THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XLI SEPTEMBER 1962 NUMBER 5

On Rearrangeable Three-Stage Connecting Networks

By V. E. BENEŠ

(Manuscript received December 26, 1961)

A class of three-stage connecting networks proven rearrangeable by D. Slepian is considered. Bounds on the number of calls that must be moved are obtained by some simple new methods.

I. INTRODUCTION

Most communications systems contain a *connecting network* as a basic functional unit. A connecting network is an arrangement of switches and transmission links through which certain terminals can be connected together in many combinations.

The calls in progress in a connecting network do not usually arise in a predetermined time sequence. Requests for connection (new calls) and terminations of connection (hangups) occur more or less at random. For this reason the performance of a connecting network when subjected to random traffic is used as a figure of merit. This performance is measured, for example, by the fraction of requested connections that cannot be completed, or the *probability of blocking*.

The performance of a connecting network for a given level of offered traffic is determined largely by its *configuration or structure*. This structure may be described by stating what terminals have a switch placed between them, and can be connected together by closing the switch. The structure of a connecting network determines what combinations of terminals can be connected together simultaneously. If this structure is

too simple, only a few calls can be in progress at the same time; if the structure is extensive and complex, it may indeed provide for many large groups of simultaneous calls in progress, but the network itself may be expensive to build and difficult to control.

The structure of a connecting network also gives rise to various purely combinatory properties that are useful in assessing performance. For example, C. Clos[1] has exhibited a whole class of connecting networks that are *nonblocking*: no matter in what state the network may be, it is always possible to connect together an idle pair of terminals without disturbing calls already in progress. We call such a network *nonblocking in the strict sense*, because it has no blocking states whatever.

If a connecting network does have blocking states, it is nevertheless possible that by suitably choosing routes for new calls one can confine the trajectory of the operating system to nonblocking states. That is, there may exist a rule whose use in putting up new calls results in avoiding all the blocking states, so that the system is effectively nonblocking. The rule only affects new calls that could be put into the network in more than one way; no call already in progress is to be disturbed. Connecting networks for which such a rule exists we call *nonblocking in the wide sense*.

In practice, the procedure of routing the calls through the network is called "packing" (the calls), and the method used to choose the routes is called a "packing rule." The use of the word "packing" in this context was undoubtedly suggested by a natural analogy with packing objects in a container. Virtually nothing rigorous is known about the effect of packing rules on network performance.

Finally, a connecting network may or may not have the property of being *rearrangeable*: given any set of calls in progress and any pair of idle terminals, the existing calls can be reassigned new routes (if necessary) so as to make it possible to connect the idle pair.

These three combinatory properties of connecting networks have been given general topological characterizations in a previous paper.[2] In this paper we consider the last property described, that of rearrangeability, and we study the extent to which it applies to a specific class of connecting networks.

Fig. 1 shows a typical member of an interesting and useful class of connecting networks that has been suggested and studied by C. Clos.[1] We refer to this class as that of *three-stage Clos networks*. Such a network consists of two symmetrical outside stages of rectangular switches, with an inner stage of square switches. It is completely determined by the integer parameters m, n, r that give the switch dimensions. In one of the few outstanding contributions to the theory of connecting networks,

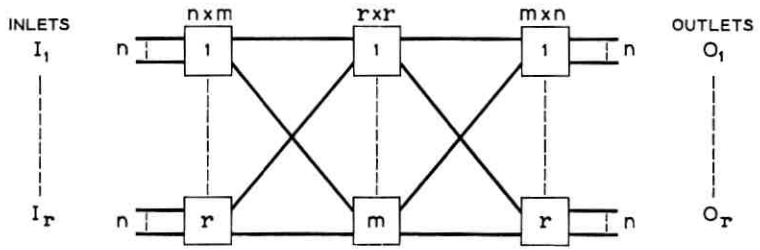


Fig. 1 — Three-stage Clos network $N(m,n,r)$.

Clos[1] showed that for $m \geqq 2n - 1$ the network is nonblocking in the strict sense. The network defined by the parameters m, n, r will be denoted by $N(m, n, r)$.

II. SUMMARY

The following two known results about rearranging three-stage Clos connecting networks are discussed:

i. The Slepian-Duguid theorem, which states that the network $N(m, n, r)$ is rearrangeable if and only if $m \geqq n$.

ii. The theorem of M. C. Paull, which states that if $m = n = r$, then at most $n - 1$ existing calls need be moved in $N(n, n, n)$ in order to connect an idle terminal pair.

The principal new result proven is a generalization (and possible improvement) of Paull's bound in (*ii*) for any m, n, r with $m \geqq n$ to $r - 1$.

The Slepian-Duguid theorem is proved in Section III by an inductive method due to Duguid[3] depending on the combinatory theorem of P. Hall on distinct representatives of subsets. We discuss Paull's theorem in Section IV, but defer our simple proof of it to Section VI, which presents simple inductive proofs of various bounds on the number of calls that must be moved. All the proofs to be given depend on a "canonical reduction" procedure that consists in removing a middle switch from the network and reducing the parameters m and n by unity.

III. THE SLEPIAN-DUGUID THEOREM

The present paper is devoted to studying the property of rearrangeability for three-stage Clos networks. We shall particularly be concerned first with the possibility of rearranging calls, and later with the number of calls that must be moved. Strictly nonblocking Clos networks will not be considered except incidentally, in view of Clos's own definitive study of them.[1]

Our first result is due essentially to D. Slepian,[4] and is

Theorem 1 (Slepian-Duguid): Every three-stage Clos network with $m \geqq n$ is rearrangeable.

Proof: The proof to be given is due to A. M. Duguid.[3*] Slepian's proof was stated for the case $m = n = r$, but actually gave an explicit procedure for rearranging the existing calls so that the additional desired call could be put up. He showed for this case that at most $2n - 2$ calls must be disturbed. This bound was subsequently improved to $n - 1$ by M. C. Paull.[5] (See Section IV.)

Duguid's proof depends on a combinatorial theorem of P. Hall, which has recently come into prominence in studies of maximal flows in networks. (See D. Gale.[6])

Hall's Theorem: Let A be any set, and let $A_1, A_2, \cdots, A_r$ be any r subsets of A. A necessary and sufficient condition that there exist a set of distinct representatives $a_1, \cdots, a_r$ of $A_1, \cdots, A_r$, i.e., elements $a_1, \cdots, a_r$ of A such that

$$a_i \in A_i \qquad i = 1, \cdots, r$$

$$a_i \neq a_j \quad \text{for} \quad j \neq i,$$

is that for each k in the range $1 \leqq k \leqq r$, the union of any k of the sets $A_1, \cdots, A_r$ have at least k elements.

The condition given is obviously necessary. The interest of the theorem, and our application of it, concern the sufficiency.

We proceed now to the proof of Theorem 1. It is obviously sufficient to consider only the case $m = n$. Let the inlets to the network be denoted by $u_1, \cdots, u_N$, where $N = nr$, and let the outlets be denoted by $v_1, \cdots, v_N$. It is sufficient to prove that every maximal assignment of inlets to outlets can be realized by a state of the network. Here "maximal" means that each inlet is to be connected to exactly one outlet, and vice versa. Such a maximal assignment is obviously equivalent to a permutation on N objects. We let $\{i \to \pi(i), i = 1, \cdots, N\}$ be such a permutation; also we denote the jth inlet switch by I_j and the jth outlet switch by O_j. It is convenient to think of I_j as the set of i for which u_i is on the jth inlet switch, and of O_j as the set of i for which v_i is on the jth outlet switch.

Let K be the set of integers $\{1, 2, \cdots, n\}$. We define the subsets $\{K_i, i = 1, \cdots, n\}$ of K by the condition

$$K_i = \{j \mid \pi(m) \in O_j \quad \text{for some} \quad m \in I_i\}.$$

* In a private communication from J. H. Déjean, the author has learned that Theorem 1 was also proved by J. LeCorre in an unpublished memorandum dated 1959.

Now let $I_{i(1)}, \cdots, I_{i(k)}$ be any k of the inlet switches, and set

$$T = \bigcup_{j=1}^{k} K_{i(j)}.$$

Suppose that there are t distinct elements in T. Then all the kr inlets in the set

$$\bigcup_{j=1}^{k} I_{i(j)}$$

are assigned by $\pi(\cdot)$ to outlets from t of the outlet switches, that is, to outlets from a set of tr outlets. But two distinct inlets are not assigned to one outlet, so $t \geqq k$. Thus any union of k sets among the K_i contains at least k elements.

Hence by Hall's Theorem there is a set of distinct representatives $\{k(i), i = 1, \cdots, n\}$ with

$$k(i) \in K_i \qquad i = 1, \cdots, n$$

$$k(i) \neq k(j) \quad \text{for} \quad i \neq j.$$

Since K contains n elements, it follows that $\{i \to k(i), i = 1, \cdots, n\}$ is a permutation. However, the interpretation of the fact that $k(i) \in K_i$ is that

$$\pi(m) \in O_{k(i)} \quad \text{for some} \quad m \in I_i.$$

In other words, to every inlet switch I_i there corresponds a unique outlet switch $O_{k(i)}$ such that $\pi(\cdot)$ maps some inlet on I_i into some outlet on $O_{k(i)}$. That is, there is a subassignment of $\pi(\cdot)$ that involves exactly one terminal on every inlet and outlet switch.

It is evident that such a subassignment can always be satisfied on a single middle switch (Fig. 1), say that numbered 1. If this subassignment is completed, that one switch is filled to capacity, and the rest of the network is essentially $N(m - 1, n - 1, r)$, i.e., that of Fig. 1 with the parameters m, n reduced by unity.

The theorem is clearly true for $m = n = 1$. As an hypothesis of induction assume that it is true for a given value of $m - 1$ $(= n - 1)$. The argument given above proves that it is then also true for m $(= n)$, for the induction hypothesis implies that the remainder of the assignment $\pi(\cdot)$ that was not put up on the first switch is satisfiable in the subnetwork, i.e., essentially in $N(m - 1, n - 1, r)$. Hence $\pi(\cdot)$ is realizable, and the theorem follows by induction on n.

IV. THE NUMBER OF CALLS THAT MUST BE MOVED: PAULL'S THEOREM

In view of the result of Slepian and Duguid that every three-stage Clos network with $m \geqq n$ is rearrangeable, it is natural to ask, for a given state x of such a network, how many calls of x need actually be changed to new routes in order to put in a given call between idle terminals. Slepian's original procedure was for the case $m = n = r$, and gave the upper bound $2n - 2$ (uniformly for all states) to the number of calls that must be disturbed. That is, he showed that if $m = n = r$, then at most $2n - 2$ calls need be rearranged. By a similar but more complicated method, M. C. Paull[5] halved this bound, proving

Theorem 2: Let $N(n, n, n)$ be a three-stage Clos network with $m = n = r$. Let x be an arbitrary state of this network. The largest number of calls in progress in x that must be rerouted in order to connect an idle pair of terminals is $n - 1$; there exist states which achieve this bound.

Since Paull's proof was involved, we have looked for and found simpler ways of proving and extending his result. In Section VI we give a simple inductive proof; the argument to be given, of course, also provides a proof of the Slepian-Duguid theorem not depending on the Hall combinatorial result used in Section III.

V. SOME FORMAL PRELIMINARIES

In order to state and prove the rest of our results, it is useful, and indeed necessary, to introduce a systematic notation. Such a notation has been described and used in a previous paper[2] by the author; the notation to be used is a consistent extension of this.

The set of inlets of a network is denoted by I, and that of outlets by Ω. The set of possible states of a connecting network is denoted by S. For a three-stage Clos network, S consists of all the ways of connecting a set of inlets to as many outlets by disjoint chains (paths) through an inlet switch, a middle switch, and an outlet switch. (See Fig. 1.) States of the network may then be thought of as sets of such chains. Variables $x, y, z, \cdots$, at the end of the alphabet, range over states from S.

A terminal pair $(u, v) \in I \times \Omega$ (with u an inlet and v an outlet) is called *idle in state* x if neither u nor v is an endpoint of a chain belonging to x. A *call* c is a unit subset $c = \{(u, v)\} \subset I \times \Omega$; c is new in a state x if (u, v) is idle in x. The *assignment* $\gamma(x)$ realized by x is the union of all calls $c = \{(u, v)\}$ such that x contains a chain from u to v. If a is an assignment, $\gamma^{-1}(a)$ is the set of all states realizing a. The cardinality of a set X is denoted by $| X |$. The states $x \in S$ are partially ordered by inclusion $\leqq$ in a natural way.

A distance between states can be defined as

$$\delta(x, y) = |\, x \Delta y \,|,$$

= the number of calls that would have to be added, removed, or rerouted to change x into y,

where Δ is symmetric difference. The distance of a state x from a set X of states is defined in the usual way as

$$\delta(x, X) = \min_{y \in X} \delta(x, y).$$

A call c new in a state x is blocked in x if there is no state $y > x$ such that $\gamma(y) = \gamma(x) \cup c$. A state x is nonblocking if no call new in x is blocked in x. The set of nonblocking states is denoted by B'. For any call c, the set of states x in which c is both new and not blocked is designated B_c'.

For a three-stage Clos network $N(m, n, r)$ with $m \geqq n$ we define

$$\begin{aligned}
\varphi_x(m, n, r) &= \max_{c \text{ new in } x} \delta(x, \gamma^{-1}(\gamma(x) \cup c)) - 1 \\
&= \max_{c \text{ new in } x} \delta(x, \gamma^{-1}(\gamma(x)) \cap B_c') \\
&= \max_{c \text{ new in } x} \min_{y \in \gamma^{-1}(\gamma(x)) \cap B_c'} \delta(x, y) \\
&= \text{the maximum number of calls that must be rerouted in order to put up a call } c \text{ new in } x.
\end{aligned}$$

We also set

$$\varphi(m, n, r) = \max_{x \in S} \varphi_x(m, n, r).$$

In this last definition, it is assumed that S is the set of states determined by the parameters m, n, r in Fig. 1.

In the notation introduced above, the Slepian-Duguid Theorem guarantees that for $m \geqq n$ and c new in x

$$\gamma^{-1}(\gamma(x) \cup c) \neq 0,$$

$$\gamma^{-1}(\gamma(x)) \cap B_c' \neq 0,$$

and Paull's Theorem may be cast as stating that

$$\varphi(n, n, n) = n - 1.$$

VI. THE NUMBER OF CALLS THAT MUST BE MOVED: NEW RESULTS

We now present some new methods for studying the number of calls that must be moved; these yield extensions of results of D. Slepian[4] and M. C. Paull.[5]

Theorem 3: $\varphi(2, 2, r) \leqq 2r - 2$.

Proof: Suppose that a blocked new call between input switch I_1, and output switch O_1 is to be put in when the network is in a state x. Consider any sequence $c_1, \cdots, c_k$ of existing calls of x with the properties

i. Either c_1 is on I_1, c_1 and c_2 are the same outlet switch, $\cdots$,

$$\begin{aligned} &c_i \text{ and } c_{i+1} \text{ are on the same outlet switch,} && i \text{ odd,} && i < k \\ &c_i \text{ and } c_{i+1} \text{ are on the same inlet switch,} && i \text{ even,} && i < k, \end{aligned}$$

or c_1 is on O_1, c_1 and c_2 are on the same inlet switch, $\cdots$,

$$\begin{aligned} &c_i \text{ and } c_{i+1} \text{ are on the same inlet switch,} && i \text{ odd,} && i < k \\ &c_i \text{ and } c_{i+1} \text{ are on the same outlet switch,} && i \text{ even,} && i < k. \end{aligned}$$

ii. c_k is the only call on some outer switch. Since neither I_1 nor O_1 is full, the largest k for which such a sequence exists is $2r - 2$. The reader can verify that a possible strategy for rearranging existing calls of x so as to put in an I_1-O_1 call is to take each call of the sequence $c_1, \cdots, c_k$ and *reverse* its route, i.e., make it go through the *other* middle switch than the one it presently uses. Thus for all x

$$\varphi_x(2, 2, r) \leqq 2 - 2.$$

Let x be a state of $N(m, n, r)$, and let M be a particular middle switch. A *canonical reduction* of x with respect to M will consist of

i. removing M,

ii. on each outer switch that has a call routed via M, removing the link, crosspoints, and terminals associated with that call,

iii. on each outer switch that has an idle link to M, removing the link, the crosspoints associated therewith, and one arbitrarily chosen idle terminal.

It is easily seen that a canonical reduction of a state x of $N(m, n, r)$ leads to a state of $N(m - 1, n - 1, r)$.

Theorem 4: $\varphi(n, n, r) \leqq 2r - 2$.

Proof: By Theorem 3, the result holds for $n = 2$, so assume it for a given value of $n - 1 \geqq 2$, and try to rearrange a given state x of $N(n, n, r)$ so as to put in a new blocked call from I_1 to O_1.

Case 1: There is a middle switch M with both an I_1 and an O_1 call on it. Perform a canonical reduction of the state x with respect to M. This yields a state of $N(n - 1, n - 1, r)$, for which the result holds.

Case 2: No middle switch has both an I_1 and an O_1 call on it. Since the call to be put in is blocked, it must be true that

$$\#\text{ (idle links out of } I_1) + \#\text{ (idle links out of } O_1) = n$$

and hence

$$\max\{\#\text{ (idle out of } I_1), \#\text{ (idle out of } O_1)\} > 1.$$

Suppose that # (idle out of I_1) > 1. There is a middle switch M with an idle link to I_1, and a busy link to O_1. Perform a canonical reduction of x with respect to M, yielding a state of $N(n-1, n-1, r)$ in which each of I_1, O_1 still has an idle terminal.

A refinement of this method suggested by M. C. Paull will halve the last two bounds. We prove

Theorem 5: $\varphi(2, 2, r) \leqq r - 1$. $(r \geqq 2)$

Proof: The result is true for $r = 2$, since in that case the network has only one blocking state (see Fig. 2), and both blocked calls can be unblocked by changing the route of one $(= r - 1)$ existing call.

Let us assume as an hypothesis of induction that the theorem holds for some value of $r - 1 \geqq 2$, and in $N(2, 2, r)$ attempt to put up a blocked new call c between input switch I_1 and output switch O_1. Since c is new and blocked, there must be an idle and a busy link on both of I_1 and O_1, and each of the busy links must pass through a different middle switch. Let c_1 be the call on I_1, and c_2 be the call on O_1. We may suppose without loss of generality that c_1 is a call from I_1 to O_2, while c_2 is a call from I_2 to O_1.

Case 1: I_2 has only one call on it, viz., c_2. Move c_2 to the other middle switch (see Fig. 3).

Case 2: I_2 has two calls on it. Remove both c_1 and c_2, so that I_1 and O_1 become empty. Consider now the state x of the subnetwork of parameter $r - 1$ obtained by removing I_1 and O_1 and reducing the dimension of the two square middle switches by unity to $r - 1$. Each of I_2 and O_2 has at least one idle terminal in x, since c_1 and c_2 were removed. Hence by the hypothesis of induction the subnetwork can be rearranged so as to put in a call from I_2 to O_2 while disturbing at most $r - 2$ existing calls. If the I_2-O_2 path thus provided is via M_1 then c_1 and c_2 can be replaced as in

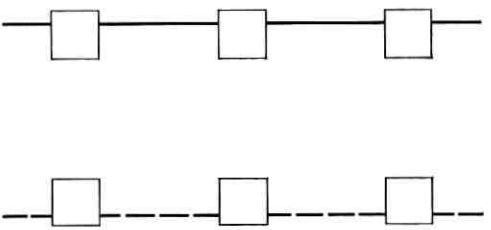


Fig. 2 — Network with only one blocking state ($r = 2$).

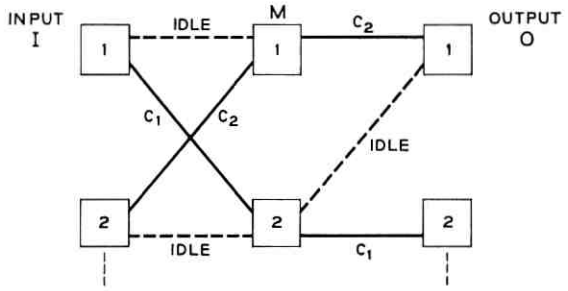


Fig. 3 — I_2 with one call, c_2 .

Fig. 4. This leaves a path for the new I_1-O_1 call c via M_2 , and shows that it was never necessary to move c_2 , and that hence at most $r - 1$ calls were disturbed. If the I_2-O_2 path provided by rearranging the subnetwork is via M_2 , then c_1 and c_2 can be replaced as in Fig. 5. This leaves a path for c via M_1 , and shows that c_1 did not really have to be moved, so that at most $r - 1$ calls were disturbed.

Theorem 6: $\varphi(n, n, r) \leqq r - 1$.

Proof: The result is true for $n = 2$. Assume that the theorem is true for a given value of $n - 1 \geqq 2$, and seek to rearrange a state x of $N(n, n, r)$ so as to put in a new call blocked in x between I_1 and O_1 . The theorem follows by induction on n by distinguishing two cases as in Theorem 4, and using a canonical reduction of x.

Theorem 7: For $m - 1 \geqq n$,

$$\varphi(m, n, r) \leqq \varphi(m - 1, n, r).$$

Proof: This is almost obvious. Remove any middle switch M of $N(m, n, r)$ and make all terminals on which there were calls routed via M idle. This gives a state of $N(m - 1, n, r)$; in this state the desired call

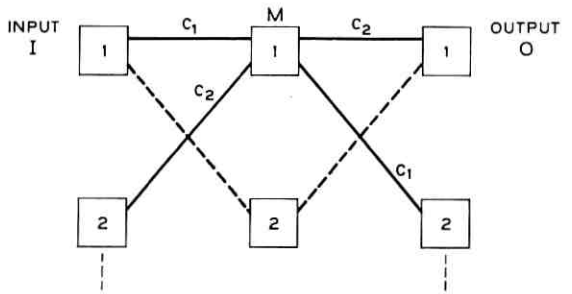


Fig. 4 — Calls c_1 and c_2 over path via M_1 .

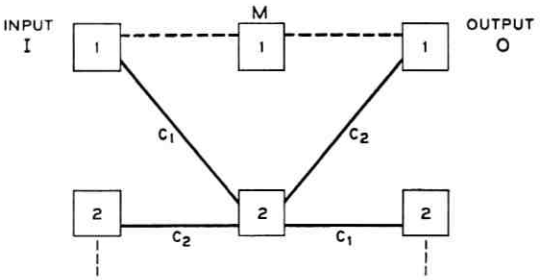


Fig. 5 — Calls c_1 and c_2 over path via M_2.

can be put in by rearranging at most $\varphi(m - 1, n, r)$ existing calls. Now replace M and the calls that were routed through it.

M. C. Paull[3] has conjectured that if $r = n$, then

$$\varphi(m, n, n) \leqq 2n - 1 - m.$$

This bound agrees with Theorem 2 if $m = n$, and with Clos' results on nonblocking networks if $m = 2n - 1$. Paull has proved the result for $m = 2n - 2$. However, no proof of the full conjecture has been found. It is tempting to try the stronger conjecture that

$$\varphi(m, n, r) \leqq 2n - 1 - m$$

for any m, n, and r. This can be disproved by the counterexample shown in Fig. 6. There is no way of connecting I_1 to O_5 without moving a call

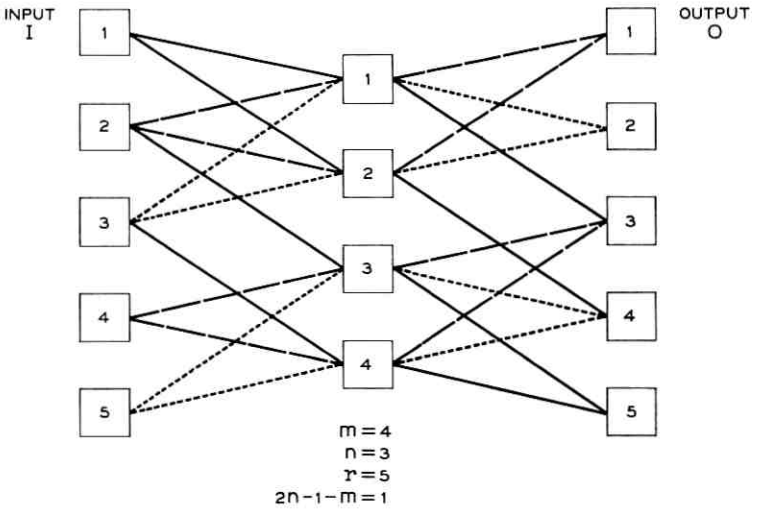


Fig. 6 — Network showing that I_1 and O_5 cannot be connected without moving a call on one of I_1, O_5.

on one of I_1 , O_5 . However, all possible alternative routes for these calls are pre-empted, so at least two calls must be moved.

VII. ACKNOWLEDGMENT

The author is indebted to M. C. Paull for helpful comments, and for suggestions that led to Theorems 5 and 6.

REFERENCES

1. Clos, C., A Study of Non-Blocking Switching Networks, B.S.T.J., **32** (1953), pp. 406–424.
2. Beneš, V. E., Algebraic and Topological Properties of Connecting Networks, B.S.T.J., **41** (1962), pp. 1249–1274.
3. Duguid, A. M., Structural Properties of Switching Networks, Brown University, Progress Report BTL-7, 1959.
4. Slepian, D., Two Theorems on a Particular Crossbar Switching Network, unpublished manuscript, 1952.
5. Paull, M. C., Reswitching of Connection Networks, B.S.T.J., **41** (1962), pp. 833–855.
6. Gale, D., A Theorem on Flows in Networks, Pacific Journal of Mathematics, **7** (1957), pp. 1073–1082.

A High-Precision Direct-Reading Loss and Phase Measuring Set for Carrier Frequencies

By J. S. ELLIOTT

(Manuscript received December 28, 1961)

A set has been developed to measure the parameters of insertion loss and phase shift of communication systems components. The design effort has been directed toward achieving laboratory precision in measurement and at the same time maintaining the ease of use and speed necessary for a large volume of measurements. Accuracies of ±0.002 db, ±0.02 degree and ±1 cycle have been attained over the frequency range from 10 to 300 kc. The entire frequency range is covered without band switching by the use of a heterodyne signal oscillator which provides frequency accuracy by locking to a frequency standard. The principle of phase detection is based on measurement of the time interval corresponding to the displacement of sine wave zero crossings caused by the unknown. This method has the advantage of good accuracy inherent in the measurement of time by counting techniques and also the ease of automatic readout of phase shift by translation from time units. In measuring loss, use of a rapid sampling technique to compare the unknown with a standard eliminates errors caused by circuit drifts.

The ever increasing complexity of communication systems and the demand for high-quality transmission have emphasized the need for close control of system components such as filters, equalizers, and repeaters. Precise instrumentation to measure the parameters of loss and phase shift must be provided for use in both the development and production areas of these components.

In the case of broadband carrier systems this need was met by development of a 3.6 mc phase measuring set.[1] Later, the set was modified to provide increases in maximum frequency from 3.6 to 20 mc, loss measurement accuracy from ±0.05 to ±0.02 db and phase measurement accuracy from ±0.25 to ±0.1 degree.

The need for instrumentation to control components in the case of

high-speed data systems has been met by the recent development of a 10 to 300 kc phase measuring set. This set will automatically measure and read out insertion phase shift to an accuracy of 0.02 degree and will measure insertion loss to an accuracy of 0.002 db.

The performance specifications of the new set are more completely stated as follows:

Frequency: 10 to 300 kilocycles; maximum accuracy ±1 cycle

Generator and network termination impedance: 75 ohms unbalanced

Test signal level: +6 dbm

Insertion loss range: 0 to 100 db, maximum accuracy ±0.002 db

Insertion phase shift range: 0 to 360°, maximum accuracy ±0.02°

The quantities measured are defined in Fig. 1. Conforming to these definitions, the measuring system compares phase and amplitude of the outputs of two transmission channels energized from the measurement frequency source, one of which serves as a standard channel while the other contains the apparatus under test. This is illustrated in the block diagram of Fig. 2.

The measuring circuit uses the heterodyne principle,[1] which provides the high degree of frequency discrimination required for precision measurements and the ease of operation by self-tuning. Heterodyning also translates the phase of the unknown from the variable frequency to a constant low intermediate frequency at which phase shift can be accurately detected. The principle of phase detection is based on measurement of the time interval corresponding to the displacement of a sine-wave zero crossing caused by the unknown, as discussed in more detail in Section IV. This method has the advantage of good accuracy inherent in the measurement of time by counting techniques and also the ease of automatic readout of phase shift by translation from time units.

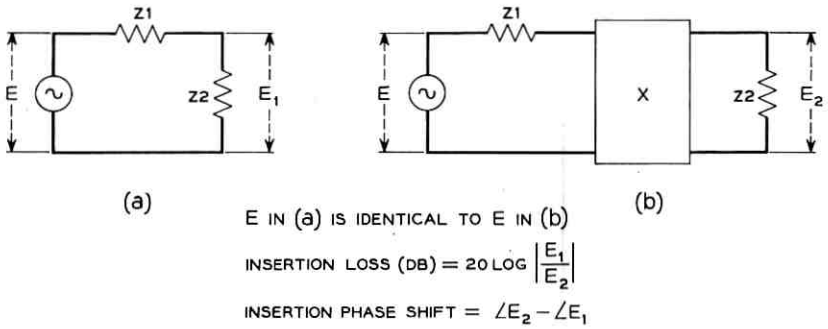


Fig. 1 — Definition of quantities measured.

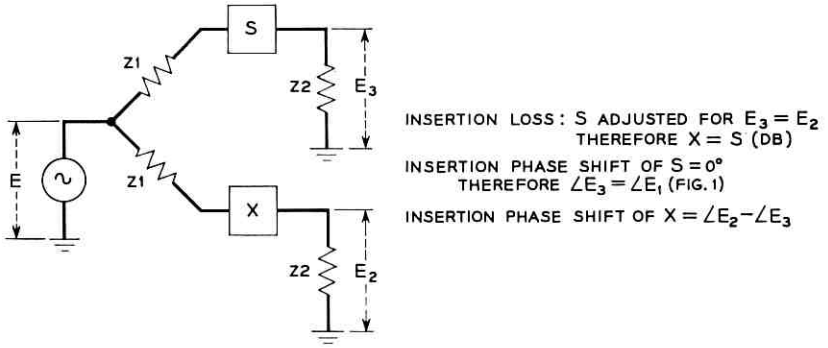


Fig. 2— Basic block diagram.

Alternate connection of the loss detection circuit to the standard and test channels[2, 3] is made at a relatively rapid rate by the use of mercury relays driven at 13 cycles per second. This fast switching method results in minimizing errors caused by magnitude instabilities in the measurement signal source which are prohibitive in manual switching methods. In phase measurements, where magnitude changes are less critical, the alternate connection of the phase detection circuit to the two channels is made at a slower rate to allow for an accurate measure of time. The overall circuit is illustrated in two sections by the block diagrams of Figs. 3 and 4.

For loss measurements a single master switch sets up the connection of Fig. 3 and the timed driving voltages for the switches. The signal at frequency F is applied through both the standard and unknown channels of the comparison unit, whose outputs are connected alternately to the input of a wideband amplifier. Approximately equal detection sensitivity for various values of loss of an unknown is attained automatically by ganging the amplifier gain controls to the standard attenuator control switches so that the level of the signals supplied to the detector is maintained constant to within 1 db. The loss detector then compares the magnitudes of the two signals and indicates their inequality on its meter. When the attenuator is adjusted for zero meter reading, the loss of the unknown is read directly from the attenuator dials. Details of the loss detection system are given in Section III.

For phase measurements, the master switch automatically connects the constant phase attenuator in place of the loss attenuator, connects the output signals of the two channels to the phase detector in place of the loss detector, as shown in Fig. 4, and changes the timing of the circuit

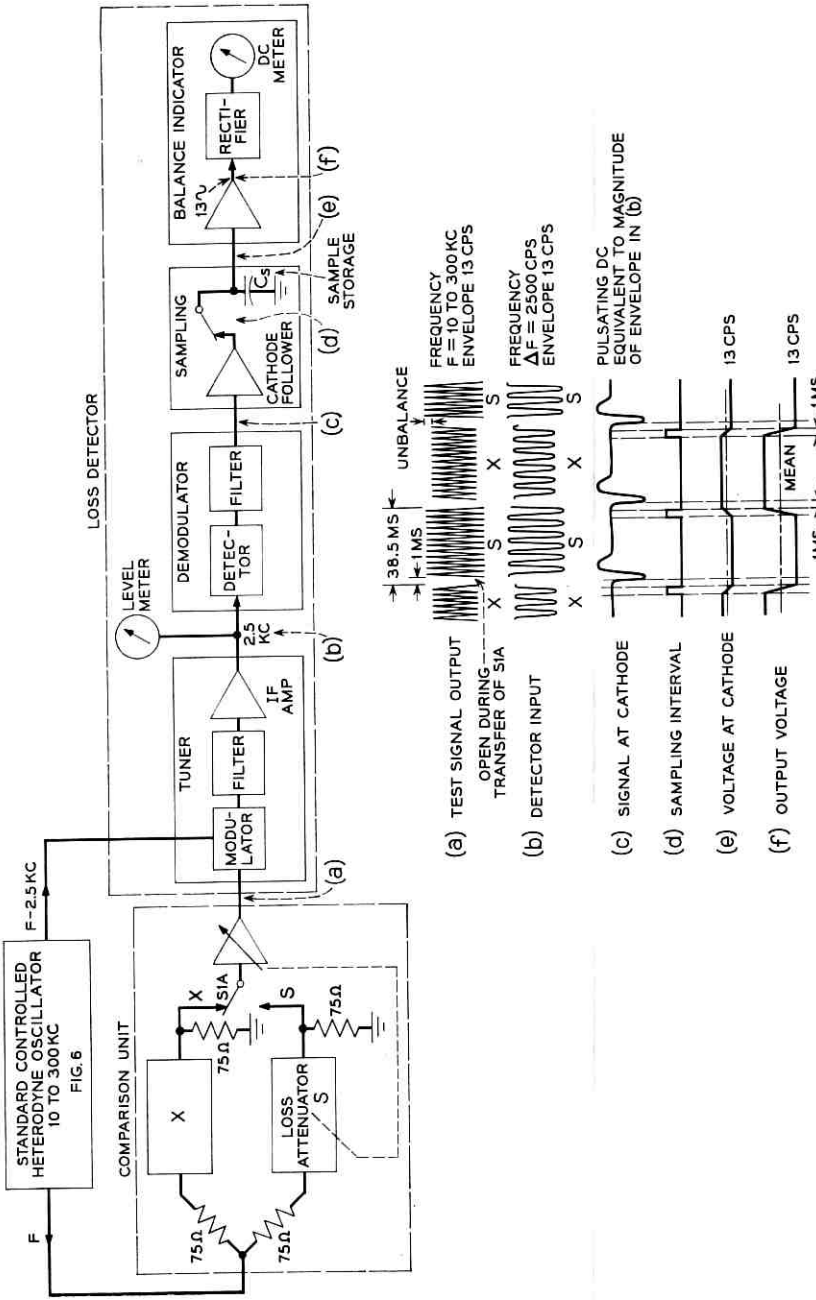


Fig. 3 — Block diagram of loss measurement circuits.

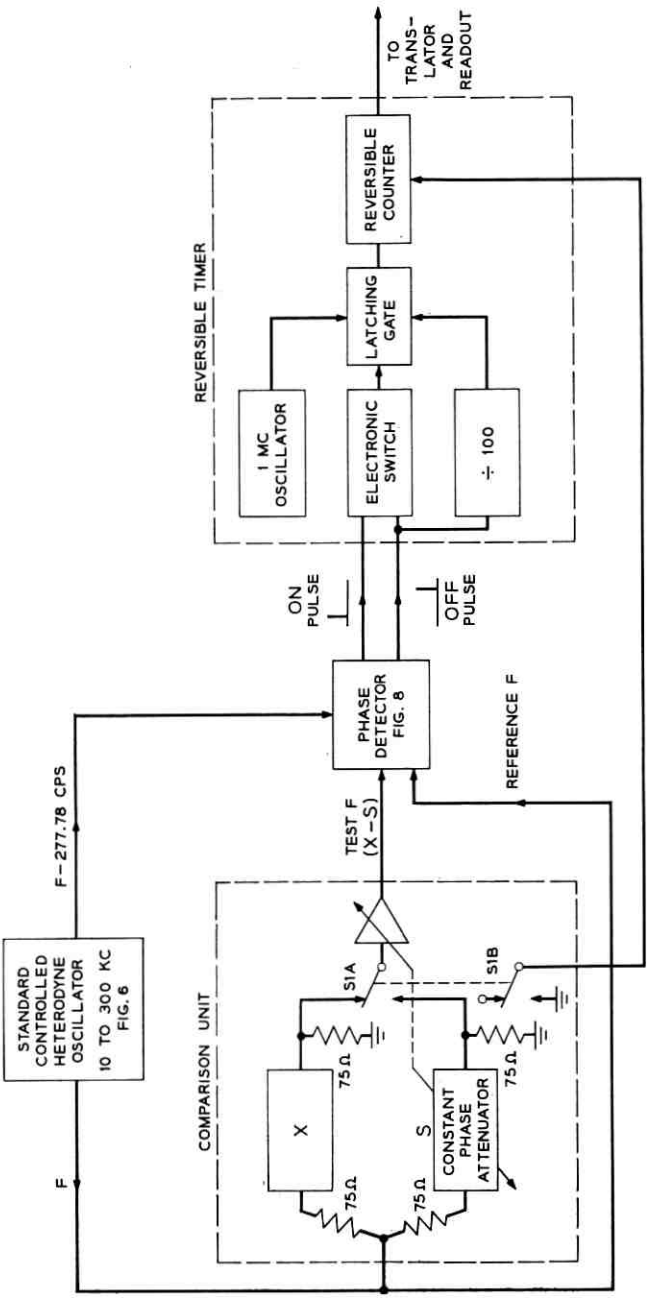


Fig. 4 — Block diagram of phase measurement circuits.

relays. The difference in phase of the two signals is now detected and indicated on an in-line readout. The constant phase attenuator is compensated so that at any setting it introduces phase shift in the standard arm exactly equal to the phase shift in the unknown arm for "zero" setting, i.e., with a patch cord inserted in place of the unknown. The indicated phase is thereby the phase shift due to insertion of the unknown. Ganging of the constant phase attenuator to the loss attenuator automatically maintains the magnitude equality of the loss measurement and thereby eliminates possible errors in phase detection due to unequal signal magnitudes. The details of the phase-detection system are also described below.

I. MEASURING SET DESIGN

The component chassis of the set are mounted in a console assembly of three relay rack cabinets as shown in Fig. 5. Particular emphasis has been placed on the arrangement of the chassis in the cabinets so that the controls of oscillators F1 and F4 and of the comparison units, which are the ones that are operated in the normal measurement procedure, are within easy reach of a seated operator. The loss balance, phase, and level indicators are also located for ease of viewing. Accessibility to chassis maintenance test points and to components for replacement is provided by mounting the most critical chassis on sliding and tilting chassis tracks. Access to rigidly mounted chassis is permitted through rear cabinet doors and in some instances through hinged front chassis covers.

The function of the individual units in relation to the overall circuit and some of the significant design considerations are discussed under the following headings:

II. STANDARD CONTROLLED HETERODYNE OSCILLATOR (SCHO)
III. LOSS DETECTOR
IV. PHASE DETECTOR
V. REVERSIBLE TIMER
VI. COMPARISON UNIT

II. STANDARD CONTROLLED HETERODYNE OSCILLATOR (SCHO)

The SCHO is the most complex part of the measuring set and includes seven separate units. Six units occupy the complete upper part of the left-hand bay and one unit is at the top of the central bay as shown in Fig. 5

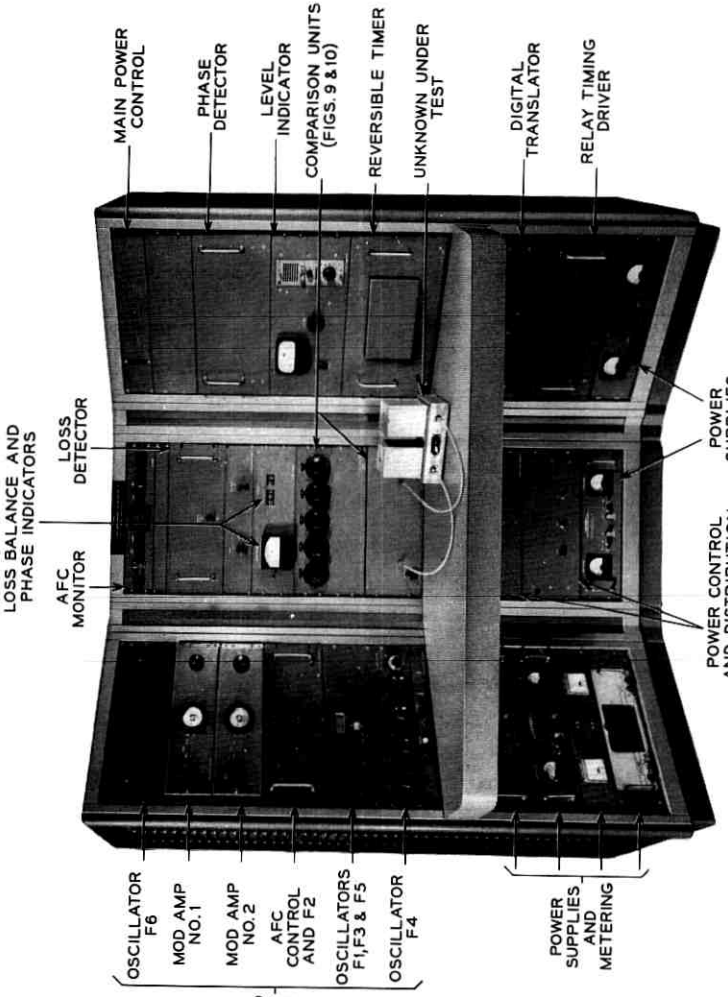


Fig. 5 — The assembled loss and phase measuring set.

It is found convenient to treat the discussion of the SCHO under the following sub-headings:

2.1 General Description and Requirements of the SCHO
2.2 Functions of Automatic Frequency Control (AFC) Loops
2.3 Design Considerations of AFC Loops

2.1 *General Description and Requirements of the SCHO*

The need to accurately measure the characteristics of networks having steep transmission and phase slopes prescribes stringent frequency accuracy requirements on the oscillator. To obtain an accuracy of ± 0.002 db in the presence of a 2 db per kilocycle slope, a frequency accuracy of ± 1 cps must be maintained. To apply the heterodyne principle, the oscillator must provide dual output frequencies: the test frequency and a slave or offset frequency equal to the test frequency minus a constant (Δ).

The design of the oscillator for this set is based on the principles of a high frequency standard controlled heterodyne oscillator (SCHO).[4] While the SCHO is a complex structure, the setting of the test signal to any desired frequency in the range of 10 to 300 kc accurate to ± 1 cycle is a simple process. The film scale is rotated to the 1 kc mark below the desired frequency and the dials of the four decades of the interpolation oscillator are set to the remaining digits of the desired frequency.

Essentially the oscillator employs local oscillators that can be set to produce the desired test and offset frequencies. These are then automatically and continuously maintained to the desired precision by reference to three standard frequency sources, two internal and one external, as shown in Fig. 6. The output frequencies, F and $F - \Delta$, are generated by a variable oscillator, F1, and two controlled oscillators, F3 and F5. F is the beat between F1 and F3 while $F - \Delta$ is the beat between F1 and F5. These beats are produced in modulator amplifiers 1 and 2 respectively.

As explained below in more detail, the frequency of F is controlled to a multiple of the external 1 kc standard through an intermediate controlled oscillator, F2. The frequency of F is then shifted to the required exact frequency by a variable interpolation oscillator, F4. Finally, the offset frequency is produced by fixed oscillators, F6 or F6′, of the appropriate Δ frequency which control the frequency difference between F3 and F5.

F1 is a stable tuned-grid oscillator[4] with a film scale 100 inches in length calibrated in terms of F at every 10 kc and further subdivided every 1 kc.

F2, F3 and F5 are tuned-grid reactance tube[1] controlled oscillators.

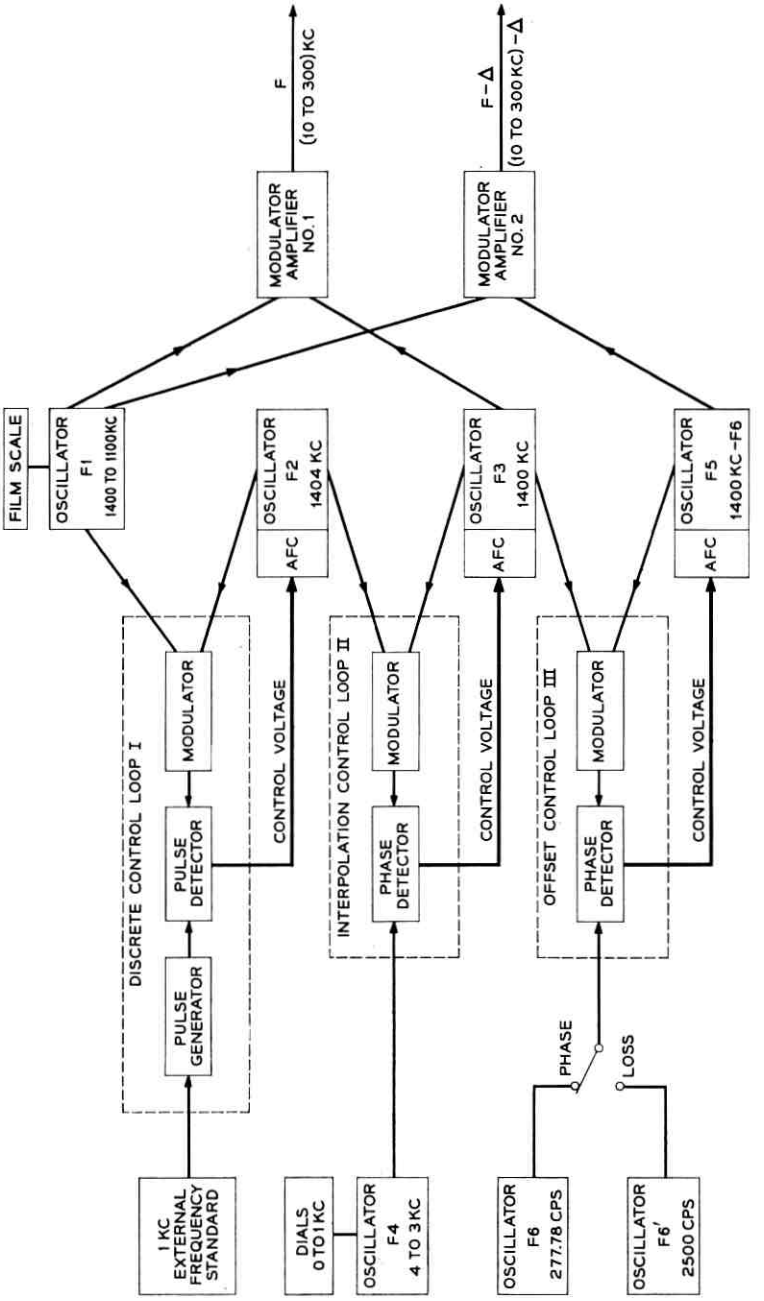


Fig. 6 — Block diagram of standard controlled heterodyne oscillator (SCHO).

F4 is a precision RC decade oscillator.

Oscillator F6 is a tuning fork oscillator, while F6′ is a crystal type. The need for the use of the two offset or Δ frequencies is discussed in Sections III and IV.

2.2 *Functions of Automatic Frequency Control (AFC) Loops*

The discrete control loop I provides voltage to control the frequency of oscillator F2 by using the 1 kc standard as a reference. F2 is controlled so that the difference F2 − F1 is an exact multiple of 1 kc. As F1 is set to within ±200 cps of any 1 kc multiple, as indicated by its calibrated film scale, F2 is shifted in frequency and phase locked to F1 by loop I as required.

The interpolation control loop II provides voltage to control the frequency of oscillator F3 by using the interpolation oscillator F4 as a reference. F3 is controlled so that the difference (F3 − F1) is the chosen 1 kc multiple plus an amount 0 to 1 kc determined by the dial settings of oscillator F4.

The offset control loop III provides voltage to control the frequency of oscillator F5 by using oscillator F6 or F6′ as a reference. F5 is controlled so that the difference F3 − F5 is equal to the frequency, Δ, of the chosen reference oscillators: F6 = 277.78 cps or F6′ = 2500 cps. Since the offset output is F5 − F1, it is equal to (F3 − Δ) − (F3 − F) or $F - \Delta$.

2.3 *Design Considerations of AFC Loops*

Modulation requirements of the SCHO oscillator are severe for both loss and phase measurements. For loss measurement errors not to exceed 0.001 db, amplitude modulation must be held to −80 db. Frequency modulation due to the operation of loop III introduces errors in phase measurement in a more complex manner. Addition of sideband frequencies to the F and F − 277.78 cycle signals fed to the modulators of the phase detector (block diagram in Fig. 7) will introduce a spurious 277.78 cycle signal in the modulator outputs. This spurious signal produces error by shifting the zero crossings, the error reaching a maximum for a 90° phase shift of the spurious component. To meet the circuit objective of 0.01 degree, the level of this interference must be 75 db down.

One source of this interference may be by frequency modulation of F5 of Fig. 6 due to insufficient filtering in loop III. This would allow transmission of frequencies Δ (277.78 cps) and 2Δ (555.56 cps) from the phase detector of loop III, Fig. 6, to the grid of the reactance tube that controls

oscillator F5. The maximum amplitude of this modulating voltage (v_m) must be less than 40 microvolts as shown in the Appendix. This low value prescribes complete, high quality shielding between the components of the frequency control circuitry of loop III and adequate filtering in the loop without introducing excessive delay.

From the description of loop functions, it will be seen that F5 must change in frequency by at least 1.5 kc, since it must compensate for frequency drifts of the order of 0.5 kc in F1, F2 and F3, as well as cover the 1 kc range of F4. In addition to this wide control range, the offset loop III must maintain the phase difference between the test and offset signals to within 0.01 degree during the phase measurement time interval to meet the objective of circuit accuracy.

The loop includes both frequency discriminating and phase detecting circuits and the reactance tube control of F5. The desired control was attained by careful design of the following factors: (a) the Q and thereby the frequency slope of the discriminator is made as high as possible with readily available components; (b) the reactance tube is operated at the point in its characteristic which produces maximum change in its output capacitance with a minimum change in grid voltage supplied by the phase detector; and (c) added gain is kept at a minimum to prevent loop sing but still sufficient to maintain control.

III. LOSS DETECTOR

The loss detector circuit is illustrated by the block diagram in Fig. 3. The detector is a self-tuned null balance type which produces an output proportional to the difference in loss of the standard and unknown branches of the comparison unit in Fig. 3 with an unbalance sensitivity of 0.001 db. The operation of the detector circuit will be described first and then the factors governing the choice of constants will be considered.

3.1 *Detector Operation*

The envelope of the test signal varies rectangularly in level (a of Fig. 3) as the comparison switch alternately connects the signal from the loss attenuator and unknown branches of the comparison circuit to the loss detector. The variable frequency signals are translated to a 2500 cps fixed intermediate frequency (b of Fig. 3), amplified, and detected with a parabolic detector. The output of the parabolic detector contains a rectangular component (c of Fig. 3), corresponding to the rectangular envelope of the carrier, and the much larger carrier component which is minimized by filtering. The 2500 cps carrier component is about 80 db

greater than the rectangular component produced by an 0.001 db inequality of the S and X signals.

Transients of variable magnitude are superimposed on the rectangular signal (c of Fig. 3) when the comparison switch is closed. Therefore, for true balance information, the output of the filter is sampled for a very short interval (d of Fig. 3) near the end of the dwell time of the comparison switch on each side. The capacitor C_S in the sampling circuit now contains a rectangular component[2, 3] (e of Fig. 3) at the switching frequency of 13 cycles. The amplitude of this component is proportional to the amount of unbalance between the standard and unknown paths, and the polarity with respect to the mean value is dependent on the sense of unbalance. This signal is then amplified, (f of Fig. 3) rectified, and the rectified voltage applied to a zero center dc meter. The meter deflection is now proportional to the loss unbalance and its direction from mid-scale indicates which of the two signals is the larger and thereby indicates whether the loss of the standard must be increased or decreased to attain balance of the signals. The balance indicator provides a relatively constant deflection for the same loss unbalance in db, independent of the absolute magnitude of loss, by maintaining the level of the input to the detector relatively constant as discussed in Section VI.

3.2 *Detector Constants and Design Considerations*

The choice of intermediate frequency, bandwidth, switching frequency and measurement period is based on the following considerations. The demodulator is a square-law rectifier detector followed by a low pass filter which must (1) transmit a rectangular envelope component, (2) provide a high attenuation to the intermediate frequency, and (3) have a rapidly decaying transient response.

The filter design problem becomes quite difficult for low values of intermediate frequency, so that a high frequency is desirable. However, if a high frequency is chosen, its low order harmonics may fall within the measurement band of 10 to 300 kc. It is difficult to provide sufficient isolation to prevent pickup of these harmonics from the IF amplifiers of the tuner through parasitic paths, such as power supplies or common grounds, back to the wide band amplifiers of the comparison unit or to units of the SCHO. Such pickup would result in measurement errors since wanted and parasitic signals could add in one position of the switch S1A of the comparison unit and subtract in the other position. An intermediate frequency of 2.5 kc was selected as a compromise value. Since only fourth and higher harmonics of 2.5 kc are in the measurement band, sufficient filtering to remove interference is less difficult.

The choice of bandwidth of the intermediate frequency is a compromise between (1) a narrow bandwidth for a high degree of frequency discrimination and thereby a high signal-to-noise ratio for good circuit accuracy and (2) a wide bandwidth to allow a fast decay of switching transients and thereby minimize their effect on circuit accuracy. To limit measurement noise below 0.001 db, a signal-to-noise ratio of 80 db is required. Bandwidth was fixed at 600 cps, which resulted in an 86 db ratio at point b of Figure 3 and a decay time of approximately 20 milliseconds.

The choice of switching rate of 13 cps is a compromise between (1) fast speed to minimize level changes of the signal source during comparison of the two paths and (2) slow speed to allow ample time for transient decay and to relax the accuracy requirement of the sampling timing in relation to the comparison time. In addition, the 13 cps rate was chosen to minimize the effect of pickup that could occur at a submultiple of the 60 cps power frequency.

IV. PHASE DETECTOR

It is convenient to divide the discussion of the phase detector into four categories: (a) General description of the principle of phase detection by measurement of time intervals, (b) Operation of the phase detector, (c) Timing of circuit measuring periods, (d) Design considerations and choice of constants.

4.1 *General Description*

As stated in the introduction, the choice of phase detection by the measurement of the time interval corresponding to the displacement of a sine wave zero crossing was based on the inherent accuracy of time measurement and the ease of automatic readout of phase by translation from time units. The circuit of the detector is illustrated by the block schematic of Fig. 7. The detector measures the phase difference between the test and reference signals of Fig. 4.

The problem of phase measurement over a wide frequency range (see Fig. 7) is simplified to the measurement at a fixed frequency by the preservation of phase in a modulation process. The considerations in the choice of 277.78 cps as the fixed frequency and the choice of a 1 megacycle counting rate will be discussed later.

The time interval, beginning with the negative-going zero crossing of the reference signal and ending with the corresponding negative-going zero crossing of the test signal, is proportional to the phase difference

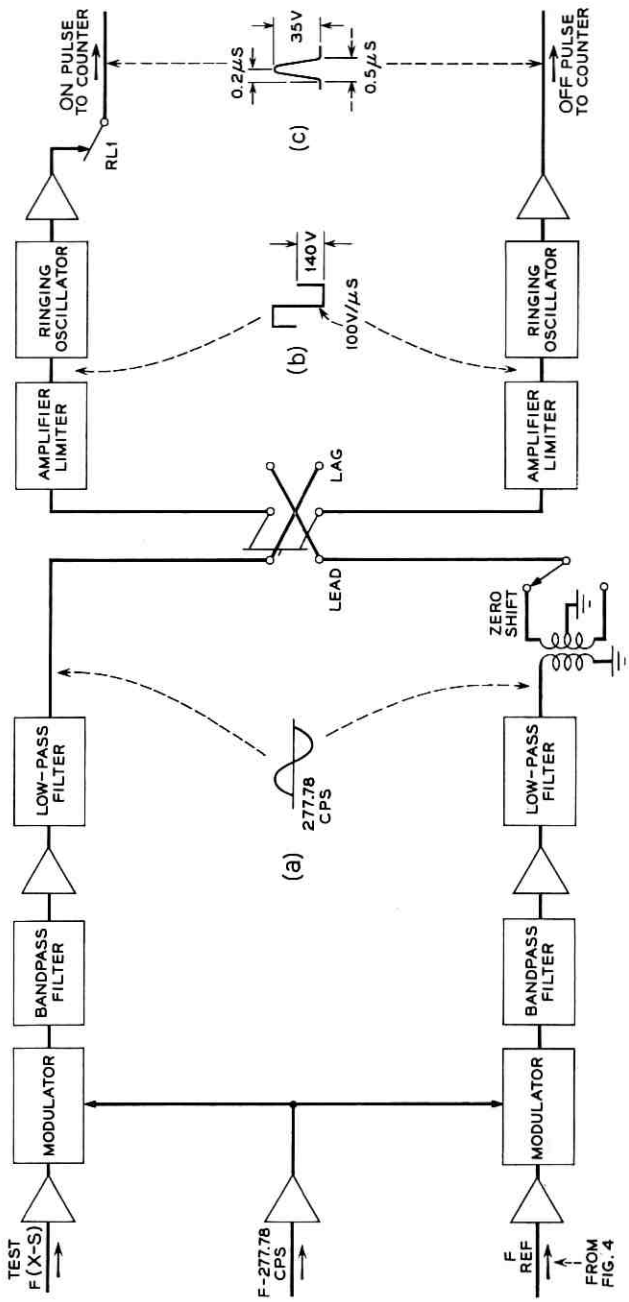


Fig. 7 — Block diagram of phase detector.

between those two channels. Since the time for 360 degree difference equals 1/277.78 second, then 1 microsecond = 0.1 degree and, at a 1 megacyle counting rate, 1 count = 0.1 degree.

Circuit zero errors, due to frequency-variable phase difference between the unknown and reference paths of Fig. 4, are eliminated by the switching of S1 of Fig. 4. The circuit compares the phases of the unknown and reference paths and then compares the phases of the attenuator and reference paths. Since the reference path is common to both comparisons, the resulting measurement is the difference in phase between the unknown and attenuator paths.

4.2 *Description of Detector Operation*

As discussed above, the detector must measure the time interval between negative-going zero crossings of two sine waves. The time of the zero crossings is established by the generation of corresponding sharp pulses at the instant of zero crossing and the time interval between pulses is measured by a reversible timer which is described in Section V.

The IF sine wave of each channel (a of Fig. 7) is clipped and amplified by three limiters in cascade which convert the signal to a fast-rise square wave (b of Fig. 7). The shock-excited triode oscillator is driven to cutoff by the fast negative-going square wave. Oscillations are produced in the tank circuit coil together with its distributed capacitance and a shunt diode damps out all but the first half sine wave. The resultant output pulses (c of Fig. 7) are 0.5 μs wide, 35 volts in magnitude and have a rise time of 0.2 μs. A relay in the ON pulse branch is controlled so that it closes approximately 0.2 second after the X-S switch of the comparison unit closes, thereby closing only after the switching transients have dissipated.

4.3 *Timing of Circuit Measuring Periods*

The direct reading of phase measurement is accomplished by the following sequence of automatic switching as illustrated by the diagram of Fig. 8:

1. The comparison switch S1A of Fig. 4 is switched to X and the timer is switched to "add" by switch S1B. No pulses reach the timer since relay RL1 of Fig. 7 is open.
2. An interval of 0.2 second is allowed for decay of transients introduced by closing S1A.
3. Relay RL1 is closed allowing transmission of pulses to the timer.
4. The timer measures the total time interval of 100 ON-OFF pulses

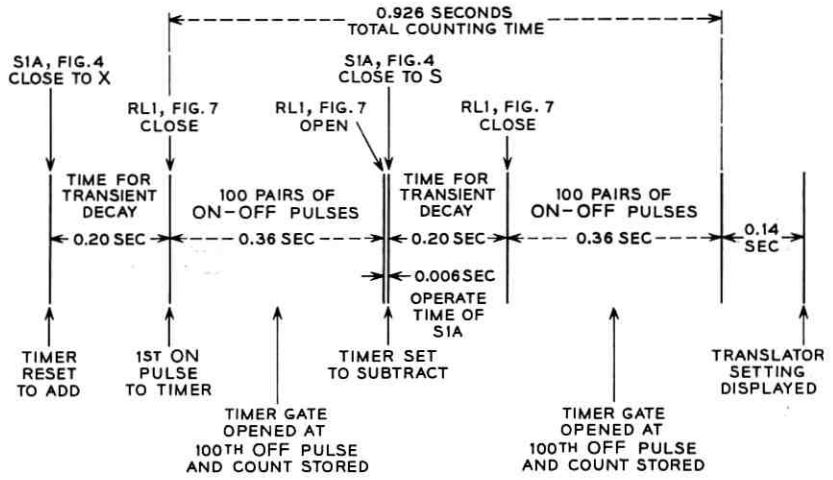


Fig. 8 — Time diagram of phase measurement circuit operation.

from the detector output. The timer gate opens at the 100th OFF pulse and the count is stored. The maximum time (and count) is 0.36 second corresponding to 360 degrees as discussed above.

5. Relay RL1 is again opened.

6. The comparison switch S1A of Fig. 4 is switched to S and the timer is switched to "subtract" by switch S1B.

7. An interval of 0.2 second is allowed for decay of transients.

8. Relay RL1 is closed.

9. The timer measures the total time interval of 100 ON-OFF pulses and subtracts it from the count stored in 4. The timer gate opens at the 100th OFF pulse and the remaining count is stored.

10. After 0.14 second, a relay in a digital translator is closed and the translator setting, corresponding to the stored count of the timer, is displayed by an in-line numerical indicator.

11. After an interval of 2 to 10 seconds, as desired, the entire procedure from 1 to 10 is repeated.

4.4 *Design Considerations and Choice of Constants*

To attain 0.01 degree accuracy of phase measurement, the choice of IF frequency (Δ) in the phase detector, counting rate, and the number of periods measured becomes highly critical because these factors are closely interrelated and also because they are subject to the performance of the SCHO oscillator.

As previously discussed, minimizing errors due to FM modulation

makes the design of the AFC loop III of the SCHO particularly difficult when Δ is small. However, since phase is to be measured by conversion to time, a large Δ requires a high counting frequency (or rate) F_R. Problems of circuit phase "zero," as previously described in Section 4.2, prescribe a reversible counter which is increasingly difficult to design for higher rates. To obtain direct reading phase measurements in decades from 0 to 360 degrees, F_R must be 10^n where n is an integer and $\Delta = F_R/360$. Based on the above considerations, the selected constants were: $F_R = 1.0$ mc and $\Delta = 277.78$ cps from which 1.0 degree = 10 microseconds. Measurement to 0.01 degree, under the ideal condition of no phase drift or phase jitter during the measurement time, requires the measurement of 10 periods. In practice there is some phase jitter, of the $F - \Delta$ signal in relation to the F signal, produced by the high gain of the AFC loop III of the SCHO which in turn is needed to control frequency over the prescribed range. To compensate for this jitter, 100 periods are measured and the last digit (0.001°) is not displayed. In effect, the displayed value of phase is an average of 100 counts for measurements of phase.

The IF bandwidth of the detector must be as narrow as possible to provide sufficient frequency discrimination and at the same time must be wide enough so that switching transients decay rapidly. The choice of 30 cps bandwidth proved adequate to achieve the aim of 0.01° phase accuracy and permits transients to decay within 0.1 second to a sufficiently low magnitude that accuracy is not affected. As previously described, 0.2 second in the switching time was allowed for transient decay which resulted in a total measurement time of 0.926 second. As previously stated, the phase lock of the SCHO is held to better than 0.01 degree shift in this total measurement time.

If the phase shift being measured is less than the phase shift through the constant phase attenuator, the reverse count will exceed the forward count and the difference will be negative. For a negative number n, however, the timer will indicate 1000.00-n: thus, $-15.37°$ is indicated as $1000.00 - 15.37°$ or $984.63°$. To avoid the necessity for arithmetic subtraction by the operator, the lead-lag switch (Fig. 7) has been provided. The switch reverses the S-X and reference pulse generator inputs to allow the count to appear as a positive number. The operator need only record the indicated count with a change of sign.

If, relative to the reference channel, the total phase shift, through standard or unknown path and through modulators and pulse generators, lies in the neighborhood of 0 or 360 degrees, the phase measuring circuit cannot operate properly. The difficulty is caused by the inability of the timer to respond to pulse pairs having a short time separation. In such

a case the phase circuit will either refuse to cycle or will give highly erratic results. Operation of the zero shift switch (Fig. 7) adds approximately 180 degrees in the reference path. Effectively, then, 180 degrees is added to both S and X measurements and the difference angle is unchanged.

As discussed in Section 2.3, sidebands of 277.78 cps on the signal applied to the phase detector may cause phase measurement error. Similarly, pickup in the detector of 277.78 cps and crosstalk between the two channels in the detector can produce phase error. The components of each channel were mounted in separate shielded compartments and at each point of entry of a power lead into a compartment heavy bypass capacitors were installed. The lead-lag switch required complete shielding including shields between each switch section, grounding of interconnecting wires in their alternate position and the use of shielded cables between the switch and the circuit connecting points.

V. REVERSIBLE TIMER

A block diagram of the reversible timer is shown in Fig. 4. The timer[5] measures the time difference between interval X, the time separation of input pulse pairs when the unknown network is in the measurement circuit, and interval S, the corresponding interval when the phase attenuator is in the circuit.

Input ON and OFF pulses, appearing on the separate input leads from the phase detector (Fig. 7), are applied to an electronic switch whose output is a rectangular pedestal voltage with leading and trailing edges coincident with the leading edges of the ON and OFF pulses. The pedestal voltage is applied to a gate which opens during the pedestal interval and transmits 1 mc voltage to a group of six tandem reversible decades. Reversal of the direction of count is accomplished by the use of interstage gates within the decade counting units actuated from the S-X switch S1B of Fig. 4. An auxiliary $\div$ 100 counter controls the electronic switch so that exactly 100 time intervals are measured for each direction of count.

The above six counting decades, measuring for 100 periods, provide a count to 0.001 degree phase. As described in Section 4.4, this last decade may jitter at random so that the output of only the first five decades are translated to decimal form and transmitted to an in-line readout display.

VI. COMPARISON UNIT

The comparison unit is designed around the parallel comparison type circuit shown in Figs. 3 and 4. The method of detecting the equality in

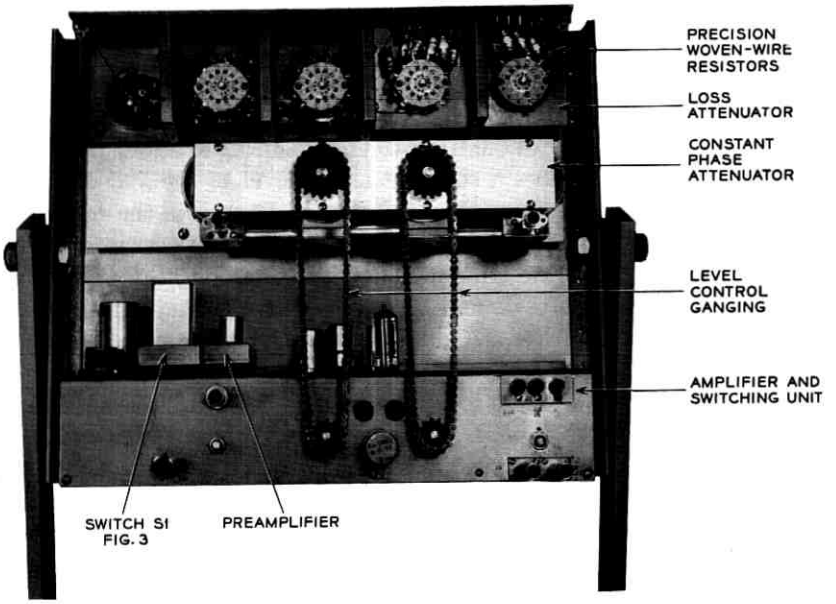


Fig. 9 — Comparison unit, showing the ganged level control.

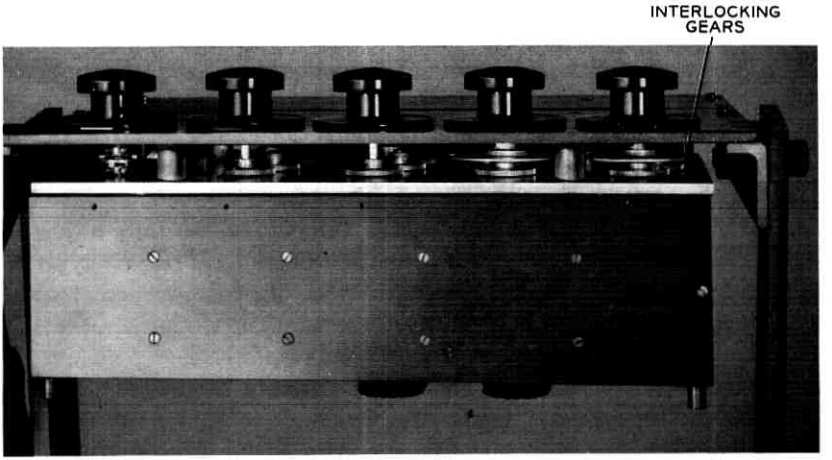


Fig. 10 — Comparison unit, showing ganging of the loss and constant phase attenuators.

magnitude and the difference in phase of the signals from the two paths was discussed in Sections III and IV. The construction used to attain the circuit objectives is shown in Figs. 9 and 10 and some of the special features are described below.

A constant sensitivity of 0.001 db when measuring losses in the range from 0 to 70 db prescribes a constant signal level to the loss detector. This is achieved by the use of step switch gain controls in the wideband amplifier, ganged by chain and sprocket to the attenuator switch shafts, as shown in Fig. 9. As the unknown loss and, therefore, the attenuator loss increases in steps of 1 db and 10 db, the gain of the amplifier is correspondingly increased over a 70 db range, thereby maintaining the level to the detector constant within 1 db.

Measurement of 40 db loss to a sensitivity of 0.001 db corresponds to 0.44 microvolt difference of the signal input to the amplifier and prescribes special precautions to minimize noise and pickup in the amplifier. Power frequency pickup is minimized by operating all heaters on filtered dc and by locating power supplies well away from areas sensitive to 60 cycle field. Microphonics are reduced by shock mounting both the mercury relay S-X switch (S1 of Fig. 3) and the input preamplifier. Thorough shielding and decoupling reduce high frequency crosstalk sufficiently to attain the desired measurement accuracy.

To attain the circuit objectives, the standard attenuator must provide insertion loss in the range of 2 to 122 db adjustable to within 0.001 db for loss measurements and to within 0.1 db for phase measurements. The insertion loss must meet accuracies to nominal of ± 0.001 to 42 db; ± 0.01 to 72 db and ± 0.5 to 122 db at frequencies from 10 to 300 kc. In addition, the insertion phase shift at any setting in the range of 2 to 42 db must be equal to the insertion phase shift at any other setting in that range to within 0.01 degree at any frequency from 10 to 300 kc. The loss accuracies are met with an attenuator embodying precisely adjusted, stable wire-wound resistors; however, the phase shift of this type attenuator varies with setting by as much as 0.5 degree. A second attenuator is used embodying deposited carbon resistors in a semicoaxial structure. Adjustment of internal capacitor trimmers for each step attains the required constant phase shift. The two attenuators are ganged, as shown in Fig. 10, so that the equality in magnitude of the signals from the two circuit arms attained in the loss measurement is held to within 0.1 db for phase measurement and the signal level from either arm to the phase detector is constant to within 1 db for loss of the unknown from 0 to 70 db.

Additional interlocking of the shafts of the 0 to 6 $\times$ 10 db and 0 to

5 × 10 db switches (2 right-hand knobs in Fig. 10) is required to avoid more complicated ganging to the amplifier gain controls. The interlock prevents moving the right-hand knob until the second switch is in the 60 db position, thereby maintaining constant detector input level for losses up to 70 db. Additional amplifier gain for the higher losses added by the right-hand switch is not needed since sensitivity requirements are comparably relaxed.

VII. ACCURACY OF MEASUREMENTS

7.1 *Loss*

The accuracy of loss measurements was verified by measuring a precision 75 ohm attenuator of 0.001 db accuracy. The accuracy of this checking standard was established by first calibrating at dc by measurement of voltage ratios with a precision potentiometer and then measuring the loss change with frequency up to 20 mc using a circuit capable of 0.01 db accuracy. Since for any value up to 40 db the loss change was less than 0.1 db from dc to 20 mc and was linear within the measurement accuracy, by interpolation the loss change of the checking standard should be less than 0.001 db from dc to 200 kc. A typical set of measurements of the checking standard using the 300 kc set is given in Table I.

As shown by the table, the design objectives were met for the measurement set of 0 to 40 db ±0.002 and 40 to 60 db ±0.02.

7.2 *Phase*

The accuracy of phase measurements was verified by comparing phase shift of coaxial cable as measured on the set with values computed from impedance measurements made on a precision bridge.

Western Electric 724 coaxial cable was found to have the necessary

TABLE I

Attenuator Setting (db)	Measured Loss (db) at				
	10 kc	35 kc	105 kc	200 kc	300 kc
14	14.001	14.000	14.000	13.999	13.999
24	23.998	23.998	23.998	23.998	23.998
34	34.000	33.999	33.998	33.998	33.998
44	43.999	43.999	43.998	43.997	43.997
54	54.001	54.001	54.000	53.998	53.996
64	64.002	64.001	63.999	63.993	63.991

time and temperature stability. Open circuit and short circuit impedances were measured on a 1200 foot length of cable. From the bridge measurements input and output image impedances were computed using:

$$Z_i = \sqrt{Z_{oc} Z_{sc}} .$$

Image impedance rather than characteristic or iterative impedance was used since the cable is not exactly symmetrical.

Insertion loss and phase shift between image impedance terminations were computed as follows:

$$\tanh \theta = \sqrt{Z_{sc} Z_{oc}} = U + jV$$

then the insertion loss is

$$\alpha = \tfrac{1}{2} \tanh^{-1} \left[\frac{2U}{1 + U^2 + V^2} \right]$$

and the phase is

$$\beta = \tfrac{1}{2} \tanh^{-1} \left[\frac{2V}{1 + U^2 + V^2} \right].$$

To avoid mismatch effects between cable and the measurement set, L type matching pads were constructed for each calibration frequency. The cable plus pads was measured and then the pads only with the cable removed. The difference in measured values then was recorded as the value of the cable.

A comparison of some measured and computed values is given in Table II.

Some errors inherent in bridge measurement and in constructing impedance matching pads may contribute as much as 0.01 degree of the above 0.03 degree maximum differences. Allowing for this uncertainty, the measurement accuracy of the circuit is considered to be ±0.02 degree over the frequency range of 70 to 150 kc. Since bridge measurement accuracies at frequencies above 150 kc do not provide sufficient accuracy of calculated phase, further checks were made in the following manner.

TABLE II

Frequency (kc)	Computed (Degrees)	Measured (Degrees)	Difference (Degrees)
70	58.94	58.91	−0.03
90	75.40	75.41	+0.01
120	99.92	99.91	−0.01
130	108.10	108.10	0.00
140	116.26	116.25	−0.01
150	124.30	124.33	+0.03

Four lengths of cable totaling 897 feet were prepared with coaxial fittings. Five impedance isolating pads were prepared each with coaxial fittings such that the cables could be inserted directly between the pads without additional fittings. Measurements of loss and phase were made of the four cable sections in tandem isolated from each other and from the set by pads and then only the pads in tandem were measured. The difference then is the total loss and phase of the four cable sections.

Each individual cable section was measured when terminated with the same pads that terminated the specific section when the tandem measurement was made. Again the pads only were measured, the difference giving the loss and phase of the individual sections.

Values of the measured sum compared to the calculated sum are given in Table III.

The above differences are derived from the individual measurements as outlined above and there is an uncertainty in each measurement of 0.001 db and 0.01 degree which may add in a random manner to 0.003 db and 0.03 degree. Since the above calculated differences are no greater than the random errors, the set capability for measuring a single value is considered to be ± 0.002 db and ± 0.02 degree.

VIII. CONCLUSION

The design effort has been directed toward achieving laboratory precision in measurement and, at the same time, maintaining the ease of use and speed necessary for a large volume of measurements. The entire frequency range of the set described in this article is covered without band switching by the use of a heterodyne signal oscillator which provides good frequency accuracy by locking to a frequency standard. The need for manual detector tuning is eliminated through the use of frequency conversion employing a locked slave (or offset) frequency source to provide a constant detection frequency. The desired test frequency can be set and read directly from the markings of a film scale type master oscillator and the dials of an interpolation oscillator. The insertion loss and

TABLE III

	150 kc		300 kc	
	Loss	Phase	Loss	Phase
Measured	0.872	79.46	1.179	156.67
Calculated	0.875	79.44	1.178	156.64
Difference	0.003 db	0.02 deg.	0.001 db	0.03 deg.

phase of an unknown may be read directly from the dials of an attenuator and from an in-line digital readout, respectively, and since the system zero is independent of test frequency, no circuit zero is required.

IX. ACKNOWLEDGMENTS

Particular acknowledgment is due Mr. G. V. Hill for much of the initial design and to Mr. M. H. Isbansky for producing a working measuring set by his most careful clearing of circuit errors. Acknowledgment is also made to Messrs. J. O. Israel, T. Slonszewski and C. H. Young for contributions to the design, and to Messrs. S. Doba and S. J. Zammataro for their assistance in the preparation of this article.

APPENDIX

Limitations of Frequency Modulation in SCHO Oscillator

The introduction of errors in phase measurement by frequency modulation of the F and $F - \Delta$ frequencies of Fig. 6 was discussed in Section 2.3. The derivation of the limitation of the amplitude of a spurious modulating signal is as follows:

The output of a frequency-modulated oscillator of very low modulation index (m_f) contains only one pair of sidebands and is well known (see Ref. 6 and similar texts) to be expressed as

$$e = E_0\left[J_0\left(\frac{\Delta f}{f_m}\right)\sin \omega t + J_1\left(\frac{\Delta f}{f_m}\right)[\sin(\omega + p)t - \sin(\omega - p)t]\right] \quad (1)$$

where

Δf = frequency deviation

f_m = modulating frequency

$J_n\left(\frac{\Delta f}{f_m}\right)$ = Bessel function of the first kind, order n.

For $\frac{\Delta f}{f_m}$ near zero (1) reduces to

$$e = E_0\left(\sin \omega t + \frac{\Delta f}{2f_m}[\sin(\omega + p)t - \sin(\omega - p)t]\right)$$

The maximum phase modulation is then:

$$\Delta\psi = 2f_m/\Delta f$$

but

$$\Delta f = kv_m$$

where

k = reactance tube oscillator sensitivity in cps per volt

$= 2.3 \times 10^3$ cps/volt for $F5$ (Fig. 6)

v_m = amplitude of modulating voltage

thus

$$\Delta\psi = \frac{f_m}{1.2 \times 10^3 v_m}.$$

For a phase error of 0.01 degree, $\Delta\psi = 0.01 \times (\pi/180)$ radians. Hence, the ratio of carrier to sideband amplitude must be greater than

$$\frac{1}{\Delta\psi} = 5.7 \times 10^3 \quad \text{or} \quad 75 \text{ db}$$

and for $f_m = 278$ cps, the maximum allowable modulating voltage is

$$v_m = \frac{278}{1.2 \times 10^3 \times 5.7 \times 10^3} = 40 \text{ microvolts.}$$

REFERENCES

1. Alsberg, D. A. and Leed, D., A Precise Direct Reading Phase and Transmission Measuring System for Video Frequencies, B.S.T.J., **28,** April, 1949, pp. 221–238.
2. Slonczewski, T., Precise Measurement of Repeater Transmission, Elec. Engr., **73,** April, 1954, pp. 346–347.
3. Slonczewski, T., Measuring Apparatus, United States Patent 2,666,100.
4. Israel, J. O., United States Patent 2,987,680.
5. Barney, K. H., The Binary Quantizer, Elec. Engr., **68,** November, 1949, pp. 962–967.
6. Terman, F. E., *Radio Engineering,* 3rd Ed., McGraw-Hill, New York, 1947.

Thermoelastic Stresses in Balanced and Unbalanced Seals

By T. D. RINEY and J. W. ELEK

(Manuscript received April 28, 1961)

This paper presents the results of a two-part analytical study of the stresses produced when ceramic cylinders are butt sealed to metal washer plates. Such stacked structures are of increasing importance in the fabrication of encapsulations for electron tubes and semiconductor devices.

Practical experience, rather than analysis, has shown the advantage of a balanced seal and it is now in common usage in the electronic industry. In some applications other requirements may prevent direct back-to-back balancing and the question arises as to how much the diameters of the opposing seals may differ and still give the desired balanced effect. In Sections 3.1, 4.1, and 5.1 this question is considered and the concept of a balanced seal is placed on a firmer basis. The results indicate that a small difference in the radii is to be preferred.

The effect of the length of an intermediate cylinder on the stresses in a three-cylinder stacked structure is considered in Sections 3.2, 4.2, and 5.2.

I. INTRODUCTION

In the electronic industry ceramic-to-metal seals are of increasing importance in the fabrication of encapsulations for electron tubes and semiconductor devices. The most common method for sealing ceramics to metals is to first "metallize" the ceramic by sintering onto it a thin coating of metallic powder; common brazing materials such as gold-copper alloys, silver-copper alloys, etc. are then used to fasten the metallized ceramic to the metal.

Temperatures in excess of 500°C are required for the brazing operation. As the assembly cools from the set-point of the brazing material to room temperature, thermoelastic stresses are induced in the bonded components by the differences in their thermal contraction rates. These stresses must be controlled if physical distortion and the resultant possibility of bond failure or fracture of the metal or ceramic are to be avoided.

Their magnitude depends on the choice of metal, ceramic, and brazing material as well as the applied heat treatment. If it is assumed that these parameters are held fixed, the magnitude of the thermoelastic stresses will still vary widely as a function of the structure geometry.

This paper presents the results of a two-part analytical study of the stresses produced when ceramic cylinders are butt sealed to metal washer plates. The problems considered are represented by the electron tube envelope assembly depicted in Fig. 1. At position A the top ceramic cylinder is used to balance the effect of the ceramic-metal seal on the other side; it serves no other purpose. Practical experience, rather than analysis, has shown the advantages of a balanced seal and it is now in common usage in the industry. In some applications other requirements may prevent direct back-to-back balancing (as at position C) and the question arises as to how much the diameters of the opposing seals may differ and still give the balanced effect. In the first part of the paper this question is considered and the concept of a balanced seal is placed on a firmer basis.

At position B the question arises as to the effect that the length of the middle cylinder has on the stresses produced. This problem is treated

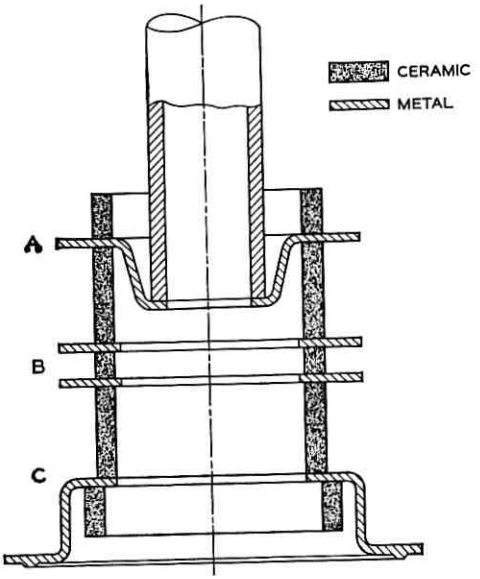


Fig. 1 — A cross-sectional view of a representative ceramic-metal electron tube envelope.

in Sections 3.2, 4.2, and 5.2. A similar study of the effects of the intermediate lengths in a graded cylindrical butt seal has been reported.[1–4] Another related problem is a long pipe reinforced by rings and submitted to the action of uniform internal pressure.[5] The thread that holds together these problems and the problem considered in Sections 3.1, 4.1, and 5.1 is that the stresses are produced by bonding together parts that would ordinarily be discontinuous at the joints under the given conditions. The continuity is produced by distributed forces and moments acting on each component along the joint.

II. ASSUMPTIONS AND NOTATIONS

Discontinuity stresses are highly localized; they are attenuated rapidly as the distance from the joint increases. Certain of the dimensions of the structure may therefore be idealized as infinite in extent without significantly affecting the stress distribution. The calculations are thereby greatly simplified. The idealized configurations for the two parts of the study are shown in Fig. 2.

The metal and ceramic bodies are imagined to be sealed together at a temperature T_0, the temperature then lowered to room value, T. It is assumed that no external restraining forces act on the structure during the sealing and cooling. The distributed forces, Q_i, and bending

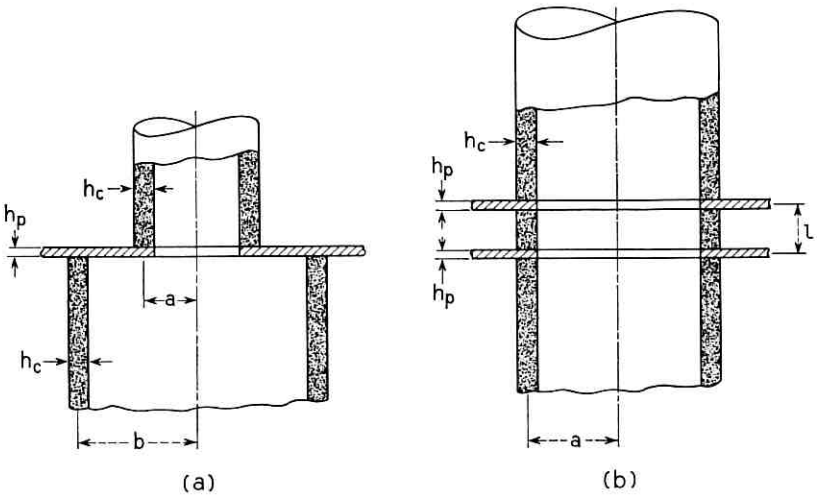


Fig. 2 — Cross-sectional view of the structures idealized to study: (a) the effects of a balanced seal; (b) the effect of varying l.

moments, M_i, required to maintain continuity at the joints of the ceramic-metal seals are depicted in the exploded diagrams of Fig. 3. The positive direction for the deflections and displacements are also shown.

To evaluate the edge loadings it is expedient to refer to the literature for the appropriate expressions for the deflections, displacements and slopes. This will be done freely. The formulas are based on the membrane theory of cylinders and plates, i.e., diameter to thickness ratio $\geqq 5$. It is also assumed that the deflections are small enough that the radial loads on the plates have negligible bending effects. In using these expressions the following notations will be employed:

$$D_c = \frac{E_c h_c^3}{12(1-\nu_c^2)} \qquad D_p = \frac{E_p h_p^3}{12(1-\nu_p^2)} \qquad R = \frac{E_c}{E_p}$$
$$\beta_1^4 = \frac{E_c h_c}{4a^2 D_c} \qquad \beta_2^4 = \frac{E_c h_c}{4b^2 D_c} \qquad \Delta = E_c(\alpha_c - \alpha_p)(T - T_0) \tag{1}$$

where E is the elastic modulus; ν Poisson's ratio, and the subscripts c and p indicate cylinder and plate, respectively.

Once the edge loadings have been calculated the results will then be used to evaluate the stresses that they produce in the ceramic cylinders. The notations used for the stresses are illustrated in Fig. 4. In calculating

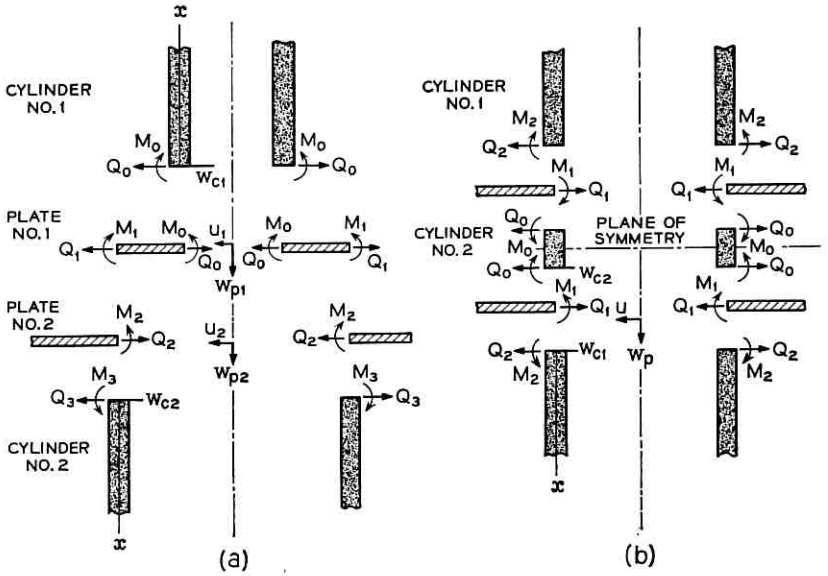


Fig. 3 — Exploded views showing the notations used for the moments and deflections.

the stress values the literature on cylindrical shell theory uses the following standard notation:

$$\varphi = e^{-\beta x}(\cos \beta x + \sin \beta x) \qquad \theta = e^{-\beta x} \cos \beta x$$
$$\psi = e^{-\beta x}(\cos \beta x - \sin \beta x) \qquad \zeta = e^{-\beta x} \sin \beta x \tag{2}$$

and

$$\chi_1(\beta l) = \frac{\cosh \beta l + \cos \beta l}{\sinh \beta l + \sin \beta l} \qquad \chi_2(\beta l) = \frac{\sinh \beta l - \sin \beta l}{\sinh \beta l + \sin \beta l}$$
$$\chi_3(\beta l) = \frac{\cosh \beta l - \cos \beta l}{\sinh \beta l + \sin \beta l}. \tag{3}$$

III. FORMULATION OF PROBLEMS

The problem considered in the first part of this and the following two sections is essentially the effect of radial unbalance in butt seals. The second is concerned with an axial unbalance resulting from the finite length of the intermediate cylinder. At this point it seems desirable to treat the problems separately.

3.1 *Radial Unbalance*

Static equilibrium at the joint $r = b$, Fig. 3(a), requires that

$$Q_2 = Q_1 + Q_3 \qquad M_2 = M_1 - M_3 \tag{4}$$

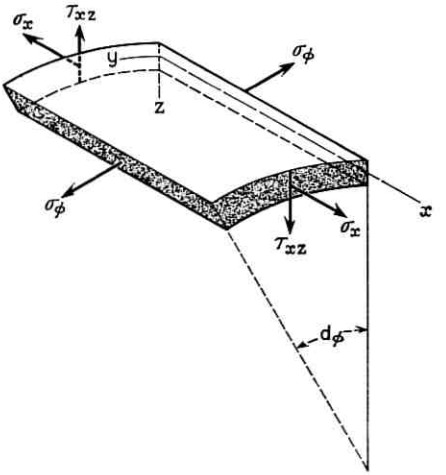


Fig. 4 — Cylindrical element showing stress directions.

and the quantities to be determined are reduced to six: Q_0, Q_1, Q_3, M_0, M_1, M_3. These will be evaluated by using the boundary conditions that at the juncture of the cylinders and plate the radial displacements are the same and the plate remains perpendicular to the cylinders. The following six conditions are obtained:

(a) The radial displacement of cylinder #1 and plate #1 are equal along the junction at $r = a$.
(b) The radial displacement of the two plates is the same at $r = b$.
(c) The radial displacement of cylinder #2 and plate #2 are equal along the junction at $r = b$.
(d) Cylinder #1 and plate #1 remain at right angles at their juncture.
(e) The slope of the two plates is the same at $r = b$.
(f) Cylinder #2 and plate #2 remain at right angles at their juncture.

The radial displacement of each part consists of a term due to the temperature change and a term due to the constraint forces resulting from the temperature change. Using the standard notations for displacements as depicted in Fig. 3(a), the first three conditions may be expressed symbolically as follows:

$$\begin{aligned} -(w_{c1})_{x=0} + a\alpha_c(T - T_0) &= (u_1)_{r=a} + a\alpha_p(T - T_0) \\ (u_1)_{r=b} + b\alpha_p(T - T_0) &= (u_2)_{r=b} + b\alpha_p(T - T_0) \\ -(w_{c2})_{x=0} + b\alpha_c(T - T_0) &= (u_2)_{r=b} + b\alpha_p(T - T_0). \end{aligned} \tag{5}$$

From the sketches in Fig. 5(a) it may be seen that the last three conditions yield the following relationships:

$$\begin{aligned} \gamma_a = -\left(\frac{dw_{c1}}{dx}\right)_{x=0} = \left(\frac{dw_{p1}}{dr}\right)_{r=a} \qquad -\gamma_b = \left(\frac{dw_{p1}}{dr}\right)_{r=b} = \left(\frac{dw_{p2}}{dr}\right)_{r=b} \\ -\gamma_b = \left(\frac{dw_{c2}}{dx}\right)_{x=0} = \left(\frac{dw_{p2}}{dr}\right)_{r=b}. \end{aligned} \tag{6}$$

It is now required to express the deflections and their derivatives appearing in (5) and (6) in terms of the unknown quantities Q_0, Q_1, $\cdots$, M_3. From Timoshenko (see Ref. 5, p. 393) it is found that

$$\begin{aligned} (w_{c1})_{x=0} = -\frac{1}{2\beta_1^3 D_c}[\beta_1 M_0 + Q_0] \quad (w_{c2})_{x=0} = -\frac{1}{2\beta_2^3 D_c}[\beta_2 M_3 + Q_3] \\ \left(\frac{dw_{c1}}{dx}\right)_{x=0} = \frac{1}{2\beta_1^2 D_c}[2\beta_1 M_0 + Q_0] \quad \left(\frac{dw_{c2}}{dx}\right)_{x=0} = \frac{1}{2\beta_2^2 D_c}[2\beta_2 M_3 + Q_3]. \end{aligned} \tag{7}$$

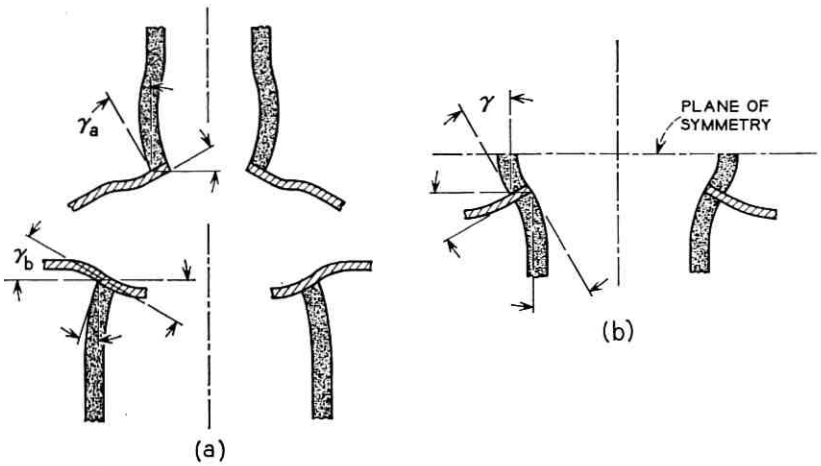


Fig. 5 — Exaggerated picture of deflections produced by discontinuity forces and moments acting at joints.

The forces Q_0 and Q_1, acting on plate #1 produce radial displacement and normal loading given by[6]

$$u_1 = Ar + \frac{B}{r} Q_{p1} = E_p h_p \left[\frac{Ar}{1 - \nu_p} - \frac{B}{1 + \nu_p} \frac{1}{r^2} \right].$$

Here the constants A and B may be determined from the conditions that Q_{p1} attains the value Q_0 and Q_1 at $r = a$ and $r = b$ respectively. If this is done and the values are substituted into the above expression for u_{p1}, one finds

$$\begin{aligned} (u_1)_{r=a} &= \frac{a}{E_p h_p (1 - a^2/b^2)} \{2Q_1 - Q_0[(1 - \nu_p)a^2/b^2 + (1 + \nu_p)]\} \\ (u_1)_{r=b} &= \frac{b}{E_p h_p (1 - a^2/b^2)} \{Q_1[(1 - \nu_p) + (1 + \nu_p)a^2/b^2] - 2Q_0 a^2/b^2\} \end{aligned} \tag{8}$$

If $a \to b$, $b \to \infty$, $Q_0 \to Q_2$, $Q_1 \to 0$, then the first of these gives

$$(u_2)_{r=b} = -\frac{b(1 + \nu_p)}{E_p h_p} Q_2 = -\frac{b(1 + \nu_p)}{E_p h_p} (Q_1 + Q_3). \tag{9}$$

The bending moments along the edges of plate #1 produce slopes (see Ref. 5, p. 63)

$$\left(\frac{dw_{p1}}{dr}\right)_{r=a} = \frac{a}{D_p(1 - a^2/b^2)} \cdot\left\{M_0\left[\frac{a^2/b^2}{1 + \nu_p} + \frac{1}{1 - \nu_p}\right] - M_1\left[\frac{1}{1 + \nu_p} + \frac{1}{1 - \nu_p}\right]\right\}$$

$$\left(\frac{dw_{p1}}{dr}\right)_{r=b} = \frac{b}{D_p(1 - a^2/b^2)} \cdot\left\{M_0\left[\frac{a^2/b^2}{1 + \nu_p} + \frac{a^2/b^2}{1 - \nu_p}\right] - M_1\left[\frac{1}{1 + \nu_p} + \frac{a^2/b^2}{1 - \nu_p}\right]\right\}. \quad (10)$$

If we let $a \to b$, $b \to \infty$, $M_0 \to M_2$, $M_1 \to 0$, then the first of these yields

$$\left(\frac{dw_{p2}}{dr}\right)_{r=b} = \frac{bM_2}{D_p(1 - \nu_p)} = \frac{b}{D_p(1 - \nu_p)}(M_1 - M_3). \quad (11)$$

The required six equations relating $Q_0, \cdots, M_3$ may now be obtained by substituting expressions (7) through (11) into conditions (5) and (6). The resulting equations are simplified and only negligible error introduced if one sets

$$\nu_p = \nu_c = \tfrac{1}{3}.$$

Then

$$D_c = \frac{3E_c h_c^3}{32}, \qquad D_p = \frac{3E_p h_p^3}{32}, \qquad \beta_1^4 = \frac{8}{3a^2 h_c^2}, \qquad \beta_2^4 = \frac{8}{3b^2 h_c^2}$$

and the six equations reduce to

$$\left[\frac{\sqrt{8/3}}{\beta_1 h_c^2} + \frac{R}{3h_p}\frac{2 + a^2/b^2}{1 - a^2/b^2}\right]\frac{Q_0}{\Delta} - \frac{R}{h_p(1 - a^2/b^2)}\frac{Q_1}{\Delta} + \frac{\sqrt{8/3}}{h_c^2}\frac{M_0}{\Delta} = -\tfrac{1}{2}$$

$$-\frac{a^2}{b^2}\frac{Q_0}{\Delta} + \frac{Q_1}{\Delta} + \frac{2}{3}(1 - a^2/b^2)\frac{Q_3}{\Delta} = 0$$

$$\frac{2}{3}\frac{R}{h_p}\frac{Q_1}{\Delta} + \left[\frac{2}{3}\frac{R}{h_p} + \frac{\sqrt{8/3}}{\beta_2 h_c^2}\right]\frac{Q_3}{\Delta} + \frac{\sqrt{8/3}}{h_c^2}\frac{M_3}{\Delta} = -\tfrac{1}{2}$$

$$\frac{2}{a\beta_1^2}\frac{Q_0}{\Delta} + \left[\frac{4}{a\beta_1} + \frac{3Rh_c^3}{h_p^3}\frac{2 + a^2/b^2}{1 - a^2/b^2}\right]\frac{M_0}{\Delta} - \frac{9Rh_c^3}{h_p^3(1 - a^2/b^2)}\frac{M_1}{\Delta} = 0 \quad (12)$$

$$\frac{a^2}{b^2}\frac{M_0}{\Delta} - \frac{M_1}{\Delta} + \frac{2}{3}(1 - a^2/b^2)\frac{M_3}{\Delta} = 0$$

$$\frac{1}{b\beta_2^2}\frac{Q_3}{\Delta} - \frac{3Rh_c^3}{h_p^3}\frac{M_1}{\Delta} + \left[\frac{2}{b\beta_2} + \frac{3Rh_c^3}{h_p^3}\right]\frac{M_3}{\Delta} = 0.$$

3.2 *Axial Unbalance*

Static equilibrium at the joint $r = a$, Fig. 3(b), requires that

$$Q_1 = Q_0 + Q_2 \qquad M_1 = M_0 - M_2 \tag{13}$$

and only the quantities Q_0, Q_2, M_0, M_2 remain to be determined. The four conditions to be satisfied are:

(*a*) The radial displacement of cylinder #1 and the plate are equal along the junction at $r = a$.

(*b*) The radial displacement of cylinder #2 and the plate are equal at $r = a$.

(*c*) Cylinder #1 and the plate remain at right angles along the junction.

(*d*) Cylinder #2 and the plate remain at right angles along the junction.

If we use the notations as defined in Fig. 3(b) the symbolic form of conditions (a) and (b) become

$$\begin{aligned} -(w_{c1})_{x=0} + a\alpha_c(T - T_0) &= (u_p)_{r=a} + a\alpha_p(T - T_0) \\ -(w_{c2})_{x=0} + a\alpha(T - T_0) &= (u_p)_{r=a} + a\alpha_p(T - T_0). \end{aligned} \tag{14}$$

From the sketch shown in Fig. 5(b) we see that the conditions (*c*) and (*d*) yield the relations

$$\gamma = -\left(\frac{dw_{c1}}{dx}\right)_{x=0} = \left(\frac{dw_p}{dr}\right)_{r=a} \qquad \gamma = \left(\frac{dw_{c2}}{dx}\right)_{x=0} = \left(\frac{dw_p}{dr}\right)_{r=a}. \tag{15}$$

The deflections and their derivatives may be written in terms of the unknown quantities Q_0, Q_2, M_0, M_2 as follows (see Ref. 5, p. 403):

$$\begin{aligned} (w_{c1})_{x=0} &= -\frac{2\beta a^2}{E_c h_c}[Q_0\chi_1(\beta l) + \beta M_0\chi_2(\beta l)] \\ \left(\frac{dw_{c1}}{dx}\right)_{x=0} &= \frac{2\beta^2 a^2}{E_c h_c}[Q_0\chi_2(\beta l) + 2\beta M_0\chi_3(\beta l)], \end{aligned} \tag{16}$$

where the notations in (3) are used.

From equations (9) and (11) it is found that in the infinite plate the following relations hold

$$\begin{aligned} (u)_{r=a} &= -\frac{a(1 + \nu_p)}{E_p h_p} \qquad Q_1 = -\frac{a(1 + \nu_p)}{E_p h_p}(Q_0 + Q_2) \\ \left(\frac{dw_p}{dr}\right)_{r=a} &= \frac{a}{D_p(1 - \nu_p)} \qquad M_1 = \frac{a}{D_p(1 - \nu_p)}(M_0 - M_2). \end{aligned} \tag{17}$$

Finally, the cylinder distortions are given by (see Ref. 5, p. 393)

$$\begin{aligned}(w_{c2})_{x=0} &= -\frac{1}{2\beta^3 D_c}[\beta M_2 + Q_2] \\ \left(\frac{dw_{c2}}{dx}\right)_{x=0} &= \frac{1}{2\beta^2 D_c}[2\beta M_2 + Q_2].\end{aligned} \tag{18}$$

The required four equations to determine the unknowns Q_0, Q_2, M_0, M_2 may now be found by substituting expressions (16), (17) and (18) into conditions (14) and (15). If we again set $\nu_p = \nu_c = \frac{1}{3}$, then

$$\beta^4 = \frac{8}{3a^2 h_c^2}$$

and the four equations reduce to

$$\begin{aligned}\left[\frac{a\beta}{h_c}\chi_1(\beta l) + \frac{2R}{3h_p}\right]\frac{Q_0}{\Delta} + \frac{2R}{3h_p}\frac{Q_2}{\Delta} + \frac{\sqrt{8/3}}{h_c^2}\chi_2(\beta l)\frac{M_0}{\Delta} &= -\tfrac{1}{2} \\ \frac{2R}{3h_p}\frac{Q_0}{\Delta} + \left[\frac{2R}{3h_p} + \frac{\sqrt{8/3}}{\beta h_c^2}\right]\frac{Q_2}{\Delta} + \frac{\sqrt{8/3}}{h_c^2}\frac{M_2}{\Delta} &= -\tfrac{1}{2} \\ \sqrt{8/3}\, h_c\chi_2(\beta l)\frac{Q_0}{\Delta} + \left[2\sqrt{8/3}\, h_c\beta\chi_3(\beta l) + 8R\frac{h_c^3}{h_p^3}\right]\frac{M_0}{\Delta} & \\ - 8R\frac{h_c^3}{h_p^3}\frac{M_2}{\Delta} &= 0 \\ \frac{2}{\sqrt{8/3}}h_c\frac{Q_2}{\Delta} - 6R\frac{h_c^3}{h_p^3}\frac{M_0}{\Delta} + \left[\frac{4}{a\beta} + 6R\frac{h_c^3}{h_p^3}\right]\frac{M_2}{\Delta} &= 0.\end{aligned} \tag{19}$$

IV. SOLUTION OF PROBLEMS

Once the edge loadings are known, the unit stresses in the ceramic cylinders may be calculated. Shell theory assumes that the normal stresses, σ_x and σ_φ, vary linearly through the thickness of the shell from a maximum value at one surface to a minimum value at the other surface. The shearing stress, $\tau = \tau_{xz}$, is assumed to vary across the thickness of the shell according to a parabolic law with a maximum at the center. The variation of these limiting values as a function of distance from the joint is given by (see Ref. 5, pp. 45 and 88)

$$(\sigma_x)_{z=\pm\frac{1}{2}h_c} = \pm\frac{6}{h_c^2} M_x$$

$$(\sigma_\varphi)_{z=\pm\frac{1}{2}h_c} = \left(\nu_c\sigma_x - \frac{E_c}{a} w_c\right)_{z=\pm\frac{1}{2}h_c} \tag{20}$$

$$\tau_{\max} = \frac{3}{2}\frac{Q_x}{h_c},$$

where Q_x, M_x and w_{cx} are the force, moment, and deflection in the cylinder at a distance x from the joint.

4.1 *Radial Unbalance*

Of interest here is the magnitude and distribution of the stresses produced in the cylinders as the radius of cylinder № 2 is increased (Fig. 3a). The cylinders are infinitely long; the expressions for the required quantities are (see Ref. 5, p. 393)

$$\begin{aligned} M_{xi} &= \frac{1}{\beta_i}\left[\beta_i M_0\varphi(\beta_i x) + Q_0\zeta(\beta_i x)\right] \\ w_{ci} &= -\frac{1}{2\beta_i^3 D_c}\left[\beta_i M_0\psi(\beta_i x) + Q_0\theta(\beta_i x)\right] \\ Q_{xi} &= -2\beta_i M_0\zeta(\beta_i x) + Q_0\psi(\beta_i x), \end{aligned} \tag{21}$$

for $i = 1, 2$. Here x is the axial displacement from the cylinder-plate joint.

4.2 *Axial Unbalance*

The effect that varying the length of the middle cylinder has on the stresses in both cylinders № 1 and № 2 is of interest (see Fig. 3b). Cylinder № 1 is infinitely long and, corresponding to (21), the required expressions are

$$\begin{aligned} M_{x1} &= \frac{1}{\beta}\left[\beta M_2\varphi(\beta x) + Q_2\zeta(\beta x)\right] \\ w_{c1} &= -\frac{1}{2\beta^3 D_c}\left[\beta M_2\psi(\beta x) + Q_2\theta(\beta x)\right] \\ Q_{x1} &= -2\beta M_2\zeta(\beta x) + Q_2\psi(\beta x). \end{aligned} \tag{22}$$

Cylinder № 2 is finite in length and the expressions for the required quantities are

$$M_{x2} = \frac{2}{\beta}\left[\Lambda \cos \beta\xi \cosh \beta\xi - \Omega \sin \beta\xi \sinh \beta\xi\right]$$

$$w_{c2} = -\frac{1}{D\beta^3}\left[\Lambda \sin \beta\xi \sinh \beta\xi + \Omega \cos \beta\xi \cosh \beta\xi\right]$$

$$Q_{x2} = -2(\Lambda + \Omega) \sin \beta\xi \cosh \beta\xi + 2(\Lambda - \Omega) \cos \beta\xi \sinh \beta\xi$$

where $\xi = x + l/2$ is the axial displacement measured from the center of the short cylinder and

$$\Lambda = \frac{Q_0 \sin \frac{\beta l}{2} \sinh \frac{\beta l}{2} + \beta M_0 \left(\sin \frac{\beta l}{2} \cosh \frac{\beta l}{2} + \cos \frac{\beta l}{2} \sinh \frac{\beta l}{2}\right)}{\sinh \beta l + \sin \beta l}$$

$$\Omega = \frac{Q_0 \cos \frac{\beta l}{2} \cosh \frac{\beta l}{2} + \beta M_0 \left(\cos \frac{\beta l}{2} \sinh \frac{\beta l}{2} - \sin \frac{\beta l}{2} \cosh \frac{\beta l}{2}\right)}{\sinh \beta l + \sin \beta l}$$

These results are obtained by superposing the solution for the case of bending by uniformly distributed shearing forces Q_0 and the solution for the case of bending by moments M_0 distributed at the ends (see Ref. 5, p. 403).

V. STRESS DISTRIBUTIONS AND MAGNITUDES

Calculations were carried out on an electronic computer to demonstrate the effects of the two types of unbalance. It was decided to use the value $R = 2.1$ for the ratio of elastic moduli corresponding to the Diamonite ceramic cylinders and the Kovar plates employed in the particular tube structure illustrated in Fig. 1. Here Δ is a positive quantity since $\alpha_p > \alpha_c$ and $T_0 > T$. The dimensions of the tube were also chosen to correspond to this practical geometry and the plate thickness was varied.

$$h_c = 0.04 \text{ in.}, \qquad a = 0.22 \text{ in.}, \qquad h_p = 0.01, 0.02, 0.04 \text{ in.}$$

Of interest are the stress distributions in the cylinders, especially the values attained at the seal junctions. These have been obtained for both radial and axial unbalanced seals for the three plate thicknesses.

5.1 *Radial Unbalance*

The stresses in the two cylinders at the cylinder-plate junctions are depicted in Fig. 6. The first graph shows the longitudinal stress at the outer wall of each cylinder; an equal tensile stress occurs at the inner

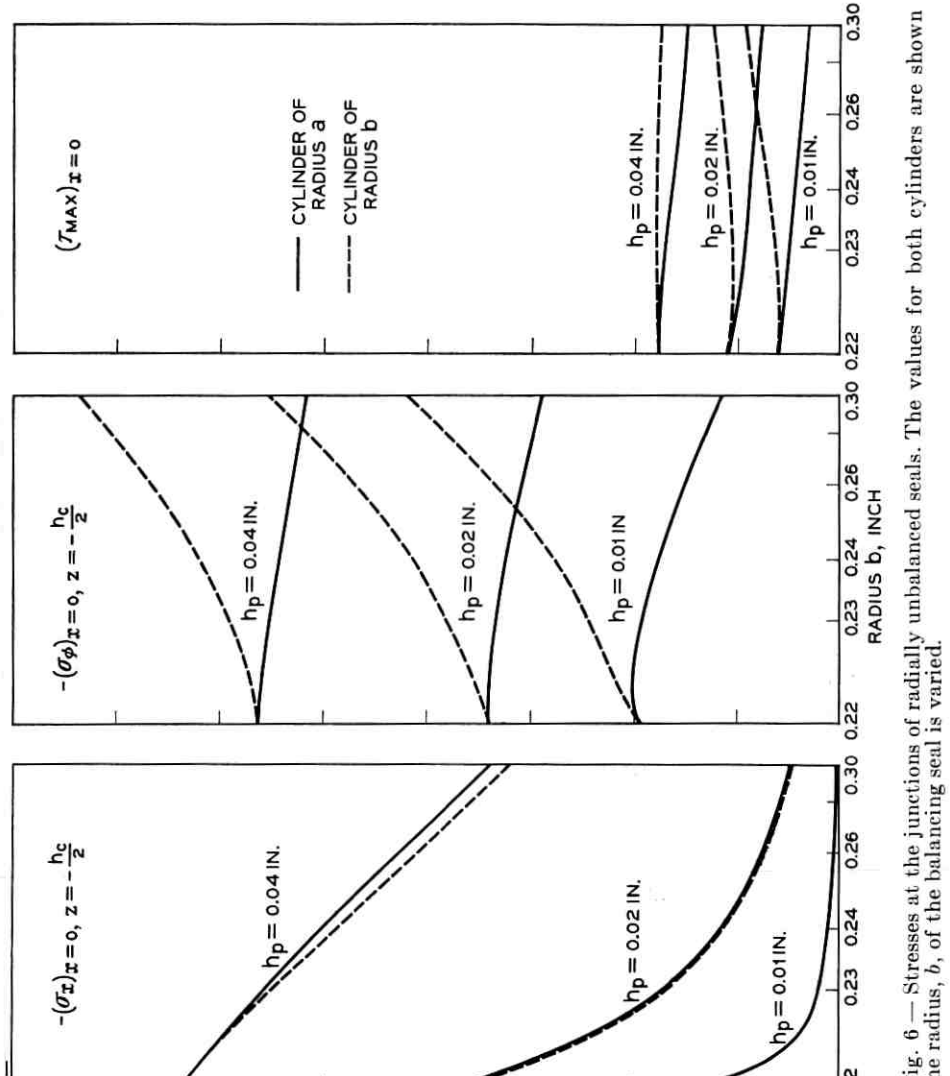


Fig. 6 — Stresses at the junctions of radially unbalanced seals. The values for both cylinders are shown as the radius, b, of the balancing seal is varied.

wall. The circumferential stresses are compressive all across the wall thickness; the maximum numerical values are attained at the outer wall and these are shown in the second graph of Fig. 6. The maximum shear stress acting on the plane of the cylinder-plate junctions is shown in the third graph.

It is seen that as the amount of the unbalance is increased, i.e., b is increased, the junction stresses in the smaller cylinder all decrease. In the larger cylinder the longitudinal stress at the junction decreases, but the circumferential and shear stresses increase. This trend actually continues until the radius of the larger cylinder is several times that of the smaller.

The variation of the stresses as a function of distance from the junction is shown in Fig. 7. It is of special interest to compare the results for completely balanced seal, Fig. 7(a), with those of a completely unbalanced seal, Fig. 7(c).

5.2 *Axial Unbalance*

Figure 8 shows that as the amount of axial unbalance is decreased, i.e., l is increased, the junction stresses in the long cylinder decrease and then increase again to a limiting value. The stresses in the short cylinder are practically the same as those in the long cylinder with the exception of the shearing stresses. The shearing stress in the short cylinder increases with decreasing unbalance to a limiting value. From the standpoint of both cylinders the optimum length for the intermediate or short cylinder appears to be about 0.125 in. The somewhat larger stresses corresponding to the perfectly balanced configuration are attained for l greater than approximately 0.3 in.

If l is allowed to tend to zero the configuration reduces to a perfectly balanced seal with a plate of twice the original thickness. The stresses for a given plate thickness at $l = 0$ must therefore equal the stresses produced in a structure using plates of twice the thickness at very large values of l, say $l = 0.4$. Figure 8 shows this very well.

Figure 9 illustrates the variations in the stress distributions for several lengths of the intermediate cylinder.

VI. DISCUSSION

Practical experience has shown the balanced seal to be stronger than the completely unbalanced seal (*i*) under routine handling conditions, and (*ii*) under thermal shock in subsequent welding operations. In the former case the unbalanced seal is more susceptible to a failure of the bond between the cylinder and plate; in the latter case the weakness is

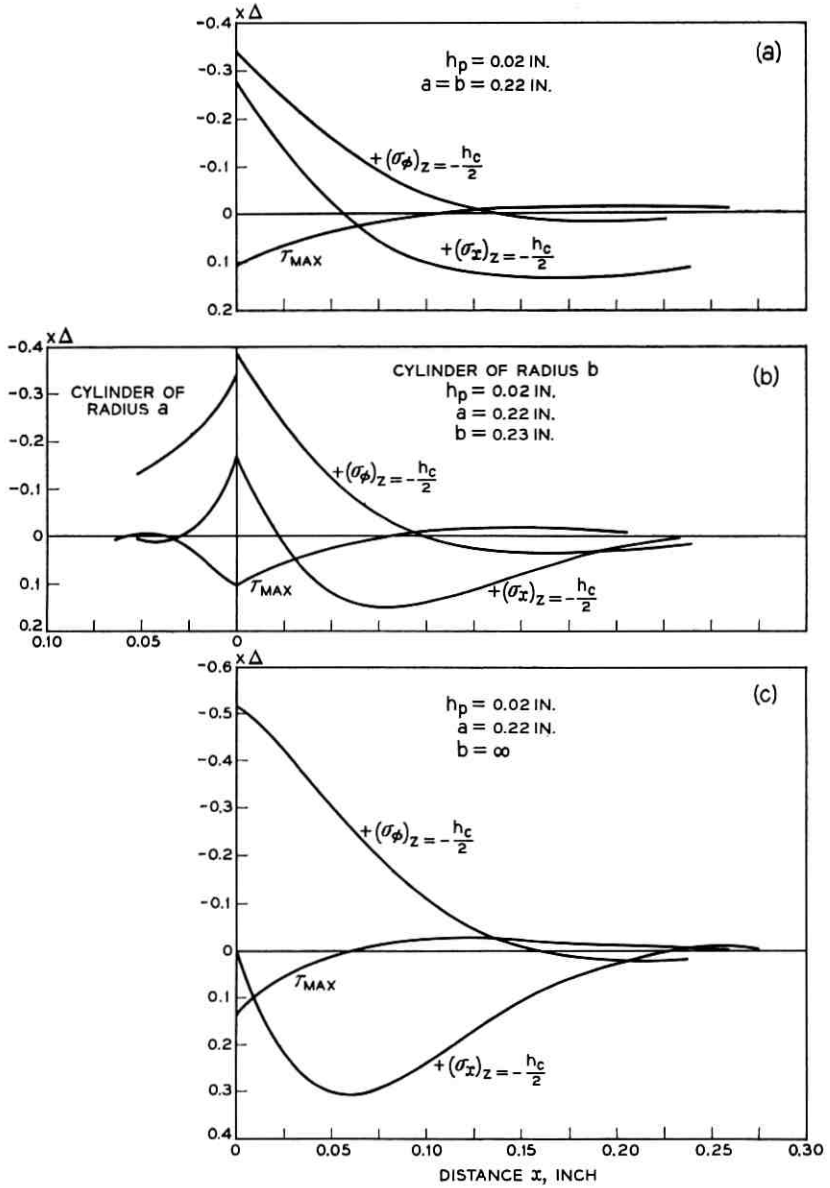


Fig. 7 — Stress distributions showing the effects of various conditions of radial unbalance; (a) completely balanced, (b) slightly unbalanced, (c) completely unbalanced seals.

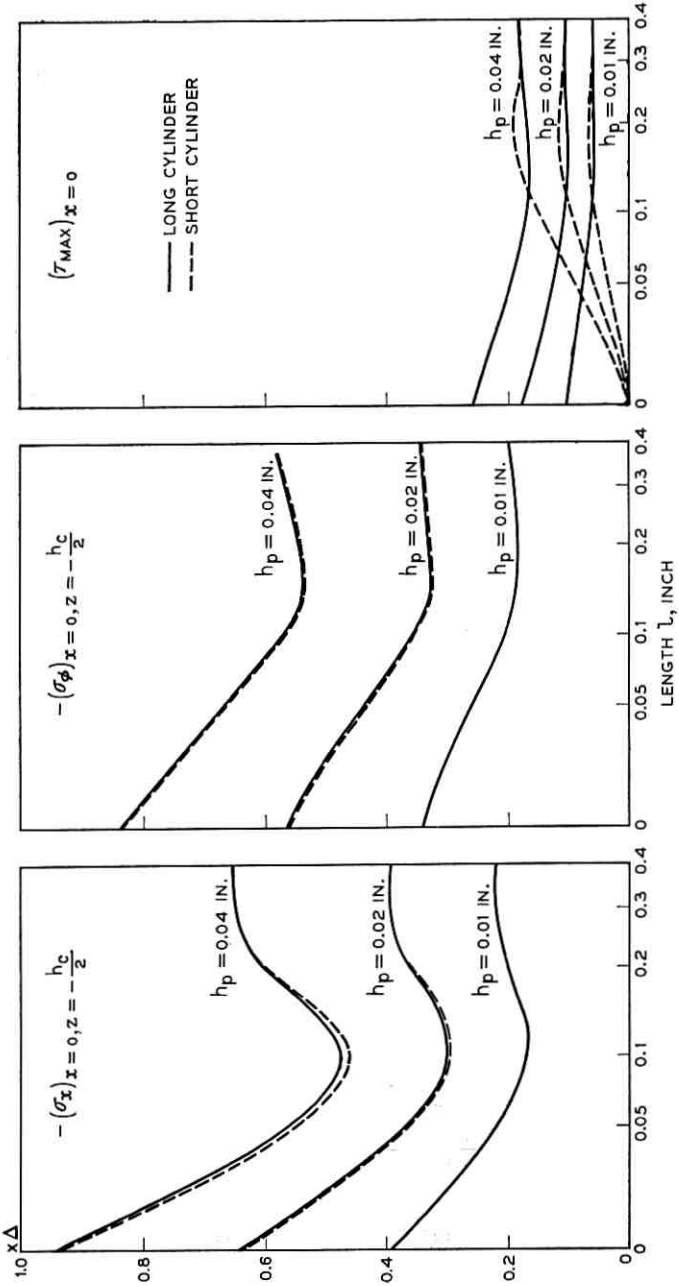


Fig. 8 — Stresses at the junctions of axially unbalanced seals. The values for both cylinders are shown as the length, l, of the intermedate cylinder is varied.

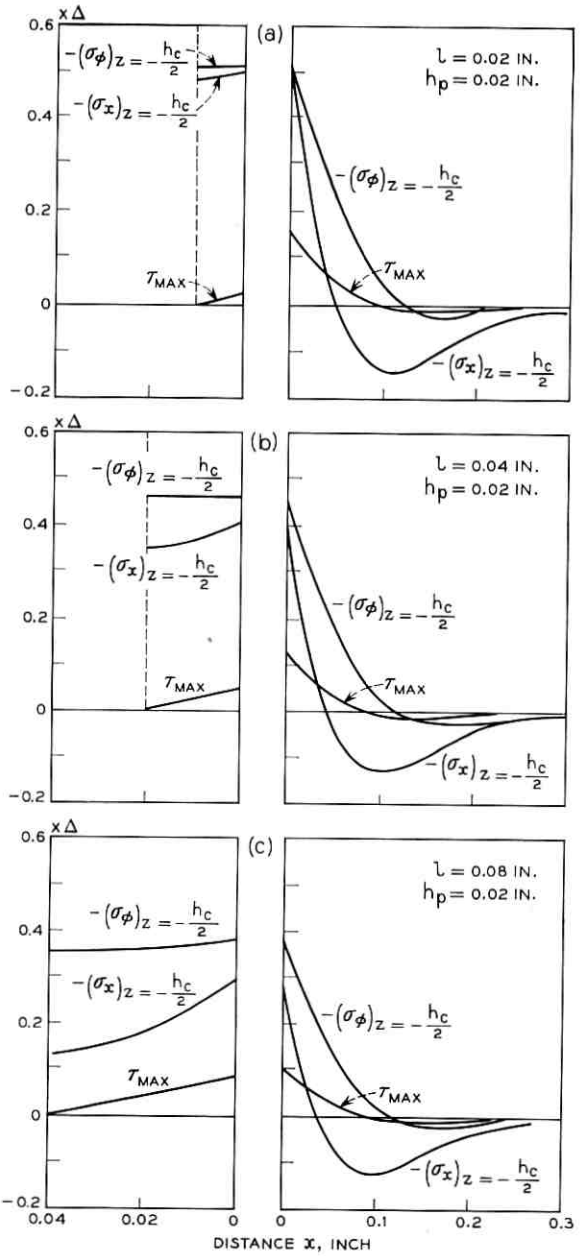


Fig. 9 — Stress distributions showing the effect of varying the length, l, of the intermediate cylinder in an axially unbalanced seal. Note the difference in scale for long (right side) and short (left side) cylinders.

manifested by longitudinal cracks in the unbalanced ceramic cylinder. It is of interest to interpret these facts in the light of the calculated stresses presented above.

Since type (*i*) failures clearly cannot result from compressive circumferential stresses and the longitudinal junction stress is essentially zero for the unbalanced seal, Fig. 7(c), the culprit must be the shear stress at the junction, i.e., the metallized interface fails in shear. Comparison of τ_{max} for the two structures, Fig. 7(a) with Fig. 7(c), shows that indeed its value is about 30 per cent greater for the unbalanced seal than for the perfectly balanced seal.

Type (*ii*) failures occur when the metal plate is raised to a higher temperature than the ceramic cylinder by uneven heating. This introduces tensile circumferential stress in the cylinder which results in longitudinal cracks if the temperature difference is great enough for the residual compressive stress to be sufficiently exceeded. A first approximation to the relative values of these transient stresses is obtained by merely reversing the sign of Δ in the calculations presented. Comparison of Fig. 7(a) and Fig. 7(c) then shows that a much larger circumferential tensile stress would be expected in the unbalanced seal under such a thermal gradient.

If these explanations are valid it may be concluded from Fig. 6 that the balancing cylinder (radius b) is even more effective for $b > a$. For, the magnitudes of the shear and circumferential stresses in the smaller (and more crucial) cylinder decrease as b is increased. The unbalance cannot be made too great, however, as the corresponding stresses in the larger cylinder increase.

In the case of axial unbalance it is clear (Fig. 8) that the intermediate cylinder must have length $l > 0.125$ if the presence of seal at the opposite end is to be negligible. If this inequality is satisfied the susceptibility to failure of either type (*i*) or type (*ii*) is greatly reduced.

VII. ACKNOWLEDGMENT

The authors are happy to acknowledge valuable discussion with W. A. Schlegel. Thanks are also due to Messrs. R. E. Caffrey and W. C. Lo for suggesting the study.

REFERENCES

1. Rawson, G., A Theory of Stresses in Glass Butt Seals, Brit. J. Appl. Phys., **2**, 1951, pp. 151–156.
2. Svenson, N. L., [Comments on Ref. 1], Brit. J. Appl. Phys., **3**, 1952, pp. 30–31.
3. Zaid, M., [Comments on Ref. 1], Brit. J. Appl. Phys., **3**, 1952, pp. 31–32.
4. Lewin, G., and Mark, R., Theory of Dissimilar Tubular Seals of Glass, Ceramics and Metals for Critical Applications, Trans. Fifth National Symposium on Vacuum Technique, 1958 pp. 44–49.
5. Timoshenko, S., *Theory of Plates and Shells*, McGraw-Hill, New York, 1940.
6. Timoshenko, S., and Goodier, J. H., *Theory of Elasticity*, McGraw-Hill, 1951.

The Tunnel Diode as a Linear Network Element

By I. W. SANDBERG

(Manuscript received April 3, 1962)

Theorems are proved which completely characterize in an explicit manner the class of immittance matrices realizable with lossless reciprocal elements and a tunnel diode represented by the three-parameter "LC, −R" model. Techniques are presented for the synthesis of any immittance matrix within this class.

Considered first, from a scattering matrix viewpoint, are the so-called degenerate cases in which the immittance matrices of the lossless network do not exist. Throughout the remainder of the discussion it is assumed that the lossless network possesses an immittance matrix. Necessary and sufficient conditions, involving in a complicated manner the existence of a certain strict Hurwitz polynomial, are derived for realization with a wide class of terminations. A study of the existence of this polynomial for the particular terminations of interest leads to explicit realizability conditions.

I. INTRODUCTION

The small signal "C, −R" model of the tunnel diode (Fig. 1) provides a fairly good representation over a wide range of frequencies, and is much simpler to use in a general study of network properties than the "LC, −R" model (Fig. 2) which includes, in addition, the series inductor. The simpler model has been used extensively by network theorists.[1-10]

The primary purpose of this paper is to define in an explicit manner the class of $n \times n$ open-circuit impedance and short-circuit admittance matrices that are realizable with lossless reciprocal elements and a tunnel diode characterized by the "LC, − R" model. The results constitute an extension of the theory presented in Ref. 10* for the "C, − R" model. The main interest in this problem to date relates to the

* Although the present paper is self-contained, some familiarity with the earlier work would be of assistance to the reader.

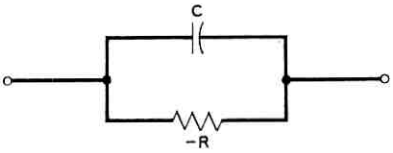


Fig. 1 — "C, −R" model of the tunnel diode.

special case $n = 2$. From a practical viewpoint our results anticipate the development of approximation techniques that lead to the specification of lossless-network tunnel-diode substructures which are to operate between prescribed sources and loads. Indeed an objective of this paper is to encourage research in this direction by presenting a complete solution to the realization problem.

The basic structure under consideration is shown in Fig. 3 in which the $(n + 1)$-port network is assumed to be a lossless reciprocal configuration containing inductors, capacitors, and ideal transformers. While we shall be particularly concerned with the case in which port $(n + 1)$ is terminated with the "LC, − R" model of the tunnel diode, many of the arguments to be presented are applicable to a much wider class of terminations. The overall network is assumed to possess either a short-circuit admittance matrix $\mathbf{Y}(s)$ or an open-circuit impedance matrix $\mathbf{Z}(s)$ relating currents and voltages at the ports $(1, 2, \cdots, n)$.

The realizability study is initiated in the following section where we discuss the cases in which the immittance matrices of the lossless network fail to exist. Throughout the remainder of the paper we consider the realizability of $\mathbf{Z}(s)$ and assume that the $(n + 1)$-port lossless network possesses an open-circuit impedance matrix $\hat{\mathbf{Z}}(s)$. This involves no loss of generality, of course, since results for the short-circuit admittance matrix $\mathbf{Y}(s)$ are identical with those for the open-circuit impedance matrix with the termination replaced with its reciprocal. In Section III necessary and sufficient conditions are presented, in terms of an unknown strict Hurwitz polynomial, for the realization of $\mathbf{Z}(s)$ with a wide class of terminations. The following sections utilize these results to obtain explicit realizability conditions for the particular termination of interest.

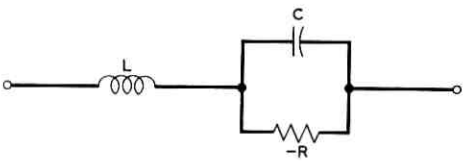


Fig. 2 — "LC, −R" model of the tunnel diode.

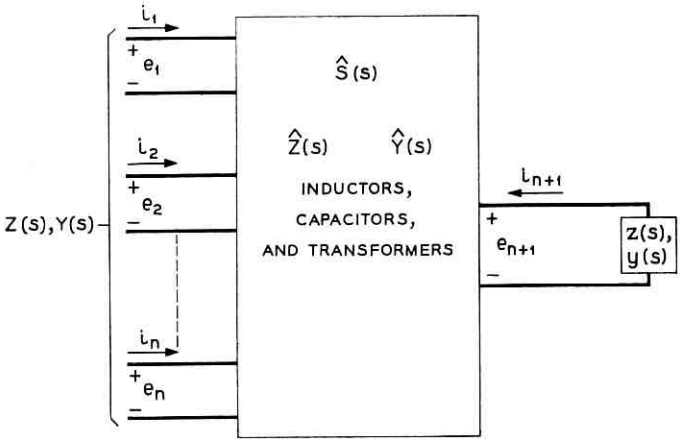


Fig. 3 — Most general structure defining $\mathbf{Z}(s)$ and $\mathbf{Y}(s)$.

In an interesting recent paper,[11] Schoeffler considers a problem similar to that discussed in Section III for the special case $n = 1$ under an assumption equivalent to supposing that the unknown polynomial is unity. In order to obtain explicit conditions, he further assumes that both $\mathbf{Z}(s)$ and the termination are regular at infinity and that $\mathbf{Z}(s)$ has no singularities on the entire $j\omega$-axis. Of course, for our purposes, these assumptions cannot be made. Indeed, for the particular problem considered here, the most interesting realizability conditions arise from a possible pole at infinity of the termination and from the influence of the unknown strict Hurwitz polynomial.

II. Realizability Conditions When the Immittance Matrix of the Lossless Network Does Not Exist

The $(n + 1)$-port lossless network in Fig. 3 invariably possesses a symmetric regular para-unitary scattering matrix[12] which we shall denote by $\hat{\mathbf{S}}(s)$. However, the corresponding short-circuit admittance matrix $\hat{\mathbf{Y}}(s)$ exists if and only if* $\det[\mathbf{1}_{n+1} + \hat{\mathbf{S}}(s)]$ does not vanish identically in s. Similarly $\hat{\mathbf{Z}}(s)$ exists if and only if $\det[\mathbf{1}_{n+1} - \hat{\mathbf{S}}(s)]$ does not vanish identically in s. In this section the following theorem is proved which completely characterizes $\mathbf{Y}$ or $\mathbf{Z}$ in the event that $\hat{\mathbf{Y}}$ or $\hat{\mathbf{Z}}$ fails to exist.

Theorem 1: If $\mathbf{Y}$ $[\mathbf{Z}]$ *in Fig. 3 exists with port* $(n + 1)$ *terminated with an admittance* y $[$*impedance* $z]$ *but* $\hat{\mathbf{Y}}$ $[\hat{\mathbf{Z}}]$ *does not exist,* $\mathbf{Y} = y\mathbf{C} + \mathbf{Y}'$ $[\mathbf{Z} =$

* The identity matrix of order $(n + 1)$ is denoted by $\mathbf{1}_{n+1}$.

$z\mathbf{C} + \mathbf{Z}'$] *where* $\mathbf{C}$ *is a nonnegative definite symmetric real matrix of constants of rank not exceeding unity and* $\mathbf{Y}'$ [$\mathbf{Z}'$] *is the short-circuit admittance matrix* [*open-circuit impedance matrix*] *of a lossless reciprocal network.*

The proof is based on the following lemma which is adapted from a result of Youla et al.[12]

Lemma 1: Let $\hat{\mathbf{S}}(s)$ *be a regular symmetric para-unitary scattering matrix of order* $(n + 1)$ *such that the normal ranks of* $[\mathbf{1}_{n+1} + \hat{\mathbf{S}}(s)]$ *and* $[\mathbf{1}_{n+1} - \hat{\mathbf{S}}(s)]$ *are* r *and* r' *respectively. Then there exist two orthogonal constant matrices* $\mathbf{T}$ *and* $\mathbf{T}'$ *such that**

$$\mathbf{T}^t\,\hat{\mathbf{S}}(s)\,\mathbf{T} = \begin{bmatrix} -\mathbf{1}_{n+1-r} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{S}}_r \end{bmatrix} \tag{1}$$

$$\mathbf{T}'^t\,\hat{\mathbf{S}}(s)\,\mathbf{T}' = \begin{bmatrix} \mathbf{1}_{n+1-r'} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{S}}_{r'} \end{bmatrix}. \tag{2}$$

where $\hat{\mathbf{S}}_r$ *and* $\hat{\mathbf{S}}_{r'}$ *are symmetric regular para-unitary scattering matrices of orders* r *and* r' *respectively. Moreover* $\det[\mathbf{1}_r + \hat{\mathbf{S}}_r]$ *and* $\det[\mathbf{1}_{r'} - \hat{\mathbf{S}}_{r'}]$ *do not vanish identically in* s.

Suppose now that $\mathbf{Y}(s)$ in Fig. 3 exists but that $\hat{\mathbf{Y}}$ does not exist. Then the normal rank of $[\mathbf{1}_{n+1} + \hat{\mathbf{S}}(s)]$ is $r < (n + 1)$. Equation (1) can be interpreted as a realization of $\hat{\mathbf{S}}(s)$ in terms of a $(2n + 2)$-port ideal transformer network, $n + 1 - r$ short-circuits, and a reactance r-port possessing a short-circuit admittance matrix $\hat{\mathbf{Y}}_r = [\mathbf{1}_r + \hat{\mathbf{S}}_r]^{-1}[\mathbf{1}_r - \hat{\mathbf{S}}_r]$, as shown in Fig. 4.

Since $\mathbf{E}_b = \mathbf{T}^t\mathbf{E}_a$, where $\mathbf{E}_b{}^t = [e_{b1}, e_{b2}, \cdots, e_{b(n+1)}]$ and $\mathbf{E}_a{}^t = [e_{a1}, e_{a2}, \cdots, e_{a(n+1)}]$, the number of independent linear relations among the components of $\mathbf{E}_a$ is equal to the number of zero components of $\mathbf{E}_b$. However, since $\mathbf{Y}$ exists but $\hat{\mathbf{Y}}$ does not exist, this number is equal to unity $(r = n)$, and the resulting single linear constraint is

$$\sum_{j=1}^{n+1} t_{j1}e_{aj} = 0, \qquad t_{(n+1)1} \neq 0 \tag{3}$$

in which the t_{j1} are the elements in the first column of $\mathbf{T}$. As a consequence, it is a simple matter to show that we may construct an $(n + 1) \times n$ matrix of real constants $\mathbf{A}$ such that

$$\begin{aligned} \tilde{\mathbf{E}}_b &= \mathbf{A}\tilde{\mathbf{E}}_a \\ \mathbf{A}^t\tilde{\mathbf{I}}_b &= \tilde{\mathbf{I}}_a \end{aligned} \tag{4}$$

* The superscript t denotes matrix transposition.

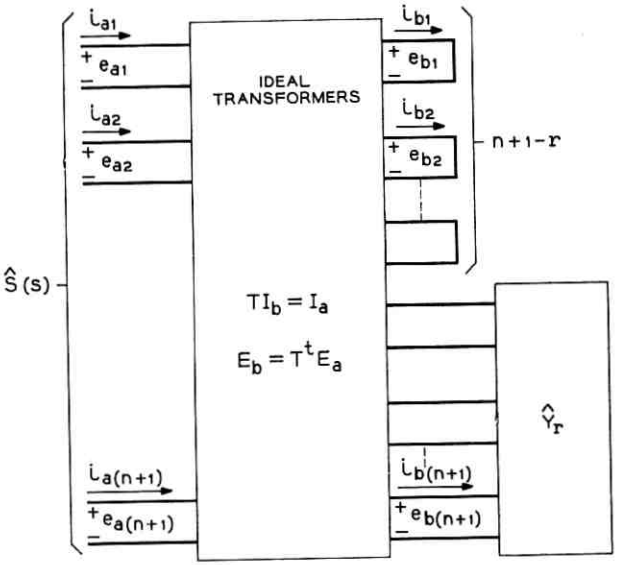


Fig. 4 — Realization of $\hat{\mathbf{S}}(s)$ when the normal rank of $[\mathbf{1}_{n+1} + \hat{\mathbf{S}}(s)]$ is r.

where

$$\tilde{\mathbf{E}}_b{}^t = [e_{a(n+1)}, e_{b2}, e_{b3}, \cdots, e_{b(n+1)}]$$
$$\tilde{\mathbf{I}}_b{}^t = [i_{a(n+1)}, i_{b2}, i_{b3}, \cdots, i_{b(n+1)}]$$
$$\tilde{\mathbf{E}}_a{}^t = [e_{a1}, e_{a2}, \cdots, e_{an}]$$
$$\tilde{\mathbf{I}}_a{}^t = [i_{a1}, i_{a2}, \cdots, i_{an}].$$

But

$$\tilde{\mathbf{I}}_b = \begin{bmatrix} y & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{Y}}_r \end{bmatrix} \tilde{\mathbf{E}}_b \tag{5}$$

which, together with (4), yields

$$\tilde{\mathbf{I}}_a = \mathbf{A}^t \begin{bmatrix} y & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{Y}}_r \end{bmatrix} \mathbf{A} \tilde{\mathbf{E}}_a . \tag{6}$$

Thus,

$$\mathbf{Y} = \mathbf{A}^t \begin{bmatrix} y & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{Y}}_r \end{bmatrix} \mathbf{A} = \mathbf{B}^t \mathbf{B} y + \mathbf{D}^t \hat{\mathbf{Y}}_r \mathbf{D} \tag{7}$$
$$= \mathbf{C} y + \mathbf{Y}'$$

where $\mathbf{B}^t$ is the n-vector of elements in the first column of $\mathbf{A}^t$ and $\mathbf{D}$ is the matrix of elements in the last n rows and columns of $\mathbf{A}$.

A very similar argument suffices to establish the result for the case in which $\hat{\mathbf{Z}}$ fails to exist.

III. REALIZABILITY OF $\mathbf{Z}(s)$ WHEN $\hat{\mathbf{Z}}$ EXISTS

Throughout the remainder of the paper we consider specifically the realizability of $\mathbf{Z}(s)$ under the assumption that $\hat{\mathbf{Z}}(s)$ exists. As mentioned earlier, this is equivalent to assuming that $\hat{\mathbf{Y}}$ exists and that the termination is replaced with its reciprocal in order to study the properties of $\mathbf{Y}(s)$.

We shall suppose that the impedance terminating port $(n + 1)$ is the positive-real function $z(s) = ab^{-1}$, where a and b are Hurwitz polynomials. Of course the impedance of the LC, $-$ R model is not a positive-real function. However, it is convenient to replace the negative resistor with a positive one of equal magnitude so that we may state that $\mathbf{Z}(s)$ is a positive-real matrix, with the understanding that $\mathbf{Z}(s) = -\overline{\mathbf{Z}}(-s)$ where $\overline{\mathbf{Z}}(s)$ is the impedance matrix of the n-port with the resistor negative.[13] Further, it is sufficient to assume that the LC, + R termination comprises an inductor of value α Henries in series with a parallel combination of a unit resistor and unit capacitor, for any other values can be accommodated by impedance and frequency scaling. Thus, we shall be particularly interested in the results for $z(s) = (\alpha s^2 + \alpha s + 1)(s + 1)^{-1}$, $(\alpha \geqq 0)$. However in this section we shall merely require that*

$$z_e = \frac{a_e b_e - a_o b_o}{b_e^2 - b_o^2}, \tag{8}$$

the even part of z, have no zeros on the finite $j\omega$-axis and that $b(s)$ is a strict Hurwitz polynomial.

It is well known that

$$\mathbf{Z} = \mathbf{Z}_{11} - \mathbf{Z}_{12}\mathbf{Z}_{12}^t \frac{1}{Z_{22} + z} \tag{9}$$

where the submatrices in (9) are defined by the following partition of $\hat{\mathbf{Z}}(s)$:

$$\hat{\mathbf{Z}}(s) = \begin{matrix} \\ n \\ 1 \end{matrix} \begin{matrix} \begin{matrix} n & 1 \end{matrix} \\ \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{12}^t & Z_{22} \end{bmatrix} \end{matrix}. \tag{10}$$

* Throughout the paper we shall denote by the subscripts o and e respectively the odd and even parts of polynomials or matrices. Thus, for example, $\mathbf{A}_e = \frac{1}{2}[\mathbf{A}(s) + \mathbf{A}(-s)]$, $\mathbf{A}_o = \frac{1}{2}[\mathbf{A}(s) - \mathbf{A}(-s)]$.

The arguments to be presented center about a study of $\mathbf{Z}_e$, the even part of the matrix $\mathbf{Z}$. This matrix is given by

$$\mathbf{Z}_e = -\tfrac{1}{2}[z(s) + z(-s)]\mathbf{Z}_{12}\mathbf{Z}_{12}^{\,t}\frac{1}{[Z_{22} + z][Z_{22}(-s) + z(-s)]} \tag{11}$$

It is convenient to introduce the notation $Z_{22} = d^{-1}n_{22}$, $\mathbf{Z}_{12} = d^{-1}\mathbf{N}_{12}$ where d is an even polynomial, n_{22} is an odd polynomial and $\mathbf{N}_{12}$ is an n-vector of odd polynomials, with the understanding that d, n_{22}, and every element of $\mathbf{N}_{12}$ may have a common simple zero at the origin.* In this way it is unnecessary to treat separately the cases in which d is even or d is odd.

Accordingly,

$$\mathbf{Z}_e = -(a_e b_e - a_o b_o)\mathbf{N}_{12}\mathbf{N}_{12}^{\,t} \cdot \frac{1}{[bn_{22} + ad][b(-s)n_{22}(-s) + a(-s)d(-s)]}. \tag{12}$$

Note that the assumptions regarding z, n_{22}, and d require that the polynomial $[bn_{22} + ad]$ be strictly Hurwitz except possibly for a simple zero at the origin. Also, as one would expect,[11] the zeros of $[bn_{22} + ad]$ cannot coincide with any of those of $(a_e b_e - a_o b_o)$. This follows from the fact that the existence of a nontrivial solution for a and b satisfying $bn_{22} + ad = 0$ and $(a_e b_e - a_o b_o) = \tfrac{1}{2}[ab(-s) + ba(-s)] = 0$ at some point $s = s_1$ requires that s_1 satisfy $b(-s_1)n_{22}(-s_1) + a(-s_1)d(-s_1) = 0$, which contradicts the fact that the zeros of $bn_{22} + ad$ are restricted to a half-plane.

It is convenient to state the following

Definition: The matrix $\mathbf{Z}_e$ is said to be in standard form if and only if

$$\mathbf{Z}_e = -(a_e b_e - a_o b_o)\mathbf{U}\mathbf{U}^t \frac{1}{v(s)v(-s)}$$

where $v(s)$ is a positive coefficient polynomial which is strictly Hurwitz except possibly for a simple zero at the origin and $\mathbf{U}^t = [u_1, u_2, \cdots, u_n]$ is a row matrix of odd real polynomials with the property that there is no factor $\eta(s)\eta(-s)$ common to all the u_i such that $\eta^2(s)$ divides $v(s)$ where $\eta(s)$ is a strict Hurwitz polynomial. Further, $v(s)$ and $(a_e b_e - a_o b_o)$ are relatively prime.

In Section 3.1 the following result is proved.†

Theorem 2: Denote by $z(s)$ the two-terminal positive-real impedance $z(s) = ab^{-1}$, with b a strict Hurwitz polynomial and z_e having no zeros on the

* With this exception, n_{22} and d are assumed to be relatively prime.

† We shall use the notation $\lim_{s\to\infty}[\cdot] = [\cdot]_\infty$ throughout.

finite $j\omega$-axis. Then the rational positive-real open-circuit impedance matrix $\mathbf{Z}(s)$ is realizable as shown in Fig. 3, with the understanding that the lossless reciprocal network possesses an open-circuit impedance matrix $\hat{\mathbf{Z}}(s)$, if and only if $\mathbf{Z}_e$ can be expressed in standard form and there exists a strict Hurwitz polynomial $\eta(s)$ defining

$$\mathbf{X} = \pm\mathbf{U}\eta(s)\eta(-s), \qquad w = v\eta^2(s)$$

such that

(i) $(w_e a_o - w_o a_e)(w_o b_o - w_e b_e)^{-1}$ *is a reactance function, the degenerate case in which* $(w_o b_o - w_e b_e) \equiv 0$ *not permitted.*

(ii) $\left[\mathbf{X}\,\dfrac{b_o a_o - b_e a_e}{w_e a_o - w_o a_e}\right]_\infty$ *exists, and*

(iii) $\left[\dfrac{1}{s}\,\mathbf{Z}\right]_\infty - \left[\mathbf{Z}_e\,\dfrac{a(w_e - w_o)}{s(a_o w_e - a_e w_o)}\right]_\infty$ *is nonnegative definite when the reactance function in (i) has a pole at infinity.*

Further, any $\mathbf{Z}(s)$ satisfying these conditions can be realized as shown in Fig. 3 with $\hat{\mathbf{Z}}(s)$ given by

$$\hat{\mathbf{Z}}(s) = \begin{bmatrix} \mathbf{Z} + \mathbf{X}\mathbf{X}^t\,\dfrac{b(a_o b_o - a_e b_e)}{(b_o w_o - b_e w_e)w} & \mathbf{X}\,\dfrac{(a_o b_o - a_e b_e)}{(b_o w_o - b_e w_e)} \\ \mathbf{X}^t\,\dfrac{(a_o b_o - a_e b_e)}{(b_o w_o - b_e w_e)} & \dfrac{(a_o w_e - a_e w_o)}{(b_o w_o - b_e w_e)} \end{bmatrix}.$$

3.1 *Proof of Theorem 2*

It is evident that a prescribed $\mathbf{Z}_e$ can be expressed in standard form if $\mathbf{Z}$ is realizable.* Suppose that $\mathbf{Z}$ and $\mathbf{Z}_e$ in standard form are given and consider the problem of determining d, n_{22}, $\mathbf{N}_{12}$, and $\mathbf{Z}_{11}$. In particular, let us consider identifying d, n_{22}, and $\mathbf{N}_{12}$ by equating the standard form expression for $\mathbf{Z}_e$ with the right-hand side of (12). A common factor may have been cancelled in the expression for $\mathbf{Z}_e$ and hence an unknown factor must be reinserted before we can proceed. However, the unknown factor must be of the form $f^2(s) = g(s)g(-s)$ where $g(s)$ is a strict Hurwitz polynomial. Therefore $f(s) = \eta(s)\eta(-s)$ where $\eta(s)$ is a strict Hurwitz polynomial. It follows that the most general expression for $\mathbf{Z}_e$ of the form:

$$\mathbf{Z}_e = -(a_e b_e - a_o b_o)\mathbf{X}\mathbf{X}^t\,\frac{1}{w(s)w(-s)} \tag{13}$$

* The problem of factoring $\mathbf{Z}_e$ into this form is discussed elsewhere.[10]

in which

1. $w(s)$ is a positive coefficient polynomial which is strictly Hurwitz except possibly for a simple zero at the origin, and
2. $\mathbf{X}^t$ is a row matrix of real odd polynomials

is expressible in terms of the standard form expression with

$$\mathbf{X} = \pm \mathbf{U}\eta(s)\eta(-s) \tag{14}$$

$$w = v\eta^2(s) \tag{15}$$

where $\eta(s)$ is an arbitrary real strict Hurwitz polynomial.

Thus if $\mathbf{Z}$ is realizable, $\mathbf{X} = \mathbf{N}_{12}$ and $w = bn_{22} + ad$ for some $\mathbf{X}$ and w generated by (14) and (15) with n_{22} and d respectively odd and even polynomials that are relatively prime, except possibly for a common simple zero at the origin, such that $n_{22}d^{-1}$ is a reactance function. Equating the even and odd parts of w and $b\bar{n}_{22} + a\bar{d}$ yields:

$$\bar{d} = (b_o w_o - b_e w_e)\,(b_o a_o - b_e a_e)^{-1} \tag{16}$$

$$\bar{n}_{22} = (a_o w_e - a_e w_o)\,(b_o a_o - b_e a_e)^{-1}. \tag{17}$$

Suppose now that $\bar{n}_{22}\bar{d}^{-1}$ is a reactance function. Then the two functions:

$$\frac{\bar{n}_{22}}{\bar{d}} + z = \frac{w(a_o b_o - a_e b_e)}{b(b_o w_o - b_e w_e)} \tag{18}$$

and

$$\frac{\bar{d}}{\bar{n}_{22}} + z^{-1} = \frac{w(a_o b_o - a_e b_e)}{a(a_o w_e - a_e w_o)} \tag{19}$$

are required to be positive-real. Since the even polynomial $(a_o b_o - a_e b_e)$ is either a constant or has zeros in the right-half plane, it is evident from (18) and (19) that (16) and (17) are polynomials. Furthermore in view of the positive-real property of (18) and (19) and the fact (16) and (17) are respectively even and odd polynomials, it follows that the zeros of (16) and (17) are restricted to the $j\omega$-axis and that these zeros are simple, except possibly for a double zero of $\bar{d}$ at the origin which can occur when v and hence w has a simple zero at the origin. With this single permissible exception, it is also the case that (16) and (17) are relatively prime, for the condition that there exists a nontrivial solution for w_e and w_o in

$$\begin{aligned} b_o w_o - b_e w_e &= 0 \\ -a_e w_o + a_o w_e &= 0 \end{aligned}$$

is $(a_ob_o - a_eb_e) = 0$, which, by assumption, is not satisfied for pure imaginary values of s. Thus, if $\bar{n}_{22}\bar{d}^{-1}$ is a reactance function, $w = bn_{22} + ad$ with $n_{22} = \bar{n}_{22}$ and $d = \bar{d}$ where, as required, n_{22} and d are relatively prime polynomials except possibly for a common simple zero at the origin.

Next consider the determination of the submatrices $\mathbf{Z}_{12}$ and $\mathbf{Z}_{11}$. For $d = \bar{d}$ given by (16),

$$\mathbf{Z}_{12} = \mathbf{X}d^{-1} = \mathbf{X}\frac{(b_oa_o - b_ea_e)}{(b_ow_o - b_ew_e)}. \tag{20}$$

Using (9), $z(s) = ab^{-1}$, $d = \bar{d}$, $n_{22} = \bar{n}_{22}$, and (20), we find

$$\mathbf{Z}_{11} = \mathbf{Z} + \mathbf{X}\mathbf{X}^t\frac{b(a_ob_o - a_eb_e)}{(b_ow_o - b_ew_e)w}. \tag{21}$$

By substituting $\mathbf{Z} = \mathbf{Z}_o + \mathbf{Z}_e$ in (21) with $\mathbf{Z}_e$ given by (13), it is easy to show that (21) is a matrix of odd functions, as it should be. Furthermore, since $(b_ow_o - b_ew_e) = d(a_ob_o - a_eb_e)$, it is evident from (21) that $\mathbf{Z}_{11}$ is regular in the entire finite strict left-half plane and consequently has finite poles only on the $j\omega$-axis.

Consider now the realizability of

$$\hat{\mathbf{Z}}(s) = \begin{bmatrix} \mathbf{Z} + \mathbf{X}\mathbf{X}^t\dfrac{b(a_ob_o - a_eb_e)}{(b_ow_o - b_ew_e)w} & \mathbf{X}\dfrac{(a_ob_o - a_eb_e)}{(b_ow_o - b_ew_e)} \\ \mathbf{X}^t\dfrac{(a_ob_o - a_eb_e)}{(b_ow_o - b_ew_e)} & \dfrac{(a_ow_o - a_ow_o)}{(b_ow_o - b_ew_e)} \end{bmatrix}. \tag{22}$$

We require the following lemma.[10]

Lemma 2: *The symmetric matrix of real constants*

$$\hat{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}{}^t & A_{22} \end{bmatrix}, \qquad A_{22} > 0$$

partitioned as in (10) *is nonnegative definite if and only if* $\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{12}{}^tA_{22}{}^{-1}$ *is nonnegative definite.*

Let $\hat{\mathbf{K}}_i$ denote the residue matrix at a pole of $\hat{\mathbf{Z}}(s)$ which arises from a zero of $(b_ow_o - b_ew_e)$ at say $s = j\omega_i$, and let the residue matrix of $\mathbf{Z}$ at that pole be $\mathbf{K}_i$. Then,*

* When $w = su = s[u_e + u_o]$, where u is a strict Hurwitz polynomial, it is necessary to replace w_e, w_o and $\mathbf{X}$ respectively with u_o, u_e and the n-vector of even polynomials $s^{-1}\mathbf{X}$ before this argument is applied to verify the nonnegative definiteness of the matrix of residues associated with the pole at the origin.

$$\hat{\mathbf{K}}_i = \left[\begin{array}{cc} \mathbf{K}_i + \mathbf{X}\mathbf{X}^t \dfrac{b(a_ob_o - a_eb_e)}{\dot{q}w} & \mathbf{X}\dfrac{(a_ob_o - a_eb_e)}{\dot{q}} \\ \mathbf{X}^t \dfrac{(a_ob_o - a_eb_e)}{\dot{q}} & \dfrac{a_ow_e - a_ew_o}{\dot{q}} \end{array} \right]_{s=j\omega_i} \tag{23}$$

where $\dot{q}$ is the derivative of $(b_ow_o - b_ew_e)$ with respect to s. Since $(\dot{q})^{-1}(a_ow_e - a_ew_o)$ evaluated at $s = j\omega_i$ is positive, $\hat{\mathbf{K}}_i$ is nonnegative definite if and only if

$$\begin{aligned} &\mathbf{K}_i + \mathbf{X}\mathbf{X}^t \frac{b(a_ob_o - a_eb_e)}{\dot{q}w}\bigg|_{s=j\omega_i} \\ &\quad - \frac{(a_ob_o - a_eb_e)^2}{(\dot{q})^2}\mathbf{X}\mathbf{X}^t \frac{\dot{q}}{(a_ow_e - a_ew_o)}\bigg|_{s=j\omega_i} \\ &\quad = \mathbf{K}_i + \mathbf{X}\mathbf{X}^t \frac{1}{\dot{q}}\left[\frac{(a_ob_o - a_eb_e)a(b_ew_e - b_ow_o)}{w(a_ow_e - a_ew_o)}\right]\bigg|_{s=j\omega_i} \end{aligned} \tag{24}$$

is nonnegative definite, a condition which is clearly satisfied since $(b_ew_e - b_ow_o)$ vanishes at $s = j\omega_i$.

Finally, we require that $\hat{\mathbf{Z}}$ have at most a simple pole at infinity and that $\hat{\mathbf{K}}_\infty = [(1/s)\hat{\mathbf{Z}}]_\infty$ be nonnegative definite. It is clearly necessary that the limit

$$[Z_{22}^{-1}\mathbf{Z}_{12}]_\infty = \left[\mathbf{X}\frac{(a_ob_o - a_eb_e)}{(a_ow_e - a_ew_o)}\right]_\infty \tag{25}$$

exist. When $[Z_{22}^{-1}\mathbf{Z}_{12}]_\infty$ does exist with Z_{22} a reactance function, it follows from (9) that $\mathbf{Z}_{11}$ has at most a simple pole at infinity, since $\mathbf{Z}$ is assumed to possess this property, and consequently that $\hat{\mathbf{Z}}$ has at most a simple pole at infinity.

Suppose now that Z_{22} has a pole at infinity. According to Lemma 2 and (22), $\hat{\mathbf{K}}_\infty$ is nonnegative definite if and only if the following matrix of constants is nonnegative definite:

$$\begin{aligned} &\mathbf{K}_\infty + \left[\frac{1}{s}\mathbf{X}\mathbf{X}^t \frac{b(a_ob_o - a_eb_e)}{(b_ow_o - b_ew_e)w}\right]_\infty - \left[\frac{1}{s}\mathbf{X}\frac{(a_ob_o - a_eb_e)}{(b_ow_o - b_ew_e)}\right]_\infty \\ &\qquad \cdot \left[\frac{1}{s}\mathbf{X}^t \frac{(a_ob_o - a_eb_e)}{(b_ow_o - b_ew_e)}\right]_\infty \left[\frac{1}{s}\frac{(a_ow_e - a_ew_o)}{(b_ow_o - b_ew_e)}\right]_\infty^{-1}. \end{aligned} \tag{26}$$

Some manipulation shows that (26) can be rewritten as simply

$$\mathbf{K}_\infty - \left[\mathbf{Z}_e \frac{a(w_e - w_o)}{s(a_ow_e - a_ew_o)}\right]_\infty. \tag{27}$$

If Z_{22} does not have a pole at infinity, $\hat{\mathbf{K}}_\infty$ is nonnegative definite if and only if $\mathbf{Z}_{12}$ does not have a pole at infinity and $[(1/s)\mathbf{Z}_{11}]_\infty$ is nonnegative definite. The first of these requirements is contained in the condition that (25) exist, while the second follows from the assumed positive-real property of $\mathbf{Z}$, for in this case $[(1/s)\mathbf{Z}]_\infty = [(1/s)\mathbf{Z}_{11}]_\infty$.

This proves Theorem 2.

3.2 *Remarks Concerning Theorem 2*

Consider first the conditions under which the degenerate situation $(b_o w_o - b_e w_e) \equiv 0$ can arise. Assume that both b_o and b_e do not vanish identically in s. Note that a strict Hurwitz w cannot satisfy the equation, for b is assumed to be strictly Hurwitz and hence if w is also strictly Hurwitz $b_o w_o$ vanishes at the origin while $b_e w_e$ does not. The alternative possibility is that $w = su$ where u is a strict Hurwitz polynomial. In this event we have $(u_e b_o - u_o b_e) \equiv 0$ and therefore* $u = b$. This leads to the following expression for $\mathbf{Z}_e$:

$$\mathbf{Z}_e = -(a_e b_e - a_o b_o)\mathbf{X}\mathbf{X}^t \frac{1}{[s(b_e + b_o)][-s(b_e - b_o)]}$$

Since each element in the matrix $\mathbf{Z}_e$ must approach zero at least as fast as z_e, which is obvious from (11), the degree of each polynomial in the n-vector of odd polynomials $\mathbf{X}$ is at most unity. Thus $\mathbf{X}$ contains no non-$j\omega$-axis factors that can be cancelled and consequently $(b_o w_o - b_e w_e) \equiv 0$ implies that η^2 is a constant, say unity, and $w = s(b_e + b_o)$ or, equivalently, $(b_o v_o - b_e v_e) \equiv 0$. However given $(b_o v_o - b_e v_e) \equiv 0$, it is not clear that a nonconstant choice of η^2 could not render $\bar{d}^{-1}\bar{n}_{22}$ realizable as a reactance function. To resolve this question assume that $w = \eta^2 s(b_e + b_o)$ and consider $z + \bar{d}^{-1}\bar{n}_{22}$ which must be a positive-real function if $\bar{d}^{-1}\bar{n}_{22}$ is realizable. Some algebra yields

$$z + \bar{d}^{-1}\bar{n}_{22} = z_e \frac{\eta^2}{2\eta_e \eta_o}. \tag{28}$$

It is clear that (28) is not a positive real function for any choice of η when b_e, $b_o \not\equiv 0$ because of the right-half plane poles of z_e. Thus,

Lemma 3: When b_e, $b_o \not\equiv 0$, condition (i) can be replaced with the statement: $(w_e a_o - w_o a_e)(w_o b_o - w_e b_e)^{-1}$ is a reactance function and $(b_o v_o - b_e v_e) \not\equiv 0$.

The discussion relating to the realization of $\mathbf{Z}$ when $z = (s + 1)$ shows that the assumption b_e, $b_o \not\equiv 0$ is necessary.

* We are ignoring a trivial constant multiplicative factor.

We wish to show now that condition (*iii*) is invariably satisfied when z is regular at infinity. This occurs because the limit

$$\left[\mathbf{Z}_e \frac{a(w_e - w_o)}{s(a_o w_e - a_e w_o)}\right]_\infty \tag{29}$$

vanishes when z is regular at infinity. To prove this it is sufficient to consider the limit obtained by evaluating (29) with $\mathbf{Z}_e$ replaced with z_e. Using $w = b\bar{n}_{22} + a\bar{d}$ to compute w_e and w_o, we find

$$\left[z_e \frac{a(w_e - w_o)}{s(a_o w_e - a_e w_o)}\right] = \frac{1}{s} z \left[\frac{\bar{d}}{\bar{n}_{22}} \left(\frac{a_e - a_o}{b_e - b_o}\right) - 1\right]$$

from which our assertion is obvious.* Thus,

Lemma 4: Condition (iii) is satisfied when $z(s)$ is regular at infinity.

In the following sections we shall use Theorem 2 and Lemmas 3 and 4 to obtain explicit realizability conditions for $z(s)$ or $z^{-1}(s) = (s + 1)^{-1}(\alpha s^2 + \alpha s + 1)$. *We assume throughout that* $\mathbf{Z}(s)$ *is known to be a rational symmetric* $n \times n$ *positive-real matrix, and that* $\mathbf{Z}_e$ *in standard form is given by*†

$$\mathbf{Z}_e = -\mathbf{U}\mathbf{U}^t \frac{1}{v(s)v(-s)}. \tag{30}$$

To further avoid repetition, the term "realizable with an impedance $z(s)$" is to be understood to refer to the realizability of the multiport matrix as a structure shown in Fig. 3 with the provision that $\hat{\mathbf{Z}}(s)$ *exists. It is convenient to treat separately the cases in which* $\alpha = 0$, *and* $\alpha > 0$.

IV. Explicit Realizability Conditions When $z = (s + 1)^{-1}$ and $z = (s + 1)$

When $z(s) = (s + 1)^{-1}$, conditions (*i*) and (*ii*) reduce to

(*i*) $w_o(w_e - sw_o)^{-1}$ must be a reactance function and $v_e \not\equiv sv_o$ (see Lemma 3).

(*ii*) $[\mathbf{X}(1/w_o)]_\infty$ must exist.

According to Lemma 4 condition (*iii*) is satisfied.

* This result can be established in a more direct fashion by observing that the nonnegative definiteness of $\hat{\mathbf{K}}_\infty$ is implied by Lemma 2 and expression (9) when $z(s)$ is regular at infinity in view of the positive-real property of $\mathbf{Z}(s)$. Nevertheless, it is instructive to consider this matter from the viewpoint presented above.

† Here $(a_e b_e - a_o b_o) = 1$.

Condition (i) requires that $[(1/s)(w_e/w_o)]_\infty \geqq 1$. Observe that w must be of even degree and that $[(1/s)(w_e/w_o)]_\infty$ is simply the reciprocal of the negative of the sum of the zeros of the polynomial $w(s)$. It follows at once that (i) is satisfied for some $\eta(s)$ if and only if $[(1/s)(v_e/v_o)]_\infty \geqq 1$. Requirement (ii) is satisfied without additional restrictions on $\mathbf{Z}(s)$, for (i) implies that

$$[\mathbf{Z}_e]_\infty = \left[-\frac{1}{w(s)\ w(-s)} \mathbf{X}\mathbf{X}^t \right]_\infty = \mathbf{0} \tag{31}$$

since $\mathbf{X}$ is an n-vector of odd polynomials and $\mathbf{Z}_e$ is bounded at infinity. This proves

*Theorem 3:**

The matrix $\mathbf{Z}$ *is realizable with an impedance* $(s+1)^{-1}$ *if and only if* $v_e \not\equiv sv_o$ *and* $\left[\frac{v_e}{sv_o}\right]_\infty \geqq 1$.

4.1 $z(s) = (s+1)$

In this more interesting case† the three conditions become

(i) $(w_o - sw_e)\ (w_e)^{-1}$ must be a reactance function and $w_e \not\equiv 0$.

(ii) $[\mathbf{X}\ (w_o - sw_e)^{-1}]_\infty$ must exist.

(iii) $\left[\frac{1}{s}\mathbf{Z}\right]_\infty - \left[\mathbf{Z}_e \frac{(s+1)\ (w_e - w_o)}{s(sw_e - w_o)}\right]_\infty$ must be nonnegative definite when $(w_o - sw_e)\ (w_e)^{-1}$ has a pole at infinity.

From (i), $k' = [w_o/sw_e]_\infty \geqq 1$. Since

$$\mathbf{Z}_e = -\mathbf{X}\mathbf{X}^t \frac{1}{w(s)\ w(-s)},$$

and (i) requires that $w(s)$ be of odd degree, (ii) is satisfied for $k' > 1$. However if $k' = 1$, (ii) is satisfied if and only if $[\mathbf{Z}_e]_\infty = \mathbf{0}$. In terms of k', condition (iii) is equivalent to the statement

$$\left[\frac{1}{s}\mathbf{Z}\right]_\infty - \frac{k'}{k'-1}[\mathbf{Z}_e]_\infty \tag{32}$$

must be nonnegative definite when $k' > 1$.

* This result was stated without proof in Ref. 10.

† This case together with the situation in which $\hat{\mathbf{Z}}$ does not exist is treated in detail from a somewhat different viewpoint in Ref. 10. It is included here primarily to illustrate the application of Theorem 2.

Consider now the influence of the strict Hurwitz polynomial $\eta(s)$. Note first that $\mathbf{Z}_e$ may be a matrix of constants; that is, $v(s)$ may be equal to γs where γ is a positive constant. Let β be the reciprocal of the negative of the sum of the zeros of $\eta(s)$. In this case $k' = \beta$ and β can be chosen arbitrarily large to minimize $k'(k' - 1)^{-1}$. Therefore when $\mathbf{Z}_e$ is a matrix of constants, (i), (ii), and (iii) reduce to the requirement that

$$\left[\frac{1}{s}\mathbf{Z}\right]_\infty - (1 + \epsilon)[\mathbf{Z}_e]_\infty$$

be nonnegative definite for some* $\epsilon > 0$.

When $\mathbf{Z}_e$ is not a matrix of constants, the most favorable choice of $\eta(s)$ is simply a constant, for then k' is maximized and $k'(k' - 1)^{-1}$ is minimized. Thus,

Theorem 4: The matrix $\mathbf{Z}$ *is realizable with an impedance* $z = (s + 1)$ *if and only if*

1. *When* $\mathbf{Z}_e$ *is a matrix of constants,* $[(1/s)\mathbf{Z}]_\infty - (1 + \epsilon)[\mathbf{Z}_e]$ *is nonnegative definite for some* $\epsilon > 0$.
2. *If* $\mathbf{Z}_e$ *is not a matrix of constants,* $k = [v_o/sv_e]_\infty \geqq 1$; *if* $k = 1$, $[\mathbf{Z}_e]_\infty = \mathbf{0}$; *if* $k > 1$, $[(1/s)\mathbf{Z}]_\infty - [k/(k - 1)][\mathbf{Z}_e]_\infty$ *is nonnegative definite.*

V. EXPLICIT REALIZABILITY CONDITIONS WHEN $z(s)$ OR z^{-1} $(s) = (\alpha s^2 + \alpha s + 1)$ $(s + 1)^{-1}$, $\alpha > 0$

In these cases, as will become clear, the polynomial $\eta(s)$ plays a central role in determining the realizability conditions. We shall consider first the case: $z(s) = (\alpha s^2 + \alpha s + 1)(s + 1)^{-1}$. Here condition (i) requires that

$$(i)\quad \frac{\alpha s w_e - w_o(\alpha s^2 + 1)}{s w_o - w_e} = -\alpha s + \frac{w_o}{w_e - s w_o}$$

must be a reactance function and, using Lemma 3, $sv_o \not\equiv v_e$.

It is clearly necessary that $[w_e/sw_o]_\infty = 1$. Thus we may assume that v and w are of even degree. Let

$$v(s) = \sum_{k=0}^{2m} v_k s^k \tag{33}$$

* If the lossless network is not required to possess an open-circuit impedance matrix, ϵ can be taken to be zero.[10]

$$w(s) = \eta^2(s)v(s) = \sum_{k=0}^{2p} w_k s^k \tag{34}$$

where $p = m +$ (degree of η). Then, since the reactance function must have a nonnegative "residue" at infinity,

$$w_{2p-1}(w_{2p-2} - w_{2p-3})^{-1} \geqq \alpha$$

or since $w_{2p-1} = w_{2p}$,

$$1 - \alpha(w'_{2p-2} - w'_{2p-3}) \geqq 0 \tag{35}$$

where $w_k' = w_k(w_{2p})^{-1}$.

Condition (ii) reads:

$$(ii) \quad \left[\mathbf{X}\frac{\alpha s^2 - (\alpha s^2 + 1)}{\alpha s w_e - (\alpha s^2 + 1)w_o}\right]_\infty = \left[\mathbf{X}\frac{-1}{\alpha s(w_e - s w_o) - w_o}\right]_\infty$$

must exist.

Assume first that (35) holds with strict inequality in which case (ii) becomes

$$\left[\mathbf{X}\frac{-1}{\alpha(w_{2p-2} - w_{2p-3})\, s^{2p-1} - w_{2p}\, s^{2p-1}}\right]_\infty \tag{36}$$

must exist. But $[\mathbf{X}(1/s^{2p})]_\infty = \mathbf{0}$, since w is of even degree, and therefore the limit (36) does indeed exist. Suppose now that (35) holds with the equal sign. Then (ii) requires that

$$\left[\mathbf{X}\frac{-1}{\alpha(w_{2p-4} - w_{2p-5})\, s^{2p-3} - w_{2p-3}\, s^{2p-3}}\right]_\infty \tag{37}$$

exist.* Hence the degree of $\mathbf{X}$ must be $(2p - 3)$ at most. Since the degree of w is $2p$, (37) will exist if and only if

$$[s^2\mathbf{Z}_e]_\infty = \mathbf{0} \tag{38}$$

Consider now condition (iii) which requires that

$$\left[\frac{1}{s}\mathbf{Z}\right]_\infty + [s^2\mathbf{Z}_e]_\infty \frac{\alpha}{1 - \alpha(w'_{2p-2} - w'_{2p-3})} \tag{39}$$

be nonnegative definite when (35) is satisfied with strict inequality.

From the form of conditions (35) and (39), it is clear that the most favorable realizability conditions for a given $\mathbf{Z}(s)$ are obtained when

* It can easily be shown that the reactance function property of the expression in (i) implies that the denominator in (37) does not vanish identically in s when (35) is satisfied with equality.

$\eta(s)$ is chosen to satisfy $[w_e/(sw_o)]_\infty = 1$ and to simultaneously minimize* $(w'_{2p-2} - w'_{2p-3})$.

This obviously requires that $[v_e/(sv_o)]_\infty \geqq 1$. We wish to establish

Lemma 5: The minimum value of $(w'_{2p-2} - w'_{2p-3})$ *is*

$$\varphi_1 = \frac{1}{3}\left[1 - \left(\frac{v_{2m-1}}{v_{2m}}\right)^3\right] - \frac{v_{2m-3}}{v_{2m}} + \left[\frac{v_{2m-1}}{v_{2m}}\right]\left[\frac{v_{2m-2}}{v_{2m}}\right] - \frac{1}{12}\left[1 - \frac{v_{2m-1}}{v_{2m}}\right]^3$$

and is attained when $\eta(s) = \lambda\left[s + \tfrac{1}{2}\left(1 - \frac{v_{2m-1}}{v_{2m}}\right)\right]^\delta$ *where* λ *is a positive constant and*

$$\delta = 1, \qquad \frac{v_{2m-1}}{v_{2m}} < 1$$

$$\delta = 0, \qquad \frac{v_{2m-1}}{v_{2m}} = 1.$$

5.1 *Proof of Lemma 5*

Denote by $s_1, s_2, \ldots, s_{2p}$ the zeros of $w(s)$. Using a result[14] due to Newton, we find that

$$\sum_{k=1}^{2p} s_k^3 = -(w'_{2p-1})^3 + 3(w'_{2p-1})(w'_{2p-2}) - 3(w'_{2p-3}) \tag{40}$$

Recalling that here $w'_{2p-1} = 1$,

$$\begin{aligned} w'_{2p-2} - w'_{2p-3} &= \tfrac{1}{3} + \tfrac{1}{3}\sum_{k=1}^{2p} s_k^3 \\ &= \tfrac{1}{3} + \tfrac{1}{3}\sum_{v} s_k^3 + \tfrac{2}{3}\sum_{\eta} s_k^3 \end{aligned} \tag{41}$$

where $\sum_v$ and $\sum_\eta$ denote respectively sums taken only over those indices corresponding to zeros of the polynomials v and η. Hence our problem reduces to determining the strict Hurwitz $\eta(s)$ such that $\sum_\eta s_k^3$ is minimized subject to

$$\sum_{\eta} s_k = \tfrac{1}{2}\left[\frac{v_{2m-1}}{v_{2m}} - 1\right] \leqq 0.$$

Of course when $\sum_\eta s_k = 0$, η is simply a constant. Assume then that

* Note that $[s^2\mathbf{Z}_e]$ is nonpositive definite.

$\sum_{\eta} s_k < 0$. First observe that the real-part of $(-g + jh)^3$ (g, h real constants; g positive), exceeds $-g^3$ for all $h \neq 0$. Hence the optimum η has only real zeros. Next note that $\sum_{\eta} s_k^3 \geqq [\sum_{\eta} s_k]^3$ when each s_k is a negative constant. Thus the optimum η is a single linear factor

$$\lambda\left[s + \tfrac{1}{2}\left(1 - \frac{v_{2m-1}}{v_{2m}}\right)\right]$$

in which λ is a positive-real constant and the corresponding minimum value of $(w'_{2p-2} - w'_{2p-3})$ is φ_1, where

$$\varphi_1 = \frac{1}{3} + \frac{1}{3}\sum_{v} s_k^3 - \frac{1}{12}\left(1 - \frac{v_{2m-1}}{v_{2m}}\right)^3. \tag{42}$$

Expression (42) can be written in the more convenient form given in Lemma 5 by using (40).

The results of this section can be summarized as follows.

Theorem 5: The matrix $\mathbf{Z}(s)$ *is realizable with an impedance* $z(s) = (\alpha s^2 + as + 1)(s + 1)^{-1}$, *where* $\alpha > 0$, *if and only if*

1. $\left[\dfrac{v_e}{sv_o}\right]_\infty = \dfrac{v_{2m}}{v_{2m-1}} \geq 1, \quad v_e \not\equiv sv_o$
2. $1 - \alpha\varphi_1 \geqq 0$
3. If $1 - \alpha\varphi_1 = 0$, $[s^2\mathbf{Z}_e]_\infty = \mathbf{0}$
4. If $1 - \alpha\varphi_1 > 0$, $[(1/s)\mathbf{Z}]_\infty + [s^2\mathbf{Z}_e]_\infty[\alpha/(1 - \alpha\varphi_1)]$

must be nonnegative definite, where

$$\varphi_1 = \frac{1}{3}\left[1 - \left(\frac{v_{2m-1}}{v_{2m}}\right)^3\right] - \frac{v_{2m-3}}{v_{2m}} + \left[\frac{v_{2m-1}}{v_{2m}}\right]\left[\frac{v_{2m-2}}{v_{2m}}\right] - \frac{1}{12}\left[1 - \frac{v_{2m-1}}{v_{2m}}\right]^3$$

and the v_k *are defined by* $v(s) = \sum_{k=0}^{2m} v_k s^k$.

Further, if $\mathbf{Z}$ *satisfies these conditions,* $\hat{\mathbf{Z}}(s)$ *given by* (22) *is realizable with* $\eta = 1$ *when* $(v_{2m})(v_{2m-1})^{-1} = 1$ *and with*

$$\eta(s) = s + \tfrac{1}{2}\left(1 - \frac{v_{2m-1}}{v_{2m}}\right)$$

when $(v_{2m})(v_{2m-1})^{-1} > 1$.

5.2 *Realizability with* $z(s) = (s + 1)(\alpha s^2 + \alpha s + 1)^{-1}$

Here condition (i) requires that

$$\frac{sw_e - w_o}{\alpha s w_o - (\alpha s^2 + 1) w_e} = \left[-\alpha s + \frac{w_e}{w_o - s w_e} \right]^{-1}$$

must be a reactance function and $\alpha s v_o - (\alpha s^2 + 1) v_e \not\equiv 0$. Thus,

$$\left[\frac{w_o}{s w_e} \right]_\infty = 1, \quad \text{and} \quad 1 - \alpha (w'_{2p-1} - w'_{2p-2}) \geqq 0 \tag{43}$$

where $w = \sum_{k=0}^{2p+1} w_k s^k$ and $w_k' = w_k (w_{2p+1})^{-1}$.

Condition (ii) requires that

$$\left[\mathbf{X} \frac{1}{w_o - s w_e} \right]_\infty = \left[\mathbf{X} \frac{1}{(w_{2p-1} - w_{2p-2}) s^{2p-1}} \right]_\infty$$

exist, which is valid if and only if $[s^4 \mathbf{Z}_e]_\infty$ exists. According to Lemma 4, condition (iii) is satisfied.

A moments reflection, in view of the two expressions in (43), will show that the determination of the polynomial $\eta(s)$ which leads to the weakest realizability conditions on $\mathbf{Z}(s)$ is essentially the same problem considered in the last section. The final result reads

Theorem 6:

The matrix $\mathbf{Z}$ *is realizable with an impedance* $z(s) = (s + 1)(\alpha s^2 + \alpha s + 1)^{-1}$, *where* $\alpha > 0$, *if and only if*

1. $\left[\frac{v_o}{s v_e} \right]_\infty = \frac{v_{2m+1}}{v_{2m}} \geqq 1, \quad \alpha s v_o - (\alpha s^2 + 1) v_e \not\equiv 0$
2. $[s^4 \mathbf{Z}_e]_\infty$ *exists*
3. $1 - \alpha \varphi_2 \geqq 0$

where

$$\varphi_2 = \frac{1}{3}\left[1 - \left(\frac{v_{2m}}{v_{2m+1}} \right)^3 \right] - \frac{v_{2m-2}}{v_{2m+1}} + \left[\frac{v_{2m}}{v_{2m+1}} \right] \left[\frac{v_{2m-1}}{v_{2m+1}} \right] - \frac{1}{12}\left[1 - \frac{v_{2m}}{v_{2m+1}} \right]^3$$

and $v(s) = \sum_{k=0}^{2m+1} v_k s^k$.

Further, if $\mathbf{Z}$ *satisfies these conditions, the corresponding* $\hat{\mathbf{Z}}$ *is realizable with* $\eta(s) = 1$ *when* $(v_{2m+1})(v_{2m})^{-1} = 1$ *and with*

$$\eta(s) = s + \tfrac{1}{2}\left[1 - \frac{v_{2m}}{v_{2m+1}} \right]$$

when $(v_{2m+1})(v_{2m})^{-1} > 1$.

VI. CONCLUDING OBSERVATION

It is of interest to note that the conditions presented in Theorem 5 $[z = (\alpha s^2 + \alpha s + 1)(s + 1)^{-1}]$ reduce to those of Theorem 3 $[z = (s + 1)^{-1}]$ as α approaches zero. However, a similar situation does not occur with respect to Theorems 4 and 6, for here the behavior of the matrix of even-parts at infinity is critically dependent upon whether $\alpha = 0$ or $\alpha > 0$.

REFERENCES

1. Sommers, H. S., Jr., Tunnel Diodes as High-Frequency Devices, Proc. I.R.E., **47,** July, 1959, pp. 1201–1206.
2. Hines, M. E., High-Frequency Negative-Resistance Circuit Principles for Esaki Diode Applications, B.S.T.J., **39,** May, 1960, pp. 477–513.
3. Youla, D. C., and Smilen, L. I., Optimum Negative-Resistance Amplifiers, Symposium on Active Networks and Feedback Systems, Polytechnic Institute of Brooklyn, 1960.
4. Smilen, L. I., and Youla, D. C., Exact Theory and Synthesis of a Class of Tunnel Diode Amplifiers, Proc. Nat. Elec. Conf., October, 1960.
5. Kuh, E. S., and Patterson, J. D., Design Theory of Optimum Negative-Resistance Amplifiers, Proc. I.R.E., **49,** June, 1961, pp. 1043–1050.
6. Desoer, C. A., and Kuh, E. S., Bounds on Natural Frequencies of Linear Active Networks, Symposium on Active Networks and Feedback Systems, Polytechnic Institute of Brooklyn, 1960.
7. Kinariwala, B. K., Theoretical Limitations on the Esaki Diode as a Network Element, NEREM Record, **2,** 1960, pp. 86–87.
8. Kinariwala, B. K., The Esaki Diode as a Network Element, I.R.E.-PGCT, **CT-8,** December, 1961.
9. Weinberg, L., Synthesis Using Tunnel Diodes and Masers, I.R.E.-PGCT, **CT-8,** March, 1961, pp. 66–75.
10. Sandberg, I. W., The Realizability of Multiport Structures Obtained by Imbedding a Tunnel Diode in a Lossless Reciprocal Network, B.S.T.J., **41,** May, 1962, pp. 857–876.
11. Schoeffler, J. D., Impedance Transformation Using Lossless Networks, I.R.E.-PGCT, June, 1961, pp. 131–137.
12. Youla, D., Castriota, L., and Carlin, H. J., Bounded Real Scattering Matrices and the Foundations of Linear Passive Network Theory, I.R.E.-PGCT, **CT-6,** March, 1959, pp. 102–124.
13. Youla, D., Physical Realizability Criteria, I.R.E. Conv. Rec., **CT-7,** August, 1960.
14. Uspensky, J. V., *Theory of Equations*, New York, McGraw-Hill, 1948, p. 262.

Stimulated Emission of Bremsstrahlung

By D. MARCUSE

(Manuscript received May 31, 1962)

A formula for the probability of stimulated emission of Bremsstrahlung is derived. It is shown that stimulated emission occurs if the incident electrons travel in a direction roughly parallel to the electric field vector of the wave stimulating the emission.

Emission from free electrons is used in electron tube devices. The purpose of this paper is to show that stimulated emission occurs already in the elementary process of the encounter of one electron and one nucleus or ion. There is no need for a slow-wave structure or elaborate electron bunching and no need to consider phase relationships. This elementary effect of stimulated emission should lead to a type of oscillator and broadband amplifier working without slow-wave structures or need for the close mechanical tolerances of high-frequency klystrons. A device of this kind may be noisier than a conventional maser.

It may be that the effect discussed in this paper is responsible for some of the hitherto unexplained semiconductor diode oscillations which have been reported in the literature.

I. INTRODUCTION

Stimulated emission of radiation from atoms or molecules is the process by which a maser operates. All masers use the transitions between bound states for their operation.

However, it is also well known that radiation can be emitted from a free electron in the presence of a static electric or magnetic field. The presence of this static field—for example the field of a nucleus—is necessary to simultaneously conserve energy and momentum during the emission or absorption process. Free electrons far from any other field can neither emit nor absorb photons of an infinitely extending radiation field because if they did, conservation of energy and momentum would be violated, as can easily be shown. Free electrons can only scatter photons, a phenomenon known as the Compton effect.

The emission of radiation from an electron passing by a nucleus is

known as Bremsstrahlung. It will be shown in this paper that stimulated emission of Bremsstrahlung is possible if the incident electron travels more or less parallel to the electric field vector of the stimulating radiation field. The electron absorbs radiation if it travels more or less perpendicular to the electric field vector of the stimulating field.

Since stimulated emission of Bremsstrahlung exists, amplifiers and oscillators may be constructed using this effect. Stimulated emission of Bremsstrahlung does not require any bunched electron beams or observation of phase relationships. Moreover, it works better with slow than with fast electrons, and no traveling-wave structure is required since the fields for the stimulation process can be confined in a cavity.

II. STIMULATED EMISSION OF BREMSSTRAHLUNG

The theoretical principles required to derive the probabilities for stimulated emission or absorption from free electrons in the presence of a Coulomb field can be found in the textbook by Heitler.[1] We limit ourselves to nonrelativistic electron velocities and derive the probability for transitions between the following two states. The initial state consists of a free electron represented by a plane wave existing in the presence of a Coulomb and a radiation field with a certain number n of photons in a particular mode, while all other modes are empty. The final state consists of the same electron with different energy and momentum and with a number of $n + 1$ photons in the case of stimulated emission, or $n - 1$ photons in the case of absorption.

The transition probability per unit time is given by[2]

$$w = \frac{2\pi}{\hbar} \mid K_{F,0} \mid^2 \rho_F \tag{1}*$$

with $\hbar = 1.05 \times 10^{-27}$ erg·sec

$$K_{F,0} = \sum \left\{ \frac{V_{FI} H_{I0}}{E_o - E'} + \frac{H_{FII} V_{II0}}{E_o - E''} \right\}. \tag{2}$$

E_o is the initial energy of the whole system, while E' and E'' are the energies of the system in the intermediate virtual states. In the transitions to the intermediate states, energy need not be conserved. However, energy conservation is certainly required between the initial and final states of the whole system.

$$E_o = \frac{p_0^2}{2m} + n\hbar\omega \tag{3}$$

* A list of symbols is given below in Section VII.

$$E' = \frac{p'^2}{2m} + (n \pm 1)\hbar\omega \tag{4}$$

$$E'' = \frac{p''^2}{2m} + n\hbar\omega. \tag{5}$$

The (+) or (−) signs in (4) refer to emission and absorption respectively. $p = mv$ is the momentum of the electron and ω is the angular frequency of the electromagnetic field.

The summation in (2) extends over all possible intermediate states. The matrix elements of the interaction Hamiltonian are given by[3]*

$$H_{I0} = -\frac{1}{\sqrt{L^3}}\frac{e}{m}\sqrt{\frac{2\pi\hbar}{\omega}}\begin{Bmatrix}\sqrt{n+1}\\ \sqrt{n}\end{Bmatrix}\frac{1}{L^3}\int e^{-i\mathbf{k}'\cdot\mathbf{r}}\, p_e\, e^{\mp i\boldsymbol{\beta}\cdot\mathbf{r}}\, e^{i\mathbf{k}^0\cdot\mathbf{r}}\, d^3r$$
$$= \frac{\hbar(k_x^0 \mp \beta_x)}{\sqrt{L^3}}\frac{e}{m}\sqrt{\frac{2\pi\hbar}{\omega}}\begin{Bmatrix}\sqrt{n+1}\\ \sqrt{n}\end{Bmatrix}\delta_{k_x',k_x^0\mp\beta_x}\cdot\delta_{k_y',k_y^0\mp\beta_y}\cdot\delta_{k_z',k_z^0\mp\beta_z} \tag{6}$$

$$H_{FII} = -\frac{1}{\sqrt{L^3}}\frac{e}{m}\sqrt{\frac{2\pi\hbar}{\omega}}\begin{Bmatrix}\sqrt{n+1}\\ \sqrt{n}\end{Bmatrix}\frac{1}{L^3}\int e^{-i\mathbf{k}''\cdot\mathbf{r}}\, p_e\, e^{\mp i\boldsymbol{\beta}\cdot\mathbf{r}}\, e^{i\mathbf{k}''\cdot\mathbf{r}}\, d^3r$$
$$= -\frac{\hbar(k_x'' \mp \beta_x)}{\sqrt{L^3}}\frac{e}{m}\sqrt{\frac{2\pi\hbar}{\omega}}\begin{Bmatrix}\sqrt{n+1}\\ \sqrt{n}\end{Bmatrix} \tag{7}$$
$$\delta_{k_x'',k_x{}^F\pm\beta_x}\,\delta_{k_y'',k_y{}^F\pm\beta_y}\,\delta_{k_z'',k_z{}^F\pm\beta_z}.$$

The δ's are the Kronecker δ-symbols. The k's and β's can assume only values of the form $2\pi n'/L$ with integer n' as a result of box normalization. e and m are the charge and mass of the electron, n is the number of photons in a large box of volume L^3 (box normalization). The upper values $\sqrt{n+1}$ and (−) signs belong to the emission case, while the lower values $\sqrt{n}$ and (+) signs belong to the case of absorption, $\boldsymbol{\beta}$ is the propagation vector of the plane electromagnetic wave with $\beta = |\boldsymbol{\beta}| = \omega/c$, $\mathbf{k}$ is the propagation vector of the plane electron wave with $\hbar|\mathbf{k}| = mv$, and p_e is the component of the momentum operator of the electron in the direction of the electric field vector of the radiation field. We choose as z-direction the direction of propagation of the stimulating light wave

$$\boldsymbol{\beta} = (0,0,\beta). \tag{8}$$

The direction of polarization (direction of electric field vector) is taken to be the x-direction so that

* Electrostatic c.g.s. units will be used throughout this paper.

$$p_e = -i\hbar \frac{\partial}{\partial x}. \tag{9}$$

In comparing the matrix elements (6) and (7) with Heitler's formulas, it has to be remembered that Heitler writes all momenta multiplied by the velocity of light c. He also drops the normalization factor L^3 taking his box as being of unit volume. The integration extends over the box of volume L^3.

The matrix elements of the Coulomb potential are given by

$$V_{II,0} = \frac{1}{L^3} Ze^2 \int \frac{e^{-i\mathbf{k}''\cdot\mathbf{r}} \cdot e^{i\mathbf{k}^0\cdot\mathbf{r}}}{|\mathbf{r} - \mathbf{r}_c|} d^3r = \frac{Ze^2}{L^3} \frac{4\pi}{|\mathbf{k}^0 - \mathbf{k}''|^2} \tag{10}$$

$$V_{F,I} = \frac{1}{L^3} Ze^2 \int \frac{e^{-i\mathbf{k}^F\cdot\mathbf{r}}\, e^{i\mathbf{k}'\cdot\mathbf{r}}}{|\mathbf{r} - \mathbf{r}_c|} d^3r = \frac{Ze^2}{L^3} \frac{4\pi}{|\mathbf{k}' - \mathbf{k}^F|^2}. \tag{11}$$

The result of the integration holds in the limit $L \to \infty$ and can be found in Ref. 4. Z is the number of charges of the nucleus, and $\mathbf{r} - \mathbf{r}_c$ is the distance between the point of integration and the nucleus. The factor ρ_F in (1) is the number of final states per unit energy range of the electron after scattering. It is

$$\rho_F = \frac{mL^3k^F}{\hbar^2(2\pi)^3} d\Omega \tag{12}$$

with $\hbar k^F = mv_F$ the momentum of the electron after scattering and

$$d\Omega = \sin\psi \, d\psi \, d\alpha \tag{13}$$

the element of solid angle of the electron scattered in the direction ψ and α. The relative orientations of ψ and α and the angles θ and φ of the incident electron are shown in Fig. 1.

The form of the matrix elements (6) and (7) contains a physical approximation. We have limited ourselves to an expansion of the electron wave function in terms of plane electron waves, as is apparent from the factors $e^{\pm i\mathbf{k}\cdot\mathbf{r}}$ appearing under the integrals. Using plane waves to describe the electron corresponds to the Born approximation, which holds if

$$2\pi \frac{Ze^2}{\hbar v} \ll 1.$$

Our probability function w, of (1), describes the differential probability per unit time that an electron incident with angles θ and φ emits (or absorbs) a photon into the existing plane wave carrying n photons, and that the electron is found with energy

$$\tfrac{1}{2}mv_F^2 = \tfrac{1}{2}mv_0^2 \mp \hbar\omega \tag{14}$$

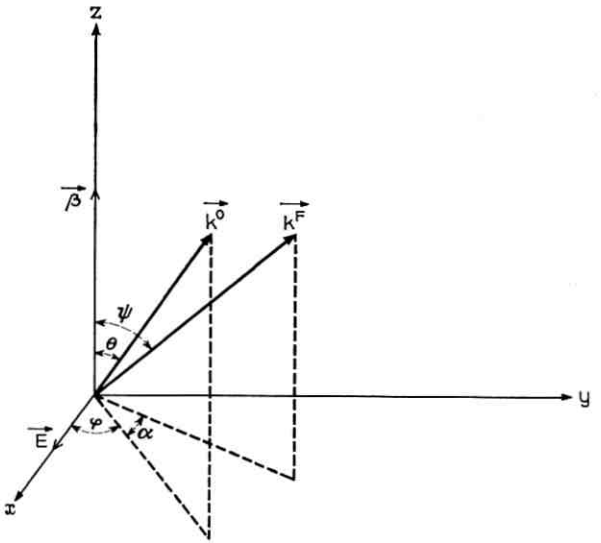


Fig. 1 — Relative orientation of the incident and scattered electron with respect to the direction of propagation and the electric vector of the radiation field.

traveling within the solid angle $d\Omega$ in the direction ψ,α. To obtain the total probability W per unit time regardless of the direction of the scattered electron, we have to integrate w over all directions of scattering.

Before proceeding further, we have to discuss the influence of the box normalization. The size of the box is arbitrary. The results become independent of the box if its sides L become infinitely long. It is apparent that as $L \to \infty$ we get $w \to 0$. To avoid this difficulty we consider that for a box of finite size we get as the number of emitted (or absorbed) photons per second

$$\Delta N = W N_e''$$

if N_e'' electrons are present in the box. Introducing the number N_e' of electrons per unit volume we get

$$\Delta N = W L^3 N_e'.$$

Calling N_e the number of incident electrons which per second fly through the unit area at speed v_o, we obtain

$$\Delta N = W \frac{L^3}{v_o} N_e = \sigma N_e. \tag{15}$$

The quantity σ defined by (15) is the scattering or interaction cross

section. This name is justified by the observation that σ has the dimension of an area.

The scattering cross section is, according to (15), defined by $\sigma = W(L^3/v_o)$. If we use the differential probability w of (1) instead of W, we obtain the differential scattering cross section

$$d\sigma = w\frac{L^3}{v_o}. \tag{16}$$

Using all the equations from (1) to (15), we obtain from (16) the differential cross section

$$d\sigma = \frac{8\pi e^6 Z^2 N k^F d\Omega}{m\omega v_o}\left\{\frac{k_x^o}{\left[\frac{p_o^2-p'^2}{2m} \mp \hbar\omega\right]\cdot\left[\mathbf{k}'-\mathbf{k}^F\right]^2} + \frac{k_x''}{\left[\frac{p_o^2-p''^2}{2m}\right]\cdot\left[\mathbf{k}^o-\mathbf{k}''\right]^2}\right\}^2. \tag{17}$$

The summation over all energy states reduces to one term because of the δ symbols in (6) and (7). We have taken

$$N = \frac{n}{L^3} \tag{18}$$

with N being the number of photons per unit volume. The normalization factor L has been eliminated from (17) and we can safely let $L \to \infty$. We see that the two factors $(n+1)/L^3$ for the emission case and n/L^3 for the absorption case become identical as $L \to \infty$. The term 1 in $n+1$ is related to the spontaneous emission of radiation. Since we calculate the transition probability (or scattering cross section) for emission of radiation into one well defined mode, without allowing for a spread in frequency or into several closely spaced modes traveling within a certain element of solid angle, the probability for spontaneous emission must be zero as $L \to \infty$.*

Only those terms give a contribution to the summation (2) for which

$$\mathbf{k}' = \mathbf{k}^o \mp \boldsymbol{\beta} \quad \text{or} \quad \mathbf{p}' = \mathbf{p}^o \mp \hbar\boldsymbol{\beta} \tag{19}$$

and

$$\mathbf{k}'' = \mathbf{k}^F \pm \boldsymbol{\beta} \quad \text{or} \quad \mathbf{p}'' = \mathbf{p}^F \pm \hbar\boldsymbol{\beta}. \tag{20}$$

* If the radiation is confined to a cavity, L^3 is the volume of the cavity and the factor $1/L^3$ does not go to zero. The 1 in $(n+1)/L^3$ gives rise to spontaneous emission of noise.

These relations, which follow from the δ-symbols in (6) and (7), mean that conservation of momentum is required even for the virtual states. We obtain with (19) and (20)

$$\mathbf{k}' - \mathbf{k}^F = \mathbf{k}^o - \mathbf{k}'' = \mathbf{k}^o - \mathbf{k}^F \mp \boldsymbol{\beta}. \tag{21}$$

With the help of Fig. 1 the following relation can be derived

$$\begin{aligned}[\mathbf{k}^o - \mathbf{k}^F \mp \boldsymbol{\beta}]^2 = {} & k^{o2} + k^{F2} + \beta^2 \mp 2k^o\beta\cos\theta \pm 2k^F\beta\cos\psi \\ & - 2k^ok^F(\cos\theta\cos\psi + \sin\theta\sin\psi\cos\alpha).\end{aligned} \tag{22}$$

Using (14) and (19) through (22) we obtain by integrating (20)

$$\begin{aligned}\sigma_{\substack{e\\a}} = {} & \frac{2\pi me^6Z^2N}{v_o\omega\beta^2k^3\hbar^4}\int_o^\pi d\psi\int_{-\pi}^{\pi}d\alpha\,\sin\psi\,\sqrt{1\mp2a\epsilon}\left[\frac{\sin\theta\cos\varphi}{a-\cos\theta\pm\frac{1}{2}\epsilon}\right. \\ & \left. + \frac{\sqrt{1\mp2a\epsilon}\,\sin\psi\cos(\varphi+\alpha)}{\sqrt{1\mp2a\epsilon}\cdot\cos\psi - a\pm\frac{1}{2}\epsilon}\right]^2 \\ & \cdot[1 \mp a\epsilon + \tfrac{1}{2}\epsilon^2 \mp \epsilon\cos\theta \pm \epsilon\sqrt{1\mp2a\epsilon}\cos\psi \\ & - \sqrt{1\mp2a\epsilon}\,(\cos\theta\cos\psi + \sin\theta\sin\psi\cos\alpha)]^{-2}.\end{aligned} \tag{23}$$

The upper subscript e of σ indicates emission and corresponds to the upper $(+)$ or $(-)$ signs in the equation. Conversely, the lower subscript a relates to absorption and corresponds to the lower signs.

We have used the abbreviations

$$\epsilon = \frac{\beta}{k^o} = \frac{\lambda_e}{\lambda} \tag{24}$$

with $\lambda_e = 2\pi/k^o$ being the wavelength of the electron wave and $\lambda = 2\pi/\beta$ the wavelength of the electromagnetic radiation.

$$a = \frac{c}{v_o} \quad \text{and} \quad a\epsilon = \tfrac{1}{2}\frac{\hbar\omega}{\frac{1}{2}mv_o^2}. \tag{25}$$

In going from emission to absorption, all we have to change is the sign of ϵ.

In order to decide if emission or absorption of radiation will actually occur, we have to take the difference

$$\sigma_T = \sigma_e - \sigma_a \tag{26}$$

between the cross sections for emission and absorption.

We do not have to evaluate the integral (23) exactly. In the range of physical interest we will always have $\epsilon \ll 1$ and also $a\epsilon \ll 1$ but $a \gg 1$

for nonrelativistic velocities. We limit ourselves to the case $a = c/v \gg 1$ and obtain approximately*

$$\sigma_{\substack{e \\ a}} = \frac{2\pi m e^6 N Z^2 v}{\omega\beta^2 k^3 \hbar^4 c^2} \int_0^{2\pi} d\alpha \int_0^{\pi} d\psi \frac{\sin\psi \sqrt{1 \mp 2a\epsilon}\,[\sin\theta\cos\varphi - \sqrt{1 \mp 2a\epsilon}\sin\psi\cos(\alpha+\varphi)]^2}{[1 \mp a\epsilon - \sqrt{1 \mp 2a\epsilon}(\cos\theta\cos\psi + \sin\theta\sin\psi\cos\alpha)]^2} \tag{27}$$

The double integral can be solved. The solutions have been expanded in terms of $a\epsilon$, and the difference between σ_e and σ_a has been taken.

We obtain to the lowest nonvanishing order of $a\epsilon = hf/mv^2$

$$\sigma_T = \frac{4e^6 N Z^2}{m^3 v^4 f^2}\left[(3\cos^2\varphi\sin^2\theta - 1)\ln\frac{2mv^2}{hf} - 2\cos^2\varphi\sin^2\theta\right]. \tag{28}$$

If $\varphi = 0$ or π and if $\theta = \pi/2$, the cross section σ_T assumes its largest positive value

$$\sigma_T = \frac{8e^6 N Z^2}{m^3 v^4 f^2}\left[\ln\frac{2mv^2}{hf} - 1\right]. \tag{29}$$

The direction of the incident electron defined by $\varphi = 0, \pi$, and $\theta = \pi/2$ is parallel to the electric vector of the radiation field (Fig. 1). Electron incidence parallel to the electric vector of the radiation field gives rise to stimulated emission of radiation.

Equation (28) shows that $\sigma_T < 0$ if $\varphi = \pi/2$. The electron is incident perpendicular to the radiation field and absorbs power from the field.

Let us take another look at (28) and determine the directions of incidence for which stimulated emission rather than absorption results.

Since an exact solution is hard to give, we will assume that the logarithmic factor in (28) is considerably larger than the remaining factor in the bracket consistent with the assumption $mv^2/hf \gg 1$.

We can then give approximate angles for which $\sigma_T = 0$.

If we let $\varphi = 0$, we obtain

$$\frac{\pi}{2} - \theta = 0.954 - \frac{0.236}{\ln\dfrac{2mv^2}{hf}}. \tag{30}$$

The angle 0.954 in radians corresponds to 54.6°.

* Terms with ϵ (not $a\epsilon$) and terms of order $1/a$ are neglected. Terms with $a\epsilon$ have to be kept because the difference $\sigma_e - \sigma_a$ is proportional to $a\epsilon$.

If we let $\theta = \pi/2$ we obtain

$$\varphi = 0.954 - \frac{0.236}{\ln \frac{2mv^2}{hf}}. \tag{31}$$

All electrons incident on the ion within a cone with a cone angle of approximately 108°, whose axis is parallel to the electric vector of the stimulating radiation field, contribute to stimulated emission.

The number ΔN of emitted photons per second is obtained if we multiply σ_T by the number N_e of electrons which per second penetrate the unit area containing the nucleus with the charge Ze. If there is more than one nucleus interacting with the electron stream, we obtain the total number of emitted photons per second by multiplying the total number of nuclei N_n with the electron flux density N_e which per second interacts with them, provided the density of nuclei is low. For electrons incident parallel to the electric field vector $\varphi = 0$ or π, and $\theta = \pi/2$ we obtain from (29)

$$\frac{\Delta N}{N} = \frac{8e^6Z^2N_eN_n}{m^3v^4f^2}\left[\ln \frac{2mv^2}{hf} - 1\right]. \tag{32}$$

III. DISCUSSION

Equation (32) shows that the ratio of emitted photons per second to the stimulating photon density decreases rapidly as the electron velocity v increases. Within the limits set by $(mv^2/hf) \gg 1$ and by $2\pi(Ze^2/\hbar v) \ll 1$ we obtain more emission with slower electrons. The number of emitted photons increases also with decreasing frequency.

It may be useful to remark on the coherence of the process. We have to keep in mind that (32) gives the number of photons added precisely to the radiation mode which stimulates the emission. Let us assume that all other possible radiation modes are empty, $N = 0$. If we ask for the number of photons emitted into one of the empty modes, our results would be proportional to $1/L^3$, but it is proportional to $(hf/mv^2)\cdot(n/L^3)$ in case of emission into the radiation mode filled with n photons. (Remember the remark about $(n + 1)/L^3$ following (18). The factor hf/mv^2 arises by taking the difference $\sigma_e - \sigma_a$. The zero-order term drops out and the result is proportional to $a\epsilon = hf/mv^2$.) The radiation from the electron occurs preferentially into the occupied mode with a probability ratio equal to $[(hf)/(mv^2)]n$ (n is a big number even for moderate power). The power added to the occupied radiation mode is coherent with the radiation present in that mode because each photon adds precisely the

energy hf. We have here the same situation which exists for stimulated emission from bound states, with the difference that the spontaneous emission of bound electrons has a well defined frequency while the free electron radiates into a very wide frequency band. Because of the ratio of spontaneous to stimulated emission, we would assume that the process discussed here is inherently noisier than stimulated emission from bound states. For the latter, the ratio of stimulated to spontaneous emission into the occupied mode is n while it is only $[(hf)/(mv^2)]n$ in our case. Apparently we get better results for slow electrons.

We want to state the oscillating condition for achieving self-sustained oscillation if the radiation is confined in a cavity. The cavity has a loaded Q defined by

$$Q_L = \omega \frac{NVhf}{n'hf} = \frac{\omega NV}{n'} \tag{33}$$

with cavity volume V, photon density N, and n' dissipated photons per second. The cavity oscillates if $n' = \Delta N$. With the help of (32), this leads to an expression for the minimum product necessary to achieve oscillation if we assume again that the electrons are incident parallel to the **E**-vector.

$$N_e N_n = \frac{\pi m^3 v^4 f^3 V}{4e^6 Z^2 Q_L \left[\left(\ln \frac{2mv^2}{hf}\right) - 1\right]}. \tag{34}$$

The oscillating frequency is determined only by the cavity. If many cavity modes can exist, oscillation will start in the mode with the smallest value of f^3/Q_L.

For use as an amplifier the bandwidth could be much larger than the bandwidth of conventional masers because there is no built-in resonance in this process to limit the frequency.

To build an amplifier or oscillator using stimulated emission from free electrons, we want to shoot dense electron beams of low energy through as dense an ensemble of ions as possible. This can be done by using ion beams rather than ions in a plasma because in a plasma scattering of the thermal plasma electrons by the plasma ions would have an adverse effect.

It is also possible to use electron scattering due to ionized impurities in a crystal to obtain stimulated emission. To apply our theory to electrons moving in conduction bands of crystals, we have to replace the electron mass m by the effective mass m_{eff} of the electron in the crystal. Furthermore, we have to multiply the scattering cross section σ by $1/\epsilon^3$ with ϵ being the dielectric constant of the crystal.

Experimentally, one has to take care that the conduction electrons move predominantly in the direction of the electric vector of the radiation field. This is achieved, for example, by lifting electrons into the conduction band with the help of the photoelectric effect using a crystal which in the dark is an insulator. The electrons will move predominantly in the direction of the electric vector of the pump light provided that this is polarized.

Another way of generating electrons moving in preferred directions is by the use of dc electric fields which exceed the breakdown voltage of an insulating or poorly conducting crystal.

The crystal will have to be cooled to increase the collision time for electron collisions with the vibrating lattice which have an adverse effect on the ordered electron motion and may not give rise to stimulated emission.*

Another way to achieve stimulated emission of Bremsstrahlung is the use of crossed electron and ion beams. These two beams can be made to cross in a capacitor which is part of an LC resonant circuit. The use of beams in vacuum will limit the application of our effect to low-frequency amplifiers and oscillators because of the limitation in available ion densities.

Equation (34) can be applied to the case of an LC circuit. In this case, V is the volume of the capacitor and Q_L the loaded Q of the LC circuit. An independent calculation has shown that (34) can be obtained if the voltage and current in the LC circuit rather than the electric and magnetic field in the capacitor are quantized. The photons in this case are the units of energy hf which are stored in the resonant circuit.

IV. NUMERICAL EXAMPLES

i. We assume an LC circuit with $Q_L = 30$ tuned to 70 mc, a capacitor volume $V = 12.5$ cm^3. The electrons are assumed to be accelerated by a potential of 10 volts, corresponding to a velocity of 2×10^8 cm/sec and to drift through the capacitor plates, which are made out of a wire mesh. To obtain oscillations, we obtain from (34) ($Z = 1$ is assumed)

$$N_e N_i = 4.6 \times 10^{29} \text{ sec}^{-1} \text{ cm}^{-2}.$$

If we assume that an electron beam with 10 ma/cm^2 current density, corresponding to $N_e = 6.3 \times 10^{16}$ sec^{-1} cm^{-2}, is employed, $N_n = 7.3 \times 10^{12}$ ions are needed in the interaction region to make the circuit oscillate.

* So far we have studied stimulated emission in the presence of Coulomb potentials. Whether other scattering potentials can be used will have to be determined by another study.

Assuming that the capacitor is made of a mesh of 5×5 cm^2 with a spacing of 0.5 cm, and assuming further that the ions are accelerated by a 10-volt potential, we find that the ion current of 54 ma has to pass through the capacitor (Cs ions assumed) in order to provide $N_n = 7.3 \times 10^{12}$ ions in the capacitor at any instant of time.

ii. We consider a solid with a dielectric constant $\epsilon = 10$, an effective electron mass $m_{\text{eff}} = 0.1\ m$ and assume that the material is contained in a cavity with volume $V = 20$ cm^3 which is resonant at 10 kmc with a $Q_L = 10{,}000$. The electron velocity is assumed to be $v = 10^7$ cm/sec. We obtain, taking $Z = 1$,

$$N_e N_n = 2.8 \times 10^{35} \text{ sec}^{-1} \text{ cm}^{-2}.$$

If we take the electron current density equal to the one used in the first example, $N_e = 6.3 \times 10^{16}$ sec^{-1} cm^{-2}, we obtain as the number of ions in the interaction region necessary to sustain oscillations, $N_n = 4.45 \times 10^{18}$ or a density of 2.2×10^{17} ions/cm^3.

The frequency could be increased further if it were possible to increase the ion and electron densities. Increasing the frequency to 100 kmc increases the product $N_e N_n$ by a factor of 1000. It may be possible to increase N_e as well as N_n by a factor of 30 and push the operating frequency to 100 kmc. Increasing the Q of the cavity would also help.

V. A COMPARISON WITH CLASSICAL THEORY

It is interesting to assume that electrons are incident from all possible directions with a uniform distribution over all angles of incidence.

We obtain from (28)

$$\bar{\sigma} = \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} d\varphi \sigma_T = -\frac{32\pi e^6 N Z^2}{3m^3 v^4 f^2}. \tag{35}$$

Electrons incident with a random distribution over all directions lead to a net absorption of radiation. This absorption process gives rise to attenuation of waves traveling through plasmas or semiconductors. However, in a semiconductor our process gives only the contribution of impurity scattering but not of lattice scattering. Equation (35) does not contain h any more and must be equivalent to classical theories.

Equation (35) could be used to derive an expression for the attenuation constant α in an ionized gas. However, the approximations implied in this work do not allow the electron velocity to become arbitrarily small. We will, nevertheless, use (35) to derive the attenuation constant per centimeter of a plane wave in a plasma with electron density n_e and

ion density n_i with the assumption that the electrons obey the Maxwell-Boltzmann velocity distribution.

$$f(v) = 4\pi v^2 \left[\frac{m}{2\pi kT}\right]^{3/2} \exp\left(-\frac{mv^2}{2kT}\right). \tag{36}$$

It is

$$\alpha = -\frac{\Delta N}{Nc} \tag{37}$$

where ΔN is the number of photons created per centimeter per second, c is the velocity of light, and N is the photon density.

We obtain with

$$\Delta N = \frac{n_i n_e}{4\pi} \int_0^\infty vf(v)\, \bar{\sigma}\, dv \tag{38}$$

from (37) with the help of (35), (36) and (38)

$$\alpha = \left\{\frac{4}{3}\sqrt{\frac{2\pi}{3}} \frac{1}{(kT)^{3/2}} \frac{n_i n_e e^6 Z^2}{m^{3/2} f^2 c}\right\} \frac{4\sqrt{3}}{2\pi} \int_0^\infty \frac{1}{x} e^{-x^2}\, dx. \tag{39}$$

The integral in (39) cannot be evaluated since it requires an integration over

$$x = v\sqrt{\frac{m}{2kT}}$$

starting from zero. The integral is logarithmically divergent. However, our theory does not allow us to apply (35) for arbitrarily small electron velocities. Nevertheless, it is interesting to note that the part of (39) inside the brackets equals exactly equation (5-48) of Ref. 5 if we follow Spitzer's indication and multiply his formula with

$$1 - e^{\frac{hf}{kT}} \approx \frac{hf}{kT}$$

to take stimulated emission into account. The divergent factor outside the brackets should be one. We cannot derive its value because of the limitations of our approximation.

VI. SUMMARY

We have presented an approximate theory of stimulated emission of Bremsstrahlung. The theory considers the process of stimulated emission which occurs if a stream of electrons is scattered by one individual ion

or nucleus in the presence of a radiation field. The approximations require that

$$\frac{c}{v} \gg 1, \quad \frac{hf}{mv^2} \ll 1, \quad \text{and} \quad 2\pi \frac{Ze^2}{\hbar v} \ll 1.$$

We have found that stimulated emission occurs if the directions of the incident electrons lie within a cone whose axis is parallel to the electric vector of the stimulating radiation and whose cone angle is approximately 108°.

The effect is fairly weak, so that rather high ion and electron densities are required to achieve substantial amplification or oscillation.

It is predicted that the effect will work best in a solid when the conduction electrons are scattered by ionized impurities.

It may be that this effect is responsible for the recently reported oscillations which were obtained with the use of semiconductor diodes and which seem presently to be unexplained.[6,7,8]

It is possible that other scattering potentials may lead to stimulated emission. We have restricted ourselves to Coulomb scattering. It may be worthwhile to study other scattering potentials, for example electron scattering by the vibrating lattice in semiconductors.

VII. LIST OF SYMBOLS

α	used as azimuth of the scattered electron relative to the incident electron; also attenuation constant
$\beta = \omega/c$	propagation constant of electromagnetic wave
c	phase velocity of electromagnetic wave
$e = 4.803 \times 10^{-10}$ e.s.u.	charge of the electron
f	frequency of the electromagnetic wave
φ	azimuth of the incident electron
$h = 6.624 \times 10^{-27}$ erg·sec.	Planck's constant
$\hbar = h/2\pi = 1.054 \times 10^{-27}$ erg·sec.	
H_{10}	matrix element of radiation field
k	propagation constant of electron wave; also Boltzmann's constant $= 1.38 \times 10^{-16}$ erg·degree^{-1} (appears only in the combination kT)
L	length of fictitious box used for box normalization

$m = 9.107 \times 10^{-28}$ gram	mass of the electron
n_i	ion density
n_e	electron density
N_e	number of electrons penetrating the unit area per unit time
N_n	number of ions or nuclei
N	photon density
Ω	solid angle
$\omega = 2\pi f$	angular frequency of the electromagnetic wave
$\mathbf{p} = \hbar\mathbf{k}$	momentum of the electron
ψ	polar angle of the scattered electron
Q_L	loaded Q of the resonant cavity
ρ_F	density of physical states per unit energy
σ	interaction cross section
T	absolute temperature
θ	polar angle of the incident electron
v	electron velocity
V_{F0}	matrix element of the Coulomb potential
w	differential transition probability per unit time
W	transition probability per unit time
Z	number of elementary charges of the ion.

REFERENCES

1. Heitler, W., *The Quantum Theory of Radiation*, 3rd Ed., Oxford, 1954.
2. Ref. 1, p. 243, equation (10).
3. Ref. 1, p. 143, equations (23a) and (23b).
4. Mandl, F., *Introduction to Quantum Field Theory*, Inter-Science Publishers, Inc., 1959, p. 119, equations (16.5), (16.6) and (16.16).
5. Spitzer, L., *Physics of Fully Ionized Gases*, Inter-Science Publishers, Inc., New York, 1956.
6. Kikuchi, M., and Abe, Y., Observation of Coherent Oscillation in Germanium, J. Phys. Soc., Japan, **17**, 1962, pp. 241–242.
7. Kikuchi, M., Coherent Oscillation from Forward-Biased Point Contact on Silicon Single Crystals, J. Phys. Soc., Japan, **17**, 1962, p. 240.
8. Larrabee, R. D., and Stelle, M. C., The Oscillistor, J. Appl. Phys., **31**, 1960, pp. 1519–1523.

Diffusion Length Measurement by Means of Ionizing Radiation

By W. ROSENZWEIG

(Manuscript received June 6, 1962)

Penetrating radiation in the form of high-energy electrons, heavy particles, and gamma rays may be used to determine minority-carrier diffusion lengths in semiconductor materials containing junctions by measuring the radiation-induced short circuit current. The electron-beam method yields an accurate absolute determination of diffusion length once the carrier-generation rate as a function of depth in the material has been measured. A series of such experiments is described for silicon solar cells utilizing electrons ranging in energy from 0.61 Mev to 1.16 Mev. A resultant maximum generation rate of 2.25 $\times$ 10^6 $\pm$ 5 per cent carriers/cm per incident 1 Mev electron is obtained at a depth of 0.096 gm/cm^2. Measurements with 16.8 Mev and 130 Mev protons, and Co^{60} gamma rays are found to be in good agreement with the electron-beam measurements. An experimental arrangement is described which yields rapid and accurate diffusion-length measurements of solar cells under conditions in which radiation damage is negligible.

I. INTRODUCTION

The minority-carrier diffusion length is an important quantity characterizing the properties of a semiconductor material or device, particularly in any situation involving the transport of charge by minority carriers. It is related to the steady-state minority carrier lifetime through the expression $L = \sqrt{D\tau}$, in which D is the diffusion coefficient and τ is the lifetime. A variety of methods have been employed to measure diffusion length in the steady state, or lifetime under transient conditions and are reviewed by Bemski.[1] The steady-state methods have a particular advantage in that the measurement is uncomplicated by trapping effects. One particular steady-state method suggested by Gremmelmaier[2] involves the use of gamma radiation to generate excess carriers at a known rate in a pn junction device. The radiation-generated

short-circuit current may then be used to calculate diffusion lengths when the device geometry is known.

It is the purpose of this paper to discuss, more generally, the use of penetrating ionizing radiations for diffusion-length measurement. Experimental results will be presented for energetic electrons and protons as well as gamma rays. It will be shown that of the three, electrons lend themselves readily to an accurate absolute determination, with an accuracy of ± 5 per cent having been achieved. Even though the method is quite general it will be treated in terms of one device, the silicon solar cell. It is the radiation damage study of this device which led to the development of the technique.

In the text which follows, Section II will be devoted to the theoretical background for the diffusion-length determination. Section III will be a description of the energetic electron measurements and the analysis which yields the absolute calibration. Sections IV and V will deal with protons and gamma rays, respectively, and give experimental results which are compared with the electron beam measurements. A brief discussion of the radiation damage problem will be given in Section VI along with some quantitative information which shows that measurements can be made in most practical cases with negligible damage.

II. THEORY

Radiation incident on the surface of a material having a shallow diffused junction, such as a solar cell shown schematically in Fig. 1, will generate excess carriers at a rate which is a function of the depth from the surface. The function depends on the nature of the radiation. That portion of the carriers which is produced deep in the bulk of the material can be collected only if the carriers diffuse to the junction. Thus, for deeply generated carriers, the bulk diffusion length is of primary importance in determining the collection efficiency. In fact, because of the very shallow junction in solar cells, the diffusion length is the only parameter which governs the collection efficiency for penetrating radiation, the front layer becoming important only for nonpenetrating radiation. For the case in which excess carriers are generated uniformly throughout the material at a rate of g_0 carriers per cm^3 per second, the collected current density, J_{sc}, under short circuit conditions is given by

$$J_{sc} = eg_0L \tag{1}$$

where e is the electronic charge. (Derivation can be found in the Ap-

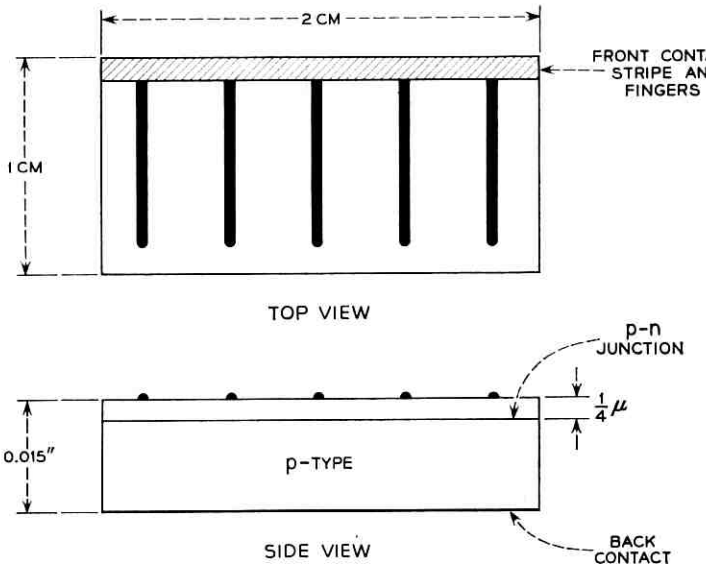


Fig. 1 — Schematic representation of the silicon solar cell.

pendix.) Equation (1) indicates that the quantitative determination of L requires a measurement of current, under short circuit conditions, when a known area of the device is "illuminated" with radiation which is capable of producing excess carriers uniformly throughout the bulk of the material and at a *known* rate. Light which has been filtered with material which is identical to that of which the cell is made (silicon in this instance) would satisfy the requirement of uniform generation rate. However, the requirement of knowing quantitatively the generation rate is difficult to meet, because the amount of reflection of the incident light depends on the condition of the surface, and the generation rate due to the transmitted beam depends strongly on the exact width of the energy band gap of the filter relative to that of the device.

The types of radiation which satisfy both needs are energetic nuclear particles, electrons, and gamma rays. As these radiations penetrate the material, a major portion of their energy is spent in causing transitions of valence electrons to the conduction band, i.e., the generation of hole-electron pairs. A small fraction of the energy produces damage by creating lattice vacancies and interstitials through nuclear collisions resulting in a reduction of carrier lifetime and change in carrier concentration and mobility. The rate at which silicon is damaged by the various types of radiation will be briefly discussed in Section VI with

the conditions necessary to produce negligible damage during a measurement of diffusion length.

The average amount of energy required to produce a hole-electron pair in silicon has been measured to be 3.6 ev, a number which appears to hold well over a wide range of particle energy and types.[3] The determination of g_0 for a particular case thus requires a knowledge of the rate at which energy is given up to the material. Moreover, to assure oneself of the uniformity of the generation rate, it is also necessary to know its rate of change with depth into the material. If the latter quantity be designated by g_1, then it is shown in the Appendix that the nonuniformity can be neglected to the extent that $(g_1/g_0)L$ is small compared to one.

For the cases of the charged particles it is convenient to introduce one other quantity, the average specific ionization, which is the average spatial rate of hole-electron pair production per incident particle. If the generation rate, at some depth in the material, is g, due to N charged particles normally incident per second per cm^2 of surface, then the average specific ionization, S, at that depth is given by

$$S = \frac{g}{N} \text{ pairs/cm.} \tag{2}$$

For the situation of uniform carrier generation, (1) may be rewritten in the form

$$J_{sc} = JS_0L \tag{3}$$

where J is the incident current density for singly charged particles, and S_0 is the uniform specific ionization.

III. ELECTRONS

Of the charged particle radiations, electrons can be handled most easily experimentally. This is due to the fact that accelerators for 1 Mev electrons, which are sufficiently penetrating for the purpose at hand, are relatively simple in design and can be operated quite reliably. Beam intensity calibrations can be carried out accurately, for electrons, with Faraday cups of not too complicated design. This situation is in contrast to heavy particle accelerators which are large and complex installations.

The theoretical calculation of the average specific ionization as a function of depth per incident electron is difficult.[4] However, it is easy to evaluate this function experimentally, as was done, in the following way: A pre-bombarded silicon solar cell was exposed to a broad beam of

monoenergetic electrons normally incident to its surface. The reason for the pre-bombardment was two-fold: (*i*) the diffusion length of the minority carriers was reduced to only a few microns so that the spatial variation of ionization in the sensitive layer was negligible, and (*ii*) the diffusion length did not change appreciably during the measurements. The solar cell short-circuit current density, J_{sc}, was measured and the electron beam calibrated to obtain the incident electron current density, J_e. By forming the ratio J_{sc}/J_e one has determined the average ionization per incident electron in the spatial interval a distance L from the junction. This quantity was then obtained as a function of depth by placing absorbers, approximately equivalent in atomic number to the cell material, on top of the cell and measuring the resultant short circuit current. The placing of the absorbers on the cell was equivalent to moving the sensitive layer, in which the ionization was measured, through the volume of the material. Care was taken in this procedure that the absorbers overlapped the edges of the cell sufficiently so that electrons scattered out at the periphery of the cell were compensated by equivalent electrons scattered in.

The results of a series of such experiments are shown in Figs. 2–5 for

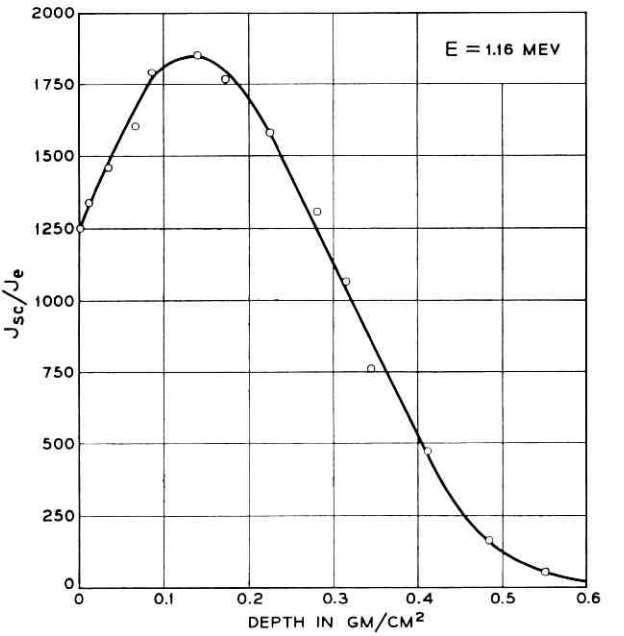


Fig. 2 — Ratio of collected current density to incident electron current density vs depth for 1.16 Mev electrons in a silicon solar cell of 8.75 μ diffusion length.

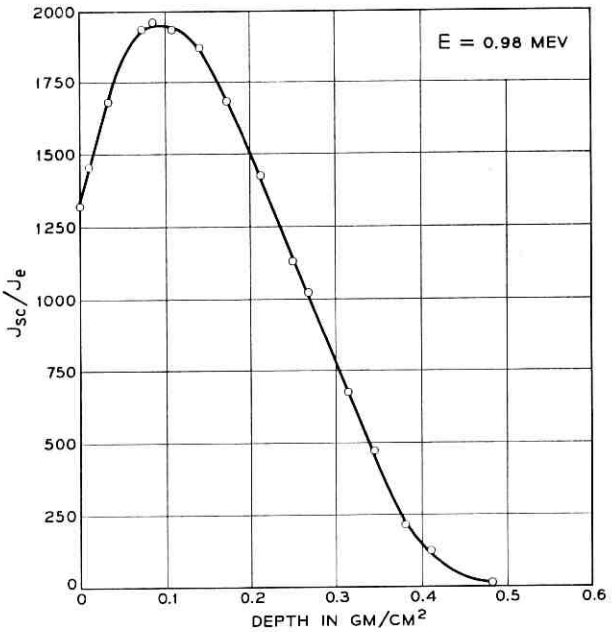


Fig. 3 — Ratio of collected current density to incident electron current density vs depth for 0.98 Mev electrons in a silicon solar cell of 8.75 μ diffusion length.

incident electrons of 1.16 Mev, 0.98 Mev, 0.80 Mev and 0.61 Mev energies respectively. They are plots of the ratio J_{sc}/J_e versus depth in silicon. (The actual absorbers used were made of aluminum, which is sufficiently close in atomic number to silicon to require only a density correction to yield equivalent thickness of silicon.) Since the integral over-all depths of the average specific ionization, S_e , per incident electron is equal to the total ionization which the electron produces in the silicon, then by means of (3) one gets the following result:

$$\int_0^\infty S_e \, dx = \frac{1}{L} \int_0^\infty \frac{J_{sc}}{J_e} \, dx = \frac{E}{\epsilon} \tag{4}$$

where E is the average energy absorbed in the silicon per incident electron and $\epsilon = 3.6$ ev/pair.

This average absorbed energy is very nearly equal to the energy of the incident electron. The latter quantity was measured with a combination vacuum Faraday cup-calorimeter, under scattering conditions of 2 mils of aluminum and 2 inches of air. The absorption measurements were made with additional scattering of 3.5 mils of aluminum and 3

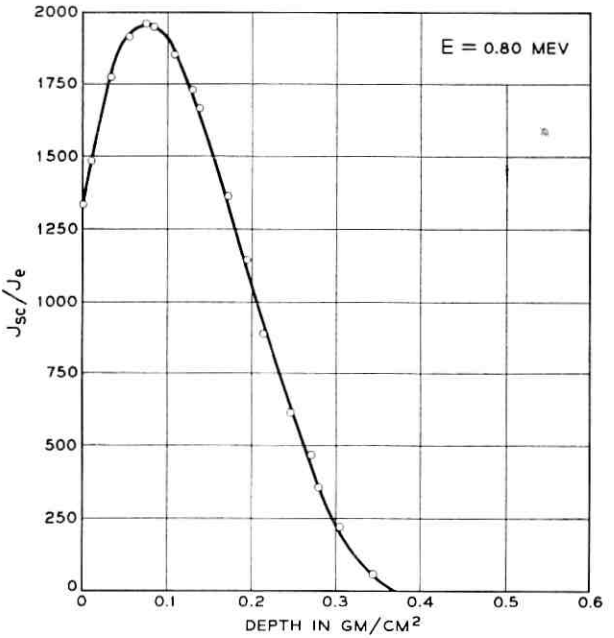


Fig. 4 — Ratio of collected current density to incident electron current density vs depth for 0.80 Mev electrons in a silicon solar cell of 8.75 μ diffusion length.

inches of air. The extrapolated ranges obtained from these curves, when corrected for the additional air and aluminum, give initial electron energies, through the range-energy relationship,[5] which agree with the calorimeter measurements within 3 per cent. Using the energies obtained by the extrapolated ranges and evaluating the integrals under the smooth curves, one obtains a set of L values for the cell used in these measurements as in Table I under the heading Uncorrected values. One observes that the calculated L values decrease with decreasing energy. The cause for this is not clearly understood, but it is possible to

TABLE I — DIFFUSION LENGTH VALUES COMPUTED FROM ABSORPTION CURVES FOR ELECTRONS OF VARIOUS ENERGIES

Electron Energy (Mev)	Uncorrected $L(\mu)$	Backscatter Corrected $L(\mu)$
1.16	8.34	8.54
0.98	8.25	8.48
0.80	7.84	8.15
0.61	7.64	8.04

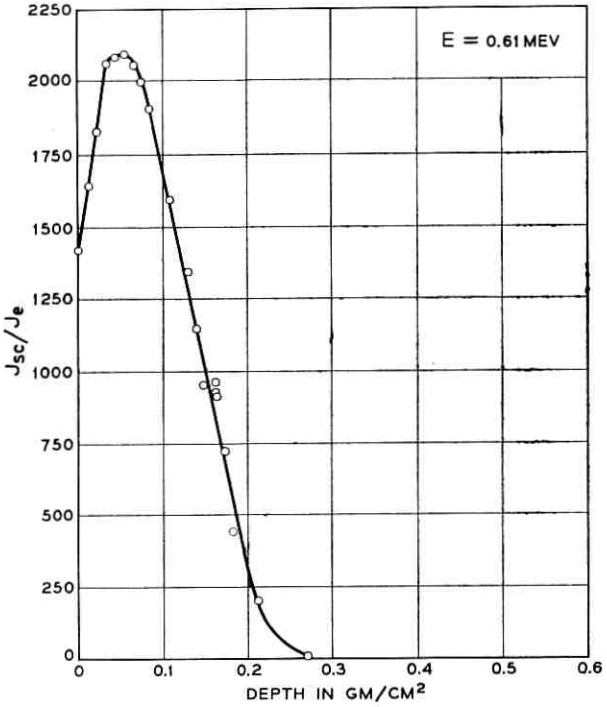


Fig. 5 — Ratio of collected current density to incident electron current density vs depth for 0.61 Mev electrons in a silicon solar cell of 8.75 μ diffusion length.

enumerate two contributing factors. The average energy of the incident electrons is less than that determined by the calorimeter under less severe scattering conditions. Similarly, the extrapolated range is only a measure of the range of the most energetic components, the less energetic ones manifesting themselves only in the shallower portion of the absorption curve. Secondly, a certain fraction of the incident energy flux is reflected back by way of secondary electron emission. The latter effect can be taken into account quantitatively by making use of some recently published data by Wright and Trump.[6] They give curves for the fraction of the incident electron beam energy which is scattered backward as a function of energy. When one makes use of the curve for aluminum, one arrives at L values shown in Table I under the heading Backscatter Corrected values. The variation with energy has been reduced but the trend still remains. If one makes the assumption that the correct L value is approached as the electron energy goes to infinity then one can attempt to obtain this value by plotting L against the

reciprocal of the energy and extrapolating to zero. Such an extrapolation yields $L = 9.0\ \mu$ which is about 5 per cent higher than the value at 0.98 Mev. If one now takes a value halfway between these two, namely 8.75 μ, any remaining uncertainty will introduce errors of less than ± 3 per cent.

With the L value of the experimental cell thus determined, it is possible to normalize the ordinates of the curves in Figs. 2–5 to give the average specific ionization as a function of depth for the various initial electron energies. For example, 0.98 Mev electrons have a maximum average specific ionization in silicon of 225 pairs/μ and, for the scattering conditions of this experiment, have a surface ionization of 152 pairs/μ. The initial slope of the curve is 2.85 pairs/μ^2 so that for a cell of 100 μ diffusion length, $g_1 L/g_0 = 0.3$ (see Appendix).

The above nonuniformity in the carrier generation rate is large enough to make it desirable to make diffusion length measurements in the vicinity of the maximum where the variation in the generation rate is less than 2 per cent over a range of 250 μ. An experimental arrangement for readily doing this has been developed and takes the form of a double-aperture Faraday cup, which is shown schematically in Fig. 6. The cup is placed in the 1 Mev electron beam in a position such that

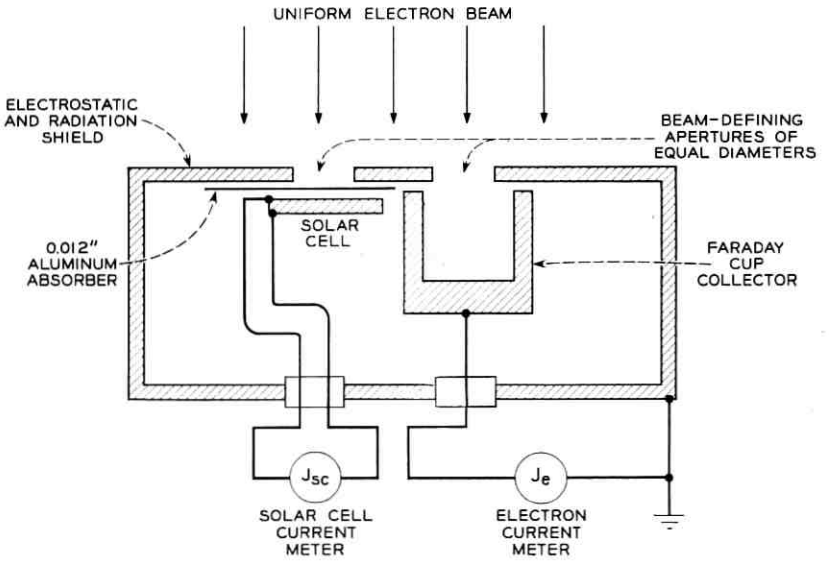


Fig. 6 — Schematic representation of the double-aperture vacuum Faraday cup for measuring J_{sc}/J_e.

equal electron fluxes enter both apertures. This is assured by making repeat measurements with the cup rotated 180° around its own axis. Under one aperture is a 12 mil aluminum foil, bringing the carrier generation rate near the maximum in silicon, which is followed by the cell whose diffusion length is to be measured. The beam entering the other aperture is collected in a Faraday cup. The ratio of the cell short circuit current to the cup current when divided by 225 will give the diffusion length in microns. Such measurements are carried out with sufficiently sensitive amplifiers so that an L determination can be made with an exposure of 10^{10} electrons/cm^2. The ratio is formed electronically and is displayed in digital form. The measurements are reproducible to within ± 3 per cent.

IV. HEAVY PARTICLES

For the case of energetic protons, deuterons, and alpha particles the situation is fairly simple. Since the particle mass is much greater than the electron mass, the incident particle passes through the material with only negligible deflection except for the extremely rare instance in which it undergoes a large angle nuclear collision. At any depth in the material, the specific ionization may be obtained by dividing the rate of energy loss, dE/dx, by the average energy required to produce a pair; i.e., 3.6 ev for silicon.

As an example, Fig. 7 is a plot of the specific ionization and its spatial rate of change in silicon as a function of proton energy. The values were computed from a tabulation of dE/dx versus energy in aluminum as given by Sternheimer.[7] These curves show that for minority carrier diffusion lengths of the order of 100 microns, a proton of initial energy greater than about 18 Mev will generate carriers uniformly to within 2 per cent in a depth of one diffusion length from the surface. This implies a 2 per cent accuracy in (1) when a uniform generation rate of g_c is used. The uniformity of carrier generation is improved with increased proton energy.

Similar considerations may be applied to other heavy nuclear particles. For quantitative information on dE/dx and range-energy relationships, references such as Bethe and Ashkin[8] or Allison and Warshaw[9] are useful.

Diffusion length measurements were made with 16.8 Mev protons (Princeton Cyclotron) and 130 Mev protons (Harvard Cyclotron). The ionization currents were measured on cells of 1 cm^2 area and the beam current densities measured with Faraday cups available at each of the two installations. When (3) is solved for L with $S_0 = 1.46 \times 10^7$ pairs/

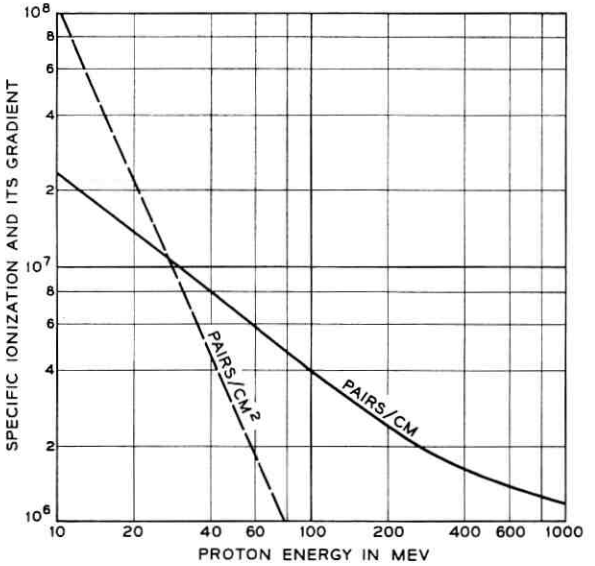


Fig. 7 — Specific ionization and its spatial rate of change vs energy for protons in silicon.

cm at 16.8 Mev and $S_0 = 3.03 \times 10^6$ pairs/cm at 130 Mev, the results in Table II are obtained. For comparison, the table also contains the diffusion length measurements obtained with the electron beam. As may be seen, agreement is obtained between the two methods for a large range of diffusion length values for both conductivity types and at two greatly different proton energies.

V. GAMMA- AND X-RAYS

Measurement with this type of radiation is discussed in the previously mentioned article by Gremmelmaier.[2] He shows that to a good degree of approximation,

$$g_0 = N_\gamma K \frac{\bar{E}}{\epsilon} \tag{5}$$

where

N_γ = incident radiation flux in quanta per cm²-sec,
K = the absorption coefficient,
$\bar{E}$ = the average electron energy per absorbed quantum,
ϵ = average energy per pair (3.6 ev for silicon).

TABLE II — DIFFUSION LENGTH MEASUREMENTS OBTAINED BY MEANS OF PROTONS COMPARED WITH 1 MEV ELECTRON MEASUREMENTS

Cell Type	Cell No.	Proton Energy (Mev)	$L(\mu)$ by Protons	$L(\mu)$ by Electrons
n on p	211	16.8	21.0	21.5
n on p	622	16.8	86.0	88.0
n on p	1112	16.8	29.3	32.5
n on p	1021	16.8	13.8	11.5
p on n	5011	16.8	5.2	4.8
p on n	5021	16.8	22.4	19.5
p on n	280R	16.8	8.7	7.9
p on n	283R	16.8	5.6	3.9
n on p	8112	130	5.4	5.5
n on p	10802	130	5.4	5.35
n on p	B-1	16.8	41.2	40.4
n on p	6112	16.8	12.0	11.5
n on p	C-2	16.8	20.4	21.2
n on p	D-2	16.8	8.2	7.8
n on p	7112	16.8	2.3	2.3
n on p	7111	16.8	2.14	2.36
n on p	6911	16.8	2.2	2.4
n on p	7012	16.8	2.5	2.4

For a given quantum energy, K and $\bar{E}$ can be calculated from tabulated absorption and scattering cross sections.[10] However, the determination of N_γ for a given source and particular geometry is quite difficult.

The use of this radiation does offer definite advantages. The rate of degradation by the radiation is substantially lower than for the case of energetic particles since the secondary electron flux is weighted toward the low energy end, with a sizable portion below the damage threshold energy. It is highly penetrating, so that measurements can be made through appreciable thicknesses of absorber. For the same reason, the carrier generation rate is uniform to great depth. The advantages of a gamma- or X-ray source can be utilized without the evaluation of the various constants in (5) by calibrating with a cell of known diffusion length, the latter quantity having been determined with electrons.

Solar cell short circuit currents were measured in a 10 kilocurie Co^{60} source. The radiation level within the source was about 0.87×10^6 r/hr. A number of cells whose diffusion lengths had previously been determined using the electron beam were placed inside the source and their short circuit currents measured. The ratio of current to diffusion length for each of the cells is shown in Table III. As may be seen, the ratio undergoes variations up to 20 per cent. This is caused primarily because the cells did not occupy positions of equal intensity in the γ-ray source. It is thus reasonable to conclude that there is good agreement between a diffusion length measurement in the γ-ray source and under

TABLE III — GAMMA RADIATION INDUCED SHORT CIRCUIT CURRENT NORMALIZED TO 1 MEV ELECTRON BEAM DETERMINED DIFFUSION LENGTH

Cell Type	Cell No.	$L(\mu)$	$\frac{I_{sc}/L}{(\mu a/\mu)}$
n on p	2A	18.6	.174
n on p	2B	19.0	.159
n on p	3A	17.7	.141
n on p	3B	17.9	.154
p on n	5A	4.3	.176
p on n	5B	4.4	.171
p on n	6A	7.0	.163
p on n	6B	5.2	.170

the electron beam, regardless of whether the minority carriers are holes or electrons. Since the area of each of the cells used was 1 cm², the average value of the ratio gives a calibration for the particular γ-ray source of 0.162 $\mu a/\mu$-cm². This number may be converted to an absorbed dose rate[11] and yields 0.88×10^6 rad/hr. By means of the definition of the rad and the roentgen, and using the mass stopping power of air relative to silicon one can convert this number to correspond to 0.81×10^6 r/hr which is in reasonable agreement with the existing calibration.

VI. RADIATION DAMAGE

The damage produced by energetic radiation causes the minority carrier diffusion length to decrease in a well defined way. Fig. 8 gives the experimental results for the measured diffusion length as a function of 1 Mev electron flux for an n-diffused, p-type 1 Ω-cm base silicon solar cell. The solid curve is a plot of the theoretical formula

$$\frac{1}{L^2} = \frac{1}{L_0^2} + K\Phi \tag{6}$$

in which L is the diffusion length after some bombardment flux Φ, and L_0 the initial diffusion length, with $L_0 = 140\ \mu$ and $K = 1.8 \times 10^{-10}$.

The value of K depends on many parameters such as the energy and type of bombarding particle, the resistivity and conductivity type of the material, the temperature, and the impurity concentration. The variation of K with particle type and energy will be the subject of other papers. Table IV summarizes some of the results for the radiations mentioned in this paper and for silicon solar cells made from pulled crystals and irradiated at room temperature.

From this one can determine, for example, the per cent change in

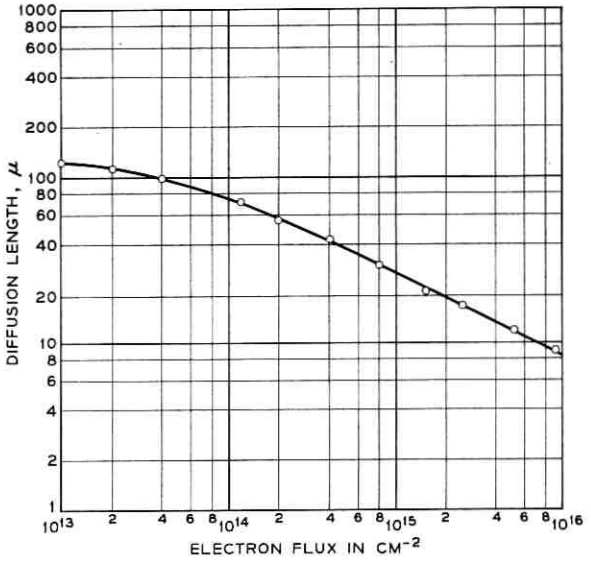


Fig. 8 — Degradation of diffusion length vs electron flux.

diffusion length of an n-type base solar cell (1 Ω cm) during a measurement with 1 Mev electrons for an exposure of 10^{10} cm^{-2} when the initial L value is 100 μ. Differentiation of (6) yields

$$\frac{dL}{L} = -\frac{K}{2} L^2 \, d\Phi \tag{7}$$

which gives, for the above case, a change of less than 0.2 per cent.

VII. CONCLUSION

It has been demonstrated that the method for measuring diffusion length by means of ionizing radiations is particularly suitable for the case of the solar cell and is expected to be useful for devices of similar geometry. Measurements with three types of radiations—electrons, pro-

TABLE IV — RADIATION DAMAGE PARAMETER $K = (d/d\Phi)(1/L^2)$ FOR SOME RADIATIONS AND CONDUCTIVITY TYPES IN SILICON

Base Conductivity Type and Resistivity		1 Mev Electrons	16.8 Mev Protons	130 Mev Protons	Co^{60} γ-rays
p-type	1 Ω cm	1.8×10^{-10}	8.3×10^{-7}	3.3×10^{-7}	$2.6 \times 10^{-3} r^{-1} cm^{-2}$
p-type	10 Ω cm	5.8×10^{-11}			
n-type	1 Ω cm	2.6×10^{-9}	5.1×10^{-6}	2.0×10^{-6}	$4.0 \times 10^{-2} r^{-1} cm^{-2}$.

tons and gamma rays—are shown to give consistent results over a wide range of diffusion lengths and for both conductivity types in silicon. The electron beam technique is shown to lend itself most readily to an absolute calibration and to routine measurements of diffusion length. By this technique the average specific ionization of electrons as a function of penetration depth is measured. For 0.98 Mev electrons in silicon this ionization reaches a maximum value of 225 pairs/μ $\pm$ 5 per cent at a depth of 0.096 gm/cm^2.

The consistency found among the various types of radiation allows one to reverse the argument and suggest that calibrated solar cells be used as solid-state ionization chambers for measuring radiation intensity. A range of 10^3 rad/hr to 10^9 rad/hr should be achievable.

VIII. ACKNOWLEDGMENT

The author wishes to express his appreciation for many helpful discussions with W. L. Brown, F. M. Smits, and W. M. Augustyniak, and for the careful assistance given by J. A. O'Sullivan in carrying out many of the measurements. Thanks are also due to P. J. Kamps for the mechanical design of the double-aperture Faraday cup.

APPENDIX

For a solar cell of the type shown in Fig. 1, for which the junction depth is much smaller than the diffusion length of the minority carriers, one needs to solve the diffusion equation only in the semi-infinite region beyond the junction for the case of penetrating radiation. If the junction is assumed to be located at $x = 0$ and the carrier generation rate as a function of distance from the junction to be given by $g(x)$, then the steady-state excess carrier concentration will be given by the solution of the equation:

$$D\frac{d^2n}{dx^2} - \frac{n}{\tau} = -g(x) \tag{8}$$

where

D = minority carrier diffusion coefficient,
τ = minority carrier lifetime,
n = excess minority carrier concentration.

Under short circuit condition, i.e., zero bias on the junction, the equation is solved with boundary conditions $n = 0$, $x = 0$, and $dn/dx = 0$, $x = \infty$. If the carrier generation rate is reasonably uniform within a

diffusion length, $L = \sqrt{D\tau}$, of the junction, then $g(x)$ may be represented by the first two terms of its Taylor series expansion as follows:

$$g(x) = g_0 + g_1 x. \tag{9}$$

The solution to (8) with $g(x)$ given by (9) under the above boundary conditions is given by:

$$n(x) = g_0\tau(1 - e^{-x/L}) + g_1 x\tau \tag{10}$$

which yields for the magnitude of the short circuit current density:

$$J_{sc} = eg_0L\left(1 + \frac{g_1}{g_0}L\right). \tag{11}$$

From (11) it may be seen that under strictly uniform carrier generation, i.e., $g_1 \equiv 0$, the short circuit current density is given by $J_{sc} = eg_0L$. This magnitude is increased or diminished in the nonuniform case by an amount eg_1L depending on whether the generation rate increases or diminishes with depth in the material, i.e., depending on whether g_1 is positive or negative.

REFERENCES

1. Bemski, Proc. I.R.E., **46,** June, 1958, p. 990.
2. Gremmelmaier, Proc. I.R.E., **46,** June, 1958, p. 1045.
3. Chynoweth, A. G., private communication.
4. Spencer, L. V. S., Phys. Rev., **98,** June 15, 1955, p. 1597.
5. Katz, L., and Penfold, A. S., Rev. Mod. Phys., **24,** Jan., 1952, p. 28.
6. Wright, K. A., and Trump, J. G., J. Appl. Phys., **33,** Feb., 1962, p. 687.
7. Sternheimer, R. M., Phys. Rev., **115,** July 1, 1959, p. 137.
8. Bethe, H. A., and Ashkin, J., *Experimental Nuclear Physics,* I, New York, John Wiley & Sons, 1953.
9. Allison, S. K., and Warshaw, S. D., Rev. Mod. Phys., **25,** Oct., 1953, p. 779.
10. Grodstein, G. W., National Bureau of Standards Circular 583 (1957), and Supplement (1959).
11. Rosenzweig, W., Rev. Sci. Inst., **33,** March, 1962, p. 379.

Imperfections in Lined Waveguide

By H. L. KREIPE and H. G. UNGER*

(Manuscript received May 11, 1962)

Thickness variations of the lining and deformations of the cross-section couple circular electric waves to unwanted modes and degrade transmission characteristics. Generalized telegraphist's equations for lined waveguide with these imperfections are found. The most critical interaction is caused by lining variations of circular symmetry between TE_{01} and higher circular electric waves. Because of such interaction the average TE_{01} loss is increased and signal transmission is distorted. More serious than signal distortion is the increase in average loss. In 2″ I.D. copper waveguide with a 0.01″ lining, the rms of a typical thickness variation should stay below 0.002″ for the TE_{01} loss at 55.5 kmc not to be raised more than 10 per cent. Cross-sectional deformations in lined waveguide cause nearly the same increase in loss and signal distortion as in plain waveguide. Tolerances for such deformations should therefore be the same as in plain waveguide.

I. INTRODUCTION

Lined waveguide shows promise as a communication medium.[1] Circular electric wave loss in bends is reduced by a low-loss lining. A lossy lining reduces the degrading effects of mode conversion and reconversion.

In straight circular waveguide with a perfectly uniform lining, circular electric waves will only suffer a slight increase in attenuation due to the loss factor of the lining and the slightly increased wall currents. Also, the phase constant will be only slightly shifted. Otherwise the transmission characteristics will remain smooth and undistorted.

A perfect waveguide with a perfectly uniform lining, however, cannot be realized in practice. The waveguide will be slightly deformed, and any lining as it is applied by spraying, dipping, or other methods, will show variations in thickness along the circumference as well as longitudinally. These variations will, in general, be distributed randomly.

* Technische Hochschule Braunschweig; work done under letter contract.

Deformations of the cross section or an imperfect lining couple the circular electric wave to other unwanted modes and by mode conversion and reconversion further increase the loss and otherwise degrade the transmission characteristics. To keep these degrading effects small, the waveguide has to be made sufficiently round and straight and the lining uniform. In order to specify tolerances, a theory is needed of circular electric wave propagation in a deformed waveguide with varying thickness of the lining.

Attempts at such a theory have been made elsewhere, but they were limited to a first-order approximation[2,3,4]. These approximations are by far not adequate to describe the cases of practical interest.

Imperfect lining in a perfect waveguide will be analyzed first. Subsequently, the effects of cross-sectional deformations will be taken into account.

II. GENERALIZED TELEGRAPHIST'S EQUATIONS FOR A WAVEGUIDE WITH IMPERFECT LINING

The lined waveguide, Fig. 1, will be considered in cylindrical coordinates (r, φ, z). The thickness t of the lining will for the moment be assumed a function of φ only. Any such function, being periodic in φ, may be expanded into a Fourier series

$$t = t_1(1 + \sum_q \nu_q \cos q\varphi). \tag{1}$$

The sine terms have been omitted in (1); they would only add terms of orthogonal polarization. t_1 is the nominal thickness of the lining.

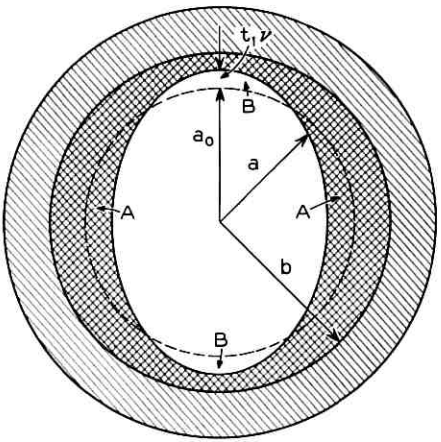


Fig. 1 — Round waveguide with lining of varying thickness.

The relative permittivity will be written

$$\epsilon_1 + \epsilon_2$$

where ϵ_1 is the relative permittivity of the waveguide with a perfectly uniform lining. ϵ_2 takes into account the actual permittivity distribution in a waveguide with varying thickness of the lining.

More specifically:

$$\begin{aligned} \epsilon_1 &= 1 \qquad & 0 < r < a_0 \\ \epsilon_1 &= \epsilon_r \qquad & a_0 < r < b \end{aligned} \tag{2}$$

where ϵ_r is the relative permittivity of the lining. Moreover,

$$\begin{aligned} \epsilon_1 + \epsilon_2 &= \epsilon_r\,; \qquad \epsilon_2 = \epsilon_r - 1 \quad \text{in areas A (Fig. 1)} \\ \epsilon_1 + \epsilon_2 &= 1\,; \qquad \epsilon_2 = 1 - \epsilon_r \quad \text{in areas B.} \end{aligned} \tag{3}$$

In other cross-sectional areas $\epsilon_2 = 0$. Only the thickness of the lining varies; the permittivity is assumed to be constant.

The electromagnetic field in the waveguide is described in terms of normal modes of the round waveguide with a perfectly uniform lining. These modes are derived from two sets of scalar functions T_n and T_n' given by (6.1) and (6.2).*

The transverse field components are written in terms of the following wave functions:

$$\begin{aligned} E_r &= \sum_n V_n \left[\frac{\partial T_n}{\partial r} + d_n \frac{\partial T_n'}{r\partial\varphi} \right] \\ E_\varphi &= \sum_n V_n \left[\frac{\partial T_n}{r\partial\varphi} - d_n \frac{\partial T_n'}{\partial r} \right] \\ H_r &= -\sum_n I_n \left[\frac{\partial T_n}{r\partial\varphi} - d_n \frac{h_n^2}{\epsilon_1 k^2} \frac{\partial T_n'}{\partial r} \right] \epsilon_1 \\ H_\varphi &= \sum_n I_n \left[\frac{\partial T_n}{\partial r} + d_n \frac{h_n^2}{\epsilon_1 k^2} \frac{\partial T_n'}{r\partial\varphi} \right] \epsilon_1 . \end{aligned} \tag{4}$$

The individual terms in (4) represent normal modes of the lined waveguide when d_n is chosen according to (6.26). $k = \omega\sqrt{\mu_0\epsilon_0}$ is the free-space wave number; h_n is the axial propagation constant; and

$$k_n \equiv \chi_n a_0 = a_0(k^2 - h_n^2)^{\frac{1}{2}} \tag{5}$$

is the radial propagation constant of the mode n.

* These equations are listed in Ref. 6; the terminology (6.1) refers to Ref. 6 equation (1), (6.2) to Ref, 6 equation (2), etc.

In Maxwell's equations the distribution of permittivity $(\epsilon_1 + \epsilon_2)$ must be taken into account:

$$\frac{1}{r}\left[\frac{\partial}{\partial\varphi}(E_z) - \frac{\partial}{\partial z}(rE_\varphi)\right] = -j\omega\mu_0 H_r \tag{6}$$

$$\left[\frac{\partial}{\partial z}(E_r) - \frac{\partial}{\partial r}(E_z)\right] = -j\omega\mu_0 H_\varphi \tag{7}$$

$$\frac{1}{r}\left[\frac{\partial}{\partial r}(rE_\varphi) - \frac{\partial}{\partial\varphi}(E_r)\right] = -j\omega\mu_0 H_z \tag{8}$$

$$\frac{1}{r}\left[\frac{\partial}{\partial\varphi}(H_z) - \frac{\partial}{\partial z}(rH_\varphi)\right] = j\omega(\epsilon_1 + \epsilon_2)\epsilon_0 E_r \tag{9}$$

$$\left[\frac{\partial}{\partial z}(H_r) - \frac{\partial}{\partial r}(H_z)\right] = j\omega(\epsilon_1 + \epsilon_2)\epsilon_0 E_\varphi \tag{10}$$

$$\frac{1}{r}\left[\frac{\partial}{\partial r}(rH_\varphi) - \frac{\partial}{\partial\varphi}(H_r)\right] = j\omega(\epsilon_1 + \epsilon_2)\epsilon_0 E_z\,. \tag{11}$$

Substituting for the transverse field components from (4) into (8) and (11) and taking advantage of (6.3), the longitudinal field components are obtained:

$$H_z = j\omega\epsilon_0 \sum_n V_n d_n \frac{\chi_n^2}{k^2} T_n' \tag{12}$$

$$H_z = j\omega\mu_0 \sum_n I_n \frac{\epsilon_1}{\epsilon_1 + \epsilon_2} \frac{\chi_n^2}{k^2} T_n\,. \tag{13}$$

To find relations for the current and voltage coefficients, the series representations (4) and (12) and (13) for the field components are substituted into Maxwell's equations. Then by multiplying with orthogonal field functions, combining some of these equations, and integrating over the cross section, generalized telegraphist's equations are obtained.

For example,

$$-\left(\frac{\partial T_m}{r\partial\varphi} - d_m \frac{h_m^2}{\epsilon_1 k^2} \frac{\partial T_m'}{\partial r}\right)\epsilon_1 \qquad \text{times (6)}$$

is added to

$$\left(\frac{\partial T_m}{\partial r} + d_m \frac{h_m^2}{\epsilon_1 k^2} \frac{\partial T_m'}{r\partial\varphi}\right)\epsilon_1 \qquad \text{times (7)}$$

and the result is integrated over the total cross section. Using the orthonormality condition (6.29)[5] and the wave equation (6.3), one obtains:

$$\frac{dV_m}{dz} + j\frac{h_m^2}{\omega\epsilon_0} I_m = -j\omega\mu_0 \sum_n I_n \int_S \frac{\epsilon_1\epsilon_2}{\epsilon_1 + \epsilon_2} \frac{\chi_n^2 \chi_m^2}{k^2} T_n T_m \, dS. \quad (14)$$

Similarly,

$$-\left(\frac{\partial T_m}{\partial r} + d_m \frac{\partial T_m'}{r\partial\varphi}\right) \qquad \text{times (9)}$$

is added to

$$-\left(\frac{\partial T_m}{r\partial\varphi} - d_m \frac{\partial T_m'}{\partial r}\right) \qquad \text{times (10)}$$

and the result is integrated over the cross section

$$\frac{dI_m}{dz} + j\omega\epsilon_0 V_m$$

$$= -j\omega\epsilon_0 \sum_n V_n \int_S \epsilon_2 \left\{\left[\frac{\partial T_n}{\partial r} + d_n \frac{\partial T_n'}{r\partial\varphi}\right]\left[\frac{\partial T_m}{\partial r} + d_m \frac{\partial T_m'}{r\partial\varphi}\right]\right. \quad (15)$$

$$\left. + \left[\frac{\partial T_n}{r\partial\varphi} - d_n \frac{\partial T_n'}{r\partial r}\right]\left[\frac{\partial T_m}{r\partial\varphi} - d_m \frac{\partial T_m'}{\partial r}\right]\right\} dS.$$

Equations (14) and (15) are generalized telegraphist's equations for round waveguide with imperfect lining.

Introducing traveling waves,

$$\begin{aligned} V_m &= \sqrt{K_m}(a_m + b_m) \\ I_m &= \frac{1}{\sqrt{K_m}}(a_m - b_m) \end{aligned} \quad (16)$$

with the characteristic impedance

$$K_m = \frac{h_m}{\omega\epsilon_0}, \quad (17)$$

the more convenient form of generalized telegraphist's equations in terms of amplitudes of forward (a_m) and backward (b_m) traveling waves is obtained:

$$\begin{aligned} \frac{da_m}{dz} + jh_m a_m &= j \sum_n [c_{mn}^{+} a_n + c_{mn}^{-} b_n] \\ \frac{db_m}{dz} - jh_m b_m &= -j \sum_n [c_{mn}^{+} b_n + c_{mn}^{-} a_n]. \end{aligned} \quad (18)$$

The coupling coefficients in (18) are

$$c_{mn}^{\pm} = \mp \frac{k^2}{2\sqrt{h_m h_n}} \int_S \frac{\epsilon_1 \epsilon_2}{\epsilon_1 + \epsilon_2} \frac{\chi_n^2 \chi_m^2}{k^2} T_n T_m \, dS$$
$$- \frac{\sqrt{h_m h_n}}{2} \int_S \epsilon_2 \left\{ \left[\frac{\partial T_n}{\partial r} + d_n \frac{\partial T_n'}{r \partial \varphi} \right] \left[\frac{\partial T_m}{\partial r} + d_m \frac{\partial T_m'}{r \partial \varphi} \right] \right. \tag{19}$$
$$\left. + \left[\frac{\partial T_n}{r \partial \varphi} - d_n \frac{\partial T_n'}{\partial r} \right] \left[\frac{\partial T_m}{r \partial \varphi} - d_m \frac{\partial T_m'}{\partial r} \right] \right\} dS.$$

To analyze circular electric wave propagation, it is sufficient to consider only coupling between circular electric and other waves. Let m denote the TE_{0m} wave, then

$$T_m = 0$$

and

$$E_{z_m} = E_{r_m} = H_{\varphi_m} = 0.$$

In this case the coupling coefficients reduce to

$$c_{mn}^{\pm} = \frac{d_m \sqrt{h_m h_n}}{2} \int_S \epsilon_2 \left[\frac{\partial T_n}{r \partial \varphi} - d_n \frac{\partial T_n'}{\partial r} \right] \frac{\partial T_m'}{\partial r} \, dS. \tag{20}$$

Also the generalized telegraphist's equations may now be written shorter:

$$\begin{aligned} \frac{da_m}{dz} + j h_m a_m &= j \sum_n c_{mn}(a_n + b_n) \\ \frac{db_m}{dz} - j h_m b_m &= j \sum_n c_{mn}(a_n + b_n). \end{aligned} \tag{21}$$

To find the z-dependence of the wave amplitudes a_m and b_m for certain initial conditions requires the solution of the generalized telegraphist's equations (18) or (21). They are a system of simultaneous first-order and linear differential equations and can be solved by standard methods.

III. COUPLING COEFFICIENTS FOR LINING IMPERFECTIONS

First of all, the coupling coefficients have to be evaluated. Under practical conditions certain approximations may be made. The maximum deviation of the lining from its nominal thickness or also any of its Fourier components are assumed to be small compared to the nominal thickness:

$$\nu_p \ll 1. \tag{22}$$

A first approximation is obtained by substituting for the wave functions T_n and T_n' in the range of thickness deviations their values at the nominal boundary $r = a_0$. Thus the integration in (20) is facilitated. ϵ_2 is different from zero only in regions A and B of Fig. 1, and in these regions the integrand is assumed independent of r.

According to the boundary condition at $r = a_0$ for the internal and external field components

$$E_{\varphi_n}{}^i E_{\varphi_m}{}^i = E_{\varphi_n}{}^e E_{\varphi_m}{}^e$$

and consequently, because of (4):

$$\left[\frac{\partial T_n{}^i}{r\partial\varphi} - d_n \frac{\partial T_n{}^{i'}}{\partial r}\right] \frac{\partial T_m{}^{i'}}{\partial r} = \left[\frac{\partial T_m{}^e}{r\partial\varphi} - d_n \frac{\partial T_n{}^{e'}}{\partial r}\right] \frac{\partial T_m{}^{e'}}{\partial r}. \tag{23}$$

The various terms of the integrand in (20) are:

$$\begin{aligned} \frac{\partial T_n{}^i}{r\partial\varphi} &= N_n J_p(k_n) \frac{p}{a_0} \cos p\varphi \\ \frac{\partial T_n{}^{i'}}{\partial r} &= N_n \chi_n J_p'(k_n) \cos p\varphi \\ \frac{\partial T_m{}^{i'}}{\partial r} &= N_m \chi_m J_0'(k_m). \end{aligned} \tag{24}$$

Introducing these terms into (20) and using (23), the coupling coefficients are:

$$\begin{aligned} c_{mn}{}^{\pm} = \tfrac{1}{2} \sqrt{h_m h_n}\, d_m N_m N_n k_m J_1(k_m) J_p(k_n)(p - d_n k_n{}^2 y_n) \\ \cdot(\epsilon_r - 1) \frac{1}{a_0} \int_0^{2\pi} \int_{a_0}^{b-t} \cos p\varphi \, dr d\varphi \end{aligned} \tag{25}$$

$$\text{where } y_n = \frac{J_p'(k_n)}{k_n J_p(k_n)}. \tag{26}$$

Integrating over the radial coordinate r, the upper limit $b - t$ is a function of φ. According to (1)

$$b - t = a_0 - t_1 \sum_q \nu_q \cos q\varphi. \tag{27}$$

Integrating in (25) over r and φ the contributions from all Fourier components of (27) disappear except those with $q = p$. One obtains for the coupling coefficients:

$$\begin{aligned} c_{mn} = -\tfrac{\pi}{2} \sqrt{h_m h_n}\, d_m N_m N_n k_m J_1(k_m) J_p(k_n) \ (p - d_n k_n{}^2 y_n) \\ \cdot(\epsilon_r - 1)\delta_1 \nu_p \end{aligned} \tag{28}$$

where for the nominal thickness the relative measure

$$\delta_1 = \frac{t_1}{a_0} \tag{29}$$

has been introduced.

For a thickness deviation described by a single coefficient ν_p, there is only coupling between circular electric waves and waves of circumferential order $p = q$.

For the term corresponding to $q = 0$ the integration results in

$$c_{mn} = -\pi\sqrt{h_m h_n}\, d_m N_m d_n N_n k_m k_n J_1(k_m) J_1(k_n)(\epsilon_r - 1)\delta_1 \nu_0. \tag{30}$$

This uniform thickness deviation only causes mutual coupling between circular electric waves.

For a very thin lining, the modes of lined waveguide may be considered perturbed modes of plain waveguide. It has been found elsewhere[6] that a first-order perturbation of this kind is a good approximation only in a very limited range. Nevertheless it will be quite informative to study the approximations for the various expressions when this first-order perturbation is introduced.

The asymmetric modes are either perturbed TE or perturbed TM modes. The coupling coefficient between circular electric and perturbed TE modes reduces to

$$c_{mn} = -\frac{k^2 k_{m0} k_{n0}\left[1 - \dfrac{p^2}{\epsilon_r k_{n0}^2 \dfrac{k^2}{h_{n0}^2}\left(1 + \dfrac{k_{n0}^2}{k^2 b^2(\epsilon_r - 1)}\right)}\right]}{\sqrt{2h_{m0}h_{n0}}\sqrt{1 - \dfrac{p^2}{k_{n0}^2}}}(\epsilon_r - 1)\delta_1^3 \nu_p \tag{31}$$

where the second subscript 0 indicates the value of the corresponding quantity in plain metallic waveguide.

The coupling coefficient between circular electric and perturbed TM modes reduces to

$$c_{mn} = \sqrt{\frac{h_{n0}}{2h_{m0}}}\, k k_{m0} p \frac{\epsilon_r - 1}{\epsilon_r}\delta_1^3 \nu_p . \tag{32}$$

The uniform thickness deviation causes coupling only between circular electric waves. For a very thin lining the corresponding coupling coefficients (30) reduce to

$$c_{mn} = -\frac{k^2 k_{m0} k_{n0}}{\sqrt{h_{m0}h_{n0}}}(\epsilon_r - 1)\delta_1^3 \nu_0 . \tag{33}$$

In all approximate expressions (31) to (33) the coupling coefficients vary with the third power of the relative thickness δ_1 of the lining and are proportional to the relative deviation ν_p of the thickness.

To evaluate the exact expressions (28) and (30) for the coupling coefficients as well as the approximate expressions (31) to (33), the propagation characteristics of the coupled modes must be known first. The characteristic equation for the normal modes in lined waveguide of perfect geometry has been solved numerically by an iterative procedure. The evaluation was programmed for automatic execution on a digital computer.

Also included in this program was an evaluation of the exact formula (28) for coupling between TE_{01} and any asymmetric mode.

For the typical values $b/\lambda = 4.70$ and 7.62 corresponding to 2″ I.D. waveguide at 55.5 kmc and 90 kmc, and for a permittivity $\epsilon_r = 2.5$, the phase constants of a number of modes are plotted versus the relative thickness of the lining in Figs. 2 to 5.

The lining changes phase constants most effectively for TE_{p1} modes and all TM_{pn} modes. All TE_{pn} modes with $n > 1$ show very little change for a thin lining. First-order approximations for the phase constants in case of very thin lining would be represented in these plots by straight lines tangent to the curves at $\delta_1 = 0$. Note that these first-order approximations are gravely in error for most modes and any substantial thickness of the lining.

Some of the coupling coefficients for thickness deviations have been plotted in Figs. 6 to 9 as a function of the thickness of the lining. In the log-log plot of these figures the first-order approximations for thin lining appear as straight lines with slope 3. Note that these straight lines are fairly good approximations for coupling between TE_{01} and TE_{pn} modes with $n > 1$ and a relative thickness of the lining smaller than 1 per cent. For coupling between TE_{01} and all other modes, however, these approximations are again seriously in error.

Note also that coupling between TE_{01} and TE_{pn} modes is always larger than coupling between TE_{01} and TM_{pm} modes. Furthermore, coupling between TE_{01} and TE_{pn} increases with increasing order n of the radial dependence of coupled modes, while it decreases in case of TM_{pn} modes.

In Fig. 10 the coefficients of coupling between TE_{01} and higher circular electric waves for circular symmetric components of lining imperfection corresponding to $p = 0$ have been plotted. First-order approximations for thin linings are shown in this diagram only. They are adequate for the entire thickness range to be considered. As in the general case of

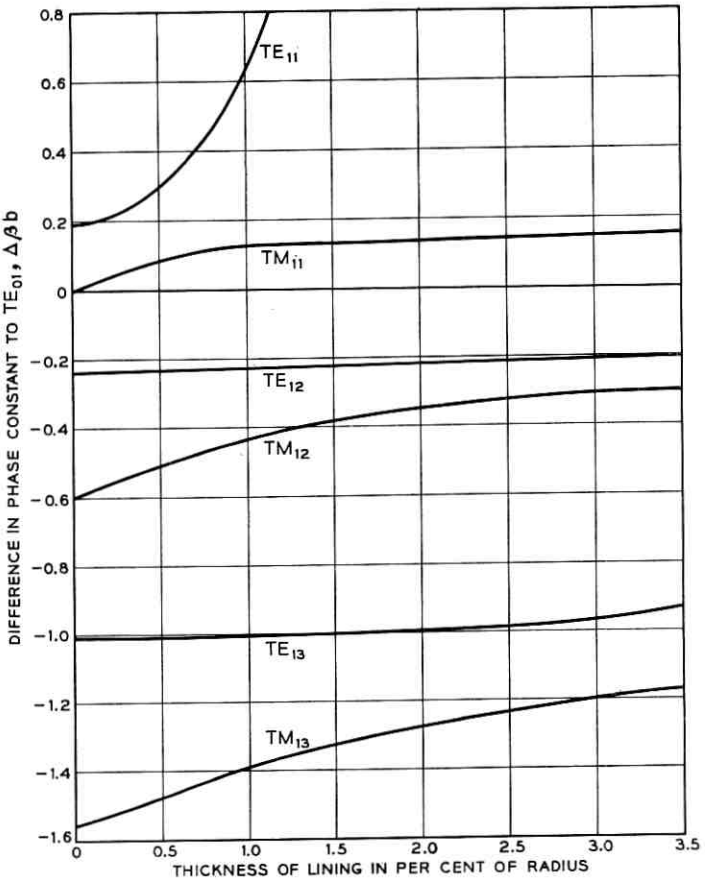


Fig. 2 — Phase constants of TE_{1n} and TM_{1n} modes in lined waveguide. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

coupling to TE_{pn} the coupling coefficient increases with the radial order n of the higher circular electric modes.

IV. MODE CONVERSION AT IMPERFECT LININGS OF VARIOUS PERMITTIVITIES AND AT VARIOUS FREQUENCIES

Lined waveguide will be used for circular electric wave transmission over wide frequency bands extending up to 100 kmc and more. In addition, for technological reasons the lining might be made of materials of various permittivities.

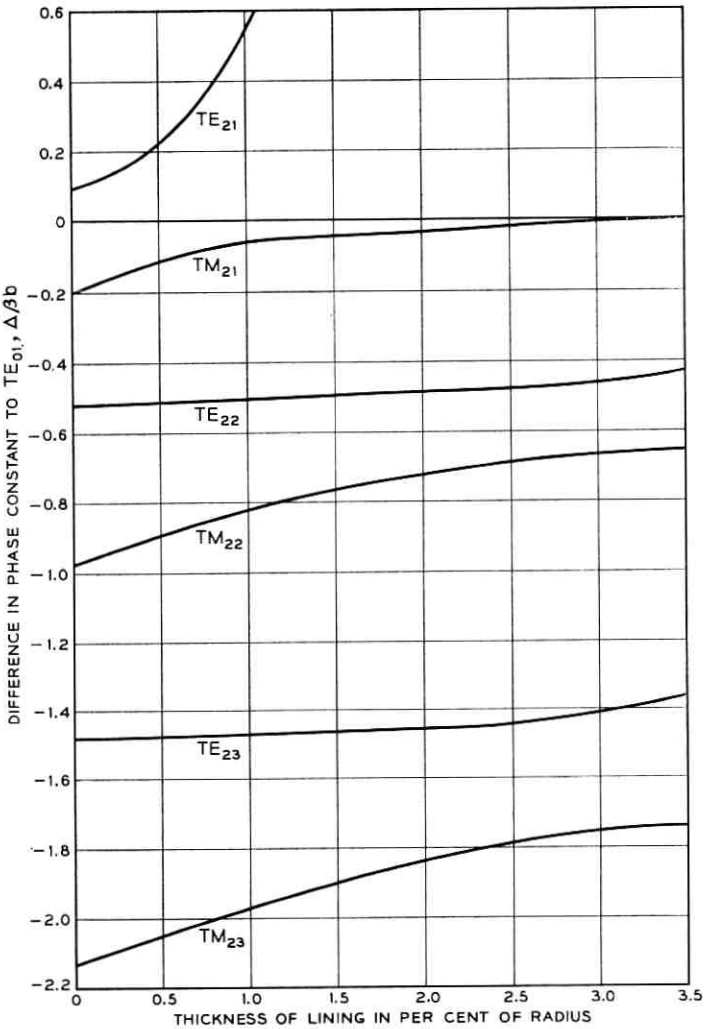


Fig. 3 — Phase constants of TE_{2n} and TM_{2n} modes in lined waveguide. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

A first indication as to how the coupling coefficients depend on frequency and permittivity of the lining is obtained from the approximate formulae (31), (32) and (33) for coefficients of coupling at lining imperfections.

For interaction between waves which are sufficiently far from cutoff we have

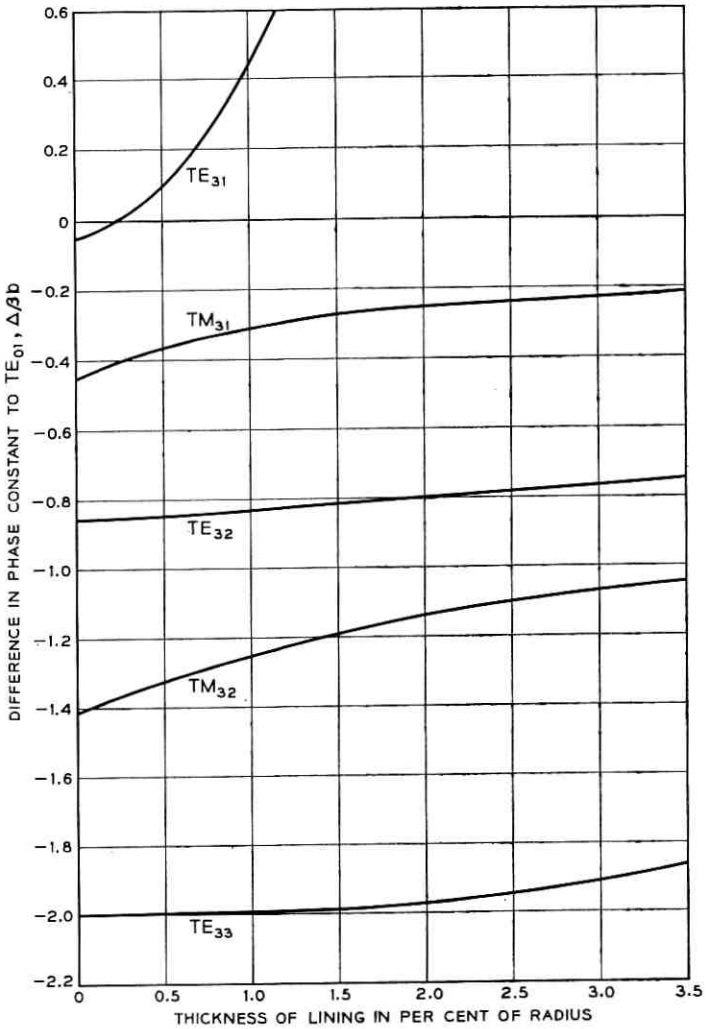


Fig. 4 — Phase constants of TE_{3n} and TM_{3n} modes in lined waveguide. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

$$h_{m0} \cong h_{n0} \cong k$$

and

$$(\epsilon_r - 1)k^2a^2 \gg k_{n0}^2$$

so that for interaction between TE_{0m} and TE_{pn} (31) may be replaced by

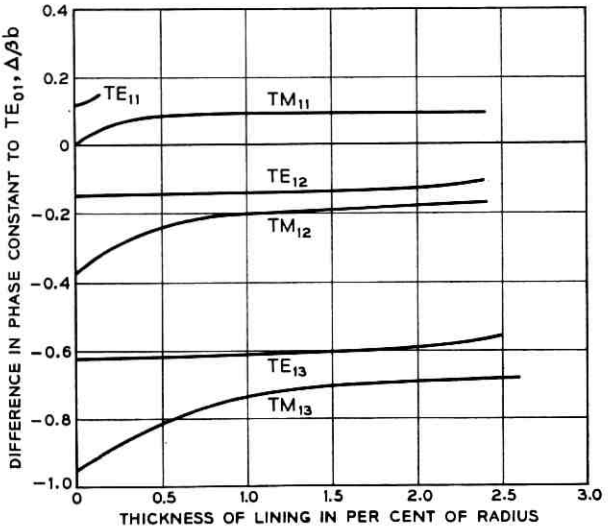


Fig. 5 — Phase constants of TE_{1n} and TM_{1n} modes in lined waveguide. $b/\lambda = 7.62$, $\epsilon_r = 2.5$.

$$c_{mn} = -\frac{k_{n0}k_{m0}}{\sqrt{2}} \frac{1 - \dfrac{p^2}{\epsilon_r k_{n0}^2}}{\left(1 - \dfrac{p^2}{k_{n0}^2}\right)^{\frac{1}{2}}} k(\epsilon_r - 1)\delta_1^3 \nu_p. \tag{34}$$

Equation (32) for interaction between TE_{0m} and TM_{pn} may be replaced by

$$c_{mn} = \frac{p}{\sqrt{2}} k_{m0} k \frac{\epsilon_r - 1}{\epsilon_r} \delta_1^3 \nu_p. \tag{35}$$

Equation (33) for interaction between circular electric waves may be replaced by

$$c_{mn} = -k_{m0}k_{n0}k(\epsilon_r - 1)\delta_1^3 \nu_0. \tag{36}$$

These approximate expressions indicate a linear dependence on frequency of all coupling coefficients. As functions of the permittivity the coupling coefficients are proportional to $(\epsilon_r - 1)/\epsilon_r$ for interaction between TE_{0m} and TM_{pn} waves, but nearly proportional to $(\epsilon_r - 1)$ for interaction between TE_{0m} and all TE waves. The term $p^2/\epsilon_r k_{n0}^2$ may for most TE modes be neglected with respect to unity.

For substantial linings the normal modes cannot be regarded as per-

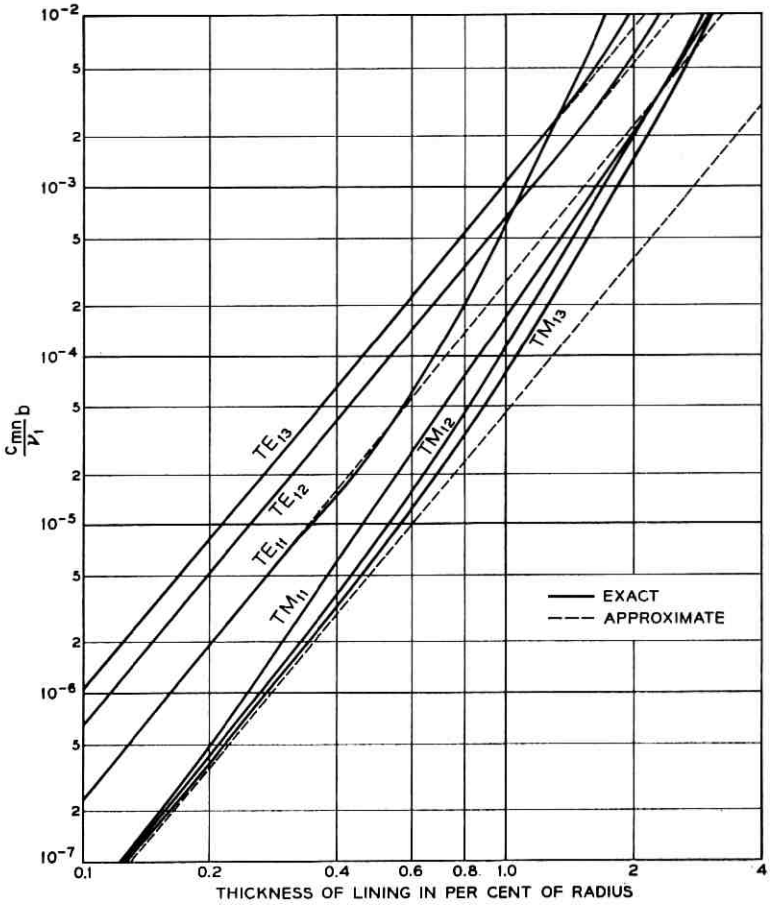


Fig. 6 — Coefficients c_{mn} of coupling between TE_{01} and TE_{1n} and TM_{1n} waves at lining imperfections of order $p = 1$. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

turbed modes of plain waveguide. In this case an indication of the dependence of mode interaction on permittivity and frequency can be obtained only by evaluating the exact expressions.

Comparing the curves of Figs. 6 and 9, it is found that the exact values for coupling coefficients show nearly the same dependence on frequency and permittivity as indicated by the approximations. Only for interaction between TE_{01} and lower-order TM modes are the coupling coefficients appreciably larger at higher frequencies even for thin linings.

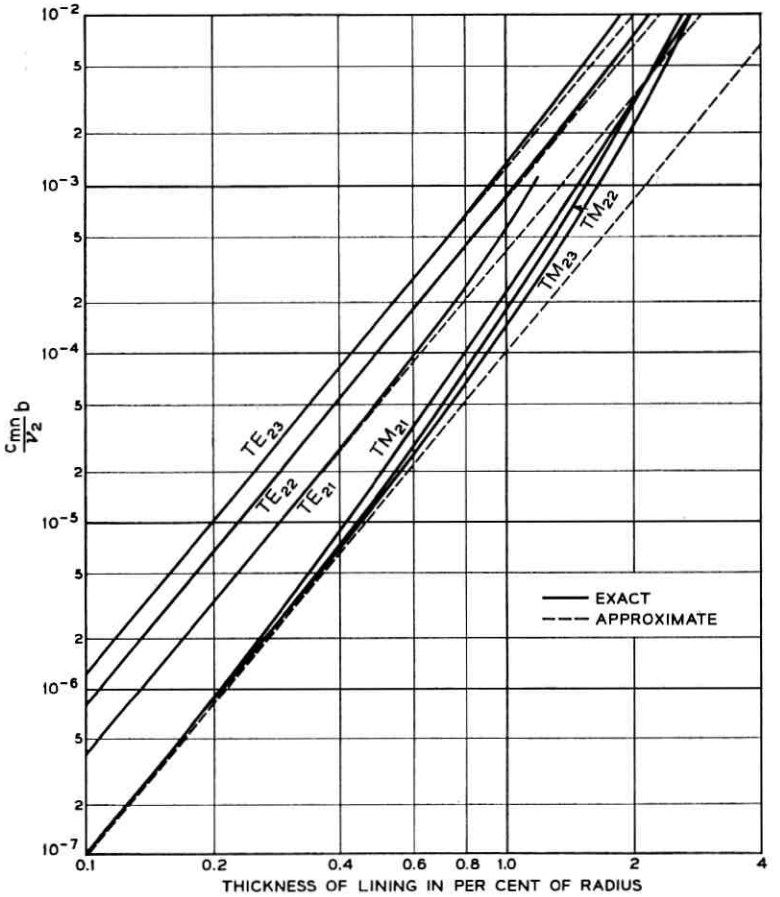


Fig. 7 — Coefficients c_{mn} of coupling between TE_{01} and TE_{2n} and TM_{2n} waves at lining imperfections of order $p = 2$. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

All these characteristics of coupling due to imperfect linings will have to be considered when the coupled line equations are solved for TE_{01} propagation.

V. LINING TOLERANCES

In manufacturing lined waveguide, tolerances for irregularities must be specified. These irregularities are randomly distributed, and at best some of their statistical properties are known. By solving the generalized

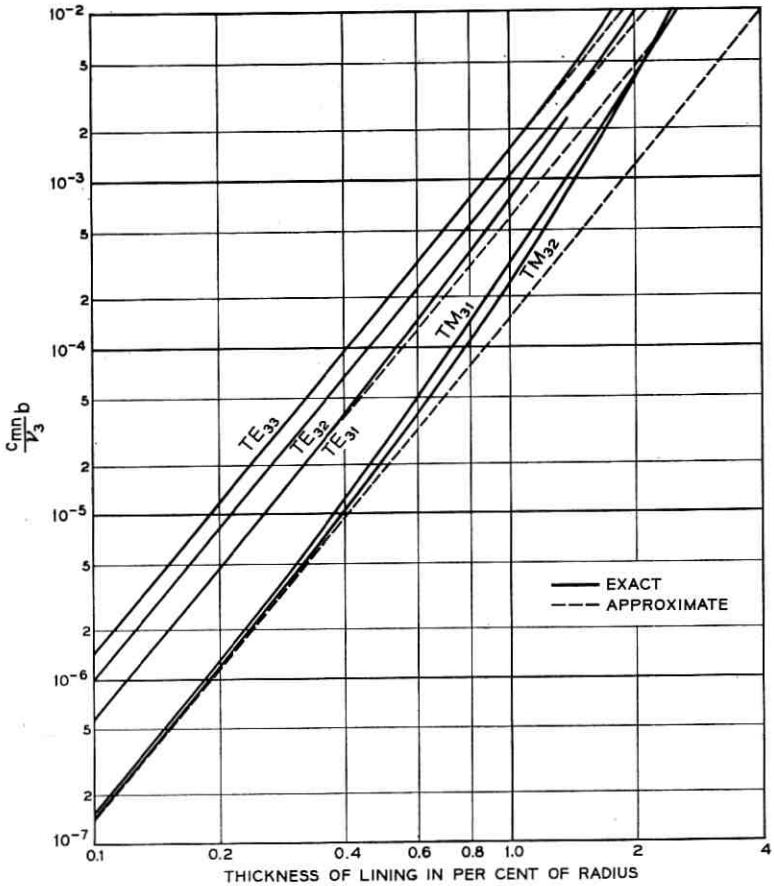


Fig. 8 — Coefficients c_{mn} of coupling between TE_{01} and TE_{3n} and TM_{3n} waves at lining imperfections of order $p = 3$. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

telegraphist's equations, statistics of the transmission characteristics may be expressed in terms of the statistics of lining imperfections.

In particular, the average added loss for circular electric waves due to mode conversion at lining imperfections is[7]

$$\langle \Delta\alpha_m \rangle = \frac{1}{L} \sum_n \int_0^L R(z)(L - z)\, C_{nm}{}^2 \cos \Delta\beta_{nm} z \, dz \qquad (37)$$

where

$$R(z) = \langle \nu_p(z_1) \cdot \nu_p(z_1 + z) \rangle$$

is the covariance of a component $\nu_p(z)$ of thickness deviation. $\nu_p(z)$ is

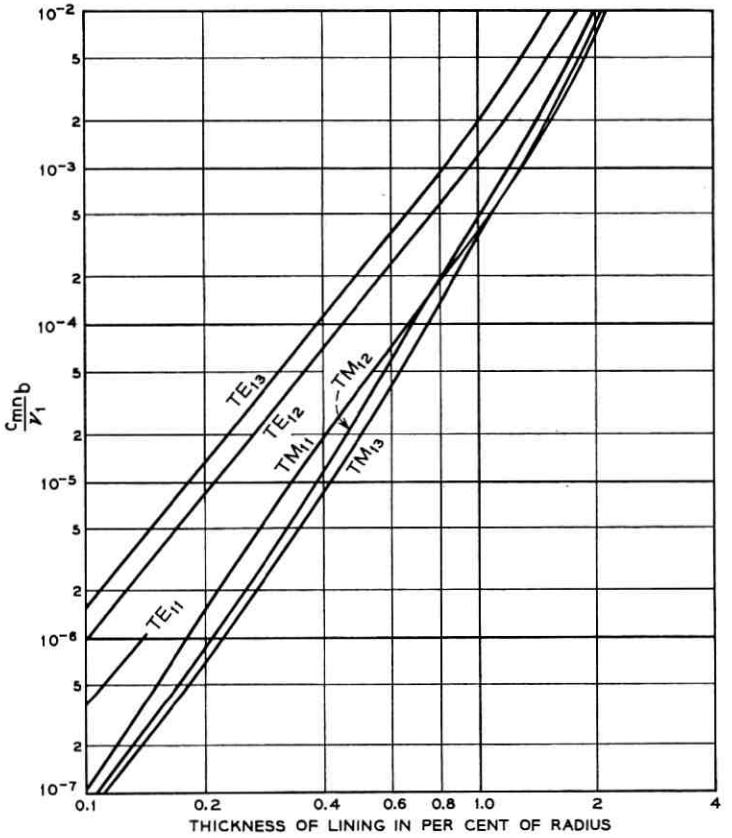


Fig. 9 — Coefficients c_{mn} of coupling between TE_{01} and TE_{1n} and TM_{1n} waves at lining imperfections ($p = 1$). $b/\lambda = 7.62$, $\epsilon_r = 2.5$.

assumed to be a stationary random process in z. C_{nm} is a coupling factor defined according to

$$c_{nm} = C_{nm}\nu_p(z).$$

L is the total length of the waveguide. $\Delta\beta_{nm}$ is the difference in phase constant between the circular electric wave m and the mode n coupled by c_{nm} to m. The difference in attenuation $\Delta\alpha_{nm}$ between mode m and n is assumed small enough so that

$$e^{-\Delta\alpha_{nm}z} \cong 1$$

for all z for which $R(z)$ has any substantial values.

The covariance will be assumed to drop off exponentially

$$R(z) = \langle \nu_p^2 \rangle \, e^{-2\pi \frac{|z|}{L_0}} \tag{38}$$

where L_0 is a correlation distance. Substituting (38) for $R(z)$ into (37) and performing the integration for $L_0 \ll L$ the average added loss is:

$$\langle \Delta\alpha_m \rangle = \langle \nu_p^2 \rangle \sum_n \frac{2\pi C_{nm}^2 L_0}{4\pi^2 + \Delta\beta_{nm}^2 L_0^2}. \tag{39}$$

To evaluate (39) for the average added loss in the range of relative thickness δ to be considered here, coupling to all TE_{pn} which are propa-

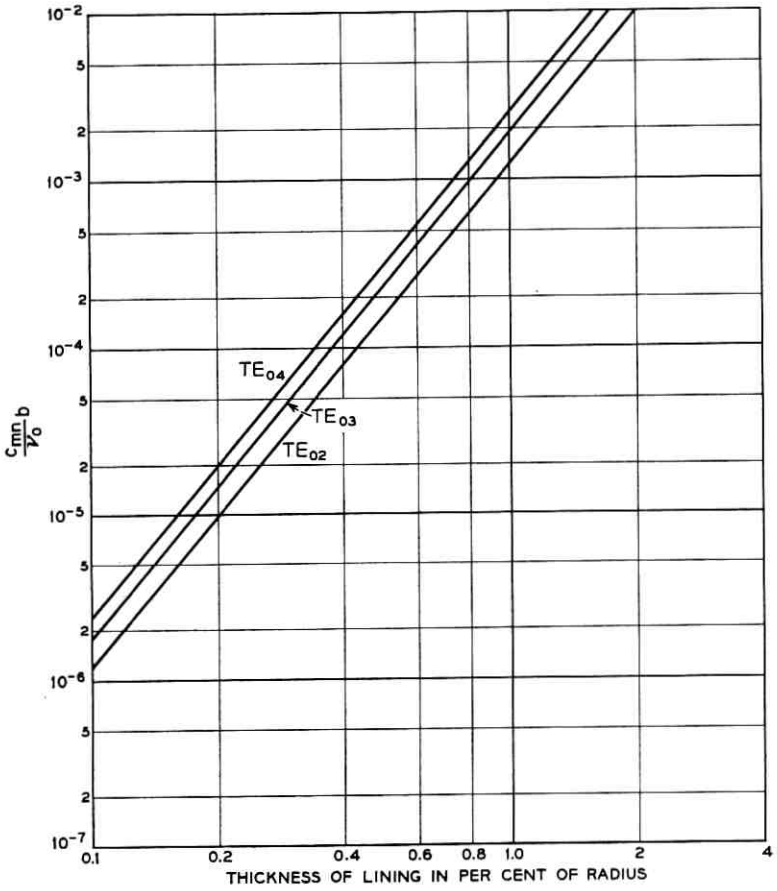


Fig. 10 — Coefficients c_{mn} of coupling between TE_{01} and TE_{0n} waves at lining imperfections of circular symmetry. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

gating must be taken into account. Of the coupling to TM_{pn} modes only lower-order modes with $n < 4$ need be taken into account. The contribution of coupling to TM_{pn} modes with $n \geqq 4$ to the average added TE_{01} loss is small enough to be neglected.

For the coupling to TE_{pn} modes, first-order approximations for phase constants and coupling coefficients may be used in case of higher-order modes with $n \geqq 4$. The contributions from coupling to these higher-order TE_{pn} modes to the average TE_{01} loss is small enough so that small errors in these approximations will cause no appreciable error in the final result.

For a numerical evaluation of (39) the phase constants of Figs. 2, 3 and 4 and coupling coefficients of Figs. 6, 7, 8 and 10 have been used. For higher-order TE_{pn} modes the approximate expression (31) is adequate. Interaction between circular electric waves at lining imperfections with no circumferential dependence may be described for all modes by approximation (33).

As a result of numerical evaluations, the diagrams in Figs. 11 and 12 have been drawn. They show as a function of the correlation distance L_0 the rms value of ν_p, the particular component of the thickness deviation, which by itself would increase the average TE_{01} loss by 1 per cent

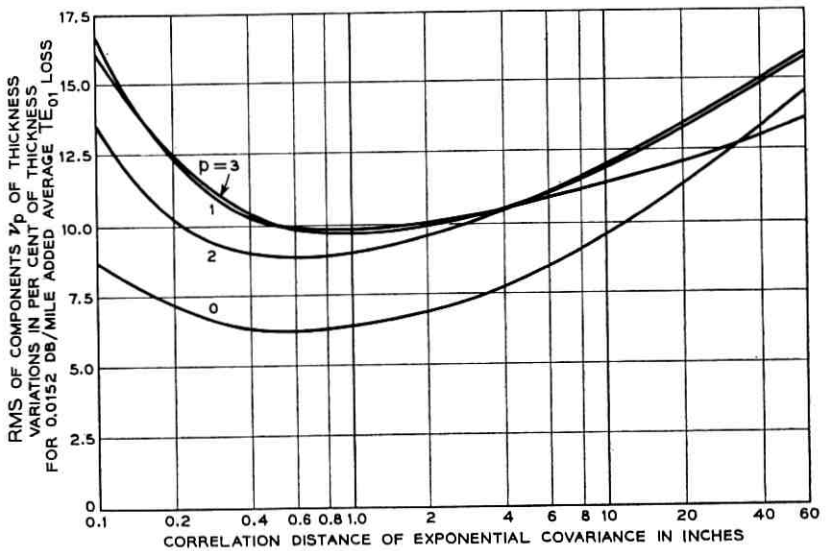


Fig. 11 — TE_{01} loss in lined waveguide with random lining imperfections of exponential covariance: 2-inch inside diameter, at 55.5 kmc, lining $\epsilon_r = 2.5$, $\delta =$ 1 per cent.

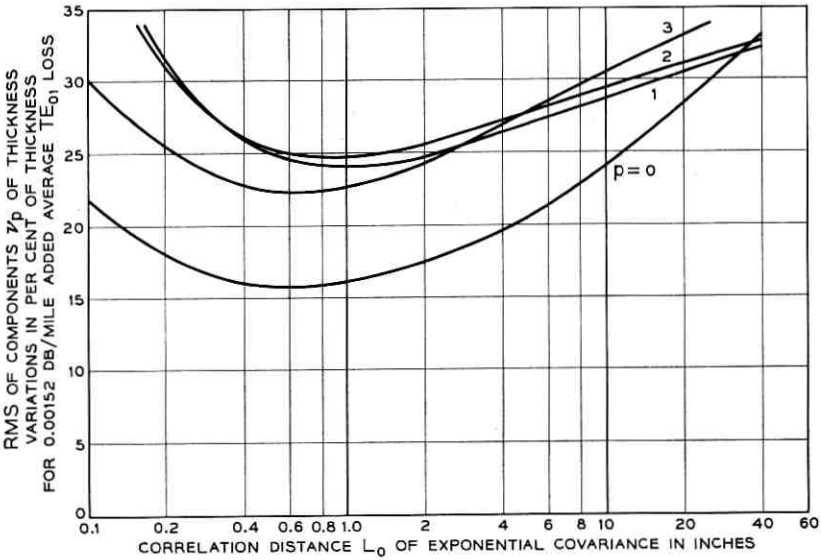


Fig. 12 — TE_{01} loss in lined waveguide with random lining imperfections of exponential covariance: 2-inch inside diameter at 55.5 kmc, lining $\epsilon_r = 2.5$, $\delta = 0.5$ per cent.

in Fig. 11 and 0.1 per cent in Fig. 12 of its value in perfect copper waveguide.

All curves show a minimum. Among all the ν_p the strictest tolerance is imposed on ν_0 by the interaction between circular electric waves. In this case for the critical correlation distance $L_0 = 0.6''$ the rms of thickness deviations of a lining of $\delta = 1$ per cent should, according to Fig. 11, be less than 6 per cent. In absolute values the 0.01″ thick lining should be made uniform to $\pm 0.0006''$.

When the lining is thinner, the tolerances are considerably eased. According to Fig. 12 and noting that the extra loss varies as the square of the thickness deviations, a 0.005″ thick lining needs only to be made uniform to $\pm 0.0025''$ for the circular electric wave interaction to increase the average TE_{01} loss not more than by 1 per cent.

Mode interaction at lining imperfections not only increases the average TE_{01} loss, but also degrades the transmission characteristics which in perfect waveguide would be smooth functions of frequency. Signals transmitted through imperfect waveguide will suffer amplitude and phase distortion.

A radio-frequency pulse of rectangular shape might be considered a standard signal. Mode conversion-reconversion effects will cause recon-

version forerunners to precede the output pulse and tails to follow it.[8] If too large in amplitude, these forerunners and tails will in a *PCM* system produce errors in regeneration.

The *rms* amplitude of the reconversion tail is largest after a long signal pulse. Immediately after this pulse, the contribution to the rms value by reconversion from one coupled mode is given by[9]

$$\sqrt{\overline{|q|^2}} = C_{mn}{}^2 \varphi(\Delta\beta_{mn}) \sqrt{\frac{L}{2\Delta\alpha_{mn}}}. \tag{40}$$

In (40) $\Delta\alpha_{mn}$ is the difference in attenuation constant between TE_{0m} and the coupled mode n. φ is the spectral distribution of ν_p given in terms of the covariance by

$$\varphi(\xi) = \int_{-\infty}^{+\infty} R(z) e^{-j\xi z}\, dz. \tag{41}$$

For the exponential covariance of (38) the spectral distribution is

$$\varphi(\xi) = \langle \nu_p{}^2 \rangle \frac{4\pi L_0}{4\pi^2 + \xi^2 L_0{}^2}. \tag{42}$$

The contributions to the reconversion tail from different modes are caused by components of the spectral distribution $\varphi(\xi)$ at the corresponding $\xi = \Delta\beta_{mn}$. From one coupled mode to the next these are quite different spectral components. Their contributions might therefore be assumed to be uncorrelated. Then the total contribution to the rms value is obtained from the sum of the squares of each single contribution:

$$\overline{|q_t|^2} = \sum_n C_{mn}{}^4 \varphi^2(\Delta\beta_{mn}) \frac{L}{2\Delta\alpha_{mn}}. \tag{43}$$

The most critical signal distortion will undoubtedly be caused by circular electric wave interaction at lining imperfections of order $p = 0$. Not only does this interaction increase the average TE_{01} loss most strongly, but higher circular electric waves are also propagating with extremely low loss. They will, therefore, contribute terms to the sum of (43) which are large because the denominator $\Delta\alpha_{mn}$ is small. Signal distortion due to circular electric wave interaction should therefore be considered first.

In Fig. 13 the quantity $\frac{|\overline{q_t{}^2}|}{L\,\overline{\nu_0{}^2}}$ has been plotted as a function of the correlation distance L_0. Mode interaction between TE_{01} and TE_{02} contributes most strongly to the reconversion tail. Therefore the curve has

a maximum where the $TE_{01} - TE_{02}$ term in (43) has its maximum value. Since in (42) the spectral distribution assumes its largest value when $L_0 = \frac{2\pi}{\xi}$, the terms in (43) will show their largest value when

$$L_0 = \frac{2\pi}{\Delta\beta_{mn}}$$

that is, when the correlation distance is equal to the beat wavelength between the two coupled modes. In Fig. 13 the maximum occurs at the beat wavelength between TE_{01} and TE_{02}.

For this most critical spectral distribution of random imperfections the total reconversion tail is still quite small. For example, let $\sqrt{\overline{\nu_0^2}} = 0.2$ be the rms of thickness deviations; then for a waveguide length of $L = 20$ miles the rms of the reconversion tail amplitude is still more than 30 db down compared to the signal pulse. Tolerances on lining imperfections, therefore, need not be as strict for signal distortion as they must be for the increase in average TE_{01} loss. In the present example the increase in average TE_{01} loss would be nearly 4 per cent.

At higher frequencies an increase of mode conversion is indicated by the linear frequency dependence of coupling coefficients according to

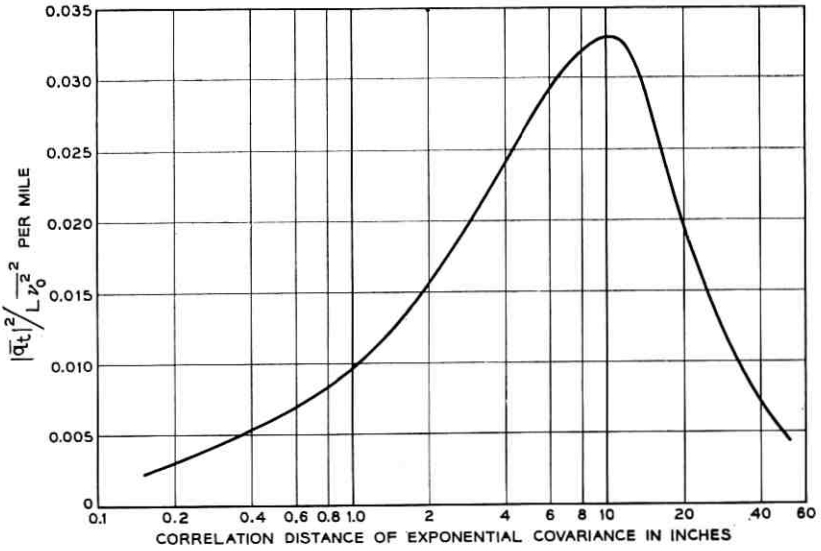


Fig. 13 — TE_{01} — pulse distortion in lined waveguide with random imperfections of the lining. Mode conversion causes a reconversion tail of rms amplitude $\sqrt{|\bar{q}_t|^2}$ to follow the pulse. $b/\lambda = 4.70$, $\epsilon_r = 2.5$, $\delta = 1$ per cent.

(34), (35) and (36). In addition the beat wavelength between coupled modes increases linearly with frequency. The coupling coefficients will also be larger for higher permittivities of the lining.

For operation at higher frequencies or with linings of larger permittivities, tolerances will have to be chosen correspondingly tighter.

VI. GENERALIZED TELEGRAPHIST'S EQUATIONS FOR DEFORMED WAVEGUIDE

Cross-sectional deformations of the lined waveguide will couple circular electric waves to unwanted modes and like imperfections of the lining degrade the transmission by mode conversion and reconversion.

To analyze wave propagation in deformed waveguide, a wall impedance representation will be used to formulate the boundary conditions at the surface of the lining.[6] In perfectly normal waveguide the boundary conditions of a perfectly uniform lining are

$$E_z = -Z_z H_\varphi \tag{44}$$

$$E_\varphi = Z_\varphi H_z . \tag{45}$$

Small deformations cause the radius a of Fig. 14 to be a function of φ,

$$a = a_0[1 + \sigma(\varphi)]. \tag{46}$$

Any function σ, being periodic in φ, may be expanded into a Fourier series:

$$a = a_0[1 + \Sigma_q \sigma_q \cos q\varphi]. \tag{47}$$

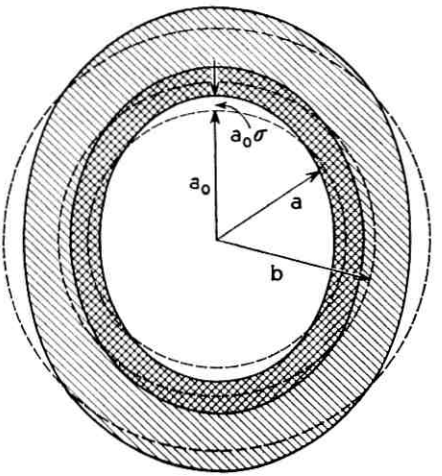


Fig. 14 — Lined waveguide with cross-sectional deformation.

Sine terms have been omitted; they would only add terms of orthogonal polarization. The deformation is assumed to be small and smooth:

$$\sigma \ll 1, \qquad \frac{d\sigma}{d\varphi} \ll 1. \tag{48}$$

For the deformed waveguide the boundary conditions will now have to be satisfied at $r = a$ and not at the nominal radius a_0. In (44) and (45) the tangential components of the magnetic field require the tangential electric field to have a certain value. The magnetic field may be considered to excite the electric field. In the deformed waveguide this excitation stays nearly the same, but due to the slight displacement of the wall it now occurs at $r = a$.

The electric field at $r = a$ can by Taylor series expansion be written in terms of the field at $r = a_0$. Neglecting higher-order terms

$$E(a) = E(a_0) + \frac{\partial E(a_0)}{\partial r} a_0\sigma \tag{49}$$

and the boundary conditions are:

$$E_z(a_0) + \frac{\partial E_z(a_0)}{\partial r} a_0\sigma = -Z_z H_\varphi(a_0) \tag{50}$$

$$E_\varphi(a_0) + \frac{\partial E_\varphi(a_0)}{\partial r} a_0\sigma + E_r(a_0) \frac{d\sigma}{d\varphi} = Z_\varphi H_z(a_0). \tag{51}$$

Maxwell's equations for $r < a$ are given by (6) to (11) with $\epsilon_1 + \epsilon_2 = 1$. Also, the representation of the transverse field components for $r < a$ in terms of normal modes of the perfect lined waveguide is as in (4). Substituting for the transverse field components from (4) into (11), the longitudinal electric field is obtained:

$$E_z = j\omega\mu_0 \sum_n I_n \frac{\chi_n^2}{k^2} T_n. \tag{52}$$

The longitudinal magnetic field, however, cannot be obtained from (8) since the series expression for E_φ in (4) is nonuniformly convergent and must not be substituted for differentiation into (8). We only substitute for E_r from (4), multiply (8) by T_m', and integrate over the cross section. After partial integration:

$$-j\omega\mu_0 \int_s H_z T_m' \, dS = \int_0^{2\pi} E_\varphi T_m' a_0 d\varphi \tag{53}$$

$$- \sum_n V_n \int_0^{2\pi} T_m' \left[\frac{\partial T_n'}{r\partial\varphi} - d_n \frac{\partial T_n'}{\partial r} \right] a_0 \, d\varphi + \sum_n V_n d_n x_n^2 \int_s T_n' T_m' \, dS.$$

To find relations for the current and voltage coefficients, substitute (4) for the field components into Maxwell's equations. Then add

$$-\frac{\partial T_m}{r\partial\varphi} + d_m \frac{h_m^2}{k^2}\frac{\partial T_m'}{\partial r} \quad \text{times (6)}$$

and

$$\frac{\partial T_m}{\partial r} + d_m \frac{h_m^2}{k^2}\frac{\partial T_m'}{r\partial\varphi} \quad \text{times (7)}$$

and integrate over the nominal cross section. The result is

$$\begin{aligned}\frac{dV_m}{dz} + j\frac{h_n^2}{\omega\epsilon_0} I_m = {} & \int_S (\text{grad } E_z)(\text{grad } T_m)\, dS \\ & + d_m \frac{h_m^2}{k^2}\int_S (\text{grad } E_z)(\text{flux } T_m')\, dS \\ & - j\omega\mu_0 \sum_n I_n \frac{\chi_n^2}{k^2}\left[\int_S (\text{grad } T_n)(\text{grad } T_m)\, dS\right. \\ & \left. + d_m \frac{h_m^2}{k^2}\int_S (\text{grad } T_n)(\text{flux } T_m')\, dS.\right.\end{aligned} \tag{54}$$

After partial integration on the right-hand side of (54),

$$\begin{aligned}\frac{dV_m}{dz} + j\frac{h_m^2}{\omega\epsilon_0} I_m = {} & \int_0^{2\pi} E_z\left(\frac{\partial T_m'}{\partial r} + \frac{d_m}{a_0}\frac{h_m^2}{k^2}\frac{\partial T_m'}{\partial\varphi}\right) a_0 d\varphi \\ & + \chi_m^2 \int_S E_z T_m\, dS - j\omega\mu_0 \sum_n I_n \frac{\chi_n^2}{k^2}\left[\int_0^{2\pi} T_n\left(\frac{\partial T_m}{\partial r} + \frac{d_m}{a_0}\frac{h_m^2}{k^2}\frac{\partial T_m'}{\partial\varphi}\right) a_0\, d\varphi\right. \\ & \left. + \chi_m^2 \int_S T_n T_m'\, dS\right].\end{aligned} \tag{55}$$

In special cases when the lining is very thin or when there is no lining, the individual terms for E_z in (52) are zero for $r = a_0$, while E_z itself, because of the boundary condition (50), is different from zero. Then (50) is a nonuniformly convergent series, which describes E_z only in the open interval $0 < r < a_0$. Term-by-term differentiation will make the series diverge. Therefore the series had not been substituted for E_z in (54). In (55) it may now be substituted in the integral over the cross section. In the line integral, E_z from the boundary condition (50) may be substituted. Thus, instead of (55),

$$\frac{dV_m}{dz} + j\frac{h_m^2}{\omega\epsilon_0} I_m = -a_0 \int_0^{2\pi} a_0\sigma \frac{\partial E_z}{\partial r}\left[\frac{\partial T_m}{\partial r} + \frac{d_m}{a_0}\frac{h_m^2}{k^2}\frac{\partial T_m'}{\partial\varphi}\right] d\varphi. \tag{56}$$

Similarly, add

$$-\left(\frac{\partial T_m}{\partial r}+\frac{d_m}{r}\frac{\partial T_m'}{\partial \varphi}\right) \quad \text{times (9)}$$

$$-\left(\frac{1}{r}\frac{\partial T_m}{\partial \varphi}-d_m\frac{\partial T_m'}{\partial r}\right) \quad \text{times (10)}$$

and integrate over the cross section. The result is

$$\begin{aligned}\frac{dI_m}{dz}+j\omega\epsilon_0 V_m = &-a_0\int_0^{2\pi} H_z\left(\frac{\partial T_m}{r\partial\varphi}-d_m\frac{\partial T_m'}{\partial r}\right)d\varphi \\ &+d_m\int_S \chi_m^2 H_z T_m' dS+j\omega\epsilon_0\sum_n V_n d_n\frac{\chi_n^2}{k^2} \\ &\cdot\left[a_0\int_0^{2\pi} T_n'\left(\frac{\partial T_m}{r\partial\varphi}-d_m\frac{\partial T_m'}{\partial r}\right)d\varphi-d_m\int_S \chi_m^2 T_n' T_m' dS\right].\end{aligned} \tag{57}$$

The boundary condition $E_{\varphi n}=Z_\varphi H_{zn}$ of the perfect waveguide may be used for the normal mode term $\left(\frac{\partial T_n}{r\partial\varphi}-d_n\frac{\partial T_n'}{\partial r}\right)$ in (53) and (57), and (53) may be substituted into (57). The result is

$$\frac{dI_m}{dz}+j\omega\epsilon_0 V_m = -j\omega\epsilon_0 a_0 d_m\frac{\chi_m^2}{k^2}\int_0^{2\pi} T_m'\left[E_\varphi-Z_\varphi H_z\right]d\varphi. \tag{58}$$

Introducing the boundary condition (51) for E_φ we get instead of (58)

$$\frac{dI_m}{dz}+j\omega\epsilon_0 V_m = -j\omega\epsilon_0 a_0 d_m\frac{\chi_m^2}{k^2}\int_0^{2\pi} T_m'\left[a_0\sigma\frac{\partial E_\varphi}{\partial r}+\frac{d\sigma}{d\varphi}E_r\right]d\varphi. \tag{59}$$

Partially integrating the last term under the integral

$$\int_0^{2\pi} T_m'\frac{d\sigma}{d\varphi}E_r d\varphi = -\int_0^{2\pi}\sigma\left[\frac{\partial E_r}{\partial\varphi}T_m'+E_r\frac{\partial T_m'}{\partial\varphi}\right]d\varphi$$

and substituting for E_r and E_φ from (4) and for $\left[\frac{1}{a_0}\frac{\partial T_n}{\partial\varphi}-d_n\frac{\partial T_n'}{\partial\varphi}\right]$ from the boundary condition in perfect waveguide, the other set of generalized telegraphist's equations is obtained:

$$\begin{aligned}\frac{dI_m}{dz}+j\omega\epsilon_0 V_m = &\ j\frac{a_0}{\omega\mu_0}d_m\chi_m^2\int_0^{2\pi} E_r\frac{\partial T_m'}{\partial\varphi}\sigma d\varphi-j\sum_n V_n d_n d_m\frac{k_n^2 k_m^2}{\omega\mu_0 a_0^2} \\ &\cdot\int_0^{2\pi}\left[1-j\frac{Z\varphi}{\omega\mu_0 a_0}\right]T_n' T_m'\sigma d\varphi.\end{aligned} \tag{60}$$

The interest is limited here to propagation characteristics of circular electric waves. Therefore, only terms that describe direct interaction between circular electric and other waves need to be retained in (56) and (60). When V_m and I_m are amplitudes of other modes, then E_z, H_φ and E_r of circular electric waves are zero. Thus (56) and (60) reduce to

$$\begin{aligned} &\frac{dV_m}{dz} + j\frac{h_m^2}{\omega\epsilon_0} I_m = 0 \\ &\frac{dI_m}{dz} + j\omega\epsilon_0 V_m = -j \sum_n V_n d_n d_m \frac{k_n^2 k_m^2}{\omega\mu_0 a_0^2}\left(1 - j\frac{Z_\varphi}{a_0\omega\mu_0}\right)\int_0^{2\pi} \sigma T_n' T_m' d\varphi. \end{aligned} \tag{61}$$

Introducing traveling waves as in (16) and (17), the more convenient form (18) of generalized telegraphist's equations is obtained.

The coupling coefficients are

$$c_{nm}^{\pm} = \frac{1}{2}\sqrt{h_n h_m}\, d_n d_m \frac{k_n^2 k_m^2}{k^2 a_0^2}\left(1 - j\frac{z_\varphi}{ka_0}\right)\int_0^{2\pi} \sigma T_n' T_m'\, d\varphi \tag{62}$$

where $z_\varphi = \sqrt{\epsilon_0/\mu_0}\, Z_\varphi$ is the wall impedance with respect to free space.

Substituting from (6.1) for T_n and T_n' and from (47) for $\sigma(\varphi)$,

$$c_{nm}^{\pm} = \frac{\pi}{2}\sqrt{h_n h_m}\, d_n d_m \frac{k_n^2 k_m^2}{k^2 a_0^2} N_n N_m J_0(k_n) J_p(k_m) \sigma_p \left(1 - j\frac{z_\varphi}{ka_0}\right). \tag{63}$$

A component σ_p causes coupling only between circular electric waves and waves of circumferential order p.

The wall impedance in (63) may for all cases of present interest be approximated by:

$$z_\varphi = \frac{1}{\sqrt{\epsilon_r - 1}} \tan ka_0\sqrt{\epsilon_r - 1}\, \delta_1. \tag{64}$$

For a very thin lining the expression (63) for the coupling coefficients reduces to the coefficient for coupling at the corresponding deformation in plain metallic waveguide. We obtain $c_{nm} = 0$ for interaction between TE_{0m} and TM_{pn}. For interaction between TE_{0m} and TE_{pn} the coupling coefficients are

$$c_{[0m][pn]} = \frac{k_{m0}k_{n0}}{a_0^2\sqrt{2h_{m0}h_{n0}}} \frac{k_{n0}}{\sqrt{k_{n0}^2 - p^2}}\, \sigma_p. \tag{65}$$

The numerical evaluation of the general expression (63) has been included in the program for automatic execution. Typical results have been plotted in Fig. 15 for a waveguide with a continuous axial offset ($p = 1$),

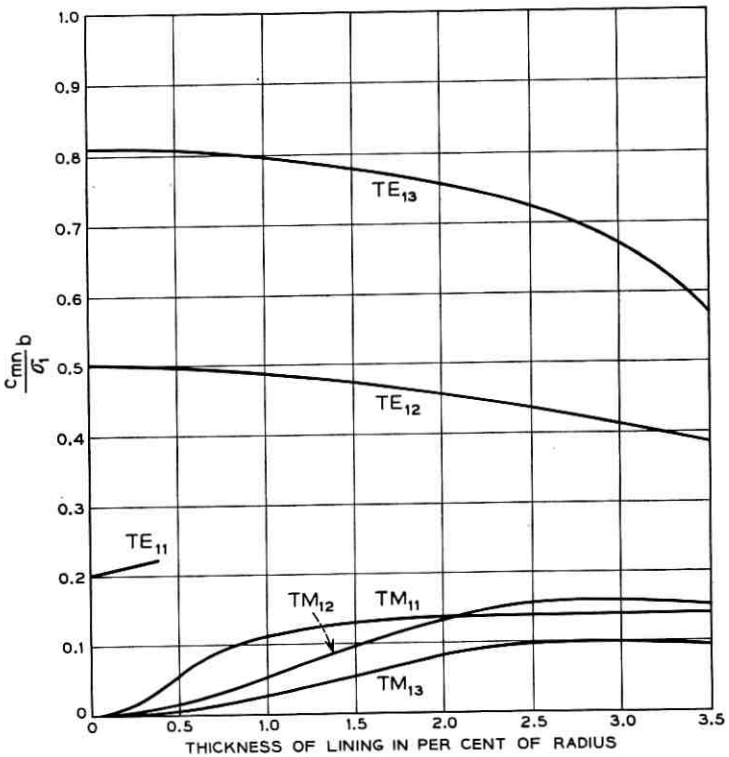


Fig. 15 — Coefficients c_{mn} of coupling between TE_{01} and TE_{1n} and TM_{1n} in axially offset lined waveguide. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

in Fig. 16 for elliptical deformation, and in Fig. 17 for trifoil deformation.

In all three cases the coupling coefficients have the following characteristics in common: The coupling between TE_{01} and TM waves is much weaker than the coupling between TE_{01} and TE waves even for a substantial thickness of the lining. For a particular deformation the coefficient of coupling between TE_{01} and TE_{pn} waves increases with the index n of the waves. Higher-order waves are coupled more strongly. The relative change of coupling between TE_{01} and TE_{pn} with the thickness of the lining is only slight in particular for coupling to higher-order waves.

The most significant consequence of these general characteristics is that mode conversion effects due to coupling at waveguide deformations will be nearly the same in lined waveguide as they are in plain wave-

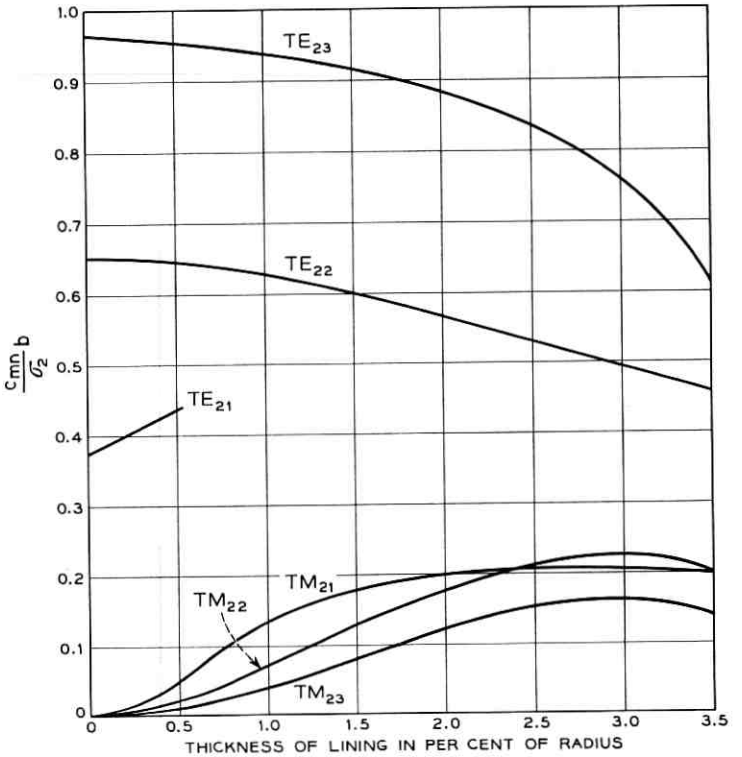


Fig. 16 — Coefficients c_{mn} of coupling between TE_{01} and TE_{2n} and TM_{2n} in lined waveguide with elliptical deformation. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

guide. Lined waveguide should therefore be manufactured to the same cross-sectional tolerances as is plain waveguide.

VII. CONCLUSIONS

TE_{01} transmission in lined waveguide is degraded by thickness variations of the lining and cross-sectional deformations. These imperfections couple the TE_{01} wave to unwanted modes. The coupling to higher circular electric waves at thickness variations of circular symmetry is strongest. Asymmetric thickness variations couple TE_{01} to the corresponding asymmetric modes.

Random thickness variations increase the average TE_{01} loss and cause signal distortion. The increase in loss is much more pronounced than signal distortion. In a typical example of random thickness variations of

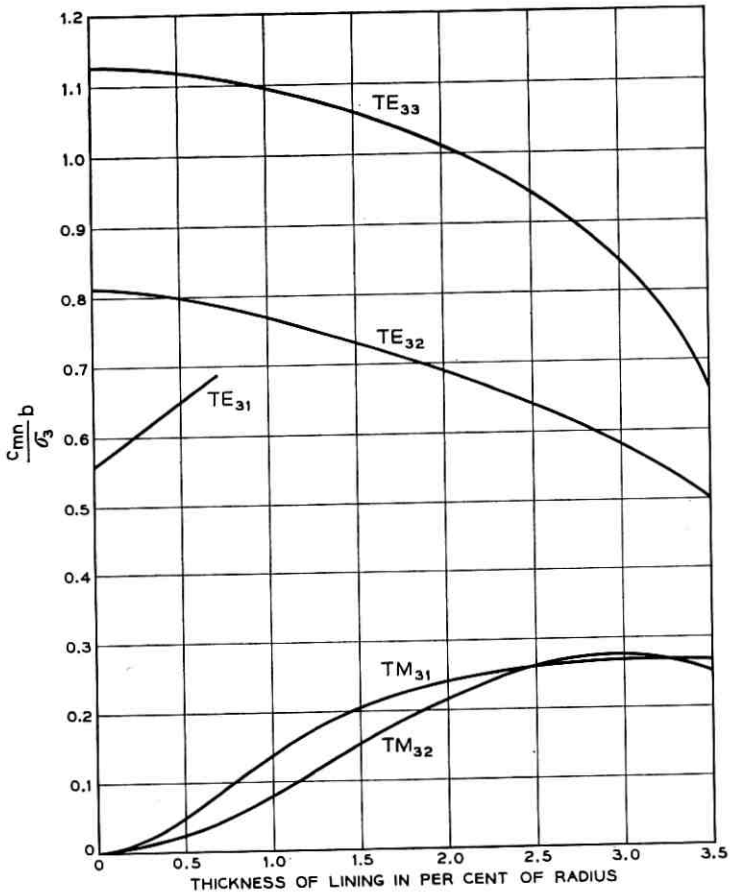


Fig. 17 — Coefficients c_{mn} of coupling between TE_{01} and TE_{3n} and TM_{3n} in lined waveguide with trifoil deformation. $b/\lambda = 4.70$, $\epsilon_r = 2.5$.

a 0.01″ lining of $\epsilon = 2.5$ in a 2″ I.D. waveguide at 55.5 kmc, the rms of the symmetric component of this variation should remain smaller than 0.002″ for the average TE_{01} loss, not to be raised by more than 10 per cent of its value in perfect copper pipe. Under these conditions a pulse signal after traveling through 20 miles of this waveguide is distorted only by a reconversion tail nearly 30 db smaller than the signal. The signal distortion, being caused by circular electric wave interaction, can not, however, be reduced by ordinary mode filters.

Cross-sectional deformations in lined waveguide cause nearly the same interaction of TE_{01} with unwanted modes as corresponding defor-

mations in plain waveguide. Lined waveguide should be manufactured to the same cross-sectional tolerances as plain waveguide.

REFERENCES

1. Unger, H. G., Circular Electric Wave Transmission in a Dielectric Coated Waveguide, B. S. T. J. **36,** September, 1957, pp. 1253–1278.
2. Katsenelenbaum, B. Z., The Effect of a Dielectric Film on the Attenuation of H_{01} Waves in a Straight Nearly Circular Waveguide, Radiotekhnika i Elektronika, 1958, pp. 38–45.
3. Noda, Ken-ichi, Circular Electric Wave Transmission through Hybrid-Mode Waveguide, Review of the Electrical Communications Lab. (Tokyo) **8,** Nos. 9–10, 1960.
4. Janssen, W., Bestimmung der Modenumwandlung im Rundhohlleiter mit unregelmässigem dielektrischen Belag. Archiv Elektr. Übertragung **15,** 1961, pp. 525–536.
5. Bressler, A. D., Joshi, G. H., Marcuvitz, N., Orthogonality Properties for Modes in Passive and Active Uniform Waveguides, J. Appl. Phys. **29,** pp. 794–799.
6. Unger, H. G., Lined Waveguide, B. S. T. J., **41,** March, 1962, pp. 745–768.
7. Rowe, H. E., and Warters, W. D., Transmission in Multimode Waveguide with Random Imperfections, B. S. T. J., **41,** May, 1962, pp. 1031–1170.
8. Miller, S. E., Waveguide as a Communication Medium, B. S. T. J., **33,** November, 1954, pp. 1209–1265.
9. Unger, H. G. Regellose Störungen in Wellenleitern, Archiv Elektr., Übertragung **15,** pp. 393–401 (1961).

Characteristics of the Service Provided by Communications Satellites in Uncontrolled Orbits

By J. D. RINEHART and M. F. ROBBINS

(Manuscript received May 22, 1962)

In a communications system which uses satellites in uncontrolled orbits, there would be times when the communication between two stations would be interrupted because no satellites would be in view. This paper provides material which can be used to describe the service which a random system of satellites would provide. An example system which could be used to provide initial world-wide service is also discussed. This system has certain nonrandom characteristics, and the effect of these characteristics on the service to various parts of the world is examined by computer simulation.

TABLE OF CONTENTS

I. INTRODUCTION

One of the several types of systems which have been proposed for world-wide communications would use numerous satellites in circular orbits 3000–8000 miles* above the earth's surface. The satellites would contain active repeaters through which any pair of stations on earth could communicate whenever a satellite is mutually visible to the stations. In this system, the satellites would move independently of each other, and the motions of the satellites relative to the earth's rotation would not be controlled. The manner in which the satellites are distributed about the earth at a time chosen at random would be much the same as if the satellites were randomly placed into position at that time.

Two ground stations will be able to communicate a large part of the time if a random system of satellites is used, but there will be times when no satellites will be mutually visible. This paper describes the service which satellites in random, or uncontrolled, orbits would provide. Service, in this context, refers to the availability of satellites for communication and does not pertain to the type or subjective quality of communication. Some of the material in this paper is an extension of a study reported earlier by Pierce and Kompfner.[1]

Sections II and III describe the geometrical properties of satellite orbits and satellite visibility. Section IV presents some methods and results which have been used to describe the service provided by a random system. Section V discusses a particular system of satellites, which approximates a random system, and describes the service which would be furnished to points throughout the world. The nonrandom

* The nautical mile is used throughout this paper.

characteristics which this system might have and the effect which these characteristics would have on the service are studied in Section VI.

II. SATELLITE ORBITS

This section describes some basic geometrical properties of satellite orbits. The discussions of service in later sections are based on the orbit properties defined here.

2.1 *Orbit Altitude*

The general orbit of an earth satellite is an ellipse which has the earth centered at one focus. Two quantities which are popularly used to describe the altitude of the elliptical orbit are the height of perigee and the height of apogee. As shown in Fig. 1, the perigee height is the distance of closest approach to the earth's surface, and the apogee height is the maximum altitude reached by the satellite. The circular orbit, which is of prime interest here, is a special case in which the apogee and perigee heights are equal. The circular-orbit altitude will be designated by H.

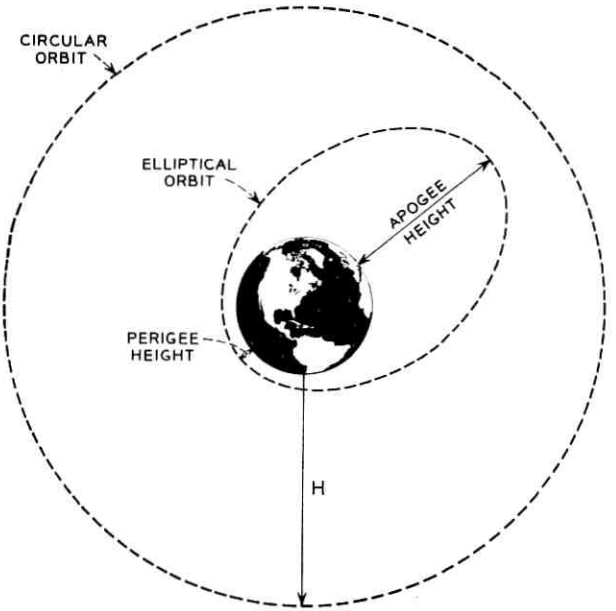


Fig. 1 — Elliptical and circular orbits.

2.2 *Satellite Period*

The period of an earth satellite, i.e., the time required for a satellite to move through 360° in its orbit, is given by

$$P = 2\pi \sqrt{\frac{a^3}{MG}} \tag{1}$$

where M is the mass of the earth, G is the universal gravitational constant, and a is the semimajor axis of the orbit. The period is plotted as a function of H in Fig. 2.

2.3 *Orientation of the Orbit Relative to the Earth*

A satellite in an equatorial orbit moves in a plane which contains the earth's equator, and the plane of a polar orbit contains the earth's

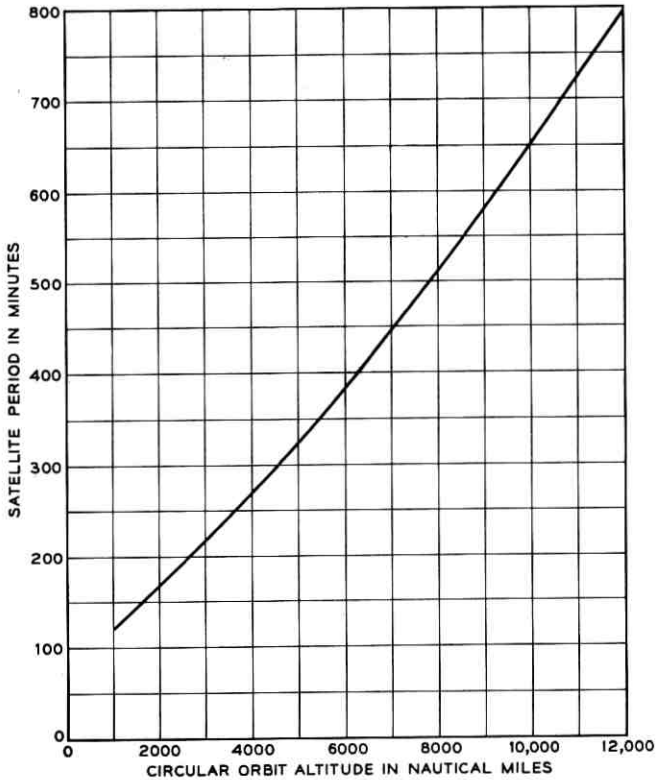


Fig. 2 — Satellite period vs circular orbit altitude. The period of elliptical orbits is found by using a circular orbit altitude which is the average of the apogee height and the perigee height of the elliptical orbit.

rotation axis. The orientation of other orbits relative to the earth is defined by the inclination angle between the equatorial plane and the orbit plane.

As shown in Fig. 3, the inclination angle is measured counter-clockwise from the equatorial plane to the orbit plane. The direction of the motion of the satellite is important, so the measurement is specified at the point at which the satellite crosses the equator from south to north (the ascending node). Satellites in orbits with inclinations less than 90° move in the direction of the earth's rotation. These orbits are easier to achieve than retrograde orbits (inclination larger than 90°) because of the inertial velocity added by the rotation of the earth.

The plane which contains a satellite orbit remains fixed in space

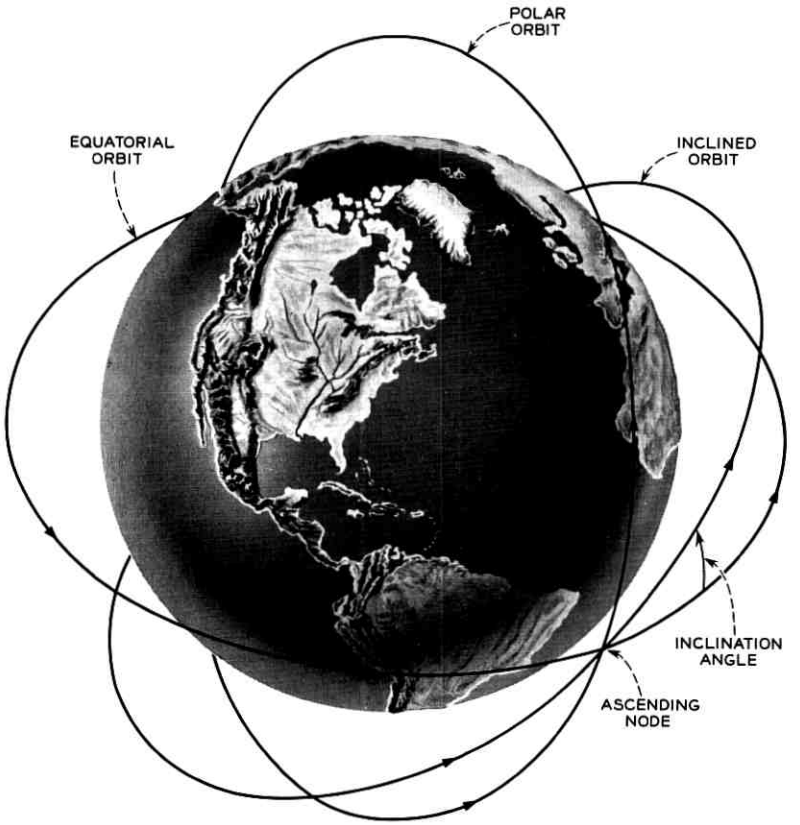


Fig. 3 — Polar, inclined, and equatorial orbits. The orientation of the orbit plane relative to the earth is given by the inclination angle.

while the earth rotates under the orbit.* The orientation with the earth's axis is constant, but the orientation of the orbit relative to a point on the earth continually changes.

2.4 *Relative Orientation of Orbit Planes and Relative Positions of Satellites*

The orientation of a system of several orbit planes is not adequately described by the inclination angle, for the planes can have any relative orientation around the earth's axis. This orientation will be specified by giving the angular position of the ascending nodes. The angle is measured in the equatorial plane from an arbitrary reference. Thus, if the ascending nodes of three polar orbits are at 0°, 120°, and 240°, the planes are uniformly spaced.

The position of a satellite in its orbit at a particular time will be given by a phase angle. This phase is the angular separation between the ascending node and the satellite. The angle is measured at the center of the orbit and is positive in the direction of satellite motion. A satellite over the north pole would thus have a phase of 90°.

The relative positions of several satellites can be completely described by giving an altitude, inclination, relative ascending node, and phase for each satellite. In a random system, all satellites usually have the same altitude, and one or several inclinations may be specified. At a time chosen at random, all ascending nodes and phases would be equally probable. In some systems one degree of randomness is removed by specifying the relative ascending nodes. Also, several satellites with the same inclination may have the same ascending node and would therefore be in the same plane.

III. SATELLITE VISIBILITY

Satellite communication can take place between two ground stations only if a satellite is visible simultaneously from both stations. Thus, mutual visibility conditions determine the service which a system of satellites provides to the stations.

3.1 *Visibility Geometry*

A satellite is considered usable for communication if its elevation angle above the theoretical horizon is larger than some specified α.† This

* This is true only for polar and equatorial orbits. The effect of the small motion of inclined orbits will be taken up in Section VI.

† For convenience in terminology, the satellite will be said to be *visible* only when this condition is met.

angle limitation is influenced by local terrain, performance of the radio system,[1] and considerations of interference with terrestrial systems.[2] A value of 7.5° will be used for α in the studies presented here, since this value has been used in most previous discussions.

The angle α defines a cone of satellite visibility as shown in Fig. 4a. The orbit of a satellite at altitude H intersects this cone in the two points at which visibility begins and ends. If the orientation of the orbit relative to the ground station is arbitrary, the locus of all possible extremes of visibility is a circle on the cone. Projecting these points radially onto the earth results in a circle with angular-great-circle radius ψ as shown in Fig. 4b, where

$$\Psi = \cos^{-1}\left(\frac{R}{R+H}\cos\alpha\right) - \alpha \tag{2}$$

A station could see a satellite which is at an altitude H whenever it is over this circular region. Two stations can communicate only if their circles of visibility intersect and form a mutual visibility region.

Some mutual visibility regions for 6000-mile orbits are shown in Fig.

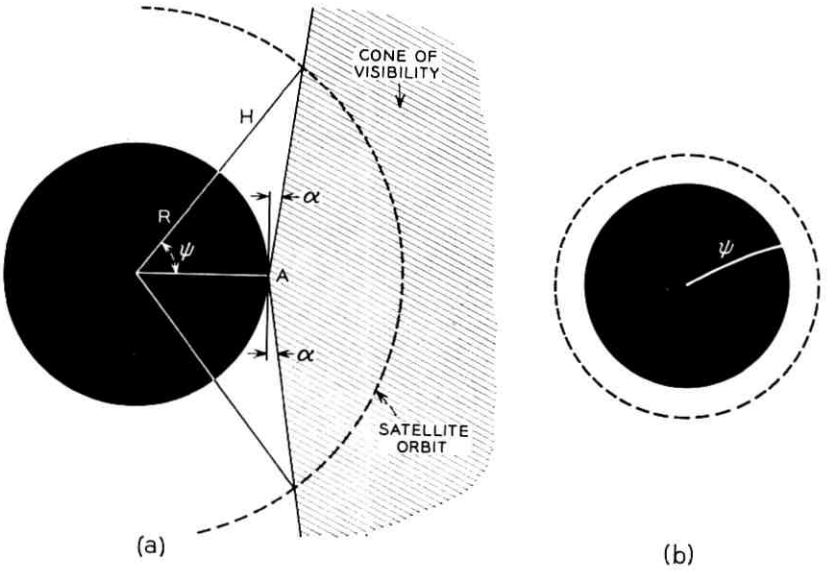


Fig. 4 — Satellite visibility geometry: (a) Plane view of the cone of visibility for a ground station at A. α is the minimum usable elevation angle. (b) Circle of visibility. A satellite of altitude H would be more than α degrees above the horizon at the ground station whenever it is over this circular region centered on the ground station.

5. Because of the polar projection used, the parts of the regions which are south of the equator are not shown. Satellites in polar orbits could pass over any part of these regions, but those with lower inclinations would not move further north than the latitude corresponding to the inclination angle. Thus if the motion of the satellites is also considered, the area of mutual visibility would be reduced for lower-inclined orbits, and it would degenerate to a line for equatorial orbits.

Fig. 5 shows a region in which a satellite in a polar orbit could be

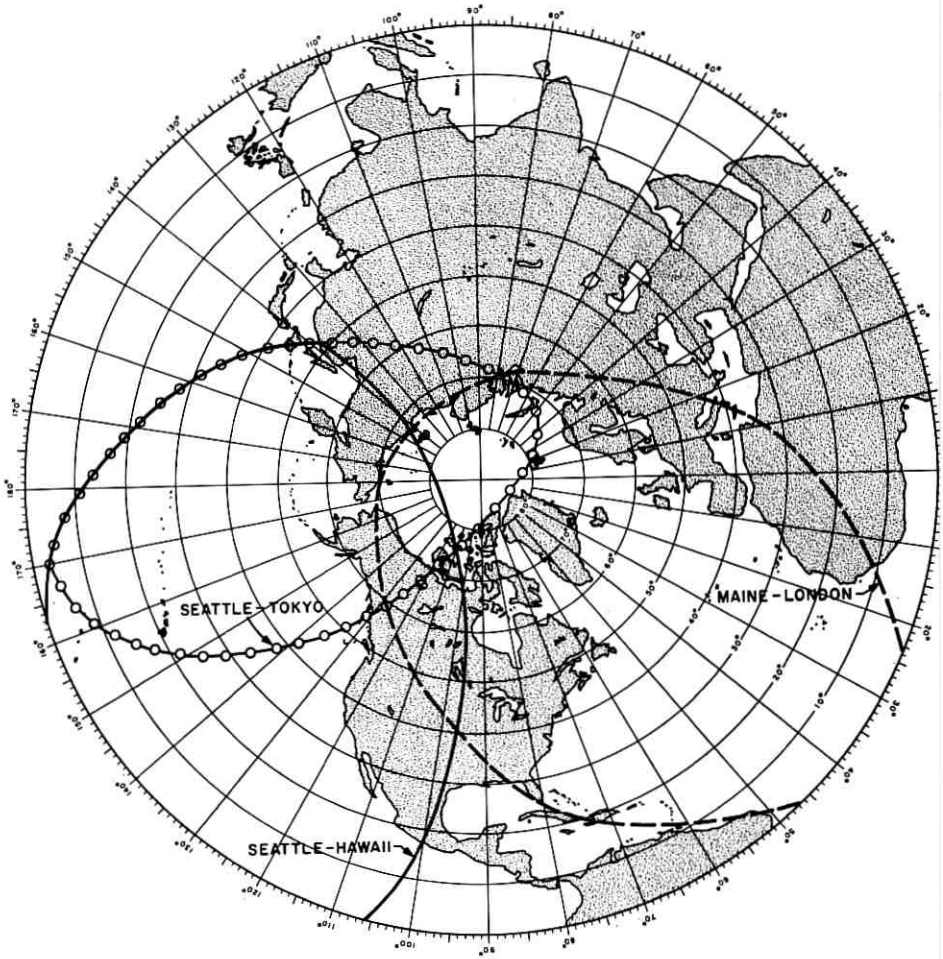


Fig. 5 — Mutual visibility regions for 6000-mile polar orbits. $\alpha = 7.5°$. Portions of the visibility regions below the equator are not shown on this projection.

seen simultaneously by all three pairs of stations. This region and the other regions of overlap present a problem of competition for the use of satellites which are over these regions. A discussion of this problem is given in Section 4.1.3.

3.2 *The Probability that a Satellite is Mutually Visible*

The description of random-orbit service requires a knowledge of the probability that a satellite in a particular orbit is mutually visible to two stations at a randomly chosen time. A numerical method is used to obtain an estimate of this probability, which will be designated as p.

3.2.1 *Numerical Determination of p*

Fig. 5 can be used to determine p for Seattle and Tokyo for 6000-mile polar orbits. The earth rotates under the orbit, so at a randomly-chosen time the orbit plane can intersect the earth along any meridian. The meridian lines in the figure thus represent a uniform sample of all orientations of the plane. All phases are equally likely for the satellite at the chosen time. For a particular orientation of the plane, the fraction of the corresponding meridian line which is contained in the mutual visibility region is the probability that the satellite can be seen. The average of the fractions determined for all meridians in the sample is an estimate of p.

For an equatorial orbit, p is simply the fraction of the earth's equator which is contained in the visibility region. The value of p for the general inclined orbit was determined by a computer program which uses a method similar to the one described for polar orbits. A method which uses numerical integration of an expression for p is given in Ref. 3.

3.2.2 *Variations in p*

For a given orbit altitude and inclination, p is a function of the latitudes of the two ground stations and the distance between them. For pairs of points whose mutual visibility regions include one of the earth's poles, polar orbits provide a higher p than do other orbits. Regions which are centered on the equator are best served by equatorial orbits. Intermediate inclinations give the highest p for regions which are between these extremes.

For a given pair of ground stations and a given inclination, p increases with the altitude of the orbit. An example of the nature of this increase is shown in Fig. 6 which gives p vs altitude for Maine–London for polar orbits. The rate of increase of p with altitude begins to diminish

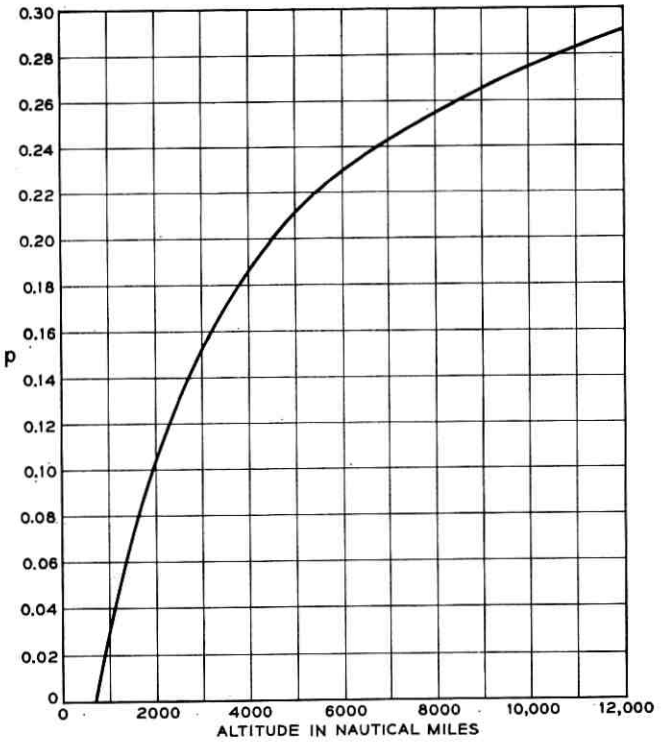


Fig. 6 — Fraction of time (p) a satellite in a circular polar orbit of altitude H is visible to Maine and London. $\alpha = 7.5°$.

rapidly at altitudes above about 8000 miles. The reason for this is apparent from (2). For high altitudes in this equation, the radius of the visibility region is approaching $(\pi/2)-\alpha$ which is the maximum value which can be attained, and the size of the visibility region directly influences the value of p. The maximum value of p for this pair of points is 0.368.

Equation (2) shows that the size of the visibility circle is also influenced by the value of α. Thus the minimum elevation angle affects the value of p for a given orbit and a given pair of ground stations. Fig. 7 gives p vs α for Maine–London and 6000-mile polar orbits. The value of p is read from the left scale. If α is zero, i.e., if the satellite were used whenever it was above the theoretical horizon, p would be 0.28. An α of 20° would result in a p of about half of this maximum value. So if α were 20°, the satellite could be used during only half of the time it is visible. The use of an α of 7.5° results in about 82 per cent utiliza-

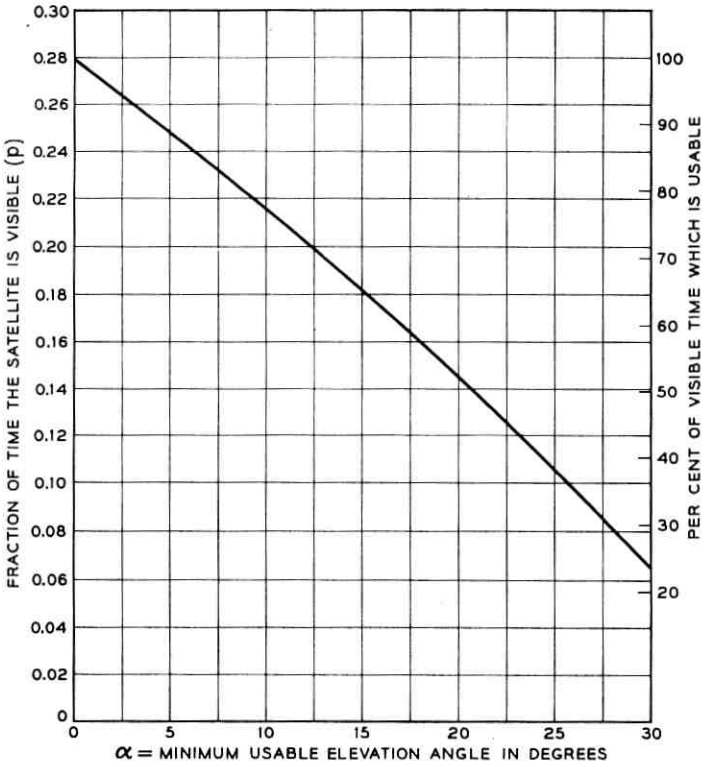


Fig. 7 — The effect of the minimum elevation angle on the usefulness of a satellite in a 6000-mile polar orbit. Ground stations at Maine and London.

tion of the visible time, so there is something to be gained in reducing this angle if it is possible.

3.2.3 *A Partial Table of p*

During the course of previous studies, values of p have been determined numerically for many ground-station locations, orbit altitudes, and inclinations. These values are given in the table in the Appendix. The table is by no means complete, but many of the pairs of points which are of interest are included.*

The first two columns in the table are the colatitudes of the ground stations. These numbers may be measured from the north pole or the south pole of the earth as long as the same convention is used for both points. The third column is the longitude separation of the two stations.

* Ref. 4 gives values of p in a more general form.

The other columns are headed by altitudes and inclinations. Linear interpolation may be used between successive altitudes and successive inclinations with reasonable accuracy.

IV. SERVICE PROVIDED BY SATELLITES IN RANDOM ORBITS

4.1 *Quality of Service, q*

A random-orbit system will use a multiplicity of satellites to provide service to many pairs of ground stations. For a particular pair of stations, the quality of service is defined as the fraction of time the required number of satellites is available for use. The symbol q is used to denote the quality of service.

4.1.1 *Satellites in Orbits of the Same Altitude and Inclination*

(i) Single-Satellite Visibility

Since p is the probability that a satellite in a certain orbit will be visible to a pair of ground stations at a time chosen at random, $(1 - p)$ is the probability that the satellite will not be visible at that time. If there are n satellites with identical p's, and if the satellites have random ascending nodes and phases, the probability that none of these satellites would be useful at a random time is $(1 - p)^n$. The fraction of time at least one satellite would be useful is then

$$q_1 = 1 - (1 - p)^n \tag{3}$$

Fig. 8 gives values of q_1 for a range of values of n and p. If p is known, the number of satellites needed to furnish a certain quality of service may be found. For example, for stations in Maine and London, and 6000-mile polar satellites, $p = 0.23$. For a quality of service of 0.999, Fig. 8 shows 27 satellites would be required. For $q_1 = 0.99$, $n = 18$; and for $q_1 = 0.90$, $n = 9$.

Equation (3) may also be written

$$n = \frac{\log (1 - q_1)}{\log (1 - p)}$$

If n is changed to n', q_1 would change to q_1', and the ratio of n'/n is given by

$$\frac{n'}{n} = \frac{\log (1 - q_1')}{\log (1 - q_1)}$$

which is independent of p. This equation was used to derive a table (next page) which contains the number of satellites needed for various q_1 relative to the number needed for $q_1 = 0.999$.

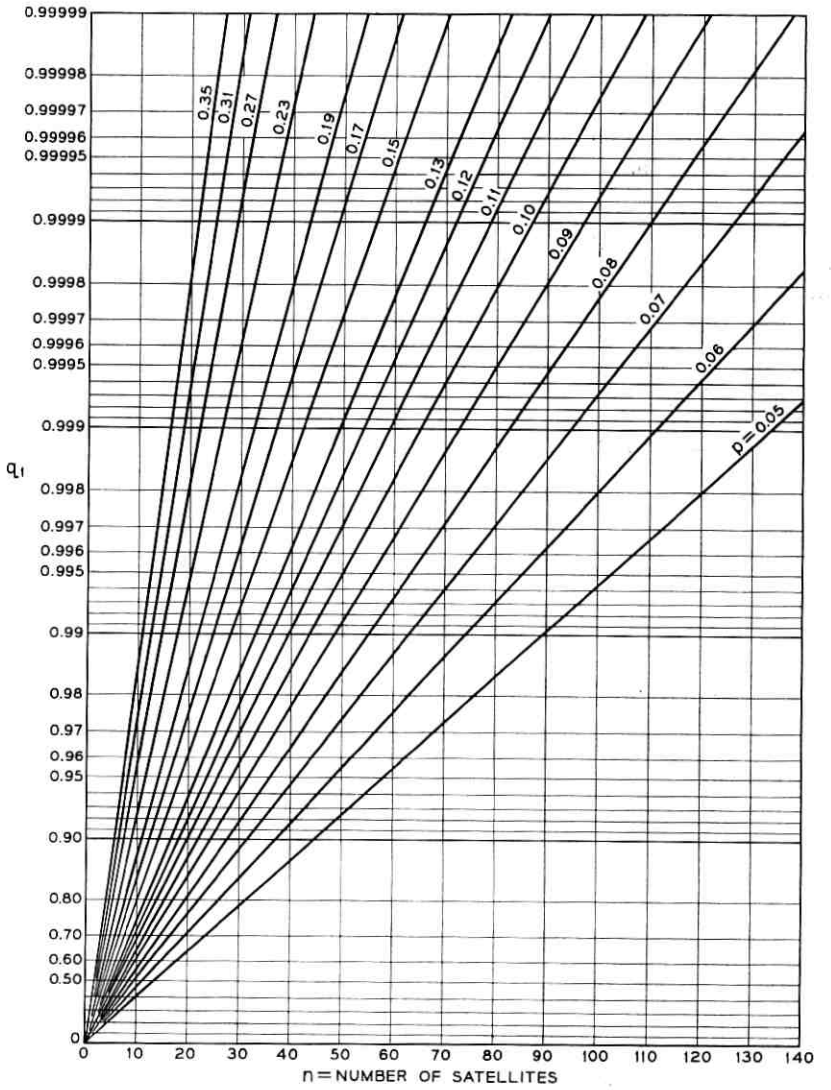


Fig. 8 — Quality of service for single satellite visibility, q_1, for a range of values of p and n.

q_1	Number of Satellites Needed
0.999	n
0.99	$0.67n$
0.95	$0.43n$
0.90	$0.33n$
0.75	$0.20n$
0.50	$0.10n$

(ii) Multiple-Satellite Visibility

Some pairs of ground stations may require more than one satellite at a time if the traffic between the two stations exceeds the capacity of one satellite. If m ground antennas were used at each station, m communication paths could be established by way of separate satellites. The quality of service on path m, which is the fraction of time m or more satellites are usable by the ground stations, is given by

$$q_m = 1 - \sum_{i=0}^{m-1} \binom{n}{i} p^i (1 - p)^{n-i} \tag{4}$$

Figure 9 gives values of q_m as a function* of q_1. For a value of $q_1 = 0.999$ on the abscissa, the intersections of the ordinate line with the curves give $q_2 = 0.991$, $q_3 = 0.964$, etc. q_1, q_2, and q_3 would be the qualities of service for three paths between Maine and London if the paths were operated on a priority basis, i.e., if path 3 operated only when three or more satellites were visible, path 2 only when two or more were visible, and path 1 when one or more was visible. Figs. 8 and 9 may be used to determine the number of satellites needed for a particular q_m. Suppose $q_2 = 0.98$ is desired. From Fig. 9, $q_1 = 0.9975$. If we again use the example of $p = 0.23$, Fig. 8 shows 23 satellites would be needed.

Multiple paths between two points may be operated with equal priority simply by assigning periods of no service to the paths in turn. The lengths of these periods are not equal, but over a long period of time all paths would get equal qualities of service. Suppose there are three paths ($m = 3$). One of the paths would be out of service whenever only two satellites are visible, two would be out whenever only one satellite is visible, and all would be out when no satellites are visible. The fractions of these times for which a particular path would be out of service are $\frac{1}{3}$, $\frac{2}{3}$, and 1 respectively. The general fraction is $(m - i)/m$, where m is the number of paths and i is the number of satellites visible ($i < m$). The average quality of service on each of m paths with equal priority is

$$\bar{q}_m = 1 - \sum_{i=0}^{m-1} \frac{m - i}{m} \binom{n}{i} p^i (1 - p)^{n-i} \tag{5}$$

* This relationship varies slightly with p. The curves in Fig. 9 result from using $p = 0.18$. The values of p which are of interest in this work range between 0.05 and 0.30. The changes in the curves of Fig. 9 for these extreme values of p would be very small, so the curves can be used with reasonable accuracy for any p in this range.

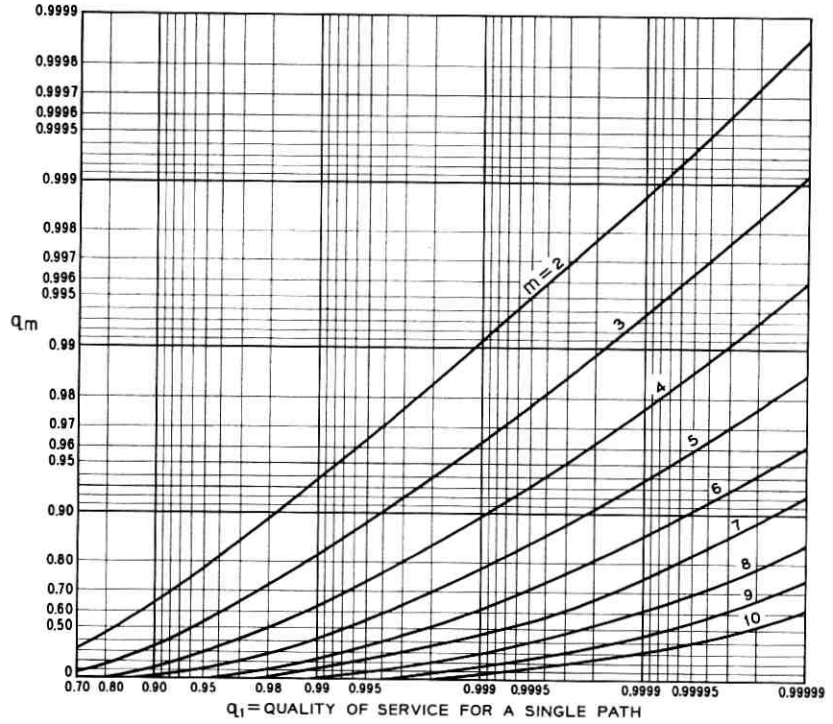


Fig. 9 — Fraction of time q_m that m or more satellites are visible. In a system with multiple paths between two ground stations q_m is the quality of service on a path with priority m. The number of satellites needed to furnish q_m is found by entering Fig. 8 with q_1 and p.

Fig. 10 gives values $\bar{q}_m$ as a function* of q_1. For $q_1 = 0.999$, $\bar{q}_2 = 0.995$, $\bar{q}_3 = 0.985$, etc. A previous example showed 27 satellites in 6000-mile polar orbits would provide $q_1 = 0.999$ between Maine and London. The above results show three equal-priority paths between these points would have qualities of service of 0.985. If 0.999 service were desired on each path, Fig. 10 shows $q_1 = 0.99996$, so 39 satellites would be needed (Fig. 8 with $p = 0.23$ and $q_1 = 0.99996$). This procedure can be used to determine the service from one point to several other points when the mutual visibility regions are almost identical. For example, stations in England, France, and Germany could be considered to be at about the same latitude and longitude, so 39 satellites would provide a

* The previous comment on the effect of different values of p also applies here.

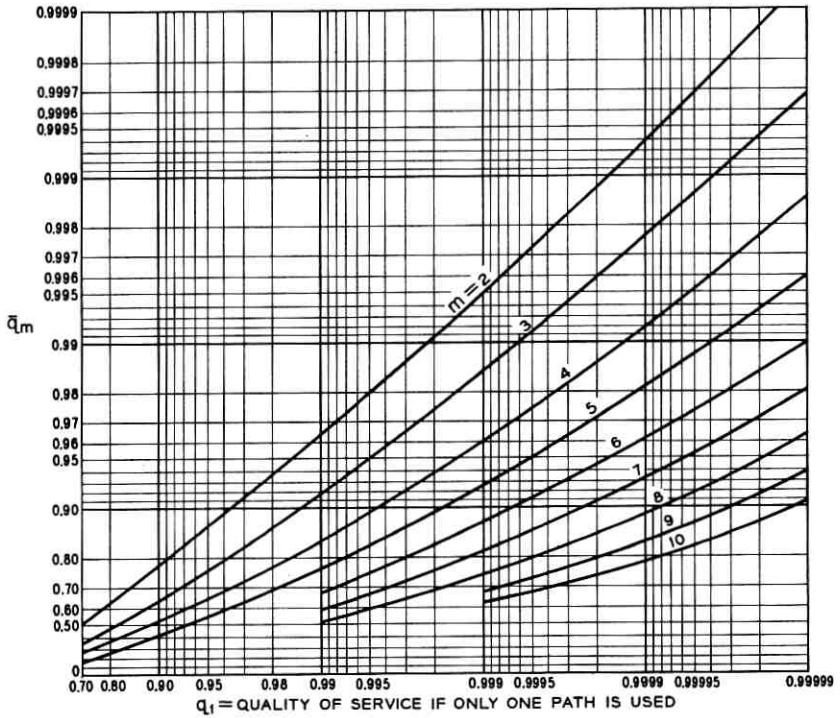


Fig. 10 — Quality of service $\bar{q}_m$ on each of m paths with equal priority. The number of satellites needed to furnish $\bar{q}_m$ is found by entering Fig. 8 with q_1 and p.

q_1 of about 0.999 to each of these points with a separate satellite for each path.

4.1.2 *Mixed Orbits*

A previous section indicated that no single orbit inclination is best for all pairs of stations. This may be verified by examination of the table in the Appendix. Miami–Rio de Janeiro, for example, would have a quality of service of only 0.80 if the 27 polar satellites discussed before were used. Equatorial orbits are best for this pair of stations, but these orbits provide little service to Maine–London. This implies that satellites should be placed in both kinds of orbits if both pairs of stations are to be served equally well. It is also possible that a single intermediate inclination would be best.

Suppose a system consists of a total of n satellites distributed in k different orbits. Let the jth orbit contain n_j satellites, and let p_j be the

fraction of time a satellite in this orbit is visible to a pair of stations. The quality of service for a single satellite is then

$$q_1 = 1 - \prod_{j=1}^{k} (1 - p_j)^{n_j} \tag{6}$$

$$\text{where} \quad n = \sum n_j$$

A solution of (6) for the above Miami–Rio de Janeiro example shows that 22 satellites in 6000-mile equatorial orbits would need to be added to the polar satellites to make $q_1 = 0.999$. However, Maine–London also gets some service from the equatorial satellite, so less than 27 polar satellites are now required to give 0.999 service to this path. Thus the solution must be adjusted until it converges to the point where 0.999 service results for both paths.

The problem of mixed orbits is concerned with the specification of the number of satellites which should be placed in each of k different orbits in order to satisfy the usability requirements of several ground stations. If the satellite altitudes are the same, the best set of orbits might be considered to be the one which requires the smallest number of satellites.* The solution may be approximated by determining the number of satellites in each orbit individually and in sequence as in the above example. This method is difficult and time-consuming for complex problems, and a more direct approach is to use a linear-programming model. Such a model has been developed, but it is beyond the scope of this paper.

A more general problem of mixed orbits involves the determination of the set of satellites, in orbits of all possible altitudes and inclinations, which will furnish required qualities of service for a minimum cost. The linear-programming model represents a large step toward the solution of this problem; however, other problems related to costs and functional relationships between other system parameters must be thoroughly investigated before the problem can be solved.

4.1.3 *The Effect of Overlap of Mutual Visibility Regions*

In Fig. 5, nearly all of the mutual visibility region for Seattle–Tokyo is contained within the regions for Maine–London and Seattle–Hawaii. When a satellite is mutually visible to Seattle and Tokyo, this satellite

* This is really an oversimplification. Depending on the location of the launching sites, the payload which a given vehicle can put in orbit depends on the inclination angle. With the possibility of launching several satellites with a single vehicle, minimizing the number of satellites does not necessarily minimize the total cost.

may already be in use by one of the other pairs of points. The previous discussion on multiple-satellite visibility indicates the problem could be solved most of the time by using other satellites which are also visible. The effect of this overlap on the qualities of service for each pair of points will be determined to illustrate the magnitude of the problem. For simplification, the overlap of the Seattle–Hawaii region with the Maine–London region will be ignored.

The Seattle–Tokyo path would be out of service if either of the following three conditions prevails:

(*i*) There is no satellite in the Seattle–Tokyo region. The fraction of time this would occur is given by $(1 - p_{S\text{-}T})^n$ where the subscript identifies the region.

(*ii*) There is exactly one satellite in region A (the area common to the Maine–London and Seattle–Tokyo regions) *and* there are no satellites in the remainder of the Maine–London region *and* there are no satellites in the remainder of the Seattle-Tokyo region. The fraction of time the event occurs is

$$[np_A (1 - p_A)^{n-1}][1 - (p_{M\text{-}L} - p_A)]^{(n-1)}[1 - (p_{S\text{-}T} - p_A)]^{(n-1)}$$

If all paths have equal priority, Seattle-Tokyo would get the satellite in one-half of the situations. Thus one-half of the above expression is the fraction of time the path is out of service because of overlap with Maine–London.

(*iii*) Same as condition (*ii*) except region B (the area common to the Seattle–Hawaii and Seattle–Tokyo regions) replaces region A and the Seattle–Hawaii region replaces the Maine–London region.

As a numerical example, a system of 27 satellites in 6000-mile polar orbits was considered. Table I summarizes the effect of the overlap.

The effect of overlap is difficult to determine in a more complex situation; but because the effect is small, one can usually account for it by nominally reducing the quality of service or nominally increasing the number of satellites. In this example, the addition of one satellite would compensate for the overlap, i.e., q_1 for 28 satellites with overlap considered would be as good as q_1 for 27 satellites with overlap not in-

TABLE I

Path	q_1 if Overlap is not Considered	q_1 including the Effect of Overlap	Fraction of Service Lost Because of Overlap
Seattle–Tokyo	0.9876	0.9841	0.0035
Seattle–Hawaii	0.9904	0.9874	0.0030
Maine–London	0.9991	0.9987	0.0004

cluded. Complex overlap situations can better be evaluated by a computer simulation which has been developed to describe the service furnished by systems of satellites.

4.2 *Description of the Periods of Service and No-Service by Simulation*

Knowledge of the fraction of time a path is out of service is not sufficient to evaluate the service provided by a system of satellites. The description of the length and frequency of these periods of no service is at least equally important. For a given quality of service, the no-service periods could be seconds long and occur frequently, or they could be days long and occur only occasionally. Certainly the way the no-service periods would be evaluated and administered would be different for these two extremes.

To describe the periods of service and no service, a computer program which simulates a world-wide satellite communications system was used. The program assigns satellites to the several ground-station pairs in a system in a way which maximizes the service on a priority basis (the paths may have equal or ordered priority). The motions of the system may be simulated for any length of time, and periods of several years may be studied with modest amounts of computer time. The periods of no service are recorded for each ground-station pair, and a statistical description of the service periods and periods of no service is provided. Both random and nonrandom systems can be studied. The results presented in this section assume random ascending nodes and phases for the satellites.

4.2.1 *Description of Service for Various Qualities of Service*

Fig. 11 displays the periods of no service on a path from Maine to London for a system of 12 satellites in 6000-mile random polar orbits (a quality of service of 0.95). This figure covers a representative 30-day period from a one-year simulation. The darkened areas, which are drawn to scale, represent the times when the path is out of service because no satellites are visible. The average length of the periods of no service (t_o) and the average time between these periods (t_i) are given for the one-year period. Fig. 12 presents the periods of no service for the Maine–London path if the 6000-mile polar orbit system has 18 satellites ($q = 0.99$), and Fig. 13 gives the same information for a system of 27 satellites ($q = 0.999$).

The communication system will probably need to be operated in such a way that periods of no service do not interrupt communication which is in progress. This would be accomplished by keeping new traffic off

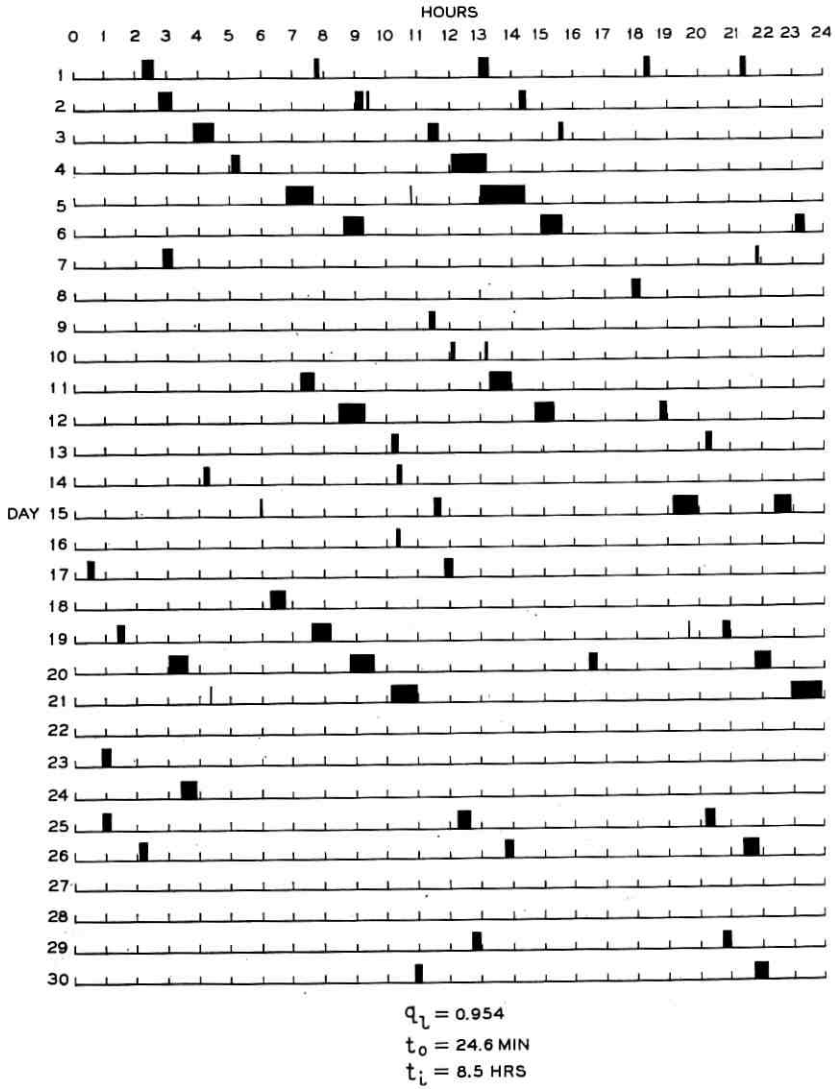


Fig. 11 — Maine–London service from 12 satellites in 6000-mile polar orbits.

the system during a "guard interval" which precedes the no-service period. The guard interval would be long enough to insure, with a reasonable probability, that all customers would be finished by the time the no-service period starts.

With this method of operation, the no-service periods can be inter-

preted simply as delays in service to new customers. The periods are predictable weeks in advance, and they are short; therefore, they are not the same as catastrophic system failures such as cable breaks and repeater failures.* As far as ordinary toll customers are concerned, a short no-service period is the same as a temporary saturation of the system by heavy traffic.† In view of the above comments, the no-service periods, which are popularly called "outages," might more appropriately be called "service delays."

The daily period of heavy traffic between most countries is only several hours long. The service delays which would occur during the busy period would probably be of primary importance, and those outside this period would be secondary. As an example assume the busy period is four hours long. Then, for the 0.99 service shown in Fig. 12, there would be no days in which more than one service delay would occur during the busy period. The absolute time in the figure is arbitrary, so the busy period could occur at any time. If all possible locations of the busy period are considered, the 30-day period shown could have as many as eight or as few as zero service delays during busy periods. The maximum number for the 0.95 service shown in Fig. 11 would be between six and twenty for that 30-day period, and the 0.999 service of Fig. 13 would have between zero and two delays during busy periods. In comparing these three services, the numbers of satellites needed in each case should also be considered since the cost of satellites in orbit will be a large part of the total system cost.

4.2.2 *The Effect of Orbit Altitude on the Service*

Fig. 14 is a representative display of the delays which would occur on the Maine–London path if a system of 18 satellites in 3000-mile random polar orbits were used. The quality of service is 0.95, which is the same quality as that of Fig. 11. In the 3000-mile system the service delays would be more frequent but of shorter duration. These differences are caused by the shorter orbital period and shorter visibility times of the 3000-mile satellites.

If the operational method discussed previously is used, a fixed guard interval would be added to each of the delays in the figures. The length of the guard interval is determined by lengths of telephone calls, so it would be the same for all cases. It is apparent that the effect of adding

* There will also be repeater failures in satellites, but the effect of these failures is only to reduce the quality of service. For example, if a system started with 27 satellites and 0.999 service (Fig. 13), failure of one-third of the satellites would only reduce the service to 0.99 (Fig. 12).

† The effect may be more severe for private line customers who require a circuit continuously.

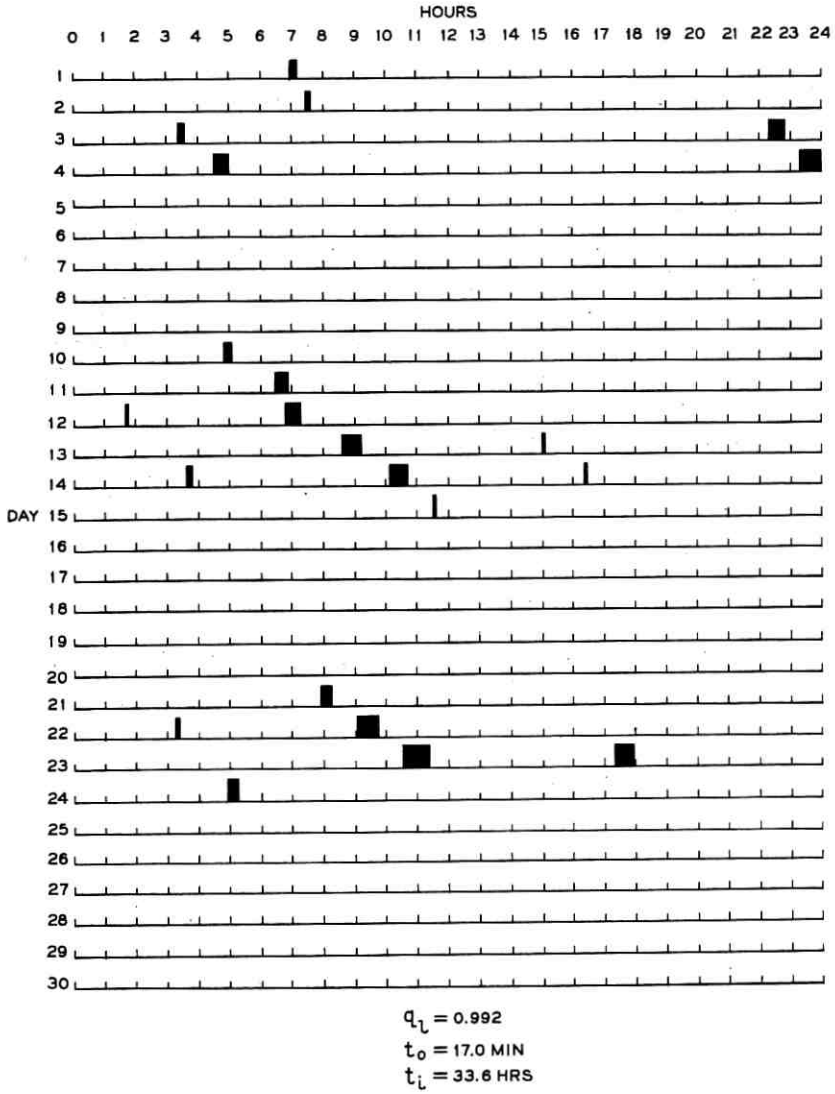


Fig. 12 — Maine–London service from 18 satellites in 6000-mile polar orbits.

the guard interval to each of the delays in Fig. 14 would be more severe than if it were added to the delays of Fig. 11.

Although it may appear that for a given quality of service a high orbit would always give better service than a lower orbit, this is not strictly true. As the satellite periods and visibility times become longer,

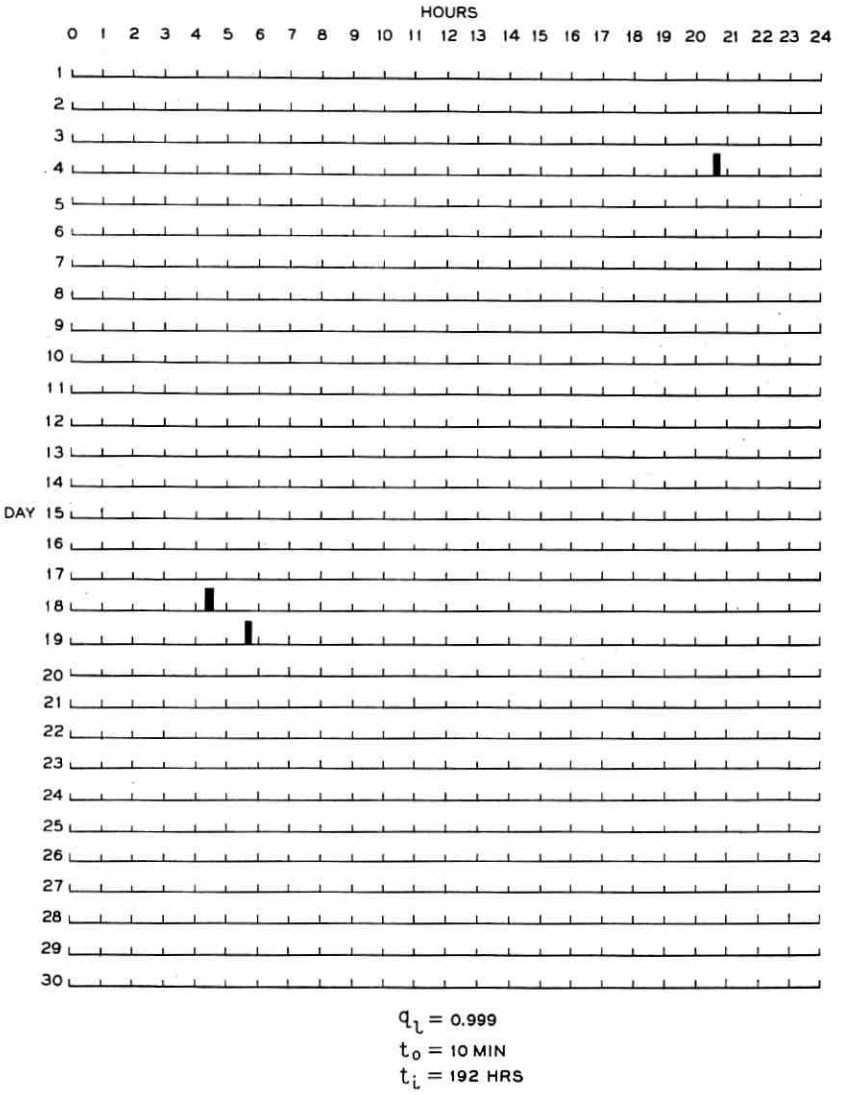


Fig. 13 — Maine–London service from 27 satellites in 6000-mile polar orbits.

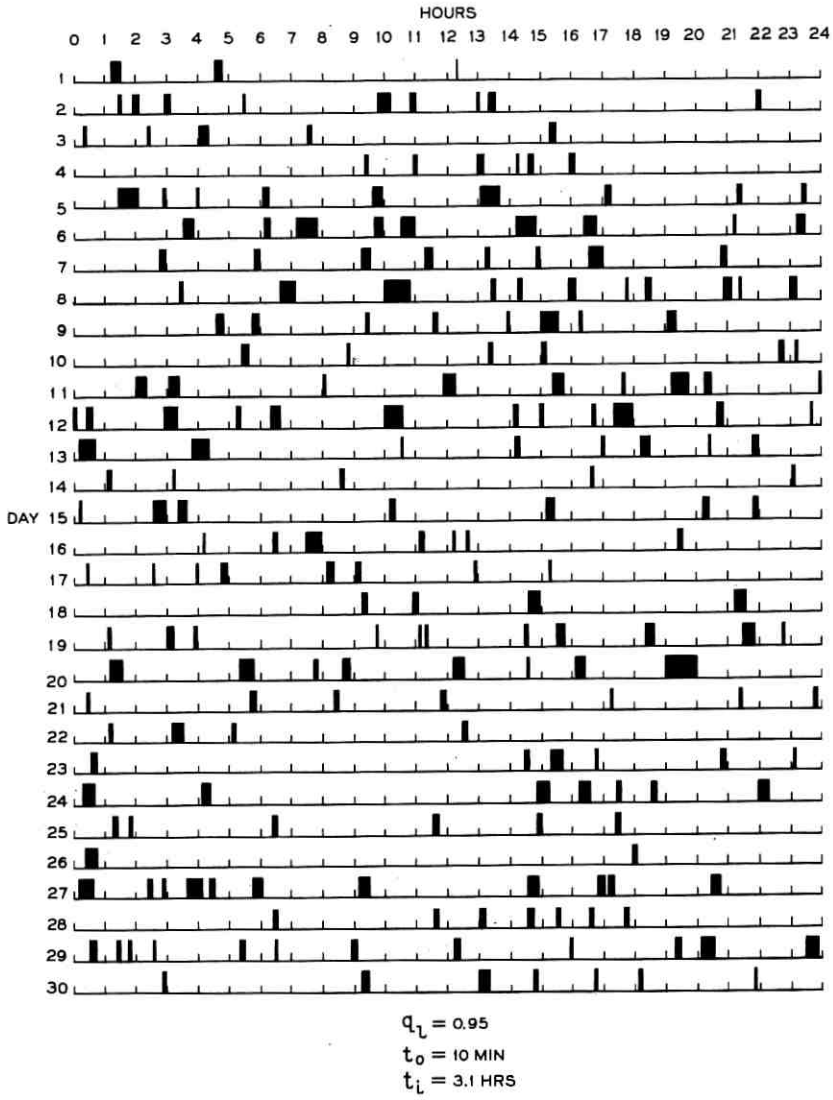


Fig. 14 — Maine–London service from 18 satellites in 3000-mile polar orbits.

the delays occur less frequently, but they also become longer. Eventually a point would be reached where the delays in service would be excessive and would approach the severity of catastrophic failures.

4.3 *Analytical Description of Periods of Service and No-Service*

A number of persons have made analytical studies of the lengths of service delays (outages) and the time between these delays. Some of these expressions are included here for completeness. Additional results are given in a companion paper by S. O. Rice[5] who makes use of a traffic model. The results predicted by the model compare favorably with simulated results described in the previous section.

The average length of the service delays is given by

$$t_o = \frac{(1 - p)T}{n} \tag{7}$$

n and p have been defined previously, and T is the average time between successive appearances of a particular satellite. If the visibility region were circular and centered on the north pole, T for polar orbits would be equal to the orbital period P, which is given in Fig. 2. For polar orbits, P is a good approximation for T for regions which include one of the earth's poles such as the Maine–London region in Fig. 5. For other regions, such as Seattle–Hawaii, only a fraction of all possible orientations of the orbit plane with the earth result in visibility, so there will be some times longer than P between successive appearances. An expression for T for these cases has not been determined. For equatorial orbits, T would equal P if the earth were not rotating. For the kind of equatorial orbits assumed here, the satellite moves in the direction of the earth's rotation, so the time between successive appearances above a point on the earth is longer than P. In this case, the expression for T is

$$T = \frac{P}{1 - \frac{P}{24}} . \tag{8}$$

The average time between service delays is given by

$$t_i = \frac{q_1}{1 - q_1} t_o . \tag{9}$$

Some useful approximations for the distribution of the lengths of delays and times between delays have also been given by Ref. 5. The fraction

of delays which would be longer than t is approximately $\exp(-t/t_o)$, and the fraction of delays which are separated by a time longer than t is approximately $\exp(-t/t_i)$.

Equations (7) and (9) and the exponential approximations for the distributions have been compared with many simulated results, and the agreement testifies favorably for the usefulness of these expressions in describing random-orbit service analytically.

V. AN EXAMPLE OF A WORLD-WIDE SYSTEM

The material given in the previous sections discusses techniques which can be used to describe the service provided by a set of satellites. The application of these techniques is illustrated in this section by describing the world-wide service which would be provided by a particular system of satellites in 6000-mile orbits. This example system provides essentially uniform service throughout the world with a modest number of satellites.

5.1 *Description of the Orbits*

This system would consist of 24 satellites in 6000-mile circular orbits. It is assumed that four satellites would be placed into orbit with a single rocket, so each plane would contain four satellites. Three of these planes would be polar, and three would have an inclination of 28°.* The three planes in each group would be uniformly spaced (their ascending nodes would be 120° apart) as shown in Fig. 15. The lines on the figure represent the intersections of the orbit planes with the surface of the earth at an arbitrary time. A discussion of the motion of the orbit planes and its effect on service is given in Section VI. A possible scheme for effecting the separation of the multiply-launched satellites is also discussed in Section VI. For the present, it is assumed that the 24 satellites are randomly phased.

5.2 *Service from the United States*

Fig. 16 shows the qualities of service which would be provided by the system of 24 satellites if stations were placed in Maine, Seattle, Puerto Rico, and Guam. A point within the 0.99 contour would be able

* Equatorial orbits would be more useful in this system than 28° orbits; however, a rocket launched from Cape Canaveral must use difficult, power-consuming maneuvers to achieve an equatorial orbit. It is shown in Ref. 6 that the Atlas Agena can place about 1300 pounds into a 6000-mile, 28° orbit but can place only about 200 pounds into an equatorial orbit of the same altitude. About 1000 pounds can be placed into a 6000-mile polar orbit from the Pacific Missile Range.

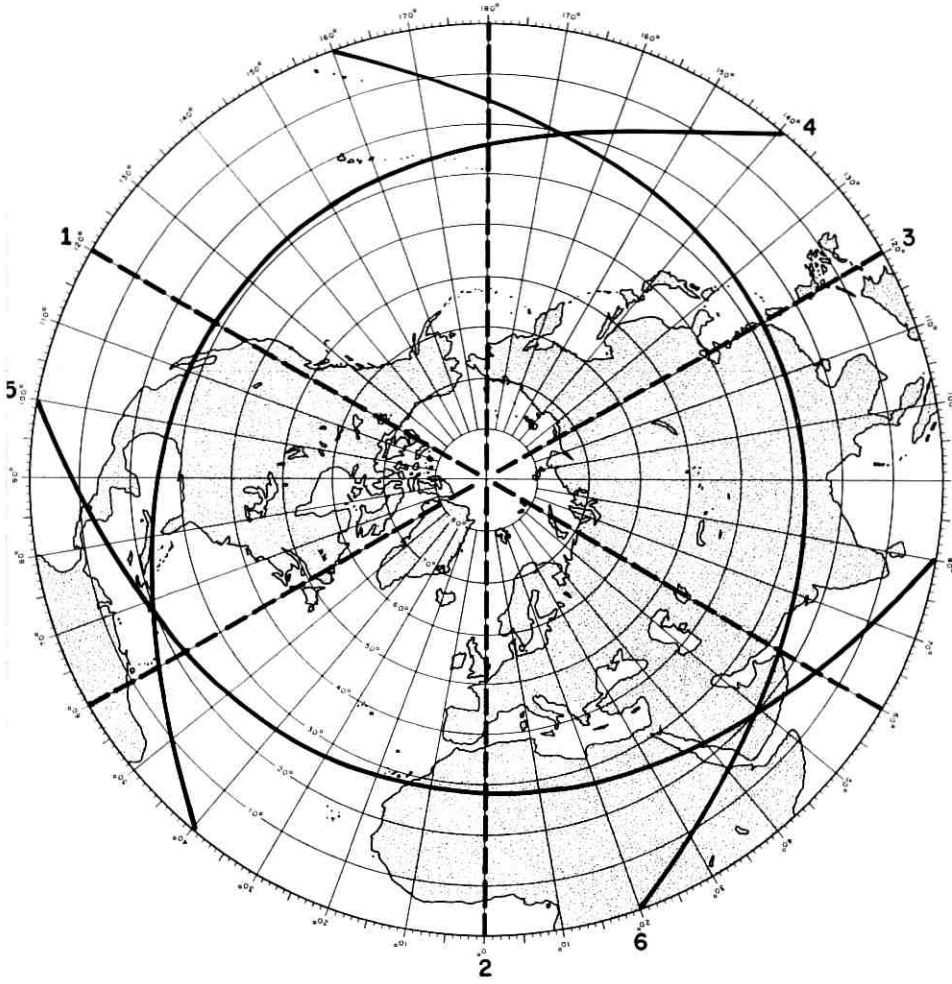
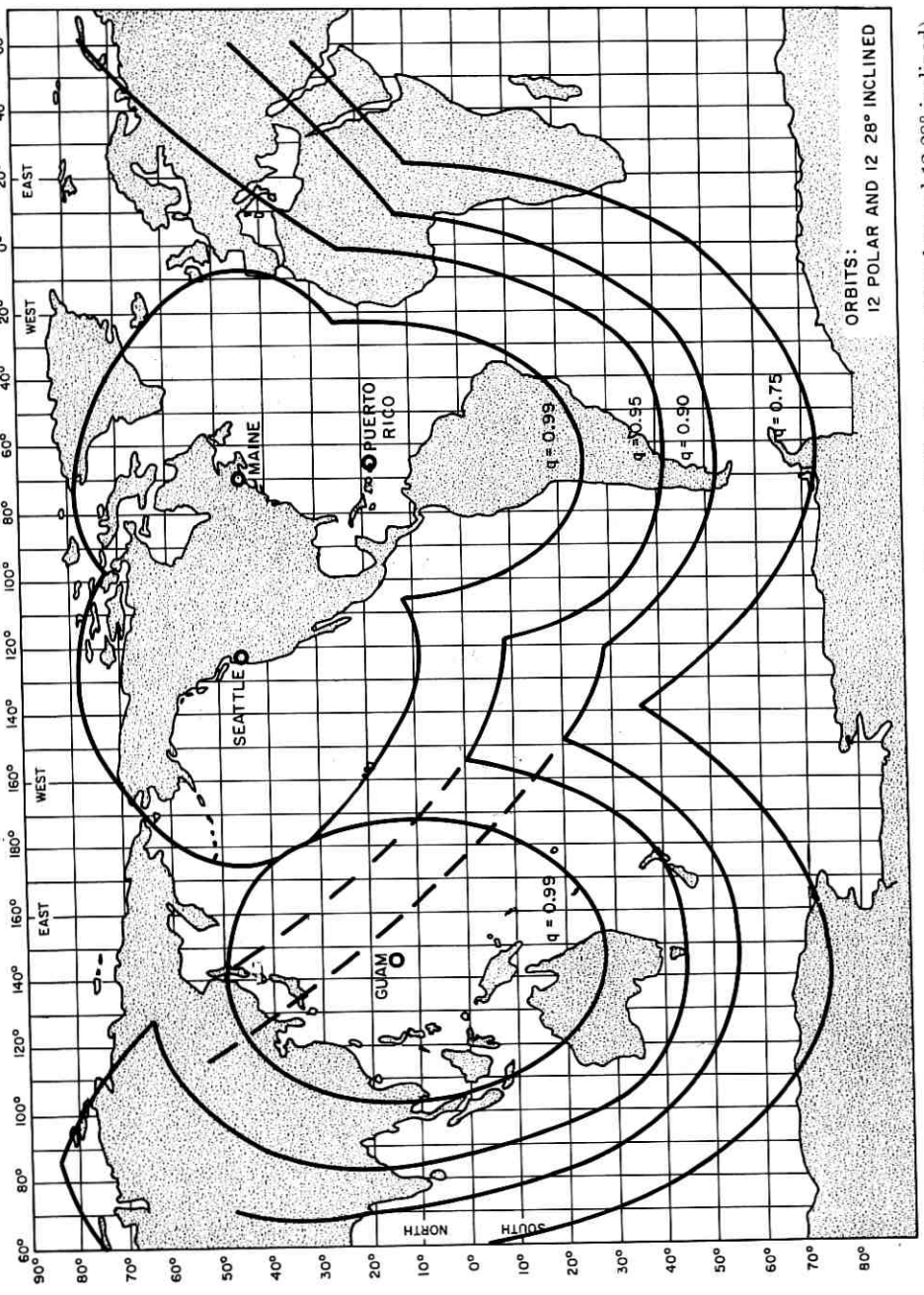


Fig. 15 — Orientation of the orbit planes of the 24-satellite system. The dotted lines are the intersections of the polar planes with the surface of the earth, and the solid lines are the intersections of the 28° planes with the earth's surface. The planes are numbered at the ascending nodes of the orbits.

to communicate with one of these stations more than 0.99 of the time. A point north of the 0.95 contour would have service with one of the stations more than 0.95 of the time, etc. These contours were determined by using (6) in Section 4.1.2 for values of p from the table in the Appendix. The location of the ground stations in the figure were chosen


ORBITS:
12 POLAR AND 12 28° INCLINED
q = 0.99
q = 0.95
q = 0.90
q = 0.75
SEATTLE
MAINE
PUERTO RICO
GUAM
NORTH
SOUTH
EAST
WEST

in an attempt to reach as much of the world as possible from points in the United States and its outlying territories.

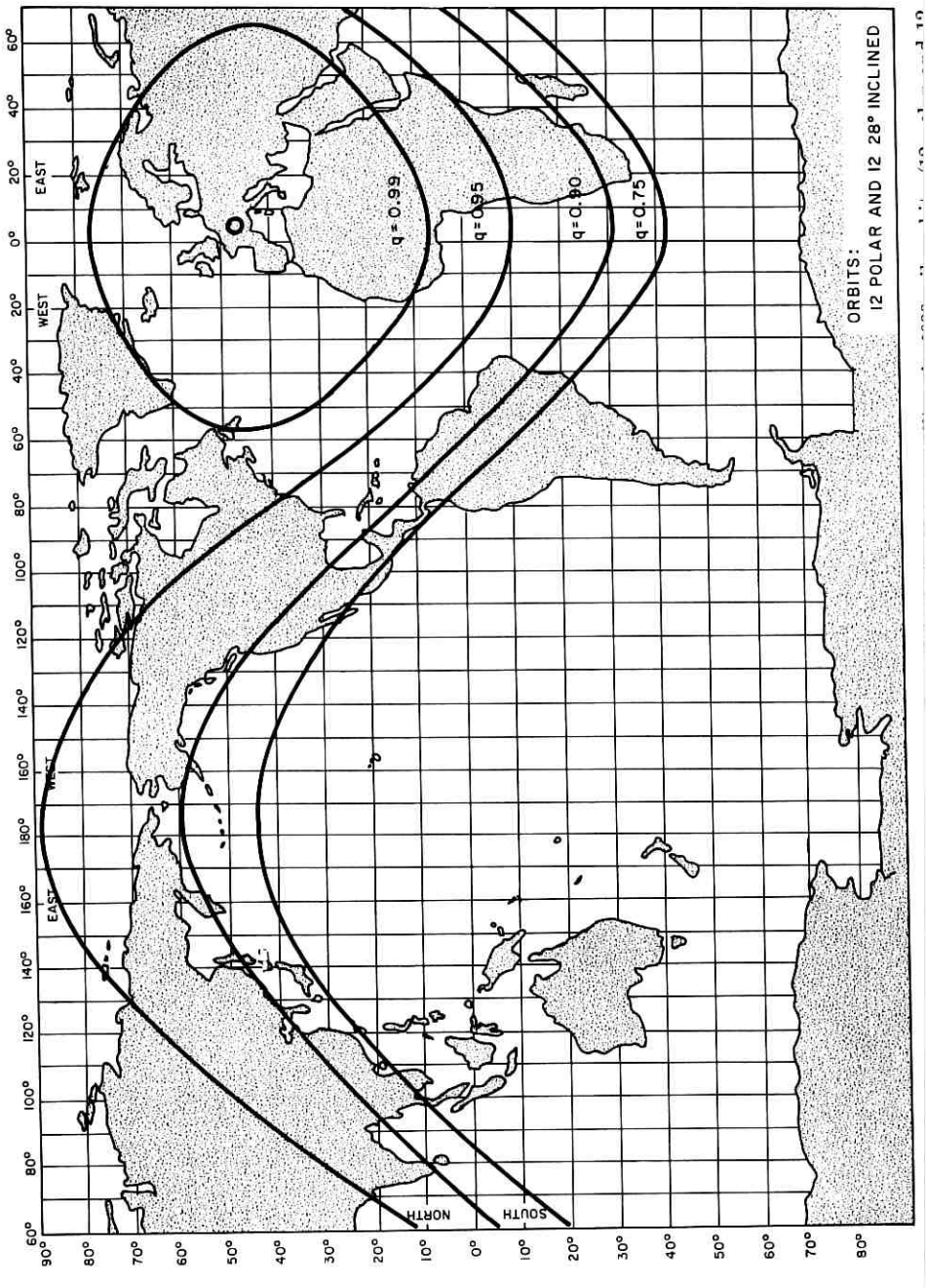
In order to be useful, the stations in Puerto Rico and Guam would have to be connected to circuits in the United States. For Puerto Rico, this connection could be made with Maine by way of a satellite or with Florida by submarine cable. The length of the path through a single satellite link, i.e., the distance from one ground station through the satellite to another ground station, can be as much as 17,000 miles for certain positions of a satellite at a 6000-mile altitude. Thus a circuit from South America to Maine consisting of two satellite links in tandem could be as long as 34,000 miles. It is well known that the performance of telephone circuits is adversely affected by the time delay associated with long path lengths, and ideally the path should be as short as possible. Thus for telephone communication with South America and Africa by way of Puerto Rico, the cable connection from Puerto Rico to Florida would probably be more desirable than a satellite connection with Maine; however, both alternatives may be desirable to provide diversity.

Guam could be connected to the United States by using a satellite link from Guam to Hawaii and another satellite link from Hawaii to Seattle. Circuits from Asia to the United States would then consist of three satellite links in tandem which could be as much as 50,000 miles long. A submarine cable which will connect Guam to the United States is presently planned for service in 1964. This cable would probably be a more attractive way of relaying telephone circuits to the United States than would the two satellite links, but again it may be desirable to have both means available.

The plan of Fig. 16 resulted from an attempt to reach as much of the world as possible through facilities under control of the United States. Most of the world could be reached with reasonably good qualities of service, but the service to Africa is probably not adequate. A plan which results in better service to this area will be given in Section 5.4.

5.3 *Service from Europe*

Fig. 17 shows contours of constant q_1 from an arbitrary point in western Europe. Much of the Pacific area cannot be reached from Europe by a single satellite link because the two regions are on opposite sides of the world. Africa could be served from Europe about as well as South America could be served from the U. S., and Europe–South American service would be limited to about the same extent as U. S.–African serv-


ORBITS:
12 POLAR AND 12 28° INCLINED
q = 0.99
q = 0.95
q = 0.90
q = 0.75
NORTH
SOUTH
EAST
WEST
WEST
EAST

ice. Service to all of these areas could, of course, be improved by establishing additional ground stations at appropriate points and using multiple satellite links in tandem.

Other contours could be drawn with respect to other parts of the world; however, the contours in Figs. 16 and 17 are probably adequate to describe world-wide service. These contours can be moved in longitude or they can be inverted and centered at the corresponding latitude in the southern hemisphere. The next section presents a plan in which consolidation of traffic could be used to provide more efficient communication between all parts of the world.

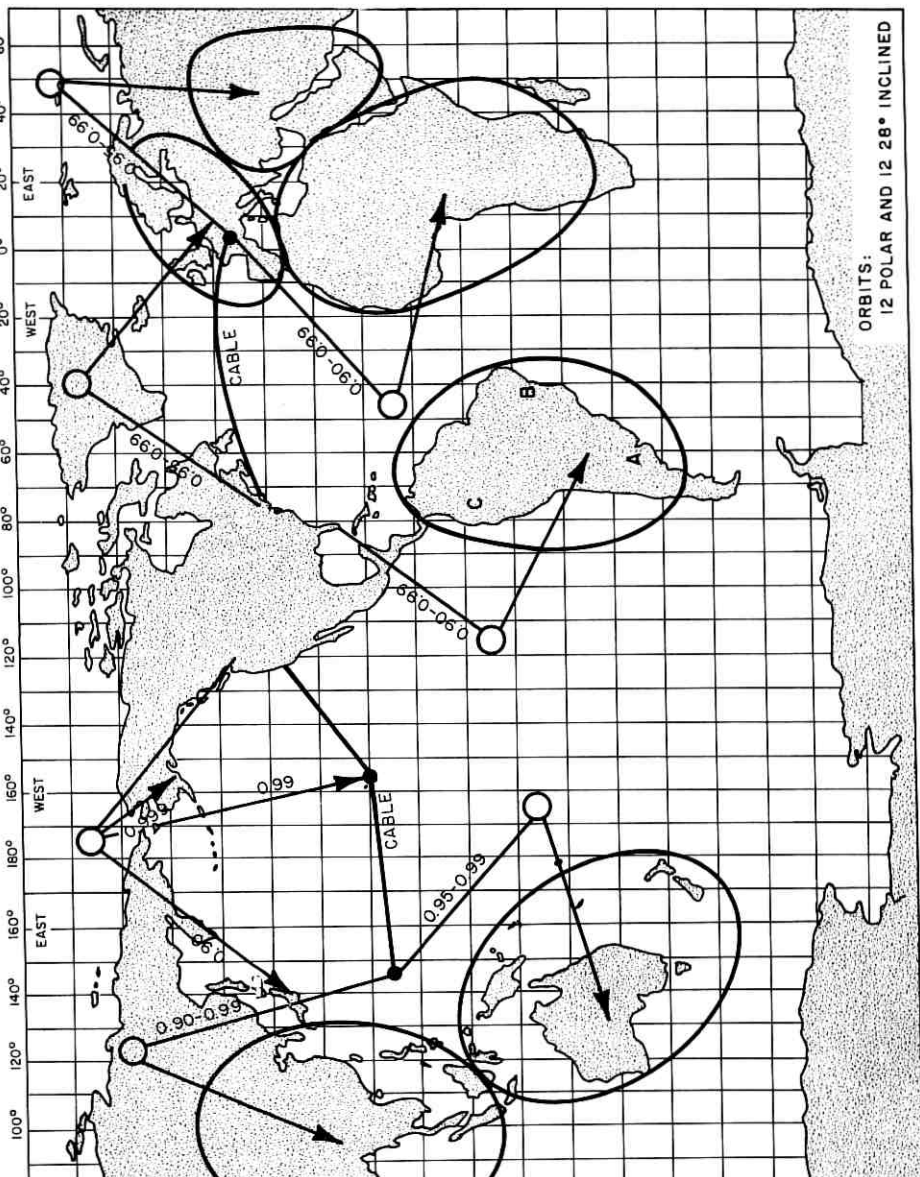
5.4 *A Possible Method of Operation for World-Wide Service*

North America has an extensive land-line network which would enable all points to communicate with overseas points through one of the assumed ground stations in Fig. 16. Europe also has an extensive network, but land lines between countries on other continents are not as well developed. Each of these countries would need a direct connection, by way of a satellite, with a network through which other countries of interest could be reached.

A possible scheme for world-wide communication is shown in Fig. 18. In this plan, countries within the encircled regions would communicate with stations in western Europe, Guam, Seattle, Maine, or Miami.* Since these stations would be interconnected by submarine cables and land lines, any of the encircled regions could be reached from any of these stations by using cable facilities and one satellite link, e.g., U. S.–African circuits could go by cable to Europe and then by satellite to Africa. Any two encircled regions could communicate by using two satellite links which would be connected by cable facilities, e.g., circuits from Asia to South America could go by satellite to Guam, then by cable and land lines to Miami, and then by satellite to South America.

The ranges in quality of service shown for each path in the figure assume that all countries within the encircled region could use the same satellite. The South American service will be discussed to point out some implications of this kind of operation. In Fig. 18, stations A, B, and C in South America would always use the satellite which is mutually visible to Miami and station A. If no satellite satisfied this condition, station A would be out of service, and a satellite which was

* This plan does not require direct communication by satellite between the U. S. and Africa, and the service to South America from a station in Puerto Rico would not be greatly different than the service from Miami would be. For this reason, Miami was chosen as the U. S. station for this part of the world.


ORBITS:
12 POLAR AND 12 28° INCLINED
A
B
C
D
0.90-0.99
0.95-0.99
0.99
CABLE
CABLE
NORTH
SOUTH
EAST
WEST
EAST
WEST

mutually visible to Miami and B would be used, etc. Thus stations close to Miami would have high qualities of service and those further from Miami would have lower qualities of service. As shown on the figure, the range of q from Miami to countries in South America is 0.90–0.99.

Communication between countries in South America could also be accomplished by way of the satellite which is used to reach Miami. The quality of service between two countries would be approximately the product of the qualities which the individual countries would have with Miami. The service between these countries could be improved if they also used mutually visible satellites which were not visible to Miami, but this would require each country to have another ground antenna.

According to Fig. 18, South America and Africa would communicate over a long network of land lines, cables, and two satellite links, but it is apparent that direct communication by way of one satellite link between these regions would be possible. However, this direct communication would require all South American countries and all African countries to have another ground antenna specifically for this purpose. This would not be true if each of these areas had a land-line network which interconnected all countries. Direct communication between Europe and South America could also be achieved if each South American country had still another ground antenna. The discussion of South American service can be extended to the other regions which are encircled in Fig. 18.

The method of operation illustrated in Fig. 18 minimizes the number of ground antennas needed by individual countries by consolidating traffic through central points which are interconnected by terrestrial facilities. With this method, routes which are frequently used could be served by single-satellite links, and other routes which are used less frequently could be served by two satellite links in tandem.

5.5 *Alternate Routing by Way of Two Satellite Links*

During the times when no satellites are mutually visible to a pair of stations, each station will often be able to see a satellite which is out of view of the other station. If there were a relay point which could see both of these satellites, a path with two satellite links could be established by way of the relay station. For example, Anchorage, Alaska might be a useful relay point for Maine and Europe. Figs. 16 and 17 give qualities of service of about 0.95 and 0.90 for Maine–Anchorage and Anchorage–Europe respectively. Another location which is well-situated for relays between Maine and Europe is the Cape Verde Islands off the

west coast of Africa. This relay point could also provide an alternate route for Miami–South America and Europe–Africa.

The 24-satellite system was simulated for a four-month period in order to obtain an appreciation for the usefulness of the alternate-routing concept. In this simulation the Maine–Europe path had 98 periods during which a satellite was not mutually visible. During 53 of these periods, satellites were mutually visible between Maine and Anchorage and between Anchorage and Europe, so an alternate route could have been established. A relay in the Cape Verde Islands would have been useful during 31 of the remaining 45 periods. Thus about 86 per cent of the service delays during this simulated period could have been avoided by using these two relay points. Of course some penalty would remain, during these periods, in the form of degraded communication performance, but it is possible that this degradation could be tolerated for the short and infrequent intervals which would be involved.

5.6 *System Growth and Service Improvement*

It is expected that over a period of years an increase in demand and a desire to improve qualities of service would require the expansion of an initial system. The orbits in which satellites are added to the system would depend on the global distribution of the increase in demand for service. It is likely that polar satellites would be added at a faster rate than would 28° satellites, and if this were the case, the shapes of the contour lines would change from those given in Figs. 16 and 17. The service improvement offered by increases in the number of satellites will be illustrated by a simple example in which the satellites in each orbit are increased by equal numbers.

As noted in Section 5.2, the contours of Figs. 16 and 17 represent qualities of service for the 24-satellite system determined by the equation

$$1 - q_1 = (1 - p_1)^{12} (1 - p_2)^{12}.$$

If the number of satellites in each of the two orbit inclinations is changed to n, it follows that

$$(1 - q_1') = (1 - q_1)^{n/12},$$

where q_1' is the new quality of service for a contour line. Therefore, these contours remain valid, with the appropriate changes in q_1, for different numbers of satellites in this type of system. The following table presents the values of q_1 for each of the contour lines for several n.

n =	*12*	*16*	*20*	*24*
Inner contour =	0.99	0.998	0.9995	0.9999
Second contour =	0.95	0.982	0.9934	0.9975
Third contour =	0.90	0.954	0.978	0.99
Outer contour =	0.75	0.842	0.901	0.937

An increase in q_1 which results from an expanded system may be considered solely as an improvement in service quality. However, this increase also affects the number of paths which can be made available between two points. For example, from the above table, for $n = 20$, $q_1 = 0.9995$ for the inner contour. Inspection of Fig. 9 shows that 2 paths could be established with ordered priority with $q_2 = 0.995$; similarly, for three paths, $q_3 = 0.978$ etc. From Fig. 10 similar information may be obtained for paths with equal priority. For the above example $\bar{q}_2 = 0.9973$, $\bar{q}_3 = 0.9915$, etc.

The communication scheme of Fig. 18 is also valid for an expanded system of an equal number of satellites in 28°-inclined and polar orbits. The qualities of service indicated would be changed by replacing them by their corresponding values in the above table.

VI. SOME NONRANDOM CHARACTERISTICS AND THEIR EFFECT ON SERVICE

The system described in the previous section would be nonrandom in two respects. In the first place, the four satellites in each plane are not controlled in phase, so these satellites could bunch together. If this should occur, the four satellites would be of little more use than a single satellite. Secondly, the orbit planes will rotate slowly about the earth's axis at different rates, so the three orbit planes in each inclination could coincide after a long period of time. Certain parts of the world could then see the satellites only during parts of the day. This section discusses these two nonrandom characteristics and presents simulation results which show their effect on service.

6.1 *Relative Satellite Velocities and Phasing of Successive Launches*

If several satellites were placed into the same orbit plane, the most uniform coverage would result if the satellites were uniformly spaced. However, even if uniform spacing could be achieved initially, it is not possible to give all satellites exactly the same velocity. Thus the satellites would drift from the desired relative positions, and to preserve the uniform spacing, a propulsion system would be needed on each of the satellites. This would probably represent a severe requirement on a system which must have an appreciable lifetime.

Fig. 19 illustrates one method of separating four satellites in such a

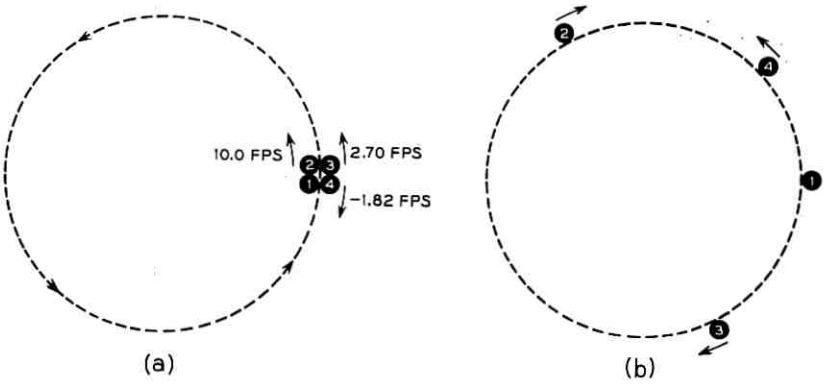


Fig. 19 — Geometry of satellite separation: (a) Incremental velocities relative to satellite #1. (b) Relative positions after 3 months in orbit.

way that the effect of future bunching of the satellites would be small. In Fig. 19(a) the satellites initially would be moving counter-clockwise with the same velocities (they would be placed into orbit as one package). Then the satellites would be separated by springs in order to give three of them the indicated velocities relative to satellite number 1. Figure 19(b) shows the relative directions of motion and the relative positions after three months in orbit. A satellite which is given an increase in velocity goes into a slightly higher elliptical orbit with a slightly longer period, and its angular velocity therefore decreases. This is the reason satellites 2 and 3 fall behind satellite 1 even though they have higher velocities. The reverse is true of satellite 4.

The relative velocities were determined by a computer simulation. Many combinations of velocities (all small enough to be within the capability of springs and large enough to effect early separation) were tried, and the relative positions of the satellites were followed for a period of ten years. The set shown in Fig. 19 is the one which resulted in the most uniform coverage over the study period; however, errors of ± 3 per cent could be tolerated in each of these velocities without significantly changing the coverage.

The three planes of satellites in each group would no doubt be launched a month or so apart because of preparation time and availability of launch facilities. There is also a coverage advantage to be gained from phasing the launchings. Some pairs of stations could see all polar orbits at all times, and other pairs could see all 28° orbits at all times. If the satellites in each of the planes were given the set of velocities shown in Fig. 19 at different times (different launch dates),

the satellites in one plane could be made to be well-spaced at times when the satellites in another plane were bunched.

The desired intervals between launchings were also determined by computer simulation. The first set of satellites was allowed to disperse for one month, and then many combinations of second and third launching times over a six-month period were tried. For each combination, the coverage was examined for a ten-year period. The best combination which was found specified launching the second plane 1.5 months after the first, and the third plane 4.5 months after the first. If either attempt were to fail, there would be other times several weeks later which would serve equally well. The two groups of planes would be handled separately from different launching points, and no phasing between groups would be attempted.

The above procedures are not necessarily optimum, but they do represent a way of achieving orbits which do not result in severe bunching of the satellites within the planes.

6.2 *Motion of the Orbit Planes*

The plane of the orbit of an earth satellite will not in general maintain a fixed orientation in inertial space. Instead, there will be a slow rotation of the plane about the earth's axis. This rotation is caused by the nonspherical shape of the earth, and for circular orbits the rate of rotation, in radians per orbit, is given by[7]

$$\Delta\Omega = \frac{2\pi J R^2 \cos i}{(R + H)^2} \tag{4}$$

where

R = radius of the earth

H = altitude of the satellite above the earth

J = 1.625×10^{-3}, the first oblateness coefficient[8]

i = inclination of the orbit plane.

The direction of the rotation of the plane is opposite to that of the earth's rotation for inclinations less than 90°, and is in the same direction as the earth's rotation for inclinations greater than 90°. The rotation vanishes for true polar orbits ($\cos i = 0$), and a 6000-mile, 28° orbit rotates at a rate of about 100° per year.

In the system shown in Fig. 15, the three orbit planes in each group would have no relative motion if each plane has the same inclination.

Small errors in inclination can be expected, however, and the resultant relative motions could cause the planes to bunch together after a number of years. The magnitudes and directions of the inclination errors which are made will determine the extent of the bunching and the time of its occurrence.

If it is assumed that the accuracy of the inclination angle will be within 5°, which is probably a pessimistic assumption, the most severe bunching of the near-polar planes would result for inclinations of 85°, 90°, and 95° respectively for planes 1, 2, and 3 in Fig. 15. With these inclinations, plane 2 would not rotate, plane 1 would rotate clockwise at a rate of about 10° per year, and plane 3 would rotate counter-clockwise, also at a rate of 10° per year. Thus the three planes would be nearly coincident* after about 6 years.

The worst combination of inclination angles for planes 4, 5, and 6 would be 33°, 28°, and 23° respectively. All of these orbits would rotate in a clockwise direction, but plane 4 would rotate about 5° per year less than plane 5, and plane 6 would rotate about 5° per year more than plane 5. The planes would be closest to coincidence when their ascending nodes were at the same position. Since the ascending nodes would be 120° apart initially, nearest-coincidence of the three planes would occur after 24 years.

The bunched system which will be discussed here is shown in Fig. 20. The orbits have the inclinations discussed in the preceding paragraphs, and the figure represents the orientation of the planes six years after the satellites are placed into orbit. Note that the assumed inclinations represent a rather pessimistic example since 5° errors would need to be made in exactly the right combination. The maximum error may also be pessimistic, and a maximum error of one-half this value would delay this bunching until twelve years after the system was started.

Even if the original system bunched together in this way after six years, it is quite likely the effect would not be severe if satellite failures and system growth are considered. The logical way to replace satellites which fail would be to establish a new plane with four satellites in it rather than try to replace individuals in the original planes.† The new plane could be placed in a gap which was caused by bunching. It is quite likely that more than 24 satellites would be needed by this time to improve the quality of service and to establish multiple paths between some stations. These additional satellites would provide another way to fill gaps caused by bunching of planes. If one considers such

* The planes would not be exactly coincident because they would have different inclinations.

† Assuming one rocket will place four satellites into orbit.

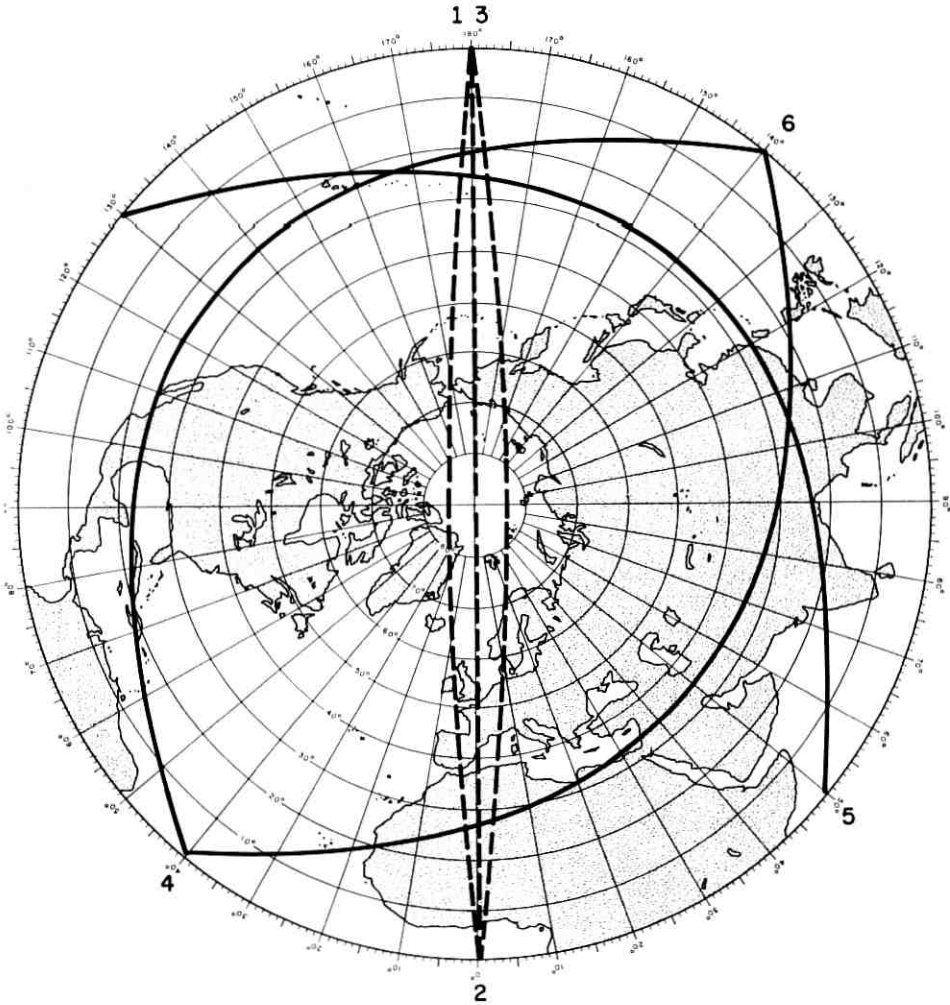


Fig. 20 — Orientation of the orbit planes of the 24-satellite system for a severe case of plane bunching. This bunching would occur after six years if a certain combination of 5° errors were made in the inclinations of four of the planes.

things as satellites failing at random, new satellites being added with random errors in inclination, and motions of orbit planes in different directions and at different rates, it is not difficult to postulate the evolution of a random system after a number of years.

6.3 *A Study of the Effect of Plane Bunching*

This section presents simulation results for three different systems of satellites. The first of these is the uniformly spaced system represented

by Fig. 15, the second is the bunched system of Fig. 20, and the third is this same bunched system with four satellites added in another 28° plane (the ascending node of the plane is 180° from that of plane 5). The services each of these systems would furnish to Maine–London, Seattle–Hawaii, and London–Johannesburg were simulated for six-month periods.

In the uniform system, the phases of the satellites were specified by using the relative velocities and phasings of successive launches discussed previously. The bunched system was described by advancing these phases through six years and arranging the ascending nodes and inclinations to correspond to those of Fig. 20.

6.3.1 *Maine–London Service*

Fig. 21 shows ten-day portions of the simulated services of the three systems for the Maine–London path. Fig. 21(a) is a representative ten-day period for the uniform system. The statistics q_1, t_o, and t_i (see Section IV) are given for the ten days and for the six months to provide a means of comparing the services quantitatively. It can be seen from these statistics that the ten-day period shown for the uniform system has longer periods of no service and longer times between these periods than the averages for the six months.

Fig. 21(b) shows the poorest service which resulted for the bunched system. The poorest ten days were chosen so that the effect of bunching of satellites within the planes would also be included. Although this group of days cannot be compared directly with the first group, the six-month statistics can be compared. For the six-month periods, q_1 is 0.984 for both systems, and the periods of no service are longer but less frequent for the bunched system than for the uniform system. The ten-day statistics for the two systems differ by similar amounts from the six-month statistics of the uniform system. From these comparisons, it is difficult to say that one system is better than the other for this path.

It is not surprising that this path would be only slightly affected by the bunching because most of the service is furnished by the near-polar satellites and Maine and London always have mutual visibility of these three orbits as shown in Fig. 5. Because the earth rotates under the orbits, there would be two times (12 hours apart) each day when the portions of these orbits which could be seen by Maine and London would be a minimum. The periods of no service would most likely occur during these times, and a 12-hour pattern in these periods would result. This pattern is evident in Fig. 21(b).

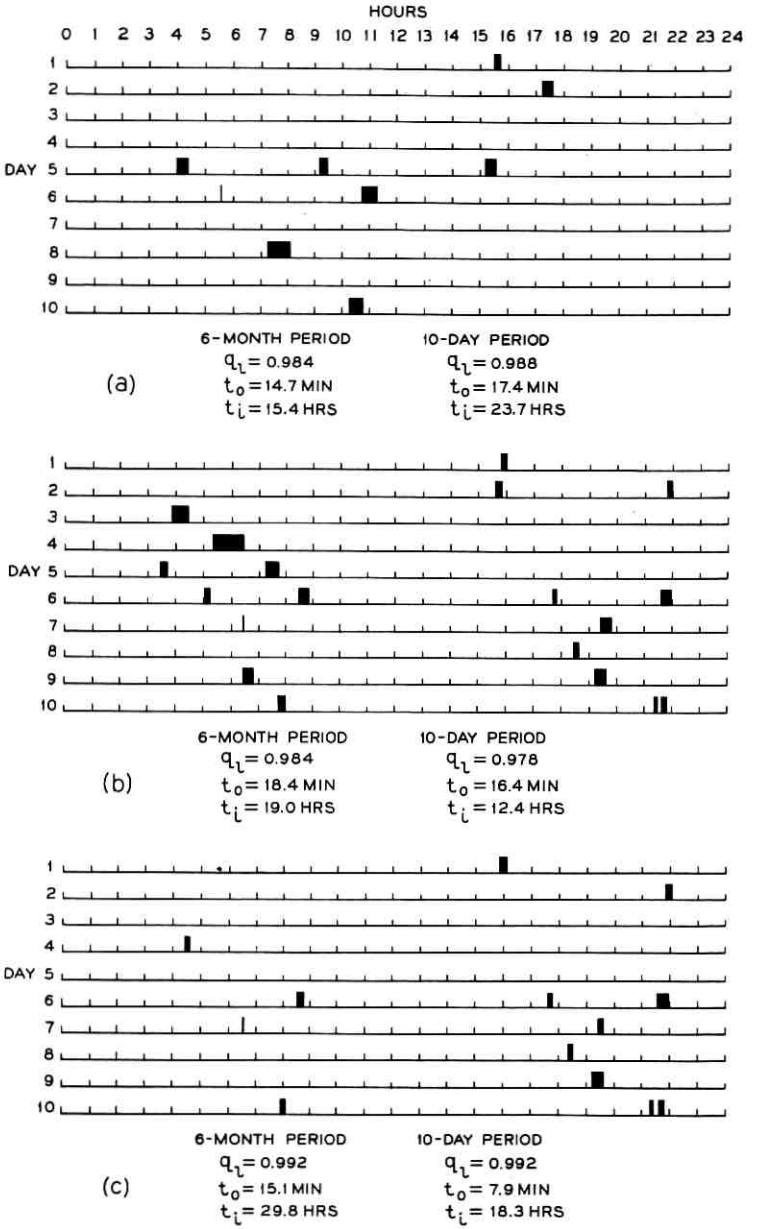


Fig. 21 — Maine–London service: (a) Representative service furnished by the uniformly spaced system shown in Fig. 15. (b) Service during the poorest 10-day period from the 6-month simulation of the bunched system shown in Fig. 20. (c) Service during the period shown in Fig. 21(b) if four satellites in a 28° plane were added to the bunched system.

It will be shown later that the other two paths in the study would be more seriously affected by the bunching, and the third system with 28 satellites will be studied as a means of relieving the effect of the bunching. This system was also simulated for the Maine–London path for completeness, and the results are given in Fig. 21(c). The same ten-day period is shown for Figs. 21(b) and 21(c), so a direct comparison can be made. It is clear from this comparison that the four additional satellites more than compensate for the possible, small effect of bunching on this path.

6.3.2 *Seattle–Hawaii Service*

Fig. 22 gives similar ten-day displays for the Seattle–Hawaii path. Comparison of the six-month statistics for the first two systems shows the service to be affected considerably by the plane bunching. In Figure 22(b), the periods of no service tend to occur during the same part of each day, and this again is caused by the rotation of the earth under the orbits. Consideration of Fig. 20 and the visibility region of Fig. 5 shows there would be a part of each day for which only orbit plane 6 would be over the visibility region and it is during this time that the no-service periods would most likely occur. Twelve hours later orbits 4 and 5 are over the visibility region but orbits 1, 2, and 3 are not; since the service at this time is still rather good, it is apparent that the effect of the bunching of the polar orbits is not large.

These characteristics suggest the addition of another 28° plane with an ascending node 180° from that of plane 5, and the service which would result from this system of 28 satellites is shown in Fig. 22(c). In this figure the no-service periods are considerably reduced from those of Fig. 22(b), and they occur during the two parts of the day when the polar orbits are not visible. Comparison of the six-month statistics for the 28-satellite system with those of the uniform 24-satellite system suggests that both systems would provide about the same service to the Seattle–Hawaii path. It might be said, then, that the four added satellites represent the penalty of the bunching for this path.

6.3.3 *London–Johannesburg Service*

The service to the London–Johannesburg path is represented in Fig. 23. Comparison of the first two groups shows that a severe pattern in the periods of no service is caused by the bunching of the planes, and comparison of the six-month statistics shows a definite degradation in the service to this path. The nature of the reduction in service is the

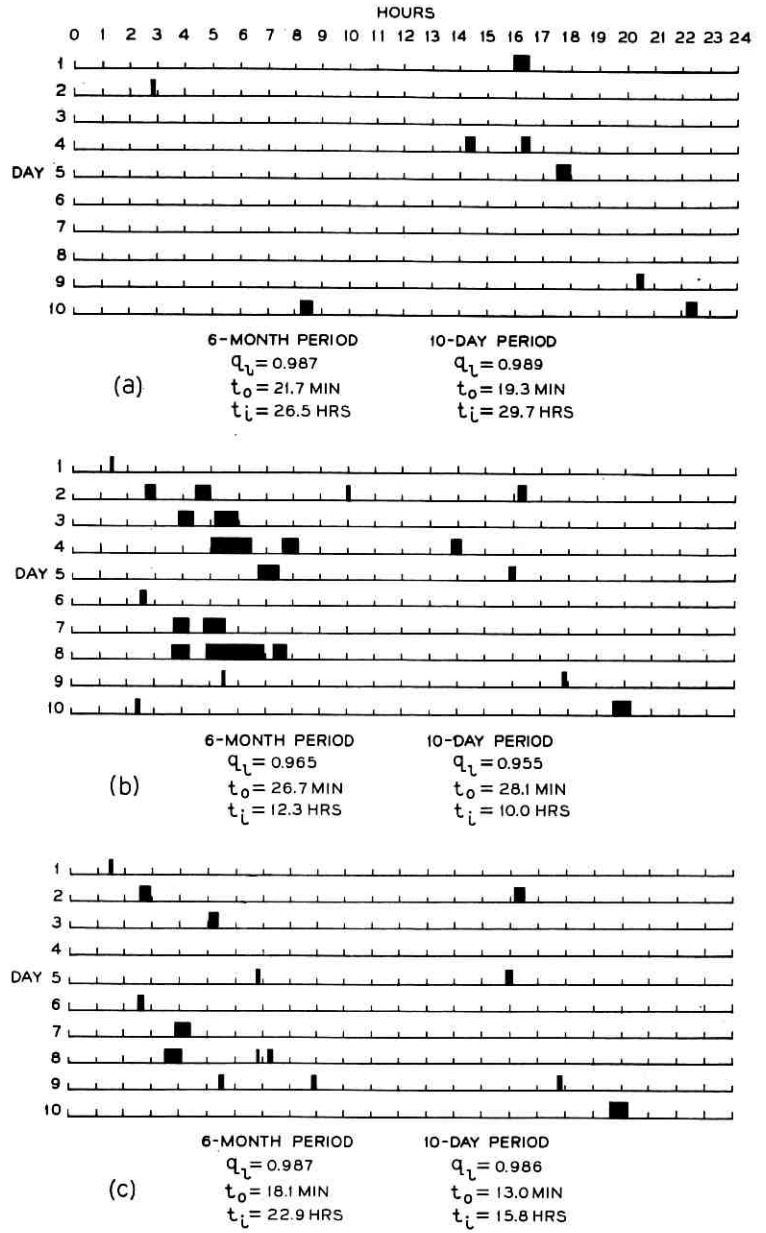


Fig. 22 — Seattle–Hawaii service: (a) Representative service furnished by the uniformly spaced system shown in Fig. 15. (b) Service during the poorest 10-day period from the 6-month simulation of the bunched system shown in Fig. 20. (c) Service during the period shown in Fig. 22(b) if four satellites in a 28° plane were added to the bunched system.

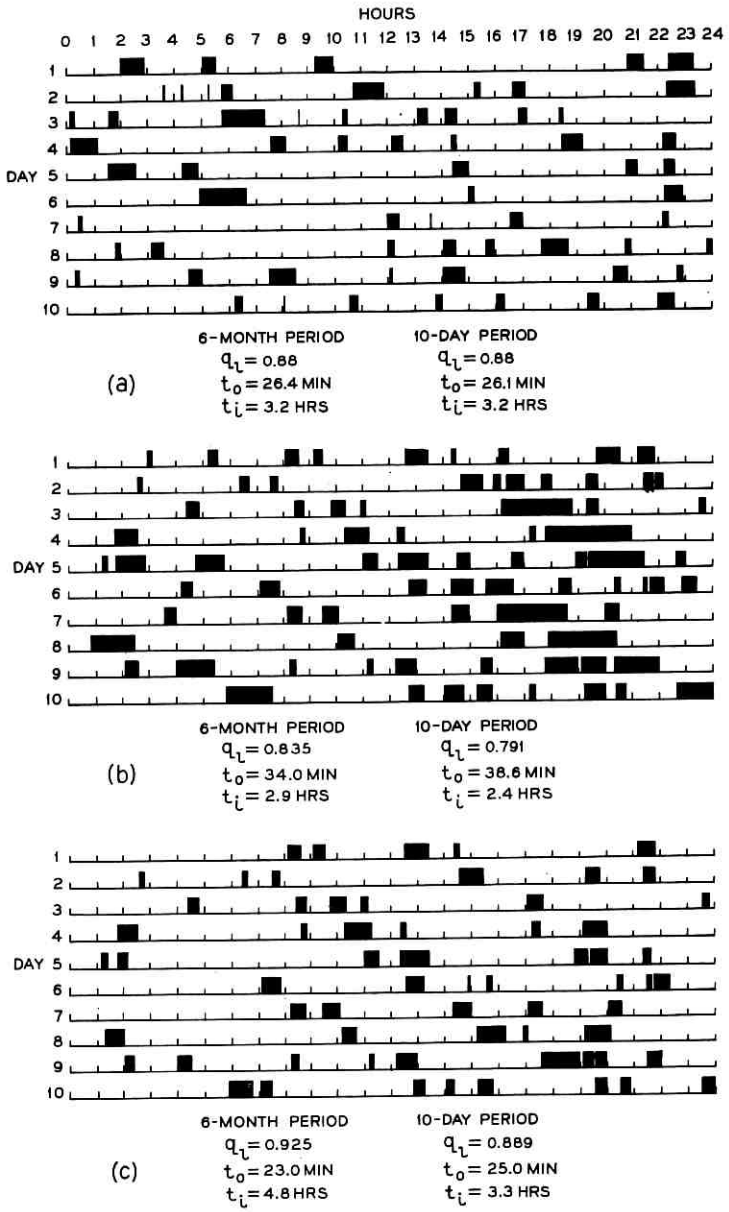


Fig. 23 — London–Johannesburg service: (a) Representative service furnished by the uniformly spaced system shown in Fig. 15. (b) Service during the poorest 10-day period from the 6-month simulation of the bunched system shown in Fig. 20. (c) Service during the period shown in Fig. 23(b) if four satellites in a 28° plane were added to the bunched system.

same as it was for the Seattle–Hawaii path, and Fig. 23(c) shows that the four additional satellites would be more than adequate to overcome the effect. This statement is best supported by comparing the six-month statistics for the three systems.

6.3.4 *Summary*

The service furnished to Maine–London by the bunched system of Fig. 20 would be essentially the same as the service furnished by the uniform system. The services furnished to the other two paths would be degraded by the assumed bunching of planes, but this degradation could be compensated by adding four satellites in a 28° orbit. Two of the paths would have better service from this bunched system of 28 satellites than they would have from the uniform system of 24 satellites. Thus the penalty of the bunching is less than four satellites.

The assumed bunching would result after six years if large (5°) errors were made in the inclinations in exactly the right combination. A reduction of the error by a factor of two would double the time required for bunching. Also, if errors in the range of 0°–5° were made on a random basis, the bunching after six years would not be as severe as that which was assumed.

Satellite failures and system growth were not included in the study, but they represent important factors in the control of bunching. If four satellites were to fail from each group of twelve during the six years, their replacements could easily fill in the gaps caused by the bunching. These gaps could also be filled by satellites which might be added to increase the qualities of service from their initial values. Because the bunching is slow, the constant redistribution resulting from replacements and additions would be expected to make the bunching of planes of minor significance, and through these processes a random system would no doubt evolve.

VII. CONCLUSION

The service which would be provided by satellites in uncontrolled orbits has been described analytically and by simulation. Section IV provides equations and graphical and tabular material which can be used to describe the service of systems which are random, and the same material can be used to estimate the service which would be provided by systems which have certain nonrandom characteristics. Results of computer simulations of several systems have been presented to develop an intuitive feeling for various qualities of service.

The example system which was studied shows that rather good world-wide service could be started with a modest number of satellites. The usefulness of existing facilities in extending this service and in providing for efficient consolidation of traffic has also been illustrated. A method of improving the continuity of service by using alternate relay points has been included.

The nonrandom characteristics of the example system were investigated by computer simulation to determine their effect on service. A method of phasing multiply-launched satellites was presented, and the effect of bunching of orbit planes was studied. A severe case of bunching was shown to degrade the service to some representative pairs of stations, but this degradation was shown to be more than compensated by a small increase in the number of satellites. Since the bunching of planes in ordered systems would occur at a slow rate, it is felt that replacements of satellites which fail and the additions of satellites for growth would provide an adequate means of controlling the effect of this bunching. It is also felt that these replacements and the additions would result in the eventual evolution of a random system in which significant bunching would be less probable.

VIII. ACKNOWLEDGMENTS

The authors have benefited greatly from the many helpful suggestions and ideas offered by Messrs. D. F. Hoth and J. L. Glaser. We are also grateful to Mr. O. Benediktsson for his contributions to the study of the phasing of multiply-launched satellites.

APPENDIX

Table II presents values of $p \times 10^3$ for numerous ground station locations, orbit altitudes and inclinations. The first three columns give the co-latitudes of the ground stations and their longitude separation in degrees. Because of symmetry with respect to the equator, two co-latitudes may be measured from either the south pole or the north pole. Some ground stations of interest are named next to their corresponding locations. Where one station is paired with several others, the first name is not repeated and the second is indented.

The other columns are headed by H, the circular orbit altitude and i, the inclination. Linear interpolation may be used between successive altitudes and inclinations with reasonable accuracy.

TABLE II

	Colat 1	Colat 2	Longitude Difference	H = 2500 N. Mi					H = 4000 N. Mi						H = 6000 N. Mi						H = 8000 N. Mi					
				i = 0°	35°	50°	70°	90°	i = 0°	28°	35°	50°	70°	90°	i = 0°	28°	35°	50°	70°	90°	i = 0°	28°	35°	50°	70°	90°
	10	45	0													088				252						
	20	45	100													037				221						
	20	45	160													000				199						
Anchorage, Alaska–																										
Seattle	28	42	28		075	110	149	184			115	152	198	239			150	186	242	266			177	217	283	304
Tokyo	28	55	70		033	059	088	075			063	093	130	142			090	123	164	190			114	149	192	230
	30	45	130													014				189						
	30	45	140													007				184						
	30	45	180													000				176						
	34	45	108											154						196						222
Lands End, England–																										
Maine	39	45	65	000	047	069	094	124	000		078	102	137	187	030	060	107	133	185	230	080		127	154	220	256
New Delhi	39	62	83		026	040	051	040			052	069	090	082			076	096	123	132			095	115	152	180
Leopoldville	39	94	20							123						161						187				
Johannesburg	39	115	30		016	011	009	009		077	053	036	029	027	81	122	094	062	050	046	188	151	132	085	066	062
	40	76	0																	141						
Seattle–																										
	42	48	76	000				086	000					160						200	085					231
Tokyo	42	55	98	000	013	027	041	038	000		036	053	076	102	000		058	075	111	150	033		074	094	143	188
Hawaii	42	70	33	000	089	100	081	070	122		127	137	134	110	175		158	166	171	148	219		184	193	196	179
	42	71	71																	086						
	42	76	93																	060						
Maine–																										
	45	45	47						060					196	116					242	152					270
Rome	45	48	83							045				142		080		101	134	165		098				170
Madrid	45	49	65							077				143		108				153		122				182
	45	55	100													045				150						

TABLE II—*Continued*

	Colat 1	Colat 2	Longitude Difference	H = 2500 N. Mi					H = 4000 N. Mi						H = 6000 N. Mi						H = 8000 N. Mi					
				i = 0°	35°	50°	70°	90°	i = 0°	28°	35°	50°	70°	90°	i = 0°	28°	35°	50°	70°	90°	i = 0°	28°	35°	50°	70°	90°
Dakar	45	76	52													128				112						
	45	80	100													041				046						
	45	90	0													191					107					
	45	90	0						200					080	263					111	305					131
Rio de Janeiro	45	112	27	077	033	023	018	017	138		077	052	041	038	188		119	082	063	057	225		159	105	082	077
	45	120	0													138				054						
	48	51	4	136				155	218					237	263					281	302					317
	48	58	21	113				112	184					171	228					224	282					295
	51	56	40	068				093	136					154	178					207	233					272
Tokyo, Japan–																										
Guam	55	76	5							193						235				165		277				
	56	58	23	137				109	197					166	238					223	286					291
	56	69	41	105				068	160					109	239					153	227					180
	58	80	27	158				069	209					105	242					138	266					164
Cairo, Egypt–																										
Puerto Rico	60	71	98							033						067				052		084				
New Delhi, India–																										
Sydney, Australia	62	123	74		001	001	001		068		023	017	014	013	100		060	040	032	029	126		091	059	046	043
Miami, Florida–																										
	64	91	2							219						232				140		298				
Rio de Janeiro	64	112	37	135	055	039	031	029	181		108	073	058	054	214		153	105	084	078	241		198	134	103	097
Santiago, Chile	64	123	9							133						191				077		234				
	68	76	57							118						155				103		178				
Hawaii–																										
San Francisco	70	53	32	129	104	107	089	076	181		146	146	143	118	216		182	178	178	164	240		209	203	206	189
Manila	70	75	84															077	068							
Sydney	70	123	53	081	028	020	016	016	128		077	048	039	036	166		121	079	061	056	200		154	105	080	074
Puerto Rico–																										
Dakar	71	76	48							140						180				116		202				
	71	89	63																	077						

Leopoldville	71	94	81							070						102				045		120					
	71	100	56																	078							
	71	120	36																	079							
	71	130	0																	079							
	71	150	0																	050							
Guam–																											
	76	25	0																	119							
Bangkok	76	76	44							150						195				121		214					
	76	76	50																	109							
	76	100	55																	082							
Sidney	76	123	7							169						223				099		259					
	76	130	0																	088							
	90	90	33						205					098	233					129	256						154

REFERENCES

1. Pierce, J. R., and Kompfner, R., Transoceanic Communication by means of Satellites, Proc. I.R.E., **47,** March, 1959, pp. 372–380.
2. Curtis, H. E., Interference between Satellite Communication Systems and Common Carrier Surface Systems, B.S.T.J., **41,** May, 1962, pp. 921–943.
3. Bennett, F. V., et al., Determination of the Required Number of Randomly Spaced Communication Satellites, NASA, TN-0619, January, 1961.
4. Bennett, F. V., Further Developments on the Required Number of Randomly Spaced Communication and Navigation Satellites, NASA TN D-1020, February, 1962.
5. Rice, S. O., Intervals between Periods of No Service in Certain Satellite Communication Systems — Analogy with a Traffic System, B.S.T.J., this issue, p. 1671.
6. Rosen, M. W., and Johnson, V. L., Launch Vehicles for Communication Satellites, NASA, August 10, 1961.
7. Ewart, D. G., On the Motion of a Particle about an Oblate Spheroid, Journal of British Interplanetary Society, **17,** 1959–1960, pp. 162–168.
8. Smith, D. E., Determination of the Earth's Gravitational Potential from Satellite Orbits, Planetary Space Science, **8,** October, 1961, pp. 43–48.

Intervals between Periods of No Service in Certain Satellite Communication Systems — Analogy with a Traffic System

By S. O. RICE

(Manuscript received May 11, 1962)

In satellite systems in which the relative positions of satellites are allowed to vary, there will be periods during which no service will be provided between given ground stations. Such periods are called "outages," and the intervals between successive outages are called "innages." Here the outage and innage time distributions are studied with the help of an analogy between a satellite system and a traffic system. The arrival of a customer in the traffic system corresponds to a satellite coming into view, and the service time of the customer corresponds to the time the satellite remains in view. In particular, the methods of analysis developed for traffic systems are applied to determine an approximation for the distribution of innage lengths.

I. INTRODUCTION

Several types of communication systems have been proposed which would use repeaters orbiting the earth as artificial satellites. The problem considered in this paper arises in systems employing a number of satellites at altitudes of several thousand miles. Typically, the orbit altitude might be of the order of 5000 miles with a period of revolution of about 5 hours.

The companion paper by Rinehart and Robbins[1] discusses the conditions under which a particular satellite will be visible to a given pair of ground stations. For the orbit altitudes considered here, the satellite will be visible intermittently. Conceivably the relative positions of the satellites could be maintained so that at least one satellite is mutually visible from the two ground stations at all times. However, at least for some of the early systems proposed, it is of interest to consider the case

in which small differences in orbital period cause the relative positions of the satellites to vary with time. Attention is therefore directed to the statistical characteristics of satellite visibility.

We are especially interested in those periods during which no satellite is available for communication between a given pair of ground stations. For convenience these events are called *outages* although, as pointed out by Rinehart and Robbins, these occurrences need not imply an interruption of calls in progress. By analogy the intervals between outages are called *innages*.

For any particular system the most effective method of obtaining outage information appears to be that of simulation. The course of the system for a year or more is computed, with the help of a high-speed digital computer, and the outages and innages recorded. [A method for doing this and some sample results are presented in the paper by Rinehart and Robbins].

The present paper is concerned with determining the distribution of outage and innage lengths. The theory of outage length distribution has been studied by a number of people and some of their results, namely those which are needed here, are summarized in Section II. The traffic model is described in Section III. Sections IV and V contain results predicted by the model. In Section VI the predicted results are compared with those obtained by simulation. The work of Appendix B gives the basis for a model which is simpler but less accurate than the one described in Section III. The material in both appendices also appears to be of interest from the standpoint of traffic theory.

The general conclusion is that, for the cases examined, both the outage and innage distributions are approximately exponential. Their averages are related by the rather simple equation (1).

II. PRELIMINARY RESULTS

First note that the average innage length $\bar{t}_i$ and the average outage length $\bar{t}_o$ ("i" for innage and "o" for outage) are connected by the relation

$$\bar{t}_i = \frac{q}{1 - q}\,\bar{t}_o \tag{1}$$

where q is the quality of service; i.e., the fraction of time transmission is possible. This relation is always true and follows almost immediately from the definition of q. Typical values are $q = 0.99$ and $\bar{t}_o = 0.25$ hours; and it follows that the corresponding average innage length is $\bar{t}_i = 24.75$ hours.

Next, let k be the number of satellites in the system. Consider a particular satellite. Every now and then it will pass over the region of mutual visibility. Let

T = average length of time between its reappearances. For polar orbits, T is the orbit time if the region of visibility includes one of the poles.

b^{-1} = average length of time the satellite is visible during one pass. The quantity b is a rate which occurs frequently in the analysis.

p = fraction of time the satellite is visible. From these definitions it follows that

$$b^{-1} = pT. \tag{2}$$

The satellite systems considered here are restricted to those for which T and b^{-1} are almost the same for all k satellites.

With this notation we have

$$q = 1 - (1 - p)^k, \tag{3}$$

$$\bar{t}_o = (1 - p)\, T/k, \tag{4}$$

$$P(t) = \left[1 - \frac{t}{(1 - p)T}\right]^{k-1}, \qquad 0 \leqq t \leqq (1 - p)T \tag{5}$$

where $P(t)$ is the probability that the length of an outage will exceed t. Also the expected number of satellites visible at a given time is kp. These formulas are based upon the assumption that the phase angles of the satellites are independent random variables and are distributed uniformly over the interval $(0,2\pi)$. Equation (3) is due to J. R. Pierce and R. Kompfner.[2] Equations (4) and (5) have been given by R. E. Mosher and R. I. Wilkinson, respectively, in unpublished memoranda. For values of t and k of interest, (5) may be approximated by

$$P(t) = \exp\,(-t/\bar{t}_o). \tag{6}$$

It will be observed that the constants of the satellite system enter the right-hand sides of (3) to (5) only through the three parameters k,p,T. Thus, as far as these formulas are concerned, the satellite system is specified by k,p,T.

III. TRAFFIC SYSTEM MODEL

The satellite system will be represented by a traffic system model which consists of k independent servers, each having an average service

time b^{-1} and a service time distribution function $B(t)$. Customers are supplied to the servers by k independent Poisson sources, each producing customers at an average rate α. When a source produces a customer, that particular source is removed from the system while the customer is being served. Thus, when n servers are busy serving n customers, the average arrival rate is $(k - n)\ \alpha$. This type of input is a special case of a more general type (the limited source or "Engset input") which has been studied in traffic theory.

The instant a satellite becomes visible from both the receiver and transmitter corresponds to the arrival of a customer. The length of time a satellite remains visible corresponds to the time required to serve a customer. After a customer has been served he leaves the system. This corresponds to the satellite leaving the region of mutual visibility. The state in which n satellites are visible simultaneously (state n) corresponds to the state in which n servers are busy serving n customers. An outage corresponds to an idle period (state 0), i.e., a period during which all servers are idle. An innage corresponds to a busy period, i.e., to a period when one or more servers are busy.

Note that the constant orbit time of the satellite may introduce a regularity in the satellite arrivals. The traffic model has the shortcoming that there is no corresponding regularity in the customer arrivals.

The analogy between the satellite system and the model is established by taking k and b^{-1} to have the same values in both and setting

$$\alpha = 1/(1 - p)T. \tag{7}$$

To justify this choice for α, note that if a particular satellite is not visible at a time t selected at random, the chance that it will become visible in $t, t + dt$ is $dt/(T - pT)$. Comparison with the corresponding probability $\alpha\ dt$ for the traffic system gives (7). When the three satellite system parameters k,p,T are known, the three model parameters k, b^{-1}, α follow at once from (7) and $b^{-1} = pT$.

IV. OUTAGE AND QUALITY OF SERVICE PREDICTED BY MODEL

The values of the quality of service q and average outage length $\bar{l}_o$ predicted by the model agree exactly with (3) and (4) while the predicted outage length distribution is the exponential approximation (6) to Wilkinson's polynomial expression. A sketch of the proof of these statements is given in the following paragraphs.

First assume the service distribution $B(t)$ to be exponential, i.e., equal to $1 - e^{-bt}$. Then since the sources are Poisson, the behavior of the system is governed by the $k + 1$ state equations (Ref. 3, p. 30)

$$\begin{aligned}
P_0' &= -k\alpha P_0 + bP_1 \\
P_1' &= k\alpha P_0 - [(k-1)\alpha + b]P_1 + 2bP_2 \\
P_2' &= (k-1)\alpha P_1 - [(k-2)\alpha + 2b]P_2 + 3bP_3 \\
P_3' &= (k-2)\alpha P_2 - [(k-3)\alpha + 3b]P_3 + 4bP_4 \\
&\cdots \\
P_k' &= \alpha P_{k-1} - kbP_k
\end{aligned} \tag{8}$$

where $P_n \equiv P_n(t)$ is the probability the system is in state n at time t and primes denote time derivatives. The steady state probability p_n that exactly n customers are present at a time picked at random is

$$p_n = \binom{k}{n}(\alpha b^{-1})^n p_0, \qquad p_0 = (1 + \alpha b^{-1})^{-k}. \tag{9}$$

This follows upon setting the derivatives in (8) to zero, taking $P_n(t) = p_n$, and solving the equations step by step. The expected number $\bar{n}$ of customers present is $k\alpha b^{-1}/(1 + \alpha b^{-1})$ and the average arrival rate is

$$\alpha(k - \bar{n}) = k\alpha/(1 + \alpha b^{-1}). \tag{10}$$

The quality of service is now

$$q = 1 - p_0 \tag{11}$$

and expression (3) for q may be obtained by using $\alpha = 1/(1 - p)T$, $b^{-1} = pT$. Since, (*i*) an outage corresponds to state 0, (*ii*) state 0 can end only through an arrival, and (*iii*) the arrivals in state 0 are Poisson with rate $k\alpha$, it follows that $\exp(-k\alpha t)$ is the probability (predicted by the model) that the length of an outage will exceed t. This agrees with the exponential approximation (6), and the average value $\bar{t}_0 = 1/(k\alpha)$ agrees with (4).

It is known that expression (9) for the steady state probability p_n holds not only for exponential service but also for the general service distribution $B(t)$ (Ref. 3., p. 90). Hence the model predicts that expressions (3), (4) and (6) still hold when the length of time a particular satellite stays in view has an arbitrary distribution $B(t)$.

V. INNAGE DISTRIBUTION PREDICTED BY MODEL

The average innage length (i.e., average busy period) predicted by the model when the service distribution $B(t)$ is arbitrary follows from $\bar{t}_o = 1/k\alpha$ and $q = 1 - p_0$:

$$\bar{l}_i = \frac{1 - p_0}{p_0}\bar{l}_o = [(1 + \alpha b^{-1})^k - 1]/k\alpha.$$

This much is easy. It is much more difficult to get the complete distribution, as the following work shows.

When the service lengths are exponentially distributed, the busy period distribution may be obtained by solving a "first passage" problem. State 0 is made absorbing and the system is started with an arrival at time 0. Thus the system starts in state 1 at time 0 and jumps from state to state in accordance with the arrivals and departures of customers. The system eventually lands in state 0 and stays there. This corresponds to the end of the busy period or innage.

When state 0 is made absorbing, the first two of the $k + 1$ state equations (8) are replaced by

$$\begin{aligned} P_0' &= bP_1 \\ P_1' &= -[(k-1)\alpha + b]P_1 + 2bP_2 . \end{aligned} \tag{12}$$

The modified equations (8) are to be solved subject to $P_1 = 1$ and $P_0, P_2, \cdots, P_k = 0$ at time 0. The probability that the length of an innage will exceed the length t is

$$G(t) = 1 - P_0 . \tag{13}$$

Step-by-step computation of the derivatives of P_0 at $t = 0$ from the differential equations gives the power series

$$\begin{aligned} G(t) = 1 - \frac{bt}{1!} + \frac{[(k-1)\alpha + b]bt^2}{2!} \\ - \frac{[(k-1)^2\alpha^2 + 4(k-1)\alpha b + b^2]bt^3}{3!} \\ + [(k-1)^3\alpha^3 + (k-1)(9k-11)\alpha^2 b \\ + 11(k-1)\alpha b^2 + b^3]\frac{bt^4}{4!} + \cdots \end{aligned} \tag{14}$$

which is useful for small values of t.

Since P_1 is determined by the last k differential equations of the modified set, and since the coefficients in these equations are constants, we may expect P_1 [and hence $G(t)$] to be expressible as the sum of k exponential terms. Indeed, when $\varphi_n(s)$ is used to represent the Laplace transform of $P_n(t)$, the k differential equations for $P_1(t), \cdots, P_k(t)$ go

into k linear equations for $\varphi_1(s)$, $\cdots$, $\varphi_k(s)$. By introducing the generating function

$$\Phi(x, s) = x\varphi_1(s) + \cdots + x^k\varphi_k(s)$$

in the usual way, the linear equations may be combined into

$$b\varphi_1(s) - x = (1 - x)(b + \alpha x)\frac{\partial\Phi(x, s)}{\partial x} - (s + \alpha k - \alpha k x)\Phi(x, s).$$

To obtain $\varphi_1(s)$, rewrite this equation as

$$[b\varphi_1(s) - x](1 - x)^{z-1}(b + \alpha x)^{-z-k-1} = \frac{\partial}{\partial x}(1 - x)^z(b + \alpha x)^{-z-k}\Phi(x, s) \tag{15}$$

where $z = s/(\alpha + b)$. Assume Re $(z) > 0$ and integrate (15) from $x = 0$ to $x = 1$. The right-hand side vanishes because $\Phi(0, s)$ is zero. Changing the variable of integration on the left-hand side from x to $y = (1 - x)/(b + \alpha x)$ gives

$$b\varphi_1(s)\int_0^{1/b} y^{z-1}(1 + \alpha y)^k dy = \int_0^{1/b}(1 - yb)y^{z-1}(1 + \alpha y)^{k-1} dy. \tag{16}$$

Expanding $(1 + \alpha y)^k$ and integrating termwise shows that the coefficient of $b\varphi_1(s)$ is $b^{-z}F_k(z)$ where

$$F_k(z) = \sum_{n=0}^{k}\binom{k}{n}\frac{(\alpha b^{-1})^n}{z + n}. \tag{17}$$

When the integrand on the right-hand side of (16) is replaced by its equivalent

$$\left[1 + \frac{z(\alpha + b)}{k\alpha}\right]y^{z-1}(1 + \alpha y)^k - \frac{\alpha + b}{k\alpha}\frac{d}{dy}y^z(1 + \alpha y)^k$$

(which is suggested by an integration by parts and adding and subtracting various terms) we obtain

$$b\varphi_1(s)F_k(z) = \left(1 + \frac{s}{k\alpha}\right)F_k(z) - \frac{b}{k\alpha}(1 + \alpha b^{-1})^{k+1}.$$

From $P_0' = bP_1$ it follows that $s\varphi_0(s) = b\varphi_1(s)$ and hence

$$\varphi_0(s) = \frac{1}{s} + \frac{1}{k\alpha} - \frac{b(1 + \alpha b^{-1})^{k+1}}{k\alpha s F_k(z)}.$$

As $s \to \infty$, $F_k(z)$ tends to $(1 + \alpha b^{-1})^k/z$ and $\varphi_0(s)$ is $O(1/s)$. Writing

$\varphi_0(s)$ as the sum of partial fractions, inverting to obtain $P_0(t)$, and using (13) gives the expression we seek:

$$G(t) = \sum_{m=0}^{k-1} \frac{b(1 + ab^{-1})^{k+1} \exp[(\alpha + b)z_m t]}{k\alpha z_m F_k'(z_m)} \tag{18}$$

where $F_k'(z) = dF_k(z)/dz$ and $z_0, z_1, \cdots, z_{k-1}$ are the zeros of $F_k(z)$. These zeros lie between the poles at $0, -1, -2, \cdots, -k$. When t is large $G(t)$ is given, effectively, by the term corresponding to $m = 0$ in (18). Usually z_0 is close to $z = 0$ and may be obtained by successive approximations from $F_k(z) = 0$, i.e., from

$$\frac{1}{z} = -\sum_{n=1}^{k} \binom{k}{n} \frac{(\alpha b^{-1})^n}{z + n}.$$

The foregoing method is the one originally used to obtain $G(t)$ as the sum of exponential terms. Subsequently, a more elegant method of obtaining (18) for $G(t)$ was suggested by L. Takács. His method removes the restriction that the service time distribution be exponential.

Takács' result is the following: Let $\beta(s)$, $\gamma(s)$ be the Laplace transforms of $B'(t)$, $-G'(t)$, the service time and busy period probability densities, respectively. The equation to determine $\gamma(s)$, given $\beta(s)$, is

$$\frac{1}{s + k\alpha - k\alpha\gamma(s)} = \int_0^\infty e^{-st}[P_{aa}(t)]^k \, dt \tag{19}$$

where $P_{aa}(t)$ has the Laplace transform $1/[s + \alpha - \alpha\beta(s)]$. It will appear later that the subscript a refers to the idle state (state 0). The expression (18) for $G(t)$ may be obtained from (19) by starting with $\beta(s) = b/(s + b)$. It turns out that $P_{aa}(t)$ is given by $[b + \alpha e^{-(\alpha+b)t}]/(b + \alpha)$ so that the Laplace transform of $[P_{aa}(t)]^k$ is not difficult to compute.

One way to establish (19) is to regard the model as composed of k independent simple systems, each consisting of a source connected to a server. Consider a simple system. The lengths of the idle and busy periods have the respective probability densities $\alpha e^{-\alpha t}$, $B'(t)$ with Laplace transforms $\alpha/(\alpha + s)$, $\beta(s)$.

From (26) of Appendix A, the probability $P_{aa}(t)$ that an idle period will be in progress at time t, given that one is in progress at time 0, may be determined by inverting its Laplace transform $1/[s + \alpha - \alpha\beta(s)]$. The probability that all k servers are idle at time t, given that they are idle at time 0, is $[P_{aa}(t)]^k$. Equation (19) now follows upon using (26) again. This time the type (a) intervals correspond to the periods (outages) during which all k servers are idle. The arrival rate is αk and $p_a(t)$ is $\alpha k e^{-\alpha k t}$. The type (b) intervals correspond to the innages with probability density $-G'(t)$ and Laplace transform $\gamma(s)$.

VI. COMPARISON WITH SIMULATION

It is interesting to compare the results predicted by the model with those obtained by simulation. As an example, we shall take results obtained by Rinehart and Robbins for a Maine-Western Europe link. This link was assumed to have 18 satellites in random polar orbits at a height of 6,000 nautical miles.

The positions of the orbits, the locations of the receiver and transmitter, and some computations involving a number of representative passages over the region of mutual visibility lead to

1. the value $T = 6.35$ hours for the average interval between reappearances of a particular satellite,
2. the distribution function $B(t)$ for l, the length of time it is visible,
3. an average value of l equal to $b^{-1} = 1.46$ hours,
4. the value $p = b^{-1}/T = 0.230$ for the fraction of time the satellite is visible.

It turns out that the probability density $B'(t)$ can be approximated by the rectangle

$$B'(t) = \begin{cases} 0 , & t < 0.86 \text{ hours} \\ 0.833, & 0.86 < t < 2.06 \\ 0 , & 2.06 < t. \end{cases} \tag{20}$$

The system was required to furnish a quality of service close to 0.99. This requirement together with the expression $q = 1 - (1 - p)^k$ and the value $p = 0.230$ gives $k = 18$ and $q = 0.99094$. The model parameters are therefore taken to be

$$k = 18, \qquad b^{-1} = 1.46 \text{ hours}, \qquad \alpha = \frac{1}{(1-p)T} = \frac{1}{(0.770)(6.35)} = 0.205 \tag{21}$$

with the understanding that $p = 0.230$.

The values of $q, \bar{t}_o, \bar{t}_i$ obtained by simulation are compared with those predicted by the formulas of Section II (and also by the model) in Table I.

The second column gives values obtained by Rinehart and Robbins by a simulation which followed the system for 18 months. The third column gives values computed from the model parameters (21). It is

TABLE I

Quantity	Simulation	Model, $p = 0.230$	Model, $p = 0.234$
q	0.9918	0.99094	0.9918
$\bar{t}_o$	0.291 hours	$1/k\alpha = 0.271$	0.270
$\bar{t}_i$	35.1 hours	$q(1-q)^{-1}\bar{t}_o = 29.6$	32.6
b^{-1}	—	1.46 hours	1.48

seen that the values of the average outage length $\bar{t}_o$ agree fairly well. The discrepancy in $\bar{t}_i$ reflects the shortcomings of the model. The last column shows what happens when we hold k and T at their former values of 18 and 6.35, and fudge the value of p so as to make q have the simulation value 0.9918. This changes p from 0.230 to 0.234 and $\bar{t}_i$ from 29.6 to 32.6. It is seen that the value of $\bar{t}_i$ is very sensitive to such changes.

Fig. 1 shows three curves for the innage length distribution. The ordinate is $G(t)$, the probability that an innage length will exceed t. Curve A is the curve predicted by the model assuming exponential service, and is obtained by substituting the parameter values (21) in expressions (14) and (18) for $G(t)$. Curve B is the exponential approximation

$$G(t) = e^{-t/\bar{t}_i} \tag{22}$$

with $\bar{t}_i$ equal to the simulation value 35.1 hours. Curve C is the result obtained by simulation. During the 18 months simulated there were 122 innages, the longest lying between 250 and 260 hours, the next longest between 190 and 200 hours, and so on. The average innage lengths corresponding to curves A and C are $\bar{t}_i = 29.6$ and $\bar{t}_i = 35.1$, respectively, in agreement with the table above.

The curves shown in Fig. 1 agree moderately well. The agreement would be improved if curve A could be shifted so as to give an average value of 35.1 instead of 29.6. Some of the discrepancy between curves A and C around $t = 0$ can be ascribed to the assumption of exponential service. Better agreement in this region could be obtained by taking the model to have the (almost) true service distribution (20). The procedure for doing this is indicated by (19) but the task of carrying through the work seems to be difficult. Again, it should be possible to use the "Erlang service" approximation $4b^2t \exp(-2bt)$ for $B'(t)$ in (19) and also to solve the corresponding first passage problem [i.e., solve the equations corresponding to (8) and (12)]. However, this was not attempted.

Some idea of the change produced in $G(t)$ when exponential service is replaced by other kinds of service may be obtained from Fig. 2. The curves of Fig. 2 show $G(t)$ for the simplified model based upon the results of Appendix B. Exponential service (curve D) and constant service

time (curve E) are assumed, and Poisson arrivals are taken for both cases. The average arrival rate is $a = k/T = 18/6.35 = 2.84$ per hour and the average service time is $b^{-1} = 1.46$ hours. Substitution of these values in the expressions for $G(t)$ given in examples (a) and (b) of Appendix B give curves D and E, respectively.

The expanded scale at the top of Fig. 2 shows the behavior of $G(t)$ around $t = 0$. Both distributions predict the same average innage length, namely

$$\bar{t}_i = [e^{ab^{-1}} - 1]a^{-1} = 22.0 \text{ hours} \tag{23}$$

which is (43) in Appendix B. The discrepancy between 22.0 and the value $\bar{t}_i = 29.6$ given by the model of Section III shows the shortcomings

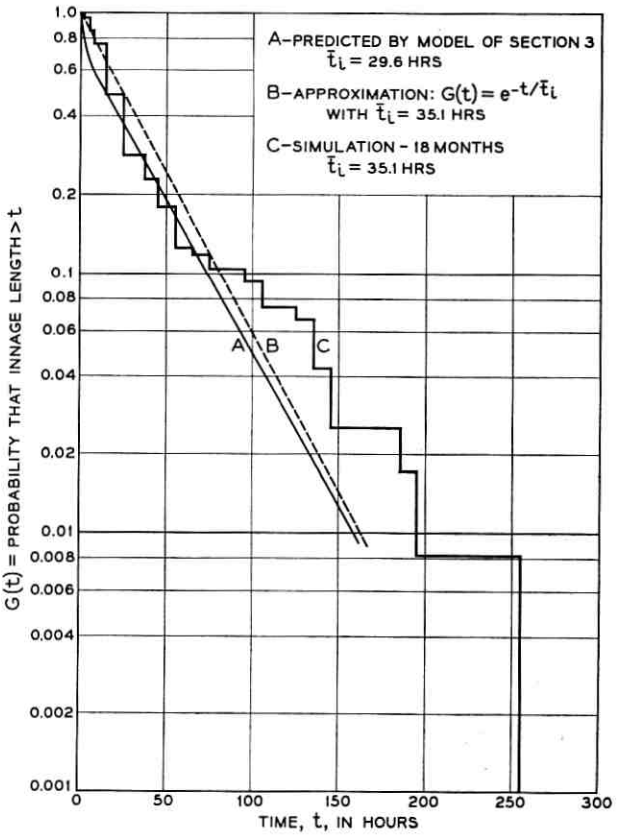


Fig. 1 — Innage length distribution for 18-satellite system with random polar orbits (Maine–Western Europe).

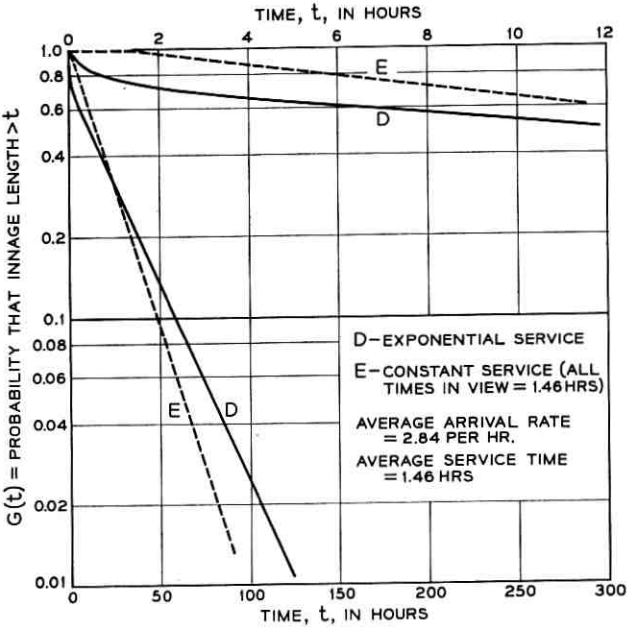


Fig. 2 — Innage length distributions.

of the simplified model. Nevertheless, it appears that the difference in shape between curves D and E illustrates the change in $G(t)$ produced by the different kinds of service. Support for this belief comes from the fact that curves A and D, both of which correspond to exponential service, have the same shape.

In view of the inaccuracies of the models and of the relatively good agreement shown by the curves B and C of Fig. 1, it seems that the simple exponential approximation for $G(t)$ is quite good.

ACKNOWLEDGMENT

I wish to express my thanks for the help I have received from my colleagues, especially J. D. Rinehart and M. F. Robbins, whose work on simulation gave rise to the present paper, and L. Takács, whose remarks led to a number of simplifications and generalizations. Many improvements in the presentation resulted from suggestions made by John Riordan. Additional information and suggestions were obtained from conversations with F. W. Sinden and W. H. Wise.

APPENDIX A

Probabilities Associated with Alternating Sequences

Expressions are recorded here for the Laplace transforms of two conditional probabilities. These transforms are of use in establishing Takács' result (19).

Consider a sequence comprised of two kinds of intervals which alternate with each other (for example, innages and outages). Let $p_a(t)$, $p_b(t)$ be the probability densities for the lengths of the two types of intervals, and let their respective Laplace transforms be $\alpha(s)$, $\beta(s)$. Suppose that the interval lengths are independent, let an interval of type (a) start at time 0, and let $P_{aa}(t)$ be the probability that an interval of type (a) is in progress at time t. Then the Laplace transform of $P_{aa}(t)$ is

$$\int_0^\infty e^{-st} P_{aa}(t)\, dt = \frac{1 - \alpha(s)}{s[1 - \alpha(s)\beta(s)]}. \tag{24}$$

Similarly, if an interval of type (b) starts at time 0, the Laplace transform of the probability $P_{ba}(t)$ that an interval of type (a) is in progress at time t is

$$\frac{[1 - \alpha(s)]\beta(s)}{s[1 - \alpha(s)\beta(s)]}. \tag{25}$$

These results are reminiscent of the relations between generating functions in the theory of recurrent events (Ref. 4, ch. 12).

The expression (24) may be obtained by noting that

$$\begin{aligned} P_{aa}(t) &= \Pr(t_{a1} > t) + \Pr(t_{a1} + t_{b1} + t_{a2} > t \\ &\qquad \text{and} \quad t_{a1} + t_{b1} < t) + \cdots \\ &= \Pr(t_{a1} > t) + [\Pr(t_{a1} + t_{b1} + t_{a2} > t) \\ &\qquad - \Pr(t_{a1} + t_{b1} > t)] + \cdots \end{aligned}$$

where t_{a1}, t_{b1}, t_{a2}, $\cdots$ are the lengths of the successive (a) and (b) type intervals. The Laplace transforms of $\Pr(t_{a1} > t)$, $\Pr(t_{a1} + t_{b1} > t)$, $\cdots$ are $[1 - \alpha(s)]/s$, $[1 - \alpha(s)\,\beta(s)]/s$, $\cdots$, and summation gives (24) when $|\alpha(s)\,\beta(s)| < 1$ (as it certainly is when $\mathrm{Re}(s) > 0$). Expression (25) may be obtained in a similar fashion or by convoluting $p_b(t)$ and $P_{aa}(t)$.

When $p_a(t) = a \exp(-at)$, the condition that a type (a) interval

start at time 0 may be replaced by the condition that a type (a) interval be in progress at time 0. In this case $P_{aa}(t)$ is also the probability that a type (a) interval is in progress at time t, given that one was in progress at time 0. From (24) and $\alpha(s) = a/(a + s)$, the Laplace transform of this probability is

$$\frac{1}{s + a - a\beta(s)}. \tag{26}$$

APPENDIX B

Systems with an Infinite Number of Servers and Recurrent Inputs

This appendix will be concerned with systems containing an infinite number of servers and a "recurrent" input, i.e., an input in which the lengths of the intervals between successive arrivals are independent of each other and of the state of the servers. In many respects these systems are simpler than the limited source input introduced in Section III. Although they do not represent the satellite system as well, their greater simplicity enables us to estimate the shape of the innage distribution for cases which are difficult to handle by Takács' result (19).

In the following list of results, $A(t)$ is the distribution function for the distances between the arrivals of the recurrent input and $B(t)$ is the service time distribution for each one of the infinite number of servers. The expected distance between arrivals is a^{-1} and the expected service time is b^{-1}.

B.1 *The Conditional Probability* $P_{10}(t)$

Let all of the servers be idle at time -0 and let the first customer arrive at time 0 making one server busy at $+0$. Denote by $P_{1n}(t)$ the conditional probability that n servers are busy at time t. Consideration of the first arrival following time $+0$ leads to an integral equation which (in theory) may, be solved for $P_{10}(t)$, namely

$$P_{10}(t) = [1 - A(t)]B(t) + B(t) \int_0^t A'(t - v)P_{10}(v)\, dv, \tag{27}$$

where $A'(u) = dA(u)/du$. A corresponding equation for the generating function for $P_{1n}(t)$ is given by (44).

Example (a). For Poisson input $1 - A(t)$ is exp $(-at)$. This corresponds to an unlimited source input. Substituting in (27), multiplying through by $[\exp(at)]/B(t)$, and differentiating with respect to t gives a

differential equation for $P_{10}(t)$. Using the fact that $P_{10}(t) \to B(t)$ as $t \to 0$ to fix the constant of integration leads to

$$P_{10}(t) = B(t) \exp\left[-a \int_0^t [1 - B(\tau)]\, d\tau\right], \tag{28}$$

a result given by Refs. 5 and 6. The work may be simplified by starting with the assumption that $P_{10}(t)$ is of the form $B(t)P(t)$.

Example (*b*). For regularly spaced arrivals, $A'(u) = \delta(u - a^{-1})$ where $\delta(t)$ is the unit impulse. Equation (27) then gives

$$\begin{aligned} P_{10}(t) &= B(t), && 0 < t < a^{-1} \\ P_{10}(t) &= B(t)P_{10}(t - a^{-1}) && t > a^{-1} \\ P_{10}(t) &= B(t)B(t - a^{-1}) && a^{-1} < t < 2a^{-1} \\ P_{10}(t) &= B(t)B(t - a^{-1})B(t - 2a^{-1}) && 2a^{-1} < t < 3a^{-1} \end{aligned} \tag{29}$$

and so on.

Example (*c*). The case when $A(t)$ is arbitrary and $1 - B(t) = \exp(-t)$ has been studied by Takács[7] (see also Ref. 3, p. 33 et seq.). Multiplying (27) by $\exp(-st)$, integrating t from 0 to ∞, and introducing the Laplace transforms $\alpha(s)$, $\theta(s)$ of $A'(t)$, $P_{10}(t)$ leads to a recurrence relation between $\theta(s)$ and $\theta(s + 1)$ which in turn gives

$$\begin{aligned} \theta(s) = \frac{1}{s} &- \frac{1}{(s+1)(1-\alpha_s)} + \frac{\alpha_{s+1}}{(s+2)(1-\alpha_s)(1-\alpha_{s+1})} \\ &- \frac{\alpha_{s+1}\alpha_{s+2}}{(s+3)(1-\alpha_s)(1-\alpha_{s+1})(1-\alpha_{s+2})} + \cdots \end{aligned} \tag{30}$$

where α_{s+n} is written for $\alpha(s + n)$. When $\theta(s)$ is known $P_{10}(t)$ may be obtained by inversion.

B.2 *The Busy Period Distribution $G(t)$ for Poisson Arrivals*

Let $-G'(t)$ be the probability density for the lengths of the busy periods (corresponding to innages) and consider the case of Poisson input and arbitrary service. The Laplace transform $\gamma(s)$ of $-G'(t)$ is given by

$$\gamma(s) = \frac{(s + a)\theta(s)}{1 + a\theta(s)} \tag{31}$$

where $\theta(s)$ is the Laplace transform of $P_{10}(t)$.

My original derivation of (31) was based on taking the Laplace transform of

$$-G'(t) = f(t) - a\int_0^t ds\, P_{10}(t-s)f(s) + a^2\int_0^t ds\, P_{10}(t-s)\int_0^s dr\, P_{10}(s-r)f(r) - \cdots, \tag{32}$$

$$f(t) = P_{10}'(t) + aP_{10}(t), \tag{33}$$

where $P_{10}(t)$ is given by (28) and $f(t)\,dt$ is the probability that the system will jump from state 1 to state 0 in $(t, t+dt)$, given an arrival at time 0 which ends an idle period. The series (32) was obtained by an application of the method of inclusion and exclusion. Subsequently, Takács obtained a formula equivalent to (31) by an elegant method based upon results of the type stated in Appendix A. When this method is applied to obtain (31) the type (a) intervals are taken to be idle periods (outages) with $p_a(t) = a\exp(-at)$, $\alpha(s) = a/(a+s)$. The type (b) intervals become the busy periods so that $p_b(t)$, $\beta(s)$ become $-G'(t)$, $\gamma(s)$, respectively; and $P_{ba}(t)$ goes into $P_{10}(t)$. Expression (25) of Appendix A then says that the Laplace transform of $P_{10}(t)$ is

$$\theta(s) = \frac{\gamma(s)}{s + a - a\gamma(s)} \tag{34}$$

from which (31) follows.

Example (a). For Poisson arrivals and exponential service with $B = 1 - e^{-bt}$, (28) becomes (Ref. 3, p. 26)

$$P_{10}(t) = (1 - e^{-bt})\exp[-\rho + \rho\, e^{-bt}], \qquad \rho = ab^{-1}.$$

The change of variable $y = e^{-bt}$ carries the integral for the Laplace transform of $P_{10}(t)$ into

$$b\theta(s) = F(z) - F(z+1), \quad z = sb^{-1}$$

where

$$F(z) = e^{-\rho}\int_0^1 y^{z-1}e^{\rho y}\,dy = \frac{1}{z} - \frac{\rho}{z(z+1)} + \frac{\rho^2}{z(z+1)(z+2)} - \cdots \tag{35}$$

$$= z^{-1}[1 - \rho F(z+1)] = \sum_{n=0}^{\infty}\frac{\rho^n e^{-\rho}}{n!(z+n)}.$$

The Laplace transform of $G(t)$ may be shown to be

$$\frac{1 - \gamma(s)}{s} = \frac{F(z+1)}{sF(z)} = \frac{1}{azF(z)} - \frac{1}{a} \tag{36}$$

and inversion gives

$$G(t) = \sum_{m=0}^{\infty} \frac{e^{z_m bt}}{\rho z_m F'(z_m)}, \qquad \rho = ab^{-1} \tag{37}$$

where $F'(z) = dF(z)/dz$. The zeros of $F(z)$ are real and (a) occur at $z_0, z_1, z_2, \cdots$, (b) are such that $-1 < z_0 < 0, -2 < z_1 < -1, \cdots$, etc. and hence lie between the poles of $F(z)$ at $0, -1, -2, \cdots$, (c) may be computed by successive approximations with the help of the last series for $F(z)$ in (35).

A power series for $G(t)$ may be obtained by expanding $[1 - \gamma(s)]/s$ in powers of $1/s$ and then replacing s^{-n-1} by $t^n/n!$. The same series may be obtained from the corresponding series (5.3) for the limited source case by letting $k \to \infty$, $\alpha \to 0$ in such a way as to keep $k\alpha$ equal to a. Replacing a by ρb then gives

$$G(t) = 1 - \frac{bt}{1!} + \frac{(\rho+1)(bt)^2}{2!} - \frac{(\rho^2+4\rho+1)(bt)^3}{3!} + \frac{(\rho^3+9\rho^2+11\rho+1)(bt^4)}{4!} - \cdots. \tag{38}$$

Example (*b*). For Poisson arrivals and constant service time, $B(t)$ jumps from 0 to 1 at time $t = b^{-1}$. Equation (28) shows that $P_{10}(t)$ jumps from 0 to $\exp(-ab^{-1})$ at $t = b^{-1}$. The Laplace transforms are readily computed and the one for $G(t)$ gives

$$G(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{[1 - e^{-(a+s)b^{-1}}]e^{st}}{s + ae^{-(a+s)b^{-1}}} ds, \qquad c > 0. \tag{39}$$

Taking c large enough, say $c = b$, to make $|x| < 1$ where x is as^{-1} $\exp[-(a+s)b^{-1}]$ and expanding the denominator in powers of x leads to

$$G(t) = 1 + \sum_{n=1}^{N} (-\rho e^{-\rho})^n \left[\frac{(bt-n)^{n-1}}{(n-1)!\rho} + \frac{(bt-n)^n}{n!} \right]. \tag{40}$$

Here, as in example (a), $\rho = ab^{-1}$ and the upper limit of summation is determined from $N < bt < N + 1$. In particular,

$$G(0) = 1, \quad G(b^{-1} - 0) = 1, \quad G(b^{-1} + 0) = 1 - e^{-\rho}$$
$$G(2b^{-1}) = 1 - (1 + \rho)\, e^{-\rho} \tag{41}$$
$$G(3b^{-1}) = 1 - (1 + 2\rho)\, e^{-\rho} + \rho \left(1 + \frac{\rho}{2}\right) e^{-2\rho}.$$

For large values of t

$$G(t) \sim \frac{(\sigma_0 + \rho)}{\rho(1 + \sigma_0)}\, e^{\sigma_0 bt} \tag{42}$$

where σ_0 is the rightmost root of

$$\sigma + \rho\, e^{-\rho-\sigma} = 0.$$

When ρ is large σ_0 is approximately $-\rho\, e^{-\rho}$.

The innages corresponding to examples (a) and (b) have the same average length, namely

$$\bar{t}_i = [e^{ab^{-1}} - 1]/a. \tag{43}$$

To see this, note that from traffic theory, or by letting $t \to \infty$ in (28), the fraction of idle time is $p_0 = 1 - q = \exp(-ab^{-1})$. The average outage time is $\bar{t}_0 = 1/a$, and (43) follows upon using the relation (1) between $\bar{t}_i$, $\bar{t}_o$, and q.

B.3 *Miscellaneous Results for Recurrent Input*

Except for the case in which arrivals occur at multiples of some fixed spacing, $P_{10}(t)$ approaches the steady state probability p_0 as $t \to \infty$, and the Laplace transform $\theta(s)$ of $P_{10}(t)$ has a simple pole at $s = 0$ with residue p_0.

If $P_{10}(t)$ tends to a periodic function the residue gives its average value. Application of this result to the case of regularly spaced arrivals and exponential service with $B(t) = 1 - e^{-t}$ leads to the rather curious expansion

$$\int_0^1 \prod_{n=0}^{\infty} (1 - x^{-\tau-n})\, d\tau$$
$$= 1 - \frac{1}{\ln x}\left[1 - \frac{1}{2(x-1)} + \frac{1}{3(x-1)(x^2-1)} - \cdots\right].$$

Both sides represent the average value of the periodic function to which

$P_{10}(t)$ tends as $t \to \infty$. The integral on the left (with $x = \exp a^{-1}, x > 1$) is the average value as computed from (29), while the series on the right is the residue of $\theta(s)$ at $s = 0$ obtained by setting $\alpha(s) = \exp(-sa^{-1})$ in (30) and letting $s \to 0$.

The generating function

$$P_1(x, t) = \sum_{n=0}^{\infty} x^n P_{1n}(t)$$

for the conditional probability $P_{1n}(t)$ that state n exists at time t, given that an arrival at time 0 ends an idle period, [see (27)] satisfies the integral equation

$$P_1(x, t) = [x + (1 - x)B(t)] \cdot \left[1 - A(t) + \int_0^t A'(t - v)P_1(x, v)\, dv\right]. \tag{44}$$

A formal step-by-step solution may be obtained by introducing the binomial moments $M_n(t)$ defined by

$$P_1(x, t) = \sum_{n=0}^{\infty} (x - 1)^n M_n(t),$$

$$M_n(t) = \sum_{k=n}^{\infty} \binom{k}{n} P_{1k}(t).$$

The value of $M_0(t)$ is one and the higher-order moments satisfy integral equations obtainable from (44). When the Laplace transforms of $A'(t)$, $B'(t)$, $M_n(t)$ are denoted by $\alpha(s)$, $\beta(s)$, $\mu_n(s)$ it is found that the integral equations lead to

$$\mu_0(s) = s^{-1}, \qquad \mu_1(s) = \frac{1 - \beta(s)}{s[1 - \alpha(s)]}$$

$$\mu_n(s) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{[1\ \beta(s - z)]}{s - z} \frac{\alpha(z)\mu_{n-1}(z)}{1 - \alpha(s)}\, dz, \qquad n > 1. \tag{45}$$

The singularities of $\alpha(z)\mu_{n-1}(z)$ are supposed to lie to the left and those of $[1 - \beta(s - z)]/(s - z)$ to the right of the path of integration, s being chosen so as to make this possible. In theory, the successive values of $\mu_n(s)$ may be obtained step by step and thus ultimately lead to an expression for $P_1(x, t)$. For exponential service the integrals in (45) may be evaluated and lead to results given by Takács[7] (see also Ref. 3, p. 33).

REFERENCES

1. Rinehart, J. D., and Robbins, M. F., Characteristics of the Service Provided by Communication Satellites in Uncontrolled Orbits, B.S.T.J., this issue.
2. Pierce, J. R., and Kompfner, R., Transoceanic Communication by means of Satellites, Proc. I.R.E., **47,** Mar., 1959, pp. 372–380.
3. Riordan, J., *Stochastic Service Systems*, Wiley, New York, 1962.
4. Feller, W., *An Introduction to Probability Theory and Its Applications*, *v.* I, Wiley, New York, 1950.
5. Riordan, J., Telephone Traffic Time Averages, B.S.T.J., **30,** Oct., 1951, pp. 1129–1144.
6. Benes, V. E., Fluctuations of Telephone Traffic, B.S.T.J., **36,** July, 1957, pp. 965–973.
7. Takács, L., On the Limiting Distribution of the Number of Coincidences Concerning Telephone Traffic, Annals of Math. Stat., **30,** 1959, pp. 134–141.

Contributors to This Issue

Vaclav E. Beneš, A.B., 1950, Harvard College; M.A. and Ph.D., 1953, Princeton University; Bell Telephone Laboratories, 1953—. Mr. Beneš has been engaged in mathematical research on stochastic processes, traffic theory, and servomechanisms. In 1959–60 he was visiting lecturer in mathematics at Dartmouth College. Member American Mathematical Society, Association for Symbolic Logic, Institute of Mathematical Statistics, Society for Industrial and Applied Mathematics, Mind Association, Phi Beta Kappa.

J. William Elek, B.S., 1957, Case Institute of Technology; M.S., 1961, Lehigh University; Bell Telephone Laboratories, 1958—. He has worked on the mechanical reliability and environment of electron tubes and semiconductors. He also has engaged in the investigation of thermal stresses induced in electron tubes and semiconductors as a result of fabrication techniques. At present he is studying the mechanical (shock and vibration) environment of electron tubes and semiconductors to improve reliability. Member Tau Beta Pi, Sigma Xi.

John S. Elliott, Jr., B.S., 1922, Pennsylvania State College; Bell Telephone Laboratories 1922—. His first assignment was spent in routine measurements; condenser (mica paper) development and design; development of precision testing equipment and apparatus used in measuring the various electrical constants of transmission apparatus. In addition, during World War II he was responsible for the production of numerous direct capacitance test sets urgently needed by vacuum tube manufacturers in controlling the performance of vacuum tubes. Member Eta Kappa Nu, Alpha Sigma Phi.

Hans L. Kreipe, Dipl. Ing., E. E., 1960, Technische Hochschule, Braunschweig, Germany. Scientific Assistant at the Institut fuer Hoechst frequenztechnik, Technische Hochschule, Braunschweig 1961—. Mr. Kreipe is presently engaged in research in microwaves.

Dietrich Marcuse, Diplom Vorpruefung, 1952 and Dipl. Phys., 1954, Berlin Free University; Siemens and Halske (Germany), 1954–

1957; Bell Telephone Laboratories, 1957—. At Siemens and Halske Mr. Marcuse was engaged in transmission research, studying coaxial cable and circular waveguide transmission. At Bell Laboratories he has been engaged in studies of circular electric waveguides and work on gaseous masers. Member I.R.E.

STEPHEN O. RICE, B.S., 1929, D.Sc. (Hon.), 1961, Oregon State College; Graduate Studies, California Inst. of Tech., 1929–30 and 1934–35; Bell Telephone Laboratories, 1930—. In his first years at the Laboratories, Mr. Rice was concerned with nonlinear circuit theory, especially with methods of computing modulation products. Since 1935 he has served as a consultant on mathematical problems and in investigation of telephone transmission theory, including noise theory, and applications of electromagnetic theory. He was a Gordon McKay Visiting Lecturer in Applied Physics at Harvard University for the Spring, 1958, term. Fellow I.R.E.

JAMES D. RINEHART, B.S., 1957, South Dakota School of Mines and Technology; M.E.E., 1959, New York University; Bell Telephone Laboratories, 1953—. Mr. Rinehart was first engaged in systems engineering of a pulse code modulated carrier system. At present he is supervisor of a group concerned with studies of satellite communications systems. For Project Telstar he was responsible for evaluating various satellite orbits and determining the orbit which best satisfied the objectives of the experiment. Member Sigma Pi Sigma Physics Honor Society.

THOMAS D. RINEY, B.S., 1950, University of Dayton, M.S., 1952, Ph.D., 1954, Purdue University; Bell Telephone Laboratories, 1954–1960; General Electric Space Sciences Laboratory, 1961—. While at Bell Laboratories, he was engaged in work on photoelastic and theoretical studies of stresses in glass seals and other glass structures, such as electron tube envelopes. Later he was engaged in work on heat transfer problems involving klystrons and magnetrons, and design of transistor encapsulation for high pressure applications. Still later, he studied heat transfer associated with semiconductor devices and problems of infrared photoelasticity and thermal stress analysis. Member of Sigma Xi, American Mathematical Society, Society for Experimental Stress Analysis, Society for Industrial and Applied Mathematics.

MARILYN F. ROBBINS, B.A., 1960, Douglass College, Rutgers University; Bell Telephone Laboratories, 1960—. Mrs. Robbins has been

engaged in various studies pertaining to satellite communications, including work for the Telstar experimental satellite program.

W. ROSENZWEIG, B.S., 1950, Rutgers University; M.S., 1952, University of Rochester; Ph.D., 1960, Columbia University. Brookhaven National Laboratory, 1951–1953; Radiological Research Laboratory, Columbia University, 1953–1960; Bell Telephone Laboratories, 1960—. At Bell Laboratories, Mr. Rosenzweig has been mainly engaged in studies of radiation damage to semiconductors. Member American Physical Society, Radiation Research Society, Sigma Xi, Phi Beta Kappa.

IRWIN W. SANDBERG, B.E.E., 1955, M.E.E., 1956, and D.E.E., 1958, Polytechnic Institute of Brooklyn; Bell Telephone Laboratories, 1958—. He has been concerned with analysis of military systems, particularly radar systems, and with synthesis and analysis of active and time-varying networks. Member I.R.E., Eta Kappa Nu, Sigma Xi, Tau Beta Pi.

HANS-GEORG UNGER, Dipl. Ing., 1951 and Dr. Ing., 1954, Technische Hochschule, Braunschweig (Germany); Siemens and Halske (Germany), 1951–55; Bell Telephone Laboratories, 1956—. His work at Bell Laboratories has been in research in waveguides, especially circular electric wave transmission. He is now on leave of absence from Bell Laboratories while professor of electrical engineering at the Technische Hochschule in Braunschweig. Senior member I.R.E.; member German Communication Engineering Society.