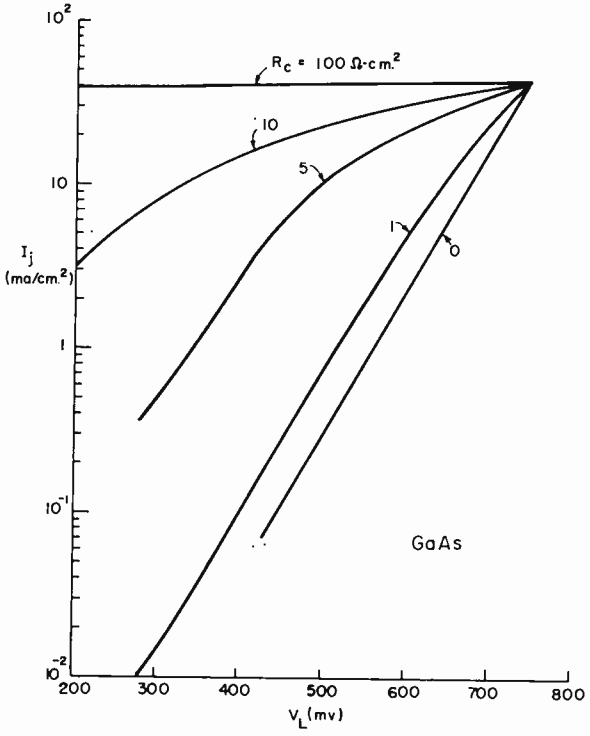


Fig. 3—I_j versus V_L with R_c as parameter.

can be expressed as a function of I_L and R_c. A plot of this function is shown in Figure 4. The effect of R_c on η is clearly seen. First, η is reduced as R_c is increased. For example, the addition of 10 ohms to a cell with an area of 1 cm^2 is sufficient to reduce η_{max} from 19 per cent to 8.8 per cent. Secondly, η_{max} occurs at smaller values of I_L as R_c is increased. This shift is a consequence of the effective increase in internal resistance of the cell.

Another parameter on which R_c has an effect is the short-circuit current density, I_{sc};

$$I_s = I_{sc} + I_0 \, (e^{\lambda R_c I_{sc}} - 1). \tag{5}$$

Figure 5 shows this effect. I_{sc} tends to saturate at a value which depends upon R_c. This value is very high for low values of R_c. For example, the curve departs from linearity at 120 ma/cm^2 when R_c is 5 ohm cm^2. This value of I_s corresponds to an input flux of 7.5×10^{17} photons/cm^2/sec.

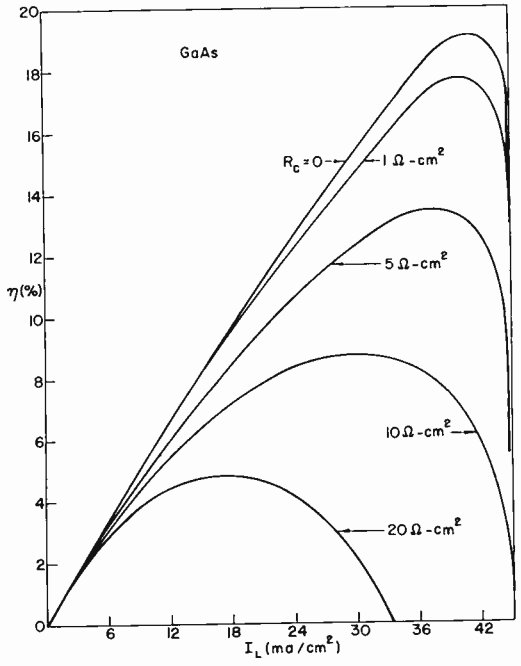


Fig. 4—η versus I_L.

The effect of R_c can also be seen by examination of the forward characteristic of the cell as a rectifier in the dark. The forward current is limited by R_c at sufficiently high biases. Also, the internal impedance of a solar cell is a strong function of R_c. All of these effects are potentially useful in determining R_c in a given cell. In general, however, the accuracy with which R_c can be determined is high only when R_c is large.

Sheet Resistance

It is sometimes necessary to place the p-n junction close to the illuminated surface to achieve high collection of the generated carriers.

The resulting thin channel between the surface and the junction gives rise to the sheet resistance, R_s. Unfortunately, R_s cannot be considered independently of the junction. Current flow in each element of R_s produces a bias which determines I_j in the vicinity of the resistance element. In other words, R_s is a distributed resistance, as shown in Figure 6. The one-dimensional geometry which will be used is shown in Figure 6a. As can be seen, a load is connected at the origin,

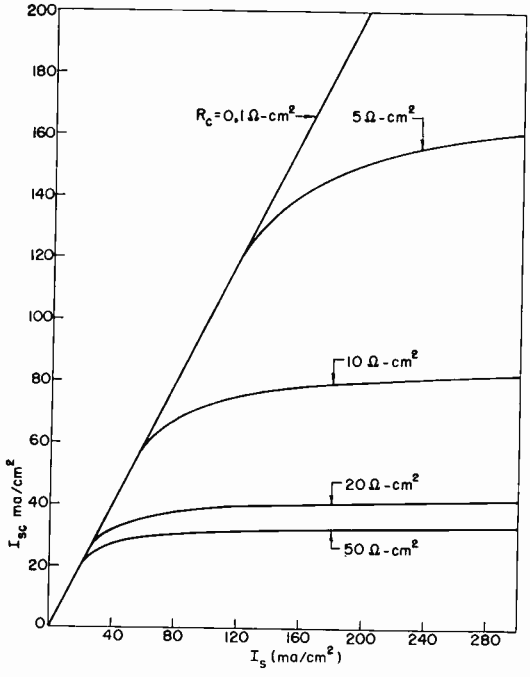


Fig. 5—I_{sc} versus I_s.

the length of the cell perpendicular to the contact is denoted L, and the current densities in the thin channel and the junction at any position x are denoted $I(x)$ and $I_j(x)$, respectively. An equivalent lumped circuit is shown in Figure 6b to illustrate the nature of the problem. The elemental sheet resistors are $R_s\Delta x/L$, while the elemental currents are as shown. As Δx is made smaller, the behavior of any section is determined by the following differential equations:[5]

[5] A similar approach is taken by B. Ya Moizhes in "On the Theory of Photocells with a p-n Junction," *Solid State Physics*, Vol. 2, p. 202, August, 1960.

$$\frac{dV}{dx'} = R_s I; \tag{6}$$

$$\frac{dI}{dx'} = -(I_s - I_j), \tag{7}$$

where x' is the dimensionless parameter, x/L.

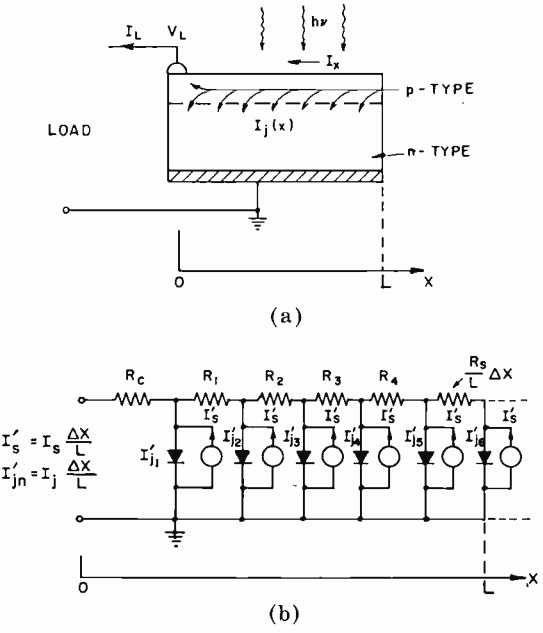


Fig. 6—(a) Photovoltaic cell and (b) lumped-circuit equivalent of photovoltaic cell.

The boundary conditions imposed on the above equations are

$$I(L) = 0; \tag{8}$$

$$I(0) = V(0)/R_L. \tag{9}$$

The differential equations are nonlinear because of the I_j term in Equation (7). A solution has been obtained by means of a computer for various values of R_s and R_L.

When R_s is small, the top face of the cell is approximately an equipotential whose value depends upon R_L. The current is a linear

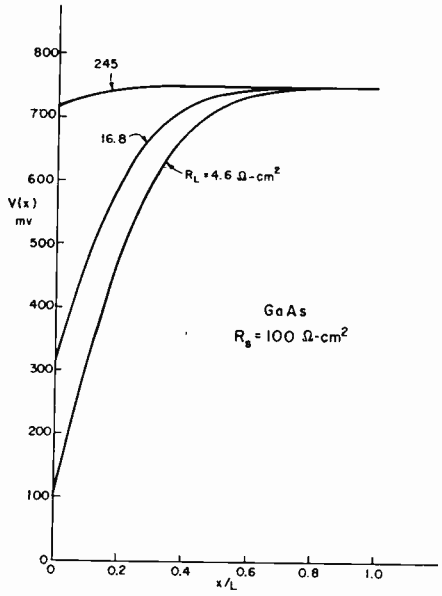


Fig. 7—$V(x)$ versus x/L with R_L as parameter.

function of x increasing from 0 at L to a value at the origin dependent upon R_L. This type of behavior is characteristic of cells with low R_s. As R_s is increased, however, $V(x)$ and $I(x)$ change as shown in Figures (7) and (8), respectively. The top face is no longer an equi-

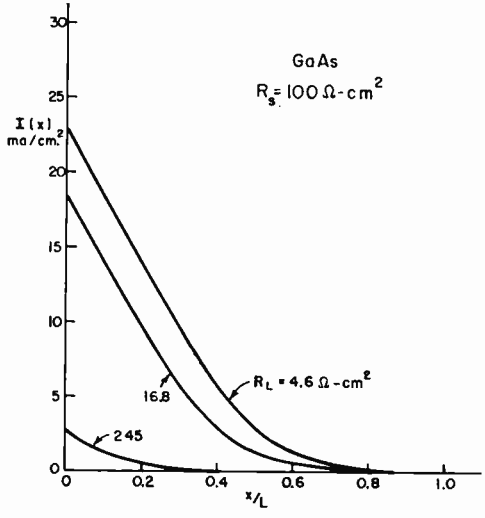


Fig. 8—$I(x)$ versus x/L with R_L as parameter.

potential. The voltage increases rapidly with x, approaching open-circuit conditions. Consequently, the effective cell length contributing current to the load is less than the actual length for large R_s.

The effective cell length is also a function of load conditions. This behavior can be seen more clearly in Figure 8 which shows $I(x)$ for the same conditions as in Figure 7. Only 0.6 of the total area is effective when R_L is 4.6 ohm cm^2. The effective area decreases to 0.4 when R_L is 245 ohm cm^2. It can be seen, therefore, that the effective area is a function of R_L when R_s is high.

The sheet resistance affects the current–voltage characteristic of the cell in the manner shown in Figure 2. The curve becomes less rectangular as R_s is increased. For any given value of resistance, however, the effect is greater when the resistance is concentrated as contact rather than as sheet resistance. This result is not unexpected. Similar remarks may be made on the effect of R_s on the I_j versus V_L plot (Figure 3) and on I_{sc} (Figure 5). In the case of I_{sc}, the saturation is less pronounced when R_s is considered.

The more undesirable resistance, in most cases, is R_c. This result is illustrated in Figure 9, where $\eta_{\max}$ is plotted as a function of resistance. The efficiency falls more rapidly with R_c than R_s. Consequently. R_c is more objectionable in cell performance.

Experimental Results

Since GaAs cells had low values of R_c when fabricated, external resistance was purposely added to the cell to see if the resultant characteristic could be used to determine R_c. Figure 10 shows a current–voltage characteristic determined in such fashion. The similarity between Figures 2 and 10 is evident. The cell has a rectangular current–voltage characteristic for low values of R_{ext}. The characteristic becomes linear when R_{ext} is 100 ohm-cm^2.

The data was used to compute the power output as a function of I_L. The results are shown in Figure 11. Here, the similarity to Figure 4 is evident. There is a discrepancy, however, in actual magnitudes. This discrepancy is characterized by the low efficiency of the experimental cell and is attributed to poor collection of the generated carriers due to low minority-carrier lifetimes.

The effect of R_c was also observed in the variation of I_{sc} with light level. The results are shown in Figure 12. The unity light level corresponds to an input of 100 mw/cm^2. The light level was changed by means of neutral density filters. Saturation does indeed occur, but the sensitivity is poor for the light levels used. No effect is seen.

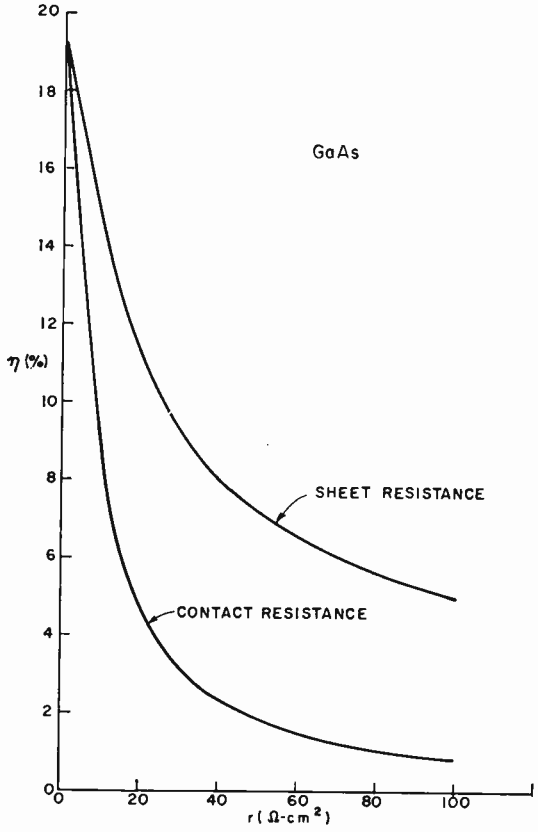


Fig. 9—η versus R_s and R_c.

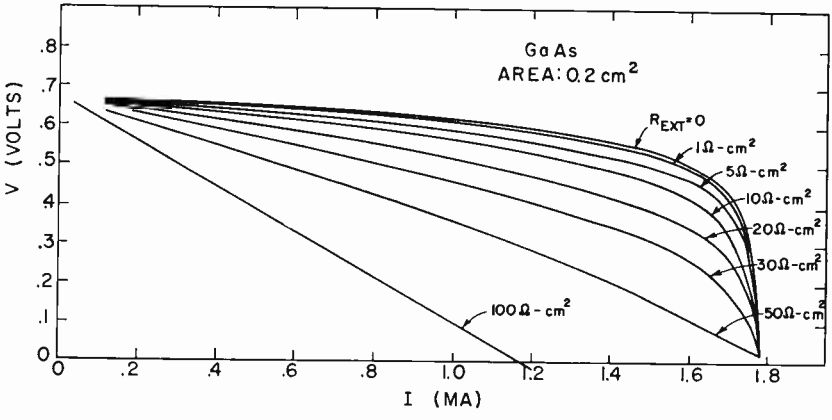


Fig. 10—Current–voltage characteristic of solar cell with external resistance added as indicated.

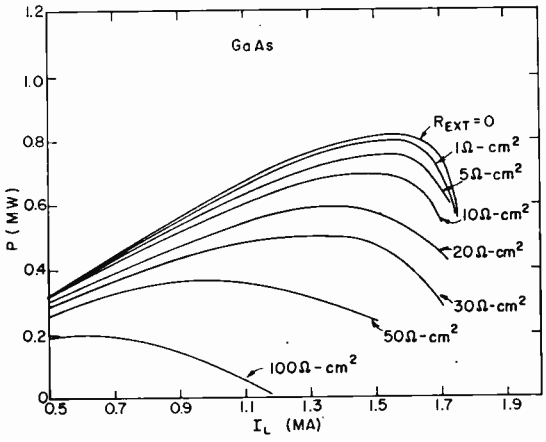


Fig. 11—Power versus I_L.

for instance, for external resistances up to 20 ohm-cm². The sensitivity could be improved by using a more intense light source. The nonlinearity in the curves near zero light level is due to infrared transmission through the filters.

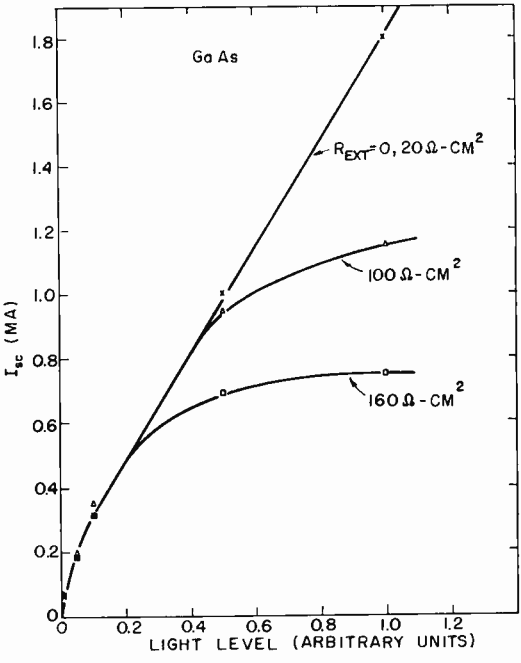


Fig. 12—I_{sc} versus light level.

Two methods were used to investigate sheet resistance. In the first, the potential of the surface with respect to the base contact was measured as a function of distance from the top contact. Secondly, I_{sc} was examined as a function of light intensity. In both cases, the sheet resistance was varied by changing the channel depth either by etching or by changes in the diffusion schedule.

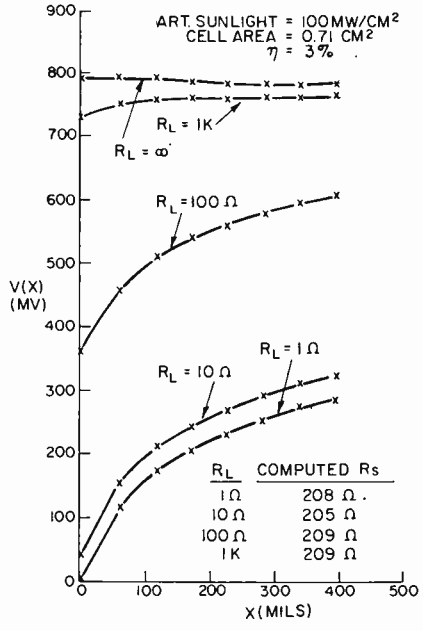


Fig. 13—$V(x)$ versus x for a GaAs cell.

The potential plot obtained on one GaAs cell is shown in Figure 13. The voltage increases with x as expected. It is possible to compute the effective value of R_s from the slope of the curve near the origin. The value so obtained is approximately 147 ohm cm². This value was constant to within 10 per cent for light levels varying from 69 to 154 mw/cm².

Similar plots were obtained when R_s was increased by reducing the channel depth through the use of chemical etching. The results were consistent with the expected variation of potential with R_s. As R_s increased, the potential drop at any point, x, increased toward the limit of the open-circuit voltage.

The short-circuit current passed through a maximum during the process of etching. This observation is interpreted as follows. As

the channel depth is reduced, more of the carriers generated near the surface are collected. The short-circuit current consequently increases. R_s increases at the same time, however, until eventually it dominates and I_{sc} falls.

Typical results obtained during the etching experiments on the variation of I_{sc} with light level are shown in Figure 14. A $CuCl_2$ filter was used to remove the infrared portion of the light from a

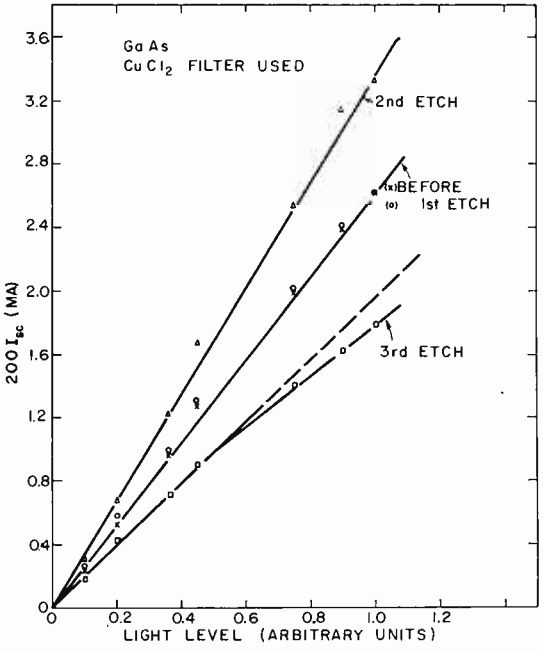


Fig. 14—I_{sc} versus light level during etching.

focused microscope lamp. Up to the third etch, I_{sc} was a linear function of light level and increased after each etch. After the third and subsequent etches, I_{sc} decreased and the curve was no longer linear. The departure from linearity is small because the light was low. Furthermore, saturation is not prominent in the case of sheet resistance as pointed out in the theoretical discussion.

CONCLUSIONS

The above experiments confirm theory qualitatively. The discrepancy in quantitive comparisons was due to several factors. First, it is not completely possible to divide the series resistance into two com-

ponents experimentally. Both resistances are present in any test. Secondly, other effects associated with the photovoltaic process (e.g. the collection of generated carriers) influence the experimental data and make a quantative comparison difficult.

The conclusions of this study are as follows. First, series resistance has an important effect on the photovoltaic process. Secondly, contact resistance reduces the conversion efficiency more than sheet resistance. Thirdly, the effective area of a cell may be less than the active area when sheet resistance is high. Furthermore, the effective area becomes a function of the load conditions.

The series resistance may be reduced by making better ohmic contacts and by using highly-doped diffused skins, deep junction depths, and conducting grids on the illuminated face.[6]

ACKNOWLEDGMENTS

The author wishes to acknowledge the help and support of P. Rappaport and J. J. Loferski. The machine calculations were performed by S. Skillman.

---

[6] M. Wolf, "Limitations and Possibilities for Improvement of Photovoltaic Solar Energy Converters." *Proc. I.R.E.*, Vol. 48, p. 1246, July, 1960.

CONSIDERATIONS OF PHOTOEMISSIVE ENERGY CONVERTERS*†

By

W. E. Spicer

RCA Laboratories
Princeton, N. J.

***Summary**—The efficiency of a solar-energy converter consisting of a [Cs]Na_2KSb emitter and an Ag-O-Cs collector is calculated taking into account the initial velocities of the photoelectrons but ignoring space charge. Efficiencies between 2 and 2.1 per cent are obtained for output voltages between 0.8 and 1.6 volts. The efficiency increases as the percentage of blue and ultraviolet radiation in the source is increased. To minimize space-charge effects, the emitter–collector spacing must be of the order of 0.01 cm or less. From our present knowledge, it appears that photoemission energy converters will be insensitive to radiation damage.*

Introduction

Currently, there is a great deal of interest in devices which convert radiant energy into electric energy. Among the devices of possible importance in such applications are photoemissive converters. Our present knowledge of photoemitters makes it possible to estimate the efficiencies which might be obtained from such devices and to discuss some of their advantages and disadvantages. The efficiencies are considered in three steps. First, a calculation is made ignoring the effect of the initial velocities of the photoelectrons and the effect of space charge; next, the effect of initial velocities is considered; and, finally, the space-charge problem is discussed.

Calculation of Efficiencies Ignoring Initial Velocities and Space Charge

The principle of operation of a photoemissive energy converter is indicated in Figure 1. Here, we have two materials of work function ϕ_e and ϕ_c facing each other with a contact potential difference, or CPD, of $V_c = \phi_e - \phi_c$. Radiation falling on the material of higher work

† The research reported in this paper has been sponsored in whole or in part by the U.S. Air Force under Contract AF33(616)-7165.

* Manuscript received 20 October 1960.

function† will produce photoelectrons which will all be collected by the anode as long as the output voltage is equal to or less than V_c. Thus, the power, P, which would be delivered into a perfectly matched output, $V_0 = V_c$, would be

$$P = I_0 V_c,$$

where I_0 is the saturated photoelectric current. The efficiency would be given by

$$E = \frac{P}{L} = \frac{I_0 V_c}{L}, \tag{1}$$

where L is the radiant-energy flux incident on the converter.

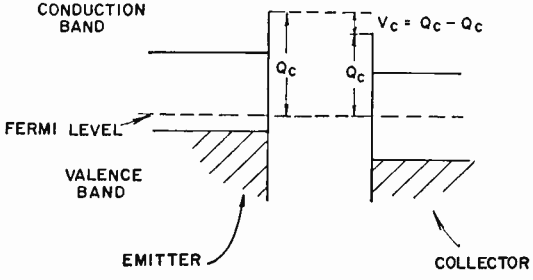


Fig. 1—Energy-level diagram for a photoemissive energy converter using semiconductors as emitter and collector. ϕ_e and ϕ_c are the work functions of emitter and collector, respectively. $V_c = \phi_e - \phi_c$ is the contact potential difference (CPD) between emitter and collector. This diagram is drawn for the case where no output power is being developed.

Values for the parameters necessary in Equation (1) can be obtained from experimental data. For solar radiation, it has been shown[1] that the most useful emitters would be [Cs]Na_2KSb, Na_2KSb, and Cs_3Sb. Since the parameters involved in the calculations are so similar for Cs_3Sb and Na_2KSb, the calculations are made only for Na_2KSb. Because of its low work function,[1,2] Ag-O-Cs appears to be the optimum

† The converter geometry must be such that the photoemission from the collector is small. Through the use of semitransparent emitters, optical filters, and/or other schemes, this should be possible.

[1] A. H. Sommer and W. E. Spicer, *Methods of Experimental Physics*, Vol. 6B, p. 376, Ed. by K. Lark-Horovitz and V. A. Johnson, Academic Press (New York and London), 1959.

[2] John E. Davey, "Thermionic and Semiconducting Properties of [Ag]-Cs_2O, Ag, Cs," *Jour. Appl. Phys.*, Vol. 28, p. 1031, Sept. 1957.

collector. It is not considered as an emitter because of its low photoelectric quantum efficiency. The spectral distribution of the incident radiation is taken to be that of a black body at 5800°K in order to approximate solar radiation. Since radiation from the sun or a fusion reaction may be rich in ultraviolet, a calculation is also made for a black body radiation of 10,000°K to indicate the effect of increased

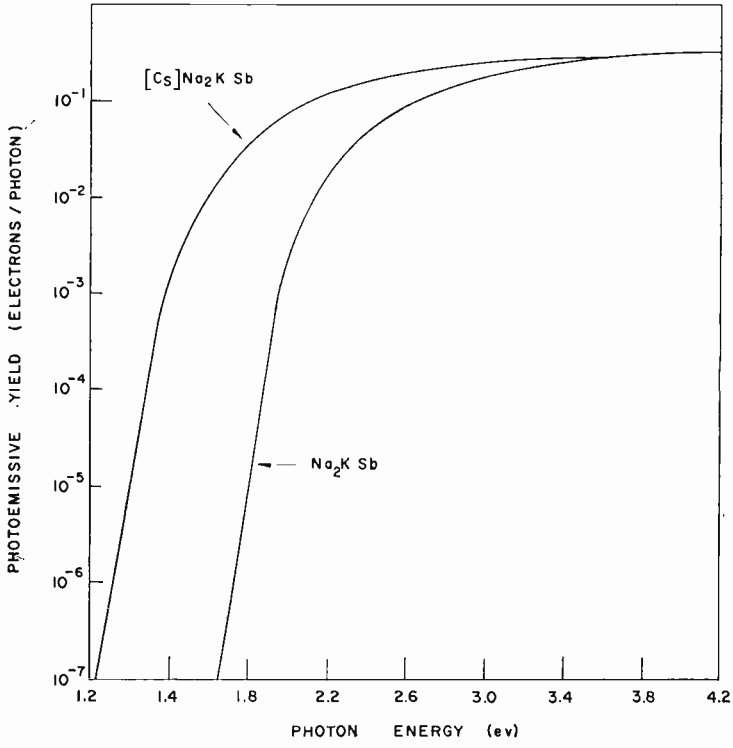


Fig. 2—The spectral distribution of photoelectric quantum yield for typical [Cs]Na_2KSb and Na_2KSb surfaces.

ultra-violet radiation. The spectral distributions of the photoelectric quantum yield used in this work are shown in Figure 2. These yield curves were taken from typical photocathodes rather than from the best which have been produced.[3] To obtain the saturated photoelectric current, I_0, a numerical integration was made using the appropriate spectral distribution curves.

[3] W. E. Spicer, "Photoemissive, Photoconductive, and Optical Absorption Studies of Alkali-Antimony Compounds," *Phys. Rev.*, Vol. 112, p. 114, Oct. 1, 1958.

To obtain values for the CPD's, it is necessary to have values for the work functions of the materials involved. For Ag-O-Cs, the work function of 0.8 ev determined by Davey[2] is used.[4] No experimental data is available on the work functions of [Cs]Na_2KSb and Na_2KSb; however, from the data on electron affinity,[3] bandgap,[3] conductivity type,[5] and impurity photoemission,[3] it is possible to make estimates of 1.2 and 1.8 ev, respectively, for the work functions of these materials.

The results obtained are given in Table I. For solar radiation (5800°K), the efficiencies are between 1.5 and 2 per cent. It is conceivable that with cathodes of maximum sensitivities and optimum work functions, sensitivities of two or three times this might be realized. The increased efficiency of the Na_2KSb surface (4.3%) for 10,000°K indicates the gain in efficiency to be expected as the ultraviolet component of the radiant spectrum is increased.

Table I—Ag-O-Cs Collector

Radiator Temperature (°K)	Emitter	Current per Watt Incident (amp)	Emitter Work Function (ev)	Collector Work Function (ev)	Contact Potential Difference (ev)	Efficiency (%)
5800	[Cs]Na_2KSb	.029	1.3	0.8	0.5	1.45
5800	Na_2KSb	.019	1.8	0.8	1.0	1.9
10,000	Na_2KSb	.043	1.8	0.8	1.0	4.3

Effect of the Initial Velocities of the Photoelectrons

The results given in Table I were calculated on the basis of an output voltage equal to the CPD between collector and emitter. However, an electron will be collected as long as the initial kinetic energy is greater than the difference between the output voltage and the CPD. The reason for this is indicated by Figure 3. If V_0 is greater than V_c, a photoelectron will experience a retarding voltage,

[4] Since Richardson plot slopes corresponding to activation energies as low as 0.5 ev have been observed (W. E. Spicer and A. H. Sommer, unpublished results), it is possible that smaller work functions might be obtained.

[5] A. H. Sommer, "N-Type and P-Type Conduction in Alkali-Antimonide Photoemitters," *Jour. Appl. Phys.*, Vol. 29, p. 1568, Nov. 1958.

$$V_r = V_0 - V_c. \tag{2}$$

If the electron is emitted with a kinetic energy larger than V_r, however, it will still be collected at the anode. Thus, the efficiency of a photoelectric energy converter can be written

$$E = \frac{I(V_r)}{L} V_0, \tag{3}$$

where $I(V_r)$ is the current collected when a retarding potential, V_r, is applied. $I(V_r)$ may be written

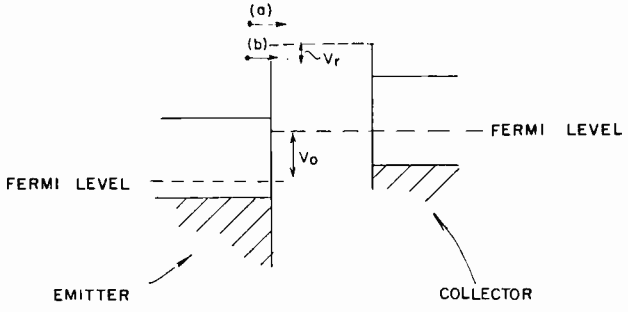


Fig. 3—Energy-level diagram for a photoemissive energy converter which is delivering an output voltage, V_o, greater than the CPD, V_c. The retarding potential, V_r, is given by $(V_o - V_c)$. (a) indicates an electron having an initial kinetic energy greater than V_r.

$$I(V_r) = I_0 F(V_r) \tag{4}$$

where $F(V_r)$ is the fraction of the saturated current collected when a retarding potential is applied. Substituting Equation (4) in Equation (3),

$$E = \frac{I_0 V_0 F(V_r)}{L}. \tag{5}$$

To determine the efficiency as a function of output voltage, therefore, it is necessary to know the function $F(V_r)$. This was determined by measuring the velocity distribution of the photoelectrons excited by solar radiation. For these measurements, cylindrical-velocity-dis-

tribution tubes similar to those described by Apker and Taft[6] were used. The results, shown in Figure 4, are as expected from the band models[3] of these materials, except that the difference between the maximum kinetic energies for the two materials is about 0.1 ev less than that expected. This discrepancy, which is not important in the calculation being made here, is probably due to a variation in the properties of the photosurfaces and/or to experimental error. From

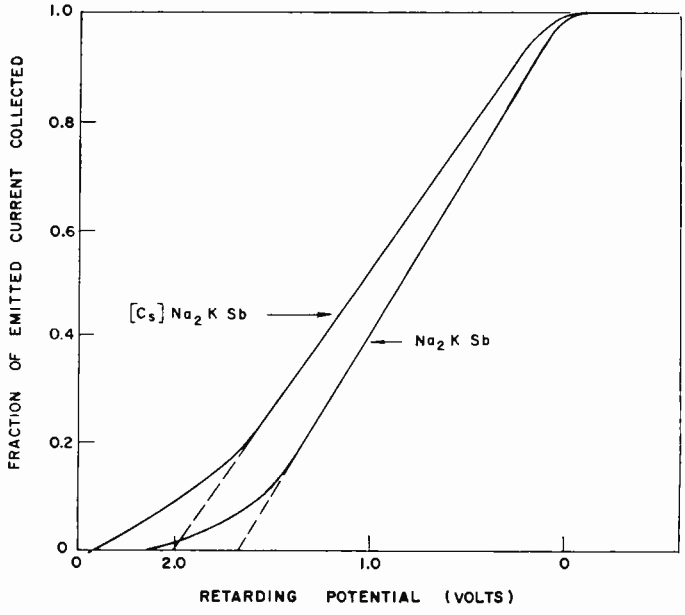


Fig. 4—Experimental retarding potential curves for [Cs]Na_2KSb and Na_2KSb for solar irradiation. The abscissa indicates the retarding potential and the ordinate indicates the fraction of the emitter electrons collected.

the data in Figure 4, the following approximations are derived for the function $F(V_r)$:

For [Cs]Na_2KSb,

$$F(V_r) = (1 - 0.5\,V_r), \text{ where } V_r \geqq 0. \tag{6}$$

For Na_2KSb,

$$F(V_r) = (1 - 0.6\,V_r), \text{ where } V_r \geqq 0. \tag{7}$$

[6] L. Apker, E. Taft and J. Dickey, "Electron Scattering and the Photoemission from Cesium Antimonide," *Jour. Opt. Soc. Amer.*, Vol. 43, p. 78, Feb. 1953.

If V_r is negative, $F(V_r)$ is unity. These approximations were made by fitting a straight line (indicated by the dashed line in Figure 4) to the curves in Figure 4. Using Equations (5), (6), and (7) and the results presented in Table I, the results presented in Figure 5 were obtained.* Two points are obvious from this figure. First, the maximum efficiency is obtained for output voltages larger than the contact difference in potential. Second, the peak of the efficiency curve is relatively broad. For [Cs]Na_2KSb, it extends over almost a volt.

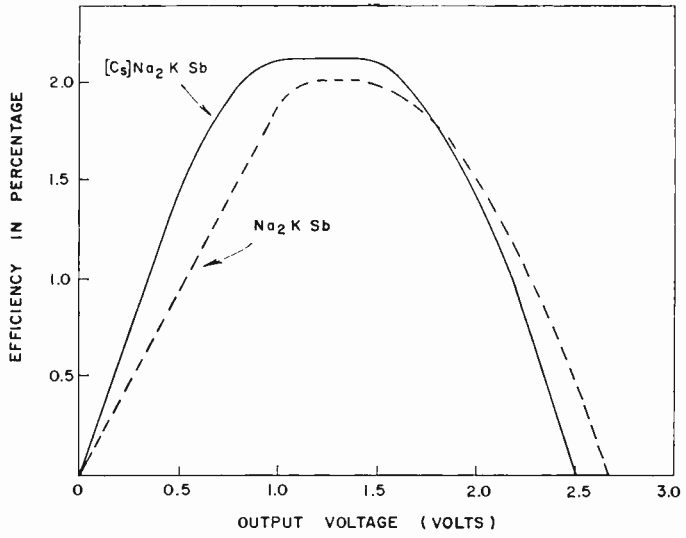


Fig. 5—Calculated conversion efficiency for solar radiation. The calculation has been made for devices with Ag-O-Cs collectors and either [Cs]Na_2KSb or Na_2KSb emitters. The effect of initial velocities has been taken into account. The efficiency is plotted as a function of output voltage.

This last characteristic should allow considerable versatility in the operation of the device.

No measurements have been made on the velocity distributions of the photoelectrons for 10,000°K radiation; however, from the velocity distributions obtained for monochromatic radiation,[7] it is apparent that the average initial velocity of the photoelectrons would be in-

* Since [Cs]Na_2KSb and Na_2KSb have identical band structures except for the reduced electron affinity of [Cs]Na_2KSb,[2] the efficiency from the former material should be greater for all output voltages. This is found to be the case except at the high output voltages. The discrepancy is due to the apparent error in the velocity distribution curves mentioned earlier and is not of importance for these considerations.

[7] W. E. Spicer, unpublished results.

creased by at least one or two electron volts with a resulting maximum efficiency approaching 10 per cent.†

One practical difficulty which might arise in the production of the converter discussed above is that of producing alkali-antimony and Ag-O-Cs cathodes in close proximity. If it proved impossible or impractical to do this, a converter could be made in which two cathodes of the same type faced each other. Here, the development of output voltage would depend solely on the initial velocities of the emitter electrons. A calculation has been made for such a system using [Cs]Na_2KSb surfaces for the emitter and collector. To do this, the

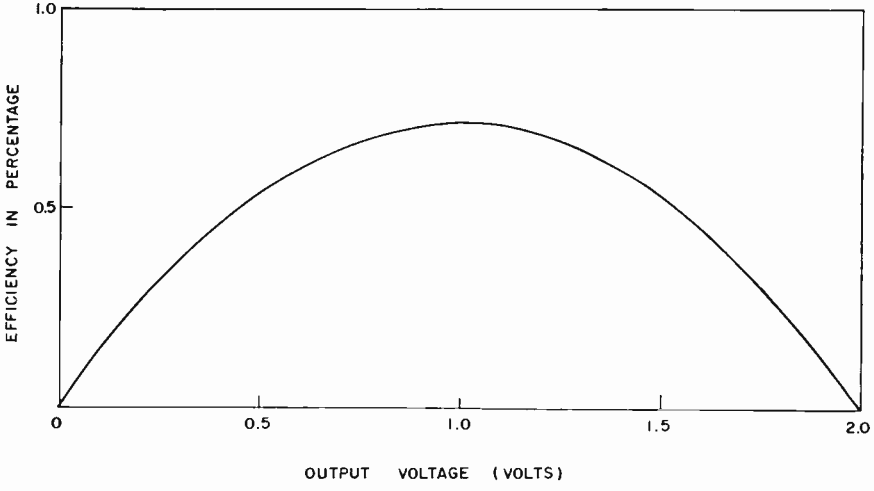


Fig. 6—Calculated conversion efficiency for solar radiation. The calculation has been made assuming the use of [Cs]Na_2KSb surfaces as both collector and emitter. The efficiency is plotted as a function of output voltage.

velocity distribution given by Equation (6) was used. The efficiency versus output voltage obtained is shown in Figure 6. The peak efficiency obtained was only 0.7 per cent. This is a drop of about a factor of three from the peak efficiency calculated using a Ag-O-Cs collector.

Space-Charge Effects

The emitted electrons form a negative space-charge cloud between emitter and collector. This retards the flow of electrons and thus tends

† Since sufficiently energetic electrons lose energy very quickly due to pair production, the effective maximum energy of the emitted electrons will be limited to about twice the bandgap. (See References 6 and 8.)

to reduce the efficiency below those calculated previously. In order to obtain an idea of the effect of space charge, the following modification of Child's Law for plane parallel geometry will be used:

$$J = \frac{2.3 \times 10^{-6}(V_c + V_n - V_0)^{3/2}}{d^2}. \tag{8}$$

Here, J is the maximum current density which can be collected (amps/cm^2), and d is the separation between emitter and collector in cm. In an attempt to take account of the initial kinetic energy of the photoelectrons in a simple way, all of the electrons have been assumed to have a single normal component of kinetic energy V_n. This component is treated in Equation (8) as if it were equivalent to an accelerating voltage. Although these approximations are relatively crude, they make it possible to obtain some feel for the seriousness of the space-charge problem and, in particular, to obtain some estimate of the dimension necessary to overcome it.

As can be seen from Equation (8), it is never possible to obtain the maximum output voltage and current. The actual current and voltage which can be realized are dependent on the emitter–collector separation, d. To obtain an estimate of the spacing needed for the multi-alkali–Ag-O-Cs system, the following values have been assumed: $V_c + V_n = 2$ volts, $V_0 = 1$ volt, and $J = 4$ ma/cm^2. The last number is obtained from the data in Table I and the value of 0.14 watts/cm^2 for the solar-energy flux. Placing these parameters in Equation (8), a spacing of 0.025 cm is found to be necessary to obtain complete collection. If the spacing were reduced to 0.01 cm, complete collection could be obtained with an output voltage of 1.7 volts. These rough calculations simply indicate that spacings of the order of 0.01 cm are necessary for photoemissive energy converters if space-charge compensation is not to be used.

It might be possible to neutralize the space charge through the use of a cesium ion vapor. However, the work of Sommer[5] indicates that a partial pressure of cesium may reduce the sensitivity of alkali-antimony cathodes.

Technological Difficulties

Two of the most important technological difficulties have already been mentioned: the necessity for close spacing and the difficulty in producing two different types of photocathodes in close proximity. Other problems which have not been touched upon are those of the cathode and cathode-interface resistance and that of drawing high

SEMICONDUCTING MATERIALS FOR THERMOELECTRIC POWER GENERATION*†

By

F. D. Rosi, E. F. Hockings, and N. E. Lindenblad

RCA Laboratories
Princeton, N. J.

Summary—*A general consideration of the thermoelectric properties of semiconductors suggests that (1) this class of materials can be useful in power-generating thermocouples operating at least up to 700°C, and (2) use of a sandwich-type arrangement or graded alloying in the construction of thermocouple branches will be necessary to achieve high figures of merit over a wide temperature range and, hence, high power-generating efficiencies. A large number of ternary compound semiconductors having the cubic structure were synthesized. Those with the rock-salt structure, such as $AgSbTe_2$, are characterized by low lattice thermal conductivities (<0.0075 watt cm^{-1} deg^{-1}). The lattice thermal conductivity as a function of composition was examined in the alloy systems of $AgSbTe_2$ with PbTe, SnTe, and GeTe. The minimum in the lattice thermal conductivity for the $AgSbTe_2$–PbTe system gives an effective mean free path for phonons which is less than unit-cell dimensions. Measurements of the temperature dependence of thermoelectric properties of a number of solid-solution alloy systems showed that (1) solid-solution alloys of Bi_2Te_3 with Bi_2Se_3, Sb_2Te_3, and Sb_2Se_3 provided the best p- and n-type material for thermocouple operation in the temperature range 25 to 250°C, (2) the ternary compound $AgSbTe_2$ and its alloys with GeTe provided the best p-type material for the range 250 to 550°C, and (3) alloys in the PbTe-SnTe system provided the best n-type material for the range 250 to 550°C. Power-generating thermocouples, constructed in the sandwich-type arrangement of materials, provided an efficiency of ~12 per cent for operation over the temperature range 20 to 550°C (i.e., $T_H - T_C = 530$°C). Continuous thermocouple operation for 300 hours resulted in no significant deterioration of material properties.*

Introduction

IT HAS BECOME APPARENT in recent years that the application of thermoelectricity to power generation is a problem which deserves the serious attention of the various scientific disciplines in solid-state research. With the advent of the transistor in 1948 and the subsequent rapid development of semiconductor technology, materials became available having a combination of thermoelectric properties which indicated the feasibility of power-generating thermo-

* Manuscript received 22 February 1961.

† Part of the work described in this paper was supported by the U.S. Navy, Bureau of Ships.

couples with much higher efficiencies than was possible with metals and metal alloys. This realization brought about considerable research activity, both in this country and abroad, on the development of new semiconductors and of certain classes of insulators and semimetals whose properties are strongly temperature dependent. Since material operation over a wide temperature range and at high temperatures is necessary for high power-generating efficiencies, the need for a broad materials program was recognized from the beginning. This application represents a much more formidable problem than that of thermoelectric refrigeration, where operating temperature ranges are much smaller and the most suitable material classes are already well defined. Although considerable data have been published in recent years on both the materials and device aspects of thermoelectric power generation, this progress represents only the early stages of technological development. A strong need still exists for (1) additional exploratory research for new and better materials, (2) a better understanding of transport phenomena in insulators, semiconductors, and semimetals at high temperatures, and (3) a demonstration of the feasibility of power-generating efficiencies competitive with conventional converters.

The present work is concerned specifically with research on semiconductor materials for thermoelectric power generation in the range 25 to 700°C. It is presently believed that semiconductors are the most promising materials for thermocouple operation up to 700°C; this temperature range coincides with potential heat sources such as focused solar and nuclear energy. The research described here is presented in three parts: (1) general considerations of efficiency factors leading to materials selection, (2) evaluation of new compound semiconductors and their solid-solution alloys, and (3) fabrication of thermocouples to demonstrate power-generating efficiencies.

General Considerations

The application of thermoelectricity to power generation dates more than a century back to its discovery by Seebeck in 1822.[1] He showed that a current is obtained when the junctions of dissimilar materials forming a thermocouple loop are maintained at different temperatures. Such a circuit is shown schematically in Figure 1. In it, two dissimilar materials, an n- and a p-type semiconductor, are joined at their ends by a metallic conductor. Heat is supplied to the

[1] T. J. Seebeck, *Reports of the Prussian Acad. Sci.* (1822-1823).

hot junction, T_H, from an external source, and the other junction is maintained at a fixed lower temperature, T_C. As a result of the temperature difference, $T_H - T_C$, a current I flows through the branches of the thermocouple in the direction indicated by the arrows. By also allowing the current to flow through an external load resistance, R_L, inserted into the circuit between the cold junctions, the arrangement in Figure 1 represents a direct conversion of heat into electrical energy with a conversion efficiency, ϕ, given by

$$\phi = \frac{\text{power supplied to load}}{\text{heat absorbed at hot junction}}. \tag{1}$$

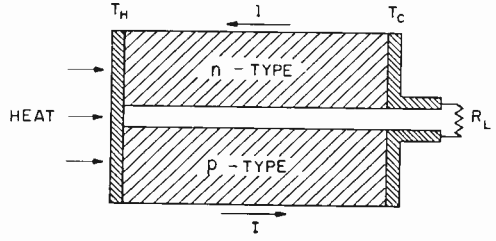


Fig. 1—Thermocouple as power generator.

The useful power is simply I^2R_L; and, if the Thomson heat is neglected, the heat absorbed at T_H consists of the Peltier heat, H_1, and the heat withdrawn from the hot junction by conduction, H_2;

$$H_1 = QIT_H \tag{2}$$

and

$$H_2 = K(T_H - T_C) - \frac{1}{2} I^2R. \tag{3}$$

Here, Q is the thermoelectric power or Seebeck coefficient of the junction, K is the thermal conductance of the two branches in parallel, and R is the sum of the branch resistances. For the value of H_2, it is necessary to subtract one-half of the Joule heat, as this quantity is returned to the hot junction. The efficiency is then given by

$$\phi = \frac{I^2R_L}{QIT_H + K(T_H - T_C) - \frac{1}{2} I^2R}, \tag{4}$$

and under optimum current-load conditions, the expression for maximum efficiency is

$$\phi_{\max} = \frac{T_H - T_C}{T_H} \cdot \frac{\sqrt{1 + z\overline{T}} - 1}{\sqrt{1 + z\overline{T}} + \frac{T_C}{T_H}}. \tag{5}$$

Here,

$$\overline{T} = \frac{T_H + T_C}{2},$$

and

$$z = \frac{Q^2}{RK}. \tag{6}$$

The quantity z, defined by Equation (6), is commonly referred to as the figure of merit of the thermocouple.[2]

The first factor in the expression for maximum efficiency in Equation (5) is simply the thermodynamic efficiency of a reversible Carnot cycle. The second factor represents the decrease in this efficiency resulting from irreversible heat conduction along the branches and power dissipation in the form of Joulean heat. This factor depends only on the average temperature $\overline{T}$ and the quantity z. If, furthermore, the resistances, R_a and R_b, of the branches are matched to give z a maximum value, i.e., if

$$\frac{R_a}{R_b} = \frac{\rho_a \kappa_a}{\rho_b \kappa_b} \tag{7}$$

where ρ_a, ρ_b and κ_a, κ_b are the resistivities and thermal conductivities, respectively, of the two branches, then the figure of merit becomes a function only of the material constants of the two branches;

$$z = \frac{(Q_a - Q_b)^2}{(\sqrt{\rho_a \kappa_a} + \sqrt{\rho_b \kappa_b})^2}. \tag{8}$$

Here Q_a and Q_b are the thermoelectric powers of the two branches referred to, e.g., lead. If

---

[2] A. F. Ioffe, *Poluprovodnikovye Termoelementy*, Moscow-Leningrad (1956).

$$Q_a = -Q_b \quad \text{and} \quad \rho_a \kappa_a = \rho_b \kappa_b,$$

then

$$z = z_a = \frac{Q_a^2}{\rho_a \kappa_a}. \tag{9}$$

Here, the figure of merit is conveniently defined in terms of the material constants of a single branch of the thermocouple. It is apparent from Equation (8) that the properties which are of primary importance in the evaluation of materials for thermoelectric power

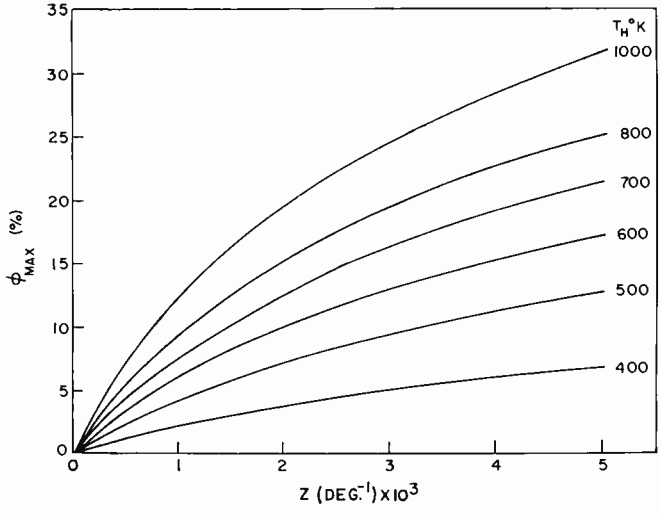


Fig. 2—Maximum coefficient of performance ϕ_{max} as a function of figure of merit z for different hot-junction temperatures T_H ($T_C = 300°K$).

generation are the same as those for thermoelectric refrigeration. Moreover, it may be seen from Equation (5) that for optimum power-generating efficiency, it is desirable to maximize the factor, $z\overline{T}$, i.e., to obtain as high a value of z over as wide a temperature range as possible. This is demonstrated in Figure 2, which shows the variation in ϕ_{max} with z for various hot junction temperatures and a fixed T_C of 300°K. It is assumed here that the resistivity, thermal conductivity, and thermoelectric power are constant within a branch over the temperature range considered and, consequently, that the Thomson coefficient can be regarded as zero.

Since this paper is concerned with research on semiconductor materials, it would be desirable to describe the figure of merit, as

defined by Equation (9), in terms of semiconductor material constants. Let us consider the usual case of an extrinsic nondegenerate semiconductor in which a single sign of charge carrier, n, predominates. From semiconductor transport theory, the electrical resistivity ρ decreases with n according to the simple expression,

$$\frac{1}{\rho} = ne\mu, \tag{10}$$

where e is the electronic charge and μ is the mobility of the charge carriers.

The thermoelectric power Q of a semiconductor represents the emf opposing the diffusion of charge carriers from the hot to the cold junction of the thermocouple. The application of the Lorentz–Sommerfeld theory gives the expression,

$$Q = \frac{\epsilon_F - \epsilon_b - \Delta\epsilon_T}{eT}, \tag{11}$$

where ϵ_F is the Fermi energy, ϵ_b is the energy of the band edge (valence or conduction), and $\Delta\epsilon_T$ is a kinetic energy term representing the average energy of the transported electrons relative to the band edge. The potential energy, $\epsilon_F - \epsilon_b$, is the same term which appears in the exponential part of the expression for the carrier concentration n; i.e.,

$$n = \left(\frac{2(2\pi m^{(N)} kT)^{3/2}}{h^3} \right) \exp \left(\frac{\epsilon_b - \epsilon_F}{kT} \right). \tag{12}$$

Here, $m^{(N)}$ is defined as the density-of-states effective mass representing the product $N_V{}^{2/3} m^*$, where N_V is the number of valleys and m^* is the effective mass and is equal to the geometric mean of the principal masses of the valleys. Expressing $(\epsilon_F - \epsilon_b)$ in terms of n and $m^{(N)}$ in Equation (12),

$$\frac{\epsilon_F - \epsilon_b}{kT} = \ln \left(\frac{2(2\pi m^{(N)} kT)^{3/2}}{h^3 n} \right). \tag{13}$$

Then by combining Equations (11) and (13), the final expression for the thermoelectric power is given by

ing by phonons. Figure 3 shows the variation in Q^2/ρ with number of carriers for different effective masses, but a constant carrier mobility of 500 $cm^2volt^{-1}sec^{-1}$; while Figure 4 shows the variation in Q^2/ρ with n for different values of μ, but a constant $m^{(N)}$ of 0.7 electron-mass units. The maxima in the curves of Figure 4 clearly show the need for carrier concentrations of $\sim 10^{19}$ cm^{-3}, and the shift in the maxima of the curves of Figure 3 to values of $n > 10^{19}$ cm^{-3} simply reflects the increase in the density of carriers corresponding to the onset of degeneracy with an increase in effective mass.

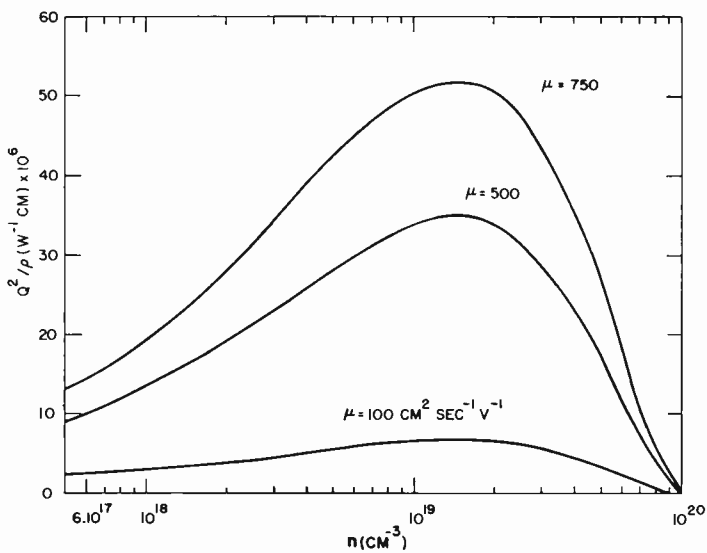


Fig. 4—Variation of Q^2/ρ with the number of carriers n for different carrier mobilities μ.

From Equation (17), it is reasonable to expect some improvement in the figure of merit of a material at higher temperatures. Since the mobility μ can be expected to decrease with temperature as $T^{-3/2}$, then the factor $\mu T^{3/2}$ in Equation (17) would be constant. Thus, the variation in z with temperature would then depend on the value of κ_{ph}, which usually has a T^{-1} dependence.

For optimum power-generating efficiency, it is desirable to obtain a high value of z over as wide a temperature range as possible (i.e., to maximize the factor $z\overline{T}$ in Equation (5)). Thus, the semiconductor material must have a large energy bandgap and a high melting point. A large bandgap is desirable, because it permits operation of the thermocouple at high hot-junction temperatures without a prohibitive

loss in thermoelectric properties. This follows from the knowledge that an increase in the bandgap of a semiconductor shifts the onset of intrinsic conduction due to thermal generation of electron–hole pairs to higher temperatures. Such electron–hole pair generation is to be avoided, because it results not only in a marked decrease in thermoelectric power, but also an increase in the thermal conductivity due to an ambipolar contribution, κ_{amb}, resulting from the diffusion of electron–hole pairs from T_H to T_C.[4]

The need for a high melting point and corresponding chemical stability is obvious. Unfortunately, it is a general rule[5] that the higher the melting point, T_m, of a material, the larger is its Debye characteristic temperature, θ, and hence κ_{ph} at a given temperature This may be seen from the expression,

$$\kappa_{ph} = \frac{\theta^2}{\gamma^2}\left[MV^{1/3}\left(\frac{\theta}{T}\right)\right]C, \qquad (19)$$

recently derived by Leibfried and Schlömann.[6] Here, M and V are the atomic weight and volume of the material, respectively, $\gamma = \partial \ln \theta / \partial \ln V$, and C is a numerical constant. Equation (19) holds for the case where phonon–phonon interactions are the main source of thermal resistance, which is the usual situation at elevated temperatures. Since θ is proportional to (T_m/M), it follows from Equation (19) that $\kappa_{ph} \propto (T_m{}^3/M)^{1/2}$. This dependence could place an upper temperature limit to the usefulness of materials which exhibit semiconductor behavior at room temperature. From available data on κ_{ph} in semiconductors and the degree to which κ_{ph} can be decreased by solid-solution alloying, it is concluded that semiconductors would be the most useful materials at least up to 700°C.

Materials Research

In research on compound semiconductors for thermoelectric power generation, it became apparent that no one material could provide the figure of merit, or $z\overline{T}$ factor, necessary for high efficiencies over the temperature range of immediate interest, 25° to 700°C. Consequently,

---

[4] P. J. Price, "Ambipolar Thermodiffusion of Electrons and Holes in Semiconductors," *Phil. Mag.*, Vol. 46, Ser. 7, p. 1252, November, 1955.

[5] J. K. Roberts and A. R. Miller, *Heat and Thermodynamics*, New York, N. Y., p. 552 (1954).

[6] G. Leibfried and E. Schlömann, "Thermal Conductivity in Electrically Insulating Crystals," *Gött. Nachr.*, Vol. IIa, p. 71 (1954).

attention was directed to a consideration of constructing thermocouple branches (1) in a sandwich-type arrangement and (2) by graded alloying. Both these techniques make use of a heterogenous distribution of materials from the hot to the cold junctions. In the sandwich-type arrangement, materials are selected to provide a high average figure of merit over relatively narrow temperature ranges. This principle is schematically illustrated by the diagrams in Figure 5, which show the temperature dependence of z for the materials, A, B and C, and their arrangement along the branch of a thermocouple.

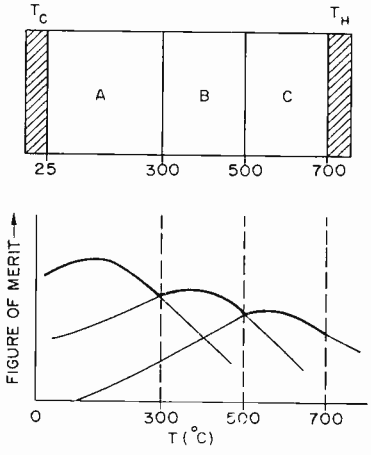


Fig. 5—Principle of sandwich arrangement of materials A, B, and C along branch of power-generating thermocouple.

It is apparent, from the temperature dependence of z, that material A provides the best average z value in the range 25 to 300°C; B provides the best average z in the range 300 to 500°C; and so on. Thus, by placing these materials in the series arrangement shown in the upper diagram of Figure 5, an envelope of z versus T curves (indicated by the heavy lines) is obtained which represents the best average z value over the entire operating temperature range $T_C \rightarrow T_H$ of the thermocouple.

The construction of thermocouple branches by the graded alloy technique is illustrated by the diagrams in Figure 6. In this principle, use is made of an alloy system, A-B, which exhibits complete solid miscibility (Figure 6b), and in which the bandgap E_g of A is less than that of B and has a general dependence on alloying as shown in Figure 6c. In addition, the dependence of κ_{ph} on alloying exhibits a broad minimum, as shown in Figure 6d. On the assumption that

loss in thermoelectric properties. This follows from the knowledge that an increase in the bandgap of a semiconductor shifts the onset of intrinsic conduction due to thermal generation of electron–hole pairs to higher temperatures. Such electron–hole pair generation is to be avoided, because it results not only in a marked decrease in thermoelectric power, but also an increase in the thermal conductivity due to an ambipolar contribution, κ_{amb}, resulting from the diffusion of electron–hole pairs from T_H to T_C.[4]

The need for a high melting point and corresponding chemical stability is obvious. Unfortunately, it is a general rule[5] that the higher the melting point, T_m, of a material, the larger is its Debye characteristic temperature, θ, and hence κ_{ph} at a given temperature This may be seen from the expression,

$$\kappa_{ph} = \frac{\theta^2}{\gamma^2}\left[MV^{1/3}\left(\frac{\theta}{T}\right)\right]C, \tag{19}$$

recently derived by Leibfried and Schlömann.[6] Here, M and V are the atomic weight and volume of the material, respectively, $\gamma = \partial \ln \theta / \partial \ln V$, and C is a numerical constant. Equation (19) holds for the case where phonon–phonon interactions are the main source of thermal resistance, which is the usual situation at elevated temperatures. Since θ is proportional to (T_m/M), it follows from Equation (19) that $\kappa_{ph} \propto (T_m{}^3/M)^{1/2}$. This dependence could place an upper temperature limit to the usefulness of materials which exhibit semiconductor behavior at room temperature. From available data on κ_{ph} in semiconductors and the degree to which κ_{ph} can be decreased by solid-solution alloying, it is concluded that semiconductors would be the most useful materials at least up to 700°C.

Materials Research

In research on compound semiconductors for thermoelectric power generation, it became apparent that no one material could provide the figure of merit, or $z\overline{T}$ factor, necessary for high efficiencies over the temperature range of immediate interest, 25° to 700°C. Consequently,

---

[4] P. J. Price, "Ambipolar Thermodiffusion of Electrons and Holes in Semiconductors," *Phil. Mag.*, Vol. 46, Ser. 7, p. 1252, November, 1955.

[5] J. K. Roberts and A. R. Miller, *Heat and Thermodynamics*, New York, N. Y., p. 552 (1954).

[6] G. Leibfried and E. Schlömann. "Thermal Conductivity in Electrically Insulating Crystals," *Gött. Nachr.*, Vol. IIa, p. 71 (1954).

attention was directed to a consideration of constructing thermocouple branches (1) in a sandwich-type arrangement and (2) by graded alloying. Both these techniques make use of a heterogenous distribution of materials from the hot to the cold junctions. In the sandwich-type arrangement, materials are selected to provide a high average figure of merit over relatively narrow temperature ranges. This principle is schematically illustrated by the diagrams in Figure 5, which show the temperature dependence of z for the materials, A, B and C, and their arrangement along the branch of a thermocouple.

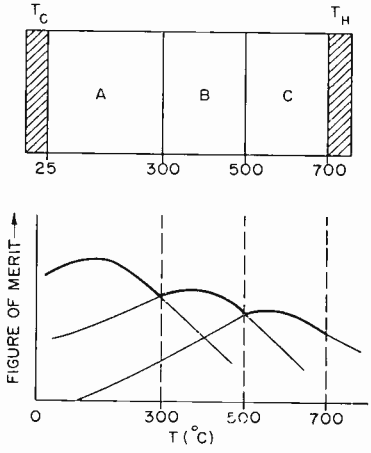


Fig. 5—Principle of sandwich arrangement of materials A, B, and C along branch of power-generating thermocouple.

It is apparent, from the temperature dependence of z, that material A provides the best average z value in the range 25 to 300°C; B provides the best average z in the range 300 to 500°C; and so on. Thus, by placing these materials in the series arrangement shown in the upper diagram of Figure 5, an envelope of z versus T curves (indicated by the heavy lines) is obtained which represents the best average z value over the entire operating temperature range $T_C \rightarrow T_H$ of the thermocouple.

The construction of thermocouple branches by the graded alloy technique is illustrated by the diagrams in Figure 6. In this principle, use is made of an alloy system, A-B, which exhibits complete solid miscibility (Figure 6b), and in which the bandgap E_g of A is less than that of B and has a general dependence on alloying as shown in Figure 6c. In addition, the dependence of κ_{ph} on alloying exhibits a broad minimum, as shown in Figure 6d. On the assumption that

the A-rich alloys with lower E_g provide higher values of z at the temperature of the cold junction, T_C, and, conversely, the B-rich alloys provide a higher z at T_H, it would be advantageous to adjust the bandgap along the thermocouple branches by a uniform variation in composition in the manner shown in the upper diagram of Figure 6. The terminal alloy compositions, 80% A–20% B and 20% A–80% B,

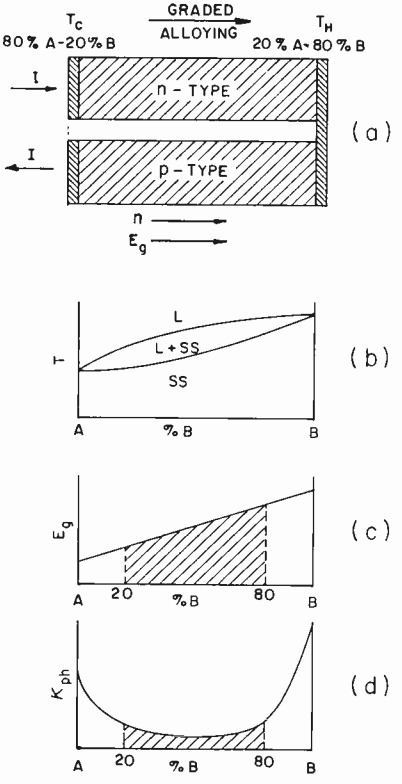


Fig. 6—Graded alloying of thermocouple branches with semiconductor alloy system, *A-B*. (a) Power-generating thermocouple with graded alloy structure; (b) phase diagram showing complete solid solubility, *SS;* (c) variation of bandgap E_g with alloying; (d) variation of lattice thermal conductivity κ_{ph} with alloying.

should be selected so as to take advantage of the decrease in κ_{ph} due to alloying, as indicated by the hashed area in Figure 6d. It also follows that the carrier concentration n must be so adjusted from T_C to T_H as to provide a thermoelectric power Q given by Equation 18 at the various temperatures along the branches. This would require an increase in n on going from T_C to T_H, since (as is apparent from

Equation (14)) for a given effective mass of current carriers, the value of Q is proportional to $T^{3/2}/n$. The graded alloy and impurity structure in Figure 6 could provide (1) higher power-generation efficiencies through operation at a larger $T_H - T_C$ for the same average value of z, or (2) operation at the same $T_H - T_C$ with a higher average value of z.

With the above principles of thermocouple construction in mind, data largely taken from the recent work of Rosi, Dismukes, and Hockings[7] are presented on the evaluation of compound semiconductors and alloys for power generation over specific temperature ranges. Emphasis is placed on the utilization of the sandwich-type arrangement shown in Figure 5. It should also be mentioned here that the compounds and their solid-solution alloys were prepared in ingot form by crystallization from the melt using either the Bridgman[8] or zone-levelling[9] technique, the details of which can be found in the literature.[7,10] In addition, the data on the thermal and electrical properties were obtained in the following manner: The electrical resistivity, ρ, was determined by an a-c scanning method; the thermoelectric power, Q, and thermal conductivity, κ, were measured simultaneously at all temperatures by the standard technique of establishing a small temperature gradient across the specimen placed between a heater and sink. The carrier concentration, n, and mobility, μ, were obtained from measurements of Hall constant. In all cases the lattice contribution, κ_{ph}, to the thermal conductivity was obtained from the measured total thermal conductivity, κ, from Equation (15).

Bi_2Te_3 Alloys

In the course of research on materials for thermoelectric cooling,[10] it became apparent that solid-solution alloys of Bi_2Te_3 with other V-VI compound semiconductors might also be useful in power-generating thermocouples operating with hot-junction temperatures up to 300°C. Consequently, the thermal and electrical properties of selected n- and p-type alloys were examined in the temperature range 25 to 500°C. For the n-type material, emphasis was placed on the 75%

---

[7] F. D. Rosi, J. P. Dismukes and E. F. Hockings, "Semiconductor Materials for Thermoelectric Power Generation up to 700°C," *Electrical Engineering*, Vol. 79, No. 6, p. 450, June 1960.

[8] P. W. Bridgman, "Certain Physical Properties of Single Crystals of Tungsten, Antimony, Bismuth, Cadmium, Zinc and Tin," *Proc. Amer. Acad. Arts and Sci.*, Vol. 60, p. 305 (1925); Vol. 58, p. 165 (1923).

[9] W. G. Pfann, "Principles of Zone Melting," *Trans. Amer. Inst. Min. Met. Engrs.*, Vol. 194, p. 747, July 1952.

[10] F. D. Rosi, B. Abeles and R. V. Jensen, "Materials for Thermoelectric Refrigeration," *Jour. Phys. Chem. Solids*, Vol. 10, Nos. 2 & 3, p. 191, July 1959.

Bi_2Te_3–25% Bi_2Se_3 alloy doped with either CuBr or AgI, while alloys of Bi_2Te_3–Sb_2Te_3 with excess Te or Se were examined for p-type material. Examples of these alloys with their thermoelectric properties at room temperature are listed in Table I. These alloy compositions were selected because at 300°K they provide (1) a lattice thermal conductivity κ_{ph} significantly lower (< 0.010 watt cm^{-1} deg^{-1}) than that of Bi_2Te_3 (0.013 watt cm^{-1} deg^{-1}), (2) a $\mu[m^{(N)}/m]^{3/2}$ product similar to Bi_2Te_3 as suggested by Q^2/ρ values, and (3) an energy bandgap greater than Bi_2Te_3 as indicated by the variation

Table I—Thermoelectric Properties of Bi_2Te_3 Alloys at Room Temperature

Alloy Composition	Conductivity Type	Impurity Added	Q (μv deg^{-1})	ρ (ohm-cm)	κ (watt cm^{-1} deg^{-1})	κ_{ph} (watt cm^{-1} deg^{-1})
% Bi_2Te_3–25% Bi_2Se_3	n	CuBr	166	10.3×10^{-4}	0.0133	0.0090
% Bi_2Te_3–25% Bi_2Se_3	n	CuBr	125	5.6×10^{-4}	0.0221	0.0091
% Bi_2Te_3–70% Sb_2Te_3 (2 wt. % Te)	p	——	195	9.2×10^{-4}	0.0148	0.0100
% Bi_2Te_3–75% Sb_2Te_3 (2.3 wt. % Te)	p	——	196	7.8×10^{-4}	0.0149	0.0093
% Bi_2Te_3–75% Sb_2Te_3 (1.75 wt. % Se)	p	——	210	9.8×10^{-4}	0.0127	0.0082
% Bi_2Te_3–80% Sb_2Te_3 (6 wt. % Te)	p	——	174	6.5×10^{-4}	0.0164	0.0096

of the thermoelectric power with electrical resistivity. These requirements are consistent with the general considerations discussed in the previous section.

The results of the measurements of thermoelectric power, Q, and electrical resistivity, ρ, at various temperatures up to 550°C for the n- and p-type alloys in Table I are summarized in Figures 7-11. In all cases, both Q and ρ initially increase with increasing temperature. The rise in Q is to be expected from Equation (14), and the occurrence of a maximum followed by a rapid decrease at the higher temperatures is associated with the onset of intrinsic conductivity. The appearance of a maximum in ρ is a direct result of the nature of the temperature dependence of the carrier mobility, μ, and carrier concentration, n. The initial sharp increase in ρ is due to the decrease in

carrier mobility with increasing temperature, while the maximum and subsequent decrease in ρ reflect the exponential increase in n with temperature. It is further noteworthy that the maximum in Q always occurs at a temperature 100 to 200°C lower than that for ρ. This could be an indication that Q is a more sensitive parameter of the onset of intrinsic conductivity in these degenerate, or near-degenerate materials.

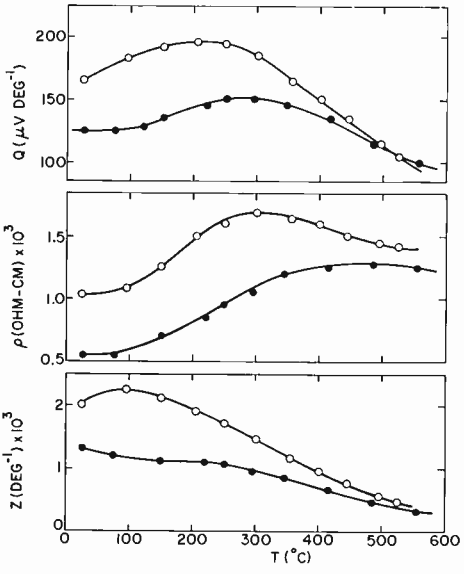


Fig. 7—Temperature dependence of thermoelectric properties in n-type 75% Bi_2Te_3–25% Bi_2Se_3 alloy.

In the case of the two n-type 75% Bi_2Te_3–25% Bi_2Se_3 alloys (see Figure 7), the observed maximum in the thermoelectric power and resistivity is broader, occurs at a lower value, and is shifted to higher temperatures for the more heavily doped degenerate alloy ($\rho = 0.55 \times 10^{-3}$ ohm-cm and $Q = 125$ μv deg^{-1} at 25°C). The shift to higher temperatures is to be expected, since the maximum in the temperature dependence of Q is usually associated with the onset of intrinsic conduction, which would occur at a higher temperature in the more heavily doped sample. The temperature dependence of z in the n-type materials of Figure 7 was determined from the expression, $z = Q^2/(\rho\kappa)$. Here, the total thermal conductivity, κ, at any given temperature was estimated from Equation (15) by calculating the electronic contribution, κ_{el}, from the measured value of ρ, and assum-

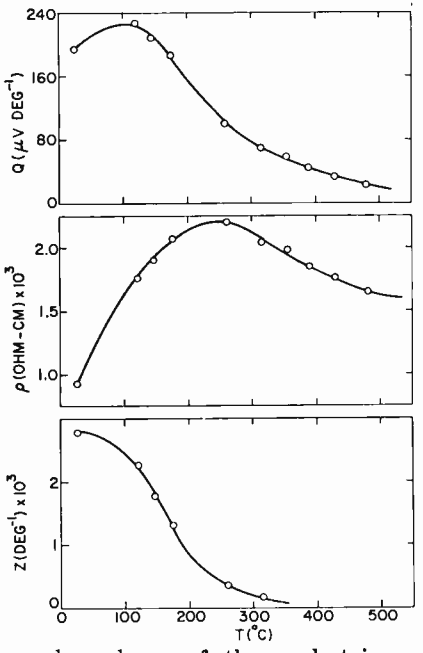


Fig. 8—Temperature dependence of thermoelectric properties in p-type 30% Bi_2Te_3–70% Sb_2Te_3 (2% Te) alloy.

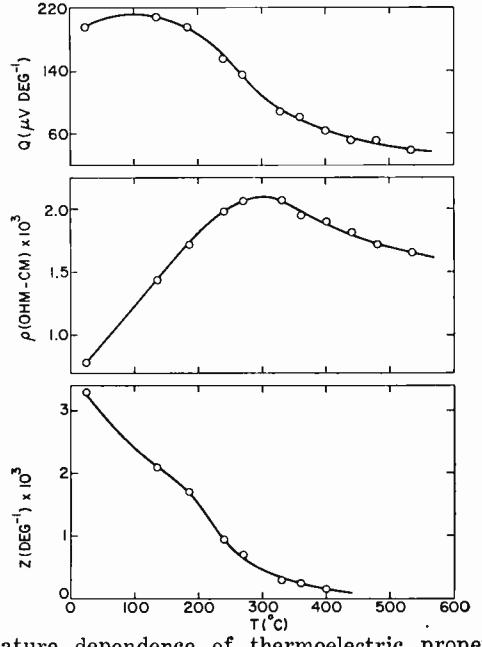


Fig. 9—Temperature dependence of thermoelectric properties in p-type 25% Bi_2Te_3–75% Sb_2Te_3 (2.3% Te) alloy.

ing the room temperature value of the lattice contribution, κ_{ph} (0.009 watt cm^{-1} deg^{-1}), to be invariant with temperature. The latter assumption represents a conservative approach toward estimating the temperature dependence of z, since in general κ_{ph} varies as T^{-1} in crystalline solids. It may be seen from Figure 7 that average z values of approximately 1.9×10^{-3} deg^{-1} and 1.2×10^{-3} deg^{-1} are obtained over the temperature range 25 to 300°C for the nondegenerate and degenerate alloys, respectively. These data indicate the potential usefulness of the n-type 75% Bi_2Te_3–25% Bi_2Se_3 alloy as a branch material for thermocouple operation up to 300°C.

An analysis of the data in Figures 8 and 11 for the p-type Sb_2Te_3–Bi_2Te_3 alloys reveal a significant shift in the maximum for ρ from 250°C to 425°C as the Sb_2Te_3 content is increased from 70 to 80 per cent. By contrast, a much smaller effect was obtained on the position of the maximum for Q, which would suggest only a slight increase in the energy bandgap in this composition range. It is also interesting to compare the data for the p-type material in Figures 8 and 9. At room temperature, both alloys have the same value of Q and, hence, similar density of carriers. The 75% Sb_2Te_3–25% Bi_2Te_3 alloy (Figure 9) has a lower resistivity, however, which would necessarily indicate a higher hole mobility. An analysis of the competing variations in mobility and carrier concentration with temperature suggests that this higher mobility should be manifested by a shift to a somewhat higher temperature in the maximum for ρ. This is indicated by a comparison of the curves in Figures 8 and 9.

It should be noted that the excess tellurium in solid solution in these alloys represents a convenient method of controlling the resistivity at room temperature. It was found that, for a given alloy, an increase in the tellurium content results in a corresponding increase in ρ, and that the amount of tellurium required for a given ρ increases with the content of Sb_2Te_3 in the alloy. The same general behavior also applies to additions of selenium in solid solution, as in the case of the p-type 75% Sb_2Te_3–25% Bi_2Te_3 alloy in Figure 10. A comparison of the Q and ρ data for this alloy with the same alloy containing 2 per cent excess Te (see Figure 8) shows a similar dependence of both ρ and Q on temperature, with the exception that the maximum in Q occurs at a somewhat higher temperature for the selenium addition. This could suggest a higher bandgap in the alloys with the selenium additions in solid solution; in this connection, it should be noted that addition of selenium in these alloys, unlike excess tellurium, should result in the preferential formation of Sb_2Se_3, which has a higher free energy of formation than either the Bi_2Te_3, Bi_2Se_3, or

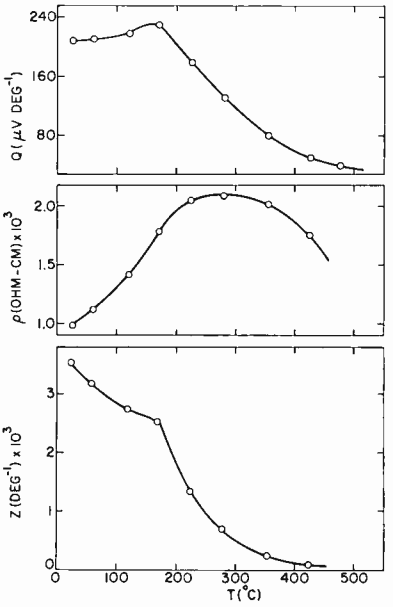


Fig. 10—Temperature dependence of thermoelectric properties in p-type 25% Bi_2Te_3–75% Sb_2Te_3 (1.75% Se) alloy.

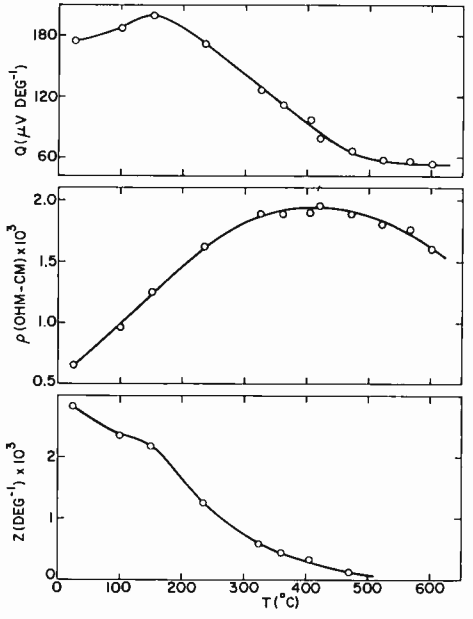


Fig. 11—Temperature dependence of thermoelectric properties in p-type 20% Bi_2Te_3–80% Sb_2Te_3 (6% Te) alloy.

Sb_2Te_3. Hence, the alloys containing selenium may actually be ternary solid-solution alloys of Bi_2Te_3–Sb_2Te_3–Sb_2Se_3 with excess tellurium.

The variation of the figure of merit, z, with temperature in the p-type materials of Figures 8-11 was determined in the manner described above for the n-type materials of Figure 7. It may be seen that these alloys are most useful as thermocouple branch materials in the temperature range 25 to 250°C. The best average z values over this temperature range were approximately 2.1×10^{-3} and 2.5×10^{-3} deg^{-1} obtained with the 80% Sb_2Te_3–20% Bi_2Te_3 (6% excess Te) and 75% Sb_2Te_3–25% Bi_2Te_3 (1.75% Se) alloys, respectively. In addition, these alloys showed a very similar gross dependence of z on temperature. It is noteworthy that the 25% Bi_2Te_3–75% Sb_2Te_3 alloy with 1.75% Se (see Figure 10) provided a z value of 3.6×10^{-3} deg^{-1} at 25° C.

Ternary Compounds

In the search for new and improved thermoelectric materials, attention was given to a class of compounds[11,12] having the generalized formula, ABX_2. Here A is a Group I element such as Cu, B is a Group III or Group V element such as In or Sb, and X is a Group VI element such as Te. These ternary compounds represent one of the simpler compositional deviations from the more familiar III-V compound semiconductors, and in these compounds the ratio of bonding electrons to atoms is 4, as is the case for the Group IV elemental semiconductors.

The work on these ternary compounds was initially concerned with those having a cubic structure, in particular, the rock-salt and zinc-blende types. More than twenty compounds of these cubic types were prepared, and their thermoelectric properties determined. The data, summarized in a previous work,[7] showed that this new class of materials includes semiconductors that could be useful for power generation at intermediate temperatures. Of particular significance was the observation that the compounds which crystallize in the rock salt structure[12,13] are all characterized by low lattice thermal conductivities (< 0.0075 watt cm^{-1} deg^{-1}). Examples of these are $AgSbS_2$, $AgSbSe_2$, $AgSbTe_2$, $AgBiS_2$, and $AgBiSe_2$. However, when the Group V element

[11] C. H. L. Goodman and R. W. Douglas, "New Semiconducting Compounds of Diamond Type Structure," *Physica*, Vol. 20, No. 11, p. 1107 (1954).

[12] J. H. Wernick and K. E. Benson, "New Semiconducting Ternary Compounds," *Jour. Phys. Chem. Solids*, Vol. 3, Nos. 1 & 2, p. 157 (1957).

[13] P. Ramdohr, "Schapbuchite, Matildite, and the Silver and Bismuth Contents of Certain Galenites," *Sitzungsber Preuss. Akad. Wiss. Phys. Math. Kl*, VI, p. 71 (1938).

in these compounds is replaced by a Group III element, such as In and Ga, the resulting compounds have the zinc-blende structure, and much higher values of κ_{ph} (0.01-0.07 watt cm^{-1} deg^{-1}.) In this connection, it is noteworthy that in the zinc-blende structure each atom has four nearest neighbors arranged tetrahedrally, as against six neighbors arranged octahedrally for the rock-salt structure. This reduction in co-ordination number is believed to be largely responsible for the marked increase in κ_{ph} in the zinc-blende compounds, since this might be expected to lead to a decrease in the lattice anharmonicity at any given temperature for materials with similar melting points, or Debye temperatures.

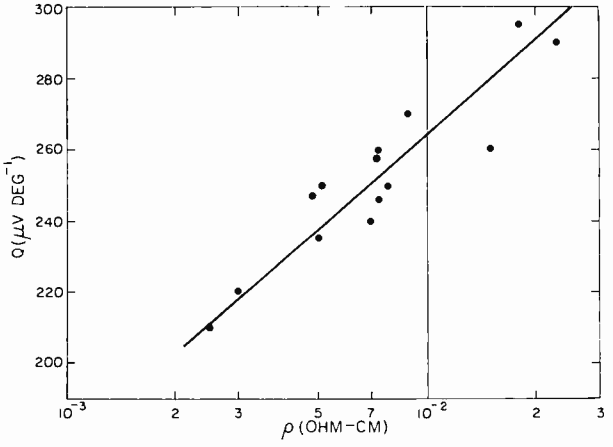


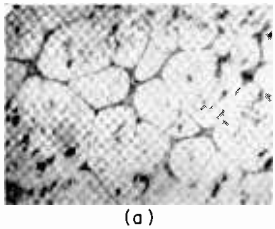
Fig. 12—Variation of thermoelectric power Q with resistivity ρ in p-type $AgSbTe_2$ at 300°K.

A large number of ternary compounds have been examined, and of these $AgSbTe_2$ with the rock-salt structure[12] was found to be the most promising for power generation. This material has a lattice thermal conductivity[14] at room temperature of only 0.0063 watt cm^{-1} deg^{-1}, a hole mobility of $\sim$140 $cm^2 volt^{-1} sec^{-1}$ for a carrier concentration of $\sim 5 \times 10^{18}$ cm^{-3}, and an energy bandgap of $\sim$0.3 ev estimated from resistivity versus temperature data. From measurements of thermoelectric power and carrier concentration at 300°K, the density-of-states effective mass ratio, $m^{(N)}/m$, for holes was calculated from Equation (14) to be 0.36 or 0.88, depending on whether a predominantly impurity or lattice-scattering is assumed, respectively, in assigning a value to the kinetic energy term, $|\Delta\epsilon_T|/(kT)$. Finally,

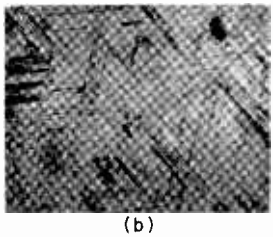
[14] E. F. Hockings, "The Thermal Conductivity of Silver Antimony Telluride," *Jour. Phys. Chem. Solids*, Vol. 10, No. 4, p. 341, August 1959

data[7] on the variation of thermoelectric power Q with resistivity ρ at 300°K (see Figure 12) show that Q values of 200-220 μv deg^{-1} can be expected in the resistivity range, 2-3 $\times$ 10^{-3} ohm-cm.

The principal drawback to the proper evaluation and utilization of $AgSbTe_2$ as a material for power generation is the present inability to adjust its electrical properties through controlled impurity additions. For example, it has not yet been possible to obtain either p-type material with resistivities below 2 $\times$ 10^{-3} ohm-cm, or n-type material. Attempts to obtain these electrical characteristics have failed in spite of a large number of experiments with elemental and compound doping. Perhaps this can be attributed, in part, to the



(a)



(b)

Fig. 13—Appearance of second-phase structures (× 100) in $AgSbTe_2$: (a) Ag_2Te, (b) Sb_2Te_3.

presence of second phases in the as-prepared $AgSbTe_2$, which were recently reported by Armstrong *et al*[15] and Wolfe *et al.*[16] In agreement with this recent work, the phases were determined by x-ray analysis to be Ag_2Te and possibly Sb_2Te_3, and metallographic examination revealed a peritectic-type distribution for the Ag_2Te phase, and a typical Widmanstätten precipitate pattern for the Sb_2Te_3.

---

[15] R. W. Armstrong, J. W. Faust, Jr. and W. A. Tiller, "A Structural Study of the Compound $AgSbTe_2$," *Jour. Appl. Phys.*, Vol. 31, No. 11, p. 1954, November 1960.

[16] R. Wolfe, J. H. Wernick and S. E. Hasko, "Anomalous Hall Effect in $AgSbTe_2$," *Jour. Appl. Phys.*, Vol. 31, No. 11, p. 1959, November 1960.

Examples of these second-phase structures are shown in the micrographs of Figure 13. It is possible that the difficulty encountered in controlling the conductivity in $AgSbTe_2$ by impurity additions results from a relatively higher solubility of such impurities in these second phases.

Measurements of thermoelectric power, Q, and electrical resistivity, ρ, were made on p-type $AgSbTe_2$ over the temperature range 25 to 525°C. The results are shown in Figure 14, along with the calculated temperature dependence of the figure of merit, z. It may be seen that ρ decreases linearly with increasing temperature, while Q increases with temperature and exhibits a broad maximum. The decrease in ρ would seem to suggest an anomalous increase in hole

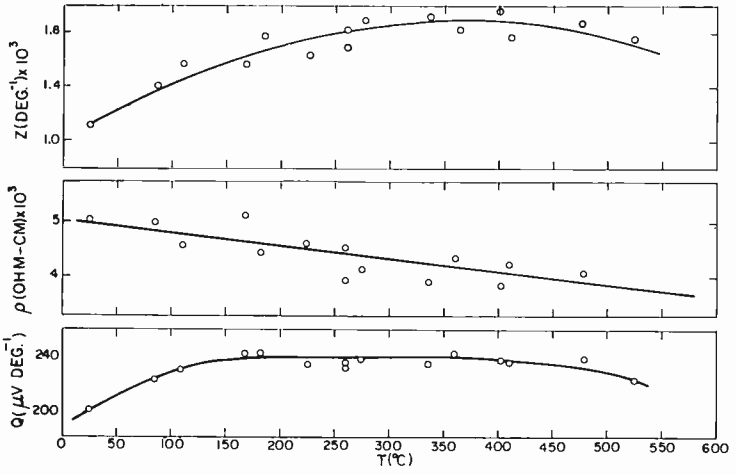


Fig. 14—Temperature dependence of thermoelectric properties in p-type $AgSbTe_2$.

mobility with temperature. It must be remembered, however, that in these specimens a second phase exists whose concentration might be expected to vary with temperature. The increase in Q with temperature follows from Equation (14), and the fact that Q does not decrease sharply at the high temperatures suggests that intrinsic conduction due to thermal generation of electron–hole pairs is not encountered even at 525°C.

The temperature dependence of z in Figure 14 was determined from Equation (9) using the data on Q and ρ. The total thermal conductivity, $\kappa = \kappa_{el} + \kappa_{ph}$, at any given temperature was estimated by calculating the electronic contribution κ_{el} from the measured value of ρ, and assuming the room-temperature value of the lattice con-

tribution, κ_{ph}, (0.0063 watt cm^{-1} deg^{-1}) to be invariant with temperature. The latter assumption again represents a conservative approach toward estimating the temperature dependence of z. Although preliminary measurements of κ_{ph} in $AgSbTe_2$ up to 500°C by the thermal diffusivity technique[17] showed a complex function of κ_{ph} with temperature, there was nevertheless an over-all decrease in κ_{ph}. No account was taken here of a possible ambipolar contribution to the thermal conductivity, which is reasonable in view of the Q values at the high temperatures. Figure 14 shows that the broad maximum in Q versus T is also manifested in the temperature dependence of z. It is noteworthy that an average z value of approximately 1.6×10^{-3} deg^{-1} is obtained in p-type $AgSbTe_2$ over the temperature range 100 to 500°C.

Solid-Solution Alloys

It has been shown by a number of investigators[7,10,18-20] that the figure of merit of semiconductors can be improved by solid-solution alloying in the carrier-concentration range where the carrier mobility is limited by impurity scattering. Under these conditions, the lattice strains due to alloying are effective in scattering phonons, but are ineffective in scattering electrons which have longer wavelengths. Thus, at a constant resistivity corresponding to near-degeneracy, solid-solution alloying could improve the figure of merit by decreasing the lattice thermal conductivity, κ_{ph}, without at the same time affecting the carrier concentration and, hence, thermoelectric power (i.e., κ decreases while the ratio Q^2/ρ remains constant). This, of course, assumes no change in the density-of-states effective mass of carriers with alloying. In addition to this beneficial effect on the figure of merit, solid-solution alloying makes it possible to adjust in a controlled manner such important physical properties as energy bandgap. In view of these considerations, the following alloy systems were examined: (1) $AgSbTe_2$ with the IV-VI compounds, PbTe, SnTe

[17] B. Abeles and G. D. Cody, Private Communication (June 1960).

[18] A. F. Ioffe, S. V. Airapetyants, A. V. Ioffe, N. V. Kolomoets, L. S. Stil'bans, "Increase in the Effectiveness of Semiconducting Thermocouples," *Doklady Akademii Nauk. SSSR*, V. 106, No. 6, p. 981 (1956); translation No. 3, *Infosearch Russian Literature Survey*, SEM-1-56; A. V. Ioffe and A. F. Ioffe, "Thermal Conductivity of Semiconductors," *Izvestiya Akademii Nauk. SSR*, Ser. Fiz. V. 20, No. 1, p. 65 (1956); *Infosearch Translation* IA-3-56.

[19] M. C. Steele and F. D. Rosi, "Thermal Conductivity and Thermoelectric Power of Germanium–Silicon Alloys," *Jour. Appl. Phys.*, Vol. 29, No. 11, p. 1517, November 1958.

[20] G. N. Gordyakova, G. V. Kokosh and S. S. Sinani, "Investigation of Thermoelectric Properties of Solid Solutions Bi_2Te_3–Bi_2Se_3," *Zhurnal Tekhnicheskoi Fiziki*, Akad. Nauk. SSSR, B. 28, No. 5, p. 977 (1958).

and GeTe, (2) PbTe-SnTe and (3) InAs-GaAs. Thermoelectric data on these systems are presented in the following sections.

1. $AgSbTe_2$ Alloys

The alloys of $AgSbTe_2$ with PbTe, SnTe and GeTe were selected for examination, because the similarity in the structures of the component compounds in these systems suggested the existence of extensive ranges of solid solution. The system $AgSbTe_2$–PbTe has been reported upon by several authors[21–23] who have indicated that solid

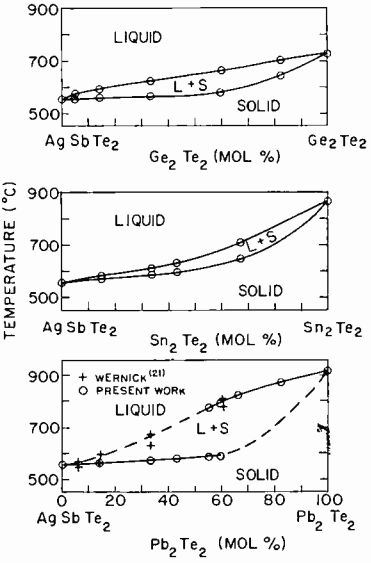


Fig. 15—Phase equilibrium diagrams for the binary systems of $AgSbTe_2$ with PbTe, SnTe, and GeTe.

solubility occurs over the entire system. For the system containing SnTe, it has been stated[22] that complete solid miscibility occurs, but no details were presented. In the present work the three systems of $AgSbTe_2$ with PbTe, SnTe, and GeTe were examined, and their tentative phase diagrams, shown in Figure 15, were obtained by differential thermal analysis during heating of small specimens in sealed silica ampules. The results indicate that a complete series of solid

---

[21] J. H. Wernick, in *Properties of Elemental and Compound Semiconductors*, Interscience Publishers, New York, N. Y., p. 69 (1960).

[22] H. Fleischmann, C. G. Folberth and H. Pfister, "Semiconducting Mixed Crystals of the Type $(A_{x/2}{}^{I}B_{1-x}{}^{IV}C_{x/2}{}^{V})\ D^{VI}$," *Z. Naturforschg*, V. 14a, p. 999 (1959).

[23] H. Rodot, "Study and Properties of the System $AgSbTe_2$–PbTe," *Compt. Rend.*, V. 249, p. 1872 (1959).

solutions can be obtained in all three systems, although for the system $AgSbTe_2$–PbTe the solidus curve is not firmly established.

The lattice parameters of the materials in the three alloy systems were determined at 300°K by x-ray powder-diffraction methods, and their variation with respect to alloy composition is shown in Figure 16. Cubic symmetry was observed in all cases at room temperature, except for GeTe and one alloy containing a high proportion of GeTe. These two exceptions showed a rhombohedral structure, with angles of 88.35° and 89.2°, respectively, which represents only a slight distortion from cubic symmetry. The approximately linear dependence of the lattice parameters in Figure 16 indicates complete solid

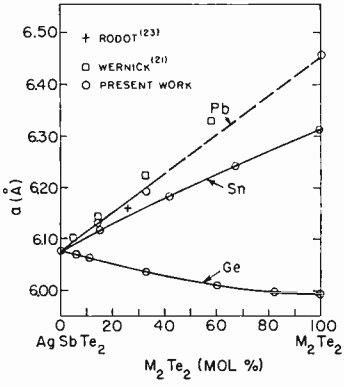


Fig. 16—Variation of lattice parameter with composition in alloy systems of $AgSbTe_2$ with PbTe, SnTe, and GeTe (M = Pb, Sn, or Ge).

solubility for all three systems in agreement with the phase diagrams shown in Figure 15. However, close examination of the three curves reveals a significant difference between the system $AgSbTe_2$–SnTe and the system $AgSbTe_2$–GeTe. In the former, the variation of lattice parameter with alloying shows a positive deviation from linearity, which is to be expected for a disordered solid-solution system. By contrast, a small negative deviation is observed in the $AgSbTe_2$–GeTe system, where attractive interatomic forces of a partially ionic character can be expected.

The variation of lattice thermal conductivity with composition for the three alloy systems at 300°K is shown in Figures 17 and 18. Again κ_{ph} was obtained from the measured total thermal conductivity κ by Equation (15). The electronic component, κ_{el}, was determined from Equation (16), assuming that the carriers had either a Boltzmann energy distribution ($S = 2$) or were degenerate ($S = \pi^2/3$). A

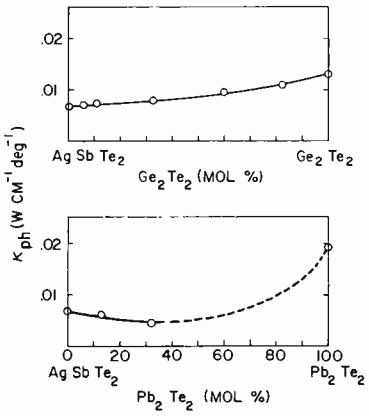


Fig. 17—Variation of lattice thermal conductivity κ_{ph} with composition in alloy systems of $AgSbTe_2$ with PbTe and GeTe.

thermoelectric power of 160 μv deg^{-1} was arbitrarily chosen to indicate the onset of degeneracy. It may be seen from Figures 17 and 18 that κ_{ph} increases monotonically with increasing SnTe or GeTe in $AgSbTe_2$. The absence of a minimum in κ_{ph}, which is commonly observed in alloy systems exhibiting complete solid miscibility,[10,24,25] would suggest that the lattice strains due to alloying in both the $AgSbTe_2$–SnTe and $AgSbTe_2$–GeTe systems are of too small a magnitude to decrease the already low κ_{ph} of $AgSbTe_2$ (0.0063 watt cm^{-1} deg^{-1}). On the other hand, the data in Figure 17 show that in the

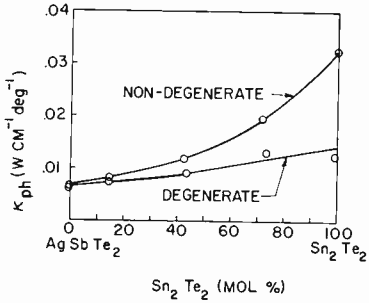


Fig. 18—Variation of lattice thermal conductivity with composition for nondegeneracy $(\kappa_{ph})_{ND}$ and degeneracy $(\kappa_{ph})_{D}$ in the $AgSbTe_2$–SnTe system.

[24] A. F. Ioffe, "Heat Transfer in Semiconductors," Proceedings International Conference Electron Transport in Metals and Solids, *Canadian Journal of Physics*, Vol. 34, No. 12a, p. 1342 (Dec. 1956).

[25] M. S. Abrahams, R. Braunstein and F. D. Rosi, "Thermal, Electrical and Optical Properties of (In, Ga) As Alloys," *Jour. Phys. Chem. Solids*, Vol. 10, Nos. 2 & 3, p. 204 (July 1959).

system $AgSbTe_2$–PbTe a minimum κ_{ph} of 0.0045 watt cm^{-1} deg^{-1} is obtained corresponding to an alloy composition of approximately 50% $AgSbTe_2$–50% PbTe. In this connection, it is noteworthy that the solidus and liquidus curves in the $AgSbTe_2$–PbTe system are more widely separated than in the $AgSbTe_2$ system with SnTe or GeTe. This wider separation could be a consequence of a much higher degree of lattice strain due to alloying. The very low lattice thermal conductivity in the $AgSbTe_2$–PbTe system demonstrates that the lower limit to the "mean free path" of phonons in semiconductors can be less than a unit-cell dimension. This is important when attempting to estimate a maximum figure of merit for a given class of materials.

Table II—Electrical Resistivity and Thermoelectric Power of $AgSbTe_2$ Alloys at 300°K

Alloy Composition	ρ (ohm-cm)	Q (μv deg^{-1})
75% $AgSbTe_2$—25% PbTe	3.1×10^{-2}	300
50% $AgSbTe_2$—50% PbTe	5.0×10^{-2}	350
40% $AgSbTe_2$—60% PbTe	3.0×10^{-2}	380
75% $AgSbTe_2$—25% SnTe	3.6×10^{-3}	110
50% $AgSbTe_2$—50% SnTe	4.8×10^{-4}	30
40% $AgSbTe_2$—60% SnTe	1.0×10^{-3}	71
20% $AgSbTe_2$—80% SnTe	6.4×10^{-4}	52
90% $AgSbTe_2$—10% GeTe	7.0×10^{-3}	235
75% $AgSbTe_2$—25% GeTe	5.5×10^{-3}	230
50% $AgSbTe_2$—50% GeTe	3.3×10^{-3}	212
25% $AgSbTe_2$—75% GeTe	1.1×10^{-3}	131
10% $AgSbTe_2$—90% GeTe	5.2×10^{-4}	70

Data from electrical resistivity and thermoelectric power measurements made at room temperature for the three $AgSbTe_2$ alloy systems are shown in Table II. On the basis of these results, as well as thermal conductivity measurements, the alloys with GeTe were considered the most promising materials for use in thermoelectric power generation. Consequently, detailed measurements were made on a selected number of p-type $AgSbTe_2$–GeTe alloys, as well as on GeTe itself. The results of the thermoelectric power and electrical resistivity measurements in the temperature range 25 to 600°C are shown in Figures 19 and 20. It may be seen from the curves in Figure 19 that ρ initially increases with temperature for all the alloys, and this can be attributed to a decrease in hole mobility with increasing tempera-

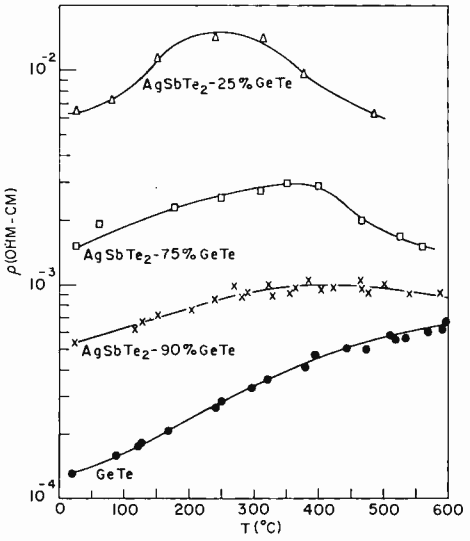


Fig. 19—Variation of resistivity ρ with temperature for p-type $AgSbTe_2$–GeTe alloys.

ture. It also appears that the maximum in resistivity is shifted to higher temperatures with increasing GeTe content in the alloys; in pure GeTe the maximum was not observed even at 600°C. The curves in Figure 20 show that the dependence of Q on temperature is similar to that for ρ, in that it also increases with temperature and reaches a maximum value which is shifted to higher temperatures with increasing amounts of GeTe. These data strongly suggest that with increasing GeTe content in the alloys, the onset of intrinsic conductivity occurs at higher temperatures.

To determine the possible usefulness of the $AgSbTe_2$–GeTe alloys

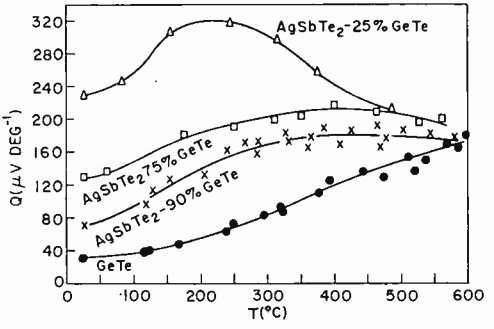


Fig. 20—Variation of thermoelectric power Q with temperature for p-type $AgSbTe_2$–GeTe alloys.

as power-generating materials, the temperature dependence of the figure of merit for the various p-type alloys was calculated from Equation (9) using the data in Figures 19 and 20. The results are shown in Figure 21; for comparison, the data for pure GeTe are also included. The total thermal conductivity was obtained by calculating κ_{el} from the measured values of ρ at the different temperatures, and by assuming a T^{-1} dependence for κ_{ph} with a lower limiting value of 0.0063 watt cm^{-1} deg^{-1} corresponding to the κ_{ph} of $AgSbTe_2$.

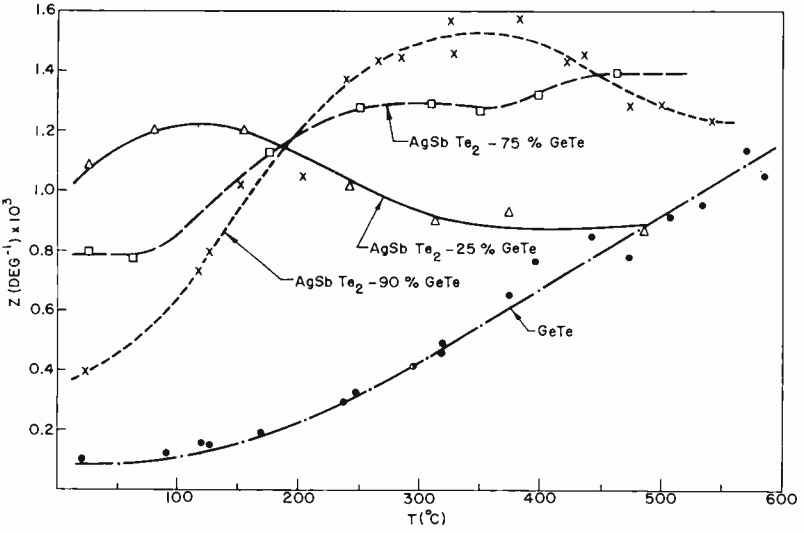


Fig. 21—Variation of figure of merit z with temperature for p-type $AgSbTe_2$–GeTe alloys.

It may be seen from the curves in Figure 21 that the 10% $AgSbTe_2$–90% GeTe alloy appears to be the most promising for thermoelectric power generation in the temperature range 250 to 550°C where an average z value of $\sim 1.5 \times 10^{-3}$ deg^{-1} is obtained. Although this does not represent an improvement over the average z value for $AgSbTe_2$ in the same temperature range (see Figure 15), it should be noted that the useful temperature range of the 10% $AgSbTe_2$–90% GeTe alloy will be 50 to 100°C higher than that of pure $AgSbTe_2$. Moreover, the thermoelectric properties have not been optimized for the intended temperature range of application. One point of interest in connection with the data in Figure 21 is the marked improvement that is made in the figure of merit of GeTe over the temperature range 25 to 500°C by simply alloying with 10% $AgSbTe_2$. Moreover, the data would suggest that pure GeTe is superior to any of the alloys above 600°C.

2. PbTe–SnTe Alloys

Interest in an n-type thermoelectric material for use in the temperature range above that of the Bi_2Te_3 alloys led to an examination of the system PbTe–SnTe, which had been shown to exhibit complete solid miscibility.[26] It was known from previous work[27] that PbTe is useful in this temperature range, and it was considered that solid-solution alloying with SnTe would provide a higher figure of merit by decreasing the lattice thermal conductivity.

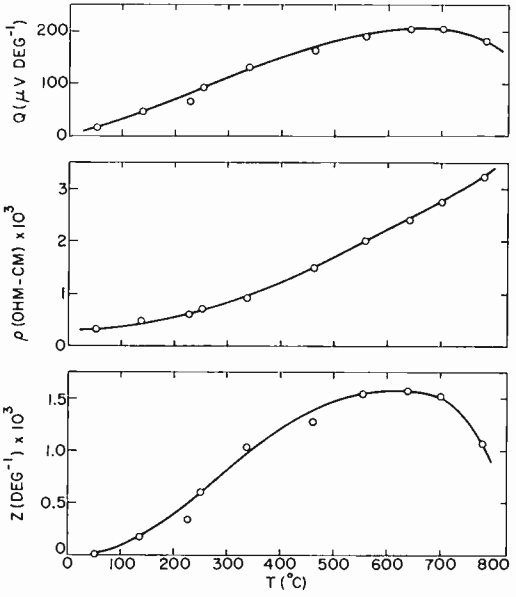


Fig. 22—Temperature dependence of thermoelectric properties for n-type 75% PbTe–25% SnTe alloys.

A study of a number of n-type PbTe-SnTe alloys showed that one of the most promising compositions was 75% PbTe–25% SnTe, which has electrical properties comparable to PbTe. In addition, this alloy has a lattice thermal conductivity of 0.012 watt cm^{-1} deg^{-1} at 25°C, which is significantly lower than that of PbTe (0.017 watt cm^{-1} deg^{-1}).[27] Detailed measurements were made of the temperature dependence of Q and ρ for the 75% PbTe–25% SnTe alloy, and

[26] N. K. Abrikosov, K. A. Dyuldina and T. A. Danilyan, "On the System SnTe–PbTe," *Zhur. Neorg. Khim.*, V. 3, p. 1632 (1958).

[27] R. W. Fritts, in *Thermoelectric Materials and Devices*, Reinhold Pub. Corp., New York, N. Y., 1960, p. 143.

the data are summarized in Figure 22. It may be seen that the onset of intrinsic conduction occurs in these specimens at about 700°C, which probably represents the upper temperature limit for the use of this alloy in power generation. This is also indicated in Figure 22 by the dependence of z on temperature which was obtained from Equation (9), where κ was again estimated by calculating κ_{el} from the measured ρ, and by assuming a T^{-1} dependence for κ_{ph}. The latter assumption is reasonable on the basis of recent thermal diffusivity measurements on solid-solution alloys of semiconductors.[28] The variation of z with temperature also shows a broad maximum in z, which provides an average z value of $\sim 1.3 \times 10^{-3}$ deg^{-1} over the temperature range 300 to 700°C.

3. InAs–GaAs Alloys

To extend thermocouple operation above 550°C, solid-solution alloys of InAs–GaAs were examined. This alloy system had been found to exhibit complete solid miscibility,[25] and materials could be obtained with melting points ranging from 960°C to 1242°C and bandgaps from 0.35 ev to 1.35 ev. Equally important, the lattice thermal conductivity of both InAs and GaAs is considerably reduced by alloying, as may be seen from Figure 23 which is taken from the work of Abrahams, Braunstein, and Rosi.[25] A minimum in κ_{ph} of 0.047 watt cm^{-1} deg^{-1} was obtained for an alloy composition of approximately 50% InAs–50% GaAs, which corresponds to an energy bandgap of 0.6 ev.

The variation of figure of merit with resistivity at 300°K for the n-type 40% InAs–60% GaAs alloy is shown in Figure 24. For comparison, data are also included for the component compounds, InAs and GaAs. It may be seen that the maximum z value for the alloy is significantly larger than that for either of the compounds, and this improvement in z with alloying is again a manifestation of the large decrease in κ_{ph} resulting from solid-solution alloying. The maximum z of 2.5×10^{-4} deg^{-1} for the 40% InAs–60% GaAs alloy is considerably lower than that for other materials at 300°K. Equation (17) shows, however, that an improvement in this value can be expected at higher temperatures. This can be seen from the data in Figure 25, which shows the variation of Q, ρ, and z with temperature in the range 25 to 650°C for a specimen with a room temperature resistivity of 1.17×10^{-3} ohm-cm corresponding to an electron concentration of 1.8×10^{18} cm^{-3}. The temperature dependence of z was determined from Equation (9), where the total thermal conduc-

[28] B. Abeles and G. D. Cody, private communication (Jan. 1961).

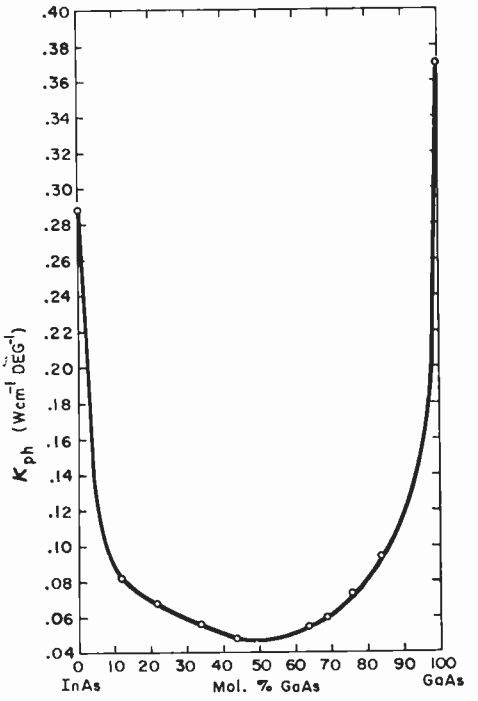


Fig. 23—Variation of lattice thermal conductivity κ_{ph} with composition in InAs–GaAs system.

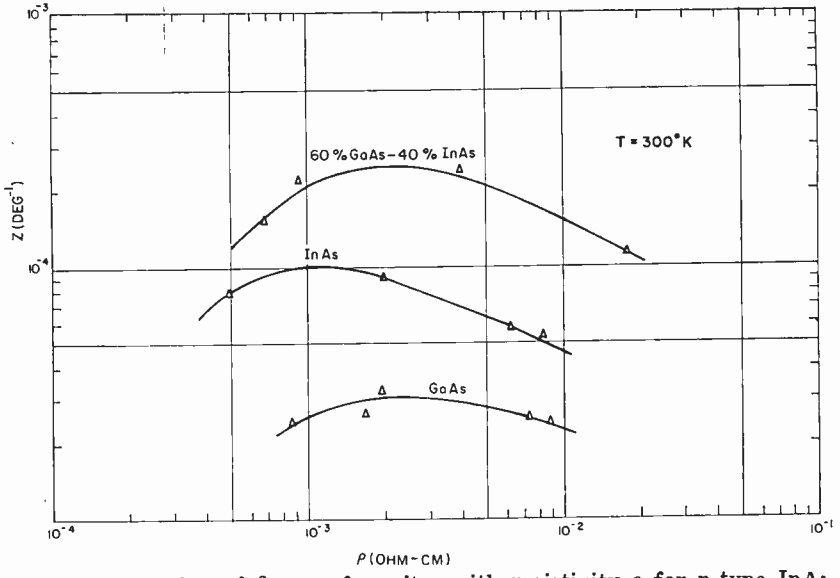


Fig. 24—Variation of figure of merit z with resistivity ρ for n-type InAs GaAs, and 60% GaAs–40% InAs alloy.

tivity at a given temperature was estimated by calculating κ_{el} from the measured value of ρ, and by assuming the κ_{ph} (0.05 watt cm^{-1} deg^{-1} at 25°C) to be invariant with temperature. The maximum z value of 4.7×10^{-4} deg^{-1} at 425°C for the 40% InAs–60% GaAs alloy is approximately 1.9 times that at room temperature. Moreover, if in estimating κ the reasonable assumption was made that κ_{ph} varies as T^{-1}, expected figures of merit would be $\sim 0.9 \times 10^{-3}$ deg^{-1} at 425°C and $\sim 1 \times 10^{-3}$ deg^{-1} at 625°C.

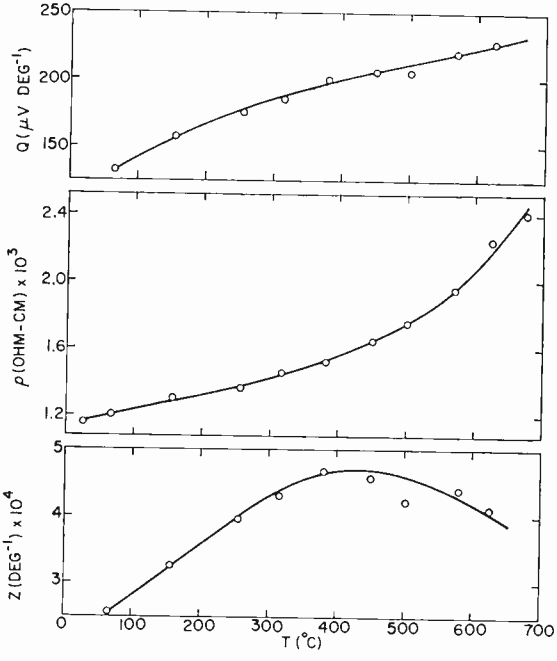


Fig. 25—Temperature dependence of thermoelectric properties for n-type 40% InAs–60% GaAs alloy.

It is significant that the data in Figure 25 show that Q increases approximately as log T, which indicates that the 40% InAs–60% GaAs alloy remains extrinsic up to 625°C. The increase in ρ with temperature is probably a result of a decrease in electron mobility, since the carrier concentration for this degenerate (at 25°C) material can be regarded as temperature independent.

Figure 26 shows the variation in thermoelectric power and resistivity with temperature up to 700°C for an 80% InAs–20% GaAs alloy with a room-temperature resistivity and thermoelectric power of 1.37×10^{-3} ohm-cm and 173 μv deg^{-1}, respectively. From the

temperature dependence of both Q and ρ, it may be seen that the specimen becomes intrinsic in the vicinity of 425°C. The variation of z with temperature is also shown in Figure 26, and z was again calculated from Equation (9), where κ at any given temperature was obtained in the same manner as described above for the 40% InAs–60% GaAs alloy using a κ_{ph} at 25°C of 0.07 watt cm^{-1} deg^{-1}. The observation that z rises continuously with increasing temperature follows from the fact that Q does not decrease rapidly with onset of

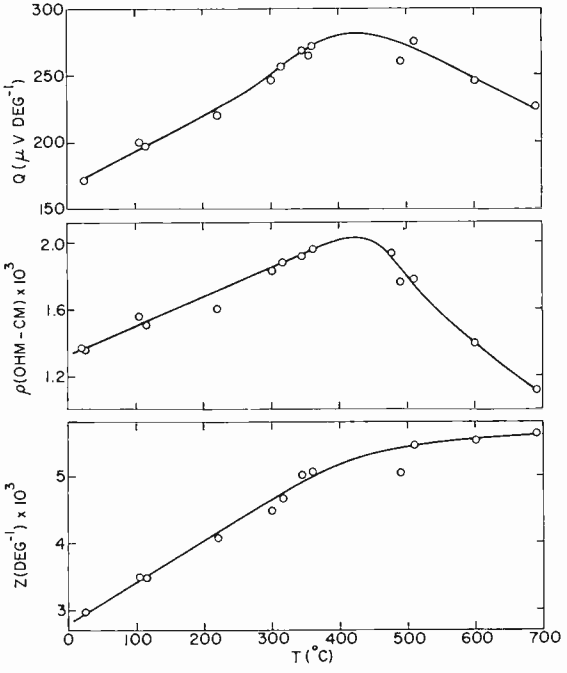


Fig. 26—Temperature dependence of thermoelectric properties for n-type 80% InAs–20% GaAs alloy.

intrinsic conductivity (presumably due to the large mobility ratio for current carriers in this system), and from Equation (17). In the temperature range 25 to 425°C, where the specimen is extrinsic, the z values can be improved still further by optimum doping so as to provide Q values in accordance with Equation (18) over this temperature range. This is generally true for all cases of the variation of z with temperature in extrinsic semiconductors.

A comparison of the data in Figures 25 and 26 clearly demonstrates the need for a higher bandgap to shift the onset of intrinsic

conductivity to higher temperatures, since the bandgaps of the 80% InAs–20% GaAs and 40% InAs–60% GaAs alloys are approximately 0.45 ev and 0.70 ev, respectively.[25] In addition, the higher Q values for the 80% InAs–20% GaAs alloy in the extrinsic temperature range 25 to 425°C reflect for the most part the higher electron mobility in this alloy. For a similar carrier concentration, the electron mobility of the 80% InAs–20% GaAs alloy at 300°K is more than twice that of the 40% InAs–60% GaAs alloy. Since the resistivity of the two alloys are similar and no significant change is to be expected in the density-of-states effective mass for electrons with alloying,[25] the difference in the Q values of the two alloys up to 425°C simply manifests a difference in the electron concentration.

Table III—Room-Temperature Properties of Semiconductor Branches of Power-Generating Thermocouple

Thermocouple Branch Material	Conductivity Type	Impurity Added	Q (μv deg^{-1})	ρ (ohm-cm)	κ (watt cm^{-1} deg^{-1})
75% Bi_2Te_3—25% Bi_2Se_3	n	CuBr	150	6×10^{-4}	0.020
50% Bi_2Te_3—40% Sb_2Te_3 —10% Sb_2Se_3	p	Bi	155	8×10^{-4}	0.015

Power-Generating Thermocouples

25 to 325°C Range

It was shown in the previous section (Figures 7-11) that Bi_2Te_3 alloys could be useful for power generation at least up to 300°C. Consequently, a power-generating thermocouple, similar to the arrangement in Figure 1, was constructed from n-type 75% Bi_2Te_3–25% Bi_2Se_3 and p-type 50% Bi_2Te_3–40% Sb_2Te_3–10% Sb_2Se_3 alloys, whose thermoelectric properties are given in Table III. The resistivity of each alloy was adjusted to provide not the highest value of z at room temperature, but instead the best average value of z over the desired operating temperature range of the thermocouple. In constructing the thermocouple, the metal–semiconductor junctions were made by soldering with an 80% Bi–20% Sb alloy, which provided a junction resistance at 25°C of at least one order of magnitude less than the sum of the branch resistances. Heat was supplied to the hot junction by a wire-wound resistance heater, and the cold junction was fixed at 25°C by circulating water.

Figure 27 shows the variation in power-generating efficiency, ϕ, (power output/heat input) with temperature difference of operation, as measured in a vacuum. It may be seen that a ϕ_{max} of approximately 6.5% is obtained for a $T_H - T_C = 300°C$ (i.e., $T_C = 25°C$ to $T_H = 325°C$). From the measured Q, ρ, and κ of the alloys in Table III over this temperature range, an efficiency of ~8 per cent was expected. The difference of 1.5 per cent from the measured and calculated values of ϕ can be attributed to an increase in the metal–semiconductor contact resistance at the hot junction. From the efficiency

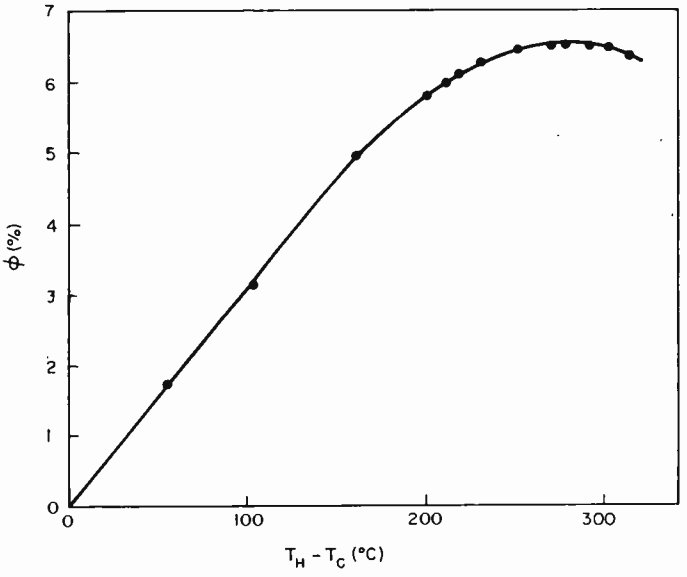


Fig. 27—Variation of power-generating efficiency ϕ with temperature range of operation ($T_C = 25°C$).

measurements in Figure 27, it was possible to calculate the actual temperature variation of the average figure of merit for the n- and p-type alloys of Table III. These data are given in Figure 28, which shows a continuous decrease in z with increasing temperature of thermocouple operation in a manner very similar to the dependence of z on T for the alloys in Figures 7 through 11.

20 to 550°C Range

Power-generating thermocouples for operation over the temperature range 20 to 550°C were constructed from semiconductor alloys in the sandwich arrangement shown schematically in Figure 29. It

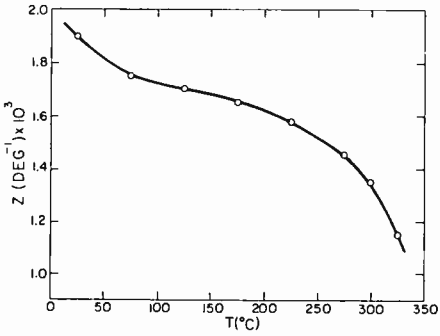


Fig. 28—Temperature dependence of average figure of merit z for n-type 75% Bi_2Te_3–25% Bi_2Se_3 and p-type 50% Bi_2Te_3–40% Sb_2Te_3–10% Sb_2Se_3 alloys.

may be seen that the n-type branch consisted of alloys of 75% Bi_2Te_3–25% Bi_2Se_3 and 75% PbTe–25% SnTe, while the p-type branch was constructed from alloys of 25% Bi_2Te_3–75% Sb_2Te_3 (1.75 weight per cent Se) and 10% $AgSbTe_2$–90% GeTe. The metal–semiconductor bonds at the cold and intermediate junctions for both the p- and n-type materials were made by soldering, while pressure contacts were used for the hot junctions. Heat was again supplied to the hot junction by a wire-wound resistance heater, and the cold junction temperature was maintained at 20°C by circulating water.

Measurements of Q, ρ, and κ as a function of temperature on the individual materials indicated that a power-generating efficiency of approximately 14 per cent could be expected with the arrangement shown in Figure 29. The variation of ϕ with hot junction tempera-

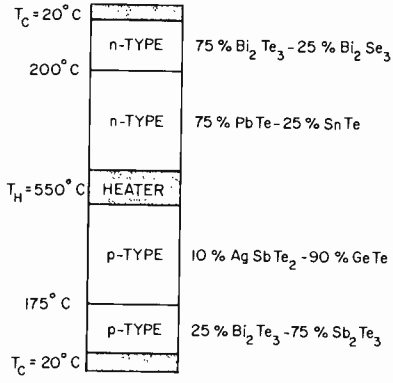


Fig. 29—Schematic representation of sandwich-type power-generating thermocouple for operation in the temperature range 20 to 550°C.

ture T_H in Figure 30 shows that an efficiency of ~12 per cent was actually realized on operating in air up to a T_H of 550°C (i.e., a $T_H - T_C = 530°C$). The curves were obtained from two thermocouples, which were cycled over the temperature range of operation.

The efficiencies shown in Figure 30 were measured in the following manner. The heater was surrounded by a metallic collector, which was calibrated so that its temperature could be directly related to the component of the power input arriving at the collector. This

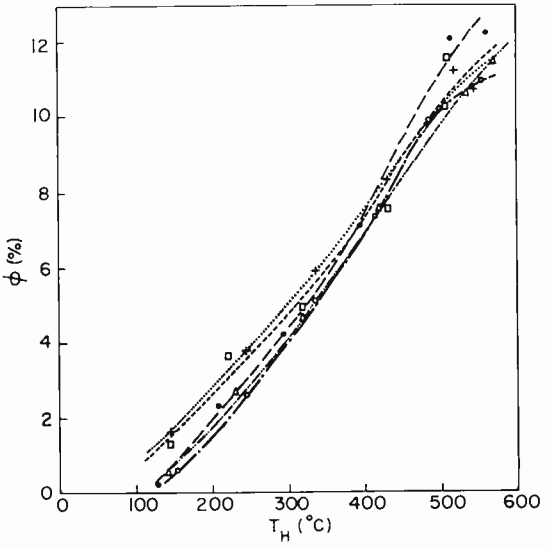


Fig. 30—Variation of power generating efficiency ϕ with temperature of hot junction T_H. ($T_C = 20°C$.)

calibration was in close agreement with calculations based upon the various thermal-conduction paths from the heater to the collector. The thermocouple efficiency was then determined by considering the difference between the total power input to the thermocouple and the heat that was transferred to the collector box. This difference represents the heat flow into the thermocouple, which, together with the measured electrical output, enabled the efficiency to be calculated. A second method for estimating efficiencies involved a direct measure of the heat flow into the cold junctions by the temperature rise of the circulating water. This calorimetric determination of the heat flow from T_H to T_C was then combined with the thermocouple power output to compute the efficiency. The data obtained for two rates of water flow are listed in Table IV. It is seen that the efficiency deter-

minations by this method are in substantial agreement with the data in Figure 30.

Studies of the operating lifetime of these thermocouples were made in an argon atmosphere; after 300 hours of continuous operation the efficiency was 20 per cent lower than its original value. This drop in efficiency was attributed to an increase in the resistance of the metal–semiconductor bonds, since there was no evidence of sufficient impurity diffusion along the branches of the thermocouples to affect materials properties.

Table IV—Generator Efficiency Determination by Calorimetric Method

Property	Run 1	Run 2
Heater Power (watts)	22.1	22.1
Heater Temperature (°C)	557.5	557.5
Water Flow (cc per sec)	0.668	0.45
Temperature Difference (°C)	2.2	3.9
Water Uptake (watts)	6.18	7.4
Generator Output (watts)	0.75	0.9
Efficiency (%)	10.9	10.9

Conclusions

A general consideration of the thermoelectric properties of semiconductors suggests that (1) this class of materials can be useful in power-generating thermocouples operating at least up to 700°C, and (2) construction of thermocouple branches in a sandwich-type arrangement or by graded alloying will be necessary to achieve high figures of merit over a wide temperature range and, hence, high power-generating efficiencies.

A large number of ternary compound semiconductors having the cubic structure were synthesized; and those, such as $AgSbTe_2$, with the rock-salt structure are characterized by low lattice thermal conductivities (<0.0075 watt cm^{-1} deg^{-1}).

The lattice thermal conductivity as a function of composition was examined in the alloy systems of $AgSbTe_2$ with PbTe, SnTe, and GeTe. The minimum in the lattice thermal conductivity for the $AgSbTe_2$–PbTe system gives an effective "mean free path" for phonons which is less than unit-cell dimensions.

Measurements of the temperature dependence of thermoelectric

properties on a number of solid-solution alloy systems showed that (1) solid-solution alloys of Bi_2Te_3 with Bi_2Se_3, Sb_2Te_3, and Sb_2Se_3 provided the best p- and n-type material for thermocouple operation in the temperature range 25 to 250°C, (2) the ternary compound $AgSbTe_2$ and its alloys with GeTe provided the best p-type material for the range 250 to 550°C, and (3) alloys in the PbTe–SnTe system provided the best n-type material for the range, 250 to 550°C.

Power-generating thermocouples constructed from alloys in the $(Bi, Sb)_2$ $(Te, Se)_3$ system have given an efficiency (power output/heat input) of ~6.5 per cent on operating over the temperature range, 25 to 325°C; i.e., a $T_H - T_C = 300°C$.

Power-generating thermocouples were constructed in the sandwich-type arrangement of materials from the p-type alloys, 25% Bi_2Te_3–75% Sb_2Te_3 and 10% $AgSbTe_2$–90% GeTe, and the n-type alloys, 75% Bi_2Te_3–25% Bi_2Se_3 and 75% PbTe–25% SnTe. These thermocouples provided an efficiency of ~12 per cent on operating over the temperature range 20 to 550°C; i.e., a $T_H - T_C = 530°C$. Continuous thermocouple operation for 300 hours resulted in no significant deterioration of material properties.

ACKNOWLEDGMENTS

The authors are especially indebted to M. Allen, E. V. Fitzke and W. Mularz for their valuable assistance in materials preparation and physical measurements; to N. J. DiGuiseppe for his assistance in constructing the power-generating thermocouples; and to W. C. Roth for x-ray analyses. Finally, the authors wish to thank G. D. Cody for his constructive comments during the course of this work.

THE PRESENT OUTLOOK FOR CONTROLLED THERMONUCLEAR FUSION*

By

George Warfield†

RCA Laboratories
Princeton, N. J.

Summary—Extensive effort has been expended in the last few years toward the development of controlled thermonuclear fusion based on the confinement of a very-high-temperature plasma. It is the opinion of the author that, barring a new approach, a useful power-producing fusion reactor is several decades away. In the immediate future, the concentration in fusion research will probably shift from a study of fusion itself to a study of basic plasma physics. Some of the reasons for this conclusion are discussed.

Introduction

Approximately ten years ago, work was started in several countries toward finding a practical method for the controlled release of energy from the fusion of light nuclei to form heavier nuclei. The main incentive for this work was the realization that success in this effort would assure mankind of an inexhaustible energy supply. The energy stored in the deuterium contained in the waters of the world is estimated to be 10^{24} kilowatt hours. At the estimated world-wide consumption rate for the year 2000 of 10^{14} kwh per year, there is sufficient fusion fuel to last 10^{7} years assuming an efficiency of utilization of 0.1 per cent. For all practical purposes this is an inexhaustible supply.

The major problems to be solved — heating and containment of the fuel — were recognized at the very start of the work, and several different proposals were advanced for their solution. While these problems appeared quite formidable at the time, it is probably fair to say that there was rather general optimism that they would be solved satisfactorily within perhaps a decade if an intensive effort were made. This optimism was natural at the time, since two apparently equally formidable problems, the development of fission and fusion weapons, had just recently succumbed to large-scale efforts. Perhaps the best indication of this optimism is the fact that the work

* Manuscript received 16 January 1961.

† Dr. Warfield is also Associate Professor of Electrical Engineering, Princeton University.

in all countries was carried on in secret, ostensibly because of the military value of the copious neutron source which would be one of the by-products of a successful fusion reactor.

However, as work in all countries progressed, there came a gradual realization that this problem was more complex than had first been supposed, and that a solution, if indeed a solution did exist, might be many years in the future. It became apparent that the best efforts of many people would be needed to attack the problem, and that a duplication of effort among various groups isolated because of secrecy could not be afforded. As a result, all secrecy restrictions were removed several years ago and there is now a relatively free exchange of information among the different groups working in this field.

Despite this interchange, the work done so far has not yet established the technical feasibility of a controlled fusion reaction. Furthermore, should the research of the next decade or so indicate technical feasibility, the economic feasibility of controlled fusion would remain to be demonstrated. Thus, in spite of the copious and relatively cheap fuel supply, there is no guarantee that a controlled-fusion process could compete with other possible sources of energy.

Although the work of the past decade has not achieved the results that many people would have hoped, it has served to define more clearly the problems to be solved and to establish a firmer scientific base from which to carry on. For example, much has been learned about the behavior of plasmas at elevated temperatures, about some of the instabilities which can develop in these plasmas, of some possible methods for suppressing the growth of these instabilities, and about methods for diagnosing plasmas. However, much more basic scientific information about the behavior of plasmas must be learned.

At the moment, then, the outlook is that a power-producing fusion reactor is probably many decades away. As with all predictions, of course, some completely new approach to the problem may change this estimate considerably. Most of the present work is concerned with continued studies of the approaches proposed near the start of the fusion program. A new approach may produce success more quickly.

There are many reasons for coming to this somewhat pessimistic conclusion.* As just one sample of the problems involved, we might look briefly at that of containment.

---

* It should be pointed out that this is the personal opinion of the author who is not actively working in this field, but who has been following the work rather closely.

CONTAINMENT

From the start, containment of the hot fuel was assumed to be the central problem in controlled fusion. The cross sections for the various fusion reactions as functions of energy were well known, and it was tacitly assumed that there would be no real problem in giving the fuel particles sufficient energy to reach ignition temperatures provided the fuel could be confined for a long enough time.

The energies needed to reach ignition temperatures are in the range of 10,000 to 50,000 electron volts, and it is relatively easy to accelerate ions through 50 kilovolts to achieve these energies. Unfortunately, a process which competes with fusion — elastic scattering — is much more probable for particles of this energy. The elastic scattering cross section is about 10^8 times the fusion cross section. Thus, it is necessary to devise a system in which any one particle will have approximately 10^8 interactions with other particles before it leaves the system or before it loses its energy. Some type of reflector is needed to keep a particle in the reaction region after it undergoes an elastic collision. This is the crux of the containment problem.

Perhaps the one point of real agreement among people working in fusion is that the only practical reflector must be based on magnetic fields, since particles of either negative or positive charge tend to rotate about field lines. Other containment methods (for example, r-f fields) have been proposed from time to time, but these have not shown as much promise as magnetic fields. Here the agreement ends. Many different magnetic field configurations have been proposed, and the different groups studying controlled fusion have essentially been established to explore these different configurations.

Most of the configurations proposed thus far, and the devices which have been constructed, may be divided into two broad classes according to the source of the confining field. In one class are those devices in which the primary magnetic field is produced by the current in the plasma itself — the so-called pinch devices. In the second class are those devices which rely on a field produced by currents external to the plasma. Actually this division is not as clear cut as it might appear since many pinch devices include externally produced stabilizing fields, while the fields produced by the plasma currents in the second class of devices may have significant effects on the behavior of these devices.

The second class may be further subdivided into those devices with ends from which particles may escape the confining field and

those devices without ends. Devices with ends are usually called mirror machines, while those without ends are usually called stellarators.

Finally, the mirror machines may be divided into two more groups. One group includes those devices in which an appropriate density of cool fuel is placed in the reaction region and is subsequently heated to the ignition temperature (hopefully) while the concentration is maintained approximately constant. In the other group are those devices in which the fuel is injected from an external heating device into the reaction region to gradually increase the particle concentration to the point at which a net power output from the device is achieved (again, hopefully).

Other classifications are possible based on heating schemes or other characteristics. However, let us examine the containment problem in the machines as we have classified them.

Pinch Machines

The pinch machines as a class are basically pulsed devices. Containment, in principle, arises from a magnetic pressure equal to the magnetic energy density $B^2/2\mu_0$.† The magnetic flux density is, in turn, produced by the current in the plasma. At the edge of the current sheath, the flux density is $B = \mu_0 I/(2\pi a)$ where I is the current in the pinch and a is the radius of the pinch. To achieve a pinch, and hence containment, the magnetic pressure must exceed the kinetic pressure of the plasma, $2nkT$, where n is the concentration of either type of charge and T is the temperature of the plasma, assuming the electron and ion temperatures are the same. Thus for a pinch to occur,

$$I^2 > 16\pi^2 a^2 nkT/\mu_0 = 16\pi ANkT/\mu_0$$

$$I > 2 \times 10^{-8} \sqrt{ANT} \text{ amperes}$$

where A is the geometric area of the discharge tube and N is the initial gas concentration before the pinch occurs.

To get some idea of the order of magnitude of currents needed for a power-producing pinch, we must examine the generated power density. This power density is proportional to the square of the particle concentration and to the temperature. For a D-T reaction at a temperature of 10^8°K (approximately the ignition temperature for this mixture), a fuel concentration of $10^{21}/m^3$ gives a power density of 60 megawatts/meter^3; a concentration of 10^{23}/meter^3 gives

† The M.K.S. system of units is used throughout this paper.

a power density of 600,000 megawatts/meter^3, which is an extremely powerful explosive; while a concentration of 10^{19}/meter^3 gives a power density of only 6 kilowatts/meter^3, which is too small to be useful as a main power source. Assuming a concentration of 10^{21}/meter^3 at the ignition temperature of 10^{8}°K, the generated power per unit length of the reactor will be

$$P = 60\,\pi a^2 = 60\,AN/n = 6 \times 10^{-20}\,AN \text{ megawatts/meter.}$$

For a pinch to occur and, hence, for the fuel to be confined under these conditions, the plasma current must satisfy the relation

$$I > 10^6\,\sqrt{P} \text{ amperes.}$$

A conventional power station may have a capacity of 500 megawatts. If we want the same output from a pinch reactor, we can determine the current needed for pinching after we pick a length for the tube. The length is limited, since prohibitive voltages would be needed to initiate the discharge if the tube were too long. In a reactor 5 meters long, which is probably optimistic, the current needed for confinement is greater than 10^7 amperes. Continuous currents of this magnitude are beyond the range of present technology. Consequently, a continuous power producing pinch can not be achieved in practice.

In addition to this limitation, pinch devices have been plagued with various instabilities which destroy the confinement and which have prevented really significant levels of fusion from being achieved. It is hoped that methods will be found for preventing all these instabilities from developing, and, in fact, methods have been found for inhibiting several of them. Others have not succumbed to attack, however, and it is not certain that they will.

Despite this outlook, work with pinch devices is continuing. The main effort is aimed at increasing the rate at which the pinch current builds up so that fusion may occur in significant amounts before the instabilities have time to build up to intolerable levels.

Even should a successful fusion pinch be achieved, energy would be produced in explosive bursts. It is extremely doubtful if a reactor could be designed to withstand more than one of these bursts. Overall then, at the moment, the prospect that controlled fusion will be achieved on a practical scale with a pinch device is not promising.

External-Field Devices

Devices in the second broad class do not suffer from the current limitation of the pinch devices. The flux density inside a toroidal

solenoid is $B = \mu_0 n_t I$, where n_t is the number of turns per unit length of solenoid. To balance the kinetic pressure of the plasma

$$n_t I > \sqrt{2nkT/\mu_0} = 2 \times 10^6 \text{ ampere turns/meter}$$

for the same fuel concentration and temperature used for the pinch calculations. In principle, for a solenoid of 50 turns/meter, a current of 40,000 amperes is sufficient for containment. Continuous currents of this magnitude are within the range of present technology, so that this second class of devices could be operated continuously. However, no continuously operating device has yet been constructed. Although continuous containment is possible from the current standpoint, there are other effects which make the containment problem difficult.

Mirror Devices

The basis of mirror devices is that charged particles tend to be reflected back and forth between two regions of strong magnetic field next to a region of weaker field. If a particle has a sufficiently high ratio of transverse to axial velocity (determined by the ratio of maximum to minimum field intensity), it will be confined to the region between the two strong fields. If its axial velocity is too large, the particle will not be reflected but will leak out one of the ends. Actually, for a given ratio of maximum to minimum field intensity, there is an escape cone in velocity space. Any particle with a velocity in this cone will not be confined by the mirror. A particle which initially has a velocity outside the escape cone may be scattered into the escape cone by a collision. Hence it is not certain that a particle will be confined even though its initial velocity is outside the escape cone.

Particles injected into the system transversely to the main field direction will certainly have an initial velocity proper for confinement. However, most of the collisions between particles are elastic collisions so that the particles in the system soon take on a random velocity distribution. A significant fraction of the particles will then have velocities in the escape cone. The probability of a particle having a collision increases with the concentration of particles in the system. At the relatively high concentrations needed for appreciable generated power density, the end losses may be quite large. This is a serious problem in both classes of mirror devices.

If one starts with a high concentration of cool fuel and heats it in the reactor by magnetic compression or by some other means, it is the more energetic particles which escape from the ends. This makes it difficult to attain the high average energy, and, hence, the temperature needed for ignition. If one attempts to build up the

concentration of fuel particles to a significant level by injecting hot fuel, there is again appreciable leakage from the ends which tends to prevent the concentration from reaching the proper level.

Up to the present time, the end problem in mirror devices has not been solved satisfactorily. However, some of the most promising fusion experiments have been performed with mirror devices.

In most of the mirror devices constructed to date the magnetic field is concave toward the central axis. This field configuration is inherently unstable. A simple way to look at this instability is that magnetic field lines tend to become as short as possible, while the plasma tends to expand. Both these effects can be achieved if the plasma and the field exchange places so to speak. To overcome this instability, a new field configuration commonly called a "picket fence" configuration is being studied. In this configuration, the field lines are convex toward the central axis and this tends to keep the field outside the plasma. While this is a promising approach, there are still other problems associated with this configuration. As yet not enough work has been done with the picket fence arrangement to assess its future usefulness.

Stellarators

The end problem does not exist in the stellarator class of devices since there are no ends. However, there are other problems with containment which have not yet been solved.

The basic principle of these devices is that the magnetic field lines do not close upon themselves after one transit through the system as they do in a simple toroid. This effect is achieved either by twisting the toroid into a figure eight configuration or by wrapping a special set of helical windings, in addition to the main solenoid, on a race-track structure to produce an angular component to the field. As a result, a given field line generates a magnetic surface after repeated transits through the device.

The important consequence of this effect is that any charge separation which tends to be produced by the inhomogeneous field may be compensated by a current flow along the magnetic surface. This prevents the build up of large transverse electric fields in the device, which would immediately destroy the confinement.

There exists an instability peculiar to this class of devices. The plasma current produces an azimuthal magnetic field. If this angular component, in addition to that from the helical windings, is sufficient to produce a net rotation of 360° or 0° in a field line after one transit, the field line will close upon itself and the magnetic surface will not

be generated. It is no longer possible for counter-currents to flow to oppose the charge separation due to the inhomogeneous field. Hence, electric fields build up which destroy the confinement. To prevent this instability, the current must be kept below the so-called Kruskal limit. This means that the concentration of fuel particles must be kept relatively low. Since the generated power density depends upon the square of this concentration, the generated power density may be lower than one would wish.

Experimental work to date has confirmed the existence of this instability which was first predicted theoretically. In addition, this experimental work has given indications of other instabilities which prevent containment even before the Kruskal limit is reached. There are some tentative explanations for the origins of these instabilities, but these theories have not been confirmed experimentally.

In fact the whole question of instabilities in the stellarator has not been studied experimentally in a satisfactory way. A large experimental stellarator is now under construction to carry on these studies, as well as other studies of high-temperature plasmas. Until these studies have been made it will not be possible to assess the future of stellarator type devices as practical energy sources.

Other Problems

In addition to these and other difficulties with containment, there are difficulties in heating the fuel to its ignition temperature. As an example of the heating difficulties, there is the problem of "runaway" electrons. The fusion reaction is between the ions in the plasma, and it is the ions which must be heated to the ignition temperature. If the energy-input mechanism heats the ions, but not the electrons, interactions between the hot ions and cool electrons will result in an energy flow from the ions to the electrons. This will make it difficult to reach sufficiently high ion temperatures.

If on the other hand the input mechanism heats the electrons preferentially, energy must be transferred from the electrons to the ions. Collisions between electrons and ions are elastic. Because of the small electron-to-ion mass ratio, only a very small fraction of an electron's energy can be transferred to an ion during a collision. As the electron energy increases beyond a certain point, the probability of a collision with an ion decreases so that the coupling of energy between electrons and ions becomes weaker. As a consequence, it is possible for the electron temperature to become much higher than the ion temperature with relatively little energy flow from electrons

to ions. Thus, while energy is supplied to the plasma, most of it is given to the electrons and the ions may never reach the ignition temperature.

This is but one of many heating problems which must be solved before controlled fusion is realized.

Assuming that both the containment and heating problems are solved satisfactorily, the problem of just how to extract and utilize the fusion energy must be studied. To date very little effort has gone into this phase of the over-all problem, and reasonably so. People have speculated on different methods for utilizing the energy, including direct conversion to electrical energy. However, to the author's knowledge no really serious study has been undertaken.

Conclusion

The work done in fusion to date indicates the need for a much better understanding of the basic properties and the behavior of plasmas in general before any real progress toward practical controlled fusion can be realized. Consequently, one can expect to see a shift in emphasis in the fusion program from a concentration on various fusion devices to a concentration on basic plasma physics. Even should empirical work with some of the present devices demonstrate controlled fusion, a better understanding of basic plasma physics will be needed to optimize the process. One can discern that this shift in emphasis is already underway. Further, since practical controlled fusion may be many decades away, it is imperative that other new energy sources be studied vigorously. Fortunately this study is already underway.

POWER SUPPLY FOR THE TIROS I METEOROLOGICAL SATELLITE*†‡

By

SEYMOUR H. WINKLER, IRVING STEIN, AND PAUL WIENER

RCA Astro-Electronics Division,
Princeton, New Jersey

Summary—*The power supply system employed in the TIROS I satellite is discussed systemwise and in detail. First, the influence of the system performance requirements on power-supply design and thermal control methods is shown. Detailed discussion covers the problems of thermal control, solar-collector orientation with respect to the sun vector, the method of mounting solar cells to the vehicle surface, the electrical arrangement of the solar-cell circuitry, the solar-cell testing program, and the design factors employed in determining the size and shape of the solar collector. The storage battery is discussed with references to the battery testing program, electrical characteristics of the battery, and the basic factors which governed the battery design. Protective and regulatory circuit functions and specifications are given. Data is given on the performance in orbit of the solar cells, batteries, and thermal control. In addition there is a discussion of the factors affecting spatial orientation of the TIROS I satellite in orbit. Graphical evidence is provided showing how telemetered solar-cell data compared with computed data on the satellite-to-sun orientation.*

INTRODUCTION

EARLY SYSTEMS CONCEPTS for the TIROS I meteorological satellite were much simpler than those incorporated in the vehicle launched on April 1, 1960. The basic function of the satellite was to take television pictures of the earth's cloud cover and relay them back to earth.

Initially it was decided to have no unfolding, latching, or extending mechanisms on this early satellite. The solar-cell power supply, therefore, had to be part of the fixed skin which covered the vehicle instrumentation. It was also decided that the satellite should be cylindrical in shape and spin-stabilized with a disc-shaped mass dis-

* Developed under the auspices of the National Aeronautics and Space Administration and under the Technical Direction of the U.S. Army Signal Corps, Contract No. DA 36-0390SCO78902.

† Presented to the American Rocket Society, Sept. 27-30, 1960.

‡ Manuscript received 9 January 1961.

tribution, so that it would continue to spin about its figure axis. Furthermore, all thermal control[1] was to be passive and was to be achieved by the selection of proper materials to obtain optimum thermal characteristics for the fixed outer surface, and by the thermal inertia and coupling of the over-all system.

As the system concept became more involved and additional power-consuming equipments were added, it became evident that the maximum envelope size permitted by the launching rocket would determine the total solar-cell power capability. The design of the power supply[2] was complicated by the fact that the output power of a tight-fitting solar-cell collector is a function of area, while equipment power requirements are roughly a function of volume. As equipments were added, the collector surface began to enclose a progressively greater volume than that of the equipment it was to power. The total vehicle equipment power requirements increased until the solar collector required an area which enclosed all the volume that the nose shroud of the launching vehicle would allow.

An earlier decision on the TIROS I operational regime had fortunately lifted the restraints on the addition of power-consuming hardware. This decision allowed for a portion of the equipments, such as cameras and tape recorders, to be programmed from the ground for immediate or delayed execution of commands, and for variable durations, making it possible to obtain an infinite number of power-versus-time profiles and thus permitting the power-supply output to vary considerably. The result of these conditions was to place restrictions on the duty cycle of the instrumentation rather than on the functions it performed. It was also decided that the 300 pictures per day (associated with the 20-watt orbit-averaged power from the size-limited cylindrical collector and storage system) represented a sufficient amount of information per day to be meaningful to meteorologists.

Preliminary Considerations

The general location of the solar cells on the TIROS "hat" was determined by the optimum position of the cameras for picture taking. (The cameras can take good pictures only when the sun is behind them and when they are pointed toward the earth.) With both camera

---

[1] M. Ritter, "Thermal Design for TIROS," *Astronautics*, Vol. 5, No. 7, p. 40, June 1960.

[2] C. C. Osgood and S. H. Winkler, "Optimizing the Design of a Solar Power Supply System," American Astronautical Society Meeting, January 18-20, 1960.

axes parallel to the spin axis and both pointing in the same direction, the solar cells should cover the flat top surface of the cylinder behind the cameras. Since good pictures of the earth can also be obtained when the sun's rays are considerably off-parallel with the camera axes, it became necessary to cover the lateral surface of the cylinder with solar cells, in order to maintain a less-fluctuating projected area. There was no need to cover the bottom flat surface with solar cells, because if they were illuminated, the sun would be looking into the cameras, rendering the satellite temporarily inoperative.

Because the solar cells cover a great proportion of the top and sides of the cylinder, their thermal surface properties (solar absorptivity and infrared emissivity) heavily influence the mean temperature of the components within. Thermal analysis determined a required value of α/ϵ ratio of around 1.0 with ϵ of 0.85. Three parallel development programs were started to develop coatings for the solar cells, (whose α/ϵ ratio when uncoated is between 2 and 3) to bring the emissivity from around 0.3 up and equal to their absorptivity of 0.9. Transparent plastics, vacuum-deposited inorganics, and bonded transparent covers were all investigated. The use of organic plastics by themselves was dropped for radiation and vacuum-sensitivity reasons. A development program with a leading optical firm had not yielded either sufficient sensitive-region transparency or high enough emissivity in time for the decision on coating materials. The result was the selection of the now-familiar, bonded Microsheet* glass platelet coating, using a vacuum-deposited blue reflector coating on the glass to protect the transparent epoxy adhesive from darkening in the ultra-violet and near ultra-violet, as well as an anti-reflective coating to maximize solar transmission at the cell's peak response wavelength of 0.8 micron. Some time after this coating had been selected, the vacuum-deposition program yielded a direct coating of SiO_2–Si_2O_3 on the solar cells, which dropped the electrical output power of the solar cells no more than 3 per cent, while raising the emissivity to between 0.8 and 0.85. The blue reflector coating has the advantage over the direct-emissivity coatings of permitting some variation in net absorptivity, with the fixed emissivity of the glass. On the other hand, the direct-emissivity coating permits a variable value of ϵ as a function of thickness, and can be augmented with a multi-layer blue reflector which gives the flexibility of varying both α and ϵ, and permits darkside as well as illuminated-side thermal

* Registered trade mark.

control. The thermal performance of TIROS I in orbit was within 5°C of the temperatures predicted for the actual launch condition.

The silicon solar cells are located on the top and lateral surfaces of the cylindrically shaped TIROS I structure (see Figure 1); the total instantaneous active silicon area projected normal to the sun's rays determines the power-collecting capabilities of the satellite. This area is expressed in terms of the silicon area on the top and lateral

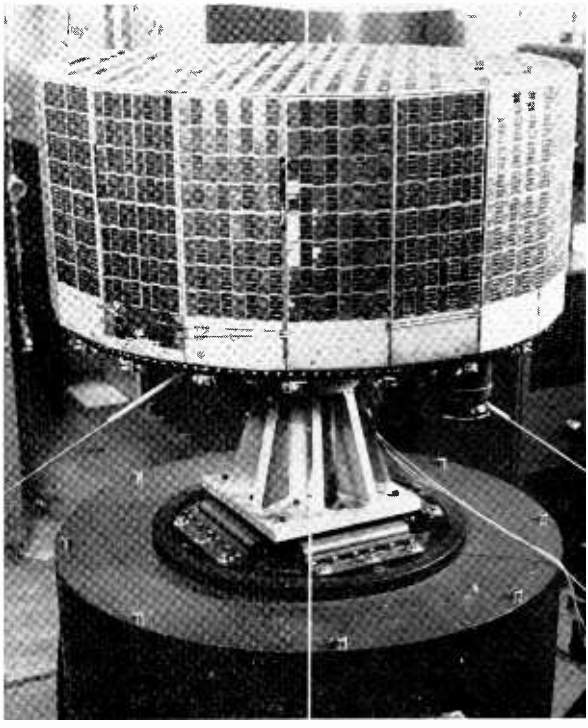


Fig. 1—The TIROS I satellite showing solar-cell location.

surfaces and the angle between the sun's rays and the satellite spin axis, denoted as the angle alpha. Thus, if alpha is zero, only the top is illuminated, while at an alpha value of 90° only the lateral area is illuminated. Maximum instantaneous solar-cell output is derived at alpha equal to about 45°.

The original concept of the TIROS I system was to place in orbit a spin-stabilized satellite whose spin axis orientation remained essentially fixed in inertial space as the satellite orbited the earth and as

the orbit precessed in space.[3,4] Prelaunch calculations demonstrated that the effect of differential gravity would perturb the spin axis only slightly. Further investigation indicated that eddy currents in the satellite structure and the presence of soft iron in the satellite would similarly affect the orientation. Under these conditions, the major variation in alpha is attributed to the relative motion of the sun and spin axis of the satellite as the earth travels in its orbit about the sun. For the pre-selected spin axis orientation, the time of year and time of day of launch were to be chosen so that during the initial three months of satellite life, alpha would be favorable for television coverage of the earth and for power collection and conversion.

Immediately after launch, pictorial and attitude information obtained from TIROS I showed a greater perturbation in alpha than had been predicted. The perturbation was caused by the interaction of the earth's magnetic field and the magnetic fields set up by current loops in the maze of satellite wiring. Calculations were performed to determine the magnitude of this magnetic dipole of the satellite and the results employed to predict future variations in alpha. Although alpha varied in a manner different than originally predicted, its magnitude was such that illumination conditions of the earth and the spin axis orientation remained favorable for television coverage and solar energy collection during the useful life of the satellite. Additional information concerning alpha is presented below in the discussion of the post-launch performance of TIROS I.

Silicon Solar-Cell Energy Converter

The basic power supply of the TIROS I consists of a silicon solar-cell energy converter, a storage battery, a charge current limiter, and a voltage regulator. The energy converter is an assemblage of approximately 9200 silicon solar cells which converts solar radiation energy into electrical energy during orbital daylight. The solar-cell output divides between supplying the equipment loads and charging the storage battery. When peak equipment loads exceed the solar-cell capacity, all of the output of the solar cells is supplied to the load and none is supplied to the batteries for storage. During the orbital night, when the solar cells are passive, silicon diodes in each series

---

[3] W. R. Bandeen and W. P. Manger, "Angular Motion of the Spin Axis of the TIROS I Meteorological Satellite Due to Magnetic and Gravitational Torques," *Journal of Geophysical Research*, Vol. 65, No. 9, p. 2992, September 1960.

[4] H. Perkel, "TIROS I Spin Stabilization," *Astronautics*, Vol. 5, No. 7, p. 38, June 1960.

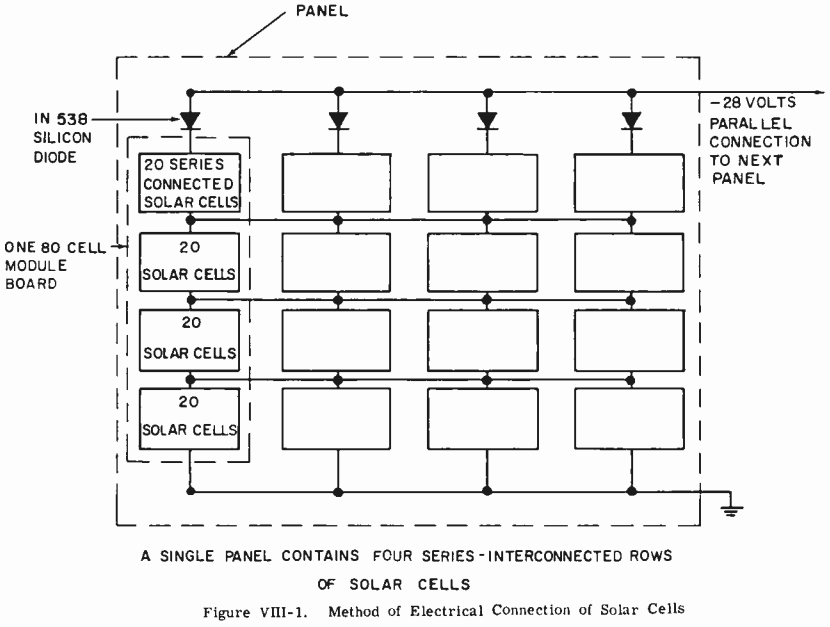


Figure VIII-1. Method of Electrical Connection of Solar Cells

Fig. 2—Method of electrical connection of solar cells.

row of solar cells (see Figure 2) prevent the storage batteries from discharging into the solar cells. Similarly, diodes are provided for each series row of solar cells located on the lateral surface of the satellite. As the satellite spins about its figure axis and the solar-cell rows are alternately illuminated and darkened, these diodes also serve to prevent the rows which are darkened at any given time from loading those which are illuminated. Table I summarizes some of the more important design features of the solar array.

The solar converter includes solar cells, bonded-glass covers, adhesives, standoffs, diodes, wiring, and module boards, but excludes the aluminum surface which was part of the structure. The weight of the converter is 24.5 pounds.

The side and top surfaces of the satellite were almost completely covered with solar cells to take full advantage of the variation of

Table I

	Top	Side	Total
Number of Solar Cells	3560	5560	9120
Square feet of active silicon area	6.90	10.8	17.7

orientation of the spin axis and the sun vector. The solar cells used were manufactured by the International Rectifier Corporation. They are 1×2 centimeter boron-doped silicon, assembled in a shingle form. The shingle comprises five of these individual cells in series. A minimum conversion efficiency of 7.5 per cent at a voltage of 1.95 volts plus or minus 5 per cent at a temperature of 30°C is specified for the basic shingle. A transparent coating, which is bonded to each individual cell of the five-cell shingle, greatly improves the thermal emissivity. This coating permits radiative self-cooling in the spectral region corresponding to the desired cell temperature. It consists of a platelet of Microsheet glass 0.006-inch thick, which is coated with an optical vacuum-deposited layer. The upper, or outer surface, has an anti-reflective coating which permits the maximum transmission of light in the 800-millimicron range, where the solar cells are most responsive. The inner, or lower surface, has a 15-layer, sharp-cutoff reflector coating which reflects or absorbs light of all wavelengths shorter than 450 millimicrons and transmits 90 to 95 per cent of the wavelengths over 450 millimicrons and beyond the solar cells' upper response limit of 1100 millimicrons. The platelets are bonded to the individual cells with a transparent epoxy adhesive.

Two small copper tabs, one on the positive strip of the first cell, and one on the negative underside of the fifth cell were used to enable simple and reliable interconnections of the shingles. A flat, compact epoxy fiberglass board was used as the basic building block to simplify the handling and assembly techniques. This board, referred to hereafter as a module, had imbedded copper wiring or printed circuitry so that the solar-cell shingles mounted on the module could be series connected. Two module boards with shingles mounted are shown in Figure 3. For the operating temperature predicted and the load voltage required, 80 cells, or sixteen series-connected five-cell shingles, were needed.

The sixteen solar-cell shingles are mounted on the module boards with a low-temperature cure, scrim-cloth epoxy adhesive of uniform thickness. Shallow wells and locating holes accurately position the shingles on the module board. The copper tabs of the bonded shingles are carefully soldered to the module printed circuitry. In addition, paralleling tabs are provided in the module so that every series set of four shingles may be paralleled. A module is 3.4×7.5 inches, weighs 80 grams, and has an electrical output of 32 to 38 milliamperes at 28 to 32 volts at a solar intensity of 100 mw/cm^2 normal incidence.

The module boards are paralleled in clusters of four, and are electrically connected to one another at the one-quarter, one-half,

and three-quarter voltage points, corresponding to 20, 40, 60, and 80 cells in series. This method of interconnection tends to minimize the loss of solar converter output in the event of failure resulting from open or short circuits in one shingle.

The top surface of TIROS I was very efficiently packed with these modules. The modules were bonded directly to the vehicle skin in electrical groups of four with the aforementioned scrim-cloth epoxy

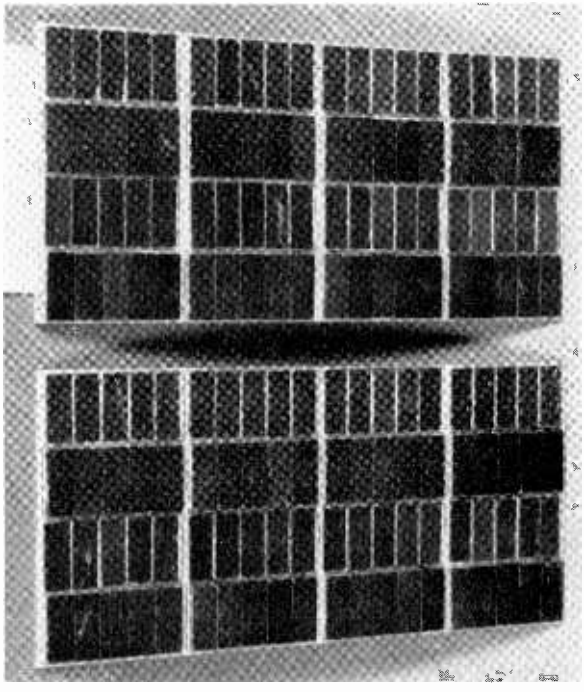


Fig. 3—Module boards with shingles mounted.

adhesive. To simplify mounting of the modules, the lateral area of the satellite was modified from a continuous curved surface to eighteen flat panels. Four module boards were epoxy bonded to each aluminum panel, and these panels were then bolted to the aluminum vehicle frame.

The solar-cell testing program was quite extensive. An indoor light-testing fixture was built that permitted testing of great quantities of five-cell shingles. At a controlled temperature and uniform light intensity, the following pertinent characteristics of the shingles were recorded: short-circuit current, open-circuit voltage, and conversion efficiency at various voltages. With this information, the

shingles were catalogued and sorted. By judicious matching of these characteristics, each shingle was assigned to a specific module. The assembled modules were then light tested and the same characteristics recorded as for the individual shingles. Again, the module characteristics determined the proper matching for the panel assembly. Following the final bonding operation of module to panel, another light test was made. After the wiring operation (paralleling every four modules and paralleling every four shingles on each of the four modules), an outdoor sun test was performed. It is interesting to note here that by this extensive indoor testing program and subsequent careful matching of the solar-cell characteristics, the final value of solar-cell-collector conversion efficiency was very nearly equal to the conversion efficiency of the individual shingles.

Calculations were made of the energy available to supply the load demands on a day-by-day basis, so that the programming rate of the vehicle could be closely controlled. Data was available on the variation of the angle between the sun vector and the vehicle spin axis for the specific day and hour of launch. Calculated data was also available for the top- and side-temperature variations on a day-by-day basis. The following additional factors were taken into consideration in computing the total energy available from the power supply on a daily basis:

a) Radiation spectral shift at satellite altitude. In going from a standard-day, sea-level solar spectrum to a satellite-altitude spectrum, a spectral shift toward the shorter wavelengths is encountered. Because of this shift, the solar-cell responsive wavelength band contains a smaller proportion of the total solar flux and a poorer distribution of this flux at satellite altitudes than at sea-level. At the satellite altitude, therefore, solar-cell efficiencies are reduced.

b) Time-degradation effect on output due to micrometeorities.

c) Loss in transmission through the cover glass of the solar cell.

d) Loss due to the presence of the protective diodes in the solar-cell circuit.

e) Effect upon the solar-cell output of the non-normal incidence of solar radiation on solar-cell surfaces.

f) Loss in output caused by the shift in the operating voltage off the optimum voltage for the specific design temperature. This shift is due to (1) the battery voltage fluctuation, which varies from a higher voltage during charge to a lower voltage during discharge; and (2) the fact that the solar cells are

not at the constant temperature for which they are designed, but rather experience a temperature range as the orientation of the vehicle changes with respect to the sun vector.

g) Losses introduced by series-parallel connections of the solar cells.

h) Losses due to the inefficiency of the energy-storage system, the voltage regulators, and current limiters.

The instantaneous power output of the solar collector is a function of the angle alpha as described earlier. As the angle alpha changes, the total silicon area projected normal to the sun changes, and, therefore, the solar-cell temperature also changes. As alpha increases, the

Table II

Alpha	Instantaneous Power (watts)	Average Solar Collector Power Output (watts)	Remarks
0	35	23	Side solar cells not illuminated.
45	51	33	Maximum collector output.
90	25	16	Top solar cells not illuminated.

top area projected normal to the sun decreases, but the average solar-cell temperature also decreases. Simultaneously, the side area projected normal to the sun increases as does its average operation temperature. The combined effect of these area and temperature changes is shown in Table II in terms of the power output of the solar collector corresponding to the average operation temperatures at each value of alpha.

The average power output of the solar collector is obtained by multiplying the instantaneous power output by the ratio of illuminated time to the duration of orbital period ("per cent sun time"). This ratio is 0.65. The average power to the load is approximately 0.68 of the average output of the solar collector, because of losses which occur in the energy-storage system, voltage regulators, and current limiters.

On the basis of the aforementioned degradation and correction factors, an average daily energy in watt-minutes available to the load was calculated for a given solar-cell conversion efficiency and exposed area. With this information, a daily power budget was submitted to the operational team that programmed the satellite.

Storage-Battery Supply[5]

The storage battery provides a relatively constant voltage to the solar cells thereby isolating them from variations in the electrical load. The battery is charged by the solar cells during the orbital day and supplies all equipment loads during orbital night. During the orbital day, the battery also supplies the difference between peak equipment power drains and the solar-cell output by the amount that the solar-cell capacity is exceeded by equipment power demands.

At the outset of the TIROS I program, the nickel-cadmium battery[6] received primary consideration as the energy-storage device because its trouble-free cycle life is potentially greater than that of other storage-battery systems. After much effort in comparing commercially available nickel-cadmium storage cells, the glass-to-metal hermetically sealed 4-ampere-hour Sonotone* storage cell was selected. Its sealed construction obviates problems associated with operation in a vacuum environment and also prevents gas escapage to the outside environment, which would result in loss of electrolyte and a reduction in cell-energy capacity. The rolled plate and separator construction permits packaging of the active cell materials in a cylindrical container, which is a better pressure vessel shape than the rectangular container.

During the TIROS I program, over 1000 such storage cells were subjected to thermal, thermal-vacuum, vibration, acceleration, and shock environment tests. Full charge–discharge tests were run at 0°C, 25°C, and 60°C with constant-current charging and discharging. Various charging rates were employed to determine the safe maximum charge rate at each temperature. Finally, partial charge–discharge or cycling tests were performed between —10°C and +60°C under vacuum conditions in a simulation of the charging and loading conditions to be experienced in actual orbit. This extensive testing and evaluation program permitted acquisition of considerable data concerning the capacity, voltage, temperature, efficiency, and life characteristics of the nickel-cadmium storage cell. When a defect was discovered in the cell operation, its cause was jointly investigated by

---

[5] S. H. Winkler, "Optimum Design of a Space Vehicle Storage System," *14th Annual Power Sources Conference Proceedings*, May 17-18-19, 1960.

[6] P. F. Bruins and A. J. Salkind, "Manufacture of Sintered Plate Nickel-Cadmium Batteries," Sandra Corp. Project Report Ref. Symbol 5311 (99), July 2, 1956. Department of Chemical Engineering, Polytechnic Institute of Brooklyn.

* Registered trade mark

RCA, Sonotone,[7] and the U. S. Army Signal Corps.[8] A redesigned version of the cell was immediately produced by Sonotone and tested by RCA. The result was that over 1,000 trouble-free charge–discharge cycles were obtained from this battery during the useful life of TIROS I.

The TIROS I storage battery is designed to provide a minimum terminal voltage of 25.5 volts throughout periods of peak power drains over a battery-operating temperature range of —10°C to +55°C. Three rows of batteries are connected in parallel. Each contains 21, 4-ampere-hour cells in series. Thus, the total nominal capacity ratings of the battery are 12 ampere-hours and 300 watt-hours. The completed battery package includes the type 63F cells, packaging, wiring, and aluminum loaded epoxy (which is used for heat sinking the storage cells). It weighs 40 pounds.

In the case of a satellite such as TIROS I, which relies on the sun as its primary energy source and where the ratio of illuminated to dark time is approximately 2:1, the total energy that can be removed from the storage battery during one orbit is not limited by its total energy capacity, but by the maximum charging rate the battery can safely accept in the relatively short time available to restore the battery to full charge. The restriction placed on the maximum charging current limits the total energy removed from each row of the battery in an orbit to about 10 per cent of its rated capacity. Therefore, three rows of batteries are required to insure that sufficient battery energy is available to support the vehicle equipment programming rates demanded by the satellite mission.

The operation of much of the TIROS I instrumentation is subject to ground control. Thus, a means is available to determine from the ground the total instrumentation energy requirements in a given orbit. During orbits in which ground illumination is unsatisfactory for taking television pictures or the satellite is out of communication range of the two ground stations, little or no equipment was programmed and the major power drain was the continuous load. During these orbits, then, considerably less than 10 per cent of the battery energy is removed. However, in the nine-to-ten orbits per day when television coverage is possible and the vehicle is ground commanded to perform its full range of functions, the average battery drain exceeds 10 per cent, but the maximum drain is limited to about 20 per

[7] Private Communication, L. Belove and H. Fields, Sonotone Battery Corporation, Elmsford, N. Y.

[8] Private Communication, P. Rappaport, USASRDL, Fort Monmouth, New Jersey.

cent of total battery capacity. Over the course of each 24-hour period during the useful life of TIROS I, the total load requirements were made to balance the calculated energy available from the power supply for that period. This insured complete battery recharge at the end of each 24-hour cycle.

The primary reason for limiting the recharge rate is to prevent excessive internal heating of the battery and excessive rates of gas production, both of which can cause material destruction of the battery. As described above, this limitation restricts the total quantity of energy that can be discharged from the battery in an orbit. This last limit on battery performance achieves the desirable results of extending battery cycle life, which appears to vary inversely as the depth of discharge, and insures that the battery terminal voltage characteristics will be maintained during peak loading periods.

Charge-current limiters prevent the charging current to each row of storage batteries from exceeding the maximum rate which the batteries can safely accept. Silicon diodes are used to isolate the three rows of storage batteries from one another so that short-circuits in one row will not affect the remaining rows.

Power is supplied to the vehicle loads directly from the unregulated battery voltage, via two 24.5-volt and two 13.0-volt d-c voltage regulators. The output voltage variation of the voltage regulators remains within $\pm$ 1 per cent over an operating temperature range of $-10°C$ to $+55°C$.

Telemetered Post-Launch Performance

To ascertain post-launch power-supply performance, the following parameters were telemetered to earth from the satellite:

1. The terminal voltage of each of three rows of batteries.
2. The voltage supplied to the input of the voltage regulators by the three rows of batteries, each supplying power to the regulator through isolating diodes.
3. The output voltage of each of two groups of solar cells electrically isolated from the solar-cell power supply and each other, but operating into accurately known resistance loads.
4. The temperature of the satellite skin immediately behind each of the solar-cell groups described above.
5. The temperature of the satellite baseplate, which supports most of the satellite components and mass, adjacent to the battery package.

6. The output voltage of two 24.5-volt and two 13.0-volt d-c voltage regulators.

Each of the above temperatures and voltages are transmitted to the ground by two telemetry channels operating independently, insuring that telemetered data is available in the event one telemetry system fails.

In the telemetry system, a 40-position switch sequentially samples each of 39 parameters which are instantaneously transmitted to earth. (There is no data storage for telemetry purposes.) Telemetry readings are automatically timed to occur when much of the satellite equipment is operative, providing a maximum of telemetered information concerning equipment performance. Additional telemetry samplings can be commanded from the ground as required.

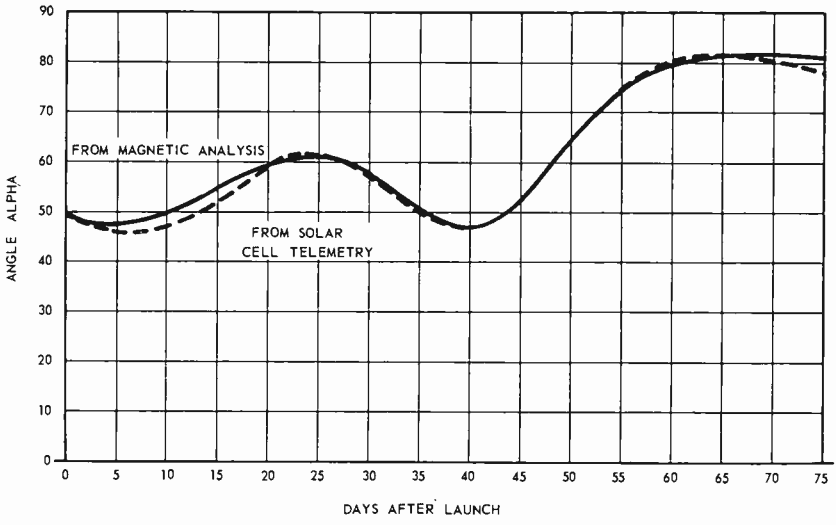


Fig. 4—Angle alpha versus days after launch.

Solar-Cell Performance

Prior to launching, each of the aforementioned solar-cell telemetry groups was calibrated to provide a plot of output voltage across a fixed load resistance versus solar-cell temperature between −5 and +70°C. Data was also obtained relating the solar-cell output voltage to the illumination angle of incidence.

The telemetered solar-cell voltage and temperature and the calibration data were employed to calculate alpha, which was previously defined as the angle between the satellite spin axis and the sun's rays. This value of alpha is compared day by day in Figure 4, with the

alpha calculated from the analysis of (1) the interaction of the earth's magnetic field and the magnetic dipole which characterizes the TIROS satellite, and (2) the effect of differential gravity upon the satellite.

Throughout the period shown the difference between these alpha curves is less than 5 degrees. Employing the alpha variation derived from the magnetic dipole and differential gravity analysis as a reference curve, one deduces that little or no degradation in solar-cell output has occurred because of the essentially constant difference between the alpha curves. Thus, the effects of particle radiations and micrometeorites upon solar-cell performance are negligible over a three-month period for an orbit at an altitude of 380 nautical miles at an inclination of about 51° to the equator.

Storage-Battery Performance

From the above, it is observed that the battery terminal voltages, the input voltage to the voltage regulators, and all the regulated voltages are sampled during peak loading periods (when the battery voltages are at their lowest levels). A great number of telemetry samples occurred when the solar cells were sharing the load with the batteries, but much data was also recorded during orbital night when the batteries supplied all the energy. Throughout the entire life of TIROS I, these battery terminal voltage readings remained between the designed limits of 25.5 and 32.0 volts. The input to the voltage regulator rarely fell below 25.8 volts.

The failure which occurred in the TIROS system and resulted in the NASA decision to halt all interrogations was not caused by a battery failure. The last telemetered battery voltages received were in the above-mentioned normal operating range. The equipment failure, which was positively identifiable, resulted in uncontrolled prolonged loads upon the power supply and depleted battery capacity. Limited operation of the remaining operational equipment is still possible if the batteries are given a sufficient opportunity to recharge.

Temperature Performance

It is noted above that the thermal performance of TIROS I was within 5°C of the predicted temperatures. The temperature of the solar cells located on the top surface varied from a maximum of 60°C for an alpha of 45° to a minimum of 0°C for alpha of about 80°. The temperature range of the solar cells located on the side surface, which is alternately illuminated and darkened, was about 0°C for alpha of 45° and 40°C for alpha near 80°. The temperature values quoted are those which occurred near the end of the orbital day and so represent

the maximum temperatures for the given spin axis orientation with respect to the sun. From the data available on the baseplate temperatures, it has been calculated that the battery temperature ranged from a minimum of about 0°C to a maximum of 30°C during the three-month period.

Programming of Operational Functions

The operation of the TIROS satellite was planned and coordinated at the NASA Goddard Space Flight Center in Washington, D. C. For each orbit, the specific programs of equipment operation were calculated and then relayed to the ground stations. The ground rules employed at the control center to determine the satellite programs were established by pre- and post-launch calculations of the daily energy capabilities of the power supply. The daily average power required by the equipment loads varied between a minimum of 15 watts and a maximum of about 25 watts.

Conclusion

On the basis of the post-launch data, it may be said that the various elements of the TIROS I power supply performed within design specifications and that the accuracy of the power-supply design was substantiated as planned pre-launch programming requirements were met or exceeded throughout most of the satellite's useful life.

BEACON TRANSMITTERS AND POWER SUPPLY FOR ECHO I*†

By

J. G. McCubbin‡ and H. B. Goldberg#

Summary—*"Echo I" is a 100-foot-diameter passive communication satellite designed to reflect radio waves. A special beacon transmitter and power supply were designed for attachment to the external surface of this satellite. The beacon permits tracking of the satellite and assists in establishing ephemeris data. A design was required which would be compatible with the folding and packing procedures for the satellite and which would operate reliably in space environment after withstanding the rigors of launching. The transmitter, which produces an output of 5 to 10 milliwatts at 107.94 megacycles, is mounted along with silicon solar cells, batteries, and antenna on a 10-inch diameter printed circuit board. The package is 0.35 inch thick and weighs 11 ounces.*

Introduction

ON AUGUST 12, 1960, the National Aeronautics and Space Administration (NASA) launched the first 100-foot-diameter passive communications satellite for Project Echo. This satellite was placed in a terrestial orbit at an altitude of 1,000 miles. It is a hollow sphere constructed from gores of 0.0005-inch thick Mylar** with the external surfaces coated with vapor-deposited aluminum to provide efficient radio-wave reflectivity. Before launch, the satellite, which weighed only about 160 pounds, was evacuated and folded into a spherical container 28 inches in diameter. When this container was placed in orbit it was explosively separated at its equator, and the satellite was inflated by the small amount of residual air entrapped within its interior. The spherical configuration was maintained by a small gas pressure derived from sublimating solids carried in the satellite.

This paper describes the two beacon transmitters and power supply which are attached to the external surface of the satellite. Each of

* Manuscript received 4 November 1960.

† This paper is based on work done under the auspices of the National Aeronautics and Space Administration, Contract NASW-97.

‡ Formerly RCA Astro-Electronics Division, Princeton, N. J.; presently General Electric Co., Philadelphia, Pa.

#RCA Astro-Electronics Division, Princeton, N. J.

** Registered trade mark.

these continuous-wave transmitters was to radiate a minimum of 5 milliwatts of power, at a frequency of 107.94 megacycles, to assist in establishing preliminary ephemeris data. A design was required which would be compatible with the folding and packing procedures for the satellite and which would operate reliably in the space environment after withstanding the rigors of launching. Since there are no means of access or connection to the satellite in its packing container, the system had to be complete within itself and to begin to operate automatically when placed in orbit.

The time a satellite is in the sun in a 1,000-mile altitude orbit may vary from 70 to 100 percent as the angle between the plane of the orbit and sun vector varies. Because the temperature of the attachments may vary over a wide range as the satellite traverses through sun and shadow periods in its orbit, some means of temperature control is required to assure that the components will operate within tolerable limits.

System Design

A solar-cell and storage-battery power supply was selected to power the two beacon transmitters. Each of the two assemblies, located diametrically opposite on the equator of the satellite, includes one of the two transmitters, its associated antenna, a group of solar cells, and one half of the storage-battery part of the power supply. Two lengths of printed cable, which are redundant for reliability, connect the assemblies, with each length traversing opposite halves of the circumference of the satellite in a direction passing over one of the poles. These conductors interconnect the storage-battery elements and place the two groups of solar cells in parallel with the storage battery. The system elements and their interconnections are shown in Figure 1.

Each solar-cell group, when it receives adequate solar radiation, will power both transmitters and charge the storage battery simultaneously. The partition of the system into two assemblies, located diametrically opposite, provides a balanced mass distribution, facility of installation, simple electrical interconnections, and optimum opportunity for collection of solar radiation by either of the two solar-cell groups.

The antenna for each transmitter is an erecting, quarter-wave, monopole. A normally open miniature switch is associated with each antenna. Either switch connects the storage battery into the circuit upon antenna erection, preventing depletion of storage-battery charge while the satellite is stored in its container.

If either of the two solar-cell groups receives adequate solar radia-

tion, operation of the transmitters is assured. The storage battery supplies the necessary power in darkness or in unfavorable positions of the sun with respect to the solar-cell groups. Discharge of the storage battery through the solar cells during dark periods is prevented by two semiconductor diodes.

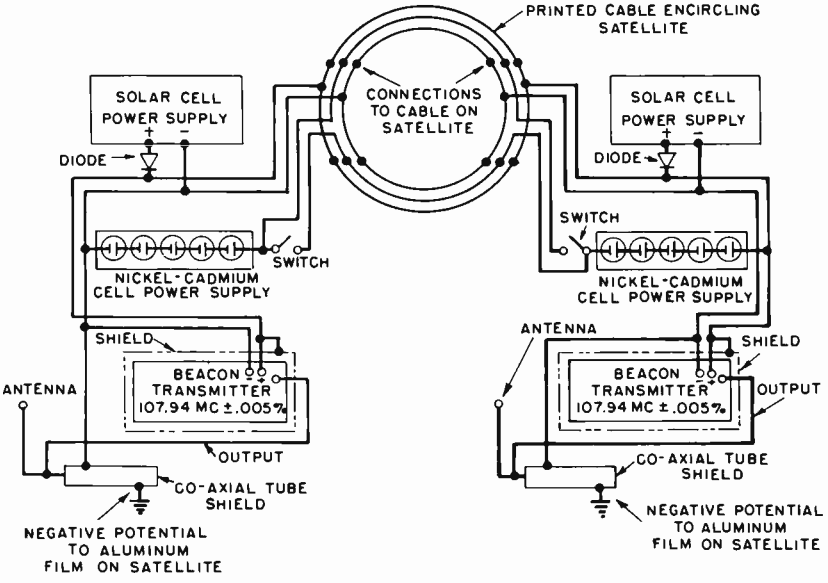


Fig. 1—Block diagram of the transmitter and power supply system.

TRANSMITTERS

The two beacon transmitters are identically constructed, crystal-controlled, continuous-wave units with the same nominal frequency. (The crystals actually chosen provide a frequency separation of 500 to 1000 cycles between the two transmitted signals.) Each transmitter consists of two 2N1143 mesa transistors and associated components mounted on a 1.4 by 2.0-inch glass epoxy printed-circuit board. A schematic diagram of the transmitter is shown in Figure 2. Transistor Q_1, used in the first stage, is a grounded-base, crystal-controlled, oscillator. The second stage transistor, Q_2, is used as an amplifier and also to isolate the oscillator from the antenna. Hand-wound, ferrite-tuned transformers are used for coupling between stages and to couple the antenna to the transmitter.

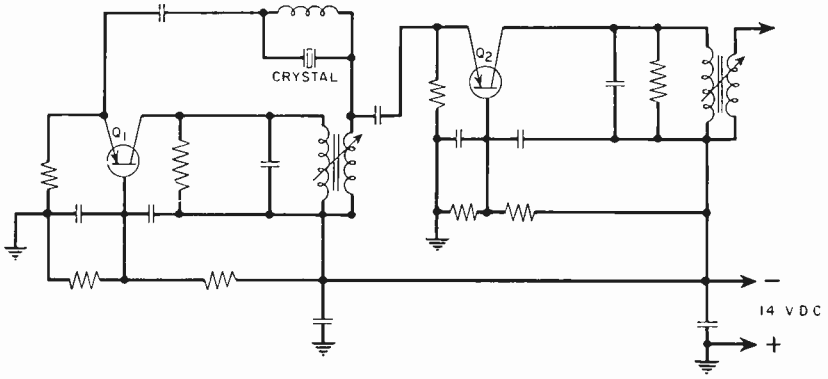


Fig. 2—Schematic diagram of beacon transmitter for inflatable satellite.

Figure 3 shows the oscillator frequency and output power of a beacon transmitter as a function of temperature. The antenna was simulated with a 40-ohm load, and the collector voltage was maintained constant at 14 volts to obtain these data. It is seen that a frequency drift of about ± 3 kilocycles from the center frequency occurs over the range from —5 to 55°C. The power output is from 10.3 to 10.6 milliwatts and is constant to better than ± 1 decibel over the same range. The power input is 100 milliwatts.

Figure 4 shows the output power as a function of collector voltage at 21°C. Collector voltage can vary from 15 volts with maximum solar cell output to 12 volts when the storage battery, in a discharged state, is supplying the power. This voltage range gives an output power variation from 13 to 6.5 milliwatts.

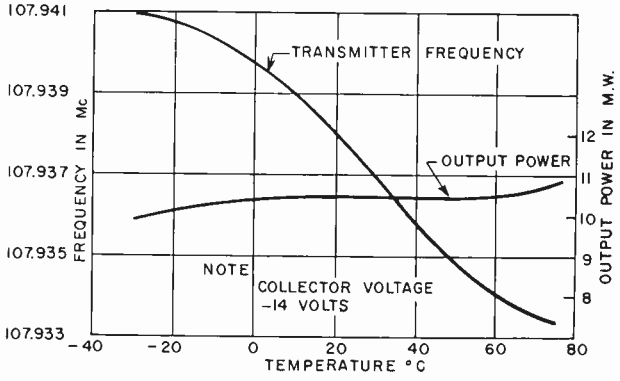


Fig. 3—Transmitter frequency and output power as a function of temperature.

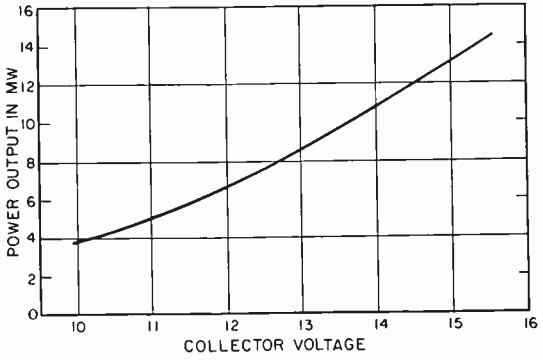


Fig. 4—Transmitter output power versus collector voltage at 25°C.

After the transmitter was assembled and tested, it was precision-molded with polyurethane foam, and 0.001-inch thick brass shields were bonded to the top and bottom of the unit. Figure 5 shows a transmitter before and after foam molding. (The aluminum tubes shown protruding from the unit provide access to the tuning elements after the transmitter is foam-molded in the complete assembly.) Precision molding provides resistance to shock and vibration and gives a repeatable range of intercomponent capacitance. Each transmitter, after foaming, weighs less than one-half ounce.

ANTENNAS

The quarter-wave, monopole antenna for each beacon transmitter

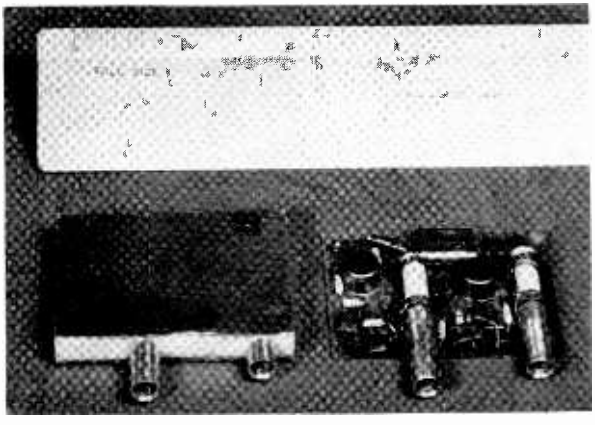


Fig. 5—A transmitter before and after foam molding.

erects normal to the satellite surface. Figure 6 shows the antenna field-intensity pattern versus azimuth measured on a 100-to-1 scale model. The pattern is similar to that of a monopole antenna above an infinite ground plane, except that the main beam is shifted about 20 degrees. Nulls were about 20 decibels down from maximum intensity.

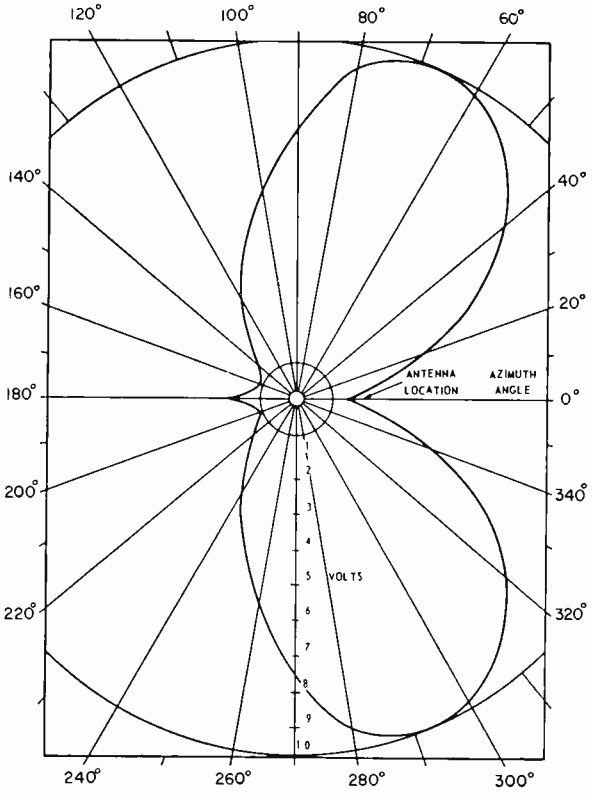


Fig. 6—Relative field intensity versus azimuth.

Impedance measurements were made on the full-scale antenna above a ground plane to determine the proper length which would give a pure resistive load. A length of 25.5 inches was found to give a resistance of 40 ohms. This is an excellent match for the transmitter impedance of 36 ohms.

The antenna can be seen mounted with its transmitter on a base in Figure 7. Its erecting force is derived from torsional energy stored

in the long length of antenna wire passing through the metal tube next to the transmitter. The signal is fed into the end of the wire opposite to the antenna where it has a right angle bend; this end is anchored by the shorter length of tube. The antenna wire is insulated with Teflon* tubing within these tubes. When folded down, torsional energy for erection is stored in the length of antenna passing through the longer metal tube. The antenna wire diameter of 0.031 inch was selected as the minimum which would support its own weight.

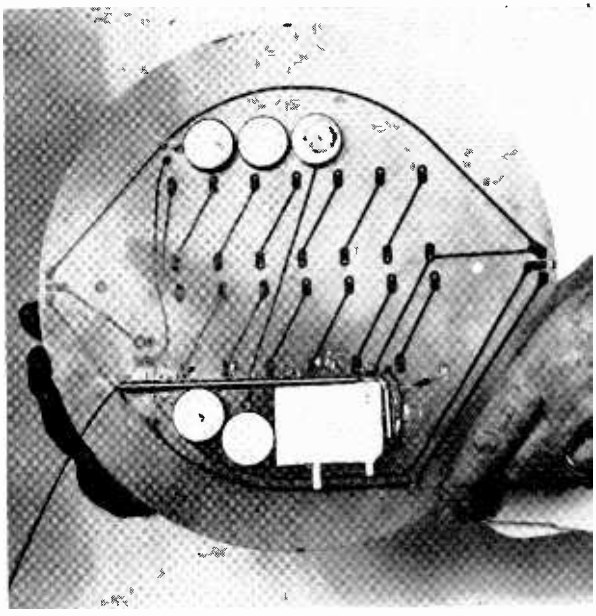


Fig. 7—Storage batteries, transmitter, and antenna mounted on the base.

Power-Supply System

The power-supply system for the beacon transmitters consists of three power sources connected in parallel — two groups of photovoltaic solar cells and a storage battery. A group of solar cells is packaged with each transmitter. Each storage battery consists of ten, series-connected, nickel-cadmium cells, five of which are packed with each transmitter. Although arranged physically in two packages, the complete system operates the beacons continuously and independently of

* Registered trade mark.

the satellite attitude with respect to the sun during sun and shadow time. When either group of solar cells receives adequate solar radiation, that particular group will power both transmitters and charge the storage battery simultaneously.

The solar cells are single-crystal, boron-doped silicon wafers 1 by 2 centimeters in size. A total of 70 series crystals is required for each group. The individual cells mounted in a rectangular pattern, 3.65 by 5.70 inches, giving a total collector area of 20.8 square inches for each group. The terminal voltage for these groups of cells is essentially regulated by the storage battery load.

The ten unit cells comprising the storage battery are each 1-inch diameter, 0.3-inch thick nickel-cadmium units. Each has a nominal capacity of 100 milliampere hours at a temperature of 21°C. They are fabricated in a welded steel container with a glass-to-metal seal at the positive terminal, features which obviate vacuum and electrolyte problems. Five of these unit storage battery cells are shown attached to a transmitter assembly base in Figure 7. The nominal terminal voltage of each unit ranges from 1.5 to 1.2 volts depending on its state of charge. The storage battery is capable of supplying transmitter power without receiving a charge for a period of 6 hours at a temperature of 21°C. For continuous discharge, the voltage would remain at about 14 volts for the first half of this period and drop to an end voltage of 12 during the remaining time. However, considering the possible temperature variations in orbit, this discharge cycle could occur in 2 to 6 hours.

The power-supply system was designed to supply 200 milliwatts of continuous power to operate both beacon transmitters. Although the maximum solar intensity at sea-level is 140 watts per square centimeter of area normal to the sun vector, only a fraction of this power is available from a square centimeter of solar-cell surface. The power available is reduced by shadow time in orbit, by the conversion efficiency of the cells, by geometric factors, and by several degradation factors.

The 70 per cent sun-time orbit is the least favorable expected and was used as a basis for the design. The shadow period in this orbit is about 35 minutes out of a total orbit period of 120 minutes. The output power in the analysis was reduced by 30 per cent to allow for this possible dark period. The solar cells used in this power supply were selected to have a minimum solar conversion efficiency of 9 per cent when normal to solar radiation at sea level at a temperature of 30°C.

The angle between the sun vector and the normal to the solar cell

collector surface is not always favorable. An average angle of 60 degrees was used for design purposes. With two diametrically opposite solar-cell groups mounted on the satellite, there was equal chance for the normal to one of the groups to form an angle greater or smaller than this value with the sun vector. Random motion of the satellite during the 85 minutes of sun time would provide sufficient solar power to replenish the storage battery for depletion of charge occurring during those times when the above angle is greater than 60 degrees for both solar-cell groups.

Since power output is proportional to the projected area of the collector surface normal to the sun vector, it varies as the cosine of this angle. Assuming an average angle of 60 degrees, the average power will be only 0.5 (cosine 60°) of the maximum available. Accordingly, the output in the analysis was reduced by 50 per cent for this factor.

The following degradation factors further reduce the solar cell output power:

1. Temperature of cells greater than 30°C,
2. Storage-battery inefficiency,
3. Attenuation of solar radiation by the cell coating,
4. Micrometeorite bombardment,
5. Spectral shift of incident radiation at orbital altitude,
6. Loss in protective diode,
7. Reflection of solar radiation,
8. Loss of effective area due to electrical connections to the collector surface of the cells.

Starting from 30°C, the power output of solar cells decreases with temperature increases at a rate of 0.6 per cent per degree centigrade. Although the solar cells may attain a temperature of 99°C when normal to solar radiation, they have maximum power output at this attitude. Because of this power-output-to-temperature relationship, it was not necessary to derate the cells a full 41.4 per cent. A reduction in power of 18 per cent for temperature considerations was adequate and was used in this analysis.

The efficiency of the nickel–cadmium storage battery for complete charge–discharge cycles was determined by laboratory tests which showed that it would not be less than 45 per cent. The loss because of battery efficiency would occur only during the 30 per cent shadow time of the orbit. Consequently, the loss in available solar power due to this factor is the product of 0.30 and 0.45, or a 13.5 per cent

reduction. The remaining degradation factors reduce the available power by an additional 53.5 per cent.

The effective power available from a single solar cell of two square centimeters under the least favorable orbit conditions can be calculated by applying the reduction factors. To obtain the total power remaining, a total reduction factor is obtained from the product of all reduction factors expressed as the fraction of power remaining.

Total Reduction Factor = (sun-time) × (conversion efficiency) × (average angle) × (temperature) × (battery efficiency) × (other degradations)
= (0.70) (0.09) (0.50) (0.82) (0.865) (0.465)
= 0.0104.

Effective Power per Cell = (solar intensity) × (area per cell) × (total reduction factor)
= (140) (2) (0.0104) = 2.9 milliwatts per cell

The total effective power required was 200 milliwatts; the number of solar cells for a group on this basis is 200 divided by 2.9, or 69. The voltage of the solar cell group must equal or exceed that of the storage battery which is 15 volts. The minimum voltage of an individual solar cell with its maximum temperature taken as 100°C is 0.242 volt. Therefore, the number of solar cells required for a group on this basis is 15 divided by 0.242, or 62. For packing considerations the total number of solar cells in a group was chosen to be 70.

Packaging

All the components of each transmitter assembly are mounted on a 1/32-inch thick glass-filled epoxy printed circuit board which provides all the electrical interconnections. This board is shown in Figure 7. The terminal connections for the solar-cell shingles are seen in the center of the board. After all the components except solar cells were mounted, this board was foam molded with polyurethane foam having a density of 12 pounds per cubic foot to bring all contours to a uniform height. The resulting mold product had a faired outer edge, a circular groove for storing the antenna, a depression for the solar cells, and two tunnels leading to the slug-tuned transformers in the transmitter circuit.

The solar-cell shingles were subsequently installed. Since the solar cells were of silicon and very brittle, a rigid mounting was provided

by bonding an 0.016-inch thick aluminum alloy mounting plate in the depression reserved for the solar cells and then bonding the shingles to this plate with epoxy resin and fiber-glass cloth. The solar-cell shingles were then covered with a thin layer of clear transparent epoxy resin leveled with the foam-molded surface. This installation of the solar cells provided a very rigid mounting because the sandwich structure, with polyurethane foam bonded between the aluminum mounting plate and the glass-filled epoxy circuit board base, resists bending in the same manner as do honey-comb structures used in aircraft. This design prevented damage to the cells during handling

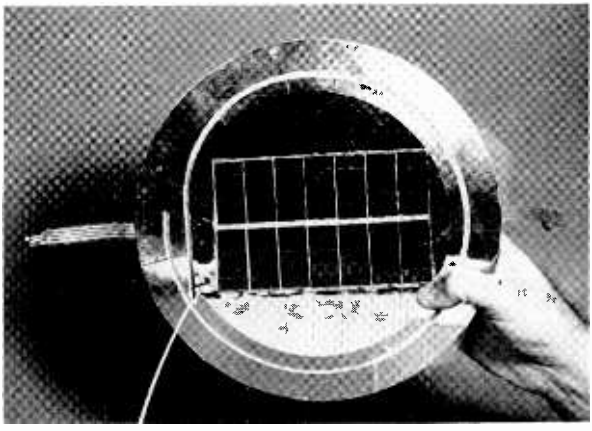


Fig. 8—Transmitter and power supply assembly.

and launch and resisted bending during the packing operation. The air was evacuated during folding of the satellite to obtain a compact package. It was determined by tests that depressions, such as that provided for storing the antenna, do not injure the satellite material during the folding and packing procedures, but that protrusions will stretch the satellite skin. The antenna was covered with Teflon sleeving and was ball-tipped as a precaution against damaging the satellite when it erected during satellite inflation. A completed transmitter assembly covered with aluminum coated Mylar is shown in Figure 8. The antenna was folded into the circular groove and held in place by the folds of the satellite when the latter was folded and packed into the container. The weight of each transmitter assembly is 11 ounces and its total thickness is 0.35 inch.

The two 157-foot lengths of printed cable were manufactured by bonding four 0.002 by 0.062-inch copper conductors between two

½-inch widths of 0.0005-inch-thick Mylar tape. The two center conductors were connected in common to reduce the resistance in series with the storage battery elements to 2.7 ohms, and may be considered as a single conductor. The two outside conductors which place the storage battery and the two groups of solar cells in parallel have a resistance of 5.4 ohms each. These resistances are negligible relative to the transmitter input impedance. The two lengths of printed cable added 18 ounces of distributed weight over the satellite.

Temperature Control

Radiation is the only mode of heat dissipation available to a satellite. Internal conduction in the Mylar is negligible. Although satisfactory operation of the solar cells and electronics is possible over a very wide range of temperature, it was necessary to limit the storage batteries to a range of —10°C to 60°C for best performance. A range of 70 per cent to 100 per cent sun-time orbit had to be considered in establishing the possible temperature extremes. Since the mass of the assemblies is small, there is little thermal inertia, and equilibrium temperatures were expected to be established quickly. The transmitter assemblies were considered to be isothermal with no rotation of the satellite in computation of the possible temperature extremes. The main sources of radiation to the transmitter assemblies included in the analysis were direct solar radiation, solar radiation reflected from the earth, and direct radiation from the earth. Radiation to the under side of the assembly from the interior of the satellite and power dissipation by the components were also considered in the analysis.

Solutions for the extreme temperatures of the beacon transmitter assemblies were obtained by using a computer. The results of the analysis were 216 different solutions for the temperature for each of five orbital positions corresponding to possible temperature extremes. Each solution was for a combination of three temperature-control parameters: the internal and external values of emissivity and the external value of solar absorptivity.

The orbital parameters which determine the possible temperatures of the transmitter assemblies are the per cent of orbit time that the satellite is in the sun, and the position relationship between the transmitter assembly, the satellite, and the sun vector. The five possible extreme temperature situations were as follows:

Case 1. The satellite is in a 100 per cent sun-time orbit with the transmitter assembly located on the sun side of the satellite normal to the sun vector. This is a steady-state

condition with the transmitter assembly receiving direct solar energy.

Case 2. The satellite is in a 100 per cent sun-time orbit with the transmitter assembly located on the shadow side of the vehicle, normal to the sun vector. This is a steady-state position with the transmitter assembly receiving no direct solar energy.

Case 3. The satellite is in a 100 per cent sun-time orbit with the transmitter assembly located parallel to the sun vector and perpendicular to the plane of the orbit. This is a transient condition with the transmitter assembly receiving no direct solar energy.

Case 4. The satellite is in a 70 per cent sun-time orbit with the transmitter assembly located on the sun side of the satellite, normal to the sun vector. This is a transient condition with the satellite receiving direct solar energy 70 per cent of the time.

Case 5. The satellite is in a 70 per cent sun-time orbit with the transmitter assembly located on the shadow side of the satellite, normal to the sun vector. This is a transient condition with the transmitter assembly never receiving direct solar energy, but receiving reflected solar radiation 70 per cent of the time.

Lack of rotation of the vehicle was assumed in the analysis. Any random rotation which the satellite might possess would be favorable since the temperature extremes computed would not be realized. (There are indications that the satellite is, indeed, rotating at a slow rate.) Temperature values for any orbit falling between the 70 and 100 per cent sun-time orbits would have temperature extremes within the computed limits.

The values of thermal emissivity and solar absorptivity which were varied to obtain the solutions were values realizable from available materials and coatings. The solution selected restricted the temperature of the transmitter assemblies to a range of —6°C to 54°C. The values of the control parameters to achieve this range are:

External emissivity: 0.04

Solar absorptivity: 0.02

Internal emissivity: 0.50

The uncoated Mylar film surface on the interior of the satellite has an emissivity value of approximately 0.5 which establishes the internal emissivity. The ratio of solar absorptivity to external emissivity was selected to give a reasonable steady-state temperature for Case 1. The absolute values forming this ratio are low in order to achieve a low value of external emissivity and thereby to increase the thermal time constant and limit the minimum temperature occurring in a 70 per cent sun-time orbit. The 54°C temperature is a constant value for Case 1 and the —6°C temperature is the minimum for Case 5.

The desired external thermal surface characteristics were obtained first by covering the entire surface of an assembly, except for the solar cells, with an aluminum-coated Mylar sheet identical to the satellite material. The desired values of emissivity and absorptivity were then obtained by placing an additional Mylar film over the areas covering the storage batteries and transmitter electronics. This latter film was separated from the first covering by a very small space to preclude conduction. To assure this physical separation, a small, flat, single-turn spring fabricated from 0.010-inch diameter wire was placed between the films. This technique is analogous to vacuum-bottle construction and resulted in very low effective values of absorptivity and emissivity for the exterior side of the transmitter assembly. This film is also identical to the satellite material, but mounted with the aluminum-coated side facing the interior. The external, uncoated, Mylar surface was covered with 0.125-inch-diameter black epoxy dots, 25 to the square inch, to produce a ratio of absorptivity to emissivity of 0.5. The need for this thermal control is apparent when it is realized that the surface of the satellite was expected to experience temperatures ranging from —109°C to 150°C.

The transparent epoxy resin coating the solar cells had an emissivity of 0.7 and a ratio of thermal emissivity to solar absorptivity of 1.0. The temperature of the solar cells can vary between —51°C and 99°C. Fortunately, these cells are not impaired by these temperature extremes and, for this reason, the degradation in power output at the higher temperature was not considered in designing the power supply.

INSTALLATION

The transmitter assemblies were installed on the deflated satellite after the final tuning adjustments were completed and sealed and the batteries were fully charged. The location for each transmitter assembly was reinforced with 16 radial strips of one-inch wide Mylar

adhesive tape attached in a spoke pattern, and each assembly was bonded to this area with an exact-size, ten-inch-diameter, double-faced adhesive sheet. Eight strips of the one-inch wide tape were used to bind the assembly to the satellite with the strips passing over the faired edge symmetrically about the assembly. The ½-inch-wide printed cable was attached to the satellite skin by centering it under one-inch-wide Mylar adhesive tape and then soldering its ends to the assemblies. Installation of the printed cable is shown in Figure 9.

The Mylar tape is 0.001 inch thick with a vapor-deposited aluminum coating on the exterior and an adhesive having 45 ounces peel adhesion to the inch on the opposite side. The adhesive slowly cures at room temperature to obtain maximum bonding strength and is suitable for high temperatures and the vacuum environment.

Fig. 9—Installing printed cable on the satellite.

CONCLUSIONS

The Echo beacons were tracked for several days after the satellite was placed in orbit. The successful operation of both transmitters was indicated by the beat note produced. For one week, the transmitters worked both when the satellite was in the sunlight and when it was in the earth's shadow. At the end of the week, a malfunction occurred in the battery supply, and the transmitters would operate only when a solar-cell pack was illuminated by the sun. It was found that both transmitters operated simultaneously when powered by either solar-cell pack. When the rotation of the satellite placed the solar cells in a position tangent to the sun's rays, both transmitters ceased to operate. This type of performance indicates that the cable connecting the two transmitters has not been broken.

equation given in Reference (1) for electron beams of *finite diameter* is subject to the condition

$$\frac{\omega_p}{\omega} \ll 1. \tag{6}$$

Here ω_p is the plasma frequency in the drift space, and ω the angular signal frequency. Unfortunately, C_5 in Equation (4) is found in klystrons to be eliminated by assumption (6). Thus the two-dimensional theory of Reference (1) cannot predict nonlinear phase distortion. However, a *one-dimensional theory* published earlier[3] is not restricted to small values of ω_p/ω and can therefore be used to calculate the phase distortion. In this one-dimensional model, the beam is assumed to be velocity modulated by a double layer at

$$\beta_p x = -\frac{\pi}{2} \tag{7}$$

with a sinusoidal r-f voltage

$$V = -V_1 \sin\left(\omega t + \frac{\pi}{2} - \frac{\omega}{\omega_p}\right) \tag{8}$$

The second-order solution yields an amplitude-dependent electric d-c field (Equation (18) in Reference (3)) given by

$$\frac{\omega_{p0}}{i_0}\bar{y}_2 = -\left(\frac{V_1}{4V_0}\right)^2 \sin 2\beta_p x. \tag{9}$$

Here $\bar{y}_2$ is the time average of the second-order axial dielectric displacement, i_0 the d-c density, and V_0 the beam voltage; ω_{p0} is the value of the plasma frequency before the beam crosses the input gap. The value in the drift space, ω_p, is slightly increased due to the reduction in average stream velocity by the r-f voltage in the gap (Equation (16) in Ref. (3)). Integration of Equation (9) leads to a d-c potential, $\overline{V}$, given by

$$\frac{\overline{V} - V_0}{V_0} = \frac{1}{2}\left(\frac{V_1}{2V_0}\right)^2 \cos^2\beta_p x. \tag{10}$$

[3] F. Paschke, "On the Nonlinear Behavior of Electron-Beam Devices," *RCA Review*, Vol. XVIII, June, 1957, p. 221.

The variation of d-c potential along the beam is caused by the r-f excitation and would adjust itself to the value given by Equation (10) provided there were no boundary which would enforce a different potential. In practice, the presence of a drift tube would most likely suppress the axial potential variation.[4] However, in the one-dimensional model used here, there is no radial boundary, but the presence of an output cavity, ignored in Reference (3), will influence the potential distribution. Suppose the output gap is located at

$$\beta_p x = \alpha, \tag{11}$$

and is kept at the beam voltage, V_0. Then the time-average part of the second-order solution given in Reference (3) has to be modified so as to take this boundary condition into account. This is done by addition of a proper solution of the homogeneous part of the differential equation for the second-order perturbation (Equation (12b) in Reference (3)). The modified time-average second-order solution is

$$\frac{\omega_{p0}}{i_0}\bar{y}_2 = \left(\frac{V_1}{4V_0}\right)^2 (-\sin 2\beta_p x + \cos\alpha \sin\beta_p x). \tag{12}$$

Equation (12) leads to the following d-c potential distribution:

$$\frac{\bar{V} - V_0}{V_0} = \frac{1}{2}\left(\frac{V_1}{2V_0}\right)^2 (\cos^2\beta_p x - \cos\alpha\cos\beta_p x). \tag{13}$$

It is seen that both at the input and output gap, $\bar{V} = V_0$.

The modification of the second-order solution requires a recalculation of the third-order perturbations (Equation (12c) in Reference (3)). The second term in Equation (12) gives rise to the following third-order terms which have to be added to the solution given by Equation (18) in Reference (3):

$$\frac{\omega_{p0}}{i_0} y_{3\text{add}} = \frac{1}{4}\left(\frac{V_1}{2V_0}\right)^3 \cos\alpha \Big[(\cos 2\beta_p x - \sin\beta_p x) \sin(\omega t - \beta_e x)$$

[4] The variation of this "rectified potential" is important for the conservation of kinetic power flow in the one-dimensional model of the beam. The reader is referred to H. Kogelnik, "Über die Energieverhältnisse in Elektronenströmungen," *Arch. d. elekt. Übertrag.*, Vol. 12, p. 419, September, 1958.

$$-\frac{\omega}{\omega_{p0}}\left(\frac{1}{2}\sin 2\beta_p x+\cos\beta_p x\right)\cos\left(\omega t-\beta_e x\right)\Bigg] \tag{14}$$

At the input gap ($\beta_p x=-\pi/2$), the velocity and convection-current density associated with this additional term are, in accordance with the boundary conditions, equal to zero.

At this point the writer finds it desirable to change to a coordinate system which is more convenient than the one used in Reference (3). In the new system the input gap is located at

$$\beta_p x=0, \tag{15}$$

and the gap voltage is assumed to be

$$V=V_1\sin\omega t, \tag{16}$$

in contrast to Equations (7) and (8) of the old system. The field distributions in the new system are obtained from the equations in Reference (3), supplemented by Equations (12)-(14) of this paper; $\beta_p x$ is replaced by $\beta_p x-\pi/2$, ωt by $\omega t-\pi\omega/2\omega_p$, and V_1 by $-V_1$. One finds for the fundamental convection current at the output gap

$$\begin{aligned}\frac{i\omega}{i_0}=&\frac{V_1}{2V_0}\frac{\omega}{\omega_{p0}}\sin\theta\cos\left(\omega t-\beta_e x\right)-\frac{1}{16}\left(\frac{V_1}{2V_0}\frac{\omega}{\omega_{p0}}\right)^3\Bigg\{\Bigg[-\frac{1}{2}\Bigg(1\\&+5\left(\frac{\omega_{p0}}{\omega}\right)^2\Bigg)\sin 3\theta+\frac{1}{2}\left(3-17\left(\frac{\omega_{p0}}{\omega}\right)^2\right)\sin\theta\\&-4\left(\frac{\omega_{p0}}{\omega}\right)^2\sin 2\theta\Bigg]\cos\left(\omega t-\beta_e x\right)+\frac{\omega_{p0}}{\omega}\Bigg[2\left(\cos 3\theta-\cos\theta\right)\\&-3\left(1-\cos 2\theta\right)\Bigg]\sin\left(\omega t-\beta_e x\right)\Bigg\}.\end{aligned} \tag{17}$$

Here θ is the distance between the cavities measured in plasma-phase angle.

Under the assumption of relatively small space-charge densities, Relation (6), one obtains the expression

$$\frac{i\omega}{i_0}=\frac{V_1}{2V_0}\frac{\omega}{\omega_{p0}}\sin\theta\left[1-\frac{1}{8}\left(\frac{V_1}{2V_0}\frac{\omega}{\omega_{p0}}\sin\theta\right)^2\right]\cos\left(\omega t-\beta_e x\right). \tag{18}$$

Note that Equation (18) is a series approximation of the familiar Bessel function found by Webster.[5] Equation (18) predicts a saturation efficiency of $(2/3)^{3/2} \doteq 54.4$ per cent, independent of the cavity distance θ. The Webster theory predicts about 58 per cent, also independent of θ. Thus the convergence of the Taylor series, Equation (18), can be considered satisfactory. It would be interesting to know how the higher-order space-charge terms in Equation (17) influence efficiency. Unfortunately, such a large gap modulation, V_1/V_0, is required to saturate the convection current for relatively large ω_p/ω, that a third-order perturbation theory is insufficient. However, one can show from Equation (17) that in the region where the one-dimensional theory can be applied[1] ($\theta \ll \pi$), space charge *improves* the saturation efficiency. This can be attributed to the bunching effect of the harmonic velocity modulation caused by the large sinusoidal r-f gap voltage.

Turning back to the prime objective of this part of the article, one derives from Equations (5) and (17) the nonlinear phase-distortion (phase-delay)

$$\phi_1 = \arctan\left[\frac{\omega_{p0}}{\omega}\frac{3}{8}\left(\frac{V_1}{2V_0}\frac{\omega}{\omega_{p0}}\sin\theta\right)^2 \frac{1+\frac{4}{3}\cos\theta}{\sin\theta}\right]. \qquad (19)$$

It must not be forgotten that the time average of the velocity with which the stream is injected into the drift space depends on the gap voltage. This d-c injection voltage is, from Equation (16) in Reference (3), equal to

$$v_0 = (2\eta V_0)^{\frac{1}{2}}\left(1-\left(\frac{V_1}{4V_0}\right)^2\right) \qquad (20)$$

Thus the drift angle $\beta_e x = \omega x/v_0$ in the argument of the cosine in Equation (17) is increased due to the gap modulation by

$$\phi_2 = \frac{\omega_{p0}}{\omega}\frac{\theta}{4\sin^2\theta}\left(\frac{V_1}{2V_0}\frac{\omega}{\omega_{p0}}\sin\theta\right)^2 \qquad (21)$$

The total amplitude-dependent phase delay is therefore given by

---

[5] D. L. Webster, "Cathode-Ray Bunching," *Jour. Appl. Phys.*, Vol. 10, p. 501, July, 1939.

$\phi_1 + \phi_2$. To obtain a practical formula, the saturation voltage, V_{1s}, is calculated from Equation (18):

$$\left(\frac{V_{1s}}{2V_0}\frac{\omega}{\omega_{p0}}\sin\theta\right)^2 = \frac{8}{3}. \tag{22}$$

It should be recalled that Equation (22) is valid only for relatively small values of ω_{p0}/ω. The total nonlinear phase delay can be written from Equations (19), (21), and (22) as

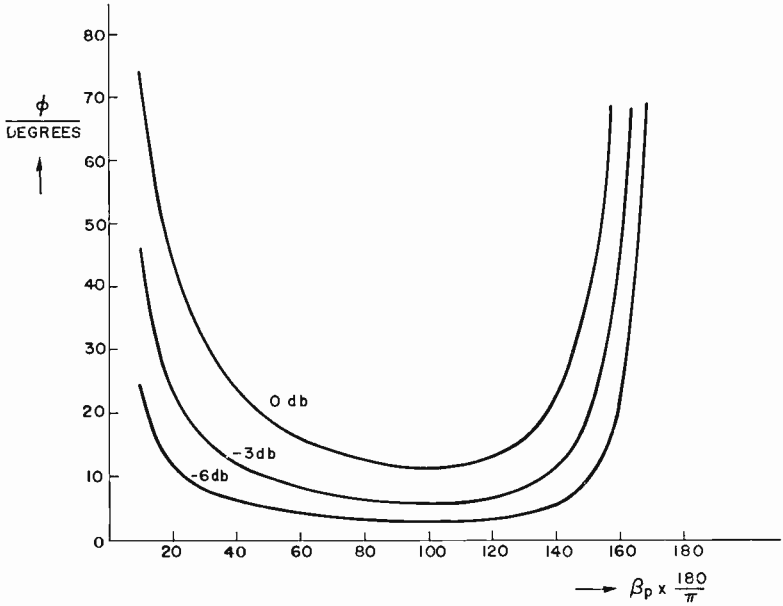


Fig. 1—Phase delay versus normalized cavity distance in a two-cavity klystron for $\omega_{p0}/\omega = 0.1$. The parameter is the input power with zero decibels corresponding to saturation.

$$\phi = \frac{\omega_{p0}}{\omega}\frac{2\theta}{3\sin^2\theta}\left(\frac{V_1}{V_{1s}}\right)^2 + \arctan\left[\frac{\omega_{p0}}{\omega}\frac{1+\frac{4}{3}\cos\theta}{\sin\theta}\left(\frac{V_1}{V_{1s}}\right)^2\right]. \tag{23a}$$

Figure 1 shows this phase delay versus the drift length for $\omega_{p0}/\omega = 0.1$. The input voltage serves as a parameter; zero decibels corresponds to saturation. It should be noted, however, that the one-dimensional

model treated here is applicable only for relatively small drift distances.[1] Thus an agreement of Equation (23) with experiments is expected only for $\theta \ll \pi$. Unfortunately, there are, to the author's knowledge, no experimental data published in the literature.

As a result of the amplitude dependence of the phase shift, amplitude modulation will be, to a certain degree, converted into phase modulation. The conversion factor k_p is defined as the change in phase shift per decibel change of input power. In most practical cases the arctan function in Equation (23a) can be replaced by its argument, so that the phase delay is a linear function of input power. Then

$$k_p = \frac{d\phi}{20d \log V_1/V_{1s}} = \frac{\phi}{10 \log e} \doteq \frac{\phi}{4.34}. \tag{23b}$$

Nonlinear Theory of the Traveling-Wave Amplifier

The theory is based on the model used in References (1) and (3). To solve the three linear differential Equations (12) in Reference (3), the following assumptions are made:

1. The beam is modulated at $z = -\infty$ by an infinitesimal monochromatic signal. Thus, in first order, only the exponentially growing wave needs to be considered.
2. The harmonics do not interact with the delay line.
3. The gain parameter[6] C is very small.

Solution of Equation (12a) in Reference (3) yields the well-known small-signal theory. One obtains for the convection-current density

$$\frac{i_1}{i_0} = \left(-\frac{1}{y}\frac{\eta_{el}}{C}\right)^{1/2} \cos(\omega t - \beta z). \tag{24}$$

Here η_{el} is the electronic efficiency of the *linear* theory (at small signals), and β the real part of the propagation constant of the growing wave. The normalized real part, y, of the propagation constant, must not be confounded with the variables y_n (displacements) in References (1) and (3). Solution of Equation (12b) in Reference (3) for $m_2 = 0$, i.e., no interaction at the second harmonic, yields the second-harmonic current

[6] The parameters used throughout this paper are from J. R. Pierce, *Traveling-Wave Tubes*, D. van Nostrand Company, Inc., Princeton, N. J., 1950.

$$\frac{i_{2\omega}}{i_0} = -\frac{3}{4}\frac{1}{y}\frac{\eta_{el}}{C}\left([(x^2+y^2)^2 - QC\,(y^2-x^2)]\cos 2(\omega t-\beta z) - 2xyQC\sin 2(\omega t-\beta z)\right)\left((x^2-y^2+QC)^2+4x^2y^2\right)^{-1} \quad (25)$$

Here x is the normalized growth constant and QC the passive-mode parameter.[6] Use has been made of the well-known relation

$$4QC^3 = (\omega_p/\omega)^2. \quad (26)$$

The third-order solution derived from Equation (12c) in Reference (3) consists of both fundamental and third-harmonic terms. To obtain the fundamental third-order term the interaction with the delay line must be taken into account. It is plausible to assume that, in complex notation,

$$\frac{\hat{y}_1 - \hat{m}_1}{\hat{y}_1} = \frac{\hat{y}_{31} - \hat{m}_{31}}{\hat{y}_{31}}, \quad (27)$$

where the index 31 designates third-order terms of the fundamental frequency. The circumflexes designate complex quantities. Physically, Equation (27) means that the complex ratio of the fundamental component of the axial electric field strength to the fundamental component of the lateral electric field strength is determined by the boundary conditions and is not a function of signal strength. This technique of determining the higher-order fringe-field terms has been successfully used in an analysis of velocity-modulated beams (Equation (8), Reference (1)). With the third-order fundamental coupling term m_{31} calculated from Equation (27), and the third-harmonic coupling term $m_{33} = 0$, the solutions of Equation (12c) in Reference (3) for the convection current are

$$\frac{i_{31}}{i_0} = -\left(\frac{\eta_{el}/C}{-y}\right)^{3/2}[A_{31}\cos(\omega t-\beta z) + B_{31}\sin(\omega t-\beta z)], \quad (28)$$

$$\frac{i_{33}}{i_0} = \left(\frac{\eta_{el}/C}{-y}\right)^{3/2}[A_{33}\cos 3(\omega t-\beta z) + B_{33}\sin 3(\omega t-\beta z)], \quad (29)$$

$$A_{31} = \frac{1}{16}\,\frac{7x^6 + 14x^4y^2 + 11x^4QC + 7x^2y^4 + 28x^2y^2QC + 4x^2(QC)^2 + 9y^4QC}{[(x^2-y^2+QC)^2+4x^2y^2]\,[4x^2+y^2]}, \quad (30$$

$$B_{31} = \frac{y}{32x}\Big[\, 27x^8 + 59x^6y^2 - 2x^6QC + 37x^4y^4 - 27x^4y^2QC$$
$$- 73x^4(QC)^2 + 5x^2y^6 - 12x^2y^4QC + 30x^2y^2(QC)^2$$
$$- 44x^2(QC)^3 - 3y^6QC + 15y^4(QC)^2 - 12y^2(QC)^3 \Big]\Big[(x^2$$
$$- y^2 + QC)^2 + 4x^2y^2 \Big]^{-1}\Big[(4x^2 + y^2)(x^2 + QC) \Big]^{-1}, \quad (31)$$

$$A_{33} = \Big(\frac{5}{2}(y^4 - x^4)^2 + 10x^2y^2(y^2 + x^2)^2 + QC(y^2 - x^2)\Big[\frac{17}{9}(y^2$$
$$+ x^2)^2 - \frac{7}{2}(y^2 - x^2)^2 - 14x^2y^2 \Big] - (QC)^2\Big[\frac{4}{9}(y^2 - x^2)^2$$
$$+ \frac{4}{3}(y^2 + x^2)^2 + \frac{32}{3}x^2y^2 \Big] + (QC)^3\frac{8}{9}(y^2 - x^2) \Big)\Big(4\Big[[x^2$$
$$- y^2 + QC]^2 + 4x^2y^2 \Big]\Big[\Big(x^2 - y^2 + \frac{4}{9}QC \Big)^2 + 4x^2y^2 \Big] \Big)^{-1} \quad (32)$$

$$B_{33} = -QCxy\Big[12x^2y^2 + 3(y^2 - x^2)^2 + \frac{74}{9}(y^2 + x^2)^2$$
$$- QC\frac{40}{9}(y^2 - x^2) - (QC)^2\frac{16}{9} \Big]\Big(4\Big[(x^2 - y^2 + QC)^2$$
$$+ 4x^2y^2 \Big]\Big[\Big(x^2 - y^2 + \frac{4}{9}QC \Big)^2 + 4x^2y^2 \Big] \Big)^{-1} \quad (33)$$

Saturation

The term with the factor A_{31} in Equation (28) is opposite in phase to the small-signal current, Equation (24), and therefore accounts for saturation of the convection current. Unfortunately, the convergence of the series is not nearly as satisfactory as in the case of the two-cavity klystron.[1,3] For example, one can see from Equation (30) that the third-order saturation term vanishes for $QC = 0$ and $x \to 0$;

the saturation efficiency would then be unlimited. The same discrepancy arises in a similar theory by Kiel and Parzen.[7] Vainshtein (Wainstein)[8,9] pointed out that the addition of some terms of higher than third order in the series does not lead to a significant improvement of the analysis. It is clear that the poor convergence must have a physical basis, and that saturation cannot be simply explained by a gradual drift of the bunches from the retarding into the accelerating phase of the electric field. This is well known since Cutler[10] demonstrated experimentally the turbulent nature of the electron flow at large signals. Because of this apparent failure of the univelocity hydrodynamic model, the question of saturation will not be further discussed.

Frequency multiplication

In the traveling-wave frequency multiplier introduced by Schnitger and Weber[11] and Uda *et al.*,[12] the electron beam passes two helices in cascade. The first helix is tuned to the fundamental and serves as modulator. The second helix is tuned to the desired harmonic, and amplifies and couples the harmonic power out of the premodulated beam. In this part of the paper, discussion is restricted to the problem of *harmonic generation.*

The second- and third-harmonic convection currents are given by Equations (25) and (29), respectively. It is seen that the n^{th} harmonic current is predicted for relatively low levels to be proportional to $(\eta_{el}/C)^{n/2}$. This agrees with experimental findings.[11-13] Of particular interest is the case where the denominator vanishes ($x = 0$, $y^2 = QC$). Since the second harmonic does not couple to the delay line, the plasma-frequency-reduction factor is

---

[7] A. Kiel and P. Parzen, "Nonlinear Wave Propagation in Traveling-Wave Amplifiers," *Trans. I.R.E. PGED*, Vol. 2, p. 26, October, 1955.

[8] L. A. Vainshtein (Wainstein), "Electron Waves in Delaying Systems. Non-Linear Equations of Travelling Wave Valves," *Radio Eng. and Elect.*, Vol. 2, No. 6, p. 22, 1957.

[9] L. A. Wainstein (Vainshtein), "Electronic Waves in Delay Lines," *L'Onde Electr.*, Vol. 38, p. 766, November, 1958.

[10] C. C. Cutler, "The Nature of Power Saturation in Traveling Wave Tubes," *Bell System Tech. Jour.*, Vol. 35, p. 841, July, 1956.

[11] H. Schnitger and D. Weber, "Die Wanderfeldröhre als Hilfsmittel der Zentimeterwellen-Messtechnik," *VDE Fachber.*, Vol. 17, No. 5, p. 40, 1953.

[12] S. Uda, K. Kamiryo, T. Fujii, S. Shoji, S. Takachita, and S. Tsuda, "Traveling-Wave Tubes as Straight Amplifiers and Frequency Multipliers," *Jour. Inst. Elect. Comm. Eng.*, Japan, Vol. 36, p. 315, June, 1953.

[13] D. J. Bates and E. L. Ginzton, "A Traveling-Wave Frequency Multiplier," *Proc. I.R.E.*, Vol. 45, p. 938, July, 1957.

$$p(2\omega) = 1. \tag{34}$$

According to Pierce[6] the reduction factor at the fundamental frequency, $p(\omega)$, is linked to the normalized propagation constant, y, by

$$p^2(\omega) = \frac{y^2}{4QC}. \tag{35}$$

For $y^2 = QC$,

$$p(\omega) = \frac{1}{2}. \tag{36}$$

From Equations (34) and (36)

$$p(2\omega) = 2p(\omega). \tag{37}$$

This relation throws some light on the physical reason for the pole† in the expression for the second-harmonic current, Equation (25). Suppose that outside of the amplification range, $x = 0$, the fundamental modulation is given by a space-charge wave;

$$i_1 = A \sin\left(\omega t - \frac{\omega \pm p(\omega)\omega_p}{v_0} z\right). \tag{38}$$

This *eigen-wave* propagates along the beam and, via the second-order nonlinearity, produces a *forced wave* which is described by a harmonic function with twice the argument of the exciting wave;

$$i_{2\omega} \sim A^2 \sin\left(2\omega t - \frac{2\omega \pm 2p(\omega)\omega_p}{v_0} z\right). \tag{39}$$

An eigen-wave at the second harmonic would be given by

$$i_{2\omega} \sim \sin\left(2\omega t - \frac{2\omega \pm p(2\omega)\omega_p}{v_0} z\right), \tag{40}$$

i.e., a space-charge wave with a plasma-frequency-reduction factor $p(2\omega)$. It is seen from Equations (39) and (40) that under condition

† The same pole occurs in the equation for the second-harmonic current of the theory of L. N. Loshakov, "Nonlinear Effects in a Waveguide in the Presence of an Electron Beam Near Synchronism Between Electron and Wave Velocities," (in Russian) *Jour. Tech. Phys. USSR*, Vol. 25, No. 10, p. 1768, 1955.

existence of these higher-order resonances appears verified by a large-signal experiment,[13] in which more than one peak in harmonic output was observed at helix voltages below the amplification region.

Nonlinear Phase Distortion

The failure of the univelocity perturbation theory in calculating saturation efficiency suggests that the nonlinear phase distortion cannot be predicted by this theory either. However, the measured

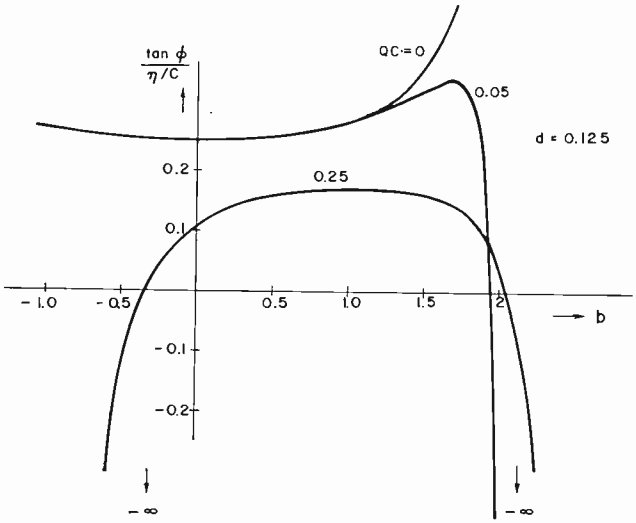


Fig. 3—Phase delay, ϕ, in a traveling-wave tube with efficiency η, gain parameter C, and passive-mode parameter QC versus velocity parameter b.

phase delay turns out to be a linear function of output power almost up to the saturation level. Because of this linearity, the third-order perturbation theory which is strictly valid only for relatively low levels, can be extrapolated up to saturation.

The deceleration of the beam associated with the power extraction is regarded as the prime cause for the amplitude-dependent phase delay. However, intricate mixing phenomena complicate the picture considerably.

From Equations (24), (28), and (5), one finds for the phase delay, ϕ,

$$\tan \phi = \frac{\eta_{el} B_{31}}{Cy}, \tag{45}$$

where B_{31} is given by Equation (31). In Figure 3, the value of B_{31}/y

is plotted versus the velocity parameter, b, for various values of the passive-mode parameter, QC. The propagation constants were taken from Reference (16). To obtain the curve for $QC = 0.05$, the following approximation of Equation (45) for $QC \ll 1$ has been used:

$$\tan \phi = \frac{1}{32x} \frac{\eta_{el}}{C} \frac{27x^6 + 59x^4y^2 + 37x^2y^4 + 5y^6 - 3y^6QC/x^2}{4x^6 + 9x^4y^2 + 6x^2y^4 + y^6 + y^6QC/x^2}. \tag{46}$$

Since x and y for $QC = 0.05$ are not given in the literature, the known values of x and y for $QC = 0$ have been used in conjunction with Equation (46) to calculate the curve for $QC = 0.05$. This approximation is inadequate for $b \ll 0$, of course.

In most practical cases $\phi \ll \pi/2$ so that

$$\tan \phi \doteq \phi. \tag{47}$$

Thus the phase delay is predicted to be a linear function of output power. Under the condition (47), and within the limits of the third-order perturbation theory, the AM-to-PM-conversion factor, k_p, measured in degrees per decibel is given by

$$k_p = \frac{\phi}{10 \log e} \doteq \frac{\phi}{4.34}. \tag{48}$$

In the following, the theoretical results are compared with measurements on traveling-wave tubes of various manufacturers. To obtain the "theoretical" value of ϕ or k_p, the *measured* value of the output power is used. Although in most practical tubes $d < 0.125$, the curves plotted in Figure 3 can be applied because the loss parameter has little influence on the phase delay.

RCA Low-Noise Tube

Figure 4a shows the phase delay versus output power measured by Beam and Blattner.[17] The saturation power is about 10 milliwatts. The theoretical tube parameters were obtained from the Pierce theory.[6] The agreement is satisfactory. Figure 4b shows the ampli-

[16] G. R. Brewer and C. K. Birdsall, "Normalized Propagation Constants for a Traveling-Wave Tube for Finite Values of C," *Hughes Aircraft Corp., Tech. Memo. No. 331*, October, 1953.

[17] W. R. Beam and D. J. Blattner, "Phase Angle Distortion in Traveling-Wave Tubes," *RCA Review*, Vol. 17, p. 86, March, 1956.

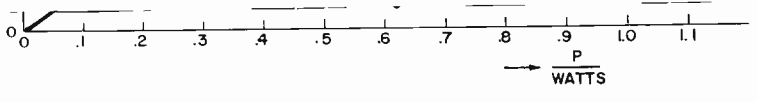


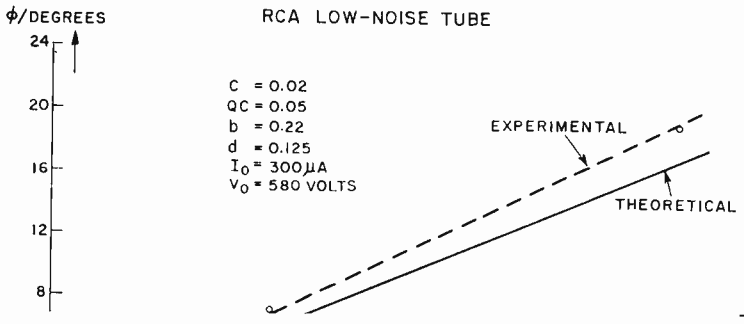
Fig. 5—Phase delay versus output power in the RCA A-1138 traveling-wave tube.

RCA A-1138

The experimental data on this tube, shown in Figure 5, were communicated to the author by D. J. Blattner. The helix voltage is adjusted for optimum small-signal gain. The saturation power is about 1.2 watts. The theoretical tube parameters were obtained from a small-signal theory[18] similar to Pierce's.[6] The Pierce theory leads to a larger QC ($\doteq 0.1$), and to theoretical phase angles about 10 per

[18] F. Paschke, "Die Wechselseitigkeit der Kopplung in Wanderfeldröhren," *Arch. d. elekt. Übertrag.*, Vol. 11, p. 137, April, 1957.

tude-dependent phase delay for constant output power of 1.4 milliwatt versus helix voltage for three values of beam current. The experimental curves actually lie at lower helix voltages but have been displaced by about 17 volts in order to match the theoretical curves and



cent below the theoretical results plotted in Figure 5. However, even for $QC = 0.1$ the agreement can be regarded as satisfactory.

RCA A-1093

The data on this tube shown in Figure 6 were communicated to the author by R. R. Lorentzen. The saturation power is about 23 watts. The theoretical tube parameters were obtained from the Pierce theory. Two curves which were measured at elevated helix voltages ($b > 1$) and therefore reduced gain deviated considerably from the predicted straight line; this is not shown in Figure 6.

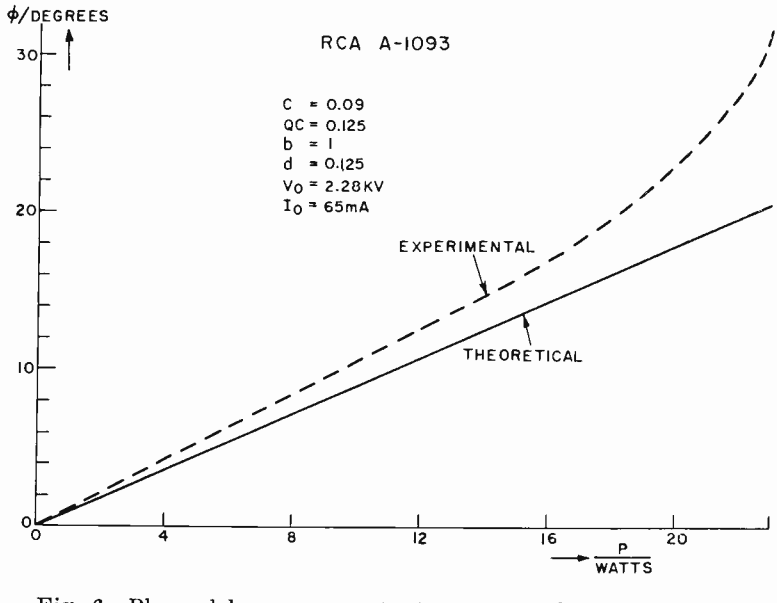
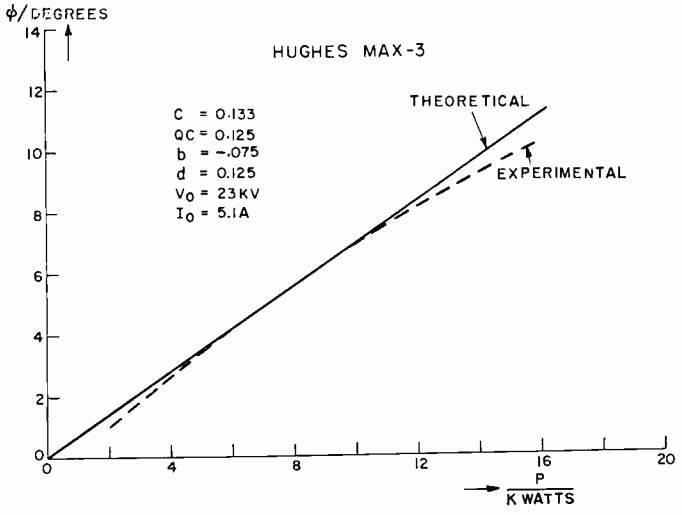


Fig. 6—Phase delay versus output power in the RCA A-1093 traveling-wave tube.

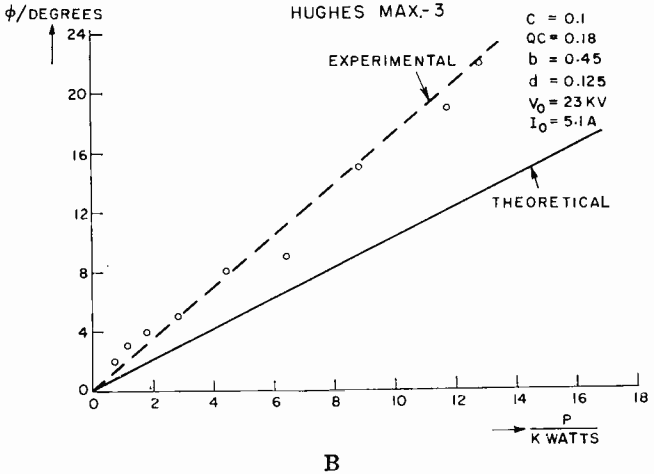
Hughes MAX-3

The experimental data and theoretical small-signal parameters of this tube were communicated to the author by D. J. Bates. The saturation power is about 16 kilowatts for $b = -0.075$, and about 13 kilowatts for $b = 0.45$. Figure 7a shows the excellent agreement of the experiments with the theory for relatively small b. For larger values of b the agreement is not so good, as can be seen from Figure 7b. For still larger values of b the measured relation between phase and output power becomes highly nonlinear in this tube. However,

the applicability of the simple theoretical model is rather limited in this case because the delay line of the MAX-3 is tapered and exhibits several severances.



A



B

Fig. 7—Phase delay versus output power in the Hughes MAX-3 traveling-wave tube for (A) relatively small values of *b* and (B) for larger values of *b*.

Bell Telephone Laboratories M-1789

Experimental data and the theoretical values for the small-signal parameters of this tube are available in the literature.[19] The experimental results shown in Figure 8a verify the predicted linear relation

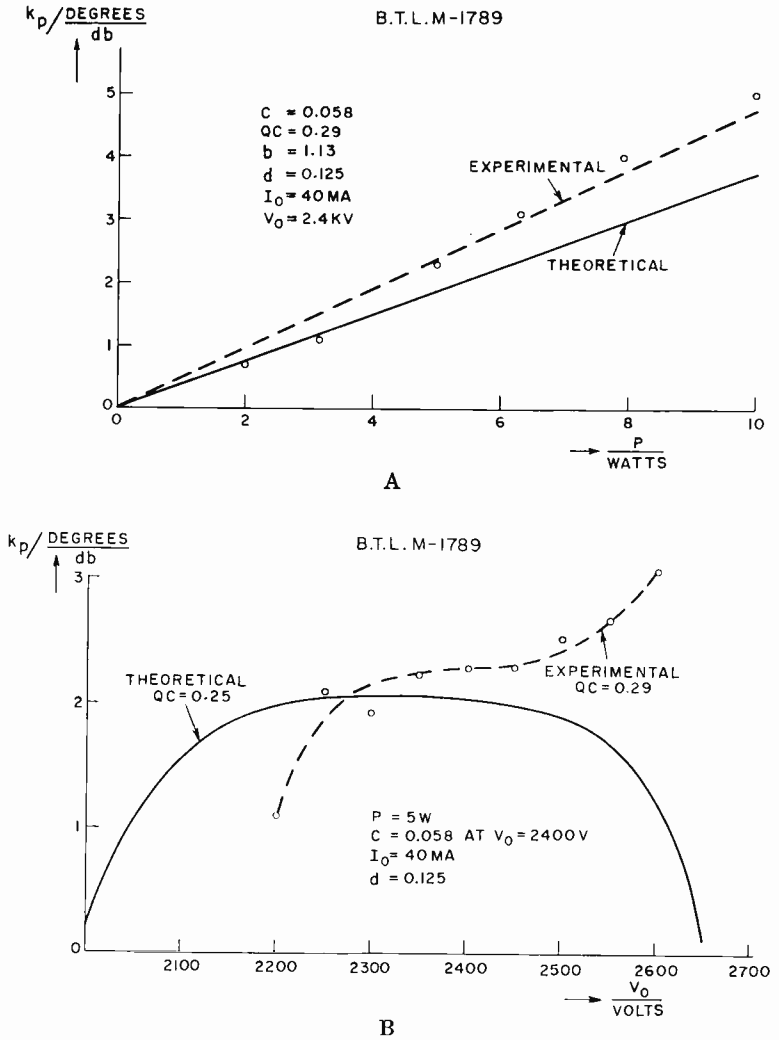


Fig. 8—AM-to-PM conversion factor in B.T.L.'s M-1789 traveling-wave tube[19] (A) versus output power, and (B) versus helix voltage for constant output power.

[19] J. P. Laico, H. L. McDowell, and C. R. Moster, "A Medium-Power Traveling-Wave Tube for 6000-Mc Radio Relay," *Bell System Tech. Jour.*, Vol. 35, p. 1285, November, 1956.

of the AM-to-PM conversion factor and output power. The saturation power is about 10 watts. Figure 8b shows the dependence of the conversion factor on helix voltage for a constant output power of 5 watts. Note that the theoretical curve is valid for $QC = 0.25$, which does not deviate appreciably from the theoretical design value, $QC = 0.29$, given in Reference (19). The agreement is quite satisfactory below and around the nominal helix voltage, $V_0 = 2400$ volts, which corresponds to maximum gain at 5 watts output power. For elevated helix voltages, the measured conversion factor increases with increasing helix voltage, in contradiction to theory. However, it is noteworthy that the experiments show a linear relation of k_p and output power even at elevated helix voltages.

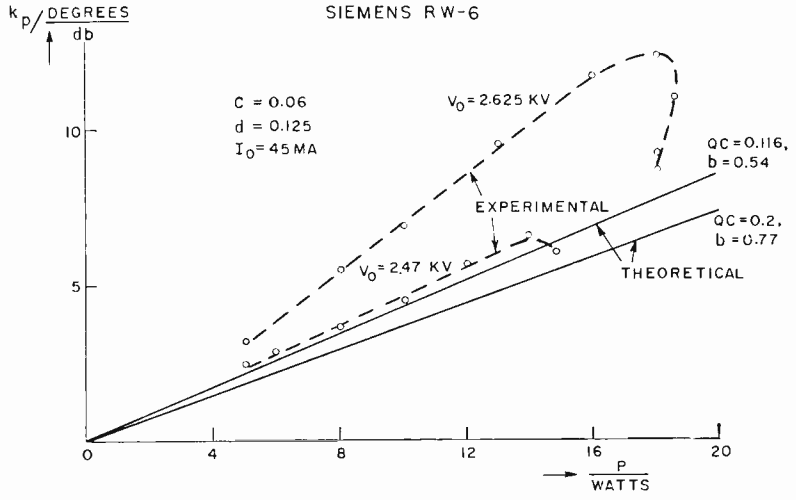


Fig. 9—AM-to-PM conversion factor versus output power in the Siemens RW-6 traveling-wave tube.[20] Both theoretical curves correspond to the nominal helix voltage, $V_0 = 2.47$ kv.

Siemens u. Halske RW-6

The data on this tube, communicated to the author by R. Liebscher, will soon be available in the literature.[20] Figure 9 shows the AM-to-PM conversion factor versus output power for two values of helix voltage. $V_0 = 2470$ volts is the nominal helix voltage of the tube and corresponds to maximum gain at 10 watts output power. To obtain

[20] R. Liebscher, W. Eichin, H. Heynisch, and P. Meyerer, "Die Siemens-Wanderfeldendröhre im J-Band RW6," International Congress on Microwave Tubes, Munich, June 7-11, 1960.

the lower and upper theoretical curves, the parameters QC and b have been calculated according to References (6) and (18), respectively. With both methods the agreement can be considered very satisfactory. However, for an elevated helix voltage of 2625 volts, the experimental conversion factor increased appreciably, while the theoretical curves, not shown in Figure 9, differ negligibly from the values for $V_0 = 2470$ volts.

In summary, the agreement of Equation (45) with experimental results is quite satisfactory. The measured phase angle, or AM-to-PM conversion factor, tends to exceed the theoretical value at elevated delay-line potentials. However, at the nominal operating voltages where large gain is achieved, the measured values agree very well with theory.

ACKNOWLEDGMENT

The author gratefully acknowledges the cooperation of D. J. Bates, D. J. Blattner, R. Liebscher, and R. R. Lorentzen, who contributed unpublished experimental results.

NOISE SMOOTHING BY REACTIVE DAMPING IN FINITE MULTIVELOCITY ELECTRON BEAMS*†

BY

J. BERGHAMMER

RCA Laboratories
Princeton, N. J.

Summary—The dispersion relation for a finite-diameter multivelocity beam having a rectangular distribution of electron velocities is derived. It is found that in addition to the usual Hahn–Ramo waves, a pair of reactively damped waves can exist. The significance of these waves for noise-smoothing in long low-velocity beams of finite geometry is discussed.

INTRODUCTION

RECENT EXPERIMENTAL SUCCESS in traveling-wave-tube noise reduction[1,2] has led to a re-examination of the noise processes in the low-voltage regions of the electron gun. In particular, attention has been focused on the multivelocity aspects of the one-dimensional beam, and it has been shown that multivelocity flow can partially correlate the current and velocity fluctuations and thereby lower the noise figure.[3,4] Similar results have been obtained by a Monte Carlo calculation which takes account of nonlinear processes.[5]

This paper is concerned with another noise reduction mechanism, reactive damping, which is particularly effective in long low-velocity beams of high density and finite transverse dimension. On the basis of the macroscopic equations for a drifting beam having a rectangular

* This work was in part supported by a Signal Corps contract.

† Manuscript received 13 February 1961.

[1] M. R. Currie and D. C. Forster, "Low Noise Tunable Preamplifiers for Microwave Receivers," *Proc. I.R.E.*, Vol. 46, p. 570, March, 1958.

[2] M. R. Currie and D. C. Forster, "New Mechanism of Noise Reduction in Electron Beams," *Jour. Appl. Phys.*, Vol. 30, p. 94, January, 1959.

[3] A. E. Siegman, D. A. Watkins, and Hsung-Cheng Hsieh, "Density-Function Calculations of Noise Propagation on an Accelerated Multivelocity Electron Beam," *Jour. Appl. Phys.*, Vol. 28, p. 1138, October, 1957.

[4] W. M. Mueller and M. R. Currie, "Noise Propagation on Uniformly Accelerated Multivelocity Electron Beams," *Jour. Appl. Phys.*, Vol. 30, p. 1876, December, 1959.

[5] C. A. Lambert and A. H. Dayem, "Noise Near the Potential Minimum of an Open Circuit Diode," paper presented at 1960 Electron Tube Research Conference, University of Washington, Seattle, June 29-July 1, 1960.

distribution of electron velocities, it is shown that in addition to the usual Hahn–Ramo waves, there exists another pair of waves. These additional waves are reactively damped (i.e., are evanescent) under certain conditions. A simple model for the low-velocity region of a finite beam is used to demonstrate the effect of the reactively damped wave on the noise figure of beam-type amplifiers.

Dispersion Relation

For the derivation of the dispersion relation, it is assumed that the charge density of a cylindrical beam is uniform and the usual linearization can be made. It follows from Maxwell's equations that the axial electric field, E, must be a solution of the wave equation. In a cylindrical coordinate system, the wave equation, neglecting azimuthal variations, is

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial E}{\partial r}\right)+\frac{\partial^2 E}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}=\mu_0\frac{\partial i_1}{\partial t}+\frac{1}{\epsilon_0}\frac{\partial \rho_1}{\partial z}. \qquad (1)$$

The small-signal current and charge density are designated by i_1 and ρ_1, respectively. Assuming a wave propagation of the form $\exp j(\omega t-\beta z)$, one obtains, from the linearized continuity equation, the relation

$$i_1=\rho_1\frac{\omega}{\beta}. \qquad (2)$$

Linearization of the current equation $i=\rho v$, with use of Equation (2), yields

$$v_1=\frac{\omega-\beta v_0}{\beta\rho_0}\rho_1, \qquad (3)$$

where v_0 is the average drift velocity. Furthermore, the linearized force equation of Reference (6) for a drifting beam becomes

$$j(\omega-\beta v_0)v_1=\eta E+j\beta\frac{v_T^2}{\rho_0}\rho_1, \qquad (4)$$

[6] J. Berghammer and S. Bloom, "On the Nonconservation of Noise Parameters in Multivelocity Beams," *Jour. Appl. Phys.*, Vol. 31, p. 454, March, 1960.

where v_T characterizes the thermal velocity, as defined later. Introducing the exponential factor $\exp j(\omega t - \beta z)$ into Equation (1) and using Equations (3) and (4), leads to

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dE}{dr}\right) + \gamma^2 E = 0, \tag{5}$$

where γ is the radial propagation constant (see Equation (6), below). The solution of this equation is

$$E = AJ_0(\gamma r) + BN_0(\gamma r),$$

where J_0 stands for the zero-order Bessel function and N_0 for the zero-order Neumann function. For a solid beam the integration constant, B, is zero and A must be determined from the boundary condition at $z = 0$. The square of the radial propagation constant is given by

$$\gamma^2 = \left(\beta^2 - \frac{\omega^2}{c^2}\right)\left[\frac{{\omega_p}^2}{(\omega - \beta v_0)^2 - {v_T}^2\beta^2} - 1\right], \tag{6}$$

where the plasma frequency is $\omega_p = (\eta\rho_0/\epsilon_0)^{1/2}$. The radial boundary condition determines the value of γ^2. For the present purpose, it is sufficient to mention that the radial propagation constant for a beam surrounded by a metallic drift tube of radius a is inversely proportional to the beam radius b, (for $a/b = 1$, $\gamma = 2.4/b$).

For an infinite beam, $\gamma = 0$, the dispersion relation, Equation (6), yields the roots

$$\beta_{1,2} = \left[\frac{\omega}{v_0} \pm \frac{\omega_p}{v_0}\sqrt{1 - \sigma^2}\right)\left(1 - \frac{\omega^2}{{\omega_p}^2}\right]\left(1 - \sigma^2\right)^{-1} \tag{7}$$

where $\sigma = v_T/v_0$ is a measure of the relative width of the distribution function. The exact treatment of multivelocity problems is based on the linearized Boltzmann equation without collisions which leads, via the open-circuit relation, to the well-known dispersion integral

$$\frac{\eta}{\epsilon_0}\int \frac{f_0(u)}{(\omega - \beta u)^2}\,du = 1. \tag{8}$$

The density of electrons having velocities between u and $u+du$ is characterized by the d-c distribution function $f_0(u)$. It is of interest to note that the dispersion integral, Equation (8), leads to exactly the same roots as those obtained above from the macroscopic equations, provided that a rectangular distribution function is used. This distribution function is defined by

$$f_0(u) = \frac{\rho_0}{2v_T}\left[S(u-u_s) - S(u-u_f)\right], \tag{9}$$

where ρ_0 is the d-c charge density of the electrons, $u_{s,f}$ signifies the slowest and fastest electron velocity, and S denotes the unit step function. The thermal velocity, $v_T = (u_f - u_s)/2$, characterizes the width of the velocity distribution function, and the average drift velocity is given by $v_0 = (u_s + u_f)/2$.

To investigate the dispersion relation for the finite beam, $\gamma \neq 0$ in Equation (6), it is convenient to use the normalized variables ω/ω_p, $\beta\lambda_D$ and

$$\theta = \gamma\lambda_D, \tag{10}$$

where λ_D stands for the Debye length,

$$\lambda_D \equiv \frac{v_T}{\omega_p}. \tag{11}$$

It is also useful to introduce the ratio of doppler-shifted frequency to plasma frequency

$$p \equiv \frac{\omega - \beta v_0}{\omega_p} = \frac{\omega}{\omega_p} - \frac{\beta\lambda_D}{\sigma}, \tag{12}$$

which is commonly known from single-velocity theory as the plasma-frequency reduction factor.

With the slow-wave assumption ($\beta \gg \omega/c$), the dispersion relation, Equation (6), can be written in terms of the normalized variables defined by Equations (10-12) as

$$p^2 = (\beta\lambda_D)^2\left[1 + \frac{1}{(\beta\lambda_D)^2 + \theta^2}\right]. \tag{13}$$

For $v_T = 0$, this relation reduces to the results known from single-velocity theory. The behavior of Equation (13) is shown schematically in Figure 1. It is seen that the one-dimensional dispersion relation, labelled $\theta^2 = 0$, is asymptotically approached by the two branches corresponding to a finite θ^2, i.e., for a finite beam diameter. Branch A lies in the right half-plane where $(\beta\lambda_D)^2 > 0$. The propagating waves associated with this branch correspond to the well-known Hahn–Ramo waves. Their propagation constants are, however, modified by the presence of the velocity distribution of the electrons. In general, in the region where $(\beta\lambda_D)^2 < 0$, a pair of complex conjugate

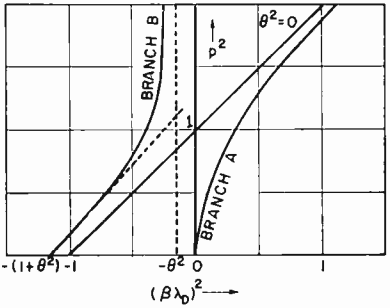


Fig. 1—Schematic representation of the dispersion relation for a finite ($\theta^2 \neq 0$) and an infinite ($\theta^2 = 0$) multivelocity beam.

propagation constants is indicated. These new waves arise only because of the assumption of a *finite diameter* beam. One of the waves associated with the propagation constants of branch B is an evanescent, or reactively damped, wave, with $\mathrm{Im}(\beta\lambda_D) < 0$, according to well-known criteria.[7] The other wave, which propagates with the complex conjugate propagation constant (increasing amplitude in the positive z-direction), corresponds to a reflected wave which is reactively damped in the negative-z direction. This wave will not be excited unless there are discontinuities in the propagating medium or in its environment. The appearance of these evanescent waves is reminiscent of the reactively attenuated waves appearing in a waveguide below cutoff. It can easily be verified from the dispersion relation, Equation (6), that for $v_0 = 0$, an infinite stationary plasma of electrons does indeed have a dispersion similar to that of a high-pass filter, with its cutoff frequency given by the plasma frequency.

[7] P. A. Sturrock, "Kinematics of Growing Waves," *Phys. Rev.*, Vol. 112, p. 1488, December, 1958.

The discussion which follows is concerned with branch B of the dispersion relation. For a *"nonmoving"* beam, $v_0 = 0$, p reduces to ω/ω_p, and branch B of Figure 1 represents a reactively damped wave for all frequencies. Its attenuation constant decreases with increasing frequency. For practical purposes, it is necessary to inquire into the condition for which the reactively damped wave exists in a *moving* beam. For the infinite beam, $\theta = 0$, the condition

$$\sigma^2 \geqq \left(1 - \frac{\omega^2}{\omega_p{}^2}\right)^{-1} > 0 \tag{14}$$

follows from Equation (7). The radicand appearing in the propaga-

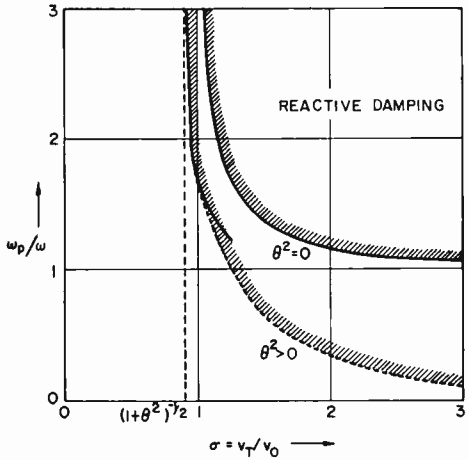


Fig. 2—Limiting curves for reactive damping in an infinite beam, $\theta^2 = 0$, and in a finite beam, $\theta^2 > 0$.

tion constant of Equation (7) becomes zero when the equality sign holds. This limiting curve is shown in Figure 2. It is seen that the necessary conditions for the existence of reactive damping are $\sigma > 1$ and $\omega < \omega_p$. How is this condition affected by the finite transverse dimension of the beam? Branch B of Figure 1 can be approximated near $p^2 = 0$ by a straight line (dotted line in the Figure). This approximation leads to the propagation constants

$$\beta = \left\{\frac{\omega}{v_0} \pm \frac{\omega_p}{v_0}\left([1+\theta^2]^2 - \sigma^2[1+\theta^2]\left[(1+\theta^2)^2 - \frac{\omega^2}{\omega_p{}^2}\right]\right)^{1/2}\right\} \cdot \left(1 - \sigma^2[1+\theta^2]\right)^{-1}. \tag{15}$$

The condition for β to be complex is consequently modified by the finite beam geometry and is now given by

$$\sigma^2 \geqq \left(1+\theta^2\right)\left[\,(1+\theta^2)^2 - \frac{\omega^2}{\omega_p{}^2}\,\right]^{-1}. \tag{16}$$

This result is applicable only for $\omega/\omega_p \ll 1$, and the limiting curve, Equation (16), is shown in Figure 2. In most practical cases, the value of $\theta = 2.4\ \lambda_D/b$ is of the order of a few tenths or smaller. The finiteness of the beam thus allows the reactive damping region to extend slightly into the region $\sigma^2 < 1$, at low frequencies.

In many practical applications, it can be assumed that the beam drifts with almost zero mean velocity. The dispersion equation, Equation (13), gives for $v_0 = 0$, (i.e., $\sigma = \infty$),

$$-(\beta\lambda_D)_0{}^2 = \frac{1}{2}\left(1+\theta^2 - \frac{\omega^2}{\omega_p{}^2}\right) \pm \frac{1}{2}\left(\left[1+\theta^2-\frac{\omega^2}{\omega_p{}^2}\right]^2 + 4\theta^2\frac{\omega^2}{\omega_p{}^2}\right)^{1/2}. \tag{17}$$

The subscript zero refers to the case of zero mean velocity. For branch B of Figure 1, the plus sign in Equation (17) applies. It is seen that $(\beta\lambda_D)_0$ is imaginary for all frequencies. The effect of a small drift-velocity, $v_0 \ll v_T$, on the propagation constant can be found from Equation (13). For this purpose the following assumption is made:

$$\beta\lambda_D = (\beta\lambda_D)_0(1+\delta). \tag{18}$$

Assuming that $\delta \ll 1$, one obtains

$$\delta = -\frac{1}{\sigma}\frac{\omega}{\omega_p}\frac{(\beta\lambda_D)_0}{1+(\beta\lambda_D)_0{}^2 - \left[\dfrac{\omega^2}{\omega_p{}^2} - (\beta\lambda_D)_0{}^2\right]^2}. \tag{19}$$

Since δ is purely imaginary, it follows that a small drift velocity does not affect the imaginary part of the propagation constant to first order in $1/\sigma$, but merely introduces a phase factor. This result leads to the conjectured continuation of Equation (16) to higher frequencies, as indicated by the dotted line in Figure 2.

Noise Smoothing by Reactive Damping

The fact that a reactively damped wave can exist in a low-velocity drift space suggests that this wave may in part be responsible for the noise-smoothing observed in some of the modern ultra-low-noise traveling-wave tubes. A complete study of the possible effectiveness of this wave for noise-smoothing purposes would require a detailed consideration of the noise excitation mechanism. However, some insight can be gained by analyzing the following simple model.

The multivelocity region in front of the cathode is approximated by a drift space of length L. The electron beam in this region drifts very slowly and furthermore includes returning electrons, $v_0 \ll v_T$. These conditions are approximately realized if, for example, a virtual cathode is established due to space-charge instability[8] in the cathode–first-anode region of the electron gun. To obtain the most optimistic estimate of reactive damping effects on the noise figure, it is assumed that only the reactively damped wave is excited by full shot-noise current fluctuations at the cathode surface, i.e., at $z = 0$.

According to these assumptions, the noise current at a distance L from the cathode, if there is no reflected wave, is given by

$$i = i_s \exp\left\{\frac{-|\beta\lambda_D|_0(1+\delta)L}{\lambda_D}\right\}, \tag{20}$$

where the imaginary quantity $(\beta\lambda_D)_0$ is given by Equation (17). The factor i_s corresponds to full shot noise at the cathode and the imaginary quantity δ is given by Equation (19). Therefore, the noise reduction factor, R, is

$$R \equiv \left(\frac{|i|^2}{|i_s|^2}\right)^{1/2} = \exp\left\{\frac{-|\beta\lambda_D|_0 L}{\lambda_D}\right\}. \tag{21}$$

Ignoring any beneficial correlation effects which arise in multivelocity flow, one finds for the minimum noise figure of a traveling-wave tube[9] (with $QC = 0 = d$),

$$F_{\min} = 1 + R\frac{T_c}{T_0}, \tag{22}$$

---

[8] J. Berghammer, "Space-Charge Effects in Ultra-Low-Noise Electron Guns," *RCA Review*, Vol. 21, p. 369, September, 1960.

[9] S. Bloom and R. W. Peter, "A Minimum Noise Figure for the Traveling-Wave Tube," *RCA Review*, Vol. 15, p. 252, June, 1954.

where T_c/T_0 is the ratio of cathode to reference temperature.

The reduction factor as calculated from Equation (21) is shown as a function of $(\omega/\omega_p)^2$ in Figure 3. It is seen that in the case of an infinite beam, $\theta = 0$, R is appreciably smaller than unity at frequencies below the plasma frequency, becomes unity at the plasma frequency, and remains at unity for higher frequencies. For a finite beam, however, the curves labelled $\theta^2 = 0.01$ show that smoothing occurs above as well as below the plasma frequency. This noise reduction becomes more effective for an increased length of the low-velocity drift space.

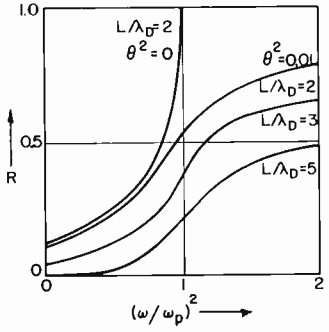


Fig. 3—Noise reduction factor, R, for a finite beam of length L, $\theta^2 = 0.01$, and for an infinite beam, $\theta^2 = 0$.

Conclusion

The propagation of space-charge waves on a beam with a rectangular velocity distribution has been considered. It is shown that four waves are possible, a pair of Hahn–Ramo waves and a pair of evanescent waves, both of which decay, one in the forward direction, the other in the backward direction. If it is assumed that only the forward evanescent wave is excited by shot noise at the cathode, the noise current is reduced by the factor

$$R = \exp\left\{\frac{-|\beta\lambda_D|_0 L}{\lambda_D}\right\} \tag{21}$$

as a function of L, the distance from the cathode. To make R small, the following conditions are necessary:

(a) The plasma frequency must be equal to or greater than the signal frequency. This condition implies a high current density at microwave frequencies,

(b) The beam must drift very slowly and include returning electrons ($v_0 \ll v_T$), as provided, for example, by a virtual cathode,

(c) The beam must be of finite transverse dimensions ($\theta^2 \neq 0$) if the signal frequency is above the plasma frequency,

(d) The low-velocity drift space must be long compared to the Debye length.

ACKNOWLEDGMENT

The author wishes to thank S. Bloom for many valuable suggestions and F. Paschke for several interesting discussions.

FERROELECTRIC SCANNING OF ELECTROLUMINESCENT DISPLAYS*

BY

MICHAEL COOPERMAN

RCA Home Instruments Division
Camden, N. J.

Summary—This paper describes a new method for scanning flat electroluminescent display devices. The basic element of this scanning system is a barium titanate ($BaTiO_3$) crystal which is activated to switch video information to the appropriate element of the display device. In order to demonstrate the principle, an array of five lines, each having five elements, was addressed with this technique. The circuits are simple in construction and low in power consumption.

This scheme, although it looks promising, is in its infancy and a number of problems remain to be solved before it can be considered practical. The immediate and most important problem is lack of brightness of the displayed information. This condition becomes more severe as the desired picture detail, and consequently the number of picture elements, are increased. As electroluminescent materials with more brightness are developed, this problem will become less serious.

INTRODUCTION

IN RECENT YEARS a considerable amount of research aimed at developing a solid-state display device has been carried out.[1-4] Such a device might replace the cathode-ray tube in industrial television and military display applications. Among the attractive features of a solid-state display are its small volume and weight for a given viewing area.

One of the most difficult problems encountered in the design of a flat display is to find a suitable means for scanning, i.e., distributing the video information to the appropriate spot on the panel. Schemes

* Manuscript received 6 January 1961.

[1] Jess J. Josephs, "A Review of Panel-Type Display Devices," *Proc. I.R.E.*, Vol. 48, p. 1380, August 1960. (This reference contains a very extensive bibliography on different type display devices.)

[2] J. A. Rajchman, G. R. Briggs, and A. W. Lo, "Transfluxor Controlled Electroluminescent Display Panels," *Proc. I.R.E.*, Vol. 46, p. 1808, November 1958.

[3] Michael Lorant, "Electroluminescent Pattern Display," *Wireless World*, p. 606, December 1958.

[4] R. C. Lyman and C. I. Jones, "Electroluminescent Panels for Automatic Displays," *Electronics*, Vol. 32, No. 28, p. 44, July 10, 1959.

to accomplish this are usually complex, power consuming, and unsuitable for mass production.

This paper describes a scheme which uses ferroelectric materials for switching sequentially the video information to a flat electroluminescent panel. This scanning scheme is low in power consumption and, due to its simplicity, is suitable for mass production. The paper also describes an experimental model employing this principle for scanning.

Principle of Operation

The basic circuit for using a ferroelectric material for sequential switching is shown in Figure 1. The ferroelectric capacitor C_f uses a $BaTiO_3$ crystal for a dielectric. The characteristic of C_f is shown in Figure 2a as capacitor charge versus applied voltage. The capacitor

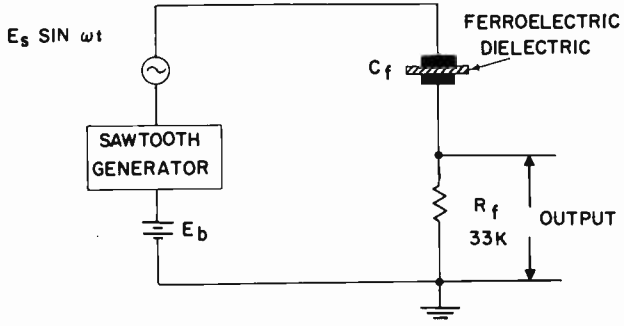


Fig. 1—Basic circuit for sequential switching.

has a hysteresis loop similar to that observed in magnetic materials. The hysteresis loop of Figure 2a has been idealized in order to simplify the analysis of switching.

Let it be assumed for the moment that in Figure 1, $E_b = 0$. C_f and R_f are so proportioned that the sawtooth produces a negligible drop across R_f. Consequently, most of the sawtooth appears across C_f. The voltage drop across C_f produced by the sawtooth and the battery voltage is shown by curve 2 in Figure 2b. As the sawtooth voltage goes through amplitudes a,b,c,d,e,f, and back to a, the hysteresis loop is traversed in the order a,b,c,d,e,f,a. As the hysteresis loop is traversed, the reactance of C_f is changed so that it is relatively high in the region a to b and c to d and relatively low from b to c. (The reactance is also changed in the region d to a, but this is of no interest here.) The ratio of high to low reactance can easily be made 100:1.

R_f and C_f are so proportioned that $E_s \sin \omega t$, whose frequency is high compared to the sawtooth fundamental and its harmonics, appears across R_f only during the time t_3 to t_4, corresponding to the interval b to c, when the reactance of C_f is low. It has been assumed in this analysis that the amplitude E_s is sufficiently small that it swings over a small portion of the characteristic of Figure 2a.

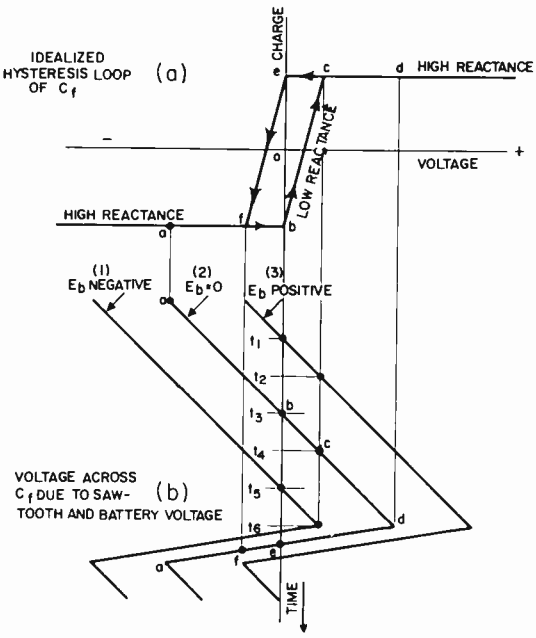
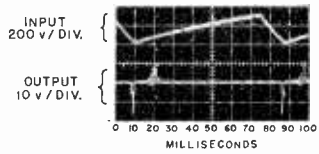


Fig. 2—Graphical analysis of sequential switching.

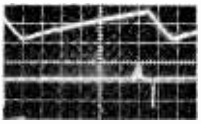
If E_b is positive, the sawtooth voltage is shifted to the right as shown in curve 3 of Figure 2b. This causes C_f to conduct earlier, namely from t_1 to t_2. If E_b is negative, the sawtooth voltage is shifted to the left, as shown in curve 1 of Figure 2b. Under these conditions C_f conducts later, namely from t_5 to t_6. Thus the time of conduction relative to the sawtooth timing is a function of the d-c bias, E_b.

Actual photographs of the input and output waveforms of the circuit in Figure 1 are shown in Figure 3. The input consists of the sawtooth in series with $E_s \sin \omega t$. The output is a burst of $E_s \sin \omega t$. Photographs (a), (b), and (c) show the relative timing between the output burst and the sawtooth for positive, zero, and negative bias conditions. The negative spikes in the output waveforms are due to the conduction of C_f during the sawtooth retrace. The frequency of



(a). BIAS E_b = + 100 VOLTS

(b). BIAS E_b = 0 VOLTS



(c). BIAS E_b = - 100 VOLTS

Fig. 3—Waveforms of input and sequentially switched output of circuit shown in Figure 1.

E_s would normally be made higher than shown in the photographs. Here the frequency was reduced so it could be observed cycle by cycle.

Experimental Circuits

Line Scanning

An experimental circuit using this principle for sequential switching and scanning one line is shown in Figure 4. This is essentially

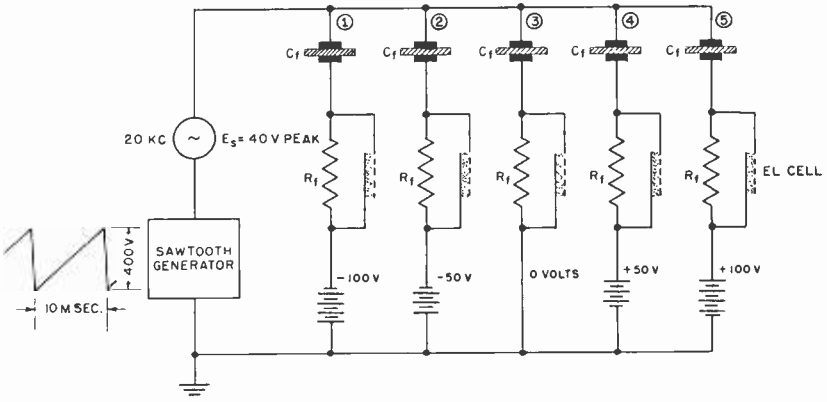


Fig. 4—Experimental circuit for a five-element scanning line.

the same circuit as that of Figure 1. Here, however, the sawtooth in series with $E_s \sin \omega t$ is applied across five $C_f R_f$ combinations connected in parallel. These are biased so that the ferroelectric capacitors conduct in the sequence of the indicated numbers. An electroluminescent cell is connected across each of the resistors. A typical brightness characteristic of an electroluminescent cell is shown in Figure 5. As each C_f conducts, its corresponding electroluminescent cell emits

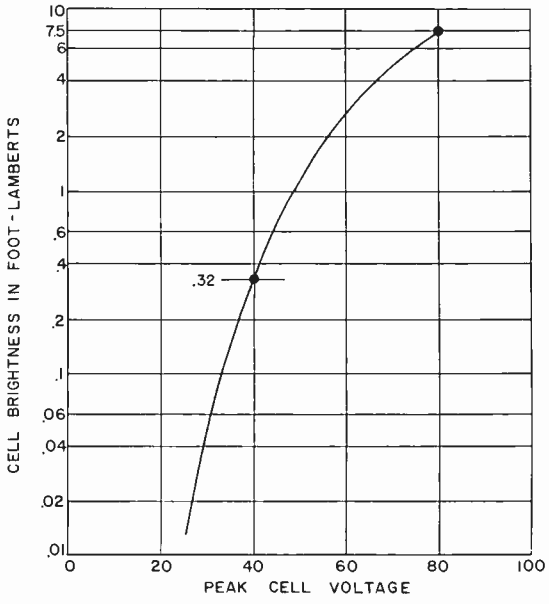


Fig. 5—20-kilocycle sinusoidal excitation of electroluminescent cell.

light for the period of conduction. The five electroluminescent cells comprising a scanning line are triggered sequentially. If $E_s \sin \omega t$ is amplitude-modulated with video information, the brightness of each cell will be proportional to the instantaneous amplitude of E_s. Thus the information with which $E_s \sin \omega t$ is modulated is displayed as the line is scanned.

A more practical arrangement of the circuit of Figure 4 is shown in Figure 6. Here the individual ferroelectric capacitors are formed by using a single strip of ferroelectric material with one common connection. Each capacitor formed by the bottom connection and the common top conductor can be operated with essentially no interaction with the rest. In the actual experiment the bottom connections were spaced 1/16 inch apart. The various bias voltages necessary for se-

quential switching are obtained from a bias distributor. The a-c impedance at the various taps of the bias distributor must be sufficiently low to prevent crosstalk between cells. This can be accomplished by making the resistors in the bias distributor small. The 4,000-ohm resistors used here were found to be quite satisfactory.

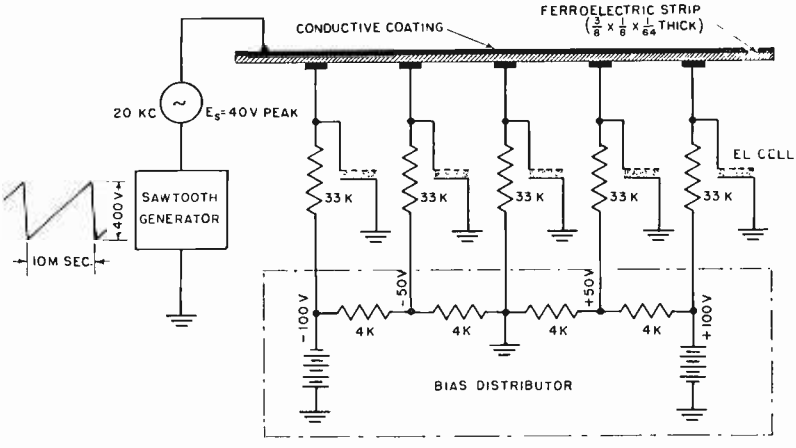


Fig. 6--Experimental circuit for a five-element scanning line.

Horizontal and Vertical Scanning

Both horizontal and vertical scanning can be accomplished by using the circuit of Figure 6 for horizontal scanning and another such circuit for vertical scanning. Such an arrangement is shown in Figure 7. The H conductors are switched at a horizontal rate and the V conductors at a vertical rate.

Electroluminescent cells, represented by equivalent capacitors, are connected between sets of conductors. The voltage across each cell is the difference between the potentials of its intersecting H and V conductors.

For the purpose of illustration, it will be assumed that the ferroelectric strips are perfect switches, (i.e., when conduction takes place, all of E_s is connected to the switched conductor). If, for example, at a particular instant the H_3 and V_3 conductors are switched, then the potential across the cell connected between them is $2E_s = 80$ volts peak at that instant. The doubling of the voltage is due to E_s in the horizontal circuit being out of phase with E_s in the vertical circuit. At the same time all remaining cells connected to the H_3 conductor and V_3 conductor can have a maximum voltage across them of $E_s = +40$

volts peak and $-E_s = -40$ volts peak, respectively. The electroluminescent cells have a typical brightness characteristic as shown in Figure 5. It can be seen that a voltage change from 40 to 80 volts peak produces a brightness change of .32 to 7.5 foot-lamberts. That is, a 2:1 change in voltage produces a 23:1 change in brightness, which is sufficient to turn a cell on and off. With proper design, the cell between two switched conductors will emit light and the remaining cells will appear dark.

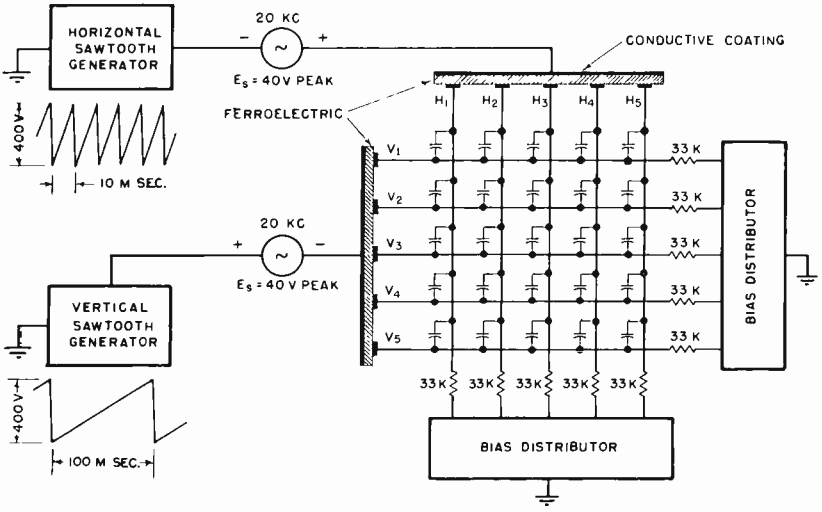


Fig. 7—Experimental circuit for horizontal and vertical scanning.

Thus, as the H and V conductors are switched in sequence, scaning proceeds from left to right horizontally, and from top to bottom vertically. Information can be displayed by modulating E_s in the vertical scanning circuit, horizontal scanning circuit, or in both.

A more practical arrangement of the interconnection between H and V conductors and electroluminescent cells of Figure 7, is shown in Figure 8. It consists of a glass plate having a set of transparent V conductors printed on it. Next comes a layer of electroluminescent phosphor which is then covered by a set of H conductors. When an a-c voltage is applied between an H and V conductor, the phosphor at their interaction emits light. Due to the high resistivity of the electroluminescent phosphor, only the area sandwiched between the two intersecting conductors emits. Such panels are available commercially.

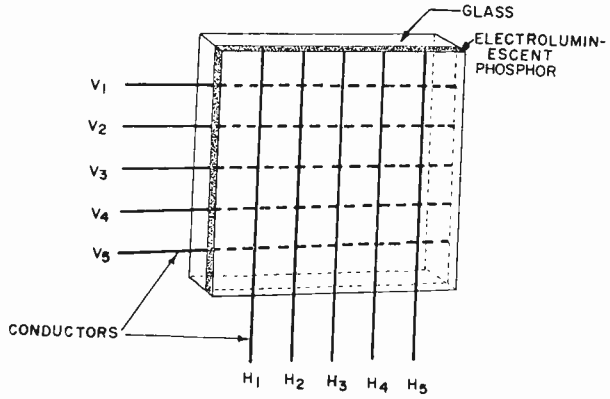


Fig. 8—Commercial electroluminescent panel for horizontal and vertical scanning.

Proposed Circuits

Line Scanning

An alternative arrangement using ferroelectric switching in conjunction with electroluminescence to form a display is shown in Figure 9. A strip of glass is covered with a transparent conductor, a layer of electroluminescent phosphor, a layer of ferroelectric material, and a coating of resistive paint. At each end of the resistive coating is a conductor for contact with terminals a and c. This arrangement can provide continuous scanning of one line. The equivalent circuit of this arrangement is shown in Figure 10. The principle of operation is the same as that of Figure 6. Resistors R_1 which represent the resistive coating provide the d-c potential distribution for sequential switching. These resistors are sufficiently small so that there is neg-

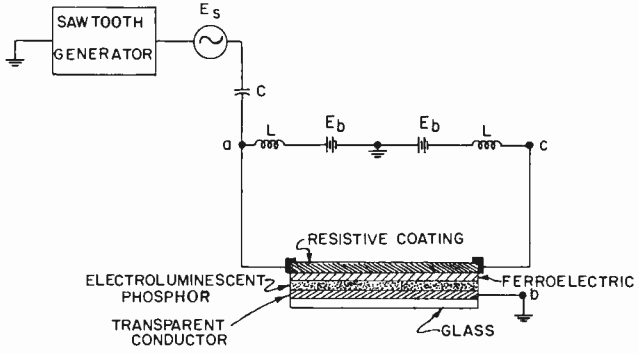


Fig. 9—Suggested construction for single-line continuous scanning.

ligible a-c potential difference between the top terminals of C_f. C_f represents the ferroelectric capacitance while R_f and R_2 represent the resistivity of the electroluminescent phosphor. The inductors L prevent the sawtooth and the a-c signal from shorting to ground.

The capacitor C prevents the d-c bias from being shorted to ground. Thus the bias distributor circuit and the scanning circuit do not interfere with each other's functions. If R_2 is considered infinite, then the remainder of the circuit and its function becomes identical

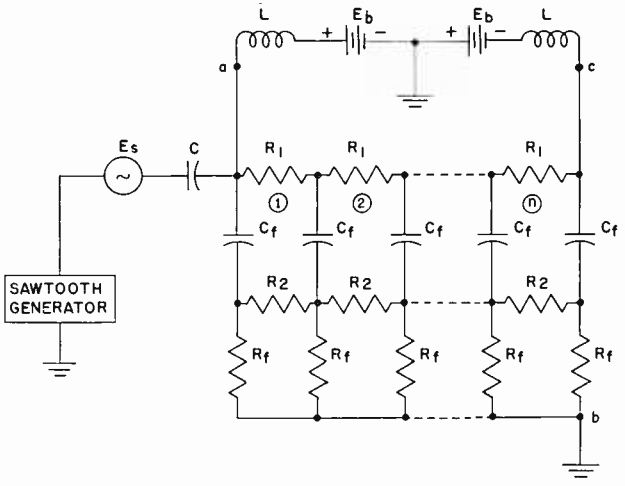


Fig. 10—Equivalent circuit for Figure 9.

to that of Figure 6. As the ferroelectric capacitors are switched, the a-c signal appears across their corresponding resistances R_f, and continuous scanning of the strip results. The presence of R_2 does not alter the principle of operation, it merely provides some interaction between adjacent R_f resistors, thus limiting the resolution.

Horizontal and Vertical Scanning

A configuration for horizontal and vertical scanning is shown in Figure 11. The horizontal strips with terminals a, b, and c are identical to those in Figure 9. The horizontal sawtooth and biasing are applied to each strip in the same manner. Terminals b are connected to a vertical scanner. This consists of a ferroelectric switch which is similar in construction and operation to that described in Figure 6. The vertical sawtooth switches E_s to each R_f and corresponding scanning line in succession. The horizontal sawtooth switches all lines at the same time. However, only the line which receives E_s from the

vertical scanner produces a visible scan. Here again information is displayed by amplitude-modulating E_s.

Brightness Limitation

With presently available electroluminescent materials, about 60 volts peak of continuous excitation is required to light a cell. This

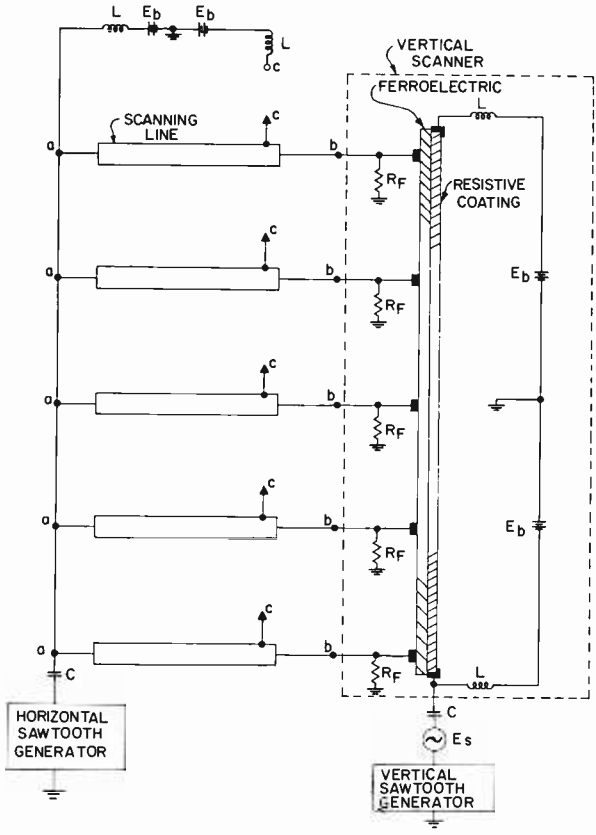


Fig. 11—Suggested scheme for horizontal and vertical scanning.

voltage is in the order of magnitude obtainable with the scanning technique described. In order to produce a larger output, the thickness of the ferroelectric material and the amplitude of the sawtooth would have to be increased.

However, for a given excitation amplitude, the brightness is also a function of the portion of time during which a cell is excited. If five elements are scanned, then each element is excited for about one-

fifth of the time. For ten elements this time is about one-tenth, etc. Consequently, as the number of elements is increased, the brightness for a given excitation voltage is reduced. One method of increasing brightness is to have some external circuitry which would maintain the excitation of each cell once it has been briefly excited.[5-7]

ACKNOWLEDGMENT

The author wishes to express his gratitude to his colleagues at Cherry Hill, Princeton, and Lancaster for their support and cooperation. Particular thanks are due to C. M. Sinnett and G. L. Grundmann under whose supervision this work was done.

---

[5] B. Kazan and F. H. Nicoll, "An Electroluminescent Light-Amplifying Picture Panel," *Proc. I.R.E.*, Vol. 43, p. 1888, December 1955.

[6] E. A. Sack, "ELF—A New Electroluminescent Display," *Proc. I.R.E.*, Vol. 46, p. 1694, October 1958.

[7] D. W. G. Ballentyne, "The Phenomenon of Electroluminescence and its Application in the Electronics Industry," *The Marconi Review*, No. 123, 4th Quarter 1956, Vol. XIX, p. 160.

RCA TECHNICAL PAPERS†

Fourth Quarter, 1960

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

"Blueprint for In-Flight Maintenance of Manned Spacecraft," T. Taylor, Jr., H. B. Matty, ad W. G. Sommer, *Space Aeronautics* (October) 1960

"Cathode-Ray Tube," M. D. Harsh, *Ecyclopedia of Science and Technology* (October) 1960

"Digital Computers and Medical Logic," R. Ebald ad R. Lane, *Trans. I.R.E. PGME* (October) 1960

"Effect of Pile Irradiation on the Dielectric Properties of Triglycine Sulfate Single Crystals," E. Fatuzzo, *Helvetica Physica Acta* (October) 1960

"Electrical and Optical Properties of Zinc Sulfide Crystals in Polarized Light," J. A. Beun and G. J. Goldsmith, *Helvetica Physica Acta* (October) 1960

"Electrode Potential," R. Glicksmann *Encyclopedia of Science and Technology* (October) 1960

"Electronic Metronome," H. Wittlinger, *Electronics World* (October) 1960

"Kinescope," C. P. Smith, *Encyclopedia of Science and Technology* (October) 1960

"A Note on a Method of Computing the Gamma Function," N. L. Gordon and A. H. Flasterstein, *Jour. of the Association for Computing Machinery* (October) 1960

"Oscillations in Germanium with an Applied Pulsed Electric Field," M. Cardona and W. Ruppel, *Jour. Appl. Phys.* (October) (Letters to the Editor) 1960

"Phototube," J. L. Weaver, *Encyclopedia of Science and Technology* (October) 1960

"Phototube Multiplier," J. L. Weaver, *Encyclopedia of Science and Technology* (October) 1960

"Reliability: an Engineering Problem," M. M. Tall, *Industrial Research* (October-November) 1960

"Spectrographic Analysis of Indium Metal for Impurities," H. M. Hyman, *Applied Spectroscopy* (October) 1960

"Stability with Simplicity," G. D. Hanchett, *QST* (October) 1960

"Storage Tube," M. D. Harsh, *Encyclopedia of Science and Technology* (October) 1960

"Television Camera Tube," R. D. Neuhauser, *Encyclopedia of Science and Technology* (October) 1960

"Theory on the Receptor Mechanism in Color Vision," A. C. Schroeder, *Jour. Opt. Soc. Amer.* (October) 1960

"Variational Treatment of Warm Electrons in Nonpolar Crystals," I. Adawi, *Phys. Rev.* (October 1) 1960

"Modern Microwave Traveling-Wave Tubes," H. K. Jenny, *Bulletin of Swiss Electrical Society* (October 8) 1960

"Practical Tubes for Bright Radar Displays," F. S. Veith, *Bulletin of Swiss Electrical Society* (October 8) 1960

"Status Report on the Environmentalization of Traveling-Wave Tubes," J. S. Posner, *Electronic Design* (October 12) 1960

"Autopilot System for Missiles," J. H. Porter, *Electronics* (October 21) 1960

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

"A Compatible Stereophonic System for the AM Broadcast Band," J. Avins, *Trans. I.R.E. PGBTR* (November) 1960
"Construction of a Thermionic Energy Converter," F. G. Block, F. H. Corregan, G. Y. Eastman, J. R. Fendley, K. G. Hernqvist, and E. J. Hills, *Proc. I.R.E.* (November) 1960
"The Effect on Color TV of Dr. Edwin H. Land's Color Experiments," C. J. Hirsch, *Trans. I.R.E. PGBTR* (November).... 1960
"Factors in the Circuit Operation of Transistorized Horizontal Deflection," R. W. Ahrons, *Trans. I.R.E. PGBTR* (November) 1960
"Measurement of Ratio of Absorptivity of Sunlight to Thermal Emissivity," G. D. Gordon, *Rev. Sci. Instr.* (November).... 1960
"Narrow-Bandwidth Video-Tape Recorder Used in the Tiros Satellite," J. A. Zenel, *Jour. S.M.P.T.E.* (November) 1960
"A Nuvistor, Low-Noise VHF Tuner," G. C. Hermeling, *Trans. I.R.E. PGBTR* (November) 1960
"Phases in the Photoelectric Sodium-Potassium-Antimony System," W. H. McCarroll, *Jour. Phys. & Chem. of Solids* (November) 1960
"Phosphor Efficiency at Very Low Excitation Current Densities," G. W. Francis and R. G. Stoudenheimer, *Rev. Sci. Instr.* (November) (Notes) 1960
"Plasma Acceleration in a Radio-Frequency Field Gradient," G. A. Swartz, T. T. Reboul, G. D. Gordon, and H. W. Lorber, *Jour. of Phys. of Fluids* (November-December) 1960
"A Small RF-Signal Theory for an Electrostatically Focused Traveling-Wave Tube," W. W. Siekanowicz, *Proc. I.R.E.* (November) 1960
"Standards for Intelligibility Testing," M. E. Hawley, *The Magazine of Standards* (November) 1960
"A Two-Color Input, Two-Color Output Image Intensifier Panel," F. H. Nicoll and A. Sussman, *Proc. I.R.E.* (November).... 1960
"Use and Abuse of Tolerances," D. Colasanto, *Production* (November) 1960
"Cartesian Tensor Scalar Product and Spherical Harmonic Expansions in Boltzmann's Equation," T. W. Johnston, *Phys. Rev.* (November 15) 1960
"Amplification-Modern Trends, Techniques and Problems," L. S. Nergaard, *RCA Review* (December) 1960
"Conditional-Switching Terminal Guidance (A Terminal Guidance Technique for Satellite Rendezvous)," A. L. Passera, *Trans. I.R.E. PGANE* (December) 1960
"Construction of Relatively Maximal, Systematic Codes of Specified Minimum Distance from Linear Recurring Sequences of Maximal Period," C. N. Campopiano, *Trans. I.R.E. PGIT* (December) 1960
"Design Characteristics of Nonvented Nickel-Cadmium Cells," L. M. Krugman, *Trans. I.R.E. PGCP* (December) 1960
"Electron-Optical Properties of a Flat Television Picture Tube," E. G. Ramberg, *Proc. I.R.E.* (December) 1960
"Generation of a Sampled Gaussian Time Series Having a Specified Correlation Function," M. J. Levin, *Trans. I.R.E. PGIT* (December) 1960
"High-Frequency Varactor Diodes," C. W. Mueller and R. D. Gold, *RCA Review* (December) 1960
"Integration of Current through Getter-Ion Pumps," J. R. Fendley, Jr., *Rev. Sci. Instr.* (December) (Notes) 1960
"The Management of Research Services," J. Kurshan, *Trans. I.R.E. PGEM* (December) 1960
"Mechanism of Photoconductivity in Microcrystalline Powders," R. H. Bube, *Jour. Appl. Phys.* (December) 1960
"On the Molecular Nature of the Dielectric Anomalies in Thiourea," C. Calvo, *Jour. Chem. Phys.* (December) 1960

"Photoemission and Related Properties of the Alkali-Antimonides," W. E. Spicer, *Jour. Appl. Phys.* (December) 1960
"A Plug-Type Image Orthicon Target," S. A. Ochs, *RCA Review* (December) 1960
"Point-Pair Reading Logig," J. M. Bailey, Jr., *Proc. I.R.E.* (December) (Correspondence) 1960
"Semiconductor Devices—Their Status and Future," E. O. Johnson, *Space Aeronautics* (December) 1960
"Test Chassis for VHF Measurements," M. Fomin, *Electronic Equipment Engineering* (December) 1960
"Time-Averaged Effects on Charged Particles in A-C Fields," T. W. Johnston, *RCA Review* (December) 1960
"Tunnel Diode Logic Circuits," R. H. Bergman, *Trans. I.R.E. PGEC* (December) 1960
"The Ultra-High-Vacuum System for the C-Stellarator," K. Dreyer and J. T. Mark, *RCA Review* (December) 1960
"Using the 7360 in the HBR-16," J. M. Filipczak, *QST* (December) 1960
"Microcircuit Binary Full Adder Uses Unipolar Transistors," M. E. Szekely, J. T. Wallmark, and S. M. Marcus, *Electronics* (December 23) 1960
"A Theory of Communications in a Research Laboratory," J. Hillier, *Research Management* (Winter) 1960
"Correlation of Semiconductivity, Photoconductivity, and Luminescence in Group II-Group VI Materials," R. H. Bube, *Solid State Physics in Electronics and Telecommunication; Semiconductors*, Vol. 2, Part 2, p. 825, Academic Press, Inc., London 1960
"Design Reliability Analysis—A Proven Technique for Product Control and Enhancement," H. L. Wuerffel and D. I. Troxel, *6th National Symposium on Reliability and Quality Control, Proceedings* 1960
"Imperfection Ionization Energies in CdS-Type Materials by Photoelectronic Techniques," R. H. Bube, *Solid State Physics*, Vol. 11, p. 223, Academic Press, N. Y. 1960
"A Note on the Applicability of Error Correcting Codes," J. E. Palmer, *4th National Convention on Military Electronics, Conference Proceedings* 1960
"Practical Maintainability Numerics," M. P. Feyerherm and H. W. Kennedy, Jr., *6th National Symposium on Reliability & Quality Control in Electronics, Proceedings* 1960
"Resistor Reliability-Capability Analysis," B. R. Schwartz, *6th National Symposium on Reliability and Quality Control in Electronics, Proceedings* 1960
"Statistical Pitfalls for the Reliability Engineer," G. H. Beckhart, *4th National Convention on Military Electronics, Conference Proceedings* 1960
"System Evaluation of Low Noise Radar Sensitivity," S. Charton and G. VerWys, *4th National Convention on Military Electronics, Conference Proceedings* 1960

Correction: In the paper entitled **"Nonlinear Theory of a Velocity-Modulated Electron Beam with Finite Diameter,"** by F. Paschke, which appeared on pages 53-74 of the March 1960 issue, Equation (17) on p. 61 contains two errors. The denominators of the last two terms, $(1+\xi)^2-(1/\sigma)^2$ and $(1-\xi)^2-(1/\sigma)^2$, should be replaced by $(1+\xi)^2-(\xi/\sigma)^2$ and $(1-\xi)^2-(\xi/\sigma)^2$, respectively. Also the curve for the third-harmonic current in Figure 5 has unfortunately been evaluated from the incorrect formula. The accompanying figure shows the correct current distribution.

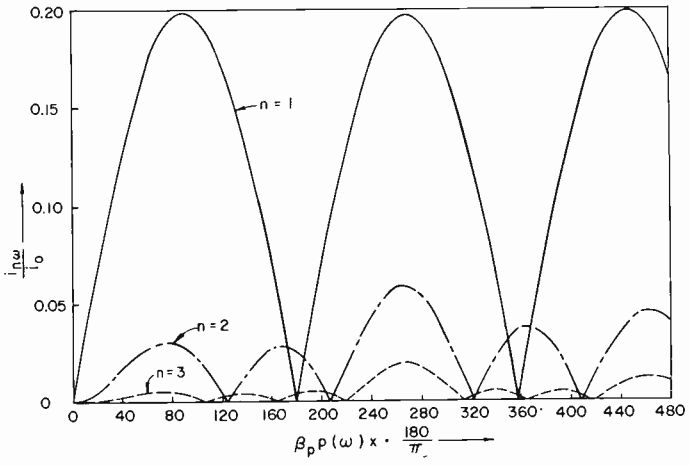


Fig. 5—Convection current of the n^{th} harmonic versus distance from the input cavity measured in plasma-phase angle. $\beta_e b \doteq 1.2$, and $b/a = 2/3$, which corresponds to $\xi \doteq 0.707$, and $\sigma \doteq 0.64$. The normalized input voltage A defined by Equation (25) is 0.1.

Correction: In the paper entitled **"High-Frequency Varactor Diodes,"** by C. W. Mueller and R. D. Gold, which appeared on pages 547-557 of the December 1960 issue, the last two column headings in Table I (p. 555) are transposed. The corrected table is given below.

Table I—Effect of Out-Diffusion on Electrical Characteristics for Diodes with the Same Breakdown Voltage

	Base wafer	V_{Br} (at 1 ma)	C^* ($\mu\mu$f)	r^* (ohms)	$f_{co}{}^*$ (kmc)	$S_c{}^*$	F^*
No Outdiffusion	6.8×10^{16} Sb	6.5	.9	10.2	18	.21	3.8
Outdiffused	2.7×10^{18} Sb	7.5	1.15	.95	150	.15	22.5

* r, C, f_{co} are evaluated at —1 volt bias; $S_c \approx (C_1 - C_2)/C_1$, $F = f_{co}\, S_c$.

AUTHORS

JOHANNES BERGHAMMER received the degree of Diplom-Ingenieur in 1954 and the Dr. Techn. in 1955 from the Technical University, Vienna. From 1954 to 1956 he was a research assistant at the Institute of High-Frequency Techniques in Vienna. Since May 1957, Dr. Berghammer has been a Member of the Technical Staff at RCA Laboratories in Princeton, N. J., and has been engaged in research on electron guns, microwave amplifiers, and electron beam noise.

MICHAEL COOPERMAN received the B.E.E. degree from the City College of New York in June 1955. From 1955 to January 1961, he was with the Advanced Development Section of the RCA Home Instruments Division, Camden, N. J. Mr. Cooperman's work has centered around the application of ferromagnetic and ferroelectric devices to television circuits and, more recently, the development of tunnel-diode circuitry and parametric amplifiers for operation at kilomegacycle frequencies. He is presently with the RCA Electronic Data Processing Division, Camden. Mr. Cooperman is a member of the Institute of Radio Engineers, Tau Beta Pi, and Eta Kappa Nu.

HAROLD B. GOLDBERG received the B.S. degree in Electrical Engineering from Lafayette College in 1956 and the M.S. degree in Electrical Engineering from Cornell University in 1958. From 1956 to 1958 he worked at Cornell University as a research assistant in radio astronomy. He joined the Astro-Electronics Division of RCA in 1958 and is, at present, working on the design of a wide-dynamic-range receiver. Mr. Goldberg is a Member of the Institute of Radio Engineers and Tau Beta Pi.

KARL G. HERNQVIST graduated in Electrical Engineering at the Royal Institute of Technology, Stockholm, Sweden, in 1945. He received the Licentiate of Technology degree in 1951 and the Doctor of Technology degree in 1959. From 1946 to 1952 he was employed by the Research Institute of National Defense in Stockholm working in the field of microwave electronics. From 1948 to 1949, Mr. Hernqvist was a trainee of the American-Scandinavian Foundation at RCA Laboratories, to which he returned in 1952; he is presently working on thermionic power conversion. Dr. Hernqvist is a Member of the Institute of Radio Engineers and Sigma Xi.

ERIC F. HOCKINGS received the B.Sc. degree in Chemistry from the University of London, England. From 1951 to 1953 he served in the British Army and then joined the Electric and Musical Industries, Hayes, England. After two years he entered Imperial College, London and received the Ph.D. degree in Chemistry in 1957. Since that time he has been with RCA Laboratories, Princeton, where he has been engaged in a study of ternary compound semiconductors with special reference to their thermoelectric properties. Dr. Hockings is an Associate of the Royal Institute of Chemistry and a Fellow of the Chemical Society.

FRED M. JOHNSON received the B.S. degree from the City College of New York in 1949. He received the M.A. in Physics in 1951 and the Ph.D. in physics in 1958 from Columbia University. He was a teaching assistant at Columbia University from 1951 to 1954 and held a research assistantship at the Columbia University Radiation Laboratory from 1954 to 1958. In 1958 he joined the technical staff of RCA Laboratories where he has been engaged in fundamental research on ion emitters and plasma synthesis. Since 1959 he has been engaged in studies on the cesium plasma, resonance ionization, and recombination phenomena. In addition, he had undertaken an experimental and theoretical investigation of the space-charge problem associated with the cesium plasma diode. Dr. Johnson is a member of Phi Beta Kappa, Sigma Xi, the American Physical Society and the American Astronomical Society.

NILS E. LINDENBLAD received an M.E. degree from Norrkoping Polytechnic Institute in 1915 and an E.E. degree from the Royal Institute of Technology, Stockholm, Sweden, in 1919. He worked in the Radio Engineering Department in the General Electric Company from February to September 1920. He then joined RCA, working on design, development, and research in the New York office until September 1926 when he was transferred to the Rocky Point laboratory. He came to RCA Laboratories, Princeton, in 1950 and retired on November 1, 1960. Since that date he has been a consultant to RCA Laboratories. He received the Modern Pioneer Award from the National Association of Manufacturers in 1940. Mr. Lindenblad is a Fellow of the Institute of Radio Engineers and a Member of Sigma Xi.

JOSEPH J. LOFERSKI received the B.S. degree from the University of Scranton in 1948, and the M.S. and Ph.D. degrees in Physics from the University of Pennsylvania in 1949 and 1953, respectively. He was a Research Associate at the University of Pennsylvania during 1952-53. He joined RCA Laboratories in 1952, where he engaged in research on semiconductor radiation converters, and radiation damage in semiconductors. Dr. Loferski is a member of Sigma Xi, American Physical Society, and the American Association for the Advancement of Science. Dr. Loferski joined the staff of Brown University, Providence, R. I., in January 1961.

JOHN G. MCCUBBIN received the B.S. degree in Mechanical Engineering from George Washington University in 1949, the M.S. degree in Mechanical Engineering from Drexel Institute of Technology in 1954, and the M.S. degree in Electrical Engineering from the same school in 1958. From 1949 to 1953 he served with the U. S. Navy. From 1953 to 1958 he was associated with Defense Electronic Products of RCA at Camden as an engineering group leader in the advanced development engineering section. He had project responsibility for development of prototype electronic equipments for data transmission, satellite signal data reduction, and new techniques for military television, magnetic recording, and heat transfer. In 1958, Mr. McCubbin joined the Astro-Electronics Division of RCA, where he worked as a project engineer concerned with satellite-vehicle thermal design integration, and design of mechanical equipment for television cameras. He is now a member of General Electric Co.

HERBERT I. MOSS received the B.S. degree in chemistry from the University of Louisville, Louisville, Ky., in 1953, and the Ph.D. degree in inorganic chemistry from Indiana University in 1960. While at Indiana University, he held a graduate teaching assistantship from 1953 to 1956, and a research assistantship from 1956 to 1959. He joined the staff of RCA Laboratories, Princeton, New Jersey, in 1959, and is currently engaged in research on photovoltaic materials. Dr. Moss is a member of the American Chemical Society, American Institute of Chemists, Electrochemical Society, and Sigma Xi.

FRITZ PASCHKE received the degree of Diplom-Ingenieur in 1953 and the Dr. techn. sc. in 1955 from the Technical University Vienna. Between 1953 and 1955 he was a research assistant at the Institute of High-Frequency Techniques in Vienna. Since April 1956, Dr. Paschke has been a member of the technical staff at RCA Laboratories, Princeton, N. J.

FRED D. ROSI received the B.E. degree in Metallurgy from Yale University in 1942, the M.E. degree in 1947, and Ph.D. degree in 1949 from the same institution. From 1943-46 he served in the United States Navy. From 1949 to 1954, Dr. Rosi was an engineering specialist in the theoretical metallurgy group at Sylvania Electric Products, Inc. where he was engaged in fundamental research on the problems of plastic flow and recrystallization in various metals and alloys. Since March 1954 he has been with RCA Laboratories where he is engaged in materials research for transistor and thermoelectric device applications and superconducting phenomena. He is a member of the American Institute of Metallurgical Engineers, the Yale Engineering Society and the Yale Chapter of Sigma Xi. He also holds the rank of Adjunct Professor at New York University where he has lectured in graduate courses on crystal mechanics and advanced physical metallurgy. Dr. Rosi was a member of the Student Guidance Committee of New York, has served as Chairman of the New York Physical Metallurgy Chapter, and has served on the Program, Mathewson Gold Medal, Metal Physics, and Semiconductor Committees of the AIME.

WILLIAM E. SPICER received the B.S. degree from William and Mary College in 1949, the S.B. degree from M.I.T. in 1951, and the M.A. and Ph.D. degrees from the University of Missouri in 1953 and 1955, respectively. He joined the RCA Laboratories at Princeton in 1955 where he has been engaged in studying electron emission processes. Dr. Spicer is a member of Phi Beta Kappa, Sigma Xi, and the American Physical Society.

IRVING STEIN received the B.M.E. degree in 1957 from the City College of New York. From 1951 to 1953, he worked for Sparkman and Stephens, Incorporated, in New York City, and from 1953 to 1955 for the Eastern Engineering Company of New York. As a project engineer for the American Machine and Foundry Company, he worked on the Talos missile from 1955 to 1956; and from 1956 to 1958, he worked at the Aerojet-General Corporation in Azusa, California, on the Polaris Missile. Mr. Stein joined the Astro-Electronics Division of RCA in September 1958, and worked on test fixture design, testing of solar cells, and assembly procedures and techniques for the TIROS I and II Satellites. He is currently engaged in the mechanical design of a solar simulator for a large environmental space chamber. Mr. Stein is a member of the American Rocket Society.

GEORGE WARFIELD received the B.S. degree in Physics from Franklin and Marshall College in 1940. From 1941-1944 he was a teaching assistant in Physics at Cornell University. From 1944-1946 he worked on the proximity fuse project at McQuay-Norris Mfg. Co. He returned to Cornell after the war and received the Ph.D. in Physics in 1949. Since that time he has been on the faculty of the Department of Electrical Engineering at Princeton University with the present rank of Associate Professor. In addition he has, since 1951, been serving as a member of the RCA Laboratories Technical Staff on a part-time basis. Dr. Warfield is a member of the American Physical Society, the American Society for Enginering Education, Sigma Xi, and Phi Beta Kappa.

PAUL WIENER received the B.S. degree in Electrical Engineering, Cum Laude, from the College of the City of New York in 1955. Upon graduation he joined the RCA Laboratories, from which he transferred to the Astro-Electronics Division. Mr. Wiener has participated in the design and development of the solar-cell storage battery power supply for the TIROS I and II Satellites and for the Echo balloon satellites. Recently he has been involved in a program for the investigation of the effects of Van Allen Belt charges particle radiations on various types of solar cells and protective coatings. Mr. Wiener is a member of Eta Kappa Nu, Tau Beta Pi, and the American Rocket Society.

SEYMOUR H. WINKLER received a B.S. degree, Cum Laude, in Mechanical Engineering from the College of the City of New York in 1947, and an M.S. degree in Mechanical Engineering from Purdue University in 1952. He has taken further graduate work towards his Ph.D. From 1947 to 1951, he was engaged as an instructor in Purdue University's School of Mechanical Engineering. From 1951 to 1955 he worked at the Kaman Aircraft Corporation and from 1955 to 1958 at Young Development Laboratories. Since 1958, he has been with Astro-Electronics Division where he developed packaging for the Project Score system. He was a Project Leader in charge of solar power systems development for the TIROS I and II, and ECHO satellite applications. He is now in charge of large solar radiation simulation development and design for aerospace chambers. Mr. Winkler is a member of the American Rocket Society and Pi Tau Sigma. He is also a member of the National Committee for standardization of solar cell measurements and is a registered Professional Engineer in the State of New York.

IRVING WOLFF was graduated from Dartmouth College in 1916 and received the Ph.D. degree from Cornell University in 1923. He served as an instructor of physics at Iowa State College and at Cornell, where he became a Hechscher Research Fellow in 1924. He joined RCA in 1924, concentrating upon acoustical development. In early 1930's, the emphasis of his work shifted to research in microwaves. He made significant contributions to the early development of radar and was instrumental in the development of the first installation of radar equipment on Navy combat vessels in 1938. Shortly after World War II, he was appointed Director of the Radio Tube Research Laboratory of RCA Laboratories at Princeton, N. J. In 1951, he was appointed Director of Research, RCA Laboratories, and in 1954 was elected Vice-President, Research. He retired in 1959, and is presently serving as Chairman, RCA Education Committee, in the role of a Consultant.

Dr. Wolff received the Distinguished Public Service Award from the U.S. Navy in 1948. He is a Fellow of the American Physical Society, the Acoustical Society of America, the American Association for the Advancement of Science, and the Institute of Radio Engineers. He is also a member of Sigma Xi and Phi Kappa Phi. He holds approximately 80 U.S. patents, relating largely to radar and acoustics.

JOSEPH J. WYSOCKI graduated from M.I.T. in 1954 with B.S. and M.S. degrees in Electrical Engineering. While at M.I.T., he was a member of the Cooperative Training Program assigned to the Naval Ordnance Laboratory, Silver Spring, Md. Upon graduation, he joined the RCA Laboratories where he has worked on semiconductor devices of various types. He is a member of the Institute of Radio Engineers and Sigma Xi.