Reliability of Components for Communication Satellites

By I. M. ROSS

(Manuscript received August 15, 1961)

This article considers the reliability of components such as transistors, diodes, and solar cells in relation to the design of a communication satellite with adequate reliability. Consideration is given to methods for determining the reliability of high-quality components and of techniques for selecting the most stable components for this application. It is concluded that, at least for a simple communication satellite, components can now be obtained that will lead to a satisfactory life.

I. INTRODUCTION

All the necessary components and circuit techniques are available to fabricate a simple communication system using low-orbit satellites.[1] Such a system would use many satellites at an altitude of a few thousand miles and be capable of global communications with a few megacycles baseband. The ground receiver portion of the system could achieve adequate signal-to-noise for very low received power by use of high-gain receiving antennas, low-noise maser receivers and FM modulation with feedback. The satisfactory performance of this type of receiver was demonstrated in the Echo I experiment.[2] In conjunction with such sensitive ground receiver equipment, it is possible to use a satellite repeater putting out only a few watts of power from an isotropic antenna, and hence avoiding the additional complexity of attitude stabilization. The components needed for such a satellite, including the traveling-wave tubes, transistors, diodes and solar cells, are all either available or achievable within the capability of existing technology. Thus a communication satellite system is feasible in principle. Whether or not it is economical and therefore practical, depends upon the life expectancy of the system, and specifically on the life of the satellite itself. It will be assumed here that a satellite life of at least five years is a reasonable target in the design of a practical communication system. By the very nature of the system, repair of the satellite is presently impossible (and if ever possible, would be exorbitantly expensive), and because of the cost penalty of additional weight in orbit, extensive redundancy is most undesirable. Thus, the practicality of the system depends critically on the reliability of the components that make up the satellite itself. This paper is devoted to a discussion of the reliability of components in relation to the design of a satellite with adequate reliability. Although the discussion is directed specifically to low-orbit (several thousand miles altitude) satellites, many of the ideas could apply equally well to higher orbits.

In Section II, below, consideration is given to the order of component reliability needed in a simple communication satellite. Section III deals with the reliability of components in general with emphasis on means for attaining highly reliable components and for determining quantitatively their degree of reliability. Section IV discusses the level of reliability that can be achieved in three critical classes of components, namely transistors and diodes, traveling-wave tubes and solar cells. Finally, it is concluded that, with careful manufacture and selection, components can be obtained for a practical communication satellite system.

II. COMPONENT RELIABILITY REQUIRED FOR COMMUNICATION SATELLITES

For the consideration of reliability it is convenient to divide the life of a satellite into three periods, namely pre-launch, launch, and orbit. It is usual practice to assume that any failure that occurs a reasonable time prior to lift-off can be corrected by replacement and that, at the worst, this could result in some delay in the launch time. For such an assumption to be valid, it is necessary that components or batches of components be accessible and removable so that failed portions of the satellite can be replaced. The design for such flexibility does necessitate some weight increase. Although the launch period is short, it is accompanied by large mechanical stresses liable to cause failure. As will be discussed later, in the section on traveling-wave tubes, experience with many launches has shown that with well designed components and equipment, failure during launch of the electronic equipment in a satellite is not a significant factor in the over-all reliability of the satellite. It is the third period, life in orbit, which dominates the reliability design of a satellite. In this section we consider the relationship between the reliability of components and the anticipated life in orbit.

In calculating the probability of survival of a system containing a large number of components, it is frequently assumed that the failure distribution of any type of component is exponential. On such an assumption, the performance of a given type of component can be characterized by a mean time to failure or a failure rate. One of the more convenient ways to represent the failure rate is in terms of a number of failures for a given number of component operating hours. A method which is in increasing use defines failure rate as the number of failures per 10^9 component hours (1 failure per 10^9 hours corresponds to a failure rate of 0.0001 per cent per 1000 hours). By way of calibration, a good resistor or capacitor has a failure rate in the range 5 to 10 per 10^9 component hours, while an entertainment receiver tube will have a rate in the neighborhood of 100,000 per 10^9 hours.

If we assume that a given system contains n_1 components of a given type, and that the failure rate for that type is f_1 per 10^9 hours, we expect statistically that there will be $n_1 f_1$ failures per 10^9 hours. Hence in a time t hours we expect $tn_1f_1/10^9$ failures. Assuming that failure probability is random and that the failure of any one of these components leads to failure of the system — that is, assuming no redundancy — the probability P_1 that the system will not fail in t hours due to failure of one of the n_1 components, is given by:

$$P_1 = \exp\left[-\frac{tn_1f_1}{10^9}\right]. \tag{1}$$

Similarly, if we have a system composed of n_1, n_2, etc., components of types having failure rates f_1, f_2, etc., and we again assume no redundancy, the probability P_m of survival for time t is given by:

$$P_m = \exp\left[-\frac{t}{10^9}\sum_1^m (n_m f_m)\right]. \tag{2}$$

This simple equation can be used to estimate probability of system's survival, provided that the following conditions are met:

a) The failure mode of the components is assumed random with recognized exceptions being treated separately.

b) The system contains no redundancy.

Assumption b) is unrealistic since a certain degree of redundancy will be featured in any good design. However, because of weight limitations in a satellite, redundancy cannot be used to correct for poor reliability performance of a majority of the devices. Hence the equation is useful in determining desired objectives.

Table I shows the results of reliability calculations for a hypothetical communication satellite. At the left of the table are listed the types and numbers of critical components used. These types and numbers, which are representative of a very simple repeater of a few megacycles baseband, do not include any allowance for redundancy, nor do they include allowance for the telemetry invariably associated with such a system.

Table I — Reliability Calculation for Simple Communication Satellite

Type of Component	Number (n)	Case I		Case II		Case III	
		Failure Rate (f) (Failures/10⁹ hrs)	P Product (nf)	Failure Rate (f) (Failures/10⁹ hrs)	Product (nf)	Failure Rate (f) (Failures/10⁹ hrs)	Product (nf)
Transistor	140	20	2800	10	1400	5	700
Diodes	161	15	2415	10	1600	5	805
Resistor	400	5	2000	5	2000	2	800
Capacitor	250	10	2500	5	1250	2	500
Inductor and Transformer	40	20	800	15	600	5	200
Relays	6	50	300	25	150	6	120
Ni-Cd Cells	20	50	1000	25	500	15	300
Totals	1017		11,815		7510		3425
Average Failure Rate		11.6		7.4		3.4	
Probability of success — 1 year		0.901		0.94		0.97	
Probability of success — 5 years		0.60		0.72		0.86	

Excluded from the list is the traveling-wave tube. The unique life properties of the single traveling-wave tube in a nonredundant satellite warrant special treatment. Also excluded are the solar cells which, as will be discussed later, will probably fail through wear-out resulting from radiation damage and thus cannot be treated with the statistics of equation (2).

The table shows three cases, each assuming somewhat different failure rates for the components. For each case the table gives the failure rate f assumed for the component, the product of the failure rate times the number n of each component, the total sum $\sum_1^m (n_m f_m)$ and the average failure rate. Also shown in the table is the probability of success of the satellite, i.e., no failure of any component as calculated using (2), for one-year operation and for five-year operation. It is seen that case 1 represents satisfactory performance for one year and poor performance for five, while case 3 represents satisfactory performance for five years. Case 2 is an intermediate case. Using some judgment as to the relative values of failure rates for various components, the failure rates were chosen in the three cases to give the above results. Thus the table shows what level of component reliability is needed to meet a given systems performance.

It must be emphasized that considerable caution is needed in the interpretation of the results shown in Table I. Implicit in the calculations are many assumptions, the validity of which could be questioned. The results should therefore be used as a guide to the order of magnitudes of reliability required and should not be considered to be precise predictions of systems performance. There are, nevertheless, a number of general conclusions to be drawn from the table. The first is that although this is a fairly simple system — 1000 components — average failure rates in the neighborhood of 10 per 10^9 component hours are required to give anything approaching economical life. As seen from (2), the life expectancy for a given probability of success varies inversely with the average failure rate. Thus, an average failure rate in the neighborhood of 100 would be intolerable, while an average failure rate in the order of 1 would permit increased design life and/or complexity. A second conclusion is that all the components that are numerous, i.e., all the transistors, diodes, resistors and capacitors, require an equally high order of reliability. This conclusion results directly from forbidding redundancy for the high-runner components. A final conclusion is that, at least for the more reliable designs, the reliability of connections between components cannot be ignored. For the 1000 components of Table I there would be several thousand connections and hence, in order that there be an insignificant probability of a connection failure, they must have failure rates substantially less than 1 per 10^9 hours. Although there is little quantitative information regarding reliability of connections, it is believed that those liable to fail are eliminated during the vibration, temperature cycle, and vacuum tests normally carried out as part of the acceptance test of a complete satellite.

III. RELIABILITY OF COMPONENTS

Fig. 1 shows a possible failure pattern for a batch of components. Such a curve could be obtained by taking a large number of new components of a specific type, operating them under typical conditions, and plotting the failure rate versus time for the batch. The distribution has two regions of relatively high failure rate, one early in life and attributable to "manufacturing freaks," one later in life attributable to "wear-out," separated by a region of low failure rate labeled "random failure." These three regions will be discussed separately.

3.1 *Wear-Out Failure*

In some manufactured products there is a mechanism or a collection of mechanisms which systematically reduces the useful performance of the product until a point is reached at which it has no further utility and is "worn out." Typical examples of wear-out mechanisms are friction of

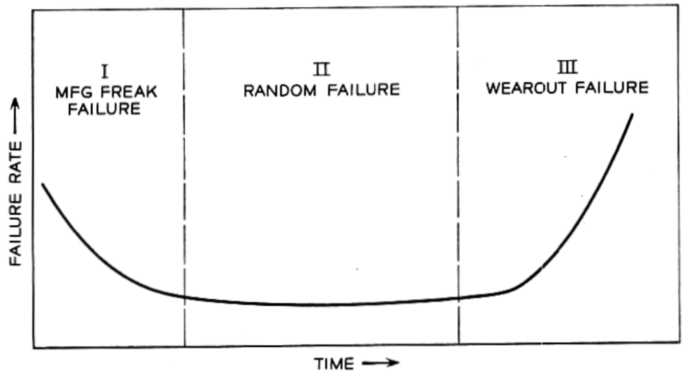

Fig. 1 — Possible failure distribution for a large number of new components.

bearings, corrosion of relay contacts, and deactivation of electron tube cathodes. If, for a given batch of components, conditions were identical during fabrication and use, then all components would fail in response to wear-out simultaneously. However, because conditions are not identical, simultaneous failure does not occur, and the failure distribution is characterized by a peak of finite width. Region III in Fig. 1 shows the onset of wear-out. Once wear-out failure commences, the failure rate of the batch of components increases vary rapidly, and effectively all components of that type must be replaced. In systems such as satellites, where replacement is not possible, the time at which wear-out becomes significant should be greater than the designed life of the satellite. Lengthening of the time to wear-out can only be achieved by understanding the wear-out mechanisms and by designing the components either to minimize or eliminate these effects.

3.2 *Manufacturing Freak Failure*

There is a certain percentage, preferably small, of any product that fails unusually early in life because of some defect in manufacture. These are, in a sense, objects that were not made according to the design. For example, such early failures can occur both in tubes and semiconductor devices as a result of defective seals or of the presence of particles inside the encapsulations. The prevalence of manufacturing freaks can be reduced drastically by quality control in manufacture. Remaining freaks can usually be detected and rejected by rigorous pre-aging tests, such as leak tests, vibration and shock tests. In addition, the product can be aged for a period longer than that corresponding to Region I, so that the remaining freaks will fail during this "pre-age period."

3.3 *Random Failure*

Even in a well designed and well manufactured product there may be a substantial period, after that exhibiting high failure rate due to manufacturing freaks and before wear-out occurs, of a continuing failure rate. These failures include components which, through presumably detectable causes, fail in response to manufacturing weaknesses much later than the majority of freaks, and others which fail through similar causes to, but earlier than, the wear-out failures. The failures that occur during this period may generally be attributable to a large number of different causes, each of which occurs so rarely that it would be exorbitantly expensive to identify all of them. This period is in essence the useful life of the product. If the frequency of such failures is sufficiently low, as indicated, these may be essentially below the noise level of identification of mechanisms, and a random failure mechanism, and hence a constant failure rate, may be assumed. Although there may be considerable doubt as to the validity of this assumption for some components, it has proved useful in the estimation of over-all systems reliability.

Fig. 2 summarizes the steps that can be taken to cope with the various modes of failure shown in Fig. 1. The region of high failure rate corresponding to wear-out can be moved further out in time by design based upon knowledge of the failure mechanisms. The number of devices subject to early failure through manufacturing freaks can be reduced by quality control, rejected after testing and annihilated by pre-aging. Hence, provided sufficient care is taken, it is possible to obtain a product which, during the intended life of the system, will exhibit substan-

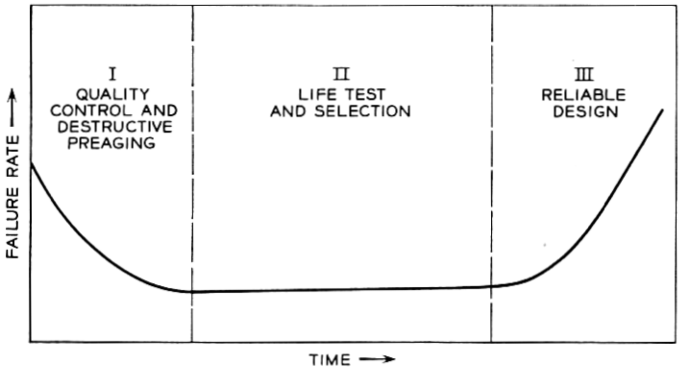

Fig. 2 — Summary of steps that can be taken to reduce failures of various types.

tially only a low failure rate corresponding to Region II. This failure rate can be determined from the results of extensive life tests involving, for the most reliable components, thousands of devices for thousands of hours.

The low failure rate of Region II is that characteristic of the product. Where reliability is of supreme importance, it is desirable to select from the product as a whole those components that exhibit the greatest degree of stability. This can be achieved by putting on life test a number of components many times that needed in the system, and after a given length of time selecting from the batch only those components which have shown the minimum change in their parameters. The duration of the life test prior to selection will depend upon a number of factors, including the life required in the system and the system's schedule which, itself, frequently limits the life-test period. In the selection of submarine cable tubes, a period of seven months is used. Although it is expected that the selected product will have a lower failure rate than the batch from which it was selected, it is difficult, if not impossible, to estimate the degree of this improvement. The consensus, however, is that a factor of 10–100 improvement could be achieved.

In order to achieve the reliability potential of a carefully designed and manufactured component, it is essential that the same care go into the design and assembly of circuits and subsystems. Circuits must be designed with adequate margins, and power dissipations must be determined so that temperatures do not reach values at which reliability of the components is no longer adequate. Assembly procedures should be arranged to avoid excessive mechanical or thermal shock. The conservative use of a component is thus an important part of the achievement of reliability.

IV. RELIABILITY OF SPECIFIC COMPONENTS

The components that appear in large number in a typical satellite and require reliabilities corresponding to 10 failures per 10^9 hours, include transistors, diodes, resistors and capacitors. Passive components, resistors and capacitors, have for many years been available with reliability in this range. However, until recently such low failure rates had not been achieved in the active components. For this reason the discussion in Section 4.1 below is restricted to transistors and diodes.

The traveling-wave tube used to generate the output power in most communication satellite designs does not require the high degree of statistical reliability called for in transistors and diodes. However, it is required to operate without failure for a period much longer than the life of ordinary tubes and also to withstand severe mechanical stress during launch. The expected performance of satellite tubes is discussed in Section 4.2 below.

The solar cells, although as numerous as the transistors and diodes, are expected to fail due to "wear-out" from radiation damage. The expected life of these components is discussed in Section 4.3.

4.1 *Transistors and Diodes*

As indicated previously, the reliability of a component in the final analysis is limited both by the design of the component and the care with which it is manufactured. The attention to design and manufacture is particularly important in the case of transistors and diodes which are both delicate and particularly sensitive to contamination, yet are required to exhibit failure rates comparable to those of the more rugged, passive components. Mechanical techniques have been developed whereby small semiconductor wafers can be bonded to headers and even smaller leads connected between the wafers and the headers, such that the resulting structure will easily withstand the mechanical shock and vibration experienced during the launch of a satellite and the temperature cycling that may be experienced while in orbit. Final cleaning and sealing techniques have also been developed which insure a degree of initial cleanliness and subsequent protection from outside contamination, such that adequate reliability for satellite applications can be achieved.

Table II outlines the complete reliability testing program proposed by Bell Laboratories for providing transistors and diodes for satellite applications. The first step is to insure that the design itself has adequate reliability potential. In order for a design to qualify for satellite use, it must pass mechanical tests which represent conditions more rugged than will be experienced during launch. The devices are further subjected to electron and proton bombardment simulating many years exposure to Van Allen radiation. Finally, devices are subjected to reliability evaluation to determine the reliability potential of the design.

The second step, that of screening and pre-aging, is designed to eliminate those few remaining freaks that were not eliminated by quality control. These tests include mechanical shock and vibration tests to eliminate weak components. In the reliability portion of these tests, a sample from the particular manufacturing lot is tested at increasing temperatures until all devices in the sample have failed. The median temperature for failure and the distribution of failures with temperature, when compared with similar figures for previous manufactured lots,

TABLE II—RELIABILITY PROGRAM FOR SATELLITE TRANSISTORS AND DIODES

1. Design Qualification Tests	
Mechanical	
Temperature cycling	−65C to +85C (−120C to +40C for blocking diodes)
Temperature-humidity cycling	
Shock	2,000 g
Centrifuge	5,000–10,000 g
Vibration	60g, 100–2,000 cycles
Radiation	
Reliability	
Accelerated aging	
Life testing	
Field experience	
2. Screening and Pre-aging	
Mechanical	
Centrifuge	2,000 g
Temperature-humidity cycle	
Tap or shock	
Reliability	
Accelerated temperature sample	
High-temperature aging	
3. Life Test and Selection	
Reliability	
System simulation and selection	

indicate whether or not there are major differences from previous lots. In addition, all the devices that may be used in satellites are subjected to a short period of high temperature aging. Since, as discussed later, aging is accelerated by raising temperature, this pre-age eliminates many devices that otherwise would have exhibited unusually early failure.

The third step consists of choosing from the components that have passed step two, a number many times greater than the number that are finally to be used, and putting them on life test for six months under power and temperature conditions simulating those anticipated in operation. The duration of this test, which ideally should be a substantial fraction of the design life of the system, is frequently limited by economic factors or by the time available prior to the system's operation. During the life-test period, the characteristics of the components are measured at frequent intervals. The components needed for the system are chosen on the basis of their performance during the life-test period. If proper choices have been made, the components used should be ones which have shown no change in characteristics.

Steps 2 and 3 in this program are intended to insure that the components selected are truly representative of the design and do not include any freaks. Assuming these steps to be successful, the most significant portion of the program in determining system performance is the evaluation in step 1 of the reliability potential of the product. Since the reliability required is in the neighborhood of a few failures per 10^9 hours, this reliability evaluation can involve tens of thousands of components for tens of thousands of hours. It is with the object of reducing the numbers and times involved that considerable emphasis has been put on the development of accelerated aging techniques.[3,4,5] The results of a typical accelerated aging experiment are shown in Fig. 3. Plotted in the figure is the median life of a germanium transistor as a function of the temperature at which the transistor is operated. The data shown as solid points were obtained for some germanium transistors manufactured by the Western Electric Company in 1958. The temperatures at which the transistors were tested range from 100°C to as high as as 350°C, while the range in time to median failure is from about 20 minutes to just over 1 year, nearly 5 decades. The fact that the points fit a straight line on a $1/T$ versus log time plot suggests that raising the temperature is accelerating some failure mode which can be characterized by an activation energy. It has been found that within experimental error, the apparent activation energy is the same for all germanium transistors and, in addition, that there is a single but slightly different activation energy for all silicon transistors and diodes. The

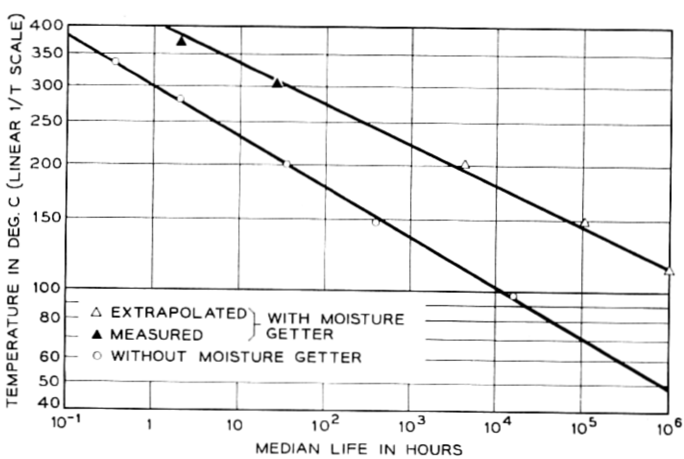

Fig. 3 — Results of a typical accelerated aging experiment on germanium transistors.

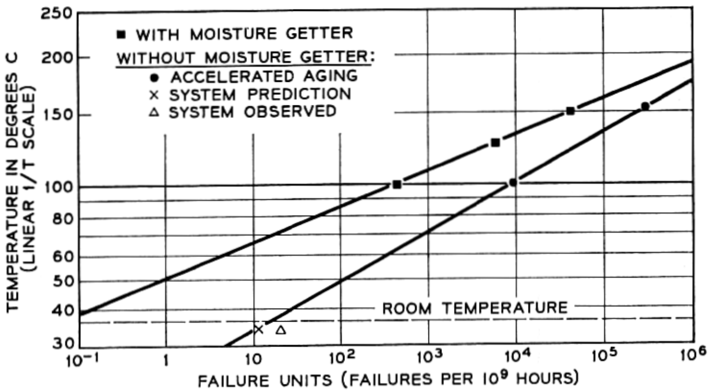

Fig. 4 — Failure rate vs temperature for germanium transistors.

triangles in Fig. 3 are for transistors manufactured by the Western Electric Company more recently. It is apparent that substantial improvements have been made at least in the high-temperature performance of the product. The data in Fig. 3 are for the median life. In performing the accelerated aging experiments, one also obtains the distribution of failures in time for a given temperature or, alternatively, in temperature for a given time. It is found that these distributions have the same shape, i.e., log normal in time* and normal in temperature, for all transistors and diodes. The widths of the distributions do not change with temperature for a given device type, that is, for fixed design and manufacturing procedure. This uniformity of failure distribution gives further confidence that raising temperature is accelerating a failure mode characteristic of the product.

Knowing the variation of median life with temperature and the distribution of failures in time for a fixed temperature, it is possible to derive a more useful plot for the systems designer, that of failure rate against temperature as shown in Fig. 4. The points are for the older transistors from the previous figure. A straight line is observed in the plot of $1/T$ against log failure rate. Extrapolating the line to room temperature, one would predict a failure rate of 10 per 10^9 hours for these transistors. The prediction of a failure rate of 10 per 10^9 hours from the acceleration curve of Fig. 4 is, however, liable to be optimistic because there is no guarantee that the curve does not dip below the straight line for times greater than the longest at which a measurement was made. There is no guarantee that in raising the temperature we are

* This is an example of a component that in the region of low failure rate does not exhibit the exponential failure distribution usually assumed.

accelerating all the failure mechanisms or even a guarantee that we are accelerating the most important failure mechanism at operating temperatures. For example, although one might expect that raising the temperature would increase the rate of reaction between the germanium surface and any water vapor inside the transistor can, one has no reason to suspect that elevated temperature would affect the occurrence of a short-circuit caused by a metal chip falling between emitter and base contact.

The accelerated aging curve, when extrapolated to room temperature, indicates the potential reliability of the design, and in the final analysis one must depend upon laboratory tests or field experience under operating conditions. The triangle on Fig. 4 shows the failure rate observed in the field trial of a new system using about 40,000 of these same transistors for about 10,000 hours. It is encouraging that the failure rate is only a factor of about 2 higher than that predicted from accelerated aging, and particularly so since the system failure rate includes failures due to mishandling and is for devices which were subjected to no special selection. It is therefore reasonable to estimate that the failure rate for these older germanium transistors, when properly handled and selected in a manner proposed for satellite use, would lie somewhere in the neighborhood of 10 to 20 per 10^9 hours.

The line through the squares in Fig. 4 is the accelerated aging curve for the more recent Western Electric product. Note again that there is a substantial improvement. The accelerated aging curve for recent silicon transistors and silicon diodes does not differ significantly from that for germanium transistors. With such an improvement in the reliability potential of the product, and with careful pre-aging and selection, one is confident that failure rates substantially lower than 10 per 10^9 hours are now achievable and that they may well be lower than 1 per 10^9 hours. However, complete confirmation of this prediction will have to await results of field trials.

The acceleration curves serve to emphasize the importance of conservative circuit design in the achievement of high reliability. It is seen from the slope of the curves that failure rate increases very rapidly with temperature. It is therefore important that power dissipation in the device be maintained sufficiently low that temperature rise above ambient does not impair reliability. It is equally important that the ambient temperature be maintained at a suitably low value.

4.2 *Traveling-Wave Tubes*

Fig. 5 is a photograph of the traveling-wave amplifier under development at Bell Telephone Laboratories for use in experimental communi-

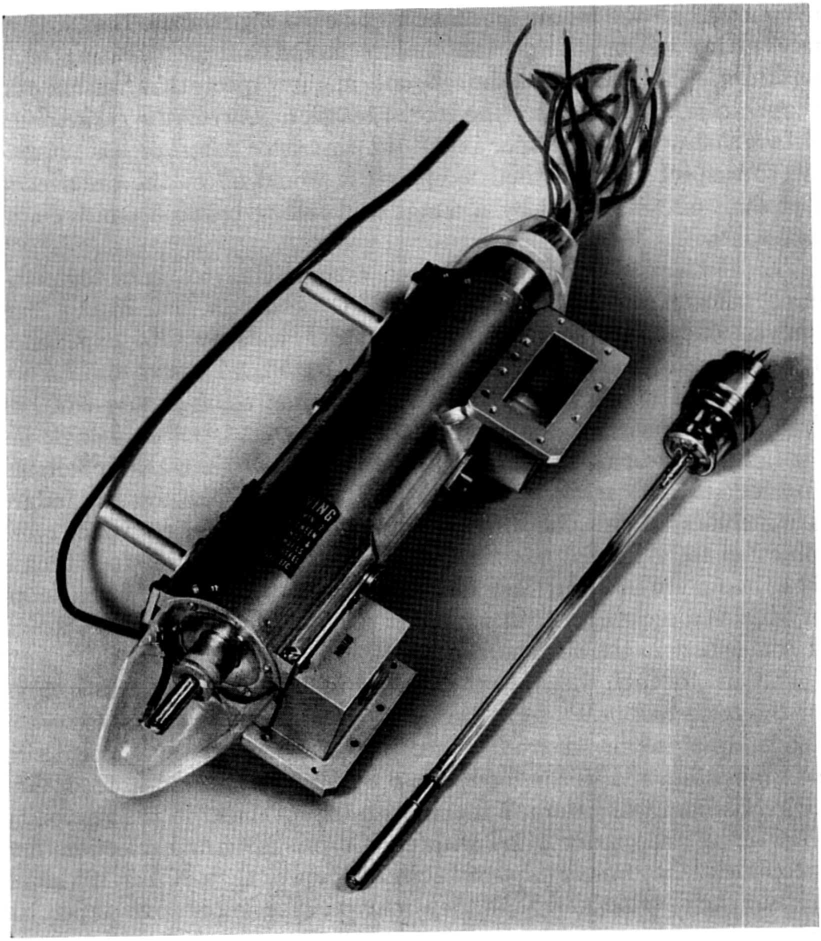

Fig. 5 — Traveling-wave amplifier under development for satellite use.

cation satellites. Table III lists the more important characteristics of this tube. Before discussing the performance and reliability of the M4041 satellite traveling-wave tube, a few words are in order on the reasons for selecting traveling-wave tubes to provide the output power in the satellite. It would appear that if a solid-state device could produce several watts at a few thousand megacycles, it would be, because of its small weight and potential reliability, an obvious choice over the traveling-wave tube. To date, however, schemes for generating power at several thousand megacycles using solid-state devices — harmonic

TABLE III—SATELLITE TUBE CHARACTERISTICS M4041 (7/7/61)

Operating point	0 dbm input saturated output
Output power (minimum)	3.5 w
Gain (at saturation)	35.5 db
Gain (low level)	41 db
Anode voltage	1770 volts
Helix voltage	1540 volts
Collector voltage	740 volts
Cathode current	17.0 ma
Cathode current density	85 ma/cm²
Collector power (including helix and anode)	12.5 w
Heater power	1.5 w
Weight	7.1 lbs.

generators, for example — operate at efficiencies very much lower than that of a traveling-wave tube, even when heater power is included. The weight of the additional solar cells needed to provide power for the solid-state device would more than offset the decrease in weight from that of a traveling-wave tube. The weight penalty for extra power is particularly severe for satellites subject to Van Allen radiation, where account must be taken not only of the weight of the solar cells and their mounting but also of the necessary protective covers. The higher gain of the traveling-wave tube gives it a distinct advantage over other tubes such as triodes, which would require at least two stages and, through consequent loss of efficiency, lead again to greater over-all weight. The high efficiency of the traveling-wave tube results from the distinct separation between the microwave interaction region and the beam formation and collection regions. After the microwave interaction takes place, the beam is allowed to enter a region of retarding field, where the beam is slowed before collection. This is usually done by depressing the collector voltage below that of the helix, as shown in Fig. 6. Since very little current is intercepted on the helix and the anode, the input power is very nearly proportional to the collector voltage. By depressing the collector voltage, efficiencies as high as 39 per cent have been achieved and 36 per cent is typical. When the power required by the cathode heater is included, this value falls to typical value of 31 per cent. A second effect of collector depression is that ions generated between the anode and the collector will flow to the collector and not to the cathode. This results in a substantial decrease in the possible ion current bombarding and consequently damaging the cathode.

The traveling-wave amplifier for a satellite must be a new design in order optimally to meet the specific needs of the system. With any

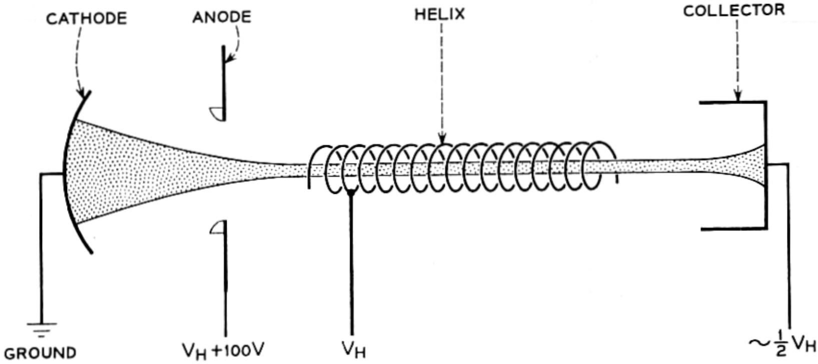

Fig. 6 — Traveling-wave tube circuit with depressed collector.

reasonable time scale, it is not possible to carry out a long-term evaluation of tube life, nor is it possible to do shorter experiments on very large numbers of models as is done with semiconductor devices. It is therefore necessary from the viewpoint of reliability to employ a design closely derived from experience gained with previous tubes and to utilize a "pedigree" approach in the assembly process. These earlier tubes include the pentodes used in telephone submarine cables,[6] the traveling-wave tubes used for microwave transmission at 6 kmc[7] and the rocket-borne traveling-wave tube used in a Bell Telephone Laboratories missile guidance system.[8] The salient features of these tubes are discussed in the next few paragraphs.

The submarine cable tube, the 175HQ, was the first tube designed to meet long-life reliability requirements somewhat similar to those encountered in satellite work. The failure pattern for this tube was found to agree with that shown in Fig. 1. The dominant wear-out mechanism in this case was determined to be the deactivation of the cathode, an effect which increases rapidly with increasing cathode temperature. Design information was developed which permitted the choice of a cathode temperature low enough to insure the desired life of the tube. The techniques of quality control to eliminate manufacturing freaks, and of life test and selection to insure the minimum random failure rate, were used extensively on this tube. As a result, the tubes that have been manufactured and put into operation in submarine cables easily meet the systems requirements. For example, Fig. 7 shows the accumulated tube life of the tubes in operation to date in submarine cables. There are now over 1600 tubes in such operation, some for as long as five years, with an accumulated life of 49 million tube-hours and no failures. It is on the basis of this evidence that it is believed possible to make long-life tubes and, in particular, to eliminate failure due to cathode deactivation.

The second tube of interest is a 6 kmc traveling-wave tube used as a ground-based microwave repeater, the M1789, now the WECo 444A. This traveling-wave tube was the first designed by Bell Telephone Laboratories specifically for long life, and it used many of the design principles and many of the selection techniques developed for submarine cable tubes. This tube also was designed to operate with a depressed collector. A little over four years ago, twelve of these tubes were placed on life test at their normal operating power of 5 watts. Table IV shows the accumulated hours on each of these tubes as of May, 1961, at which time there had been no tube failures. On the basis of this experience and the fact that the satellite traveling-wave tube has been designed to have a substantially lower cathode loading and cathode temperature than the 6 kmc tube, the satellite tube has an expectation of a life considerably in excess of four years.

The third tube is a traveling-wave tube designed to operate in the Bell Telephone Laboratories Command Guidance System, the M1958, now the 7116. In this system, the rocket to be guided contains a receiver, decoder and transmitter. There is a component count approximating 1000, including one traveling-wave tube. This system has been

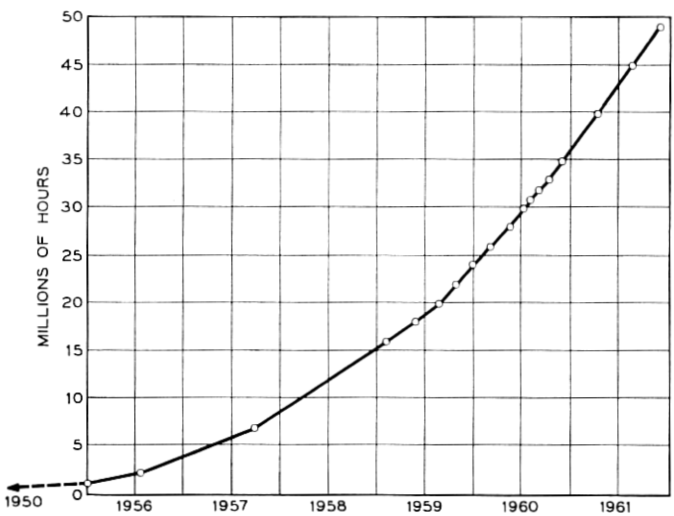

Fig. 7 — Operational life of electron tubes in undersea cable system repeaters.

TABLE IV—M1789 TRAVELING-WAVE TUBE LIFE TEST

Tube Number	Accumulated Hours 5/1/61
BC-856	39502
BC-1342	39630
BC-1363	39319
BD-14	39256
BD-660	39401
BH-69	39256
BH-208	33994
BH-413	37813
BH-464	36840
BH-559	35394
BS-41	36615
BS-102	34352

used in the guidance for about one-third of the U.S. satellites now in orbit. It was used, for example, with Echo I and with the three Tiros satellites. There have been to date over fifty successive firings using this guidance package with no failure. Since the guidance system needs only to operate for a few minutes, it gives us little information on long term reliability. However, since it not only must survive launch but must also operate during launch, this performance is a very potent demonstration that traveling-wave tubes can be made rugged enough to withstand the strains of launch. It further demonstrates that an electronic system containing roughly the number and kind of components needed in an active satellite can also survive launch.

To summarize, then, it is known from experience with the submarine cable tube and with the microwave relay tube that traveling-wave tubes can be designed with a life expectancy considerably in excess of four years. The performance of the guidance tube demonstrates that techniques are available for making a traveling-wave tube sufficiently rugged to withstand launch.

4.3 *Solar Cells*

Communication satellite designs for the immediate future rely on silicon solar cells as the prime source of power. These cells will be subject to radiation in the Van Allen belt,[9] which consists of electrons with substantial densities at energies up to 1 mev and protons at energies as high as 100 mev. Fig. 8 is a map of the Van Allen belt on a plane containing the earth's magnetic axis. There is a peak in the electron intensity at an altitude of about 2000 miles, and a second peak at about 10,000 miles with a substantial density of electrons at intermediate altitudes.

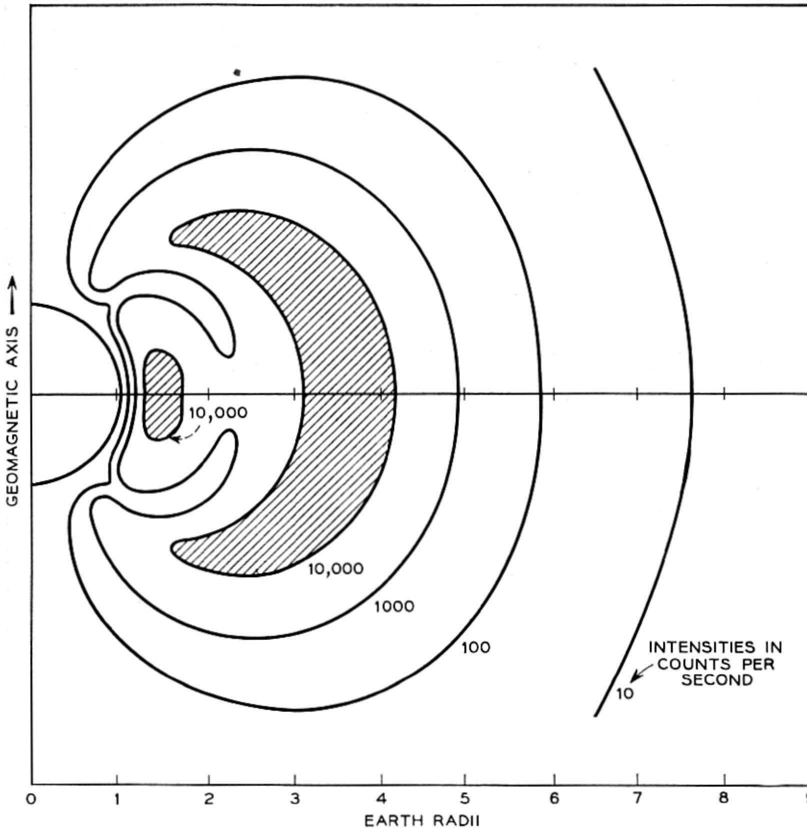

Fig. 8 — Map of Van Allen radiation belt in plane through earth's axis.

The protons, which are much less numerous, have a distribution which also peaks at around 2000 miles and falls off in some undetermined manner to negligible values beyond 10,000 miles. Bombardment of solar cells with particles of such energy results in a continual decrease of power output with time, at such a rate that this degradation could result in the failure of the power supply within the desired life of the satellite. Here then is an example of probable failure due to wear-out, in which case it is particularly important both to understand the mechanism of wear-out and to design the devices to minimize the effect. In this section, we discuss the effects of Van Allen belt radiation on solar cells, the means of designing cells to minimize the effects, and the predicted performance of such specially designed cells.

As shown in Fig. 9, a solar cell typically consists of a slice of n-type silicon with a thin p-type layer on one surface and contacts made to both surfaces. When light falls on the p-type surface, the photons penetrate the silicon to depths dependent upon their wavelengths and are absorbed with the creation of free carriers, hole-electron pairs, in the silicon. The free carriers created in response to the longer wavelength light are created deeper in the material. Some of the carriers move to the junction, and in crossing the junction create a current flow in the external circuit. Thus an illuminated solar cell is a source of electric power and has a voltage-current characteristic typically as shown in the figure.

In discussing the optimum design of solar cells, it is convenient to divide the generated carriers into two classes, namely those that are generated in the body of the material beneath the pn junction, and those that are generated in the surface layer above the pn junction. Those generated beneath the junction will reach it only if they are generated within a distance called the diffusion length, that is, the distance that generated carriers may move in the material before being annihilated by recombination. The diffusion length is a property of a

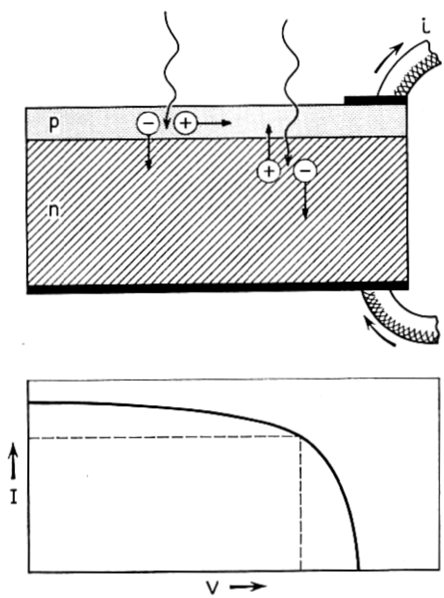

Fig. 9 — Solar cell construction and typical voltage-current characteristic.

particular material and depends critically upon its perfection and purity. For a solar cell to have the maximum efficiency, this diffusion length should be as long as possible in order that effectively all carriers generated beneath the junction may reach the junction and contribute to the output current. A somewhat different situation exists for the carriers generated in the surface layer. This layer is usually quite thin compared to a diffusion length. However, the surface of the semiconductor acts as a sink for carriers and thus competes with the junction for carrier collection. The net result is that the efficiency for collection of carriers generated above the junction is less than that for carriers generated below the junction. It is therefore desirable to minimize the thickness of the surface layer.

The perfect solar cell therefore would have a zero thickness of surface layer and an infinite diffusion length. A zero thickness surface layer, however, would lead to infinite series resistance. Obviously a compromise is necessary. Fig. 10 shows the distribution of carriers generated in silicon in response to sunlight. The plot gives the percentage of carriers generated beyond the value of the abscissa. It is seen that about 75 per cent of the carriers are generated below 1 micron depth, and that for a junction depth about $\frac{1}{4}$ micron, essentially all the carriers are generated below the junction.

When high-energy electrons or high-energy protons are incident on a silicon solar cell, they create local disorder in the crystal which results in a steady decrease of diffusion length with time. A simple theory for

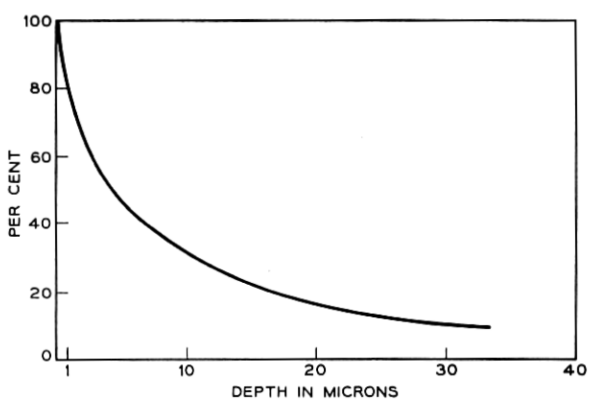

Fig. 10 — Distribution of free carriers generated in silicon in response to sunlight.

the degradation of diffusion length predicts that the diffusion length L should depend on the total flux Φ of electrons or protons according to the equation:

$$\frac{1}{L^2} = \frac{1}{L_0^2} + K\Phi \tag{3}$$

where L_0 is the value of the diffusion length before irradiation and K is a constant for a given energy of particle and for a given semiconductor. Hence, for large enough radiation fluxes, the diffusion length is inversely proportional to the square root of the flux. Fig. 11 shows a plot of diffusion length versus flux of 1 mev electrons. The experimental points were obtained by measuring the diffusion length in silicon after successive exposure to 1 mev electrons from a Van de Graaff generator. The line on Fig. 11 is a two-parameter fit of (3) to the experimental data. Similar results are obtained for proton bombardment.

As the diffusion length in a solar cell decreases with exposure to radiation, fewer and fewer of the carriers generated deep in the silicon are collected at the junction. Thus, the power output of the solar cell decreases. Since, as pointed out earlier, the depth of generation increases with the wavelength of light, the solar cell degrades initially by loss of response to the longer wavelength, i.e., the red light. This fact has a number of implications for the design of solar cells for use in the Van Allen belt. Firstly, since it is the blue response that is likely to be maintained, and this response involves the carriers generated closest to the surface, it is most important for satellite solar cells that the junction depth be minimized. Secondly, it is important that any antireflective

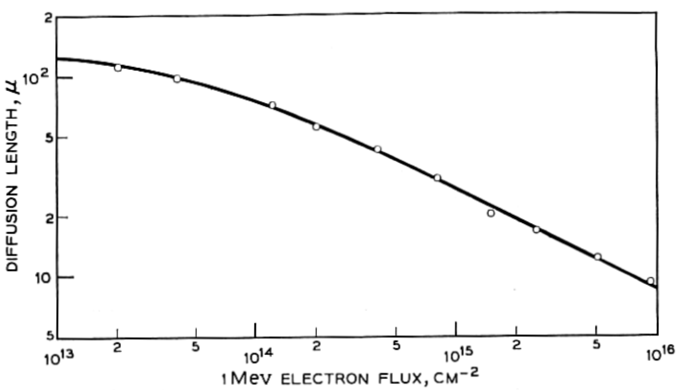

Fig. 11 — Diffusion length vs flux of 1 mev electrons.

coating be optimized for blue light, not for red. Initial good response to red light, which calls for long diffusion length, becomes of lesser importance.

It has been found by several investigators that the decrease of diffusion length in response to electron and proton bombardment is less rapid in p-type silicon than it is in n-type silicon.[10] For this reason, cells for satellite use are preferably made with a thin n-skin on a p-type body rather than the other way around. Fig. 12 is a schematic diagram of a solar cell designed at Bell Telephone Laboratories and incorporating the features just discussed.[11] It is made on a p-type silicon body with an n-layer $\frac{1}{4}$ micron thick. In order to produce such a thin layer with good properties, it is necessary to minimize surface damage. For this reason the surface used is given an optical polish. Such a thin layer tends to have high sheet resistance and calls for many contact fingers to minimize the effect of series resistance. Finally, the cell is given an antireflection coating of thickness designed to optimize the response to blue light.

Having designed a cell to minimize the effects of radiation damage, it is then necessary to consider what, if anything, can be done to shield the cells from the radiation. In the case of electrons, substantially all of which have energies of less than 1 mev, such shielding is practical using materials like quartz or sapphire. Fig. 13 shows the measured degradation of the short-circuit current of variously shielded solar cells after electron bombardment corresponding to increasing time in the Van Allen belt. The shield thicknesses are represented as g/cm². It is seen that over the range for which the measurements were made — which

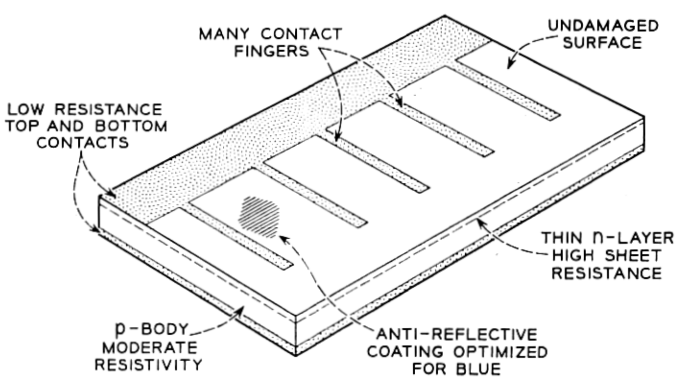

Fig. 12 — Structure of Bell Laboratories solar cell for satellite use.

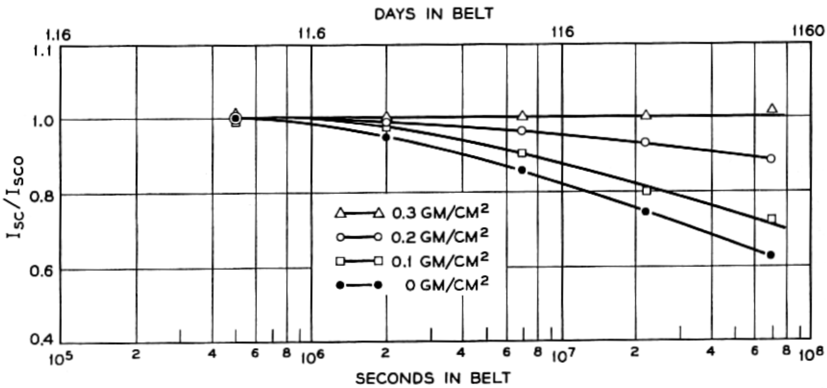

Fig. 13 — Solar cells with various shielding: measured degradation of short-circuit current after electron bombardment.

was equivalent to two years in the Van Allen belt — the effect of electrons was eliminated by the use of 0.3 g/cm² of shielding. Shielding of protons, which are much more energetic, would require intolerable weights of material. However, the 0.3 g/cm², which eliminates the electron damage, does provide some reduction in the proton damage.

Fig. 14 is a plot of the anticipated power output of the solar cells

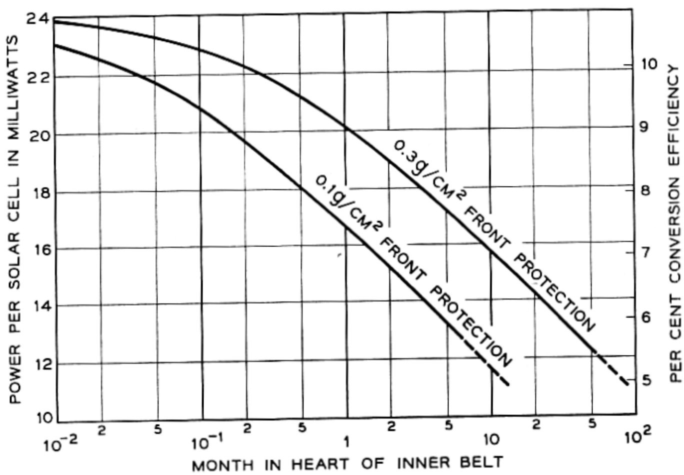

Fig. 14 — Anticipated power output of solar cells as function of time in Van Allen belt; with present data, error factor may be as great as 3 in time.

shown in Fig. 12 as a function of months in the heart of the Van Allen belt. The curves were obtained by estimating the densities and energy distributions of electrons and protons in the Van Allen belt and subjecting the cells to electron and proton bombardments simulating Van Allen conditions. There may be considerable errors in the estimation of Van Allen radiation and, as a result, the time to a given degradation may well be in error by a factor as great as 3. It should further be noted that the curves have been calculated for the case of a satellite that spends all its time in the Van Allen belt, and this is certainly pessimistic. A satellite in a circular polar orbit, for example, would spend approximately $\frac{1}{6}$ of the time in the Van Allen belt.

The most significant feature of the curves in Fig. 14 is that the plot of power output per solar cell versus log time is approximately linear after initial degradation. This dependence is consistent with the anticipated variation of diffusion length with flux, Fig. 11, and the distribution of carriers generated in the silicon, Fig. 10. The degradation with time becomes progressively less severe at longer times. Thus, for the case of 0.3 g/cm^2 protection, the output after 10 months has dropped from an initial value of 24 mw to about 16 mw while at the end of 100 months it has dropped further only to 11 mw. This additional decrease in power output for a factor of 10 increase in time could be compensated for by a 50 per cent increase in the number of solar cells. It appears then that provided there has been no gross underestimate of the nature and effect of the Van Allen belt radiation, solar cell power can be provided for a design life of five years and that the design life could be increased without excessive penalty. The curves also illustrate the design choices that can be made in selecting the mass of front protection. It is seen that for a given power output per cell, a factor of 3 increase in weight of protection yields about a factor of 5 improvement in life. However, the same improved life for a given power output could be achieved by retaining the lighter front protection but increasing the number of cells by 30 per cent. Just which is the best design of front protection thickness will depend on the particular satellite under consideration. For the case of the experimental satellite being designed at Bell Telephone Laboratories, a front protection consisting of 0.3 g/cm^2 of sapphire was found to be the best choice. Fig. 15 is a photograph of some solar cell modules with and without the sapphire protection.

The solar cell is yet another example of a component which can give adequate life performance only if the component is properly designed and used conservatively. In this case, conservative use involves paying

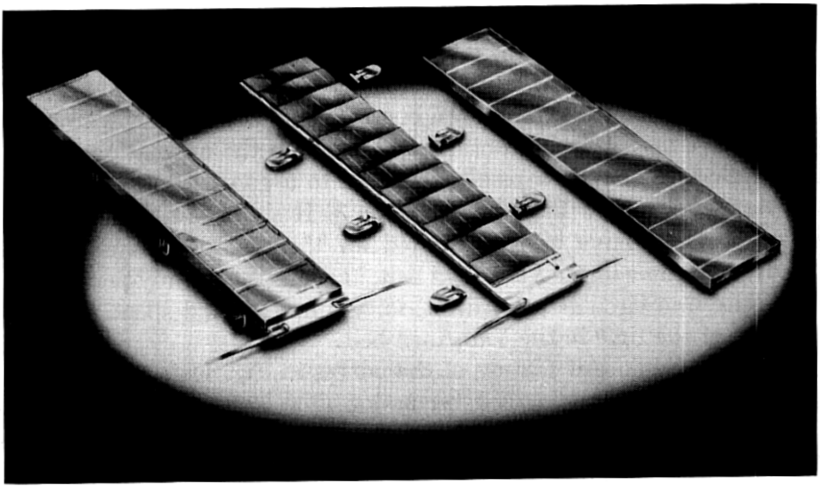

Fig. 15 — Photograph of solar cells, without protection (center) and with sapphire shields.

the weight penalty of sufficient radiation protection and increasing the number of solar cells to allow for some inevitable loss of power output per cell in response to radiation.

V. CONCLUSIONS

Returning to Table I, it is seen that the failure rate of 20 per 10^9 hours chosen for transistors in case I is probably a conservative figure. This degree of reliability has already been observed in the field on older devices that did not have the benefit of more recent design improvements and that were not life tested, selected and carefully handled as devices would be for satellite use. With proper selection and handling care, these older devices would almost certainly meet the requirements for case II and possibly for case III. The results of accelerated aging of the newer product lead to predictions of at least one order of magnitude improvement in transistor reliability. Assuming that at least some of this improvement will be realized under operating conditions, one expects that transistor performance is adequate for case III. The reliability of diodes, which approximates that for transistors, is similarly adequate for case III. Should transistor and diode failure rates indeed turn out to be in the region of one per 10^9 hours, then more complex satellites could be designed with life expectancy much longer than five years.

It further appears that traveling-wave tubes can be made that will survive launch and should not limit the life in orbit. Finally, even under the most pessimistic assumptions as to the nature of the Van Allen belt, solar cell power plants can be provided, at a weight penalty, to meet the required life. More precise design of solar cell power supplies will only be possible when more precise and extensive data are available on the nature of the Van Allen belt.

Adequately reliable communication satellites can therefore be made, provided they incorporate components of proven integrity which are used in a conservative design. The use of components of proven integrity involves expense for high-quality design, careful manufacture and painstaking selection. The use of such components does not permit the performance advantages that might be gained with use of developmental components. In the final analysis, conservative design leads to more weight per given function. Typical examples are the increased weight of a rugged traveling-wave tube, the weight of solar cell protective covers, the weight of additional solar cells to allow for the inevitable degradation in the Van Allen belt, and the additional weight of circuitry designed with ample margins.

Hence, limitations of weight in orbit and requirements of long life in orbit both result in a limit on the complexity of the satellite. Communication satellites in the immediate future must be simple. As higher component reliability is demonstrated and as improved vehicles permit greater payloads, so can the complexity of the satellites increase.

VI. ACKNOWLEDGMENTS

In preparing this paper, the author has drawn heavily on the work and advice of many colleagues at Bell Telephone Laboratories. He is particularly indebted to B. T. Howard and to D. S. Peck.

REFERENCES

1. Glaser, J. L., The Design of Medium Height Random Orbit Satellite Systems, NEREM Record — 1961, p. 39.
2. The Bell System Technical Journal, **40,** July, 1961, complete issue.
3. Peck, D. S., A Mesa Transistor Reliability Program, Solid State J., Nov./Dec., 1960.
4. Peck, D. S., Semiconductor Reliability Predictions from Life Distribution Data, in *Semiconductor Reliability*, edited by Shwop, J. E. and Sullivan, H. J., Engineering Publishers, Elizabeth, N.J., 1961.
5. Dodson, G. A. and Howard, B. T., High Stress Aging to Failure of Semiconductor Devices, Proc. Seventh National Symposium on Reliability and Quality Control, Philadelphia, Penn., Jan., 1961.
6. McNally, J. O., Metson, G. H., Veazie E. A., and Holmes, M. F., Electron Tubes for the Transatlantic Cable System, BSTJ, **36,** 1957, p. 163.

7. Laico, J. P., McDowell, H. L., and Mosler, C. R., A Medium Power Traveling-Wave Tube for 6000 mc Radio Relay, B.S.T.J., **35,** 1956, pp. 1285.
8. Bradford, C. E., and Laico, J. P., Ruggedized Traveling-Wave Tubes for Missile Use, Proc. of 1960 Electronic Components Conference, Washington, D.C., pp. 91–95.
9. Van Allen, J. A., J. Geophys. Res., **64,** 1959, p. 1683.
Holly, F. E., and Johnson, R. G., J. Geophys. Res., **65,** 1960, p. 771.
Freden, F. C., and White, R. S., Phys. Rev. Lett., **3,** 1959, p. 9 and J. Geophys. Res., **65,** 1960, p. 1399.
Naugle, J. E., and Fichtel, C. E., Am. Phys. Soc. Bull. II, **6,** 1961, p. 53.
Singer, S. F., *Space Research,* North Holland Publishing Co., Amsterdam, 1960, p. 797.
Naugle, J. E., Nucleonics, **19,** April, 1961, p. 89.
10. Madelkorn, J., et al, A New Radiation-Resistant High-Efficiency Solar Cell, delivered at a meeting on "Radiation Damage to Semiconductors by High Energy Protons," sponsored by N.A.S.A., Wash., D.C., Oct. 20, 1960. Available from U.S. Army Signal Res. and Dev. Lab., Fort Monmouth, N.J.
11. Smits, F. M., Smith, K. D., and Brown, W. L., Solar Cells for Communication Satellites in the Van Allen Belt, presented at 1961 Convention, The British I.R.E., Oxford, to be published in Proc. of the British I.R.E.