THE BELL SYSTEM

Technical Journal

Volume 50 October 1971 Number 8

THE BELL SYSTEM TECHNICAL JOURNAL

THE BELL SYSTEM TECHNICAL JOURNAL is published ten times a year by the American Telephone and Telegraph Company, H. I. Romnes, Chairman and President, J. J. Scanlon, Vice President and Treasurer, R. W. Ehrlich, Secretary. Checks for subscriptions should be made payable to American Telephone and Telegraph Company and should be addressed to the Treasury Department, Room 2312C, 195 Broadway, New York, N. Y. 10007. Subscriptions $10.00 per year; single copies $1.25 each. Foreign postage $1.00 per year; 15 cents per copy. Printed in U.S.A.

THE BELL SYSTEM

TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 50 October 1971 Number 8

On the Addressing Problem for Loop Switching

By R. L. GRAHAM and H. O. POLLAK

(Manuscript received March 25, 1971)

The methods used to perform the switching functions of the Bell System have been developed under the fundamental assumption that the holding time of the completed call is long compared to the time needed to set up the call. In considering certain forms of communication with and among computers the possibility arises that a message, with its destination at its head might thread its way through a communication network without awaiting the physical realization of a complete dedicated path before beginning on its journey. One such scheme has been proposed by J. R. Pierce and may be called "loop switching." We imagine subscribers, perhaps best thought of as computer terminals or other data generating devices, on one-way loops. These "local" loops are connected by various switching points to one another as well as to other "regional" loops which are in turn connected to one another as well as to a "national" loop. If a message from one loop is destined for a subscriber on another loop it proceeds around the originating loop to a suitable switching point where it may choose to enter a different loop, this process continuing until the message reaches its destination. The question naturally comes up, how the message is to know which sequence of loops to follow. It would be desirable for the equipment at each junction to be able to apply a simple test to the destination

address at the head of the message which would determine which choice the message should make at that junction.

In this paper we propose a method of addressing the loops which has several attractive features:

(*i*) *It permits an extremely simple routing strategy to be used by the messages in reaching their destinations.*
(*ii*) *By using this strategy, a message will always take the shortest possible path between any two local loops in the same region.*
(*iii*) *The method of addressing applies to any collection of loops, no matter how complex their interconnections.*

The addressing scheme we propose will be applied primarily to local loops where the mutual interconnections may be quite varied. If a certain amount of hierarchical structure is introduced into the regional and national loop structure, as suggested by J. R. Pierce,[1] it is possible to achieve addressings which are both compact and quite efficient.

I. INTRODUCTION

The methods used to perform the switching functions of the Bell System have been developed under the fundamental assumption that the holding time of the completed call is long compared to the time to set up the call. It is thus sensible to hold portions of a route while the attempt is made to establish the connection. In considering certain forms of communication with and among computers, as well as the consideration of many schemes for time division switching, the possibility arises that a message, with its destination at its head, might thread its way through a communication network without awaiting the physical realization of a complete dedicated path before beginning on its journey.

One such scheme has been proposed by J. R. Pierce,[1] and may be called "loop switching." We imagine subscribers, perhaps best thought of as computer terminals or other data generating devices, on one-way loops. If a meassage is destined for a subscriber on another loop it proceeds around the originating loop to a suitable switching point where it may choose to enter a different loop and continue the process until it reaches its destination.

The question now comes up, how the message is to know which sequence of loops to follow. A sufficiently complicated memory in the originating loop might, of course, look up an appropriate route, and then attempt to seize a complete path; but this is the old and perhaps in-

appropriate solution. It would be more convenient and sometimes preferable if the equipment at each junction could apply some simple test to the destination address at the head of the message which would determine which choice the message should make at that junction.

In the nation-wide loop switching system as conceived by Pierce, we can envisage local loops, regional loops, and a national loop. The simplest imaginable structure is one in which each local loop has an interchange only with its regional loop and with no other loop; similarly each regional loop interchanges with the national loop and otherwise only with its local loops. How does it work? Suppose that a message originates in local loop X, and has its destination in local loop Y, where X and Y may or may not be identical. When the message comes to the interchange between X and X's regional loop, it exits onto the regional loop if and only if $Y \neq X$. It later exits onto the national loop if Y's region is different from X's region; otherwise the message stays on X's regional loop until it reaches Y. Therefore what should addresses look like? We see that if a portion of the loop address represents the regional loop, and another portion the local loop, routing decisions will be made on the basis of identity or nonidentity of certain portions of the sending and the receiving addresses.

The loop configuration just described is perhaps too special to be practical. For example, it provides for no alternate routing, and for no special direct connections between two local loops with high mutual traffic. Pierce has shown how each of these difficulties can, to some extent, be alleviated. There remain, however, the further problems of the configuration of local loops belonging to a given region, and of the configuration of regional loops themselves. It is quite likely that the local loops attached to a given regional loop have many mutual switching points among themselves, so that calls within one region are not normally expected to use the regional loop. How should we address such local loops so as to make routing easy? Much of the rest of this paper will be devoted to this problem. We shall, in this and the next two sections, speak of "loops" generally, but mean a system of *local* loops as the most likely realization. We note in passing that a completely general national configuration of loops on which no hierarchical structure has been imposed will have the same addressing problem—but probably a much larger number of loops. We return to the hierarchical situation in Section IV.

In some very simple arrangements of loops it is easy to see how addressing might successfully be accomplished. Consider, for example, Fig. 1 showing four loops which touch as if they were circles of radius $\frac{1}{2}$ at

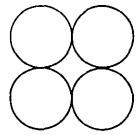


Fig. 1—Simple arrangement of loops.

the vertices of the unit square. If the address of each loop were just the two-digit numeral ij, i, $j = 0$ or 1, representing the coordinates of its center, then routing could be done in the following extremely simple manner: at each junction, go into the new loop if this decreases the Hamming distance* between where you are and your destination. If it doesn't decrease the Hamming distance, don't go. Thus, if you wish to go from loop 10 to loop 11 then the Hamming distance is 1. You will not take the exit from 10 to 00 if you reach it first, for this increases rather than decreases the Hamming distance. You will, however, exit into 11 when you reach that junction. To go from 10 to 01 either exit, to 00 or to 11, improves the Hamming distance and either routine is equally good.

A simple potential routing scheme can thus be described as follows. Each loop has a binary address, n bits long. You make an exit from one loop to another if and only if it decreases the Hamming distance between where you are and where you want to go. If several exits do the same job then each one must lead to an equally short optimal path from sending loop to receiving loop. Furthermore, the number of loops traversed should, if possible, be exactly the Hamming distance between sender and receiver, with each transfer decreasing the distance from the receiver by exactly 1.

Can such an addressing scheme be devised for every collection of loops with whatever adjacency structure? A little reflection shows that there will certainly be difficulties. Let's think of the collection of loops abstractly as a graph, with each loop a vertex, and two vertices connected if and only if the two loops have a mutual transfer point. Thus, the graph of the previous example is as shown on Fig. 2. We have numbered each vertex with a pair of binary digits so that adjacent vertices differ in exactly one position, the number of edges required to pass from one point to another is exactly the Hamming distance between the corresponding numberings, and all shortest paths between two points are achieved by following routes of decreasing Hamming distance to the destination. Another example (Fig. 3): if we wanted a collection of six

* The Hamming distance between two n-place binary numbers is the number of places in which they differ.

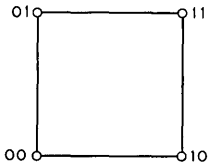


Fig. 2—Graph of Fig. 1.

loops arranged cyclically we could use the numbering 000, 100, 101, 111, 011, and 010. We see that we are looking for a closed path on the 3-dimensional cube with the additional property that two points are exactly as far apart in Hamming distance as the number of edges to be traversed between them—otherwise, the routing logic would be ruined. Thus we can use the realization for a cycle of six loops shown in Fig. 4a. The realization shown in Fig. 4b, however, would *not* be a valid solution. In this latter picture, 100 and 110 have Hamming distance 1 and therefore should be directly connected. The path between them, however, has length 3, the first link out increases rather than decreases Hamming distance, and therefore would not represent a useful addressing scheme.

We thus see a difficulty caused by points coming too close together on the cube for the addressing scheme to work, but there are even deeper difficulties. Suppose we wish to construct an addressing scheme for a system consisting of three pairwise adjacent loops (see Fig. 5). This can never be drawn on a cube of *any* dimension. For any closed path of edges on a cube has even length, and 3 is odd. Is the scheme therefore kaput?

Not quite. We can still imagine the 3-cycle embedded on a cube in an appropriate dimension (in this case a square) if we are willing to generalize what we mean. We shall attach to A the code 00, B the code 10, and to C *both* 11 and 01. We shall denote the pair 11 and 01 by the symbol $d1$, where d means "don't care." Hamming distance between two n-tuples of 0's ,1's, or d's is computed by crediting 1 for every position at

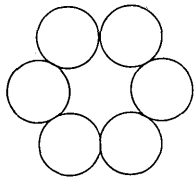


Fig. 3—Cyclic arrangement of six loops.

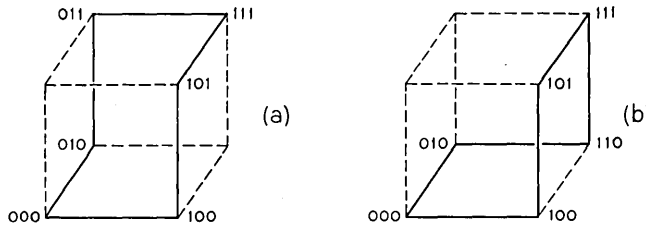


Fig. 4—Realizations for a cycle of six loops: (a) a valid solution; (b) not a valid solution.

which one n-tuple has a 0 and the other a 1, and 0 for every other position. Thus, the Hamming distance between 01d1d0 and 11d010 is 2, with the contributions coming from the first and fourth positions. With this convention, the Hamming distance between any two of the three addresses 00, 10, and d1 is certainly 1, and correct routing still consists exactly of decreasing by 1 the Hamming distance at each junction at which a transfer is made.

We now have a number of fundamental questions to answer. Can *every* collection of loops be numbered by assigning to each loop an address consisting of a sequence of 0's, 1's, and d's? We require that every shortest route between two loops can be found automatically by moving from a loop to an adjacent one if and only if this decreases the Hamming distance to the final destination by 1. How many bits long would such an address have to be? Let's state right away the fundamental theorem of this paper: Every collection of n loops, with maximum distance s between any two loops, can indeed be realized by giving each loop an address of no more than $s(n - 1)$ 0's, 1's, or d's. In fact, we know of no example where more than $(n - 1)$ "bits" are needed, and we shall give a construction that has found addresses no more than $n - 1$ bits long in every case on which it has been tried. The construction, however, is not quite an algorithm and we do not have a proof that it can always be done with as few as $n - 1$ bits.

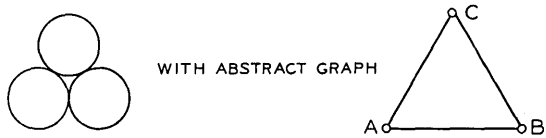


Fig. 5—Three pairwise adjacent loops.

How is this routing algorithm going to work in practice? Here we have only the very earliest and simplest suggestions. The basic idea of the scheme is to obtain the greatest possible simplicity of routing strategy at the expense of the length of the loop address. Thus, for example, you could physically realize addresses consisting of 0's, 1's, and d's by encoding 0 as 00, 1 as 01, and d as either 10 or 11. The logic then says: If the $2k$ — 1st digit of both addresses is 0 then compute the Hamming distance between the $2k$th digits. If the $2k$ — 1st digit of either address is 1, ignore it. Add up over all k, and see if going into the new loop decreases Hamming distance to the destination. This could be very easy to mechanize; the arbitrary bit following a 1 in an odd position could be used for parity checks or other purposes.

It is not immediately clear who assigns the loop address to an individual message. The "phone book" may contain a shorter code that is translated in the first junction you come to, or the sending computer itself may use the destination's correct loop address. This problem is connected with that of system growth. How many numbers do you have to change if a loop is added to the system? The consequent desire for a hierarchical loop address structure is to a large extent fulfillable and will be discussed in Section IV.

Before we proceed with the general theory, let's see how a particular and not so simple example works out. Thus, consider the system of loops in Fig. 6. The distance between pairs of vertices is given by the following (symmetric) table:

	A	B	C	D	E	F
A	0	2	1	3	1	2
B	2	0	2	1	1	2
C	1	2	0	2	1	1
D	3	1	2	0	2	1
E	1	1	1	2	0	2
F	2	2	1	1	2	0

We shall assign a sequence of five 0's, 1's, and d's to each vertex in such a way that the Hamming distance between the 5-tuples corresponding to two vertices is exactly the distance in the table. One solution, as the reader should verify, is the following:

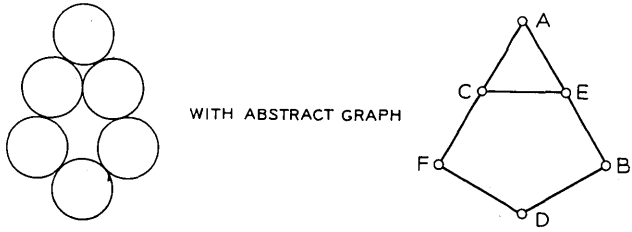


Fig. 6—System of loops.

A—1111d

B—001dd

C—11d0d

D—000d1

E—10dd0

F—010dd

In the sequel, we shall see how such a solution can in fact be found for every possible system of loops. A really surprising amount of interesting mathematics seems, at present, to be involved in the problem.

In order to see how a set of satisfactory loop addresses can always be constructed, let us analyze the previous example in more detail. The first column of the solution is:

A—1

B—0

C—1

D—0

E—1

F—0

We see that A, C, and E have the value 1 at this coordinate while B, D, and F have the value 0. Thus, this coordinate will contribute a 1 to the Hamming distance from any of ACE to any of BDF. We may denote

this as $ACE \times BDF$. Therefore, the first column makes the following contribution to the overall distance matrix.

	A	B	C	D	E	F
A	0	1	0	1	0	1
B	1	0	1	0	1	0
C	0	1	0	1	0	1
D	1	0	1	0	1	0
E	0	1	0	1	0	1
F	1	0	1	0	1	0

The second column may be written as $ACF \times BDE$ and contributes the following to the distance matrix.

	A	B	C	D	E	F
A	0	1	0	1	1	0
B	1	0	1	0	0	1
C	0	1	0	1	1	0
D	1	0	1	0	0	1
E	1	0	1	0	0	1
F	0	1	0	1	1	0

The first two columns (i.e., coordinates) then contribute the sum of the previous matrices to the distance matrix.

	A	B	C	D	E	F
A	0	2	0	2	1	1
B	2	0	2	0	1	1
C	0	2	0	2	1	1
D	2	0	2	0	1	1
E	1	1	1	1	0	2
F	1	1	1	1	2	0

The third column is:

$$A\text{—}1$$
$$B\text{—}1$$
$$C\text{—}d$$
$$D\text{—}0$$
$$E\text{—}d$$
$$F\text{—}0$$

It will contribute 1 between A or B and D or F. Since C and E have the third coordinate value d, it cannot contribute to the Hamming distance from C or E to any other point. We can write $AB \times DF$ and obtain the following contribution to the distance matrix:

	A	B	C	D	E	F
A	0	0	0	1	0	1
B	0	0	0	1	0	1
C	0	0	0	0	0	0
D	1	1	0	0	0	0
E	0	0	0	0	0	0
F	1	1	0	0	0	0

The first three columns (coordinates) thus contribute the following to the distance matrix:

	A	B	C	D	E	F
A	0	2	0	3	1	2
B	2	0	2	1	1	2
C	0	2	0	2	1	1
D	3	1	2	0	1	1
E	1	1	1	1	0	2
F	2	2	1	1	2	0

The last two columns are $A \times C$ and $D \times E$ respectively. If the cor-

responding 1's are added to the distance matrix, we obtain the matrix of our example. We see, therefore, that we can think of the distance matrix for our example as generated by the sum of the products $ACE \times BDF$, $ACF \times BDE$, $AB \times DF$, $A \times C$, $D \times E$. Notice that in each product we assign a 0 to each element of one multiplier, a 1 to each element of the other, and a d to any possible multiplier which does not occur. Which set you make 0 and which set you make 1 does not matter. If we carry this out we obtain the coordinates for A through F given previously.

The same mathematics works in general. Take the system of loops for which we wish to find an addressing scheme, and find the abstract graph in which each loop represents a vertex and two vertices are connected if and only if the loops touch. Now write down the (symmetric) distance matrix for this graph. If the vertices of the graph are $A_1, A_2, \cdots, A_n$, and the Hamming distance between A_i and A_j is d_{ij}, then we may take d_{ij} copies of $A_i \times A_j$ and then sum over all i and j. The contribution to the address of each A_k will be d_{ij} coordinates 1 to A_i, 0 to A_j, and d to all other vertices. Therefore, the total contribution to the distance matrix will be d_{ij} in the (i, j) position, and 0 everywhere else. Thus, the resulting complete set of coordinates for the A_k will consist of 0's, 1's, and d's calculated from each necessary copy of each $A_i \times A_j$ and will produce the desired distance matrix.

This proves that the addressing scheme is always possible, but we have used a ridiculously large number of coordinates, perhaps

$$\frac{sn(n-1)}{2}$$

where s is the largest point-to-point distance in the distance matrix. We can save a factor $n/2$ if we take

$$A_1 \times (A_2, A_3, \cdots, A_n) + A_1 \times (A_{i_1}, A_{i_2}, \cdots, A_{i_{k_1}}) + A_1 \times (A_{j_1}, \cdots, A_{j_{k_2}}) + \cdots$$

where $A_{i_1}, \cdots A_{i_{k_1}}$ are all those vertices for which $d_{1,i_t} \geqq 2$, $A_{j_1}, \cdots A_{j_{k_2}}$ are all those vertices for which $d_{1,j_t} \geqq 3$, etc. We then repeat for $A_2 \times (A_3, \cdots A_n) \cdots$, and so on up to $d_{n-1,n}$ copies of $A_{n-1} \times A_n$. This time we have at most $s(n-1)$ products, and therefore have found a set of at most $s(n-1)$ coordinates for the A_i such that loop addressing will work in the desired way. We have proved:

Theorem 1: Given any system of n loops so that the maximum distance between any two loops is s, a system of addresses such that every minimal path between loops is obtained by switching to an adjacent loop if and only if

the Hamming distance to the destination is decreased by 1 can always be found. The length of each address can be taken to be no more than $s(n-1)$.

Let us remark right away that we believe the right answer to be $(n-1)$ rather than $s(n-1)$. We have no proof and we have no counterexamples. We will, however, prove the following theorems in the sequel.

Theorem 2: If the abstract graph of the loop system is the complete graph on n vertices, then addresses of length $(n-1)$ are best possible.

Theorem 3: If the abstract graph of the loop system is a tree on n vertices, addresses of length $(n-1)$ are best possible.

Theorem 4: If the abstract graph of the loop system is a cycle of length n, then addresses of length $n/2$ are best possible if n is even, and addresses of length $(n-1)$ are best possible if n is odd.

II. MATHEMATICAL DEVELOPMENT

Let us summarize what we have proven so far. Let (d_{ij}) be the distance matrix of the abstract graph G with vertices A_i. Let

$$\sum_{\alpha=1}^{N(G)} (A_{i_{\alpha,1}} \cdots A_{i_{\alpha,\mu_\alpha}}) \times (A_{j_{\alpha,1}} \cdots A_{j_{\alpha,\nu_\alpha}}) \tag{1}$$

represent the graph G in the sense that A_i and A_j appear on opposite sides of products exactly d_{ij} times. The number of coordinates which we must assign to each vertex of G is the minimum of $N(G)$ over all decompositions that satisfy the above conditions.

The problem is equivalent to a problem in quadratic forms. Write

$$\sum_{1 \leqq i \leqq j \leqq n} d_{ij} x_i x_j = \sum_{\alpha=1}^{N} (x_{i_{\alpha,1}} + \cdots + x_{i_{\alpha,\mu_\alpha}})(x_{j_{\alpha,1}} + \cdots + x_{j_{\alpha,\nu_\alpha}}). \tag{2}$$

Since $d_{ii} = 0$, no single x_k can appear in both factors of any single product. The equivalence is immediate since either decomposition will immediately yield the other. Our problem then is to find the minimum number N for any given quadratic form whose coefficients d_{ij} are the distance matrix of a graph. We shall prove the following lemma due to H. S. Witsenhausen.

Lemma 1: Let n_+, n_- be respectively the number of strictly positive and strictly negative eignevalues of the distance matrix (d_{ij}). Then

$$N \geqq max\ (n_+, n_-).$$

Proof: Let $Q(x_1, \cdots, x_n)$ denote the quadratic form $\sum_{1 \leq i \leq j \leq n} d_{ij}x_ix_j$. As we have seen, the existence of a length k addressing of G is equivalent to the existence of a decomposition of Q into the sum of k products of the form $(x_{i_1} + \cdots + x_{i_r})(x_{j_1} + \cdots + x_{j_s})$. But we see that

$$Q = \sum_{u=1}^{k} (x_{i_{u,1}} + \cdots + x_{i_{u,r(u)}})(x_{j_{u,1}} + \cdots + x_{j_{u,s(u)}})$$

$$= \frac{1}{4} \sum_{u=1}^{k} \{(x_{i_{u,1}} + \cdots + x_{i_{u,r(u)}} + x_{j_{u,1}} + \cdots + x_{j_{u,s(u)}})^2$$

$$- (x_{i_{u,1}} + \cdots + x_{i_{u,r(u)}} - x_{j_{u,1}} - \cdots - x_{j_{u,s(u)}})^2\}$$

so that we have represented Q as a sum of k squares minus another sum of k squares. However, it is an easy consequence of the theory of quadratic forms (cf. Ref. 2) that for any representation of Q as a sum of p squares minus a sum of q squares, we must have

$$p \geqq \text{index } Q = n_+ ,$$

$$q \geqq \text{rank } Q - \text{index } Q = n_- .$$

Therefore, $k \geqq \max(n_+, n_-)$ and the lemma is proved. For most simple examples equality seems to hold in the above lemma. However, most unfortunately, in general $N \neq \max(n_+, n_-)$. For the graph given in Fig. 7, we have $n_+ = 1$, $n_- = 5$, but a computer search of possible decompositions has shown $N = 6$.

Lemma 1 is strong enough to settle the best N in many cases. In preparation let us prove Lemma 2, due to E. N. Gilbert.

Lemma 2: If the $n \times n$ distance matrix (d_{ij}) is cyclic (meaning $d_{ij} = a(j - i) \bmod n$), then the eigenvalues of (d_{ij}) are the values of

$$P(z) = \sum_{0}^{n-1} a_i z^i$$

at each nth root of unity.

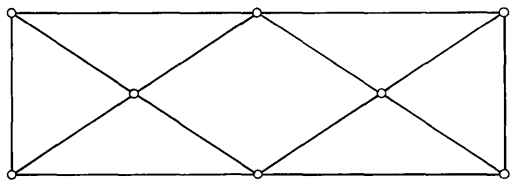


Fig. 7—Graph for Lemma 1.

Proof of Lemma 2: Let the cyclic matrix M be

$$M = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ & & \vdots & \\ a_1 & a_2 & & a_0 \end{pmatrix}.$$

If we try an eigenvector of the form

$$\begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{pmatrix},$$

then

$$M\begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{pmatrix} = \begin{pmatrix} a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} \\ a_{n-1} + a_0 z + \cdots + a_{n-2}z^{n-1} \\ \vdots \\ a_1 + a_2 z + \cdots + a_0 z^{n-1} \end{pmatrix}.$$

If $z^n = 1$, then the latter matrix equals

$$(a_0 + a_1 z + \cdots + a_{n-1}z^{n-1})\begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{pmatrix}.$$

Thus the values of $a_0 + a_1 z + \cdots + a_{n-1}z^{n-1}$ if $z^n = 1$ are eigenvalues of the matrix M. Since they are n in number, and since M has only n eigenvalues, they are *all* the eigenvalues of M.

Theorems 2, 3, and 4 may now be proved by using these lemmas. Let us prove a statement equivalent to:

Theorem 2: If G is the complete graph on n vertices, then $N(G) = n - 1$.*

Proof: For this graph, $d_{ii} = 1$, $1 \leqq i < j \leqq n$. The corresponding quadratic form is

$$\sum_{1 \leqq i < j \leqq n}^{n-1} x_i x_j ,$$

* i.e., any two vertices of G are joined by an edge.

which is equal to

$$\sum_{i=1}^{n-1} x_i(x_{i+1} + \cdots + x_n).$$

Hence $N(G) \leqq n - 1$. To obtain an inequality in the opposite direction, we examine the eigenvalues of the (d_{ij}) matrix. By Lemma 2, they are the values of

$$P(z) = z + z^2 + \cdots + z^{n-1}$$

when $z^n = 1$. But $P(z) = [z(z^{n-1} - 1)/(z - 1)]$, so that if $z^n = 1$ and $z \neq 1$, $P(z) = -1$. If $z = 1$, $P(z) = (n - 1)$. Hence $n_+ = 1$, $n_- = (n - 1)$, and, by Lemma 1, $N(G) \geqq n - 1$. Hence $N(G) = n - 1$.

Theorem 3: If the graph G is a tree with n vertices, then* $N(G) = n - 1$.

Proof: We first examine the distance matrix D_n for a tree with n vertices. Consider a terminal vertex v_i, i.e., a vertex which is distance 1 from just *one* other vertex, say v_j. By a suitable relabeling we can assume $i = n$ and $j = n - 1$.† Thus, $d_{n,k} = 1 + d_{n-1,k}$ for $1 \leqq k \leqq n - 1$. Hence, the matrix D_n has the form

$$D_n = \begin{bmatrix} 0 & d_{12} & \cdots & d_{1n-1} & 1 + d_{1n-1} \\ d_{12} & 0 & \cdots & d_{2n-1} & 1 + d_{2n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ d_{1n-1} & & \cdots & 0 & 1 \\ 1 + d_{1n-1} & & \cdots & 1 & 0 \end{bmatrix}.$$

We wish to evaluate the determinant det (D_n) of D_n . Certainly we can subtract column $n - 1$ from column n and row $n - 1$ from row n of D_n without changing det (D_n). This leaves us with a matrix D_n with the form

$$D_n' = \begin{bmatrix} 0 & \cdots & d_{1n-1} & 1 \\ & & & 1 \\ \vdots & & \vdots & \vdots \\ d_{1n-1} & \cdots & \cdots & 1 \\ 1 & \cdots & 1 & -2 \end{bmatrix}.$$

* i.e., G is connected and has no cycles.

† We have chosen $j = n - 1$ to simplify the exposition of the first part of the proof. In fact, any j, $1 \leqq j \leqq n - 1$, is acceptable. This generality is required later in the proof.

But, we now imagine removing the vertex v_n from G, forming a tree G_{n-1} with $n - 1$ vertices. The interpoint distances in G_n are given exactly by the upper-left $(n - 1)$-by-$(n - 1)$ submatrix of D_n'. As before, we can suitably relabel the vertices of G_{n-1} so that v_{n-1} is a terminal vertex adjacent only to v_{n-2}. The corresponding rearranged matrix D_n'' now has the form

$$D_n'' = \begin{bmatrix} 0 & d_{12} & \cdots & d_{1n-2} & 1 + d_{1n-2} & 1 \\ d_{12} & 0 & \cdots & d_{2n-2} & 1 + d_{2n-2} & 1 \\ \vdots & & & \vdots & \vdots & \vdots \\ d_{1n-2} & & \cdots & 0 & \cdot & 1 \\ 1 + d_{1n-2} & & \cdots & 1 & 0 & 1 \\ 1 & & \cdots & 1 & 1 & -2 \end{bmatrix}.$$

By subtracting column $n - 2$ from column $n - 1$ and row $n - 2$ from row $n - 1$ we obtain

$$D_n''' = \begin{bmatrix} 0 & d_{12} & \cdots & d_{1n-2} & 1 & 1 \\ d_{12} & 0 & \cdots & d_{2n-2} & 1 & 1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ d_{1n-2} & d_{2n-2} & \cdots & 0 & 1 & 1 \\ 1 & 1 & \cdots & 1 & -2 & 0 \\ 1 & 1 & \cdots & 1 & 0 & -2 \end{bmatrix}.$$

It is not difficult to see that this process can be continued until we reach the matrix

$$D_n^* = \begin{bmatrix} 0 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & -2 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & -2 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & -2 \end{bmatrix}.$$

The first (surprising) conclusion we draw is that det (D_n) depends

only on the number of vertices n and not on the *structure* of the tree G. By expanding det (D_n^*) along the last column it is easy to get the recurrence

$$D_n = (-1)^{n-1}2^{n-2} - 2D_{n-1}, \quad D_1 = 0, \quad D_2 = -1$$

from which it follows that

$$D_n = (-1)^{n-1}(n-1)2^{n-2}, \quad n \geqq 1.$$

We next note that if we relabel the vertices of G, according to the relabeling used to get the matrix D_n^*, in the corresponding distance matrix $\overline{D_n}$ (which is a permutation of the original distance matrix D_n) the upper left-hand k-by-k submatrix $\overline{D_k}$ of $\overline{D_n}$ is just the distance matrix for some k vertex subtree of G. Hence,

$$\det(\overline{D_k}) = (-1)^{k-1}(k-1)2^{k-2}, \quad k \geqq 1.$$

Finally, the sequence of determinants

$$1, \det(\overline{D_1}), \det(\overline{D_2}), \cdots, \det(\overline{D_n}) \tag{3}$$

is just

$$1, 0, -1, 4, -12, 32, \cdots, (-1)^{n-1}(n-1)2^{n-2}.$$

Hence, the number of *permanences of sign* of this sequence (where 0 is fixed as either positive or negative) is just *one*! By a theorem in matrix theory (cf. Ref. 2), the number of permanences in sign of the sequence (3) is exactly the number of positive eigenvalues of $\overline{D_n}$ which we have seen is just one. Since $\overline{D_n}$ is nonsingular for $n \geqq 1$, then $\overline{D_n}$ has no *zero* eigenvalues and hence, $\overline{D_n}$ must have $n-1$ negative eigenvalues. Therefore D_n also has $n-1$ negative eigenvalues and by Lemma 1, $N(G) \geqq n-1$.

The construction which gives $N(G) \leqq n-1$ has an easy recursive definition: Each time we choose the next vertex v_i in the tree to assign an address to,* make sure that it is adjacent to a vertex v_j which is already addressed, and let $A(v_i) \rightarrow A(v_j)\,1$ and $A(v_k) \rightarrow A(v_k)\,0$ for the previously addressed vertices (i.e., 1 and 0 are adjoined to the previous addresses). Thus, after all vertices have been addressed, all addresses will have length $n-1$ and, in fact, no d's are used. Therefore, $N(G) = n-1$ and the theorem is proved.

Theorem 4: If G is a cycle on n vertices, then $N(G) = n/2$ if n is even and $(n-1)$ if n is odd.

* Where we assign 0 to the first vertex and 1 to the second vertex.

Proof: If $n = 2m$, then the vertices $A_1, \cdots A_{2m}$ can be coordinatized as follows:

$$A_s = (\underbrace{1 \cdots 1}_{s-1}, \underbrace{0 \cdots 0}_{m-s+1}) \quad \text{if} \quad 1 \leqq s \leqq m+1,$$

and

$$A_s = (\underbrace{0 \cdots 0}_{s-m-1}, \underbrace{1 \cdots 1}_{2m-s+1}) \quad \text{if} \quad m+2 \leqq s \leqq 2m.$$

Clearly $d_{ij} = \min(|i-j|, 2m - |i-j|)$ is the number of places in which A_i and A_j differ, and is the correct distance on a cycle. Hence $N(G) \leqq m$. On the other hand, $d_{1,m+1} = m$, and hence A_1 and A_{m+1} must differ in exactly m coordinates. Therefore there must *be* at least m coordinates, and hence $N(G) \geqq m$. Thus $N(G) = m$.

If $n = 2m + 1$, consider the following addresses:

$$\begin{array}{ll} & \overbrace{\qquad\qquad}^{2m} \\ & \overbrace{\quad}^{m}\ \overbrace{\quad}^{m} \\ A_0 & -000\cdots00\cdots000 \\ A_1 & -d00\cdots00\cdots001 \\ A_2 & -dd0\cdots00\cdots011 \\ & \vdots \\ A_m & -ddd\cdots d1\cdots111 \\ A_{m+1} & -1dd\cdots d1\cdots110 \\ A_{m+2} & -d1d\cdots d1\cdots100 \\ A_{m+3} & -dd1\cdots d1\cdots000 \\ & \vdots \\ A_{2m} & -ddd\cdots10\cdots000 \end{array}$$

We see that:

(*i*) if $0 \leqq i \leqq j \leqq m$, $d_{ij} = j - i$;
(*ii*) if $m < i \leqq j \leqq 2m$, $d_{ij} = j - i$;
(*iii*) if $0 \leqq i \leqq m$ and $j = m + s$ where $s > 0$, then consider separately $i > s$ and $i \leqq s$. If $i > s$ then the first m coordinates contribute 0 and the second m contribute $j - i$. If $i \leqq s$ the

first m coordinates contribute 1 and the second m contribute $i + 2m - j$, so that together they give $2m + 1 - j + i$ which is the correct cyclic distance.

We thus know that $N(G) \leqq 2m$.

To prove $N(G) \geqq 2m$, we use Lemmas 1 and 2.

$$P(z) = z + 2z^2 + \cdots + mz^m + mz^{m+1} + (m-1)z^{m+2} \cdots + z^{2m},$$

and we consider z such that $z^{2m+1} = 1$.

If $z_k = \exp(2\pi ik/2m + 1)$, then

$$P(z_k) = 2\sum_{j=1}^{m} j \cos\frac{2j\pi k}{2m+1}, \qquad k = 0, 1, 2, \cdots, 2m.$$

$P(1) > 0$; we shall prove $P(z_k) < 0$ for all other k. We find that if we define

$$g(x) = \sum_{j=1}^{m} \sin\frac{2j\pi x}{2m+1},$$

then

$$g(x) = \frac{1}{2}\left[\frac{\cos\dfrac{\pi x}{2m+1} - \cos \pi x}{\sin\dfrac{\pi x}{2m+1}}\right].$$

Therefore

$$g'(x) = \frac{\pi}{2}\,\frac{-\dfrac{1}{2m+1} + \sin \pi x \sin\dfrac{\pi x}{2m+1} + \dfrac{1}{2m+1}\cos \pi x \cos\dfrac{\pi x}{2m+1}}{\sin^2\dfrac{\pi x}{2m+1}}$$

is $\pi/(2m+1)$ times the desired series if x is $1, 2, \cdots 2m$. But $g'(x) < 0$ at all of these points. Hence $n_+ = 1$, $n_- = 2m$, and $N(G) \geqq 2m$ by Lemma 1. The theorem is proved.

III. ADDRESSES OF MINIMUM LENGTH

We describe an algorithm which is guaranteed to produce a valid addressing for any graph G. This algorithm has always succeeded in finding an addressing of length $\leqq n - 1$ for every graph G on n vertices to which it has been applied. However, no proof that this will always happen is currently known.

The algorithm proceeds as follows:

(*i*) Number the n vertices of G with integers $\{1, 2, \cdots, n\}$ *so that for* $k > 1$, *the vertex numbered* k *is adjacent to some vertex with a smaller number*. Since G is connected, this is always possible. Let $v(k)$ denote the vertex to which k has been assigned.

(*ii*) Assign the (partial) addresses of 0 to $v(1)$ and 1 to $v(2)$.

(*iii*) In general, suppose we have assigned (partial) addresses to $v(1), v(2), \cdots, v(k)$, say, $A(i)$ has been assigned to $v(i)$, so that $d_{ij} = d_H(A(i), A(j))$, $1 \leqq i < j \leqq k$, where d_H denotes the Hamming distance and d_{ij} denotes the distance between $v(i)$ and $v(j)$ in G. We next search for an address $A(k+1)$ (of the same length as the $A(i)$) with the property that $\max_{1 \leqq i \leqq k} (d_{i,k+1} - d_H(A(i), A(k+1))) \equiv m_{k+1}$ is as small as possible under the constraint

$$\min_{1 \leqq i \leqq k} (d_{i,k+1} - d_H(A(i), A(k+1))) \geqq 0. \qquad (*)$$

Of course, we can always find *some* address which satisfies (*), namely the all d's address. Typically we can choose $A(k+1)$ so that $m_{k+1} = 1$. In fact, it is usually possible to do this by choosing $A(k+1)$ to be a slightly perturbed copy of some $A(l)$ where $v(l)$ is adjacent to $v(k+1)$. This is intuitively reasonable since in this case $|d_{i,k+1} - d_{i,l}| \leqq 1$.

After $A(k+1)$ has been chosen, we then *adjoin* m_{k+1} symbols to each of the partial addresses $A(i)$, $1 \leqq i \leqq k+1$, as follows. To $A(k+1)$ we adjoin m_{k+1} 1's. To $A(i)$ we adjoin $m_{k+1} - (d_{i,k+1} - d_H(A(i), A(k+1)))$ d's and $d_{i,k+1} - d_H(A(i), A(k+1))$ 0's. It is easy to check that for the new augmented addresses $A'(i)$, $1 \leqq i \leqq k+1$, we have

$$d_{i,j} = d_H(A'(i), A'(j)), \quad 1 \leqq i < j \leqq k+1.$$

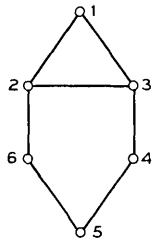
We continue in this manner until the addressing is completed. By construction, the terminal addresses will form a valid addressing for G of length $1 + m_3 + \cdots + m_n$.

As an example, we construct an addressing for the graph in Section I by this process. In Fig. 8 we show this graph with a particular "adjacent-numbering" chosen and also the distance matrix for the graph.
We start with

vertex address

1—0
2—1

Adjoining vertex 3, we see that any partial address of length one will give $m_3 = 1$. We choose 0.



	1	2	3	4	5	6
1	0	1	1	2	3	2
2	1	0	1	2	2	1
3	1	1	0	1	2	2
4	2	2	1	0	1	2
5	3	2	2	1	0	1
6	2	1	2	2	1	0

$D = (d_{ij})$

Fig. 8—Addressing example.

vertex address

1—0
2—1
3—0

We next adjoin $m_3 = 1$ 1's to $A(3)$ and augment $A(1)$ and $A(2)$ accordingly.

vertex address

1—00
2—1*d*
3—01

Now adjoin vertex 4, choose partial address 01, calculate that $m_4 = 1$, and augment the partial addresses accordingly.

vertex address

1—000
2—1*d*0
3—010
4—011

Continue this for two more steps. Each time $m_k = 1$.

vertex address	*vertex address*
1—0000	1—0000*d*
2—1*d*0*d*	2—1*d*0*dd*
3—0100	3—0100*d*
4—0110	4—01100
5—0111	5—0111*d*
	6—1*d*1*d*1

The last array gives a length 5 addressing for G. Of course, different partial addresses or a different initial vertex numbering will result in different addressings for G.

As we have previously stated, we have no general proof that $N(G) \leqq n - 1$ in all cases although a number of partial results in that direction have been given as well as a heuristic construction.

IV. ADDRESSING IN RESTRICTED LOOP SYSTEMS

The addressing scheme we have been describing has the very great power of being able to handle an arbitrary configuration of loops, and to provide alternate routing in an optimal way without any supervisory memory. The price we have paid for this generality is in the length of the address—typically $n - 1$ "bits" for n loops in the simplest encoding—and in possible complications under system growth. It is clear that if a new loop is added which greatly shortens the distance between many pairs of loops, then many addresses may change a good deal. There would be various ways of handling this, but it is obviously a problem. It arises essentially because the numbering in its full generality is not hierarchical.

Typical Bell System loop configurations, as we noted in the introduction, will not be arbitrary collections of loops, but will have a hierarchical structure.

By correspondingly restricting the allowable adjacency graphs G, it is possible to modify the routing algorithm and effectively take advantage of a natural "product" construction, as pointed out by J. R. Pierce.[1] In this system, as we saw, loops are partitioned into three classes—national, regional, and local. The address portion of the message is subdivided into three corresponding portions. The routing algorithm now consists of three steps: (*i*) First apply the previous Hamming distance algorithm to the "national" portions of the sending and the destination addresses; (*ii*) When the distance in i becomes zero, then apply the Hamming distance algorithm to the "regional" portions of the addresses; (*iii*) Finally, when the distance in *ii* is zero, apply the Hamming distance algorithm to the "local" portions of the address.

This scheme combines the efficiency of the Hamming distance algorithm with the savings in address lengths resulting from the hierarchical structure. As an example, the network in Fig. 9 has 44 local vertices. For a direct Hamming algorithm addressing we should expect addresses to have length of around 59. By distinguishing national, regional, and local loops (capital letters, lower case letters, and integers

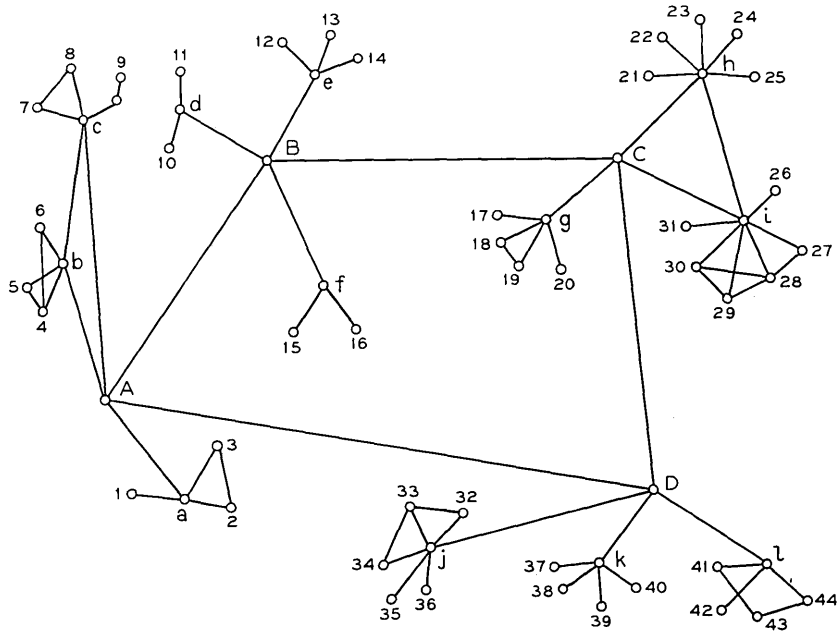


Fig. 9—Network example.

respectively), with a small additional computed cost in routing (several extra conditional transfers) we can have addresses of length $\leqq 11$. For example, let N, R, L denote national, regional, local, respectively. One possible addressing begins:

	N	R	L
A—	(00,	000	—)
B—	(01,	000,	—)
C—	(10,	000,	—)
D—	(11,	000,	—)
a—	(00,	001,	000)
b—	(00,	010,	000)
c—	(00,	1d0,	000)
d—	(01,	001,	000)
⋮		⋮	

1—(01, 001, 001)

2—(01, 001, 010)

3—(01, 001, 1d0)

⋮ ⋮

44—(11, 000, 001)

Moreover, to add additional local stations to a regional station it is a very simple matter to modify just the *neighboring* local addresses to obtain a correct addressing for the augmented network.

The restriction on local loops in the above addressing is that each one must interchange directly with one and only one regional loop. If a local loop meets no regional loop directly, but only other local loops, then the addressing must make special provision for routing calls to other regions properly. If a local loop meets more than one regional loop—really a violation of the hierarchical concept—then routing becomes more difficult, and must assure that a call to a different region exits the local loop properly. As J. R. Pierce has pointed out,[1] a special trunk loop connecting a local loop in one region to a local loop in another (i.e., a preferred alternate route in a special case to the national loop) is no problem. The exit from the local loop is just before the regional interchange, and the entrance to the local loop just after. Exit is made only if the total loop address matches exactly. Alternate routes more generally are perhaps most easily provided by duplicating portions of regional or natural loops.

V. SOME VARIANTS OF THE ADDRESSING PROBLEM

The purpose of this section is to record very briefly some other alternatives that have been considered.

(*i*) We have required that in every alternate route between the loops, Hamming distance decrease by exactly 1 at each transfer. One could consider the alternate problem in which *any* exit which decreases Hamming distance is valid—even if it decreases it by *more* than 1. Under special conditions, this can lead to shorter addresses, but we do not have a solution for this alternate problem.

(*ii*) Since the introduction of d's causes some complication of the address codes, it is interesting to consider the possibility of getting rid of them. They arose originally because of the need for *odd* cycles, as in the case of a 3-cycle. One way out of this example would be to *double* all the

distances. If these vertices were located at 000, 110, and 101 respectively, the Hamming distance between any pair is two, and correct routing would be possible without any d's in the addresses.

Unfortunately, this technique of doubling all the distances to get rid of d's does not generalize. Consider the graph in Fig. 10a. We double all

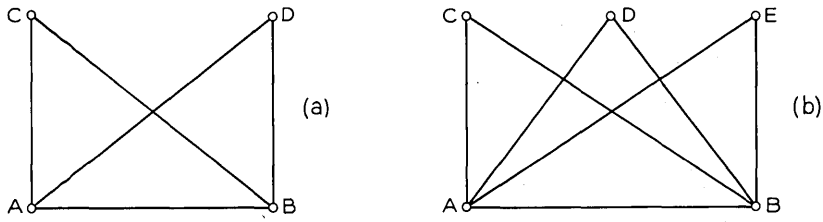


Fig. 10—Graphs to illustrate one variant of the addressing problem.

distances, so that $AB = AC = AD = BC = BD = 2$, $CD = 4$. Then $A = 00 \cdots$, $B = 11 \cdots$, where the coordinates are identical from the third onward. Now C must differ from each of A and B by 2. It therefore must differ in one of the first two columns, and in one other, say the third. Thus, we may assume $A = 000 \cdots$, $B = 110 \cdots$, $C = 101 \cdots$, where the coordinates are identical from the fourth onward. D must also differ in exactly two places from A and B and in four places from C. Hence $A = 0000 \cdots$, $B = 1100 \cdots$, $C = 1010 \cdots$, $D = 0101 \cdots$.

So far so good. If we now require yet another point E (Fig. 10b) such that $EA = EB = 2$, $EC = ED = 4$, we have no possible coordinates for E left. E's address must begin with 01 or 10 in order to differ from A and B by equal amounts, say with 10. To differ from C and D by equal amounts the first four coordinates must be 1001. But it now differs from A and B by 2 and from C and D by 2; no additional coordinates can make $EC = ED = 4$ without destroying $EA = EB = 2$. Thus doubling distances will not get rid of d's.

Similar arguments show that even if we are allowed to multiply all distances by a fixed number $m > 2$, we still cannot get along without d's.

REFERENCES

1. Pierce, J. R., "Network for Block Switching of Data," unpublished work.
2. Jones, B. W., "The Arithmetic Theory of Quadratic Forms," CARUS Mathematical Monograph No. 10, published by MAA, 1950.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Heavy Traffic Characteristics of a Circular Data Network

By B. AVI-ITZHAK

(Manuscript received April 26, 1971)

Traffic behavior in the Pierce loop for data transmission is studied under assumptions of heavy loading. A deterministic mathematical model for describing traffic flows is developed and analyzed. The mathematical problem is of a linear complementarity type which has not been dealt with in the literature of mathematical programming. An effective procedure, the load-and-shift algorithm, for determining traffic flows is proposed. The procedure yields all feasible solutions for traffic flows and reveals the possibility of stations grouping into dominating classes and preventing other stations from using the system. This property, which can be eliminated by exercising appropriate control, also may affect the stochastic behavior of the system when heavy traffic conditions do not prevail and therefore deserves careful investigation. The paper includes two numerical examples illustrating use of the load-and-shift algorithm and numerical results from a simulation showing some of the effects of dominating classes when heavy traffic conditions do not prevail.

I. INTRODUCTION

The concept of a loop network for data transmission has been proposed recently by J. R. Pierce.[1] In such a network the stations are connected to a closed loop main line on which one-way traffic is allowed. A message to be delivered from one station to another is arranged, at the sending station, into standard packets each carrying the address of the receiving station. These packets are then delivered onto the main line, one at a time, where they flow around in the allowed traffic direction. The address of each packet is checked at each station on the way until it reaches the receiving station where it is removed from the main line. Traffic on the main line cannot be delayed; therefore, a station can deliver a packet onto the main circular line only when permitted by the existence of a gap in traffic or when receiving a packet from the main line. Principal

features of the system may be explained with the aid of the four-station network shown schematically in Fig. 1.

The four-armed structure revolves around the central axis and stops briefly every time the four packet-carrying compartments at the ends of the arms are aligned with the four stations. During such a stop each station is able to check the content of the aligned compartment. If the compartment is empty, the station can load it with a packet. If there is a packet in the compartment, it will not be removed unless it is addressed to the said station, in which case the station is permitted to load the compartment again after unloading it. A Pierce loop can be represented by the mechanical analog structure shown in Fig. 1; however, the number of revolving arms in the structure is not necessarily equal to the number of stations in the loop. Rather, the number of arms is determined by the "loop time" of the system (the time needed for a bit to complete one round on the loop). Significance of the loop time is discussed in more detail in Section V.

In the Pierce loop it is currently assumed that no outside control is applied and each station strives to send its messages at the earliest. Therefore, a station will never miss the opportunity to load a compartment unless there are no messages waiting for delivery at the station.

This paper presents a study of the flow characteristics of such a system in heavy traffic, i.e., when the system is not able to deliver all messages, and infinite queues build up at some stations. In Section II the notion of "stable solutions" for the traffic flows is introduced and

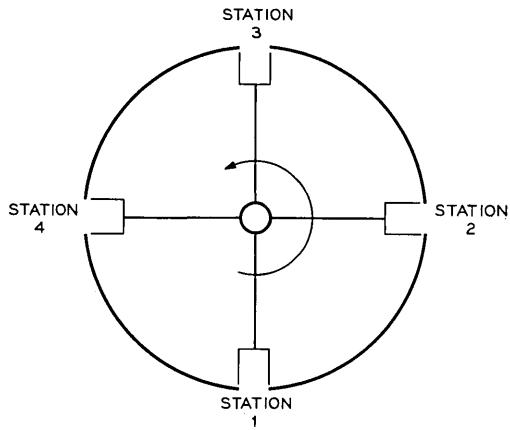


Fig. 1—A schematic description of a four-station Pierce loop.

formulated mathematically. The case of a totally saturated system is analyzed in Section III where the basis is laid for the mathematical development of the load-and-shift algorithm given in Section IV. A detailed description of a Pierce loop for data transmission is given in Section V where the alternating priorities effect due to dominating classes of stations is studied by simulation. This section also discusses the important aspect of the order of stations in the loop. Readers not interested in the mathematical elaborations may skip Section IV.

II. MATHEMATICAL FORMULATION

We assume that the flow direction on the main circular line is counterclockwise. There are n stations, connected to the main line, numbered from 1 to n in counterclockwise increasing order. The segment of main line between the ith and the $(i + 1)$th station, $i = 1, 2, \cdots, n - 1$, is called the ith branch. Similarly, the nth branch is the segment between station n and station 1. Let p_{ij} be the proportion of flow (packets) emerging from station i and destined for station j.

$$\sum_{j=1}^{n} p_{ij} = 1. \tag{1}$$

The $n \times n$ square matrix $P = \{p_{ij}\}$ possesses all the properties of a stochastic matrix. Note, however, that the elements of P are not necessarily probabilities.

The demand at station i is given by λ_i which is the average amount of flow (packets per time unit) generated at the station. The capacity of each branch equals 1, that is, each branch is capable of carrying a maximum flow of one unit. In the schematic description given in Fig. 1 assume that one full revolution takes four time units (in general it will take the number of time units equal to the number of arms) The capacity of each branch will then be one packet per time unit.

The average flow emerging from station i will be denoted by x_i and the average flow in branch i by ρ_i. Clearly

$$0 \leqq x_i \leqq \lambda_i, \quad i = 1, 2, \cdots, n \tag{2}$$

and

$$0 \leqq \rho_i \leqq 1, \quad i = 1, 2, \cdots, n. \tag{3}$$

Let a_{ij} be the proportion of flow emerging from station j and flowing through branch i, then

$$a_{ij} = \sum_{k \in S_{ij}} p_{jk}, \tag{4}$$

where

$$S_{ij} = \begin{cases} (i+1, i+2, \cdots, j) & \text{if } j > i, \\ (i+1, i+2, \cdots, n, 1, 2, \cdots, j) & \text{if } j \leqq i. \end{cases} \tag{5}$$

Note that $a_{i-1j} - a_{ij} = p_{ji} \geqq 0$ for all $j \neq i$ where $i-1$ is defined to equal n when $i = 1$. For $j = i$ we have $a_{ii} = 1$ and $a_{i-1i} = p_{ii} \leqq 1$. In most reasonable applications $p_{ii} = 0$.

The average flow in branch i may now be expressed as a linear function of $X = (x_1, x_2, \cdots, x_n)$.

$$\rho_i = \sum_{j=1}^{n} a_{ij} x_j \leqq 1. \tag{6}$$

Every X which is a feasible solution for the average flows must satisfy relations (2) and (6). The set of all feasible solutions is therefore contained in a convex polyhedral set. A central control could select a particular solution from this set to suit a given objective.

In the circular network suggested by Pierce, however, there is no central control. Rather each station is striving to maximize its own flow onto the main line. For this case we define a stable solution as a solution from which the system will not depart without outside intervention. Suppose then that X^* is a stable solution. Clearly X^* must satisfy relations (2) and (6) and the additional condition that if $\rho_i = \sum_{j=1}^{n} a_{ij} x_j^* < 1$ then $x_i^* = \lambda_i$. To show that this is a necessary condition assume that $\rho_i < 1$ and $x_i^* < \lambda_i$. However, station i strives to maximize its flow and can increase it as long as $x_i^* < \lambda_i$ and $\rho_i < 1$. Therefore, it is not possible that $\rho_i < 1$ and $x_i^* < \lambda_i$. This additional condition is not generally sufficient for assuring that x^* is a stable solution.

Insufficiency is best demonstrated by a simple numerical example. Suppose $n = 4$, $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 2$ and

$$P = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix}.$$

It is easy to verify that $X^0 = (0, 0, 1, 1)$ and $X^+ = (1, 1, 0, 0)$ both satisfy relations (2), (6), and the additional necessary condition as does $X = \alpha X^0 + (1 - \alpha) X^+$ for $0 \leqq \alpha \leqq 1$. For simplicity we assume

that the flows generated at the stations are deterministic in nature and each station generates exactly two packets each time unit.

Suppose now that the system (see Fig. 1) has one rotating arm only. In such a case there are only two stable solutions, namely X^0, ($\alpha = 1$) and X^+, ($\alpha = 0$). If the system has two rotating arms, an additional stable solution, $\alpha = \frac{1}{2}$, is added. If the system has k arms, there exist $k + 1$ stable solutions, $\alpha = m/k$, $m = 0, 1, \cdots, k$. The system shall settle for the stable solution $\alpha = m/k$ if at time zero m compartments contain packets from stations 3 and 4 and $k - m$ compartments contain packets from stations 1 and 2. The necessary condition for stability is also a sufficient condition if the flows generated by the stations are continuous (as in the case of nonmixable fluids or small particles such as vehicles and a loop consisting of a pipe or a road). In such a case the set of feasible solutions is identical to the convex polyhedral set given by relations (2) and (6).

We wish to find all solutions which satisfy relations (2), (6), and the necessary stability condition (these will be all the stable solutions in the case of a system with continuous flows). Our problem can be redefined as one of finding all feasible solutions to the following set of equations:

$$x_i + u_i = \lambda_i ,$$

$$\sum_{j=1}^{n} a_{ij}x_j + z_i = 1, \qquad i = 1, 2, \cdots, n. \tag{7}$$

$$u_i z_i = 0, \qquad x_i \geqq 0, \qquad u_i \geqq 0, \qquad z_i \geqq 0.$$

This form resembles a linear complementarity problem,[2] where u_i and z_i are slack variables. A feasible solution to the set of equations (7) is a basic feasible solution to the set of equations (2) and (6) since at least n of the variables must equal zero.

III. COMPLETE SATURATION

For sufficiently large values of λ_i (for example $\lambda_i = 1, i = 1, 2, \cdots, n$), all the branches are saturated and $\rho_i = 1$, $i = 1, 2, \cdots, n$. The set of equations (7) takes the form

$$\sum_{j=1}^{n} a_{ij}x_j = 1, \qquad i = 1, 2, \cdots, n. \tag{8}$$

This can be shown to be equivalent to the set of $(n + 1)$ linear equations.

$$x_i - \sum_{j=1}^{n} p_{ji} x_j = 0, \qquad i = 1, 2, \cdots, n, \tag{9a}$$

$$\sum_{j=1}^{n} x_j d_j = b. \tag{9b}$$

The ith equation in (9a) is obtainable by subtracting the ith equation of set (8) from the $(i + 1)$th equation of set (8). Equation (9b) may be selected as any linear combination of equations (8). Note that one equation in (9a) is redundent since $\sum_{j=1}^{n} p_{ij} = 1$ for all i.

Remark: The physical interpretation of equations (9) is that when in complete saturation a station is able to deliver only when receiving. The flow emerging from a given station must equal the flow entering the station from the main line. This equilibrium relation is expressed by the n equations of set (9a). The $(n + 1)$th equation, (9b), expresses the capacity limitation of the branches. If we select $d_j = \sum_{i=1}^{n} a_{ij}$ we have

$$\sum_{j=1}^{n} x_j d_j = n, \tag{10}$$

where d_j is the average number of branches (distance) traveled by a packet emerging from station j. Equation (10) states that the average work (in terms of packets times distance) demanded from the system per time unit must equal n, since all branches are saturated and each traverses one packet per time unit. In matrix form we have

$$X^T P = X^T,$$

$$\sum_{i=1}^{n} x_i d_i = b, \tag{11}$$

where X^T is the transpose of X, (note that all vectors are defined to be column vectors).

P is a stochastic matrix and, therefore, the problem represented by equations (11) strongly resembles one of determining the steady state probabilities of a finite state space Markov chain. The difference is that in the Markov chain problem $d_i = b$ for all j while in our problem this is not necessarily so. In the following we shall make use of this resemblance.

Definitions:

(*i*) S_i shall be used to abbreviate "station i."

(*ii*) S_i is said to be accessible from S_i , $S_i \rightarrow S_i$, if there exists a

sequence of elements of P such that $P_{ik_1}P_{k_1k_2} \cdots P_{k_{r-1}k_r}P_{k_ri} > 0$. In such a case it is also customary to say that S_i leads to S_i .

(*iii*) If $S_i \to S_j$ and $S_j \to S_i$ both stations are said to communicate ($S_i \leftrightarrow S_j$). Clearly if $S_i \leftrightarrow S_j$ and $S_k \leftrightarrow S_j$ then $S_i \leftrightarrow S_k$.

(*iv*) Let $C(i) = \{S_j : S_j \leftrightarrow S_i\}$. Clearly if $S_j \in C(i)$ then $C(j) = C(i)$.

(*v*) A nonempty class of stations, C, is called a communicating class if for some station $S_i \in C$, $C = C(i)$. It follows that two communicating classes are either identical or disjoint.

(*vi*) A communicating class is closed if no station outside the class is accessible from a station in the class.

(*vii*) A closed communicating class shall be called a dominating class or a class of dominating stations. A station not belonging to any dominating class is a dominated station.

From the theory of Markov chains we know that there exists at least one class of dominating stations in a given Pierce loop. However, it is possible that there will be no dominated stations. In such a case all the stations are dominating and may form into one or more dominating classes. It may be shown that the number of dominating classes is one if and only if there exists a station accessible from all other stations in the loop. In the case that dominated stations exist each must lead to at least one dominating station.

Each dominating class of stations is represented by a principal submatrix of P (a dominating submatrix). This submatrix is obtainable by deleting all rows and columns of P corresponding to stations not belonging to the particular dominating class. Similarly, all dominated stations may be represented by one submatrix of P.

Theorem 1. Let B be a $k \times k$ submatrix of P representing a dominating class of stations, then all vectors Y satisfying the equation,

$$Y^TB = Y^T \tag{11a}$$

form a linear space of one dimension. Furthermore, all the elements of Y must have the same sign, i.e., either all positive, or all negative, or all zero.

Proof. Theory of finite Markov chains (e.g., Kemeny and Snell[3]).

Corollary. Assume that the dominating class represented by B is $C = \{S_{i_1}, S_{i_2}, \cdots, S_{i_k}\}$ then there exists a unique vector $Y^* = (y^*_{i_1}, y^*_{i_2}, \cdots, y^*_{i_k})$ satisfying equation (11a) and the scaling equation

$$\sum_{j=1}^{k} y^*_{i_j} d_{i_j} = b. \tag{12}$$

Clearly $Y^* > 0$ and any solution to equation (11a) may be obtained by multiplying Y^* by some real number. We define a vector $X = (x_1, x_2, \cdots, x_n)$ where $x_i = 0$ if $S_i \notin C$, and $x_i = y_{i_i}^*$ if $i = i_i$. The vector X is called a dominating solution to equations (11) corresponding to the matrix B and the class C.

Theorem 2. Let Q be a $k \times k$ submatrix of P representing dominated stations, then the only solution to $Y^TQ = Y^T$ is $Y = 0$.

Proof. Theory of finite Markov chains.

Theorem 3. Suppose that the matrix P contains exactly m dominating submatrices $B_1, B_2, \cdots, B_m$ representing dominating classes $C_1, C_2, \cdots, C_m$ and suppose that X_i is the dominating solution corresponding to B_i, $i = 1, 2, \cdots, m$. X is a solution to equations (8) if and only if

$$X = \sum_{i=1}^{m} \alpha_i X_i ,$$

and

$$\sum_{i=1}^{m} \alpha_i = 1, \tag{13}$$

where α_i, $i = 1, 2, \cdots, m$ are real numbers.

Proof. This is an immediate result of Theorems 1 and 2.

We conclude that in the case of complete saturation, ($\rho_i = 1$, $i = 1, 2, \cdots, n$), there always exists a nonnegative solution to our problem. If there exists only one dominating class of stations in the loop, there exists a unique solution to equations (8). If there is more than one dominating class in the loop, then there exist infinitely many nonnegative solutions obtainable as convex combinations of the dominating solutions.

IV. THE LOAD-AND-SHIFT ALGORITHM

Returning to the more general case, we describe in this section the load-and-shift procedure for solving the set of equations (7) when $\rho_i \leqq 1$, $i = 1, 2, \cdots, n$. The algorithm is based on the results obtained in the foregoing analysis. Assuming that the capacity of the branches, ϵ, may be varied between 0 and 1, we start with $\epsilon = 0$ and increase it until $\epsilon = 1$. While so doing we simultaneously load the system and obtain the feasible solutions for any given value of $0 \leqq \epsilon \leqq 1$.

We shall start by outlining the procedure for finding just one solution.

Start: $r := 0$, $\epsilon := 0$, $P^{(1)} := P$, $\lambda_i^{(1)} := \lambda_i$.

Step 1: Increase r by 1. In $P^{(r)}$ find a square submatrix $B^{(r)}$ representing a dominating class of stations $C^{(r)}$, $C^{(r)} = \{S_{i_1}, S_{i_2}, \cdots, S_{i_k}\}$.†

Step 2: Find the unique positive vector $Y^{(r)} = \{y_{i_1}^{(r)}, y_{i_2}^{(r)}, \cdots, y_{i_k}^{(r)}\}$ satisfying the set of equations

$$(Y^{(r)})^T B^{(r)} = (Y^{(r)})^T,$$

$$\sum_{m=1}^{k} y_{i_m}^{(r)} d_{i_m} = 1.$$

To determine the value of d_{i_m} we select an i such that $S_i \in C^{(r)}$ and let $d_{i_m} = a_{i,i_m}$.

Enlarge $Y^{(r)}$ to the form $X^{(r)} = (x_1^{(r)}, x_2^{(r)}, \cdots, x_n^{(r)})$, where $x_i^{(r)} = 0$ if $S_i \notin C^{(r)}$ and $x_i^{(r)} = y_{i_m}^{(r)}$ if $j = i_m$.

Step 3 (Load): Find a number $\Delta^{(r)}$ such that

$$\Delta^{(r)} = \operatorname{Min}\left\{\frac{\lambda_i^{(r)}}{x_i^{(r)}}\right\} = \frac{\lambda_j^{(r)}}{x_j^{(r)}},$$

where the minimization is over all i such that $S_i \in C^{(r)}$.

If $\Delta^{(r)} \geqq 1 - \epsilon$, set $\Delta^{(r)} = 1 - \epsilon$, set $N = r$, and go to "Last Step." Otherwise increase ϵ by $\Delta^{(r)}$ and continue.

Step 4 (Shift): If $r = n$, set $N = n$ and go to "Last Step." Otherwise construct the $(n - r + 1) \times (n - r + 1)$ square matrix $P^{(r+1)}$ by adding the jth column of $P^{(r)}$ to its $(j + 1)$th column (if j is the last column, it is added to the first one) and then deleting the jth column and the jth row. $\lambda_i^{(r+1)} := \lambda_i^{(r)} - \Delta^{(r)} x_i^{(r)}$, $i = 1, 2, \cdots, n$. End of rth iteration. Go back to "Step 1."

Last Step: The solution is

$$X = \sum_{r=1}^{N} \Delta^{(r)} X^{(r)}. \tag{14}$$

STOP.

Theorem 4: The procedure described above will always yield a vector $0 \leqq X \leqq \lambda$ *in at most n iterations.*

Proof: All possible matrices $P^{(r)}$ are stochastic. Therefore there always exists at least one dominating submatrix of $P^{(r)}$ denoted by $B^{(r)}$. Since $X^{(r)} \geqq 0$ and $\Delta^{(r)} \geqq 0$ then $X = \sum_{r=1}^{N} \Delta^{(r)} X^{(r)} \geqq 0$. From the

† A procedure for determining $B^{(r)}$ is described in the Appendix.

algorithm we have $\lambda = \sum_{r=1}^{N} \Delta^{(r)} X^{(r)} + \lambda^{(N)} = X + \lambda^{(N)}$. The algorithm also ensures that $\lambda^{(r)} \geqq 0$, $r = 1, 2, \cdots, N$, and therefore $X \leqq \lambda$. Since P is $n \times n$ the number of iterations cannot exceed n.

Theorem 5 (Existence): X obtained by the Load-and-Shift procedure is a feasible solution to equations (7).

Proof: We enlarge $Y^{(r)}$ by adding zero elements corresponding to columns of $P^{(r)}$ not included in $B^{(r)}$. The enlarged vector, denoted by $Y^{0(r)}$ is a dominating solution to

$$Y^T P^{(r)} = Y^T$$

$$\sum_{j \in R_r} y_j d_j = 1, \tag{15}$$

where R_r is the set of indices of columns included in $P^{(r)}$. These equations are equivalent [see equations (8) and (9)] to

$$\sum_{j \in R_r} a_{ij}^{(r)} y_j = 1, \qquad i \in R_r . \tag{16}$$

It is easily verified, by the use of equation (4), that

$$a_{ij}^{(r)} = a_{ij} . \tag{17}$$

The vector $X^{(r)}$ is obtained by adding to $Y^{0(r)}$ zero elements corresponding to columns of P not included in $P^{(r)}$. Therefore

$$\Delta^{(r)} \sum_{j=1}^{n} a_{ij} x_i^{(r)} = \Delta^{(r)}, \quad \text{for all} \quad i \in R_r . \tag{18}$$

From the definition of a_{ij}, equations (4) and (5), we know that $a_{i-1j} \geqq a_{ij}$ for all j except $j = i$, (where $i - 1 = n$ if $i = 1$). If $i \notin R_r$ then $x_i^{(r)} = 0$. It follows that

$$\Delta^{(r)} \sum_{j=1}^{n} a_{ij} x_i^{(r)} \leqq \Delta^{(r)}, \quad \text{for all} \quad i \notin R_r . \tag{19}$$

Summing equations (18) and (19) with respect to r and then substituting equation (14) we obtain (for $N < n$)

$$\sum_{j=1}^{n} a_{ij} x_j = \sum_{r=1}^{N} \Delta^{(r)} = 1, \quad \text{if} \quad i \in R_r \quad \text{for all} \quad r = 1, 2, \cdots, N,$$

$$\sum_{j=1}^{n} a_{ij} x_j \leqq \sum_{r=1}^{N} \Delta^{(r)} = 1, \quad \text{otherwise.} \tag{20}$$

From the algorithm we know that if $i \in R_r$ for all $r = 1, 2, \cdots, N$ then $x_i \leqq \lambda_i$. Otherwise $x_i = \lambda_i$. This completes the proof for $N < n$.

In the case $N = n$ it is possible that $\sum_{r=1}^{N} \Delta^{(r)} < 1$ (nonheavy traffic). The theorem still holds since $x_i = \lambda_i$, $i = 1, 2, \cdots, n$. In the following we illustrate the use of the algorithm in solving for the flows in a 5-station loop:

Numerical Example 1

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 2/3 & 0 & 1/3 & 0 & 0 \\ 1/4 & 3/4 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 0 & 1/4 \\ 1/4 & 0 & 1/2 & 1/4 & 0 \end{bmatrix} \qquad \lambda = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.5 \\ 1.0 \\ 0.8 \end{bmatrix}$$

Table of Results

r	$C^{(r)}$-dominating Class	$x_1^{(r)}$	$x_2^{(r)}$	$x_3^{(r)}$	$x_4^{(r)}$	$x_5^{(r)}$	$\Delta^{(r)}$	ϵ
1	S_1, S_2, S_3.	0.6000	0.7000	0.5444	0	0	0.5714	0.5714
2	S_1, S_3.	0.2500	0	1.0000	0	0	0.1953	0.7667
3	S_1, S_4, S_5.	0.3123	0	0	1.0000	0.2500	0.2333	1.0000

The procedure terminated in three iterations yielding a solution:

$$x_1 = 0.4646$$

$$x_2 = 0.4000 = \lambda_2$$

$$x_3 = 0.5000 = \lambda_3$$

$$x_4 = 0.2333$$

$$x_5 = 0.0508$$

Actually the algorithm finds the value of one variable in each iteration. When all variables equal to their respective λ_i's have been determined, the algorithm finds in one iteration the values of all the remaining variables. It is therefore advisable to test the possible solution $X = \lambda$ beforehand.

We note that in the example there is a single dominating class of stations in each iteration (i.e., each matrix $P^{(r)}$ has only one dominating submatrix $B^{(r)}$). It will be shown later that if P has only one dominating submatrix then there exists a unique solution to equations (7). In the general case, however, $P^{(r)}$ may have more than one dominating

submatrix. This may give rise to the existence of infinitely many solutions. We would now like to improve the algorithm in order to be able to determine all feasible solutions.

Theorem 6: If $P^{(r)}$ has exactly one dominating submatrix, say $B^{(r)}$, then $P^{(r+1)}$ will also have exactly one dominating submatrix.

Proof: Suppose i is such that $S_i \in C^{(r)}$, then $S_j \rightarrow S_i$ for all j. We apply the shift operation from i to k, (adding column i to column k and then deleting row and column i), thus creating $P^{(r+1)}$. Clearly now $S_j \rightarrow S_k$ for all j. However, if there exists a station accessible from all other stations there exists exactly one dominating class in the loop. Furthermore $C^{(r+1)} = C(k)$.

Corollary: If P has one dominating submatrix so will $P^{(r)}$, $r = 1, 2, \cdots, N$.

Theorem 7: If $P^{(r)}$ has $m > 1$ dominating submatrices, $P^{(r+1)}$ will have either $(m - 1)$ or m dominating submatrices.

Proof (outline): Let $B_1^{(r)}, B_2^{(r)}, \cdots, B_m^{(r)}$ be the m dominating submatrices of $P^{(r)}$ and let $C_1^{(r)}, C_2^{(r)}, \cdots, C_m^{(r)}$ be the corresponding classes of dominating stations. $Q^{(r)}$ is the submatrix representing dominated stations. We apply the shift operation from i to k to obtain $P^{(r+1)}$. Without loss of generality we assume that $S_i \in C_1^{(r)}$. For S_k one of three alternatives must be true:

(*i*) $S_k \in C_1^{(r)}$. In this case a dominating class will be formed, containing some or all the remaining stations of $C_1^{(r)}$. Note that S_k must be in the newly formed dominating class since it is accessible from all remaining stations of $C_1^{(r)}$. Those stations of $C_1^{(r)}$ which are not included in the newly formed dominating class turn into dominated stations. The matrix $P^{(r+1)}$ will then have m dominating submatrices, namely $B_2^{(r)}, B_3^{(r)}, \cdots, B_m^{(r)}$ and a newly formed one.

(*ii*) $S_k \in C_l^{(r)}$, $l \neq 1$. In this case every station in $C_l^{(r)}$ is accessible from any remaining station of $C_1^{(r)}$. Therefore all remaining stations of $C_1^{(r)}$ become nondominating, and $P^{(r+1)}$ will have $m - 1$ dominating matrices, namely $B_2^{(r)}, B_3^{(r)}, \cdots, B_m^{(r)}$.

(*iii*) S_k is nondominating. S_k must lead to at least one dominating class. If it leads to any dominating class other than $B_1^{(r)}$ then this case becomes the same as (*ii*). If S_k leads to $B_1^{(r)}$ only, then a new dominating class is formed. This class includes all stations accessible from S_k. Remaining stations of $B_1^{(r)}$ which are not accessible from S_k

become nondominating. $P^{(r+1)}$ will have m dominating submatrices, namely $B_2^{(r)}$, $B_3^{(r)}$, $\cdots$, $B_m^{(r)}$, and a newly formed one.

This completes the outline of the proof. It is important to note that a dominating submatrix of $P^{(r)}$, say $B_l^{(r)}$, will be a dominating submatrix of $P^{(r+1)}$ if $S_i \notin C_l^{(r)}$.

An Outline of the Complete Version of the Load-and-Shift Algorithm

In the procedure for finding just one solution, outlined in the preceding, we apply one load and one shift operation in each iteration. In the complete version of the algorithm we need to apply several such operations in each iteration. To eliminate possible confusion the superscript denoting the iteration number will be placed at the upper left side. Thus, for example, ${}^{(r)}P$ is the stochastic matrix remaining by the beginning of the rth iteration.

Definitions: Suppose ${}^{(r)}P$ has m_r dominating submatrices, ${}^{(r)}B_1$, ${}^{(r)}B_2$, $\cdots$, ${}^{(r)}B_{m_r}$, representing classes ${}^{(r)}C_1$, ${}^{(r)}C_2$, $\cdots$, ${}^{(r)}C_{m_r}$. The matrix ${}^{(r)}P$ was obtained from ${}^{(1)}P = P$ by executing some sequence of load and shift operations. In order to keep track of these operations we relate to ${}^{(r)}B_i$ a set ${}^{(r)}E_i$ containing labels. If the label ${}^{(k)}\alpha_j \in {}^{(r)}E_i$ we know that ${}^{(r)}B_i$ cannot be obtained unless the appropriate shift operation is applied to ${}^{(k)}B_j$. Thus, when ${}^{(r)}E_i$, $i = 1, 2, \cdots, m_r$, are given, the exact sequence of load and shift operations that led us to ${}^{(r)}P$ is known.

The n dimensional vector ${}^{(r)}\lambda$ denotes the remaining unutilized flows at the n stations of the loop. The m_r dimensional vector ${}^{(r)}\epsilon = ({}^{(r)}\epsilon_1, {}^{(r)}\epsilon_2, \cdots, {}^{(r)}\epsilon_{m_r})$ describes the amount of branch capacity utilization. Thus ${}^{(r)}\epsilon_i$ is the amount of branch capacity utilized by the sequence of load and shift operations resulting in ${}^{(r)}B_i$.

For each ${}^{(r)}B_i$ it is possible to determine the dominating solution ${}^{(r)}X_i$ and the quantity ${}^{(r)}\Delta_i$ in the manner described in steps 2 and 3 of the procedure outlined previously.

We say that ${}^{(r)}B_i \rightarrow {}^{(r)}B_j$ if when applying a shift operation on the appropriate column of ${}^{(r)}B_i$ all remaining stations of ${}^{(r)}C_i$ lead to stations of ${}^{(r)}C_j$. If ${}^{(r)}C_i$ consists of a single station then ${}^{(r)}B_i \rightarrow {}^{(r)}B_j$ if the shift is from this station to a station in ${}^{(r)}C_j$ or to a dominated station leading to ${}^{(r)}C_j$. The submatrices ${}^{(r)}B_i$, $i = 1, 2, \cdots, m_r$, divide into three types: transient, terminal, and ring members.

Ring: The set $R = \{{}^{(r)}B_{i_1}, {}^{(r)}B_{i_2}, \cdots, {}^{(r)}B_{i_k}\}$ is called a ring if it satisfies three conditions:

(*i*) ${}^{(r)}B_{i_1} \to {}^{(r)}B_{i_2} \to \cdots \to {}^{(r)}B_{i_k} \to {}^{(r)}B_{i_1}$,

(*ii*) ${}^{(r)}B_{i_j}$, $j = 1, 2, \cdots, k$, does not lead to any dominating submatrix outside of the same ring,

(*iii*) $1 - \sum_{i=1}^{k} ({}^{(r)}\epsilon_{i_j} - {}^{(r)}\Delta_{i_j}) > 0$.

Terminal: (*i*) If conditions *i* and *ii* above are satisfied while condition *iii* is violated the matrices ${}^{(r)}B_{i_1}, {}^{(r)}B_{i_2}, \cdots, {}^{(r)}B_{i_k}$ are called terminal matrices.

(*ii*) If ${}^{(r)}\Delta_i \geqq 1 - {}^{(r)}\epsilon_i$ then ${}^{(r)}B_i$ is terminal.

(*iii*) If ${}^{(r)}B_i$ is terminal and ${}^{(r)}B_j \to {}^{(r)}B_i$ then ${}^{(r)}B_j$ is terminal too.

Transient: ${}^{(r)}B_i$ is called transient if it does not belong to a ring and is not terminal.

There is a strong similarity between the notion of a ring and the notion of a dominating class. To identify rings one may use essentially the same technique proposed in the Appendix for identification of dominating classes.

It is important to observe that when applying the appropriate shift operations to all members of R, what remains of R will form one dominating submatrix and possibly some columns and rows corresponding to dominated stations will be left out. This property follows from Theorem 7.

The following is an outline of the algorithm:

Start: $r := 0$, ${}^{(1)}P := P$, ${}^{(1)}\lambda := \lambda$, ${}^{(1)}\epsilon := 0$, ${}^{(1)}E_i := \phi$.

Step 1: $r := r + 1$.
Find all the dominating submatrices of ${}^{(r)}P$ denoted by ${}^{(r)}B_1, {}^{(r)}B_2, \cdots, {}^{(r)}B_{m_r}$. Calculate the numerical values of ${}^{(r)}\Delta_i$ and ${}^{(r)}X_i$, corresponding to ${}^{(r)}B_i$, ${}^{(r)}\lambda$, and ${}^{(r)}\epsilon_i$, $i = 1, 2, \cdots, m_r$. Find all the rings of ${}^{(r)}P$ denoted by ${}^{(r)}R_1, {}^{(r)}R_2, \cdots, {}^{(r)}R_{k_r}$. Identify the terminal and transient dominating submatrices of ${}^{(r)}P$. If all dominating submatrices of ${}^{(r)}P$ are terminal go to LAST STEP. If the number of columns of ${}^{(r)}P$ equals the number of its dominating submatrices who form into one ring, go to LAST STEP 1. Otherwise ${}^{(r)}\Delta_i := 0$ for all ${}^{(r)}B_i$ not belonging to a ring.

Step 2: Execute the appropriate shift operations to all ${}^{(r)}B_i$ belonging to rings and obtain ${}^{(r+1)}P$. Note that in ${}^{(r+1)}P$ there will be

one dominating submatrix corresponding to each ring of ${}^{(r)}P$ and, in addition, all terminal and transient dominating submatrices of ${}^{(r)}P$ will be dominating in ${}^{(r+1)}P$.

Step 3:

$${}^{(r+1)}\lambda := {}^{(r)}\lambda - \sum_{i=1}^{m_r} {}^{(r)}\Delta_i \, {}^{(r)}X_i \, .$$

$${}^{(r+1)}\epsilon_i := \sum_{\{j \,:\, {}^{(r)}B_j \in {}^{(r)}R_i\}} ({}^{(r)}\epsilon_i + {}^{(r)}\Delta_i), \qquad i = 1, 2, \cdots, k_r \, .$$

$${}^{(r+1)}\epsilon_i := {}^{(r)}\epsilon_i \text{ where } {}^{(r+1)}B_i = {}^{(r)}B_i \, , i = k_r + 1, k_r + 2, \cdots, m_{r+1} \, .$$

$${}^{(r+1)}E_i := \bigcup_{\{j \,:\, {}^{(r)}B_j \in {}^{(r)}R_i\}} ({}^{(r)}E_i \cup {}^{(r)}\alpha_i), \qquad i = 1, 2, \cdots, k_r \, .$$

$${}^{(r+1)}E_i := {}^{(r)}E_i \text{ where } {}^{(r+1)}B_i = {}^{(r)}B_i \, , i = k_r + 1, k_r + 2, \cdots, m_{r+1} \, .$$ GO TO STEP 1.

Last Step 1: The unique solution is $X = \lambda$. *STOP.*

Last Step: $N := r$.

All the solutions are given by

$$X = \sum_{r=1}^{N} \sum_{i=1}^{m_r} {}^{(r)}\alpha_i \, {}^{(r)}X_i \, , \tag{21}$$

$$\sum_{r=1}^{N} \sum_{i=1}^{m_r} {}^{(r)}\alpha_i = 1, \qquad 0 \leqq {}^{(r)}\alpha_i \leqq {}^{(r)}\Delta_i \, , \qquad {}^{(r)}\alpha_i = 0, \qquad r \geqq 2,$$

unless ${}^{(k)}\alpha_i = {}^{(k)}\Delta_i$ for all ${}^{(k)}\alpha_i \in {}^{(r)}E_i$. *STOP.*

Note that it is advisable to test the possible solution $X = \lambda$ beforehand.

Theorem 8: The load-and-shift procedure will yield vectors $0 \leqq X \leqq \lambda$ in at most n iterations. Each such vector, given by equation (21), is a feasible solution to set of equations (7).

Proof: The proof is practically identical to the proofs of Theorems 4 and 5.

The use of the algorithm is illustrated by a 10-station numerical example.

Numerical Example 2

$$P = {}^{(1)}P = \begin{array}{c} \begin{matrix} S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 & S_9 & S_{10} \end{matrix} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 & 3/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3/4 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 3/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 3/4 \\ 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 1/2 & 0 \end{bmatrix} \end{array}$$

$$\lambda = {}^{(1)}\lambda = \begin{bmatrix} 0.10 \\ 0.46 \\ 0.20 \\ 0.20 \\ 0.65 \\ 0.50 \\ 0.45 \\ 0.60 \\ 0.50 \\ 1.00 \end{bmatrix}$$

We first observe that ${}^{(1)}P$ has $m_1 = 3$ dominating submatrices

$${}^{(1)}B_1 = \begin{array}{c} \begin{matrix} S_1 & S_3 \end{matrix} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}, \quad {}^{(1)}B_2 = \begin{array}{c} \begin{matrix} S_2 & S_4 & S_5 \end{matrix} \\ \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 0 & 3/4 \\ 3/4 & 1/4 & 0 \end{bmatrix} \end{array}, \quad {}^{(1)}B_3 = \begin{array}{c} \begin{matrix} S_6 & S_8 \end{matrix} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}.$$

The calculated results for the first iteration are summarized in the following table.

Results for First Iteration

${}^{(1)}C_1 = \{S_1, S_3\}$ ${}^{(1)}C_2 = \{S_2, S_4, S_5\}$ ${}^{(1)}C_3 = \{S_6, S_8\}$	${}^{(1)}X_1 = (1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0)$ ${}^{(1)}X_2 = (0\ \frac{26}{33}\ 0\ \frac{20}{33}\ \frac{28}{33}\ 0\ 0\ 0\ 0\ 0)$ ${}^{(1)}X_3 = (0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 0)$	${}^{(1)}\Delta_1 = 0.1$ ${}^{(1)}\Delta_2 = 0.33$ ${}^{(1)}\Delta_3 = 0.50$	${}^{(1)}B_1 \rightarrow {}^{(1)}B_2$ ${}^{(1)}B_2 \rightarrow {}^{(1)}B_2$ ${}^{(1)}B_3 \rightarrow {}^{(1)}B_3$	Tr	${}^{(1)}\epsilon_1 = 0$ ${}^{(1)}\epsilon_2 = 0$ ${}^{(1)}\epsilon_3 = 0$

Tr—Transient; Te—Terminal.

$$ {}^{(1)}R_1 = \{{}^{(1)}B_2\}, \qquad {}^{(1)}R_2 = \{{}^{(1)}B_3\}, $$

$$ {}^{(1)}E_1 = \phi, \qquad {}^{(1)}E_2 = \phi, \qquad {}^{(1)}E_3 = \phi. $$

Note: ${}^{(1)}\Delta_1$ will be set to equal zero before continuing. Shift operations are performed from column 4 to 5 and from 6 to 7 in ${}^{(1)}P$. The resulting 8×8 matrix, ${}^{(2)}P$, has three dominating submatrices.

$$ {}^{(2)}B_1 = \begin{array}{c} \begin{matrix} S_2 & S_5 \end{matrix} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}, \quad {}^{(2)}B_2 = \begin{array}{c} \begin{matrix} S_1 & S_9 & S_{10} \end{matrix} \\ \begin{bmatrix} 1/4 & 3/4 & 0 \\ 1/4 & 0 & 3/4 \\ 1/2 & 1/2 & 0 \end{bmatrix} \end{array}, \quad {}^{(2)}B_3 = {}^{(1)}B_1 $$

$$ {}^{(2)}\lambda = (0.10 \quad 0.20 \quad 0.20 \quad 0 \quad 0.37 \quad 0 \quad 0.45 \quad 0.10 \quad 0.50 \quad 1.00) $$

Results for Second Iteration

${}^{(2)}C_1 = \{S_2, S_5\}$ ${}^{(2)}C_2 = \{S_7, S_9, S_{10}\}$ ${}^{(2)}C_3 = \{S_1, S_3\}$	${}^{(2)}X_1 = (0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0)$ ${}^{(2)}X_2 = (0\ 0\ 0\ 0\ 0\ 0\ \frac{20}{29}\ 0\ \frac{24}{29}\ \frac{18}{29})$ ${}^{(2)}X_3 = (1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0)$	${}^{(2)}\Delta_1 = 0.20$ ${}^{(2)}\Delta_2 = 0.50$ ${}^{(2)}\Delta_3 = 0.10$	${}^{(2)}B_1 \rightarrow {}^{(2)}B_3$ ${}^{(2)}B_2 \rightarrow {}^{(2)}B_2$ ${}^{(2)}B_3 \rightarrow {}^{(2)}B_1$	Te	${}^{(2)}\epsilon_1 = 0.33$ ${}^{(2)}\epsilon_2 = 0.50$ ${}^{(2)}\epsilon_3 = 0$

$$^{(2)}R_1 = \{^{(2)}B_1 \quad ^{(2)}B_3\},$$
$$^{(2)}E_1 = \{^{(1)}\alpha_2\}, \qquad ^{(2)}E_2 = \{^{(1)}\alpha_3\}, \qquad ^{(2)}E_3 = \phi.$$

Note: $^{(2)}\Delta_2$ will be set to equal zero before continuing. Shift operations are performed from column 1 to 2 and from column 2 to 3 in $^{(2)}P$. The resulting 6 × 6 matrix, $^{(3)}P$, has two dominating submatrices.

$$^{(3)}B_1 = \overset{S_3}{[1]}, \qquad ^{(3)}B_2 = {}^{(2)}B_2 .$$
$$^{(3)}\lambda = (0 \quad 0 \quad 0.10 \quad 0 \quad 0.17 \quad 0 \quad 0.45 \quad 0.10 \quad 0.50 \quad 1.00)$$

Results for Third Iteration

$^{(3)}C_1 = \{S_3\}$ $^{(3)}C_2 = \{S_7, S_9, S_{10}\}$	$^{(3)}X_1 = (0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0)$ $^{(3)}X_2 = (0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad \frac{20}{29} \quad 0 \quad \frac{24}{29} \quad \frac{18}{29})$	$^{(3)}\Delta_1 = 0.10$ $^{(3)}\Delta_2 = 0.50$	$^{(3)}B_1 \rightarrow {}^{(3)}B_4$ $^{(3)}B_2 \rightarrow {}^{(3)}B_2$	Te	$^{(3)}\epsilon_1 = 0.63$ $^{(3)}\epsilon_2 = 0.50$

$$^{(3)}R_1 = \{^{(3)}B_1\},$$
$$^{(3)}E_1 = \{^{(1)}\alpha_2 \ , \ ^{(2)}\alpha_1 \ , \ ^{(2)}\alpha_3\}, \qquad ^{(3)}E_2 = \{^{(1)}\alpha_3\}.$$

Note: $^{(3)}\Delta_2$ will be set to equal zero before continuing. Shift operation is performed from column 3 to 5 in $^{(3)}P$. The resulting 5 × 5 matrix $^{(4)}P$ has two dominating submatrices

$$^{(4)}B_1 = \overset{S_5}{[1]}, \qquad ^{(4)}B_2 = {}^{(3)}B_2 .$$
$$^{(4)}\lambda = (0 \quad 0 \quad 0 \quad 0 \quad 0.17 \quad 0 \quad 0.45 \quad 0.10 \quad 0.50 \quad 1.00)$$

Results for Fourth (Final) Iteration

${}^{(4)}C_1 = \{S_5\}$ ${}^{(4)}C_2 = \{S_7, S_9, S_{10}\}$	${}^{(4)}X_1 = (0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0)$ ${}^{(4)}X_2 = (0\ 0\ 0\ 0\ 0\ 0\ \frac{20}{29}\ 0\ \frac{24}{29}\ \frac{18}{29})$	${}^{(4)}\Delta_1 = 0.17$ ${}^{(4)}\Delta_2 = 0.50$	${}^{(4)}B_1 \rightarrow {}^{(4)}B_2$ ${}^{(4)}B_2 \rightarrow {}^{(4)}B_2$	Te Te	${}^{(4)}\epsilon_1 = 0.73$ ${}^{(4)}\epsilon_2 = 0.50$

$$ {}^{(4)}E_1 = \{{}^{(1)}\alpha_2\ ,\ {}^{(2)}\alpha_1\ ,\ {}^{(2)}\alpha_3\ ,\ {}^{(3)}\alpha_1\},\quad {}^{(4)}E_2 = \{{}^{(1)}\alpha_3\}. $$

The iterative procedure terminates after completing the fourth iteration since all dominating submatrices of ${}^{(4)}P$ are terminal.

The set of all feasible solutions to the 10-station problem is

$$ X = {}^{(4)}\alpha_1\,{}^{(4)}X_1 + {}^{(4)}\alpha_2\,{}^{(4)}X_2 + {}^{(3)}\alpha_1\,{}^{(3)}X_1 + {}^{(2)}\alpha_1\,{}^{(2)}X_1 + {}^{(2)}\alpha_3\,{}^{(2)}X_3 + {}^{(1)}\alpha_2\,{}^{(1)}X_2 + {}^{(1)}\alpha_3\,{}^{(1)}X_3\ , $$

where

$$ {}^{(4)}\alpha_1 + {}^{(4)}\alpha_2 + {}^{(3)}\alpha_1 + {}^{(2)}\alpha_1 + {}^{(2)}\alpha_3 + {}^{(1)}\alpha_2 + {}^{(1)}\alpha_3 = 1, $$

and

$0 \leqq {}^{(4)}\alpha_1 \leqq 0.7$, ${}^{(4)}\alpha_1 = 0$ unless ${}^{(3)}\alpha_1 = 0.10$.

$0 \leqq {}^{(4)}\alpha_2 \leqq 0.50$, ${}^{(4)}\alpha_2 = 0$ unless ${}^{(1)}\alpha_3 = 0.50$.

$0 \leqq {}^{(3)}\alpha_1 \leqq 0.10$, ${}^{(3)}\alpha_1 = 0$ unless ${}^{(2)}\alpha_1 = 0.20$ and ${}^{(2)}\alpha_3 = 0.10$.

$0 \leqq {}^{(2)}\alpha_1 \leqq 0.20$, ${}^{(2)}\alpha_1 = 0$ unless ${}^{(1)}\alpha_2 = 0.33$.

$0 \leqq {}^{(2)}\alpha_3 \leqq 0.10$.

$0 \leqq {}^{(1)}\alpha_2 \leqq 0.33$.

$0 \leqq {}^{(1)}\alpha_3 \leqq 0.50$.

It remains to show that the load-and-shift algorithm will yield all the feasible solutions to equations (7).

Theorem 9: If $X^0 = (x_1^0, x_2^0, \cdots, x_n^0)$ *is a feasible solution to equations* (7), *then*

$$\lambda_i \geqq x_i^0 \geqq \text{Min}\left\{\sum_{j=1}^{N} p_{ji}x_j^0 , \lambda_i\right\}. \tag{22}$$

Proof: The left-side inequality of (1) is part of equations (7). For the right-side inequality, suppose first that $x_i^0 < \lambda_i$, then

$$\sum_{j=1}^{n} a_{ij}x_j^0 = 1 \quad \text{and} \quad \sum_{j=1}^{n} a_{i-1j}x_j^0 \leqq 1.$$

Taking the difference we obtain

$$\sum_{j=1}^{n} (a_{ij} - a_{i-1j})x_j^0 \geqq 0. \tag{23}$$

From the definition of a_{ij} we have that

$$a_{ij} - a_{i-1j} = -p_{ji}, \quad \text{for all} \quad j \neq i,$$

and

$$a_{ii} - a_{i-1i} = 1 - p_{ii}.$$

Therefore

$$x_i^0 \geqq \sum_{j=1}^{n} p_{ji}x_j^0 . \tag{24}$$

It follows then that $x_i^0 = \lambda_i$ if $\sum_{j=1}^{n} p_{ji}x_j^0 \geqq \lambda_i$ and therefore

$$\lambda_i \geqq x_i^0 \geqq \text{Min}\left\{\sum_{j=1}^{n} p_{ji}x_j^0 , \lambda_i\right\}.$$

Theorem 10: Let X^0 *be a solution to equations* (7), *and assume that* $\lambda > 0$, *then*

(*i*) *If* $x_j^0 = 0$, *then* $x_i^0 = 0$ *for all* i *such that* $S_i \to S_j$.
(*ii*) *If* $x_j^0 > 0$, *then* $x_i^0 > 0$ *for all* i *such that* $S_j \to S_i$.

Proof: If $S_j \to S_i$ there exists a sequence $p_{j,k_1}, p_{k_1,k_2}, \cdots, p_{k_{r-1},k_r}, p_{k_r,i}$ whose product is positive. Suppose $x_j^0 > 0$, then $\min(\sum_{m=1}^{n} p_{mk_1}x_m^0, \lambda_{k_1}) \geqq p_{ik_1}x_j^0 > 0$. From relation (22) it follows that $x_{k_1}^0 > 0$ and similarly $x_{k_2}^0 > 0, \cdots, x_i^0 > 0$. This proves *ii*. The proof of *i* is immediate.

Remark: The assumption $\lambda > 0$ is not restrictive. If $\lambda_i = 0$, we apply a shift operation to the ith column of P.

Theorem 11: Let $B^{(1)}$ *be a dominating submatrix of* P *and assume, without loss of generality that* $C^{(1)} = \{S_1, S_2, \cdots, S_k\}$ *is the class of dominating stations represented by* $B^{(1)}$. $\Delta^{(1)}$ *and* $X^{(1)}$ *are then uniquely defined. Suppose* X^0 *is a feasible solution to set of equations (7) then*

Either: (*i*) $x_i^0 = \alpha x_i^{(1)}$, $i = 1, 2, \cdots, k$, $0 \leqq \alpha \leqq \Delta^{(1)}$.

or: (*ii*) $x_i^0 \geqq \Delta^{(1)} x_i^{(1)}$, $i = 1, 2, \cdots, k$, *where the inequality is strict for at least one value of* i.

(*iii*) *If* i *is the case then* $x_i^0 = 0$ *for all* $S_i \notin C^{(1)}$ *and leading to* $C^{(1)}$. *If there exists such an* S_i *and* $x_i^0 > 0$, *then ii is the case.*

Proof: Since $B^{(1)}$ is a dominating submatrix then $\sum_{i=1}^{k} p_{ii} = 1$ for $i = 1, 2, \cdots, k$.

(*i*) We assume that $x_i^0 < \lambda_i$ for $i = 1, 2, \cdots, k$ and then sum the first k inequalities (24) to obtain

$$\sum_{i=1}^{k} x_i^0 \geqq \sum_{i=1}^{k} x_i^0 + \sum_{j=k+1}^{n} x_j^0 \sum_{i=1}^{k} p_{ji} .$$

For this relation to hold it is necessary that $x_j^0 \sum_{i=1}^{k} p_{ji} = 0$ for all $j \geqq k + 1$. If $C^{(1)}$ is not accessible from S_j then $\sum_{i=1}^{k} p_{ji} = 0$, otherwise either $\sum_{i=1}^{k} p_{ji} > 0$ and therefore x_j^0 must equal zero, or $S_j \to S_m$, $m \geqq k + 1$, and $\sum_{i=1}^{k} p_{mi} > 0$ and therefore x_m^0 must equal zero. From Theorem 10 we know that in this event $x_j^0 = 0$. Concluding then that $x_j^0 = 0$ for all $\{j: S_j \notin C^{(1)}\}$ and $S_i \to C^{(1)}$ we obtain [in a manner similar to the one used in obtaining relation (24)]

$$x_i^0 = \sum_{j=1}^{k} p_{ji} x_j^0 , \qquad i = 1, 2, \cdots, k,$$

yielding

$$x_i^0 = \alpha x_i^{(1)}, \quad i = 1, 2, \cdots, k.$$

Since $0 \leqq x_i^0 < \lambda_i$ we have that $0 \leqq \alpha < \Delta^{(1)}$. It also follows that if there exists $S_j \notin C^{(1)}$ and leading to $C^{(1)}$ and $x_j^0 > 0$ then there exists at least one value of i, $i = 1, 2, \cdots, k$, for which $x_i^0 = \lambda_i$.

(*ii*) Suppose there exists a value of i, $i = 1, 2, \cdots, k$, such that $x_i^0 \geqq \Delta^{(1)} x_i^{(1)}$. From Theorem 10 it follows that $x_i^0 > 0$ for all $i = 1, 2, \cdots, k$. We select a number α, $\Delta^{(1)} > \alpha > 0$, such that

$$\operatorname*{Min}_{i=1,2,\cdots,k} \{x_i^0 - \alpha x_i^{(1)}\} = x_{i_m}^0 - \alpha x_{i_m}^{(1)} = 0.$$

Clearly if such an α exists then $x_{i_m}^0 < \Delta^{(1)} x_{i_m}^{(1)}$ and from relation (24)

we have

$$x^0_{i_m} \geqq \sum_{j=1}^{n} p_{j i_m} x^0_j \,. \tag{25}$$

On the other hand we have

$$x^0_{i_m} = \alpha x^{(1)}_{i_m} = \sum_{j=1}^{k} p_{j i_m} \alpha x^{(1)}_j \,. \tag{26}$$

Subtracting equation (26) from (25) yields

$$0 = \sum_{j=1}^{k} p_{j i_m} (x^0_j - \alpha x^{(1)}_j) + \sum_{j=k+1}^{n} p_{j i_m} x^0_j \,.$$

Since $S_{i_m} \in C^{(1)}$ there exists a value of j, $j \neq i_m$ and $S_j \in C^{(1)}$, such that $p_{j i_m} > 0$. It follows then that $x^0_i = \alpha x^{(1)}_i$ for $i = 1, 2, \cdots, m$. This is a contradiction, which completes our proof.

Corollary: Let ${}^{(1)}B_1, {}^{(1)}B_2, \cdots, {}^{(1)}B_{m_1}$ *be all the dominating submatrices of* $P = {}^{(1)}P$, *and let* ${}^{(1)}C_1, {}^{(1)}C_2, \cdots, {}^{(1)}C_{m_1}$ *be the corresponding dominating classes. If* X^0 *is a solution to equations (7) then either*

$$X^0 = \sum_{i=1}^{m_1} {}^{(1)}\alpha_i \, {}^{(1)}X_i \,, \qquad \sum_{i=1}^{m_1} {}^{(1)}\alpha_i = 1, \qquad 0 \leqq {}^{(1)}\alpha_i \leqq {}^{(1)}\Delta_i \,, \tag{27}$$

or there exists at least one class of dominating stations, say ${}^{(1)}C_1$, *such that* $X^0 \geqq {}^{(1)}\Delta_i \, {}^{(1)}X_i$.

Theorem 12: If X^0 *is a feasible solution to equations (7) it is obtainable by the load-and-shift algorithm.*

Proof: If X^0 is given by equation (27), it is obviously obtainable after one iteration of the algorithm. Otherwise there exists a class of dominating stations, say ${}^{(1)}C_i$, such that $X^0 \geqq {}^{(1)}\Delta_i \, {}^{(1)}X_i$. We execute a shift operation involving the submatrix ${}^{(1)}B_i$ and obtain a stochastic matrix $P^{(2)}$ (note that $P^{(2)} \neq {}^{(2)}P$). We let $X^{0(2)} = X^0 - {}^{(1)}\Delta_i \, {}^{(1)}X_i$ and $\lambda^{(2)} = \lambda - {}^{(1)}\Delta_i \, {}^{(1)}X_i$. Clearly $X^{0(2)}$ is a feasible solution to equations (7) when using as parameters $P^{(2)}$ and $\lambda^{(2)}$ and replacing the right side by $1 - {}^{(1)}\Delta_i$. It is possible therefore to proceed with a sequence of load-and-shift operations until X^0 is obtained. Since the algorithm takes into account all possible sequences of load-and-shift operations, X^0 is contained in the set of solutions given by equation (21).

Corollary (uniqueness): If P has exactly one dominating matrix then there exists a unique feasible solution to equations (7). Note, however,

that this is not a necessary condition for uniqueness. It is possible that P will have several dominating submatrices and the solution will be unique.

V. DISCUSSION

The heavy traffic assumption enables us to regard the system as a deterministic one. The analysis of the deterministic flows shows that stations tend to "band" into classes with the ability of dominating the system and preventing other stations from using it.

This undesirable property, which can be eliminated by exercising appropriate control, also may affect the stochastic behavior of the system when heavy traffic conditions do not exist. One can imagine two classes of stations competing for domination of the system. Since traffic is not heavy, all the stations in the system are able to deliver their messages in finite time. Nevertheless, when one dominating class controls the system, it will prevent other stations from using the belt line until the queue at one of the stations belonging to this class becomes empty. At that moment the competing dominating class is able to take over and prevent other stations from using the belt line. This may result in a situation of alternating priorities (see Ref. 4) where, while one class is served, the queues at the competing class build up. While average queue sizes may not be strongly affected, the strong fluctuations in queue lengths may be undesirable. This possibility has been explored numerically by the use of a digital simulator.[5] The operating principles of the simulated system will be explained with the aid of Fig. 2.

Each of the stations is represented by a B-box. A packet coming out of a station is first multiplexed on the line by the B-box, provided the line is free, and then is passed from B-box to B-box until its destination is reached. At each B-box on the way the address of the packet is examined. At the particular B-box of destination the packet is taken off the line. The main function of the A-box, shown in Fig. 2, is synchronization of the loop.

Assume that the packets are made of L bits each and the address is given in the first k bits of the packet. A time unit in this system is the time it takes to multiplex a bit on the main line by a B-box. Assume also that the traveling time from one B-box to the adjacent one is zero. Station i is allowed to start sending a packet at times $mL + ik$, $m = 0, 1, 2, \cdots$, providing that the main line is free. The A-box is a buffer. Bits coming out of the B-box of station n accumulate in the A-box. Suppose at time k B-box 1 starts sending a packet. At time

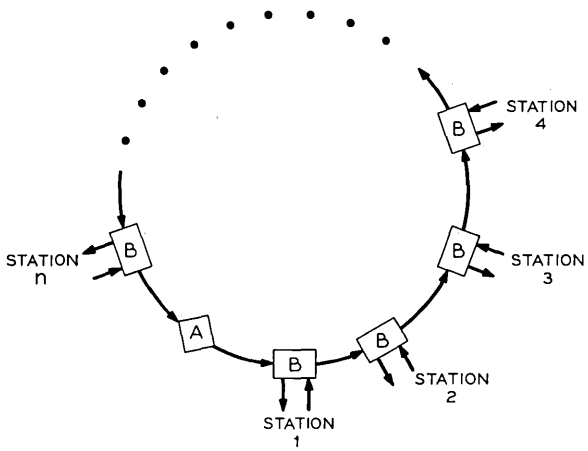


Fig. 2—A schematic description of a loop with n stations.

$k + 1$ the first bit reaches B-box 2 and is delayed there until time $2k$ when the whole address of the packet has been received. If the packet is addressed to station 2, it will be taken out and at the same time B-box 2 can start sending out its own packet. If the packet is not addressed to station 2, it will be sent from B-box 2 to B-box 3 where the same process will take place starting at time $3k$. Bits arriving at the A-box are buffered. At time L the A-box starts to send bits (at the same rate as a B-box) until a whole packet has seen sent. If the buffer is empty the A-box will wait another L time units and will start sending at time $2L$. In general, the ith B-box checks its buffer at times $mL + ik$, $m = 0, 1, 2, \cdots$, and if the buffer is empty it may start sending its own packet. If the buffer contains the address of station i, the B-box will remove the arriving packet and may, at the same time, send out its own packet. If the buffer contains an address different than station i, the B-box will pass on the arriving packet.

In a similar manner the A-box checks its buffer at times mL, $m = 1, 2, \cdots,$. If the buffer is not empty, it sends out L bits. If the buffer is empty, the A-box remains inoperative for the next L time units.

The loop time is the time it takes a bit to complete one round of the loop, and is measured in multiples of L. In the single A-box loop described here the loop time is the smallest integer greater than or equal to nk/L where n is the number of B-boxes in the loop. Clearly there is a complete analogy between the loop described here and the one presented in Fig. 1 if the number of revolving arms is taken as equal to the loop time.

It is important to note that the simulated system described here represents only one conceptual way in which the loop may be operated. It is possible, for example, that the packets will move from one B-box to another in one block rather than bit by bit. This is equivalent to placing an A-box between any two B-boxes (n rotating arms).

We have used the digital simulator to examine the queuing characteristics of several small systems. The main purpose of the simulation was to study the effects of dominating classes in nonheavy (nonsaturated) traffic situations. Numerical results are presented for an 8-station loop with two dominating classes $C_1 = (S_1, S_2, S_3, S_4)$ and $C_2 = (S_5, S_6, S_7, S_8)$. The loop time for this system was selected to equal 1 (one rotating arm) and the system was simulated for three different expected main line loads (utilization). The P matrix and average queue sizes (in packets) at the stations are shown below.

Simulated Example: 8-Station Loop (Two Dominating Classes)

$$P = \begin{array}{c|cccc:cccc} & S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 \\ \hline S_1 & 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ S_2 & 1/3 & 0 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ S_3 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ S_4 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ \hdashline S_5 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ S_6 & 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 \\ S_7 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 1/3 \\ S_8 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 \end{array}$$

No.	Line and Source Utilization	Average Queue Sizes S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Ave.	Max.
1	$\lambda = 0.239$, $\rho = 0.956$	9.7	10.1	13.8	6.7	15.6	14.6	11.2	6.8	11.1	33–58
2	$\lambda = 0.229$, $\rho = 0.916$	5.6	5.6	6.1	3.0	5.0	6.8	5.2	3.2	5.1	26–45
3	$\lambda = 0.213$, $\rho = 0.852$	2.4	2.6	1.9	1.1	2.4	2.3	2.1	1.2	2.0	16–22

The alternating priorities effect, due to domination, is demonstrated in Fig. 3. The total number of packets at the four queues of C_1 (dotted line) and C_2 (solid line) were plotted against time. For the case $\rho =$

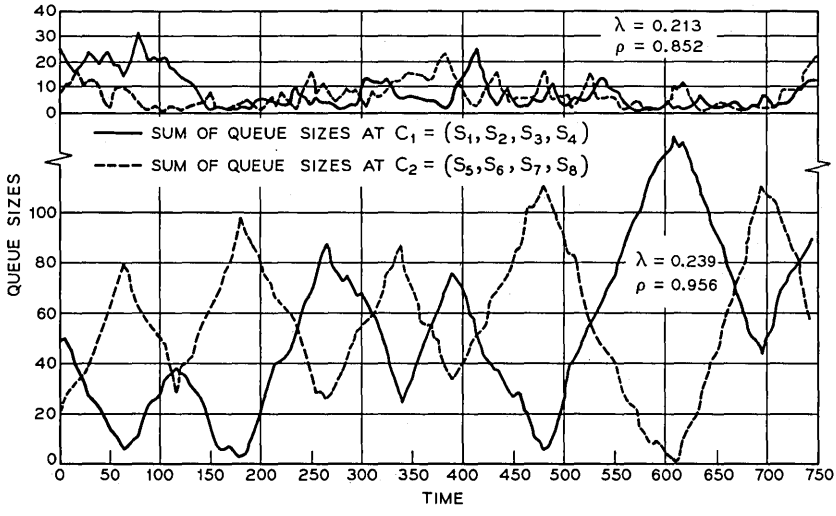


Fig. 3—Simulated queue sizes.

0.956 (high utilization) one can clearly see that only one dominating class is served at one time. When stations of C_2 take over control of the system the queues at stations of C_1 build up while at C_2 they are being depleted until the queue at S_8 reaches zero and C_1 can take over.† The average cycle time (time elapsing between two consecutive peaks of the dotted or solid lines) for this case was 110 time units. As the load on the system is decreased, the alternating priorities effect becomes less and less distinctive. For $\rho = 0.916$ (not graphed) the average cycle time reduces to 23 time units, and for $\rho = 0.852$ (see Fig. 3) alternations are very frequent and cycles are practically unnoticeable.

Notwithstanding the complete symmetry within classes the average queue sizes at S_4 and S_8 are consistently smaller than the average queue sizes at the other stations. The explanation of this phenomenon is as follows: At a moment when C_1 loses control to C_2 the queue size at S_4 is zero while at S_1, S_2, and S_3 it is greater than or equal to zero. From that moment on the queues of C_1 build up at equal average rates until the moment control returns to C_1 (peak of the dotted line in Fig. 3). At that moment the queues start being depleted at equal average

† Note that the queue size at S_8 being zero is a necessary but not sufficient condition for losing control.

rates. Since the expected queue size at S_4 is smallest, it has a higher probability of being completely depleted first. This phenomenon tends to shorten the alternations cycle.

The alternations cycle will be shorter when the number of rotating arms (loop time) is increased. Therefore large loops (local or regional loops) will be relatively more stable when high utilization occurs.

The nature of the stochastic process used for generating packets at the stations of the loop is described in Ref. 5.

An important aspect, not analyzed in this study, is the question of the order of stations in the loop. Clearly, the amount of traffic the loop can carry and the resulting congestion are strongly dependent on the specific order of the stations in the loop. In Example 1 we have assumed counterclockwise traffic direction. If we reverse the direction of traffic on the main line we shall get a different solution for the flows. The two solutions are compared in the following:

	Counterclockwise Flow Direction	Clockwise Flow Direction
	$x_1 = 0.4646$	$x_1 = 0.5000 = \lambda_1$
	$x_2 = 0.4000 = \lambda_2$	$x_2 = 0.4000 = \lambda_2$
	$x_3 = 0.5000 = \lambda_3$	$x_3 = 0.5000 = \lambda_3$
	$x_4 = 0.2333$	$x_4 = 0.0917$
	$x_5 = 0.0508$	$x_5 = 0.3667$
Total:	1.6487	1.8584

Reversing the flow direction results in an increase of 13 percent in the amount of satisfied demand.

In a practical situation not all orders are feasible. Still the number of feasible orders may be overwhelmingly large and an appropriate algorithm for determining best order is called for. An interesting possibility is a double loop system where each station is connected to two loops with opposite traffic directions. This may increase reliability and enable better utilization by allocating traffic in an efficient manner. One possible allocation rule is shortest distance allocation where the loop to be used is the one with the shortest travel distance for each particular message (this is an example of a possible rule and is not proposed as an optimal rule).

The bounded linear complementarity problem presented by equations (7) is of somewhat more general interest, bearing little relation

to the Pierce loop. In matrix form we have

$$X + U = \lambda,$$
$$AX + Z = 1,$$
$$U^T Z = 0, \quad U \geqq 0, \quad Z \geqq 0, \quad X \geqq 0.$$

Substituting $X = \lambda - U$ we obtain the same set of relations in a slightly different form.

$$AU - Z = A\lambda - 1 = q$$
$$U^T Z = 0, \quad U \geqq 0, \quad Z \geqq 0, \quad U \leqq \lambda. \tag{28}$$

It is now possible to compare our problem to the Fundamental Problem† treated by Lemke,[6] and Cottle and Dantzig.[2]

$$AU - Z = q$$
$$U^T Z = 0, \quad U \geqq 0, \quad Z \geqq 0. \tag{29}$$

The only basic difference between the two problems is that in our problem U is bounded from above while in Lemke's problem U is unconstrained. In this respect our problem is more general. The shift-and-load procedure is, however, fundamentally based on the specific structure of A and q and may not prove useful for a wider class of parameters.

VI. ACKNOWLEDGMENT

I am indebted to J. Salz, J. F. Hayes, and D. N. Sherman for introducing me to the Pierce loop concept, and to M. Segal and S. Halfin for their helpful comments and suggestions.

A special acknowledgment is due R. R. Anderson who wrote the digital simulation and ran the numerical examples used in this study.

APPENDIX

In this appendix we outline a procedure for determining all dominating submatrices of a given stochastic $n \times n$ matrix P.

Step 1: Construct a matrix $\Pi = \{\pi_{ij}\}$ where

$$\pi_{ij} = \begin{cases} 1 & \text{if } p_{ij} > 0, \\ 0 & \text{otherwise.} \end{cases}$$

† The term "Fundamental Problem" was coined by Cottle and Dantzig. [2]

Step 2: $Q = \{q_{ij}\} := \Pi$.
Do for $j = 1, 2, \cdots, n$.
Repeat until q_{kj}, $k = 1, 2, \cdots, n$, remain unchanged.

$$q_{kj} := q_{kj} + \sum_{\{i : q_{ij}=1\}} q_{ki}, \qquad k = 1, 2, \cdots, n.$$

(Note that all additions are Boolean.)
If $S_i \rightarrow S_j$ then in the resulting matrix Q the element $q_{ij} = 1$. Otherwise $q_{ij} = 0$.

Step 3: Construct a matrix $Q^{(0)} = \{q_{ij}^0\}$ such that

$$q_{ij}^{(0)} = q_{ji}^{(0)} = q_{ii}q_{jj}.$$

If S_i does not belong to a communicating class then $q_{ij}^{(0)} = 0$, $j = 1, 2, \cdots, n$. Otherwise $q_{ii}^{(0)} = 1$ and $S_i \in C(i)$ if and only if $q_{ij}^{(0)} = 1$. $C(i)$ is closed (dominating) if and only if the ith row of $Q^{(0)}$ is identical to the ith row of Q.

REFERENCES

1. Pierce, J. R., "Network for Block Switching of Data," unpublished work.
2. Cottle, R. W., and Dantzig, G. B., "Complementary Pivot Theory of Mathematical Programming," *Mathematics of the Decision Sciences*, Part 1, American Math. Soc., 1968, pp. 115–136.
3. Kemeny, G. J., and Snell, J. L., *Finite Markov Chains*, New York: D. Van Nostrand Co., 1965.
4. Avi-Itzhak, B., Maxwell, W. L., and Miller, L. W., "Queuing with Alternating Priorities," *Oper. Res.*, *13*, No. 2, 1965, pp. 306–318.
5. Hayes, J. F., and Sherman, D. N., "Traffic Considerations in a Pierce Data Loop," unpublished work.
6. Lemke, C. E., "Bimatrix Equilibrium Points and Mathematical Programming," Mgt. Sci., *11*, No. 7, 1965, pp. 681–689.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Bending Losses of the Asymmetric Slab Waveguide

By D. MARCUSE

(Manuscript received April 30, 1971)

The bending losses of the asymmetric slab waveguide are computed. The computation is based on the knowledge of the exact form of the solution of Maxwell's equations of the bent structure and the additional assumption that the field near the bent waveguide can be approximated by the field of the straight waveguide. The result of this theory is in good agreement with an existing theory. It appears that the bending loss formula can be used to estimate the bending losses of the round optical fiber if the mode parameters entering the formula are replaced by the corresponding mode parameters of the round fiber. We present curves that allow the numerical evaluation of the bending loss of the lowest order even TE mode of the symmetric slab waveguide.

I. INTRODUCTION

E. A. J. Marcatili has shown that a bent slab waveguide loses power by radiation.[1] His analysis is based on a solution of the eigenvalue equation of the bent waveguide. It is possible to derive the expression for the bending losses from an approximate theory that is much simpler than the solution of the eigenvalue equation. We use this method to derive the formula for the bending losses of an asymmetric slab waveguide. The symmetric slab waveguide is, of course, included in this treatment as a limiting case. The result of this approximate theory is in very good agreement with the theory of Marcatili. Furthermore, if the parameters of the HE_{11} mode of the round optical fiber are used in the slab waveguide formula, loss values are obtained that agree well with experimental loss values for this mode.[2]

The bending loss theory presented in this paper is based on the following idea. A bent slab waveguide can conveniently be described in a cylindrical coordinate system whose axis coincides with the center of curvature of the waveguide (Fig. 1). The solution of Maxwell's equations

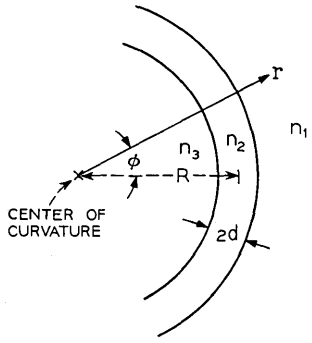


Fig. 1—Bent slab waveguide with a cylindrical coordinate system centered at the center of curvature.

in the cylindrical coordinate system is known so that the shape of the field distribution of the curved waveguide is known except for an undetermined amplitude factor and for the value of the order number of the cylinder function of the solution. Both of these unknown parameters can be obtained if we assume that the field in the vicinity of the waveguide must be similar to the field of the straight guide as long as the radius of curvature is large. The approximate solution is used to calculate the power that is radiated from the waveguide so that the power loss per unit length caused by the waveguide curvature can be determined. This procedure leads to a simple equation for the curvature loss. The theory breaks down when the curvature is so severe that the field near the waveguide can no longer be approximated by the field of the straight guide. The limits of applicability of the curvature loss theory can be expressed by inequalities for the waveguide parameters.

II. THE FIELD OF THE STRAIGHT ASYMMETRIC SLAB WAVEGUIDE

The field of the straight asymmetric slab waveguide is obtained as the solution of a straightforward boundary value problem. The geometry of the structure is shown in Fig. 2. We assume that there is no field variation in the y direction so that the waves of the structure are simple TE and TM modes. We limit our discussion to TE modes. The field is then given by the following equations:[3]

$$E_y = Ae^{-\gamma(x-d)} \qquad d \leqq x < \infty, \tag{1a}$$

$$E_y = A\cos\kappa(x-d) - \frac{\gamma}{\kappa}\sin\kappa(x-d) \qquad -d \leqq x \leqq d, \tag{1b}$$

$$E_y = A\left(\cos 2\kappa d + \frac{\gamma}{\kappa}\sin 2\kappa d\right)e^{\theta(x+d)} \qquad -\infty \leqq x \leqq -d. \tag{1c}$$

The amplitude of the field can be expressed by the power P carried by the field:

$$A = 2\kappa\left\{\frac{\omega\mu_o P}{\beta\left(2d + \frac{1}{\gamma} + \frac{1}{\theta}\right)(\kappa^2 + \gamma^2)}\right\}^{\frac{1}{2}}. \tag{2}$$

A factor $\exp[i(\omega t - \beta z)]$ has been suppressed. The constants and parameters appearing in equations (1) and (2) are defined as follows:

ω = $2\pi f$, radian frequency,
P = power carried by the mode,
μ_o = magnetic permeability of free space,
ϵ_o = electric permittivity of free space,
β = propagation constant,
$2d$ = slab thickness,

$$\kappa = (n_2^2k^2 - \beta^2)^{\frac{1}{2}}, \tag{3}$$

$$\gamma = (\beta^2 - n_1^2k^2)^{\frac{1}{2}}, \tag{4}$$

$$\theta = (\beta^2 - n_3^2k^2)^{\frac{1}{2}}, \tag{5}$$

$$k = \omega(\epsilon_o\mu_o)^{\frac{1}{2}}, \tag{6}$$

n_1 = refractive index in the region $d < x$,
n_2 = refractive index in the region $-d < x < d$,
n_3 = refractive index in the region $-\infty < x < -d$.

The magnetic field components are obtained from the equations

$$H_x = -\frac{i}{\omega\mu_o}\frac{\partial E_y}{\partial z} \tag{7}$$

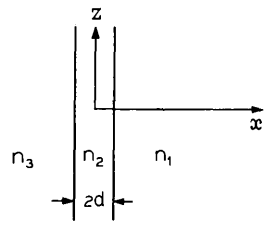


Fig. 2—Straight slab waveguide. n_1, n_2, and n_3 are the refractive indices of the three media.

and

$$H_z = \frac{i}{\omega\mu_o}\frac{\partial E_y}{\partial x}. \tag{8}$$

The boundary conditions (requirement of continuity of the transverse electric and magnetic field components at the two interfaces) lead to the eigenvalue equation for the propagation constant β,

$$\tan 2\kappa d = \frac{\gamma + \theta}{\kappa\left(1 - \frac{\gamma\theta}{\kappa^2}\right)}. \tag{9}$$

III. THE FIELD OF THE CURVED STRUCTURE

The solution of Maxwell's equations in the region $(R + d) < r < \infty$ can be expressed as follows:

$$E_y = BH_\nu^{(2)}(n_1 kr)e^{-i\nu\phi}. \tag{10}$$

The Hankel function of the second kind and of order ν represents an exact solution of Maxwell's equations in the coordinate system shown in Fig. 1. There is no field variation in the direction of the axis of the cylindrical polar coordinate system of Fig. 1. (In the coordinate system used in Fig. 2 the direction of the polar coordinate axis would be y.) The radial distance r is measured from the center of curvature of the bent waveguide. The Hankel function of the second kind is required since at infinite distance, $r \to \infty$, an outward traveling wave must result. With our time dependence, exp $(i\omega t)$, the field of equation (10) satisfies this requirement.

The order number ν need not be an integer in this case since we need not require periodicity of the field as a function of the polar angle ϕ. If we were interested in an exact solution of the problem of the mode traveling along a curved waveguide we would obtain the value of ν as the solution of an eigenvalue equation. In our approximate treatment we assume that the field near the waveguide can still be approximately described by the field of the straight structure. We can use the coordinate system of the straight guide (Fig. 2) to describe the curved guide. The z axis of the straight coordinate system becomes bent and we have the relation

$$z = R\phi. \tag{11}$$

The function exp $(-i\nu\phi)$ is equivalent to the propagation factor exp $(-i\beta z)$ of the straight waveguide so that we have the approximate

relation

$$\beta = \frac{\nu}{R}. \tag{12}$$

Since R is much larger than the wavelength of the field, ν is a very large number. Using ν as defined by (12) in (10) constitutes the first approximation. It remains to obtain a relation between the amplitude factor B and the power that is carried by the guided mode. To achieve this we use an approximation for the Hankel function that is valid for very large order number ν in the region

$$\nu > n_1 kr. \tag{13}$$

The desired approximation can be found in Ref. 4.

$$H_\nu^{(2)}(n_1 kr) = -i \frac{e^{\nu(\alpha - \tanh \alpha)}}{\sqrt{\frac{\pi}{2} \nu \tanh \alpha}} \tag{14}$$

with

$$\cosh \alpha = \frac{\nu}{n_1 kr}. \tag{15}$$

We obtain the hyperbolic tangent of α by the relation

$$u = \tanh \alpha = \frac{[\cosh^2 \alpha - 1]^{\frac{1}{2}}}{\cosh \alpha} = \frac{\left[\beta^2 - \left(n_1 k \frac{r}{R}\right)^2\right]^{\frac{1}{2}}}{\beta}. \tag{16}$$

Equation (12) was used to replace ν by β. Near the axis of the waveguide we can use the approximation $r/R \approx 1$ so that we obtain from (4), (12), and (16)

$$\nu \tanh \alpha = \gamma R. \tag{17}$$

This approximation is adequate for the denominator of (14). The expression in the exponent must be approximated more accurately. We use the x coordinate to describe the radial distance from the center of the waveguide core and write

$$\frac{r}{R} = 1 + \frac{x}{R}. \tag{18}$$

This x coordinate corresponds directly to the x axis of the straight guide as shown in Fig. 2. Using (4) and (16) we obtain

$$u \approx \frac{\left(\gamma^2 - 2(n_1k)^2 \frac{x}{R}\right)^{\frac{1}{2}}}{\beta} \approx \frac{\gamma}{\beta}\left[1 - \left(\frac{n_1k}{\gamma}\right)^2 \frac{x}{R}\right]. \tag{19}$$

We can express α by a well-known relation between the inverse hyperbolic tangent function and the natural logarithm[4]

$$\alpha = \tanh^{-1} u = \tfrac{1}{2} \ln \left(\frac{1+u}{1-u}\right). \tag{20}$$

The logarithm can be expanded in an infinite series

$$\alpha = u + \tfrac{1}{3}u^3 + \tfrac{1}{5}u^5 + \quad u^7 + \cdots. \tag{21}$$

From (19) we obtain approximately

$$u^p = \left(\frac{\gamma}{\beta}\right)^p - p\left(\frac{n_1k}{\gamma}\right)^2 \frac{x}{R}\left(\frac{\gamma}{\beta}\right)^p. \tag{22}$$

Substitution of (22) into (21) yields

$$\alpha = \frac{\gamma}{\beta} + \frac{1}{3}\left(\frac{\gamma}{\beta}\right)^3 + \frac{1}{5}\left(\frac{\gamma}{\beta}\right)^5 + \cdots - \left(\frac{n_1k}{\gamma}\right)^2 \frac{x}{R}\frac{\gamma}{\beta}\left[1 + \frac{\gamma^2}{\beta^2} + \left(\frac{\gamma^2}{\beta^2}\right)^2 + \left(\frac{\gamma^2}{\beta^2}\right)^3 + \cdots\right]. \tag{23}$$

The first series can again be expressed by the logarithmic function. The second part of (23) contains a simple geometric series. We thus obtain the approximation

$$\alpha = \tfrac{1}{2} \ln \left(\frac{1 + \frac{\gamma}{\beta}}{1 - \frac{\gamma}{\beta}}\right) - \frac{(n_1k)^2}{\gamma\beta} \frac{1}{1 - \frac{\gamma^2}{\beta^2}} \frac{x}{R}. \tag{24}$$

With the help of (4), (19), and (24) we can form the expression

$$\alpha - \tanh \alpha = \alpha - u = \tfrac{1}{2} \ln \left(\frac{1 + \frac{\gamma}{\beta}}{1 - \frac{\gamma}{\beta}}\right) - \frac{\gamma}{\beta} - \frac{\gamma}{\beta}\frac{x}{R}. \tag{25}$$

The approximations (12), (14), (17), and (25) allow us to express the E_y component (10) in the following approximate form:

$$E_y = -iB \frac{\exp\left\{\left[\left[\frac{1}{2}\frac{\beta}{\gamma} \ln \left(\frac{1 + \frac{\gamma}{\beta}}{1 - \frac{\gamma}{\beta}}\right) - 1\right]\gamma R\right\}}{\sqrt{\frac{\pi}{2}\gamma R}} e^{-\gamma x} e^{-i\beta z}. \tag{26}$$

The functional dependence of (26) on the coordinate x coincides with (1a).

For large values of R the field of the curved structure near the waveguide is approximately equal to that of the straight guide. The factor $\exp(-i\beta z)$ appearing in (26) was omitted from (1a). Comparison of (1a) and (26) allows us, with the help of (2), to determine the amplitude coefficient of the field,

$$B = 2i\kappa e^{\gamma d}\left\{\frac{\frac{\pi}{2}\omega\mu_o\gamma RP}{\beta\left(2d+\frac{1}{\gamma}+\frac{1}{\theta}\right)(\kappa^2+\gamma^2)}\right\}^{\frac{1}{2}} \cdot \exp\left\{-\left[\frac{1}{2}\frac{\beta}{\gamma}\ln\frac{1+\frac{\gamma}{\beta}}{1-\frac{\gamma}{\beta}}-1\right]\gamma R\right\}. \tag{27}$$

Very far from the waveguide, $r \gg R$, the Hankel function can be expressed by its approximation for large argument[4] so that (10) assumes the form

$$E_y = B\sqrt{\frac{2}{\pi n_1 kr}}\, e^{-in_1 kr} e^{i(2\nu+1)\pi/4} e^{-i\beta z}. \tag{28}$$

Equation (28) is very interesting. It shows that far from the waveguide the field is very different from the field of the straight structure. Whereas the field of the curved structure decays exponentially near the waveguide it assumes the form of a radiation field far from the guide. This behavior of the exact field solution (10) of the curved structure explains why curved dielectric waveguides lose power by radiation.

IV. THE BENDING LOSS FORMULA

Since we know the field far from the waveguide it is now easy to calculate the power loss caused by the fact that energy is radiated away from the waveguide. The amplitude of the radiation field is independent of the z coordinate. The z dependent factor in (28) determines only the phase of the field. The power loss suffered by the field at a given position z can thus be calculated by the radial power flow at the same position z, even though the contribution to this radiation may have come from a point z_1 with $z_1 \ll z$, because each length element of the guide contributes an equal amount of radiation. An element of unit length on the axis of the waveguide is projected on an element of arc length

$$L = \frac{r}{R} \tag{29}$$

at a distance $r - R$ from the waveguide. The power loss 2α per unit length of waveguide is thus

$$2\alpha = \frac{r}{R}\frac{S_r}{P}. \tag{30}$$

S_r is the r component of the Poynting vector,

$$S_r = -\tfrac{1}{2}E_y H_\phi^* = \tfrac{1}{2}n_1\sqrt{\frac{\epsilon_o}{\mu_o}}\,|\,E_y\,|^2, \tag{31}$$

and P is the power carried by the mode. Using (27), (28), and (31) we obtain from (30)

$$2\alpha = \frac{2\gamma\kappa^2 e^{2\gamma d}e^{-U}}{(n_2^2 - n_1^2)k^2\beta\left(2d + \frac{1}{\gamma} + \frac{1}{\theta}\right)} \tag{32}$$

with

$$U = \left\{\frac{\beta}{\gamma}\ln\left|\frac{1+\frac{\gamma}{\beta}}{1-\frac{\gamma}{\beta}}\right| - 2\right\}\gamma R \approx \frac{2}{3}\frac{\gamma^2}{\beta^2}\gamma R. \tag{33}$$

The approximation on the right-hand side of (33) holds for $\gamma/\beta \ll 1$. The relation $\kappa^2 + \gamma^2 = (n_2^2 - n_1^2)k^2$ was used to simplify (32). The range of validity of the bending loss formula (32) cannot be given precisely. We have already encountered the inequality (13) that was necessary for the approximation (14) to hold. A similar inequality can be stated for the field expressed by Bessel and Neumann functions inside of the waveguide. In order to be able to express the field inside of the curved waveguide by approximate expressions that reduce to the sine and cosine functions appearing in (1b) in the limit of large radius of curvature, we must require

$$\nu < n_2 kr \tag{34}$$

everywhere inside of the waveguide. We can express these conditions in the form

$$\beta > n_1 k\left(1 + \frac{d}{R}\right) \tag{35}$$

and

$$\beta < n_1 k\left(1 - \frac{d}{R}\right). \tag{36}$$

The validity of our theory becomes doubtful if one or both of these inequalities are violated. However, a comparison with Marcatili's theory[1] shows that our approximation is still quite good even in regions where (36) no longer holds.

The simple expression (32) for the bending loss of an asymmetric slab waveguide can be used only if the values of κ, γ, θ, and β are known. It is, of course, only necessary to determine one of these parameters from the eigenvalue equation (9) since they are all interconnected by the equations (3) through (6).

It is useful to point out that the loss equation (32) seems to be applicable to other types of waveguide than the one for which it was derived. I have compared experimental values[2] of bending losses of a round optical fiber with the loss predicted by (32). For such a comparison it is necessary to use the parameters κ, γ, etc., that apply to the waveguide to which the formula is to be applied. In case of the round fiber the parameters of the HE_{11} mode were used to compute the bending loss from (32). The reason for this choice of parameters is the fact that the parameter γ determines the decay behavior of the field outside of the waveguide. It is very important that the proper field decay is used, so that it is more logical to use the γ value of the round fiber instead of the value computed from (9), if (32) is to be used to compute the bending loss of the round fiber.

It is a curious fact that the loss formula (32) can also be obtained without use of the Hankel function appearing in (10) if we use a field of the form

$$E_y = \frac{C}{\sqrt{\gamma(r)r}} \exp\left\{-i \int_R^r \gamma(r)\, dr\right\} \tag{37}$$

with

$$\gamma(r) = \left[\beta^2 \frac{R^2}{r^2} - n_1^2 k^2\right]^{\frac{1}{2}}. \tag{38}$$

The validity of this claim can easily be checked by performing the integration. The factor in front of the exponential function is somewhat arbitrary. However, the exponential function itself admits of a physical interpretation.

The straight waveguide has a field that, outside of its core, behaves according to $\exp(-\gamma x)$. In the curved system x is naturally replaced by r. If we consider that the process of bending the waveguide is likely to lead also to a distortion of the phase fronts we can try to describe the separation of consecutive wavefronts by an r dependent wavelength

(λ_g is the wavelength of the straight guide),

$$\lambda(r) = \lambda_g \frac{r}{R}. \tag{39}$$

The propagation constant is related to the wavelength by the relation

$$\beta(r) = \frac{2\pi}{\lambda(r)} = \beta_g \frac{R}{r}. \tag{40}$$

By replacing the propagation constant in (4) with (40) we obtain (38). Since γ is now no longer a constant it is natural to replace γr by $\int \gamma \, dr$ and thus arrive at the form of the exponential function appearing in (37). By using (37) instead of (10) and proceeding exactly as shown in this paper we obtain (32) once more. It is interesting that we have thus obtained an approximation for the Hankel function that holds for the region where the order number is very nearly equal to the argument as well as for the region where the argument is much larger than the order number. Equation (38) shows clearly that $\gamma(r)$ changes from real to imaginary values as r increases.†

One might hope that a similar procedure would allow us to obtain approximate expressions for the bending loss of the round optical fiber. However, such attempts lead to equations that are not in agreement with experiment. On the other hand, the loss formula (32) agrees well with experiment[2] if the parameters of the round fiber are used.

The loss formula (32) holds for all values of the refractive indices n_1, n_2, and n_3 for which mode guidance is possible. Small index differences are not required for (32) to be valid.

V. NUMERICAL EXAMPLES

Because of the large number of variables involved it is not possible to provide graphic displays for all possible applications. The loss formula (32) is sufficiently simple (except for the need of knowing the waveguide parameters κ, γ, etc.) so that loss values for cases of interest can easily be calculated. We provide curves that aid in computing the bending loss of the even TE mode of the symmetric slab waveguide.

Figure 3 is a comparison of our theory with the results of Ref. 1. The ordinate is the function $(2\Delta)^{\frac{1}{2}}\alpha R$ while $(8\Delta)^{\frac{1}{2}}2dn_2/\lambda$ is plotted on the abscissa. The parameter Δ is defined as $n_2 - n_1$ (we are using $n_1 = n_3$). The expression $36n_2R\Delta^{\frac{3}{2}}/\lambda$ assumes the constant value 60 for the curve of Fig. 3. The agreement with Marcatili's theory[1] is re-

† A discussion of bending losses based on a similar argument is presented in Ref. 5.

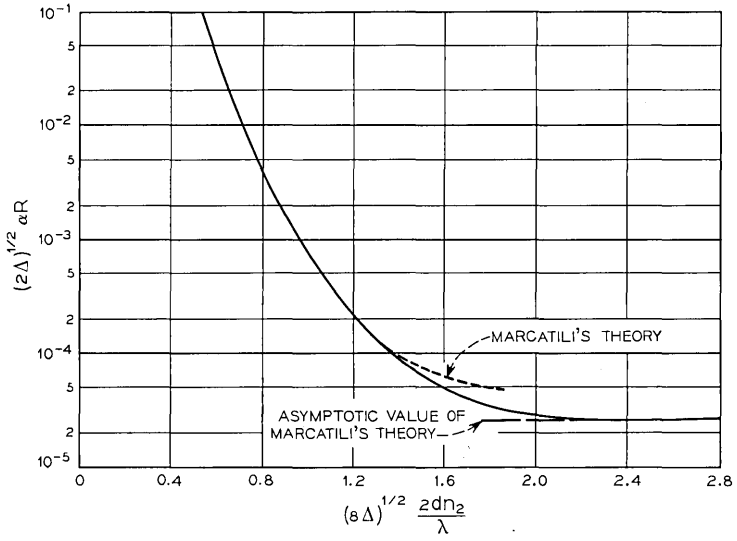


Fig. 3—Comparison between this theory and the theory of Ref. 1. $\Delta = n_2 - n_1$, $n_1 = n_3$, $36n_2R\Delta^{\frac{3}{2}}/\lambda = 60$.

markably good. The inequality (36) is violated for values of the abscissa that are larger than 1.7. This may explain the departure between the solid curve representing equation (32) and the dash-dotted curve representing Marcatili's theory. The fact that the solid curve actually touches the asymptotic value shown as a dotted line in Fig. 3 (this value is assumed for infinite values of the abscissa, its location as shown in the figure has no meaning) is probably an accident since the solid line increases again for larger values of the abscissa outside of the range shown in the figure.

We restrict ourselves to the case of the symmetric slab waveguide with $n_1 = n_3$. It is possible to express κ and γ (which equals θ in this special case) as functions of

$$V = (n_2^2 - n_3^2)k^{\frac{1}{2}}d. \tag{41}$$

The propagation constant β depends not only on V but also on the value of n_1k so that it is not possible to express the bending loss only in terms of V. However, in the special case that n_1 and n_2 are very nearly equal we have $\beta = n_1k$ so that we need not actually solve the eigenvalue equation to obtain the propagation constant. To aid in the evaluation of the bending loss formula we provide curves for $2\alpha\beta d^2e^U$ and for $d^3\beta^2U/R$ in

Fig. 4 as functions of V. The approximate expression given in (33) was used to express U. If β can be approximated by n_1k these two curves enable us to calculate the bending loss without any difficulty. For symmetric slab waveguides with a large value of $n_2 - n_1$ we can calculate β with the help of (4) from known values of γ and n_1k. A plot of γd as a function of V is provided in Fig. 5. The parameter γ is interesting in itself since it determines the exponential decay of the guided mode outside of the waveguide core.

VI. CONCLUSIONS

The bending loss of an asymmetric slab waveguide has been calculated using an approximation that is based on the assumption that the field near the bent guide is still almost identical to the field of the straight guide. The results of this approximate theory are in good agreement with the bending loss theory of Marcatili[1] in the range of applicability of our theory. It is hard to apply a similar analysis to the bent round fiber because the exact form of the solutions of Maxwell's equations for the curved structure is not known. However, the bending loss formula obtained for the slab waveguide model yields good agreement with

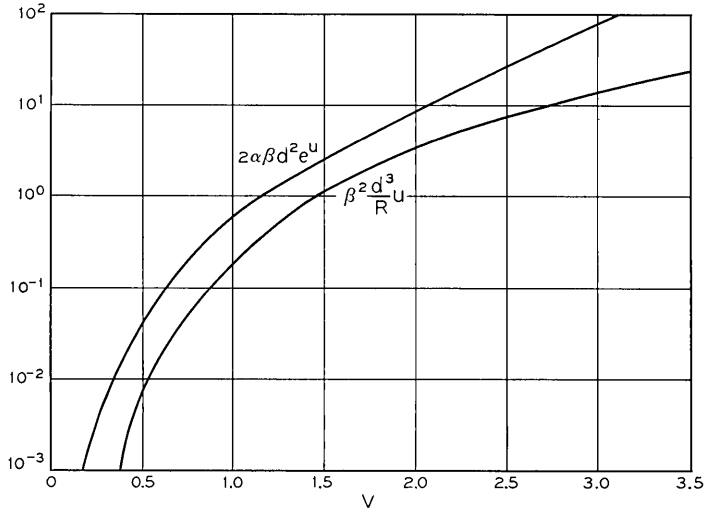


Fig. 4—The functions $2\alpha\beta d^2 e^U$ and $d^3\beta^2 U/R$ are plotted versus $V = (n_2 - n_1)^{\frac{1}{2}} kd$. ($n_1 = n_3$.)

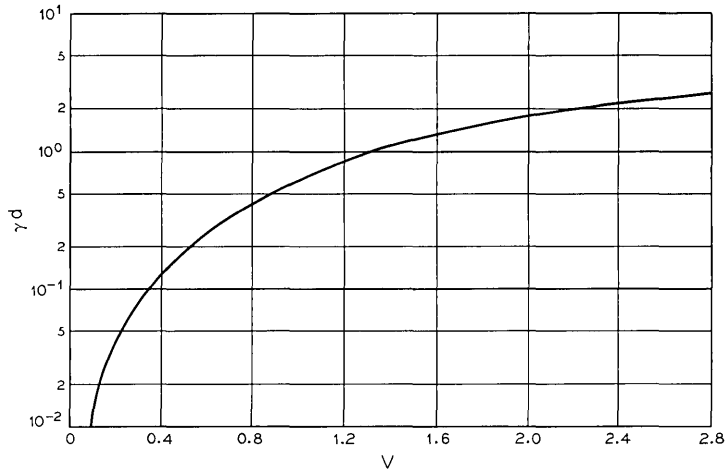


Fig. 5—Plot of γd as a function of V. ($n_1 = n_3$.)

experiment[2] if the mode parameters of the fiber mode are used in the loss formula instead of the mode parameters of the slab waveguide.

For the case of small index differences, curves that allow the determination of the bending loss of the lowest order symmetric TE mode of the symmetric slab waveguide are provided.

REFERENCES

1. Marcatili, E. A. J., "Bends in Optical Dielectric Guides," B.S.T.J., *48*, No. 7 (September 1969), pp. 2103–2132.
2. Kapron, F. P., Keck, D. B., and Maurer, R. D., "Radiation Losses in Glass Optical Waveguides," Appl. Phys. Letters, *17*, No. 10 (November 15, 1970), pp. 423–425.
3. Marcuse, D., and Marcatili, E. A. J., "Excitation of Waveguides for Integrated Optics with Laser Beams," B.S.T.J., *50*, No. 1 (January 1971), pp. 43–57.
4. Gradshteyn, I. S. and Ryzhik, I. M. *Tables of Integrals, Series and Products*, New York: Academic Press, 1965.
5. Marcatili, E. A. J., and Miller, S. E., "Improved Relations Describing Directional Control in Electromagnetic Wave Guidance," B.S.T.J., *48*, No. 7 (September 1969), pp. 2161–2188.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Attenuation of Unwanted Cladding Modes

By D. MARCUSE

(Manuscript received April 5, 1971)

This paper contains design criteria for the suppression of cladding modes in optical fibers by means of a lossy jacket. The calculations are actually based on the slab waveguide model. However, it is believed that the results are representative, at least to order of magnitude, for the round optical fiber. It is found that the cladding modes are absorbed most effectively if the real part of the refractive index of the lossy jacket equals the refractive index of the cladding material. However, slight deviations from this optimum design are not critical.

The loss of the cladding modes depends on many parameters so that general statements as to the order of magnitude of the expected loss are hard to make. However, a loss of 1 *dB/m should be easily obtainable.*

I. INTRODUCTION

Losses in dielectric optical waveguides have been discussed by several authors.[1,2] In two earlier papers[3,4] we explored the problem of crosstalk between two parallel optical waveguides. Two types of crosstalk were considered. The directional coupler mechanism couples two dielectric waveguides even if they have no imperfections of any kind. The coupling, in this case, is caused by the exponentially decaying field tail of one guide reaching to the region of the second guide. The second mechanism causing crosstalk involves light scattering from waveguide imperfections. The light that is scattered out of the guided mode of one guide can be scattered back into the guided mode of a neighboring waveguide. Both types of crosstalk can, in principle, be reduced by placing a lossy material between the two guides. However, it was concluded that isolation of the waveguides by means of a lossy surrounding medium is really not necessary to avoid crosstalk. The waveguides must be designed in such a way that their mode losses are low. The same mechanism that influences crosstalk also influences the mode loss. For example, if two guides couple appreciably by means of the directional coupler mechanism their mode fields penetrate each other to a con-

siderable extent. Adding loss to the surrounding medium thus also causes added loss to the mode field. In fact, if the loss in the surrounding medium is needed to reduce crosstalk it simultaneously adds so much loss to the guided mode as to defeat its purpose. The cladding of the fiber must thus be made sufficiently thick to protect the guided mode from losses of its environment. But a sufficiently thick cladding also provides adequate protection against crosstalk. The scattering crosstalk problem, on the other hand, does not benefit from a lossy surrounding medium for different reasons. Even if the cladding is made sufficiently thick to protect the guided mode from losses of the surrounding medium it is still possible to obtain substantial scattering crosstalk. However, the same mechanism that produces scattering crosstalk also causes scattering loss to the guided modes. Except for the unlikely case of a systematic periodic distortion that persists throughout the entire length of both waveguides, the scattering mode loss is more critical than the scattering crosstalk so that it is necessary to build a guide with sufficient accuracy to hold scattering losses down. Once this is achieved scattering crosstalk is no longer a problem.

It thus appears as though it were not necessary to surround the members of a bundle of waveguides with a lossy medium in order to prevent crosstalk provided that the waveguides are built with low scattering losses and with a sufficiently thick cladding. However, there is one more reason why losses in the surrounding medium may be desirable. We have so far ignored the possibility that power which is scattered out of the core of the waveguide is trapped in the cladding. These cladding modes are undesirable for two reasons. They couple some of their power back into the core causing problems of delay distortion of the guided core modes. In addition, these modes may reach the end of the waveguide and there enter the detector giving further rise to unwanted, delayed signals. It is thus important to suppress cladding modes by providing each waveguide with a lossy jacket.

We must keep in mind that the loss requirements for the cladding modes are more stringent than the loss requirements for the core modes. A loss factor that is prohibitively high for a core mode may be far too low for a cladding mode. The cladding mode can be harmful even if it is too lossy to propagate through the entire length of the waveguide by itself. Cladding modes are continuously excited and reconvert to core modes throughout the entire length of the waveguide so that their presence is undesirable not only at the end of the waveguide but throughout the entire length of the guide. It is thus desirable to ensure cladding

mode losses far higher than the losses that appear excessive if they should occur for core modes.

A very effective way of suppressing cladding modes would be to provide the waveguides of the cable with a surrounding medium whose index is higher than that of the cladding. In this case the light can no longer be trapped in the cladding but would rapidly escape into the surrounding medium. However, if it were not attenuated there it would still be harmful since some of it can scatter back into guided modes and some of it may reach the detectors at the end of the guide giving rise to crosstalk and delay distortion.

The next best thing, therefore, is designing a surrounding medium or jacket with a refractive index whose real part nearly equals the index of the cladding but which is sufficiently lossy to attenuate power that tries to travel in it. This paper tries to provide the design criteria necessary to build effective lossy jackets for the suppression of cladding modes.

II. CLADDING MODE LOSS FORMULA

Our interest is, of course, directed toward the cladding modes of round optical fibers. In fact, the formula for core mode losses in round optical fibers provided in Ref. 3 can be used to calculate cladding mode losses in fibers with infinitely thick jacket by letting the refractive indices of cladding and surrounding medium coincide and by shrinking the cladding thickness to zero. However, a discussion of the cladding mode losses for round optical fibers with finite jacket thickness is complicated due to the complexity of the round fiber equation. As always, the slab waveguide is much easier to treat than the round optical fiber. Results obtained for the slab waveguide model are applicable to the round fiber at least as order of magnitude estimates. Since order of magnitude estimates are all that we need in the present situation and since the formulas for the slab waveguide are so much easier to evaluate we shall base our discussion of cladding mode losses on the slab waveguide model.

The slab waveguide model to be studied is schematically shown in Fig. 1. We ignore the core of the waveguide since it does not contribute appreciably to the propagation behavior of the cladding modes. The refractive index of the fiber core is so close to the refractive index of the cladding in most fibers of interest for communications purposes that we can ignore the core altogether. The cladding mode losses are caused by the fact that the reflection of the electromagnetic energy at the boundary

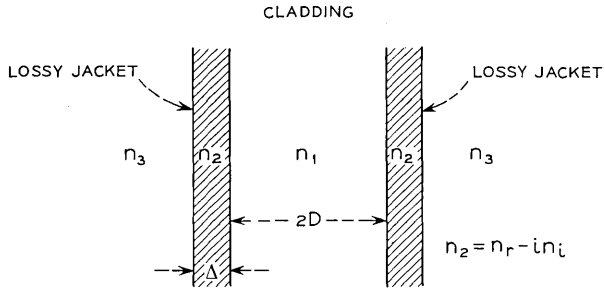


Fig. 1—Sketch of a dielectric waveguide (core not shown) with a lossy jacket.

between cladding and jacket is not 100 percent efficient. It is thus the plane wave reflection coefficient at the cladding jacket interface that determines the cladding mode losses. The reflection coefficient is higher for waves whose electric field vector is polarized parallel to the dielectric interface (TE modes) compared to waves whose magnetic vector is parallel to the interface (TM modes). The modes with the higher reflection coefficient have lower loss. The TE modes have lower loss than the TM modes. It is thus sufficient to study TE modes since the TM mode losses must be higher. A design that succeeds in suppressing TE modes does suppress TM modes even more effectively.

In order to obtain the mode losses for the structure shown in Fig. 1 we study the reflection coefficient of a plane wave incident on the structure shown in Fig. 2. The electric field vector is given by the equations:

$$E_y = Ae^{-i(\beta z+\kappa x)} + Be^{-i(\beta z-\kappa x)} \qquad x < 0, \tag{1}$$

$$E_y = Ce^{-i(\beta z+\sigma x)} + Fe^{-i(\beta z-\sigma x)} \qquad 0 \leqq x \leqq \Delta, \tag{2}$$

$$E_y = Ge^{-i(\beta z+\rho x)} \qquad \Delta \leqq x \leqq \infty . \tag{3}$$

The time dependent factor exp ($i\omega t$) has been suppressed. The parameters κ, σ, and ρ are defined by the equations:

$$\kappa = (n_1^2k^2 - \beta^2)^{\frac{1}{2}}, \tag{4}$$

$$\sigma = (n_2^2k^2 - \beta^2)^{\frac{1}{2}}, \tag{5}$$

$$\rho = (n_3^2k^2 - \beta^2)^{\frac{1}{2}}, \tag{6}$$

with

$$k = \omega(\epsilon_0\mu_0)^{\frac{1}{2}} = 2\pi/\lambda_0 . \tag{7}$$

The magnetic field components are obtained from E_y with the help of

$$H_x = -\frac{i}{\omega\mu_0}\frac{\partial E_y}{\partial z}, \qquad H_z = \frac{i}{\omega\mu_0}\frac{\partial E_y}{\partial x}. \tag{8}$$

The boundary conditions at the two dielectric interfaces determine the relations between the amplitude coefficients:

$$B = \frac{\sigma(\kappa - \rho)\cos\sigma\Delta + i(\kappa\rho - \sigma^2)\sin\sigma\Delta}{\sigma(\kappa + \rho)\cos\sigma\Delta + i(\kappa\rho + \sigma^2)\sin\sigma\Delta} A, \tag{9}$$

$$C = \frac{\kappa(\sigma + \rho)e^{i\sigma\Delta}}{\sigma(\kappa + \rho)\cos\sigma\Delta + i(\kappa\rho + \sigma^2)\sin\sigma\Delta} A, \tag{10}$$

$$F = \frac{\kappa(\sigma - 1)e^{-i\sigma\Delta}}{\sigma(\kappa + \rho)\cos\sigma\Delta + i(\kappa\rho + \sigma^2)\sin\sigma\Delta} A, \tag{11}$$

$$G = \frac{2\kappa\sigma e^{i\rho\Delta}}{\sigma(\kappa + \rho)\cos\sigma\Delta + i(\kappa\rho + \Delta^2)\sin\sigma\Delta} A. \tag{12}$$

The field that is described by the equations (1) through (12) is the cladding mode field of the waveguide. The type of mode that actually results depends on the width $2D$ of the guide and on the angle at which the plane wave travels in the slab; that is, it depends on the ratio κ/β. For the calculation of the mode losses we do not need to know this ratio very accurately. It is clear, however, that lossless modes satisfy the

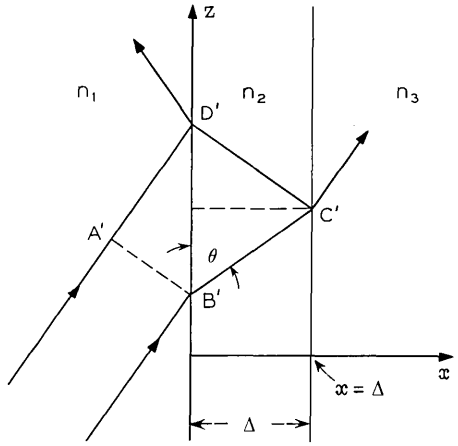


Fig. 2—This diagram shows the interference between an incident wave and the wave that has penetrated into the jacket.

requirement $| A | = | B |$. The transverse field distribution in the slab at $-D \leqq x \leqq D$ is thus a standing wave. Since it is very hard to achieve a large loss coefficient α for the cladding modes we can assume $\alpha \ll 1/\lambda$ and have

$$\frac{| A | - | B |}{| A |} \ll 1. \tag{13}$$

In order to obtain the mode loss, we calculate the power flow along the waveguide axis,

$$P_z = \text{Re}\left\{-\frac{1}{2}\int_{-2D}^{0} E_y H_x^* \, dx\right\} = \text{Re}\left\{\frac{\beta}{2\omega\mu}\int_{-2D}^{0} | E_y |^2 \, dx\right\}. \tag{14}$$

Re indicates the real part of the quantity in brackets. The integral is extended only over the cladding region of the waveguide assuming that most of the power is contained in this region. For reasonably low loss this condition is satisfied. Using (1) and (13) we obtain

$$P_z \approx \frac{2\beta D}{\omega\mu} | A |^2. \tag{15}$$

The amount of power at the cladding boundary that flows per unit area out of the waveguide is given by

$$S_x = \text{Re}\{\tfrac{1}{2} E_y H_z^*\}_{x=0} = \frac{\kappa | A |^2}{2\omega\mu}\left(1 - \left|\frac{B}{A}\right|^2\right). \tag{16}$$

The power loss coefficient 2α (α is the amplitude loss coefficient) is

$$2\alpha = \frac{2S_x}{P_z}. \tag{17}$$

The factor 2 on the right-hand side of (17) takes account of the fact that equal amounts of power are lost at the cladding jacket interface on either side of the waveguide. Substitution of (15) and (16) into (17) results in the desired formula for the loss of the cladding modes:

$$2\alpha = \frac{\kappa}{2\beta D}\left(1 - \left|\frac{B}{A}\right|^2\right). \tag{18}$$

The coefficient B/A is obtained from (9). We consider only the case that the medium outside of the waveguide jacket has the lowest refractive index, $n_3 < | n_2 |$ and $n_3 < n_1$. We thus have

$$\rho = -i\gamma \tag{19}$$

with

$$\gamma = (\beta^2 - n_3^2 k^2)^{\frac{1}{2}} \tag{20}$$

being a real quantity. We set

$$\sigma = u - iv \tag{21}$$

with u and v both real and positive. In the limit of an infinitely thick jacket, $\Delta \to \infty$, we obtain from (9) and (18)

$$2\alpha = \frac{2\kappa^2 u}{\beta D(u^2 + v^2 + \kappa^2 + 2u\kappa)}. \tag{22}$$

It is interesting to consider two different alternatives. First we assume that

$$\mathrm{Re}(n_2^2 k^2) > \beta^2. \tag{23}$$

The refractive index of the jacket material is a complex number

$$n_2 = n_r - i\frac{\alpha_j}{k}. \tag{24}$$

α_j is the amplitude loss coefficient of a plane wave traveling in the jacket. We assume

$$n_r \gg \frac{\alpha_j}{k} \tag{25}$$

and obtain from (5) and (21)

$$u = [\tfrac{1}{2}\sigma_0^2 + \tfrac{1}{2}(\sigma_0^4 + 4n_r^2 k^2 \alpha_j^2)^{\frac{1}{2}}]^{\frac{1}{2}} \tag{26}$$

and

$$v = [-\tfrac{1}{2}\sigma_0^2 + \tfrac{1}{2}(\sigma_0^4 + 4n_r^2 k^2 \alpha_j^2)^{\frac{1}{2}}]^{\frac{1}{2}} \tag{27}$$

with

$$\sigma_0 = (n_r^2 k^2 - \beta^2)^{\frac{1}{2}}. \tag{28}$$

It is apparent from (22) that α vanishes for $u = 0$ and for $u = \infty$. This indicates that 2α must assume a maximum as a function of u. By differentiation of (22) we find that the maximum is located at

$$u_{\max} = (\kappa^2 + v^2)^{\frac{1}{2}}. \tag{29}$$

The maximum value $2\alpha_{\max}$ that corresponds to (29) is given by

$$2\alpha_{\max} = \frac{\kappa^2}{\beta D[(\kappa^2 + v^2)^{\frac{1}{2}} + \kappa]}. \tag{30}$$

In order to express the condition for $u_{\max}$ in terms of the refractive

indices of the waveguide we square (29) and substitute (26) on the left-hand side and (27) on the right-hand side of the squared equation. We obtain the solution

$$\sigma_0 = \kappa. \tag{31}$$

Comparison of (4) and (28) shows that the maximum cladding mode loss for an infinitely thick jacket is obtained if

$$n_r = n_1 , \tag{32}$$

that is, for the case that the real part of the refractive index of the jacket is equal to the refractive index of the cladding. The requirement of an infinite jacket thickness means only that the effect of the boundary between the jacket and the outside medium must be negligible. As a practical matter this condition is satisfied if the product $\alpha_i\Delta$ is larger than about 10 dB. For an infinitely wide jacket and as long as the condition (23) holds we obtain maximum loss for vanishing v, that is, in the case of a lossless jacket. Increasing the loss of the jacket decreases the loss of the cladding modes slightly as long as v remains smaller than κ. For very high losses of the jacket material the cladding mode loss decreases with increasing jacket loss. This consideration shows that the loss of the jacket material must not be made too high.

In the opposite case

$$\mathrm{Re}(n_2^2k^2) < \beta^2 \tag{33}$$

we obtain

$$u = [-\tfrac{1}{2}\,\delta^2 + \tfrac{1}{2}(\delta^4 + 4n_r^2k^2\alpha_i^2)^{\frac{1}{2}}]^{\frac{1}{2}} \tag{34}$$

and

$$v = [\tfrac{1}{2}\,\delta^2 + \tfrac{1}{2}(\delta^4 + 4n_r^2k^2\alpha_i^2)^{\frac{1}{2}}]^{\frac{1}{2}} \tag{35}$$

with

$$\delta = (\beta^2 - n_r^2k^2)^{\frac{1}{2}}. \tag{36}$$

For vanishing loss in the jacket we now obtain $u = 0$ so that the cladding modes propagate without loss in this case. The waves are totally internally reflected at the interface between cladding and jacket. The condition for $u_{\max}$ given by (29) can now not be satisfied. For numerical calculation we shall always assume that the cladding modes are so tightly guided that we can use

$$\beta = n_1k. \tag{37}$$

Far from cutoff the values for κ are obtained from[5]

$$\kappa D = (\nu + 1)\frac{\pi}{2} \tag{38}$$

with integer values of ν. Even integers of ν belong to even modes while odd integers belong to odd modes. Equations (22) and (30) show that the lowest order mode ($\nu = 0$) propagates with the lowest loss. Since we are interested in maximizing the losses of all the modes it is thus sufficient to consider the loss of the lowest order even TE mode and use

$$\kappa D = \frac{\pi}{2}. \tag{39}$$

For finite thickness of the jacket material and sufficiently low loss the cladding mode losses vary as a function of cladding thickness Δ. Figures 3 and 4 show these loss fluctuations for the case $kD = 50$, $n_1 = 1.6$, $n_r = 1.65$, and $n_3 = 1$. The jacket loss was assumed to be $2\alpha_j D = 0.1$ dB in case of Fig. 3 and 1 dB in case of Fig. 4. These loss variations as a function of jacket thickness complicate the design of an optimum jacket. Since we need a certain minimum loss for the effective suppression of all cladding modes we are interested in the minimum values of the loss. It is thus necessary to find the position of the loss minima of equation (18). Since it is difficult to find the minimum values mathematically from the loss expression we use a physical argument based on ray optics for the case $n_r > n_1$.

Figure 2 shows the geometry of the problem. Minimum loss is obtained if the light rays that are reflected from the cladding-jacket boundary add in phase to the light rays that penetrate into the jacket, are reflected at its outer boundary, and reenter the cladding. The condition for minimum cladding mode loss can thus be stated as follows

$$(n_r k\eta + \phi_2) - (n_1 k\xi + \phi_1) = 2\nu\pi. \tag{40}$$

The distance from point A' to point D' in Fig. 2 is ξ, the distance from point B' to C' and on to D' is η, and the phase angles ϕ_1 and ϕ_2 are the additional phase shifts that the wave suffers on reflection from the interface between cladding and jacket and between jacket and the outside medium. The number ν must be an integer. Using simple geometry we obtain

$$\eta = \frac{2\Delta}{\sin\theta}. \tag{41}$$

The angle of the incident ray is shown vastly exaggerated in Fig. 2.

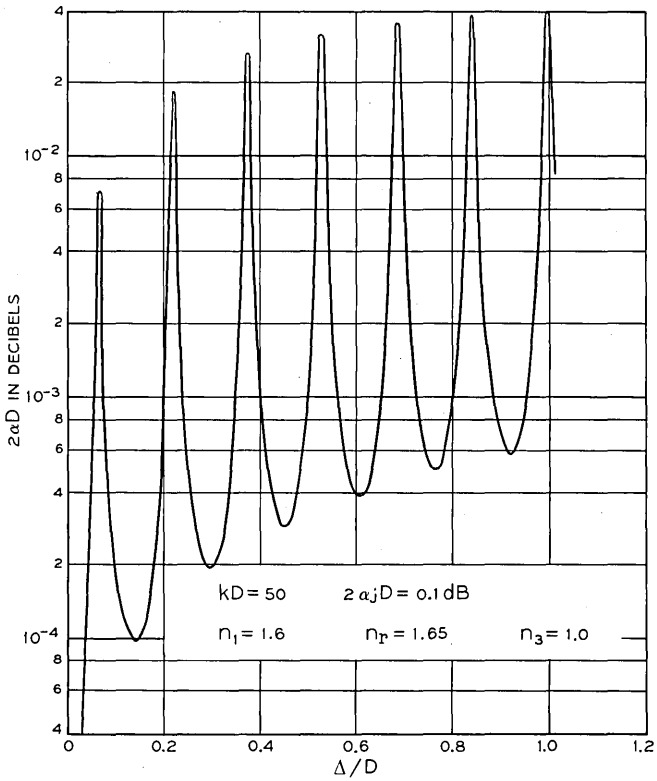


Fig. 3—Fluctuation of cladding mode loss as a function of jacket thickness Δ. (D = cladding thickness; $2\alpha_j$ = power loss coefficient of the jacket material; k = propagation constant of free space; n_1, n_r, and n_3 are respectively the refractive indices of the cladding, the (real part) of the jacket, and the outside medium). $kD = 50$, $2\alpha_j D = 0.1$ dB, $n_1 = 1.6$, $n_r = 1.65$, $n_3 = 1.0$.

For low-order cladding modes the angle between the incident ray and the interface is very nearly zero. We thus obtain from Snell's law

$$\sin \theta = \left[1 - \frac{n_1^2}{n_r^2}\right]^{\frac{1}{2}}. \tag{42}$$

Because the input ray intercepts the interface at grazing angles we can approximate ξ as follows

$$\xi = 2\Delta \cot \theta = 2\Delta \frac{n_1}{[n_r^2 - n_1^2]^{\frac{1}{2}}}. \tag{43}$$

The additional phase angle on reflection from the optically denser jacket

is

$$\phi_1 = \pi. \tag{44}$$

The phase angle ϕ_2 is more complicated. From the Fresnel formulas for the transmission of a plane wave across a dielectric interface we obtain

$$\phi_2 = -2 \arctan \left[\frac{(n_1^2 - n_3^2)^{\frac{1}{2}}}{(n_r^2 - n_1^2)^{\frac{1}{2}}} \right]. \tag{45}$$

Collecting all these equations we obtain from (40) the jacket width that minimizes the cladding mode loss:

$$\Delta_{\min} = \frac{\nu\pi + \frac{\pi}{2} + \arctan \left\{ \frac{(n_1^2 - n_3^2)^{\frac{1}{2}}}{(n_r^2 - n_1^2)^{\frac{1}{2}}} \right\}}{k(n_r^2 - n_1^2)^{\frac{1}{2}}}. \tag{46}$$

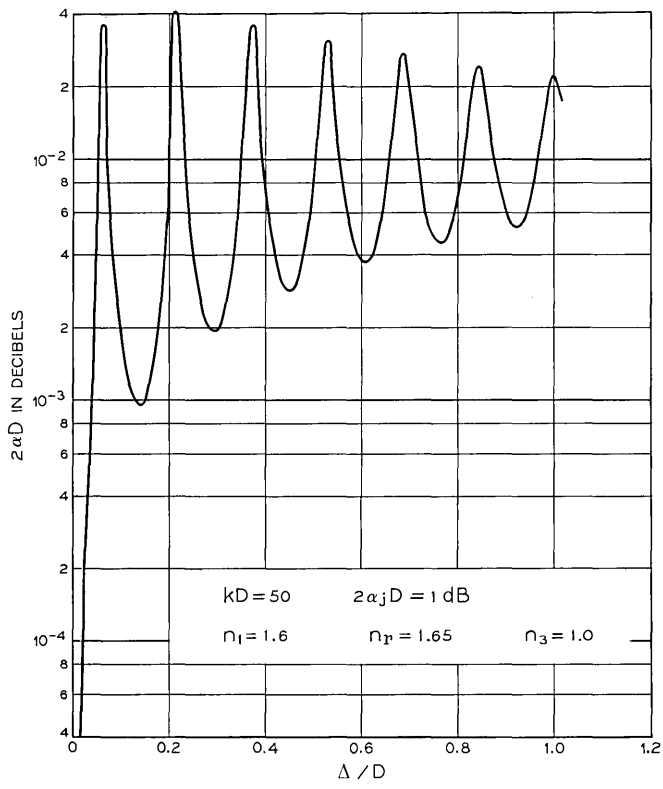


Fig. 4—Same as Fig. 3 except that $2\alpha_j D = 1$ dB.

Numerical values of (46) agree with the position of the minima appearing in Figs. 3 and 4.

III. NUMERICAL RESULTS

The theory presented on the preceding pages contains the information for the design of lossy jackets for the purpose of suppressing cladding modes. Because of the large number of independent parameters entering equations (9) and (18) it is impossible to draw curves covering all possible cases of interest. We are thus limiting the discussion of numerical results to a few cases hoping to show the trend and the order of magnitude of the cladding mode losses that may be obtainable.

All curves are calculated for a cladding index $n_1 = 1.6$ and an index of $n_3 = 1$ for the surrounding medium. Most curves apply to the case $kD = 50$. Those curves are drawn as solid lines. Broken curves are examples for $kD = 100$ and dash-dotted curves apply to $kD = 500$. A comparison between these three kD values gives some indication of the dependence of the cladding mode losses on the kD parameter.

The cladding mode loss $2\alpha D$ for a guide length equal to the half width D of the cladding is shown in Fig. 5 for $n_r/n_1 > 1$ as a function of the jacket loss $2\alpha_j\Delta$. (D = cladding half width, Δ = jacket thickness, both losses are expressed in dB). Figure 5 was computed from (9), (18), and (46) for the minima of the fluctuating loss curves (compare Figs. 3 and 4). It is important to remember that the first loss minimum occurs for $\Delta_{\min}$ of (46) for $\nu = 0$. The curves of Fig. 5 do not apply to values of the jacket thickness Δ that are much smaller than this minimum thickness. The loss curves shown in Figs. 3 and 4 drop off very rapidly for Δ values that are smaller than the value at the first minimum so that the mode losses obtained from Fig. 5 would be much higher than the actual losses if the figure were applied to Δ values that are smaller than the first minimum. For jackets that are thicker than the minimum value (46) with $\nu = 0$ the curves of Fig. 5 show the lowest possible cladding mode loss. The loss curves shown in Fig. 5 are correct for small values of $2\alpha_j D$ (note that we have just replaced Δ with D). As long as it stays below a certain critical value the curves are independent of $2\alpha_j D$. For larger values the cladding mode losses decrease with increasing $2\alpha_j D$. The critical value is different for each curve shown in Fig. 5. The dependence of the cladding mode losses on $2\alpha_j D$ is shown in Fig. 6. This figure presents the values of the cladding mode losses for infinite jacket thickness Δ as a function of $2\alpha_j D$. It is apparent from Fig. 5 that the cladding mode losses become independent of the jacket thickness for

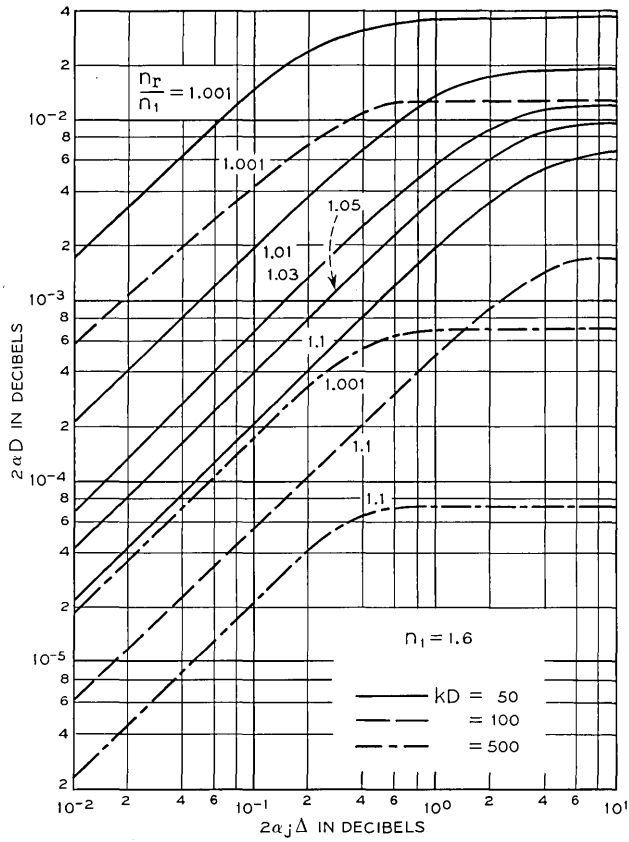


Fig. 5—The minimum cladding mode loss as a function of the jacket loss $2\alpha_j\Delta$ for different values of $n_r/n_1 > 1$.

sufficiently large values of Δ. Figure 6 shows the mode losses of Fig. 5 after they have leveled off ($\Delta \rightarrow \infty$). It can be seen that the cladding mode losses become less sensitive to the value of $2\alpha_j D$ for increasing values of the jacket-to-cladding index ratio n_r/n_1. However, the loss curves for $n_r/n_1 = 1.001$ become dependent on the jacket loss coefficient α_j already for $2\alpha_j D = 1$ dB. It is important to keep this in mind when using Fig. 5.

Comparison of the dash-dotted, broken, and solid lines in Figs. 5 and 6 shows that the cladding mode loss depends on the kD parameter as $(kD)^{-2}$ for $\Delta \rightarrow \infty$. It is very hard to see this relation from (18); however

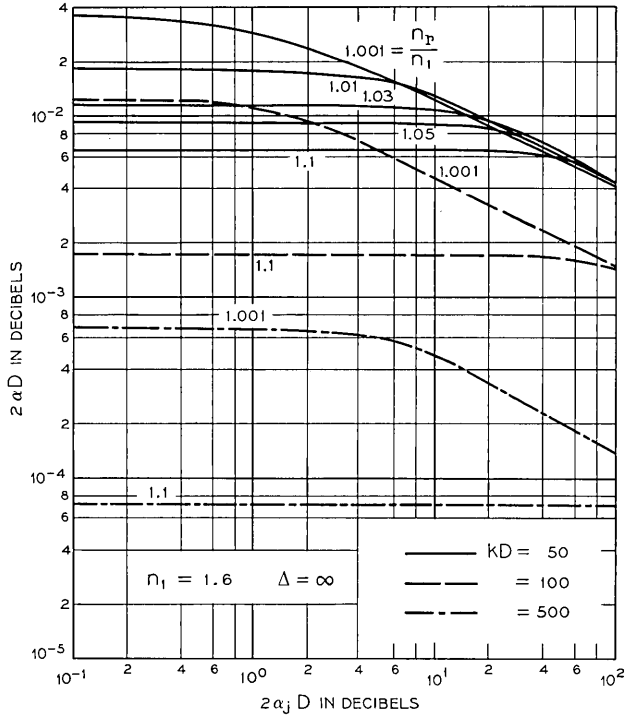


Fig. 6—Cladding mode losses for infinitely thick jacket, $\Delta \rightarrow \infty$, as functions of the jacket loss parameter $2\alpha_j D$ for $n_r/n_1 > 1$.

it can be seen from (22) and (26) for small values of v. It is interesting to observe that the maximum loss (30) depends on kD only as $(kD)^{-1}$.

In order to obtain a feeling for the absolute loss values that can be achieved let us assume that the cladding half thickness is $D = 10\mu$m. The loss factor $2\alpha D = 10^{-5}$ dB thus corresponds to a cladding mode loss of $2\alpha = 1$ dB/m. Cladding mode losses well in excess of 1 dB/m are thus easily obtainable. In fact it appears possible to design jackets that provide cladding mode losses of 100 to 1000 dB/m.

So far we have considered jackets with a refractive index whose real part is larger than the index of the cladding material. If the real part of the jacket index is smaller than the cladding index, $n_r/n_1 < 1$, the cladding mode losses are no longer even approximately independent of $2\alpha_j D$. The loss curves can thus not be drawn as functions of $2\alpha_j\Delta$. Figures 7, 8, and 9 are representations of the cladding mode losses for $n_r/n_1 < 1$ as functions of Δ/D. The parameter $2\alpha_j D$ is used to label the

different curves in each figure. Comparison shows that for equal values of kD and $2\alpha_j D$ the loss curves are nearly identical for all three figures for low values of Δ/D. The leveling off of the curves for large values of Δ/D is, however, quite different for each value of n_r/n_1. The dependence of the cladding mode loss for $n_r/n_1 < 1$ as a function of $2\alpha_j D$ for $\Delta \to \infty$ is shown in Fig. 10. The loss curves of Fig. 10 show maxima that shift as a function of n_r/n_1. The mode loss $2\alpha D$ is not an oscillating function of the jacket thickness if $n_r/n_1 < 1$. The curves shown in Figs. 7 through 9 are thus not the loss minima but the actual loss values as functions of Δ/D.

The dependence on the kD parameter for $n_r/n_1 < 1$ can be seen from

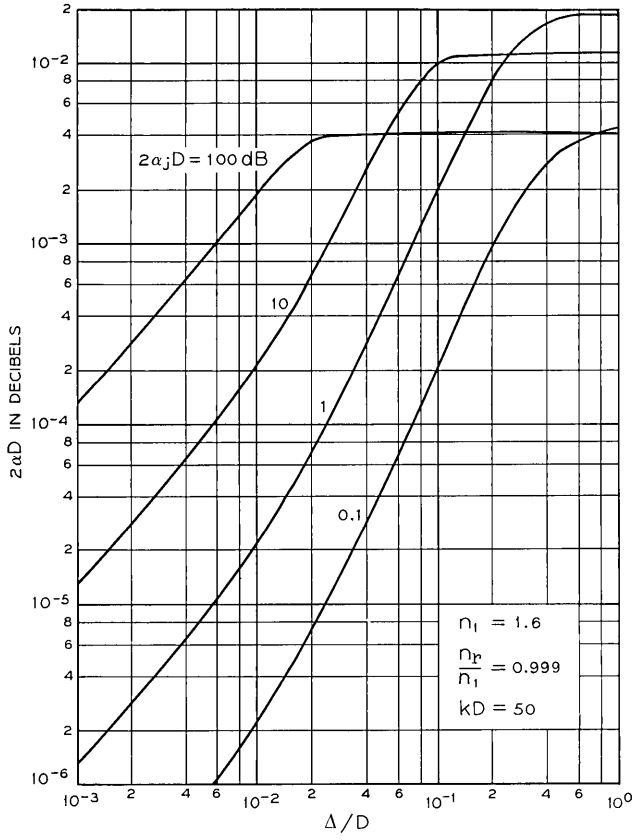


Fig. 7—Cladding mode losses as functions of the jacket thickness for $n_r/n_1 = 0.999$.

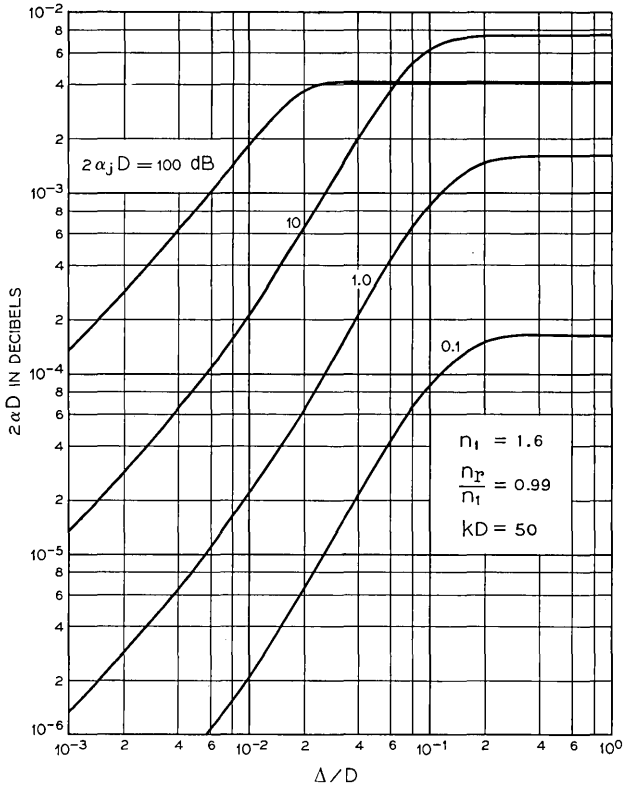


Fig. 8—Cladding mode losses as functions of the jacket thickness for $n_r/n_1 = 0.99$.

the figures to be as $(kD)^{-3}$ for $\Delta \to \infty$ while the dependence for small values of Δ/D is as $(kD)^{-2}$.

Even for jackets with a refractive index whose real part is smaller than the cladding index, losses well in excess of 1 dB/m are obtainable. However a comparison of the two cases, $n_r/n_1 > 1$ and $n_r/n_1 < 1$, shows that jackets with a refractive index higher than the cladding index provide more losses to the cladding modes.

Lossy jackets need not be very thick to be effective. For example, for $kD = 50$ and $\Delta/D = 0.1$ (for $D = 10\mu$m this would correspond to a jacket 1μm thick), we could choose $n_r/n_1 = 1.1$ and obtain a cladding mode loss of $2\alpha D = 2 \times 10^{-3}$ dB (or 200 dB/m) with a total jacket loss of $2\alpha_j\Delta = 1$ dB or $2\alpha_j D = 10$ dB. Figure 6 shows that the curve with $n_r/n_1 = 1.1$ is still constant at $2\alpha_j D = 10$ dB so that Fig. 5 is applicable

for a jacket loss of that magnitude. Increasing the jacket thickness to much larger values would triple the cladding mode losses (in dB) in this case. An increase of kD to $kD = 500$ reduces the loss to 7.2×10^{-5} dB (or 7.2 dB/m).

IV. CONCLUSIONS

Losses for cladding modes have been calculated with the help of a slab waveguide model with a lossy jacket. It has been found that the cladding mode losses are maximized if the real part of the refractive index of the jacket material equals the refractive index of the cladding material. However, high cladding mode losses can be achieved with jackets whose real part of the refractive index is either lower or higher than that of the cladding material. Losses in excess of 1 dB/m are easily

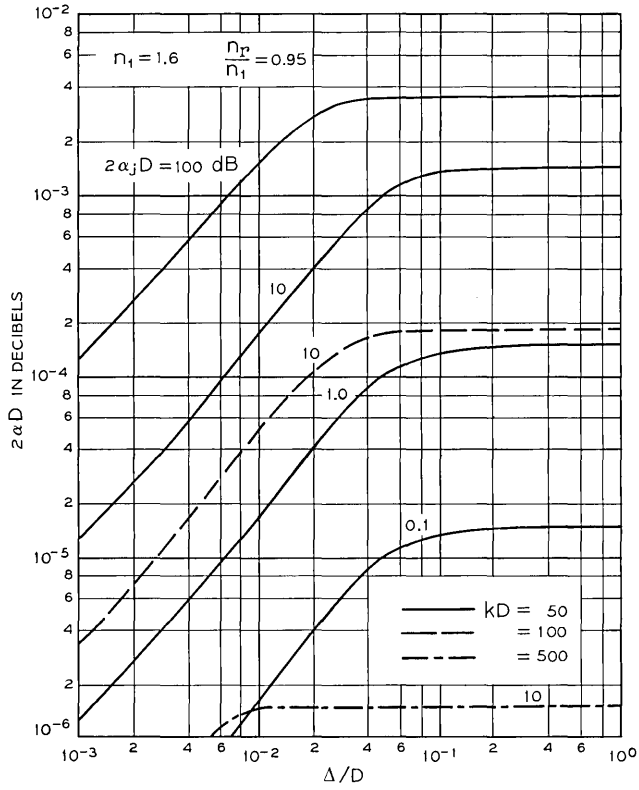


Fig. 9—Cladding mode losses as functions of the jacket thickness for $n_r/n_1 = 0.95$.

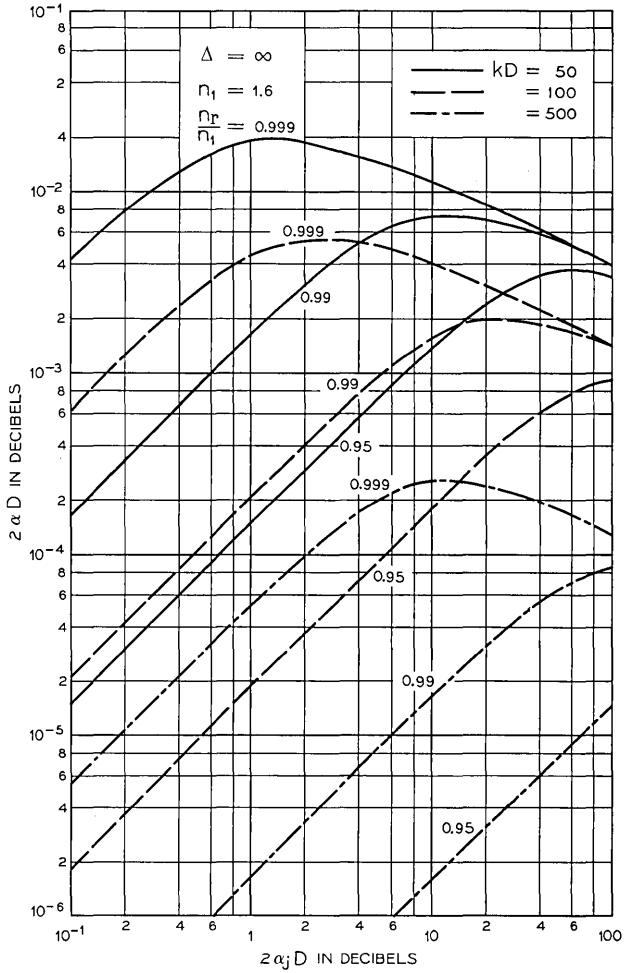


Fig. 10—Cladding mode losses for infinitely thick jacket as functions of the jacket loss parameter $2\alpha_j D$ for $n_r/n_1 < 1$.

achieved. A careful design should make it possible to obtain cladding mode losses between 100 and even up to 1000 dB/m, depending on the kD value at which the fiber is operated.

REFERENCES

1. Clarricoats, P. J. B., and Chan, K. B., "Excitation and Propagation of Modes of a Multilayer Fiber," Elec. Letters, *6*, No. 23 (November 12, 1970), pp. 750–752.

2. Roberts, R., "Propagation Characteristics of Multimode Dielectric Waveguides at Optical Frequencies," Conf. Trunk Telecommun. by Guided Waves, London, 1970, Conf. Digest, pp. 39–50.
3. Marcuse, D., "The Coupling of Degenerate Modes in Two Parallel Dielectric Waveguides," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 1791–1816.
4. Marcuse, D., "Crosstalk Caused by Scattering in Slab Waveguides," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 1817–1831.
5. Snyder, A. W., and De La Rue, R., "Asymptotic Solutions of Eigenvalue Equations for Surface Waveguide Structures," IEEE Trans. on Microwave Theory and Techniques, *MTT-18*, No. 9 (September 1970), pp. 650–651.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Improved Intersymbol Interference Error Bounds in Digital Systems

By Y. S. YEH and E. Y. HO

(Manuscript received April 21, 1971)

A thorough solution to the problem of determining the error rate of a digital communication system with intersymbol interference and additive Gaussian noise is presented in this paper. The solution achieves for the first time a combination of computational simplicity and a high degree of accuracy, and is obtained by deriving tight upper and lower bounds on the error rate. It is shown that, for a system with a normalized peak distortion less than unity, these bounds can be made to differ by an arbitrarily small amount. The numerical evaluation of the bounds takes less than one second on the GE-Mark II time-sharing system for almost all the cases.

Examples are given for 2M-ary digital systems to demonstrate the accuracy and computational efficiency of our method. The results show that our estimates of error rate are generally orders of magnitude better than the Chernoff bound. For example, in the case of an ideal bandlimited system [(sin t)/t pulse shape] with a signal-to-noise ratio of 16 dB and a sampling instant deviation of 0.05 from the optimum value, the lower and upper bounds on the error rate are 1.1 $\times$ 10^{-8} and 1.2 $\times$ 10^{-8}, respectively.

This method can also be applied to the calculation of the performance of certain phase-shift-keyed systems and certain systems with co-channel interference.

I. INTRODUCTION

In many cases the transmission efficiency of a digital system is largely limited by intersymbol interference rather than by additive noise. Intersymbol interference may result from imperfect design of the filters, distortion in the transmission channel, nonideal sampling instant, or nonideal demodulating carrier phase. In analyzing such a digital data system, it is important to determine the system error rate due to intersymbol interference and additive noise.

Various methods[1-7] to evaluate the error rate have been proposed. They provide either a loose upper bound of the error rate or the error rate of a channel with truncated impulse response.

In this paper we present a simple method to evaluate both an upper and a lower bound of the error rate without invoking the finite pulse-train approximation. Furthermore, it is shown that for a system with a normalized peak distortion less than unity, the upper and lower bounds can be made arbitrarily close thus obtaining an accurate estimate of the error rate of the system. This method can be applied to $2M$-ary AM and coherent phase-shift-keyed systems.

The data system model will be described briefly in Section II. Various proposed techniques to evaluate the error probability and their drawbacks are discussed in Section III. In Section IV, we will present new upper and lower bounds and the computation of the bounds by a series expansion. Applications and the convergence properties of the bounds are described in Section V. Throughout, additive Gaussian noise and independence of information digits are assumed.

II. BRIEF DESCRIPTION OF THE SYSTEM

A simplified block diagram of a digital AM data system is shown in Fig. 1. We assume that an impulse $\delta(t)$ having amplitude a_l is transmitted through the channel every T seconds. The system transfer function is

$$R(\omega) = S(\omega)T(\omega)E(\omega). \tag{1}$$

In the absence of channel noise, a sequence of input signals,

$$\sum_{l=-\infty}^{\infty} a_l \, \delta(t - lT), \tag{2}$$

will generate a corresponding output sequence,

$$\sum_{l=-\infty}^{\infty} a_l r(t - lT), \tag{3}$$

where $r(t)$ is the Fourier transform of $R(\omega)$, $\{a_l\}$ is a sequence of independent random variables, and $a_l = \pm 1, \pm 3, \cdots \pm (2M - 1)$ with equal probability for all integers, l. We also assume that additive Gaussian noise is present in the system. Thus the corrupted received sequence at the input to the receiver detector is

$$y(t) = \sum_{l=-\infty}^{\infty} a_l r(t - lT) + n(t), \tag{4}$$

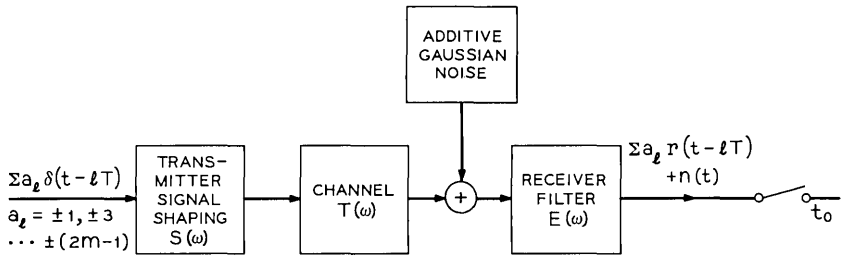


Fig. 1—Simplified block diagram of a $2M$-ary data system.

where $n(t)$ is additive Gaussian noise with power σ^2 watts. At the detector, $y(t)$ is sampled every T seconds to determine the amplitude of the transmitted signal. At sampling time t_0 , the sampled signal is

$$y(t_0) = a_0 r(t_0) + \sum_{\substack{l=-\infty \\ l\neq 0}}^{\infty} a_l r(t_0 - lT) + n(t_0). \tag{5}$$

The first term is the desired signal while the second and the third terms represent the intersymbol interference and the Gaussian noise respectively.

The set of slicing levels is[1]

$$0,\ \pm 2r(t_0),\ \pm 4r(t_0),\ \cdots,\ \pm(2m-2)r(t_0). \tag{6}$$

Based on the decision levels given by equation (6), for a particular transmitted signal level, a_0 , the conditional error probability is

$$P_r(e/a_0) = \begin{cases} P\{y(t_0) \geqq -2(m-1)r(t_0)\}, & a_0 = -(2m-1) \\ P\{y(t_0) \leqq 2(m-1)r(t_0)\}, & a_0 = 2m-1 \\ P\{[y(t_0) \geqq (a_0+1)r(t_0)]U[y(t_0) \leqq (a_0-1)r(t_0)]\}, & \\ & a_0 \neq \pm(2m-1), \end{cases} \tag{7}$$

where AUB is the union of the events A and B.

Substituting equation (5) into (7), we obtain

$$P_r(e/a_0) = \begin{cases} P\{\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \geqq r(t_0)\}, & a_0 = -(2m-1) \\ P\{\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \leqq -r(t_0)\}, & a_0 = 2m-1 \\ P\{[\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \geqq r(t_0)]U & \\ \cdot[\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \leqq -r(t_0)]\}, & a_0 \neq \pm(2m-1). \end{cases} \tag{8}$$

Since $\sum_{l\neq 0} a_l r(t_0 - lT)$ and $n(t_0)$ are equally likely to be positive or negative, equation (8) reduces to

$$P_r(e/a_0) = \begin{cases} P\{\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \geqq r(t_0)\}, & a_0 = \pm(2m-1) \\ 2P\{\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \geqq r(t_0)\}, & a_0 \neq \pm(2m-1). \end{cases} \tag{9}$$

The error rate of the system is

$$\begin{aligned} P_e &= \sum_{\text{all } a_0} P_r(e/a_0) P_r(a_0) \\ &= [(2m-1)/m] P\{\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \geqq r(t_0)\}. \end{aligned} \tag{10}$$

We notice that in equation (10) the variables m, a_l , and $n(t_0)$ have already been defined. The sequence $r(t_0 - lT)$ is assumed to be known* in the following sense:

$$r(t_0 - lT) \quad \text{is finite and known} \quad \forall l \in S_N , \tag{11}$$

where S_N is a set of $N + 1$ distinct integers (including $l = 0$) and†

$$\sum_{l\notin S_N} r^2(t_0 - lT) = \sigma_r^2 < \infty. \tag{12}$$

Define

$$X = \sum_{l\neq 0} a_l r(t_0 - lT). \tag{13}$$

From equation (12) we conclude that the infinite sum X converges absolutely to a random variable and equation (10) can be alternately written as

$$P_e \int_{\text{all } X} (2\pi\sigma^2)^{-1} \int_{-\infty}^{0} \exp\{-[y - r(t_0) + X]^2/2\sigma^2\}\, dy\, dF(X). \tag{14}$$

III. REVIEW OF EXISTING METHODS

The existing methods of evaluating equation (10) can be divided into the following categories.

3.1 *Worst Case Estimate*

A worst case sequence[1] or "eye pattern" analysis is frequently used to analyze a data system. The error probability is estimated by setting

* The sequence $r(t_0 - lT)$ is either experimentally determined or calculated through the system transfer function.

† σ_r^2 is obtained through the application of Parseval's theorem to equivalent Nyquist pulse (p. 47, Ref. 1).

$\sum_{l \neq 0} a_l r(t_0 - lT)$ to its worst case value in equation (10). In many cases, this estimate is exceedingly pessimistic since the occurrence of such a worst case sequence is extremely rare.

3.2. *Chernoff Bound*

Recently, Saltzberg[2] and Lugannani[3] applied the Chebyshev inequality to equation (10) to obtain the upper bound on error probability. We have shown in Ref. 6 that these upper bounds are in many cases still too pessimistic by orders of magnitude.

3.3. *Finite Truncated Pulse Train Approximation*[4,5]

When $r(t)$ decreases rapidly relative to the sampling period T, we may approximate the channel by a finitely truncated pulse train. The error rate can be calculated by enumerating all the possible combinations of intersymbol interference. However, since each calculation of the conditional error probability takes a great deal of computer time, the number of m^N must be held to several thousand.[1] This limitation leads to a poor approximation of the true channel, and the error probability so obtained is not very useful. Recently, Hill[8] has reported that by computer simulation of the density function of X_N , the computation time can be reduced.

3.4 *Series Expansion Method*

Recently, Ho and Yeh[6] and, independently, Shimbo and Celebiler[7] discovered that equation (10) can be calculated in terms of an absolutely convergent series involving moments of the intersymbol interference.* Furthermore, the moments can be obtained readily through recurrence relations, and the computation time is greatly reduced. A better approximation of the real channel can be obtained by increasing the number of terms in the pulse train approximation. However, the error in the P_e estimate introduced by the truncation of the system impulse response is still unknown.

IV. ERROR BOUNDS AND COMPUTATION TECHNIQUES

In this section we shall derive new upper and lower bounds on the error rates and define the range of applicability of our method. No truncation of the intersymbol interference is required. Furthermore, this method will give an accurate estimate of the error rate with a negligible amount of computation time.

* Only truncated pulse train approximations are considered.

4.1 *Upper and Lower Bound of P_e*

Let the intersymbol interference be partitioned into two disjoint sets where

$$X_N = \sum_{\substack{l \neq 0 \\ l \in S_N}} a_l r(t_0 - lT), \tag{15a}$$

and

$$X_R = \sum_{l \notin S_N} a_l r(t_0 - lT). \tag{15b}$$

Equation (14) can be rewritten as

$$P_e = [(2m - 1)/m] \int_{\text{all } X_N} \int_{\text{all } X_R} (2\pi\sigma^2)^{-\frac{1}{2}} \cdot \int_{-\infty}^{0} \exp \{-[y - r(t_0) + X_N + X_R]^2/2\sigma^2\}\, dy\, dF(X_N)\, dF(X_R). \tag{16}$$

Proposition 1: P_e is lower bounded by

$$P_L = [(2m - 1)/m] \int_{\text{all } X_N} (2\pi\sigma^2)^{-\frac{1}{2}} \cdot \int_{-\infty}^{0} \exp [-\{y - r(t_0) + X_N\}^2/2\sigma^2]\, dy\, dF(X_N), \tag{17}$$

provided the truncated system has an "open eye pattern," i.e.,

$$r(t_0) - \sum_{\substack{l \in S_N \\ l \neq 0}} | r(t_0 - lT) | \geqq 0. \tag{18}$$

Proof: The complementary error function is concave upwards for negative values of its argument and satisfies the following relationship:

$$\tfrac{1}{2} \operatorname{erfc} (z + \alpha) + \tfrac{1}{2} \operatorname{erfc} (z - \alpha) \geqq \operatorname{erfc} (z), \quad z \leqq 0. \tag{19}$$

Since X_R is symmetrically distributed around zero and X_N satisfies equation (18), we obtain, by applying equation (19), that

$$\int_{\text{all } X_R} \int_{-\infty}^{0} \exp \{-[y - r(t_0) + X_N + X_R]^2/2\sigma^2\}\, dy\, dF(X_R) \geqq \int_{-\infty}^{0} \exp \{-[y - r(t_0) + X_N]^2/2\sigma^2\}\, dy. \tag{20}$$

Substituting equation (20) into equation (16), we obtain the lower bound of equation (17).

Proposition 2: P_e is upper bounded by

$$P_u = [(2m-1)/m] \int_{\text{all } X_N} (2\pi\sigma^2)^{-\frac{1}{2}} \cdot \int_{-\infty}^{0} \exp\{-[y - r(t_0) + X_N]^2/2\sigma_1^2\}\, dy\, dF(X_N), \tag{21a}$$

where

$$\sigma_1^2 = \sigma^2(1 - \sigma_R^2/\sigma^2)^{-1}, \tag{21b}$$

$$\sigma_R^2 = (1/3)(2m-1)(2m+1)\sigma_r^2, \tag{21c}$$

and σ_r^2 is defined in equation (12).

Proof: Applying the following inequality,

$$\exp\{-X_R^2/2\sigma^2\} \leqq 1, \tag{22}$$

to equation (16), we obtain

$$P_e \leqq [(2m-1)/m] \int_{\text{all } X_N} \int_{-\infty}^{0} \int_{\text{all } X_R} (2\pi\sigma^2)^{-\frac{1}{2}} \cdot \exp\{-[y - r(t_0) + X_N]^2/2\sigma^2\} \cdot \exp\{-[y - r(t_0) + X_N]X_R/\sigma^2\}\, dF(X_R)\, dy\, dF(X_N). \tag{23}$$

Knowing from equation (15b),

$$X_R = \sum_{l \notin S_N} a_l r(t_0 - lT),$$

the average over X_R can be performed, we thus have

$$\int_{\text{all } X_R} \exp\{-[y - r(t_0) + X_N]X_R/\sigma^2\}\, dF(X_R) = \prod_{l \notin S_N} \langle \exp\{-[y - r(t_0) + X_N]a_l r(t_0 - lT)/\sigma^2\} \rangle_{a_l}, \tag{24a}$$

where $\langle g(x) \rangle_x$ means expectation of $g(x)$. It has been shown* that the following inequality holds.

$$\langle \exp\{a_l x\} \rangle_{a_l} \leqq \exp(x^2\sigma_{a_l}^2/2) = \exp\{x^2(2m-1)(2m+1)/6\}. \tag{24b}$$

Substituting equation (24b) into (24a) we obtain

$$\int_{\text{all } X_R} \exp\{-[y - r(t_0) + X_N]X_R/\sigma^2\}\, dF(X_R) \leqq \exp\{[y - r(t_0) + X_N]^2\sigma_R^2/2\sigma^4\}, \tag{24c}$$

* Appendix of Ref. 6.

where σ_R^2 is given by equation (21c). Substituting equation (24c) into (23) we obtain the upper bound of equation (21a).

It is interesting to note that the upper bound differs from the lower bound only through a modification of the noise power by the truncated terms. For a system with a peak distortion* less than unity, by taking the set S_N large enough σ_R^2 approaches zero, σ_1^2 approaches σ^2, and the upper bound converges to the lower bound. Therefore, the exact error probability can be located within a small range. The computation time involved for large enough N is rather minimal when a digital computer is used as will be illustrated in Section V.

4.2 *Evaluation of P_L and P_u*

We have already shown in Ref. 6 that equations (17) and (21) can be expanded into an absolutely convergent series involving moments of the truncated intersymbol interference.

The series expansion of equation (17) is

$$P_L = [(2m-1)/m]\ \text{erfc}\ [-r(t_0)/(2^{\frac{1}{2}}\sigma)] + [(2m-1)/m] \sum_{k=1}^{\infty} [(2k)!]^{-1}(2\sigma^2)^{-k}(\pi)^{-\frac{1}{2}} \exp(-r^2(t_0)/2\sigma^2) \cdot H_{2k-1}(r(t_0)/(2^{\frac{1}{2}}\sigma))M_{2k}, \tag{25}$$

where

H_{2k-1} is the Hermite polynomial,
M_{2k} is the $2k$th moment of the random variable X_N.

The series expansion of equation (21) is similar to equation (17),

$$P_U = (\sigma_1/\sigma)\Big\{[(2m-1)/2m]\ \text{erfc}\ [-r(t_0)/(2^{\frac{1}{2}}\sigma_1)] + [(2m-1)/m] \sum_{k=1}^{\infty} [(2k)!]^{-1}(2\sigma_1^2)^{-k}(\pi)^{-\frac{1}{2}} \exp[-r^2(t_0)/2\sigma_1^2] \cdot H_{2k-1}(r(t_0)/(2^{\frac{1}{2}}\sigma_1))M_{2k}\Big\}. \tag{26}$$

The moments (M_{2k}) can be obtained through the characteristic

* Normalized peak distortion (D_p) is defined as

$$D_p = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} |r(t_0 + nT)|/|r(t_0)|.$$

function of X_N without the explicit evaluation of the distribution function. The recurrence formula for M_{2k} is

$$M_{2k} = -\sum_{i=1}^{k} \binom{2k-1}{2i-1} M_{2(k-i)}(-1)^i \{2^{2i}[(2m)^{2i} - 1]/2i\} \mid B_{2i} \mid \cdot [\sum_{\substack{l \neq 0 \\ l \varepsilon S_N}} r(t_0 - lT)^{2i}], \tag{27}$$

where B_{2i} are the Bernoulli numbers.

4.3 *Truncation Error Bound of Series Expansion*

The error incurred by truncating the series of equation (25) at $(n-1)$ term is given by

$$R_n = [(2m-1)/m] \sum_{k=n}^{\infty} (2k!)^{-1}(2\sigma^2)^{-k}(\pi)^{-\frac{1}{2}} \exp [r^2(t_0)/2\sigma^2] \cdot H_{2k-1}[r(t_0)/(2^{\frac{1}{2}}\sigma)] \cdot M_{2k} . \tag{28}$$

Let

$$\lambda = \max \mid X_N \mid = (2m-1) \sum_{\substack{l \varepsilon S_N \\ l \neq 0}} \mid r(t_0 - lT) \mid . \tag{29}$$

It can be shown that the moments satisfy

$$M_{2k+2p} \leqq M_{2k}\lambda^{2p}, \quad p = 0, 1, 2, \cdots . \tag{30}$$

For $(2k-1) \gg x$, the Hermite polynomials are upper bounded by

$$\mid H_{2k-1}(x) \mid \leqq 2^{k-\frac{1}{2}}[(2k-3)!!] \sqrt{2k-1} \exp [x^2/2]. \tag{31}$$

Substituting equations (30) and (31) into equation (28) we obtain the following:

$$\begin{aligned} \mid R_n \mid &\leqq [(2m-1)/m](2\pi)^{-\frac{1}{2}} \exp [-r^2(t_0)/4\sigma^2] \\ &\quad \cdot M_{2n} \cdot (2\sigma^2)^{-n} \sum_{k=n}^{\infty} (k!)^{-1}(2k-1)^{-\frac{1}{2}}(\lambda^2/2\sigma^2)^{k-n} \\ &\leqq [(2m-1)/m](2\pi)^{-\frac{1}{2}} \exp [-r^2(t_0)/4\sigma^2] \\ &\quad \cdot M_{2n} \cdot (2\sigma^2)^{-n} \Bigg\{ \sum_{k=n}^{p-1} (k!)^{-1}(2k-1)^{-\frac{1}{2}}(\lambda^2/2\sigma^2)^{k-n} \\ &\quad + (P!)^{-1}(2p-1)^{-\frac{1}{2}}(\lambda^2/2\sigma^2)^{p-n}(1 - \lambda^2/2p\sigma^2)^{-1} \Bigg\} , \end{aligned} \tag{32}$$

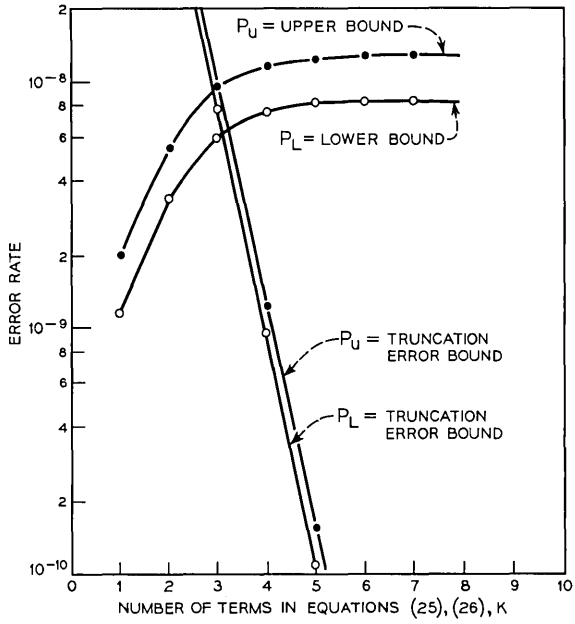


Fig. 2—The convergence of the series expansion method; $(\sin \pi t/T)/(\pi t/T)$ pulse, binary AM system, $t_0 = 0.05\ T$, SNR = 16 dB. The set S_N includes the first 12 sampling points around t_0.

where p is an integer which is chosen to satisfy $(\lambda^2/2p\sigma^2) < 1$. Similar truncation error bounds can be obtained for P_u .

V. APPLICATION

The error probability of a $2M$-ary digital AM system with an ideal band-limiting pulse signal operating over an ideal channel is calculated by equations (25) and (26) to determine the convergence of the method. The received binary pulse is assumed to be

$$r(t) = (\sin \pi t/T)/(\pi t/T). \tag{33}$$

The system SNR is defined by

$$\text{SNR} = \langle a_0^2 \rangle r^2(0)/\sigma^2. \tag{34}$$

The convergence of the series expansion method is illustrated in Fig. 2. The system is binary with the sampling instant deviated by $0.05T$ from its nominal sampling instant. The SNR is 16 dB. The set S_N includes 12 elements, i.e., $l = \pm 1, \pm 2, \cdots, \pm 6$. It is observed that

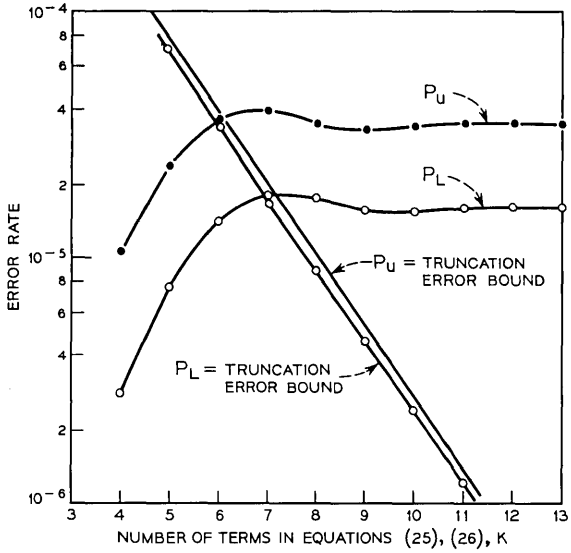


Fig. 3—Convergence of P_L as a function of number of terms in the series expansion; $(\sin \pi t/T)/(\pi t/T)$ pulse, 4-level digital AM system, $t_0 = 0.05\ T$, SNR = 23 dB. The set S_N includes the first 12 sampling points around t_0.

the series converges after 3 or 4 terms. A similar example is given in Fig. 3 for a 4-level system with SNR = 23 dB.

The convergence of the upper bound to the lower bound with increased size of S_N is illustrated in Figs. 4, 5, and 6 for binary and 4-level systems respectively.* It is observed that the two bounds indeed merge together as N is increased. The upper and lower bounds on the error rate were calculated using a program written for the GE-Mark II time-sharing system. For the examples given here, computation time was less than a second. The change of N from 6 to 30 hardly had any effect on the computation time which indicates that one should start with S_N sufficiently large such that σ_R^2 is small in comparison with σ^2, probably of the order of 0.2 σ^2 or smaller. Under this condition the upper and lower bound should be fairly close. As a comparison, the Chernoff bounds are also presented in Figs. 4, 5, and 6.

The method given here can also be applied to the calculation of the error rate of a coherent phase-shift-keyed system. The error rate calculation of a two and four phase system can be reduced to the basic

* The number of terms in the series expansion used to calculate the points in these figures are determined so that the truncation error is negligible.

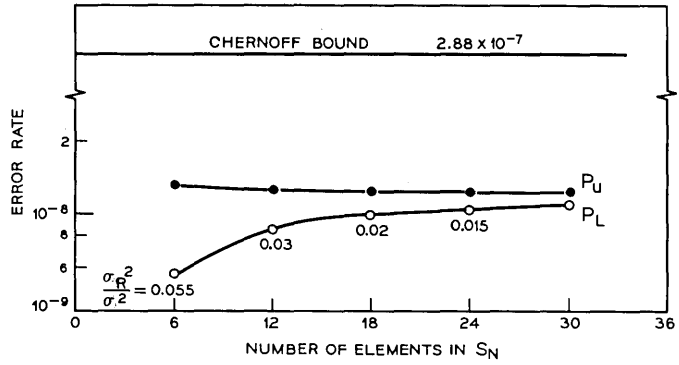


Fig. 4—Convergence of P_u to P_L; ideal band-limited pulse, binary system, SNR = 16 dB, $t_0 = 0.05\ T$.

formula, equation (10),[9] which then requires the determination of

$$P\{\sum_{l\neq 0} a_l r(t_0 - lT) + n(t_0) \geqq r(t_0)\}. \tag{35}$$

The method described here can then be applied. Similar applications can also be found in the error calculation of co-channel interference.

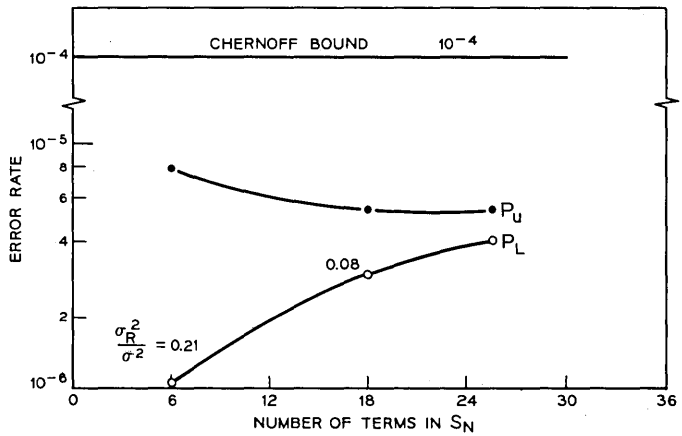


Fig. 5—Convergence of P_u to P_L; ideal band-limited pulse, binary system, SNR = 16 dB, $t_0 = 0.1\ T$.

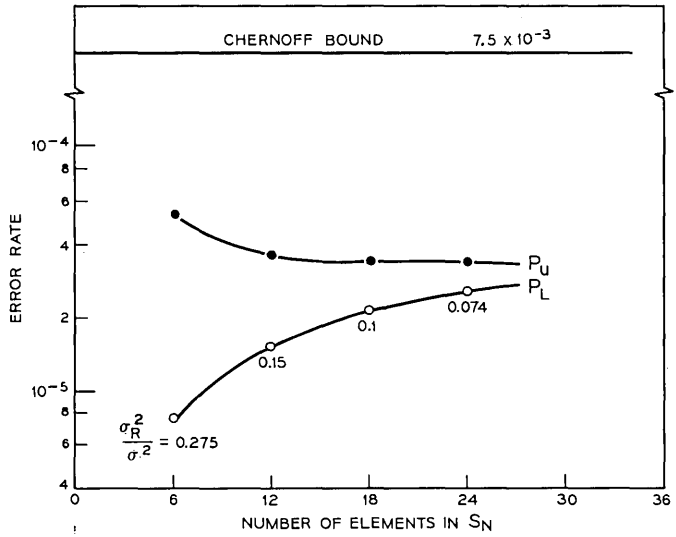


Fig. 6.—Convergence of P_u to P_L; ideal band-limited pulse, 4-level system, SNR = 16 dB, $t_0 = 0.05\ T$.

VI. CONCLUSIONS

We have presented a method to calculate the error rate of a coherent digital system subject to intersymbol interference and additive Gaussian noise. The error rate for a system with a peak distortion less than unity can be determined to arbitrary accuracy through the calculation of an upper bound and a lower bound of the error rate. The computation time involved (less than one second on the GE-Mark II time-sharing system) is many orders of magnitude shorter than the time required by the straightforward calculation of all the possible states. On the other hand, the results are generally much more accurate than the results obtained through the application of the Chernoff bound.

VII. ACKNOWLEDGEMENT

The authors wish to thank J. Salz and D. A. Spaulding for many helpful discussions.

REFERENCES

1. Lucky, R. W., Salz, J., and Weldon, E. J., Jr., *Principles of Data Communication*, New York: McGraw-Hill Book Company, 1968.

2. Saltzberg, B. R., "Intersymbol Interference Error Bounds With Application to Ideal Bandlimited Signaling," IEEE Trans. Information Theory, *IT-14*, July 1968, pp. 563–568.
3. Lugannani, R., "Intersymbol Interference and Probability of Error in Digital Systems," IEEE Trans. Information Theory, *IT-15*, November 1969, pp. 682–688.
4. Aein, J. M., and Hancock, J. C., "Reducing the Effects of Intersymbol Interference With Correlation Receivers," IEEE Trans. Information Theory, *IT-9*, July 1963, pp. 167–175.
5. Aaron, M. R., and Tufts, D. W., "Intersymbol Interference and Error Probability," IEEE Trans. Information Theory, *IT-12*, January 1966, pp. 26–34.
6. Ho, E. Y., and Yeh, Y. S., "A New Approach for Evaluating the Error Probability in the Presence of Intersymbol Interference and Additive Gaussian Noise," B.S.T.J., *49*, No. 9 (November 1970), pp. 2249–2265.
7. Celebiler, M. I., and Shimbo, O., "Intersymbol Interference Considerations in Digital Communications," ICC Record, June 1970, pp. 8-1–8-10.
8. Hill, F. S., "The Computation of Error Probability for Digital Transmission," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 2055–2077.
9. Prabhu, V. K., "Error Rate Considerations for Coherent Phase-shift Keyed Systems with Co-Channel Interference," B.S.T.J., *48*, No. 3 (March 1969), pp. 743–768.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Dual Frequency Measurements of Rain-Induced Microwave Attenuation on a 2.6-Kilometer Propagation Path

By R. A. SEMPLAK

(Manuscript received April 9, 1971)

The ratio of attenuations measured simultaneously at 18.5 and 30.9 GHz on a common 2.6-km path during 1970 in New Jersey show that, with exceptions for special cases, the Laws and Parsons drop-size distribution adequately represents the rain on terrestrial microwave communications paths. The ratio derived from cumulative distributions of attenuation is also discussed. A rain-induced attenuation of 30 dB is found to be exceeded 0.001 and 0.01 percent of the time at 18.5 and 30.9 GHz, respectively. These results are in good agreement with attenuations derived from rain rates obtained on the Holmdel rain gauge network.

I. INTRODUCTION

Rain-induced attenuation at centimeter and millimeter wavelengths is a function of drop-size; conversely, a measure of drop-size can be obtained from multi-wavelength measurements over a common propagation path. As shown in Fig. 1, a 2.6-kilometer path, shared by both 18.5 and 30.9 GHz, extends to the southeast of Crawford Hill, Holmdel, New Jersey. Both of the receivers are located in the same building at Crawford Hill; the equipment has been reported elsewhere,[1,2] it will not be discussed here.

In addition to obtaining a measure of the drop-size distributions by examining the ratio of the 30.9-GHz attenuation to the 18.5-GHz attenuation during rainstorms, sufficient data are available for a statistical examination of the rain-induced attenuations for both frequencies. In the following, the cumulative distributions of attenuation at 18.5 and 30.9 GHz will be discussed individually and then related one to the other.

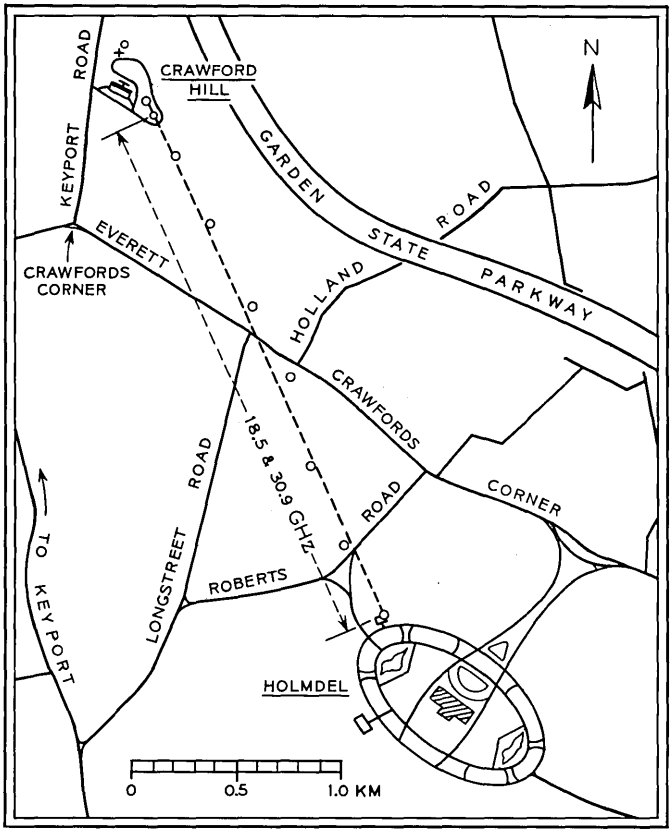


Fig. 1—Map showing the propagation path of 2.6 km used at 30.9 and 18.5 GHz.

II. VARIABILITY IN DROP-SIZE DISTRIBUTION

The variability of drop-size can be assessed by examining the data from the common propagation path, shown in Fig. 2. In this figure, the ratios of 30.9-GHz to 18.5-GHz attenuation are plotted versus the total path attenuation at 18.5 GHz (the upper plot) and versus the total path attenuation at 30.9 GHz (the lower plot). In both plots, the dashed line is the theoretical ratio of 30.9/18.5-GHz attenuation obtained using a Laws and Parsons drop-size distribution,[3,4] assuming a uniform rainfall over the path. It is apparent from the data in both plots of Fig. 2 that the theoretical ratios derived from Laws and Parsons represent the body of the data rather well, especially for path

attenuations exceeding 10 dB at 30.9 GHz; in fact the least squares fit to the data is not shown because it lies so close to the theoretical curve. However, for attenuations less than about 10 dB at 30.9 GHz there is a large scatter in the measured ratios; much of the scatter is caused by two particular storms.

The two specific sets of data are shown in the two plots of Fig. 2

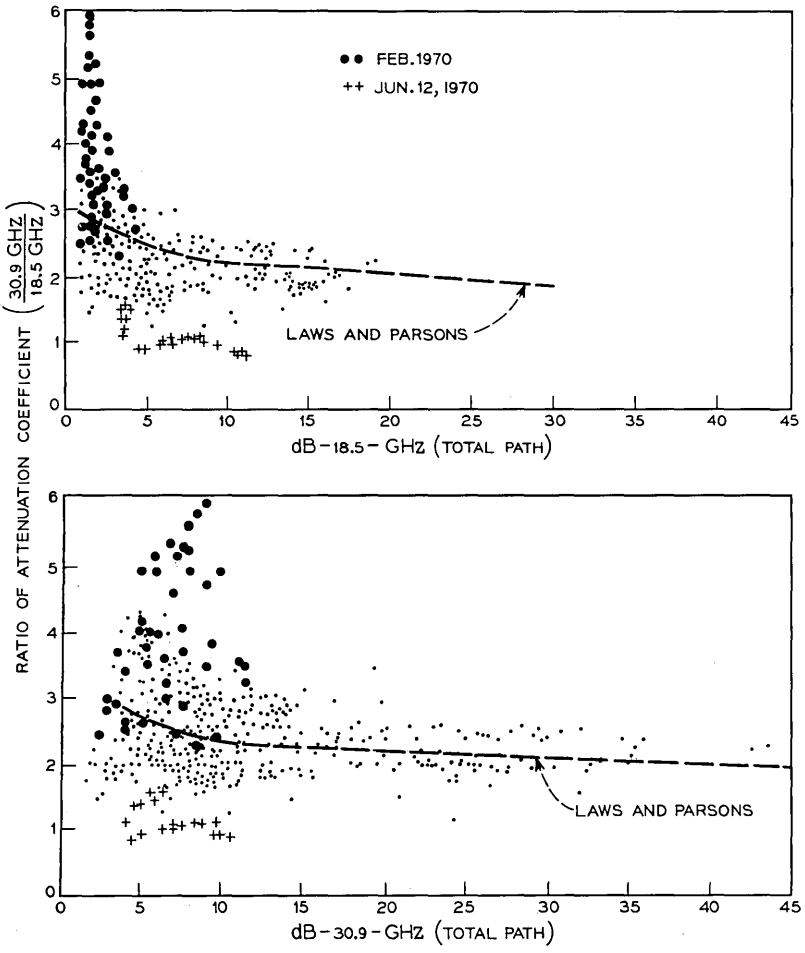


Fig. 2—The ratio of attenuation at 30.9 GHz to that at 18.5 GHz plotted, in the upper graph, versus the 18.5-GHz total path attenuation and, in the lower, versus the 30.9-GHz total path attenuation. The dashed curve in each plot is the theoretical ratio using Laws and Parsons drop-size distribution.

by distinctive symbols; these data were measured during periods of the month of February, and on the twelfth day of June, 1970. Suffice it to say that these are special cases where the large ratios obtained in February are due to light rains composed of more small drop diameters than accounted for by Laws and Parsons, whereas the lower ratios obtained in June are the results of drops larger than that given by the Laws and Parsons drop-size distribution. If these two special cases are removed, the remaining data are well represented by the line labeled Laws and Parsons. The spread in the measured data is due not only to variations in drop-size distributions but also to a lack of homogeneity in the rainfall intensity along the propagation path.

III. CUMULATIVE DISTRIBUTIONS OF 30.9- AND 18.5-GHZ RAIN-INDUCED ATTENUATION

A sample of 6432 hours of recording on the dual-frequency experiment was obtained during the period January 1, 1970 through September 25, 1970. During this period a total of thirty-six individual showers occurred. Of these thirty-six, attenuation measurements were obtained at both 18.5 and 30.9 GHz for thirty-four. The remaining two storms were missed; in one instance, there was an equipment failure at 18.5 GHz, in the other, a failure at 30.9 GHz.

Percent-of-time distributions for 18.5- and 30.9-GHz rain-induced attenuations are shown in Fig. 3 (by the solid curves) for the recording period of 6432 hours, *not* a full year. However, based on earlier work,[2] this period is known to include the months of heaviest rain rate, therefore the sample contains essentially all of the high attenuations which occur in a full year.

Examination of the solid curves of Fig. 3 shows that for this 2.6-kilometer path and a 0.01-percent probability level, the 18.5- and 30.9-GHz rain-induced attenuations exceeded 15 dB and 30 dB, respectively. Thus, these levels of attenuation were exceeded for a total of 38 minutes of the 6432-hour recording period. Similarly, if one examines the 0.001-percent probability level, which for this sample is equivalent to about 4 minutes, one finds that the attenuation at 18.5 GHz exceeded 30 dB whereas at 30.9 GHz, using some extrapolation,* the attenuation exceeds 57 dB.

As stated earlier, the likelihood that the 18.5- or 30.9-GHz attenuation would have exceeded 30 and 57 dB for the remainder of the year, beyond the 6432-hour sample, is remote. If the percentage level is

* The measuring range at 30.9 GHz is 48 dB.

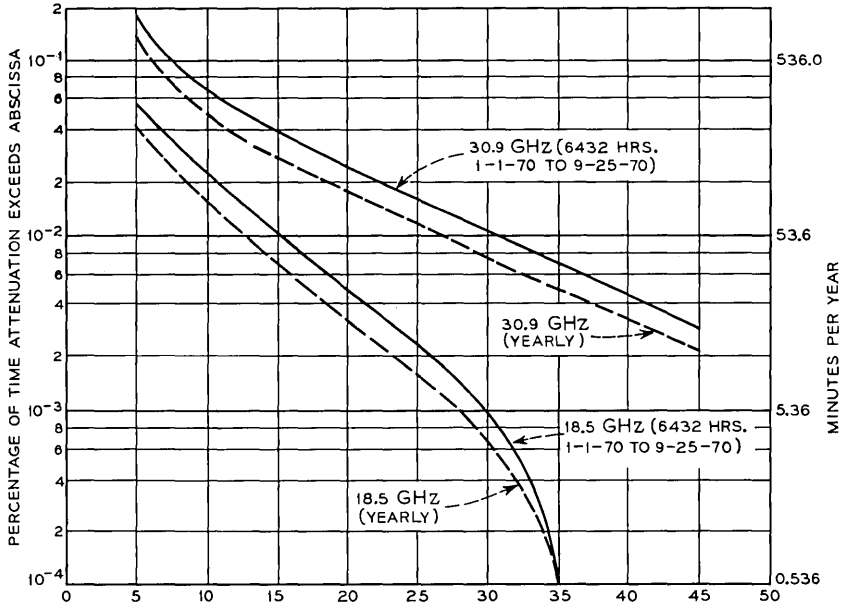


Fig. 3—Percent-of-time distributions of the rain-induced attenuations for both 18.5 and 30.9 GHz. The solid curves are based on 6432 hours of data, the dashed curves are projected yearly distributions.

computed on a yearly basis, it is found that the attenuation at 18.5 GHz exceeds 28 dB 0.001 percent of the time, approximately 5 minutes per year, and the attenuation at 30.9 GHz exceeds 54 dB for 5 minutes of the year. The yearly attenuation distributions are shown by the dashed curves of Fig. 3.

An interesting comparison can be made between the data points just discussed and the attenuation computed[5] from the rain rates measured on the Holmdel rain gauge network. Figure 4 is a reproduction from Ref. 5 of the curves of computed rain attenuations exceeded 0.001 percent of the year (Fig. 4a) and 0.01 percent of the year (Fig. 4b), as a function of path length. These computed curves are based on average path rain rates from the 1967 rain gauge network data. The current 1970 data at the 0.001-percent probability level for both 18.5 and 30.9 GHz are shown as crosses on Fig. 4a;* the agreement is quite

* Figure 4 also shows measurements for the two probability levels at 18.5 and 30.9 GHz taken on 6.4- and 1.9-km paths during 1967, 1968, and 1969.[2]

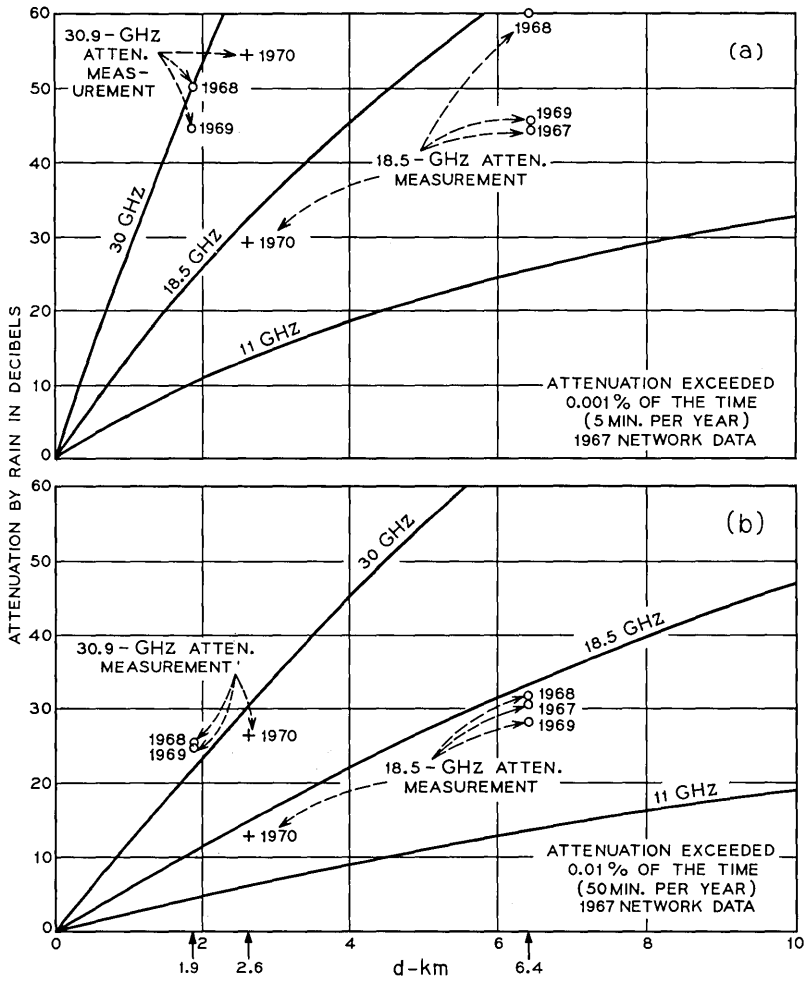


Fig. 4—Attenuation as a function of path length at 11, 18.5, and 30 GHz (1967 Holmdel rain gauge network data) for (a) 0.001 percent probability and (b) 0.01 percent probability along with measurements on paths of length 1.9 km (30.9 GHz), 2.6 km (18.5 and 30.9 GHz), and 6.4 km (18.5 GHz).

good at 18.5 GHz but at both frequencies the computed curves predict a value somewhat larger than the measurements.

A similar comparison at the 0.01-percent probability level can be made if one is willing to assume that no attenuations exceeding 15 and 30 dB at 18.5 and 30.9 GHz, respectively, occurred for the re-

mainder of 1970, beyond the 6432-hour recording period. This is not an overly hazardous assumption since experience has shown that the tail of the distribution for the measuring period discussed here contains essentially all the high attenuation data for a year. This assumption was made with the 1970 data for both 18.5 and 30.9 GHz and the results (12.5 and 27 dB) for the 0.01-percent probability level are shown by crosses on Fig. 4b. Agreement between computations and measured data is fairly good, the curves computed from the rain gauge network data again being somewhat conservative.

IV. RELATIONSHIP BETWEEN THE TWO ATTENUATION DISTRIBUTIONS

In view of the fact that the 18.5- and 30.9-GHz attenuations are measured on a common path, the received signal levels at the two frequencies are always attenuated by the same volume of rain and therefore fade together. One is justified in attaching significance to the ratio of the attenuations at a given percent-of-time level.

Ratios of 30.9- and 18.5-GHz attenuation for various percentage levels (from Fig. 3) are plotted as solid dots in Fig. 5 versus the total path attenuation at 18.5 GHz; the theoretical ratio[4] of 30.9/18.5-GHz attenuation obtained using a Laws and Parsons drop-size distribution is plotted as the solid curve on this figure. Again, as in Fig. 2, there is

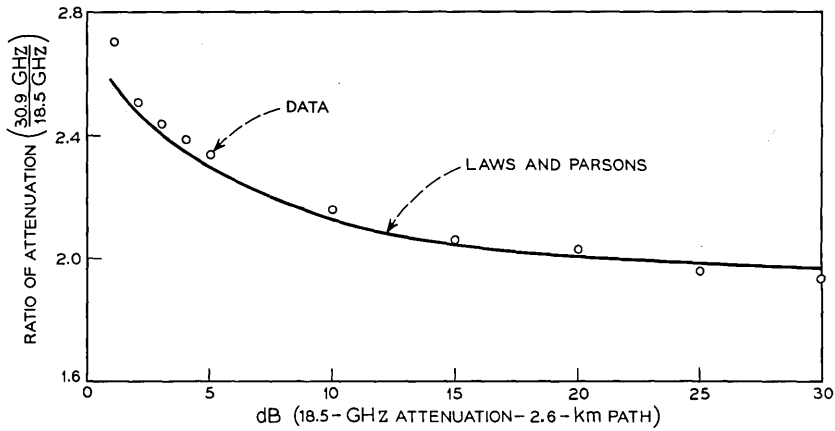


Fig. 5—The ratio of measured attenuation from the 30.9-GHz and 18.5-GHz distributions of Fig. 3, plotted as solid dots, versus the measured 18.5-GHz total path attenuation. The solid curve is the theoretical ratio using Laws and Parsons drop-size distributions.

good agreement between the ratios obtained from the measured attenuation and the theoretical distributions based on the Laws and Parsons drop-size distribution; thus, it appears that the Laws and Parsons distribution holds on the average. The least squares fit to the measured ratios has not been plotted in Fig. 5 because of its proximity to the computed curve.

V. CONCLUSION

Our measurements of the instantaneous ratio of 30.9- to 18.5-GHz attenuation appear to be consistent (Fig. 2) with computations based on the Laws and Parsons raindrop-size distribution. The attenuation ratios have also been derived from the cumulative distributions of attenuation (Fig. 3) and these are in very good agreement (Fig. 5) with the theoretical computations. Relative values obtained from cumulative distributions of attenuation at two frequencies (or more) on common paths therefore proves to be a valuable technique for probing the rain environment.

REFERENCES

1. Semplak, R. A., and Turrin, R. H., "Some Measurements of Attenuation by Rainfall at 18.5 GHz," B.S.T.J., *48,* No. 6 (July–August 1969), pp. 1767–1787.
2. Semplak, R. A., "The Influence of Heavy Rainfall on Attenuation at 18.5 and 30.9 GHz," IEEE Trans. Antennas and Propagation, *AP-18,* No. 4 (July 1970), pp. 507–511.
3. Laws, J. O., and Parsons, D. A., "The Relation of Raindrop-Size to Intensity," Trans. Am. Geophysical Union, *24,* 1943, pp. 432–460.
4. Setzer, D. E., "Computed Transmission Through Rain at Microwave and Visible Frequencies," B.S.T.J., *49,* No. 8 (October 1970), pp. 1873–1892.
5. Hogg, D. C., "Statistics on Attenuation of Microwaves by Intense Rain," B.S.T.J., *48,* No. 9 (November 1969), pp. 2949–2962.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Optical Modulation at High Information Rates

By GERARD WHITE

(Manuscript received March 18, 1971)

This paper describes an optical modulation scheme with a demonstrated information rate of 1 Gb/s. The modulation scheme employs an electronic multiplexing technique to obtain the high data rate from four 250-Mb/s channels. Efficient interfacing between the gas laser beam and the high-speed electronic circuits is achieved using a traveling-wave type modulator structure. General considerations in optical modulation systems are also included.

I. INTRODUCTION

One of the major problems in engineering is that of interfacing between two differing technologies: This problem is particularly severe in laser communications systems where solid state electronics is to be interfaced with optics. It is, perhaps, at this technological interface that the early hopes of approaching the Shannon-Weaver information rates of laser systems have been limited. While electro-optic crystals have been shown to have modulation mechanisms operating in excess of 10 GHz[1] and photodetectors with gain-bandwidth products of the order of 100 GHz have been produced,[2] laser communications systems, so far, have demonstrated information rates significantly less than the capabilities of these components. Significant in the bandwidth limitation of demonstrated laser communications systems have been the demands on solid state circuit performance[3] imposed by the optical system and, in particular, the modulator. In this paper an optical modulation scheme is described in which system variations and subsystem configurations have been adopted in an attempt to relax the severe circuit performance requirements so that system information rate capabilities approach the bandwidth capabilities of the opto-electronic transducers. The system described, using an argon laser, has a demonstrated information rate of 1 Gb/s and some

techniques for improving upon this rate are proposed. Indicated system noise levels are observed to be low so that it may be possible to transmit a multilevel type code which could at least triple the information rate.

The paper, while describing state of the art results in wide-bandwidth optical communications, also provides tutorial matter in connection with general considerations in the system design and evaluation. The paper is divided basically into four sections, namely: (*i*) system considerations and optical arrangements, in which it is shown how variations in somewhat standard optical configurations can result in a high information rate modulation scheme; (*ii*) interfacing, which is mainly concerned with the producing of efficient interaction between the electronic circuits and the optical signals with traveling-wave type modulators; (*iii*) electronic circuits, which describes how the 1-Gb/s binary pulse code modulation (PCM) signals are generated by a gating and multiplexing process; and (*iv*) system performance, with general remarks on actual and extrapolated performance data.

II. SYSTEM CONSIDERATIONS

Various mechanisms exist for impressing information onto light beams including acousto-optic,[4] magneto-optic,[5] and electro-optic,[6] to mention but a few. Applied to laser beams, electro-optic interactions have proved most widespread, and in particular, the linear Pockels'[7] effect which is exhibited by many dielectric crystals. This effect relies on electric-field-induced perturbations of the natural birefringence of certain crystals including those in the point group 3m such as lithium niobate and lithium tantalate.[8] Lithium niobate's susceptibility to optically induced inhomogeneities to the natural birefringence[9] at relatively low optical beam powers makes its use as a modulator restricted and, in this application, lithium tantalate has been used. A figure of merit for $LiTaO_3$ when used as an electro-optic modulator is the half-wave voltage; this is the voltage required to induce a phase retardance of 180 degrees between components of the beam in the ordinary and extraordinary directions. Its value is derived from the elements r_{33} and r_{13} of the electro-optic tensor;[10] i.e.,

$$V_\pi = \frac{\lambda}{2(n_e^3 r_{33} - n_o^3 r_{13})} \cdot \left(\frac{c}{a}\right), \tag{1}$$

where n_o and n_e are the refractive indices experienced by ordinary and extraordinary components of the beam respectively. The factor c/a

is the inverse aspect ratio of the crystal, c and a being the width and length of the crystal respectively. For a value of wavelength equal to 496.6 nm, the half-wave voltage for $LiTaO_3$ is 2120 volts for the unity aspect ratio case. While large aspect ratio modulator rods can be prepared, so as to reduce the voltage required for a π-radian phase retardance, the value of voltage is still high enough to create severe performance demands on modulator driving circuitry in systems which rely on an "on-off" type modulator switch. Clearly, a partial modulation scheme will relax these standards but such a system is not easily implemented. It is pertinent at this point to consider some of the problems in realizing partial modulation schemes and at the same time discuss some of the advantages.

Consider the basic optical system shown in Fig. 1. It is easy to show that the modulated optical power, I_m, at the detector, is related to the applied modulator voltage V and the relative phase retardance ϕ of the beam components incident at the modulator by

$$I_m = \frac{I_t}{2}\left[1 - \cos\left(\phi + \frac{\pi}{2}\frac{V}{V_\pi} + \alpha\right)\right], \tag{2}$$

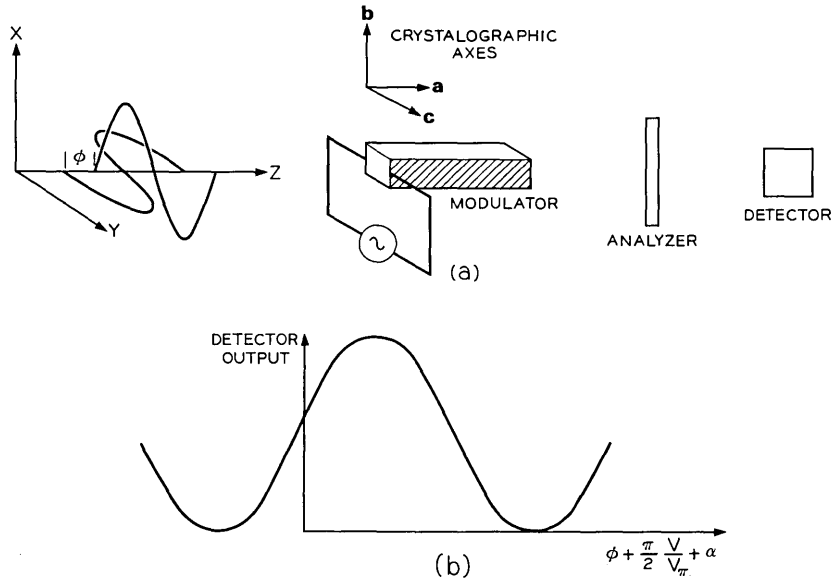


Fig. 1—(a) Basic optical modulation scheme. (b) System modulation characteristics.

where I_t is the total beam power at the modulator (assuming zero losses) and α is the retardation introduced by the natural birefringence of the crystal. Clearly, maximum beam modulation per unit modulation voltage is obtained when the x and the y components of the beam are equal in amplitude and adjusted for $\pi/2$ retardance at the analyzer. At this optimum bias and a crystal aspect ratio of 50 a small signal beam modulation depth of 8 percent per volt can be obtained. With such an arrangement satisfactory modulation depths of approximately 20 to 30 percent can be obtained with just a few volts of modulation potential.

In an actual system in which the modulated beam is detected via a photodetector which is coupled into a transmission structure of characteristic impedance Z_0, the system transfer function (defined as the differential ratio of modulation potential to potential appearing across the output transmission line) for a quasi small signal mode of operation is derived from equation (2) as

$$\frac{dV_{\text{out}}}{dV_{\text{in}}} = \frac{I_t}{V_\pi} \frac{\pi}{4} \frac{q\lambda}{hc} Z_0 \tag{3}$$

where q is the electronic charge, h is Planck's constant, and c is the velocity of light. The above relationship, based on the generation of one hole-electron pair per photon, assumes zero optical losses and a 100-percent quantum efficiency in the photodetector and works out to be approximately 0.5 per watt of beam power. Clearly, high beam powers will enhance the system transfer function coefficient, equation (3); however, a limit is imposed by optically induced inhomogeneities to the birefringence of the crystal. In an experimental system an upper limit of approximately 30 mW is realistic which gives a transfer function coefficient of 0.015 (not allowing for transmission losses) which for a modulation signal of several volts results in a satisfactory output amplitude. Satisfactory signal amplitude levels within the system, of course, depend on the system noise levels; in this respect, the presence of mode competition in a free-running laser system results in a very noisy channel. Two solutions to this problem are available: mode locking[11] and etalon tuning.[12] Mode locking effectively locks the longitudinal modes of the laser to produce a train of optical pulses at the frequency separation of the longitudinal modes, $c/2l$, where l is the length of the laser cavity. Optical communications systems incorporating mode-locked lasers offer some advantage and have been successfully demonstrated.[13,14] Mode-locked laser sys-

tems offer the availability of a stream of very narrow pulses of width inversely proportional to the oscillation bandwidth of the laser which can result in pulses of just a few picoseconds.[15]

A disadvantage of the mode-locked system is that optimum modulation depths and satisfactory base line definition are not compatible with low-voltage modulation since, in general, it is desirable in these systems to supply a modulation potential which is a substantial fraction of the half-wave retardation potential. One further disadvantage of mode-locked regime in gas lasers is the relatively low repetition rates of the optical pulses: This repetition rate is related to the frequency separation of the longitudinal modes, $c/2l$, and at short resonator spacings (required for high pulse repetition rates) the gain of gaseous lasers is insufficient to sustain oscillations. While techniques exist for effectively increasing this repetition rate by, for example, multiplexing[16] or harmonic pumping of the mode-locked frequency,[17] neither technique is a well established laboratory practice and the implementation of reasonably high rates (i.e., 1 GHz) is not, at present, conveniently realizable.

Relative noise immunity can be obtained with an etalon tuner by effectively selecting one pure optical frequency from the Gaussian line distribution. This is obtained by deforming the resonant poles of the laser so that only one conjugate pair lies within the laser gain curve.[18] The consequence of this type of operation is that mode competition is effectively eliminated and beam noise current is observed to be less than 5 percent of the total beam current. With such low noise powers, low modulation depths can be quite acceptable and a partial modulation scheme readily implemented.

There is no clear optimum in the choice of a particular optical frequency as a system carrier. While advantageous reductions in the modulator half-wave voltage can be obtained with shorter wavelengths [equation (1)], the quantum efficiency of the detector is in general reduced at these wavelengths (i.e., below 800 nm for Si). Focusing requirements are relaxed somewhat with shorter wavelengths since narrower diffraction-limited waists are obtained for a given far-field focusing angle. There is not, as yet, sufficient evidence for an optimum wavelength from the standpoint of optically induced changes to the birefringence. The absorption lines in $LiTaO_3$ are spaced such that, for gaseous laser output frequencies, the crystal is quite transparent and the choice of frequency is mainly determined by available powers from the various laser systems. In the system under discussion an ion

laser (argon) operating at 496 nm has been found convenient. This wavelength is somewhat below the characteristic peak in quantum efficiency for silicon photodetectors (800 nm), but nevertheless quantum efficiences of 0.3 A/W can still be obtained.

Coding for optical communications systems has so far received little attention. In the past the use of mode-locked lasers has given preference to a binary PCM coding technique. Recent studies[19] indicate that polarization multiplexed, phase shift keying, or pulse position modulation may be more appropriate to optical communications systems from a transmission standpoint. Problems still remain in implementing these modulation schemes. However, a frequency modulation scheme for space applications has recently been described[20] which shows promise for potentially wide bandwidths. Because of the advantages of regeneration, PCM systems appear attractive and in low-noise channels multilevel systems can be considered. The present system uses a binary type PCM coding scheme but is readily adaptable to multilevel with little increase in complexity.

III. OPTICAL ARRANGEMENTS

The optical arrangement used is of a fairly conventional nature employing a type of Senarmont compensator[21] as shown in Fig. 2. The argon laser provides a single frequency output of approximately 30 mW, plane polarized in the vertical direction. It is easy to show that a necessary condition for achieving optimum modulation is that the beam be circularly polarized at the analyzer plate. The consequence of a left- or right-hand polarization state is merely to induce a 180-degree phase inversion of the detected electrical signal. The circularly polarized state can be obtained by ensuring that the ordinary and extraordinary beam component amplitudes are equal with the necessary relative phase retardance. This is achieved by a 45-degree setting of the quarter-wave plate with an incident beam plane polarized at an angle $\theta/2$ which is adjusted by the half-wave retarda-

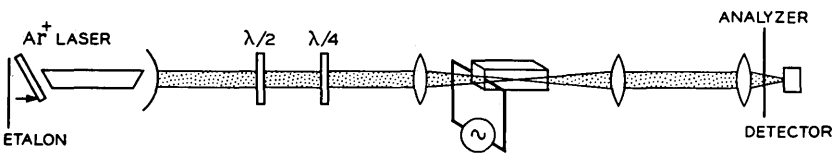


Fig. 2—Standard optical system with Senarmont type compensation.

tion plate. This particular arrangement can be easily checked by reference to the Poincaré-sphere.[22] Under these conditions, the x and y components of the beam are given by:

$$\left.\begin{aligned} E_x &= \sqrt{2}\, A\, \mathrm{Sin}\left(\frac{ct}{\lambda} - \frac{\theta}{2}\right) \\ E_y &= \sqrt{2}\, A\, \mathrm{Sin}\left(\frac{ct}{\lambda} + \frac{\theta}{2}\right) \end{aligned}\right\}. \tag{4}$$

The exact angle, $\theta/2$, is selected to compensate for the natural birefringence of the $LiTaO_3$ crystal. The electric-field-induced perturbations of the natural birefringence of the modulator rod translate to perturbations of the relative phase retardation of the x and y components of the exiting beam to produce a modulation of the beam ellipticity state. Collimation of the beam is provided to permit distant transmission. The analyzer may be placed before or after collimation depending on whether it is desired to transmit amplitude or polarization state modulation. At the receiver the beam modulation is converted to an electrical signal via the silicon avalanche photodetector. The avalanching mechanism provides a signal gain of only 2 or 3 due to the high optical power levels. In a detector operating at low levels, resulting from, for example, atmospheric attenuation of the beam, much higher gains should be possible. This is a desirable feature of the avalanche detector in that it provides a self-limiting or automatic gain control on the signal.

Efficient focusing of the beam for low-loss transmission through the modulator crystal is important. The problem can be dealt with in a number of ways and one method is given by Kogelnik and Li[23] in terms of the confocal parameters for optimum clearance. Another approach is to consider the percent transmission of power through the rod. Consider a rod of circular cross section, as shown in Fig. 3a, with an incident Gaussian beam with field distribution given by

$$E = E_0 e^{-[(x^2+y^2)/w^2]}, \tag{5}$$

where w is the width of the beam at the $1/e$ points. The intensity distribution, I, is therefore given by

$$I = I_0 e^{-2[(x^2+y^2)/w^2]}. \tag{6}$$

The total power in the beam, I_t, is computed from

$$I_t = I_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-2[(x^2+y^2)/w^2]}\, dx\, dy, \tag{7}$$

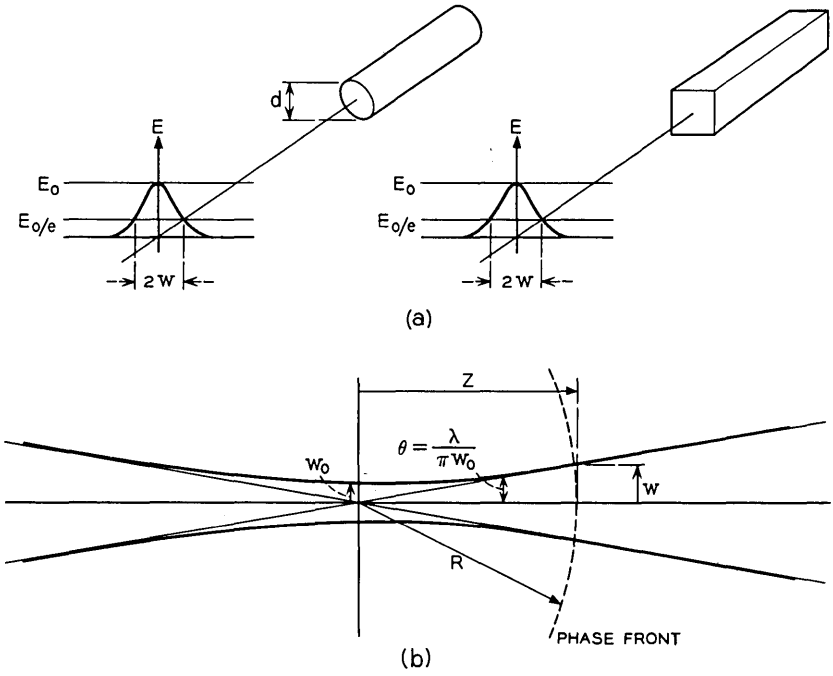


Fig. 3—(a) Modulator configurations. (b) Diffraction-limited beam waist.

which reduces to

$$I_t = \frac{I_0 \pi w^2}{2}. \tag{8}$$

Now, for the case of the circular cross section rod we can compute the rod dimension for 99-percent power transmission from

$$0.99\left(\frac{I_0 \pi w^2}{2}\right) = I_0 \int_0^{d/2} e^{-2r^2/w^2} 2\pi r \, dr, \tag{9}$$

which, after a parametric substitution, reduces to

$$\frac{d}{2} = 1.5w. \tag{10}$$

In a similar manner, the more practical case of a square cross section rod (Fig. 3a) can be dealt with, i.e.,

$$0.99\left(\frac{I_0\pi w^2}{2}\right) = I_0 \int_{-a}^{+a}\int_{-a}^{+a} e^{-2[(x^2+y^2)/w^2]}\,dx\,dy, \tag{11}$$

which, again after a parametric substitution, reduces to

$$a = 1.37w. \tag{12}$$

Clearly, the circular structure is more efficient from the standpoint of power transmission; however, the practical problem of applying uniform transverse electric field prohibits its use.

In matching the optical beam to the modulator rod, most power loss, due to dimensional limitations, will take place at the faces of the modulator due to beam expansion. Kogelnik has analyzed the expansion of focused Gaussian beams in some detail.[24] Figure 3b shows a typical beam waist at the focus point. The equations governing the beam parameters in this diffraction-limited case are:

$$w_0 = \frac{\lambda}{\pi\theta}, \tag{13}$$

where θ is the far-field diffraction angle;

$$w(z) = w_0\left[1 + \left(\frac{\lambda z}{\pi w_0^2}\right)^2\right]^{\frac{1}{2}}; \tag{14}$$

and

$$R(z) = z\left[1 + \left(\frac{\pi w_0^2}{\lambda z}\right)\right], \tag{15}$$

from which the confocal parameter, b, is obtained as the distance from the phase front radii minima as

$$b = \frac{2\lambda}{\pi\theta^2}. \tag{16}$$

The beam expansion can be completely specified, for a given wavelength, in terms of far-field diffraction angle, θ, which is, of course, the focusing cone half angle. For the case of a collimated beam† incident on the lens, θ is given by

$$\theta = \frac{w}{f} \text{ radians},$$

where w is the beam radius at the lens and f is the lens focal length.

† The description "collimated," within the context of Gaussian optics, here means a beam wavefront radius much greater than the focal length of the lens.

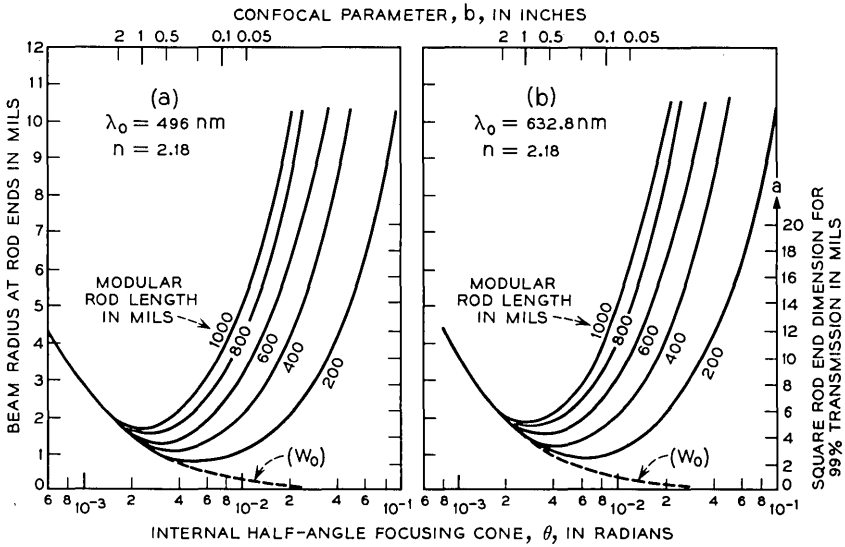


Fig. 4—Beam width of a modulator ends as a function of focusing cone angle: (a) $\lambda_0 = 496$ nm; (b) $\lambda_0 = 632.8$ nm.

Equation (14) may be used to calculate the beam radius at faces of a rod of length L for a given θ. This has been done for the case of $\lambda = 496$ nm and $\lambda = 632.8$ nm and the results are presented in Figs. 4a and b respectively. For each modulator rod length, a clear optimum cone angle is indicated. This optimum angle gives the same confocal parameter (on the upper abscissa) as Kogelnik's analysis. The ordinate on the right gives the necessary rod cross section for 99-percent optical power transmission in accordance with equation (12). Even for the case of the 1-inch rod with an aspect ratio of 100, a wide range of focusing cone angles are permitted for a high optical power transmission. In practical cases, criteria for high-efficiency transmission through the rod are easily met and other criteria assume importance. It is interesting to note that the curves for clearance in Figs. 4a and b coalesce for all rod lengths at a cone angle less than 1.5 milliradians. Thus, in this case, the transmission of optical power becomes relatively independent of the rod length.

One other criterion important in beam matching is that of nonuniqueness in the beam polarization state across the width of the beam. This arises as a result of the differing path lengths experienced by the various portions of the beam in the rod. Consider the wavefront

diagram shown in Fig. 5a. The optical path length of a particular portion of the beam in the modulator can be determined by a count of the wavefronts. Clearly, the path lengths are longer for the extremities of the beam than for the center. This is of no import in an isotropic media but in a crystal exhibiting birefringence, the result is a nonunique polarization state across the section of the beam. The problem of the propagation of Gaussian beams in anisotropic media has been analyzed in some detail[25] but it is sufficient here to adopt a somewhat simplistic argument to assess the quantitative effects of the nonuniqueness. The effect is important since it is at its maximum where the wavefront radius is at its minimum, i.e., when the length of the rod is equal to the confocal parameter, b. Consider the

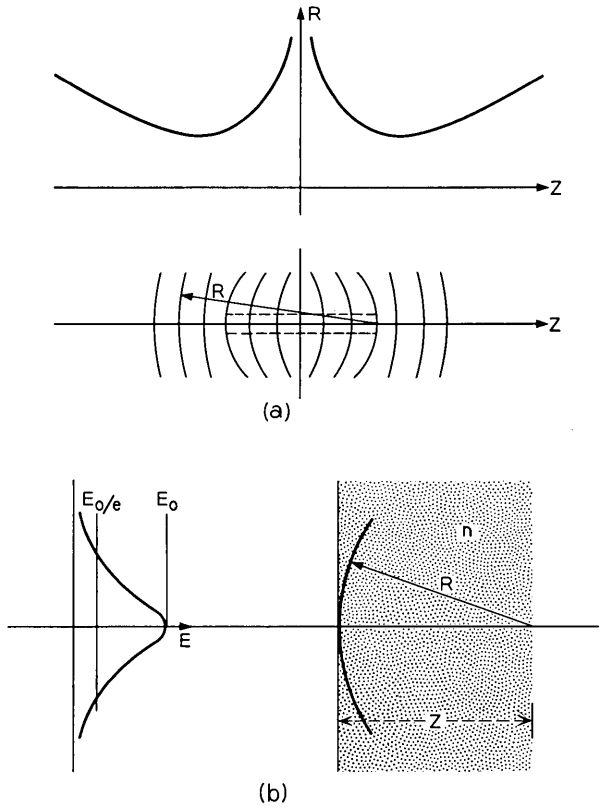


Fig. 5—(a) Wavefront radii at beam waist point. (b) Gaussian beam at surface of anisotropic media.

Gaussian beam incident at a flat face modulator crystal of length L as shown in Fig. 5b. The radius of curvature of the wavefront at the face of the crystal is given by

$$R(L/2) = \frac{L}{2}\left[1 + \left(\frac{\pi w_0^2}{\lambda L/2}\right)^2\right]. \tag{17}$$

For the important case of the rod length equal to the confocal parameter, for optimum clearance, equation (17) reduces to

$$R(L/2) = \frac{L}{2}\left[1 + \left(\frac{b}{L}\right)^2\right] = L. \tag{18}$$

A quantitative measurement of the nonunique polarization state of the beam can be obtained by determining the differences in relative phase retardation between the center and the E_0/e points of the beam. The radius of the beam at the face of the crystal is determined from

$$W\left(\frac{L}{2}\right) = W_0\left[1 + \left(\frac{\lambda L/2}{\pi w_0^2}\right)^2\right]^{\frac{1}{2}}, \tag{19}$$

which reduces to

$$W\left(\frac{L}{2}\right) = \sqrt{\frac{\lambda L}{\pi}} \tag{20}$$

for the confocal case.

The construction shown in Fig. 6a is useful in determining the nonuniqueness of polarization state for the above case. The situation is more complicated than that suggested by this simple construction in that the radius, R, is dependent on the beam polarization state as is

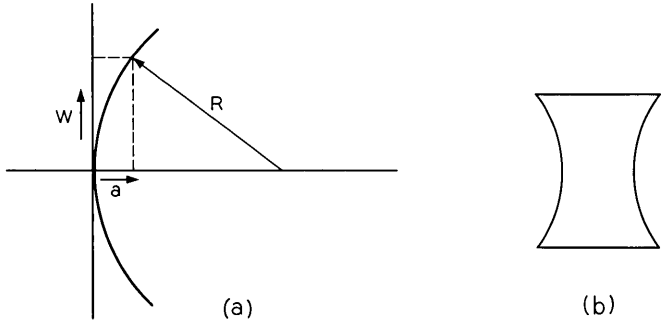


Fig. 6—(a) Evaluation of nonuniqueness in beam polarization state. (b) Modified modulator cross section to compensate for crystal birefringence.

also the beam angle in the crystal. Nevertheless, an approximation to the excess distance experienced by the E_0/e points of the beam is represented by the distance, a. The distance a is given by

$$a = R\left[1 - (\sqrt{1 - w^2/R^2})\right], \tag{21}$$

which, for $w/R \ll 1$ reduces to

$$a = \frac{w^2}{2R}. \tag{22}$$

Substituting equations (18) and (22) results in

$$a = \frac{\lambda}{2\pi} \tag{23}$$

and the relative differential path length, Δa, resulting from this excess distance is

$$\Delta a = \frac{\lambda}{\pi}(n_e - n_o) \tag{24}$$

(the effect is doubled due to the summation of the input and output excess distances). For $LiTaO_3$ this amounts to only 0.0004 λ, a negligible phase retardance. For other materials, such as $LiNbO_3$ or, more particularly, deuterated KDP, the effect will be much stronger due to the increase in the $(n_e - n_o)$ factor. The effect is analogous, in effect, to a built-in electric field, and where the polarization non-uniqueness is degrading, a modified cross section (Fig. 6b) can be adopted so as to produce a compensating nonlinear electric field across the modulator rod.

Of more practical significance in matching the beam to the modulator structure is the peak power intensity produced at the diffraction-limited beam waist which can give rise to optically induced inhomogeneities of the natural birefringence.[9] A field annealing process can reduce this susceptability[26] so that peak power intensities of 500 W/cm^2 are tolerable. The peak power intensity as a function of focusing cone angle is given in Fig. 7. This power intensity is obtained by combining equations (8) and (13) to give

$$I_0 = \frac{2n^2\pi\theta^2}{\lambda_0} \text{ W/cm}^2, \tag{25}$$

where θ is the internal focusing cone half angle. The strong dependence of power intensity on cone angle dictates a large confocal length, consistent with satisfactory clearance. A safety factor on clearance

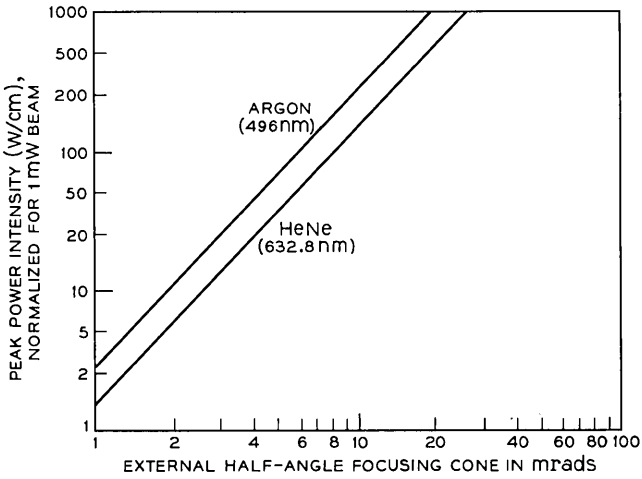


Fig. 7—Peak power intensity as a function of beam focusing angle.

has been proposed[6] which, in general, will restrict the minimum cone angle to some value less than the theoretical optimum indicated in Figs. 4a and b. For modulator rods of 10 × 10 mils cross section an internal half cone angle of 1.5 milliradians provides adequate clearance whilst maintaining satisfactory power intensities with beam powers of approximately 30 mW.

In summary, the matching of the beam to the modulator crystal is influenced by three criteria, viz.

(*i*) High-efficiency optical power transmission
(*ii*) Uniqueness of beam polarization state
(*iii*) Limitation of peak optical power intensity at the beam waist.

Separate optimization of items *i* and *ii* give rise to directly conflicting matching conditions. Fortunately, the degree of nonuniqueness of polarization state is not significant, for $LiTaO_3$, even for the worst case condition when the confocal parameter b equals the modulator rod length L. Further, the optimum exhibited for efficient optical power transmission is weak, and a wide range of focusing cone angles is allowable for very high efficiencies. Dominant in practical systems employing $LiTaO_3$ is the peak power intensity which will compromise the optimum condition for high power transmission. In cases when it is not possible to satisfy item *iii* and simultaneously maintain high-efficiency optical clearance, the crystal may be operated at an ele-

vated temperature (107°C) where the effect of optically induced birefringence changes is observed to relax at a rate faster than its creation.[9]

IV. INTERFACING—MODULATORS

Efficient interfacing of solid state circuitry with the electro-optic modulator is vital for high information capacity optical communications systems. In a practical system, the interfacing problem reduces to matching a transverse Pockels' effect crystal (lithium tantalate) to the impedance environment of the silicon driving circuits to achieve an efficient transfer of information onto the optical beam. A common approach in the past has been to regard the modulator as a lumped parameter structure which presents a reactive driving point impedance.[27,28] The major disadvantage of this approach is that not only must electrical charge be injected into the structure, but it must also be withdrawn by the driving circuits to restore the modulation field to its initial state. To accomplish this charging and discharging, at high data rates, impractically large charging and discharging currents are necessary. Further, the capacitive load will create a mismatch in the driver-modulator transmission path resulting in standing waves with consequent modification of the impressed waveform.

A different approach is to make use of the intrinsic characteristic impedance of approximately 55 ohms which is exhibited by $LiTaO_3$. This impedance can be realized by rearranging the modulator structure to make use of the inherent distributed parameter structure of the modulator crystal. The resulting modulator will be of the traveling-wave type,[29,30] which presents a real driving point impedance so that an efficient interface is obtained with the impedance levels of high-speed silicon transistor circuits (50 ohms). Previously traveling-wave type interactions have been adopted to increase the modulator bandwidth;[31] however, in this application their primary use is to provide improved impedance matching. The improved bandwidth appears as a secondary effect as will be discussed later.

The exact characteristic impedance of the modulator structure can be determined from the differential capacitance and inductance along the length of the crystal. The capacitance per unit length is simply

$$C = \epsilon_0 \epsilon_r \frac{b}{c} \text{ F/m}, \tag{26}$$

where b and c are the crystal dimensions along the **b** and **c** optical

axes respectively. This, of course, assumes zero fringing of the electric field outside the crystal which is justifiable since $\epsilon_r \gg 1$. The inductance L due to the two parallel crystal electrodes of length ℓ is[32]

$$L = 4\ell\left[\log_e \frac{2\ell}{b} - \log_e \frac{2\ell}{c} + 1.5 - \log_e e + \log_e k - \frac{c}{\ell} - \frac{1}{4}\frac{c^2}{\ell}\right], \quad (27)$$

where $\log_e e$ and $\log_e k$ are functions of b and c given in Ref. 32. For a modulator crystal of 10 mils square cross section, the inductance per unit length reduces to

$$\frac{dL}{d\ell} = \left(6.3 - \frac{c^2}{\ell^2}\right) \text{nH/cm}. \quad (28)$$

So that, combining (26) and (28), the characteristic impedance, as a function of distance, is

$$Z_0(\ell) = \sqrt{\frac{\left(6.3 - \frac{c^2}{\ell^2}\right)c}{\epsilon_0\epsilon_r b \times 10^2}} = 41 \text{ ohms} \quad (29)$$

for $c^2 \ll \ell^2$, i.e., at some distance from the ends of the modulator. The exact variation of characteristic impedance along the length of the rod is given in Fig. 8. The deviations of Z_0 at the rod ends are in fact abstractions, which will not be exhibited by a practical structure due to the presence of connections to the electrode end points which results in a more complex situation. It is assumed in this calculation that there is no dispersion of the dielectric permittivity. Experimental high-frequency measurements have indicated that this is the case[33] and further experiments have extended this result to 6 GHz.[34]

The group velocity differences of the electrical and optical signal velocities in the modulator will give rise to a bandwidth limitation. This limitation is readily analyzed by consideration of the situation depicted in Fig. 9. The voltage experienced by an optical wave at a point z in the modulator rod is

$$V(z) = V \text{ Sin}\left[\omega\left(t + \frac{z}{v_t} - \frac{z}{v_e}\right)\right], \quad (30)$$

where v_t and v_e are the velocities of the light and electrical signals respectively in the modulator rod. For a rod length, L, the total induced phase retardation of the optical beam is, therefore,

$$\phi_{\text{Tot}} = \int_{-L/2}^{+L/2} \frac{\pi V}{V_\pi L} \text{Sin}\,\omega\left(t + \frac{z}{v}\right) dz, \quad (31)$$

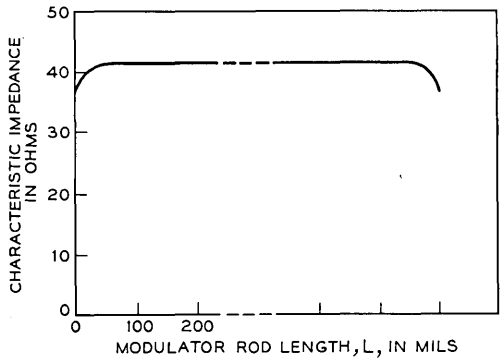


Fig. 8—Modulator characteristic impedance as a function of distance along crystal.

where

$$\frac{1}{\nu'} = \frac{1}{\nu_t} - \frac{1}{\nu_e}\,; \tag{32}$$

hence,

$$\phi_{\text{Tot}} = \frac{2\pi V \nu'}{V_\pi L} \operatorname{Sin} \frac{\omega L}{2\nu'} \operatorname{Sin} \omega t \tag{33}$$

which is of the familiar form shown in Fig. 10a. This amplitude spectrum results in the Dirac response of Fig. 10b and the electro-optic transfer function step function response time is, therefore,

$$t_r = L\left[\frac{1}{\nu_e} - \frac{1}{\nu_t}\right]. \tag{34}$$

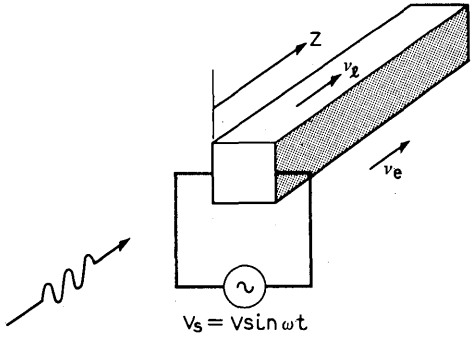


Fig. 9—Traveling-wave type modulator structure.

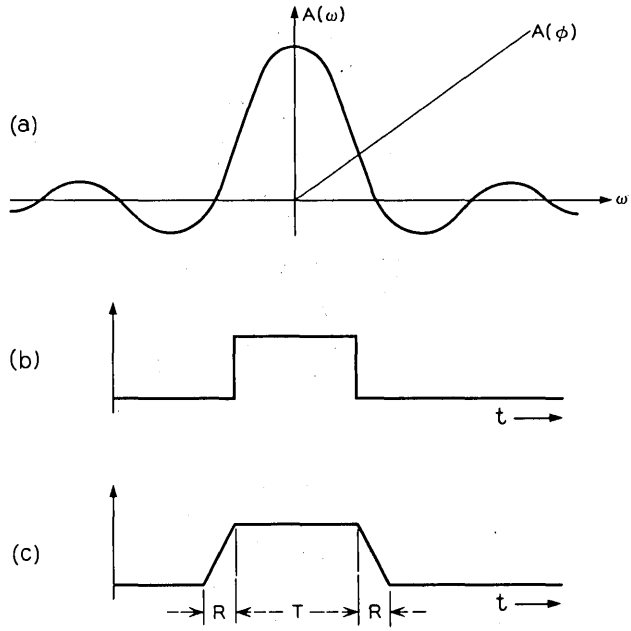


Fig. 10—(a) Amplitude spectrum of traveling-wave modulator. (b) Dirac response. (c)System response to pulse width T.

Typically, for a 1-cm-long modulator crystal a rise time of 150 ps is indicated.

The modulator bandwidth can be increased with velocity compensation of the light and electrical signals. Kaminow, et al.,[35,36] and Peters[37] have described some techniques for accomplishing this type of compensation. The proposals suffer from some disadvantages in practical realizations, including a loss of modulation power, fabrication problems in achieving high-efficiency power transmission, and, in particular, the latter proposal[37] could result in multimode propagation of the electrical modulation signal in the modulator structure. In practice, the modulator bandwidth limitation is not significant, with respect to system bandwidth, even in high information rate systems (1 Gb/s) and it is not necessary to resort to velocity compensation techniques. This is easily shown by considering an input pulse, $f_1(t)$, of width T where $2T$ is the system information rate interval, i.e.,

$$f_1(t) = A[H(t) - H(t - T)], \tag{35}$$

denoting the unit step function as $H(t)$. The system Dirac response, $f_2(t)$, has been shown to be

$$f_2(t) = B[H(t) - H(t - R)], \tag{36}$$

where

$$R = \frac{L}{C}(\epsilon_r^{\frac{1}{2}} - n). \tag{37}$$

The system response, $f_0(t)$, to $f_1(t)$ is given by the convolution

$$f_0(t) = f_1(t) * f_2(t)$$
$$= \int_0^t A[H(\tau) - H(\tau - T)]B[H(t - \tau) - H(t - R - \tau)]\, d\tau \tag{38}$$

from which it is obvious, from the graphical form of the convolution integral, that $f_0(t)$ is as shown in Fig. 10c. So that, providing $R \leqq T/2$ the modulator bandwidth is not limiting; this is clearly so at information rates of 1 Gb/s.

It is important to observe from (33) and (34) that a trade off exists between modulator bandwidth and half-wave voltages so that it is possible to employ velocity compensation to permit longer crystals (while maintaining the same bandwidth as the uncompensated structure) to achieve advantageous reduction in the half-wave voltage. This trade off is also possible with different modulator materials. In fact, a gain-bandwidth product can be assigned to a particular modulator material. With this figure of merit, the gain is related to the inverse half-wave retardation voltage and the bandwidth is related to the optical refractive index and the microwave dielectric permittivity, i.e.,

$$GB = \frac{1}{(n - \epsilon_r^{\frac{1}{2}})V_\pi}. \tag{39}$$

This figure of merit is given for a number of modulator materials in Table I. The situation is more complex, however, than is suggested by a numeric value of the gain-bandwidth product: GaAs, for example, only exhibits below bandgap transmission for $\lambda > 900$ nm and GaP incorporates a dielectric relaxation time which can be as large as 10 μs.[38] Further, the mechanical properties of the materials dictate the feasibility of fabrication of large aspect ratio crystals to effect a gain-bandwidth trade off. Other factors, including free carrier optical absorption and optically induced crystal damage, further complicate the choice of material.

TABLE I — FIGURE OF MERIT FOR DIFFERENT MODULATOR MATERIALS

Material	Type	n(av)	ϵ_r	V_π†	$GB(V^{-1} \times 10^{-5})$
$LiTaO_3$	Ferroelectric Perovskite	2.18	43	2130	10.7
$LiNbO_3$	Ferroelectric Perovskite	2.25	30	2340	13.3
KH_2PO_4	KDP[40]	1.5	21	4350	7.5
GaAs	AB-Type Semiconductor	3.5	12.5	3870	$\rightarrow \infty$
GaP	AB-Type Semiconductor	3.45[39]	11.1[38]	7830	$\rightarrow \infty$

† Extrapolated to $\lambda = 496$ nm.

$LiTaO_3$ represents a compromise of the gain-bandwidth product and the other factors while still providing significant advantages in response time when incorporated in a traveling-wave structure compared to the lumped parameter configuration. Equation (34) can be used to compare the performance of the traveling-wave $LiTaO_3$ modulator with the electro-optic transfer function rise times obtainable with lumped parameter configurations operating in the normal 50-ohm environment. The rise times are, of course, a function of modulator rod length as is seen in Fig. 11. Even for the double-pass-type lumped parameter structure,[13] the response times of the uncompensated traveling-wave structure are significantly superior. Typically, for a modulator rod length of 500 mils, the rise times are 150 ps, 250 ps, and 500 ps for the traveling-wave, double-pass lumped parameter, and single-pass lumped parameter structures respectively.

The traveling-wave modulator structure has been realized in microstrip thin film form on an alumina (Al_2O_3) substrate as shown in Fig. 12. The alumina substrate has microstrip transmission lines of 50 ohms characteristic impedance etched in the evaporated chrome-gold layers with the reverse side metalization removed below the modulator rod. Connections are brought through the substrate from the reverse side metalization to make contact with the upper contact lands. The modulator crystal has evaporated electrodes, typically chrome and palladium, on opposite faces along the longitudinal axis. Placement of the modulator crystal, with the b optical axis perpendicular to the surface of the substrate, is facilitated with low-temperature solder which also provides acoustic damping.[27] The electrical modulation signal is applied via the coaxial-to-strip-line transition, the strip line, and the rod, and exits from the substrate to be absorbed in a matching termination. The possibility exists, however, for reapplying the elec-

trical modulation signal to further cascaded modulator structures without the need for any additional electrical power. In this manner, not only can increased modulation depths be obtained but some degree of compensation can be obtained for the electrical and optical velocity differences by adjusting the relative electrical and optical delays between the two modulators as shown in Fig. 13.

A photograph of an actual modulator structure with a 500-mil-length rod is shown in Fig. 14. The time domain reflectometry (TDR) trace of the characteristic impedance throughout the modulator is presented in Fig. 15a. Small perturbations in characteristic impedance due to the coaxial-to-strip-line transitions are observed with a much larger deviation from the 50-ohm impedance level along the length of the crystal. The crystal impedance level ($\sim$ 30 ohms) is less than indicated by equation (29) and is partially attributable to the electric field deformation due to the close proximity of the high permittivity ($\epsilon_r = 9.7$) alumina substrate. The time resolution of the TDR system is 30 ps and the relative slowness of the drop to the 30-ohm level indicates the possibility of a dielectric relaxation mechanism in the crystal. The impedance mismatch results in a reflection coefficient, ρ, of 0.25 or a VSWR of only 1.65, and no significant deformations of

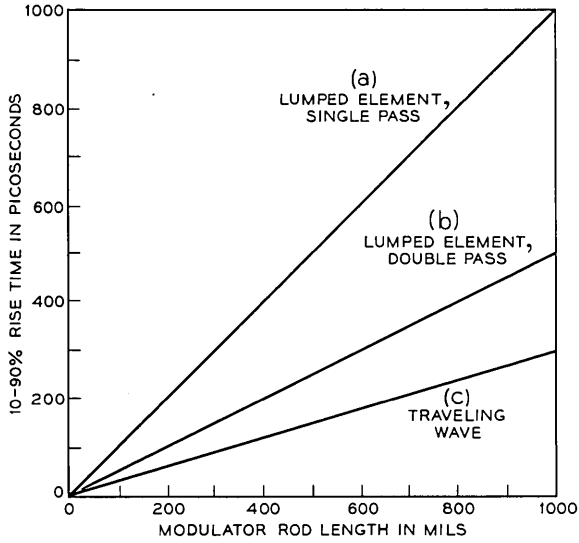


Fig. 11—Rise time as a function of modulator rod length for traveling-wave and lumped parameter single- and double-pass systems.

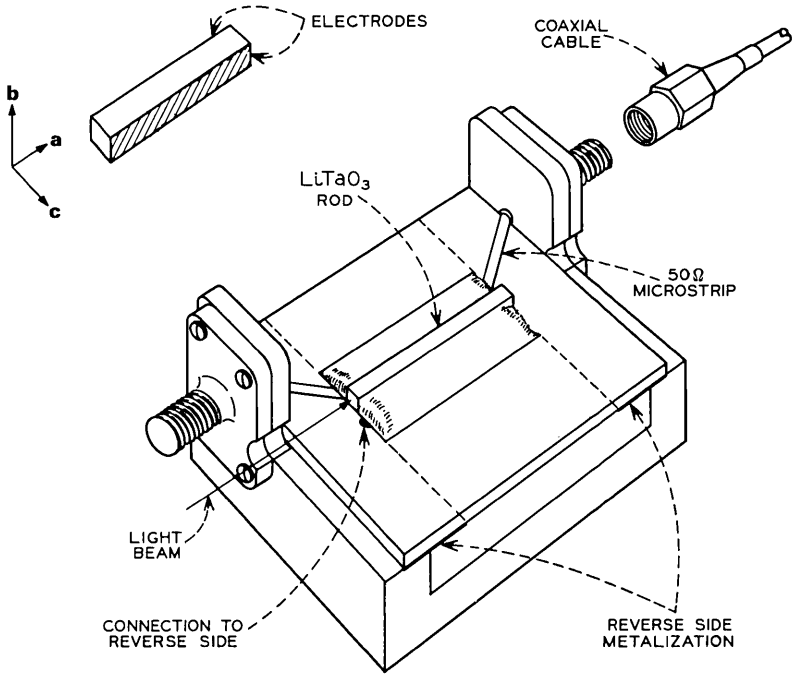


Fig. 12—Traveling-wave modulator embedded in microstrip transmission lines.

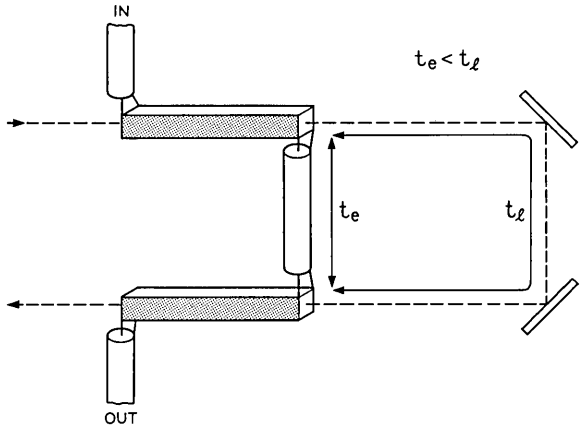


Fig. 13—Cascaded traveling-wave modulator structure.

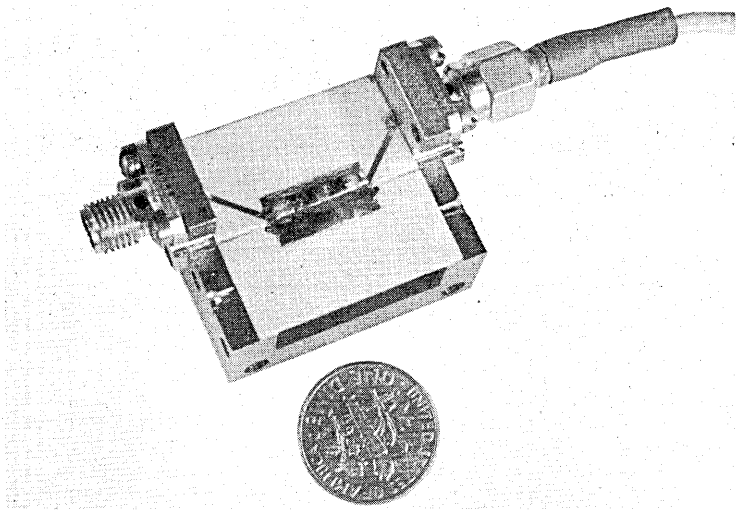


Fig. 14—Photograph of traveling-wave $LiTaO_3$ modulator.

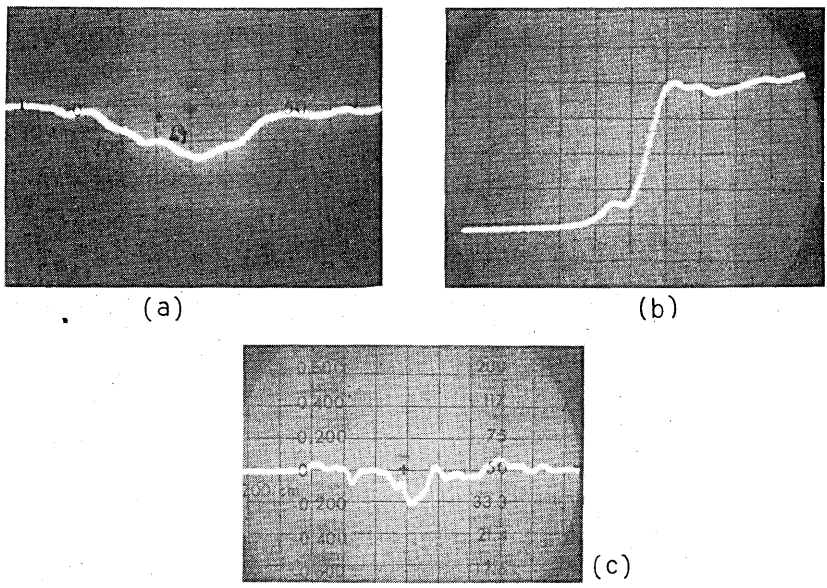


Fig. 15—(a) TDR characteristics for $LiTaO_3$ modulator (100 ps/cm). (b) Electrical pulse transmission through modulator (100 ps/cm). (c) TDR characteristics for $LiNbO_3$ modulator (200 ps/cm).

the modulation waveforms are observed in the actual system as a result of the nonunity VSWR. The high-frequency attenuation of the electrical signal after traversing the modulator structure is illustrated in the time domain response of Fig. 15b. The rise time is significantly less than 100 ps indicating an electrical band pass in excess of 4 GHz.

Lithium niobate, with a lower permittivity, produces a characteristic more nearly approximating the desired 50-ohm impedance level as indicated in the TDR trace of Fig. 15c. This is for a 360-mil-length rod with a 10 × 10-mil cross section mounted in a similar traveling-wave structure.

Modulator rods of 60:1 aspect ratios (10 × 10 × 600 mils) have been fabricated without any noticeable evidence of stress induced inhomogeneities to the birefringence and the actual system performance to be described employs this length of modulator crystal.

No attempt has been made to compensate for the strong temperature dependence of the electro-optic effect in $LiTaO_3$. Small signal modulation reduces this problem since RF heating is negligible. Optical power absorption and ambient temperature changes can be significant and in a practical system the cascaded, temperature compensated structure of Peters[41,42] can be employed. This scheme has recently been realized in a 10 × 10-mil cross section crystal in a traveling-wave type structure.[43]

V. HIGH-SPEED ELECTRONIC DRIVING CIRCUITS

At present, a major factor in restricting the information capacity of optical communications systems is the speed of operation of the electronic driving circuits. It is in this area that effort will result in information rate increases. The present system uses an electronic multiplexing and gating process to achieve the 1-Gb/s PCM word patterns necessary to produce the overall system information capacity. Optical multiplexing[16] has been proposed as a means of increasing system information capacity and it is possible that this multiplexing technology could complement the electronic multiplexing described here. Indeed, such a combination may be preferable to a wholly optical multiplexing scheme since the complexity of the optical arrangements would be reduced.

The general schema of the electronic PCM system is shown in Fig. 16. Here basic system timing is provided by a master 250-MHz sine-wave oscillator. After division by a quarter-wave transformer to

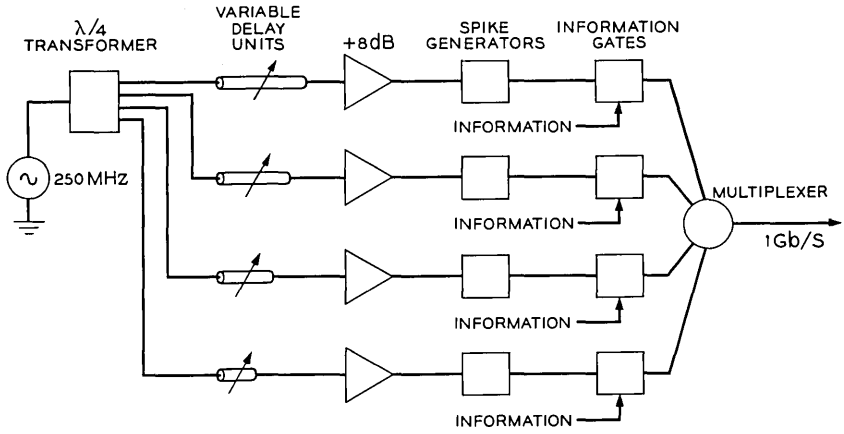


Fig. 16—Schema of electronic PCM system.

achieve impedance matching, the waveform provides system synchronization and drive to four Boff (snap-off) diode subnanosecond spike generators. Amplification and variable delay units are provided for adjustment of the phase and amplitude of the four channels. The amplifiers are standard class C tuned stages providing approximately 8 db of gain and the variable delay units are standard variable-length coaxial lines. The timing of the four waveforms is arranged so that the spike pulses fall in sequence into each of four time slots or windows, 1 ns wide. In this manner the pulses can be multiplexed in an OR gate multiplexer to produce a bit rate of 1 Gb/s. Information is impressed onto the pulse trains in the individual channels in "information AND gates." With this technique the information signals can be of a relatively slow nature (i.e., rise and fall times of 2 ns) at a bit rate of 250 Mb/s. Logic processing prior to impressing onto the spike trains is already available in standard emitter-coupled integrated circuit form.

The realization of some of the constituent parts of this system represents some difficult design and fabrication problems and it is pertinent to discuss in some detail these subsystem elements.

5.1 *Boff Diode Spike Generators*

The high-speed subnanosecond spikes are generated in a tuned circuit excited by the abrupt cessation of carrier storage in the Boff diode. The generator is of a conventional shunt-mode type which is well

documented.[44,45] The actual circuit form is shown in Fig. 17a. The width of the pulse is determined by the half period of the resonant circuit formed by L_t and C_t. L_m and C_m provide matching to the sine-wave source whilst R_b sets the dc bias.

Amplitudes of approximately 10 volts are obtained with pulse base widths of 700 ps. Significantly narrower pulse widths are easily obtainable with the promise of higher orders of multiplexing and concomitantly higher bit rates. The prime limiting factor to this type of extrapolation, however, is the transmission bandwidth of the information gate. A typical spike generator output waveform is shown in Fig. 17b. Figure 17c shows the output waveform of the four channels with the pulses positioned in their relative 1-ns time slots. The circuit is realized in conventional strip line construction with lumped parameter reactances.

5.2 *Information Gates*

The function of the information gate is to transmit or inhibit individual spike waveforms at a bit rate of 250 Mb/s in response to a "1" or a "0" on the information signal. High-speed considerations would normally suggest a passive gate employing Schottky barrier diodes; however, for the case of equal input and output impedance it is not possible to obtain a passive gate gain in excess of 0.5. In

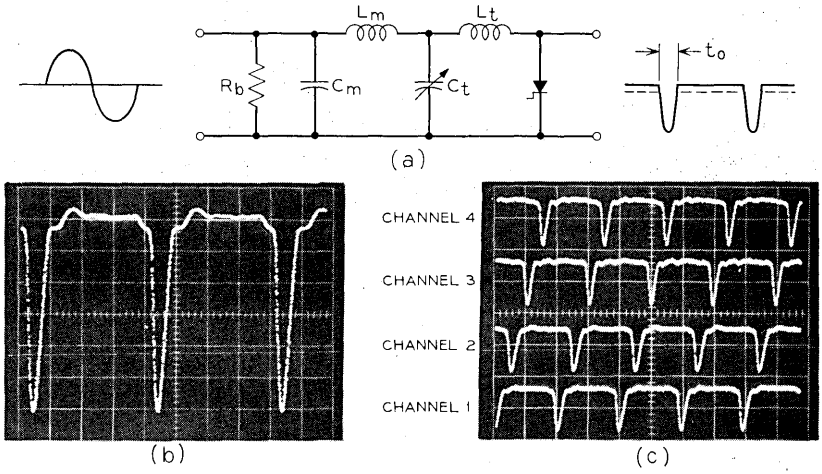


Fig. 17—(a) Boff diode spike generator. (b) Single-channel output waveform (1 ns/cm, 2 V/cm). (c) Four-channel output in relative time slots (2 ns/cm, 5 V/cm).

contrast, active gates can give gains in excess of unity although a more limiting bandwidth is imposed. The bandwidth limitation can be minimized to acceptable levels with the current routing pair configuration which has been used in this system. The simple circuit form is shown in Fig. 18a. The actual circuit incorporates a number of dc restoration networks so that little demand is made on the entropy of the information signal.

The high-speed spikes with frequency components in excess of 1 GHz are applied to the common emitter node via a terminating resistance resulting in current pulses with amplitudes of the order 100–200 mA. Current routing between transistors Q_1 and Q_2 is controlled by the base potential of Q_1 which is determined by the amplitude of the information signal. With this configuration Q_2 operates in essentially the common-base mode so that the full common-base current gain cutoff frequency of the device can be realized. In contrast, the switching action in response to the information pulses takes place

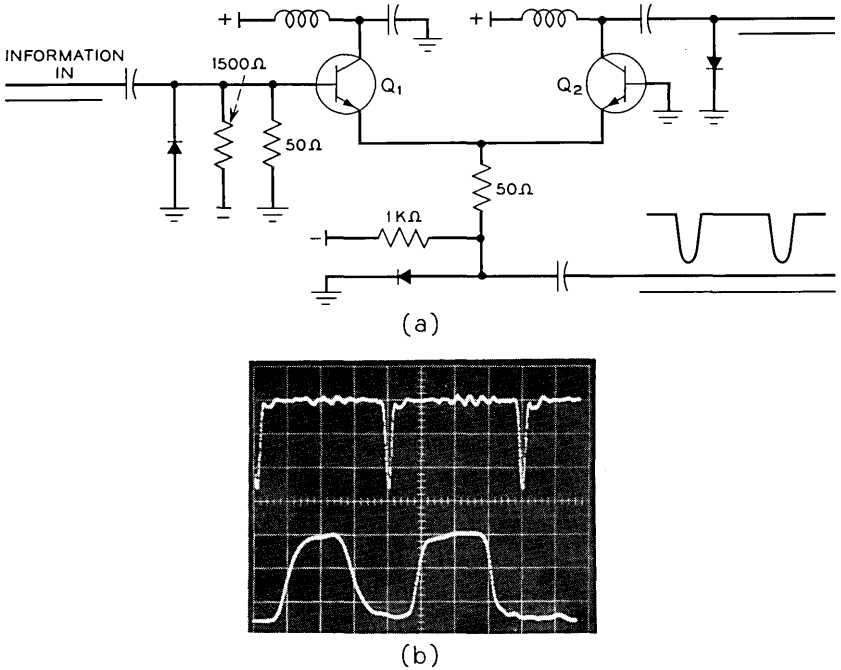


Fig. 18—(a) Information gate. (b) Information gate switching: top trace, output; bottom trace, information input (2 ns/cm, 2 V/cm).

at the base where switching speed is commensurate with the rise times of the information signal.

As a result of the large-amplitude currents being switched in the devices, high-level injection effects occur. The problem then is that phenomena, such as emitter crowding, transverse voltage drops[46] in the base, and the Kirk effect,[47] produce device parameter degradation. These high-level injection effects are still significant in the nine-stripe geometry devices used. Device parameter degradation has two effects, first, a bandwidth limitation which in turn produces a reduction in output pulse amplitude, and second, emitter-to-base voltages which become quite large so that complete current switching from one device to the other is not possible. A consequence of the latter effect is that it is not feasible to switch off Q_1 at any time (due to reverse emitter-base junction breakdown of Q_2), and Q_2 is only switched off in response to large (5 V) pulses at the base of Q_1. This degenerate switching mode of operation is satisfactory however, since only the output of Q_2 is utilized. The overall capabilities of the information gate are a switching time of less than 2 ns with transmission bandwidth in excess of 1 GHz for output pulse amplitudes of 7 volts. Typical switching action is indicated in Fig. 18b. Here, information impression on the spike pulse train is shown (upper trace) for a 10101 word pattern, in response to the relatively slow information signal (lower trace). It is to be noted that the base width of the exiting word pattern pulses is still significantly less than 1 ns so four-channel multiplexing is still efficient. The circuit is realized in thin film microstrip with thin film chrome resistors and beam leaded transistors. The actual structure is shown in Fig. 19. The devices used had a cutoff frequency of approximately 4 GHz and a collector-base capacitance of approximately 0.3 pF.

5.3 *1-Gb/s Multiplexing Gate*

Multiplexing of the four information bearing pulse trains can be accomplished with a four-input "OR" gate. "OR" gates, unlike "AND" gates, can have unity gain with passive components. The use of passive components relaxes somewhat the high-speed circuit requirements for the multiplexer. In particular, Schottky barrier diodes can be used to advantage for the multiplexer. The low carrier storage time of these devices permits the high-speed circuit performance required. The actual circuit used is a simple four-diode structure in microstrip as shown in Figs. 20a and b. This particular hybrid construction used leadless inverted device (LID) packaging of the devices.

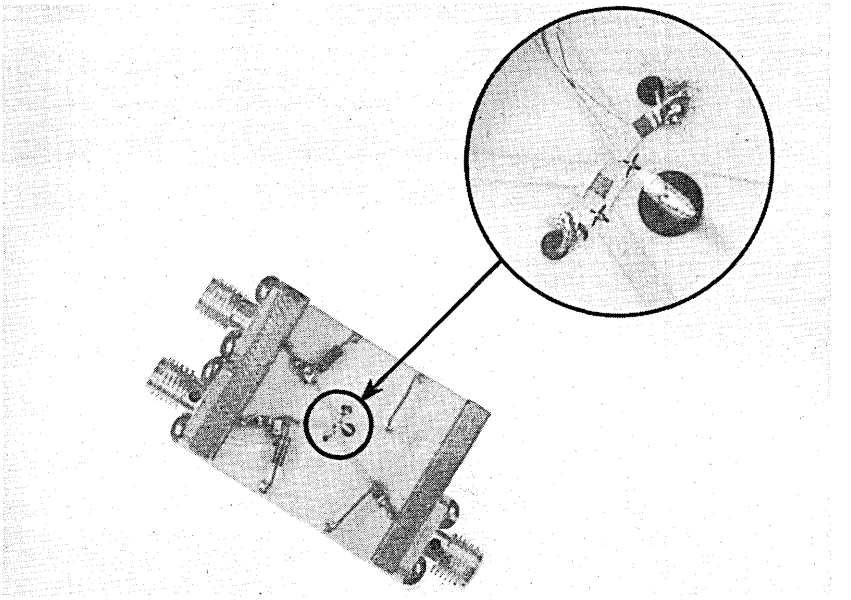


Fig. 19—Information gate realization in thin film microstrip with beam lead devices.

A typical set of input waveforms to the multiplexing gate is shown in Fig. 21a. Here, the four channels are positioned in their relative 1-ns-wide time slots with a 10101 word pattern on channel 1. The multiplexed, 1-Gb/s output is shown in Fig. 21b. Little degradation in wave form speed is apparent. The speed limiting factors are Schottky barrier diode storage time and junction capacitance. The former can be significantly less than 100 ps[48] and the latter can be a fraction of a picofarad.[49] In cases where a large number of channels are multiplexed, the junction capacitance can become the dominant limiting factor since it is a cumulative effect. Diode barrier potential and differential "on" resistance result in some amplitude attenuation which attenuates output pulses to 5 or 6 volts.

The output from the multiplexing gate is applied, via a coaxial line, to the $LiTaO_3$ traveling-wave modulator.

VI. OTHER REALIZATIONS

Other electronic realizations of the circuit functions described are possible which should provide improved performance at higher data rates with higher modulation depths. An example of an alternative

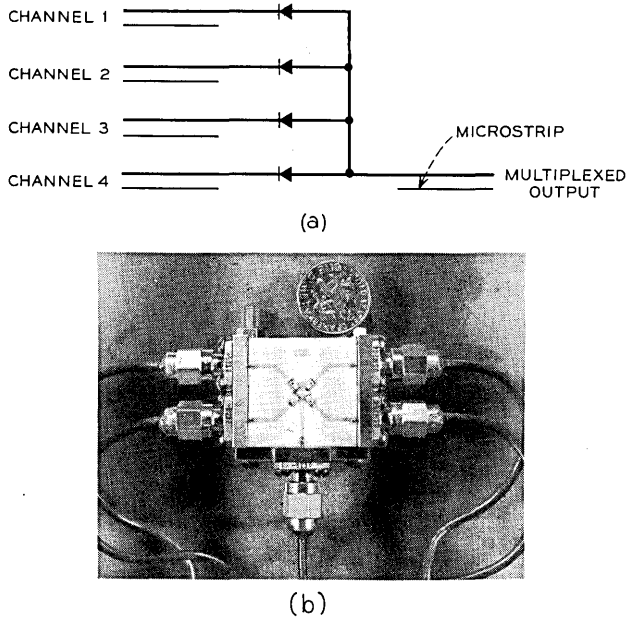


Fig. 20—(a) Multiplexing gate "OR" circuit. (b) Multiplexing gate realization in thin film microstrip.

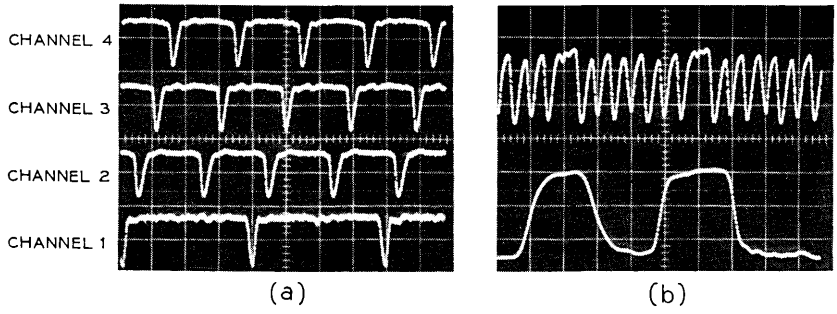


Fig. 21—(a) Input waveforms to multiplexing gate (2 ns/cm, 5 V/cm). (b) Upper trace: 1-Gb/s multiplexed output; lower trace: channel 1 input word pattern (2 ns/cm, 2 V/cm).

realization is shown in Fig. 22. In this realization the functions of the information gates and the multiplexer are combined by taking advantage of inherent multiplexing capabilities of the collector output circuits. In this manner, attenuation resulting from the multiplexer diode barrier potential and differential "on" resistance is eliminated. High-level injection effects in the information gate devices can be reduced by cascaded structures so that significantly higher amplitude modulation waveforms can be obtained. Such realizations could lead to "on-off" systems with very high data rates (in excess of 1 Gb/s). The limiting factor in these realizations is the cumulative addition of

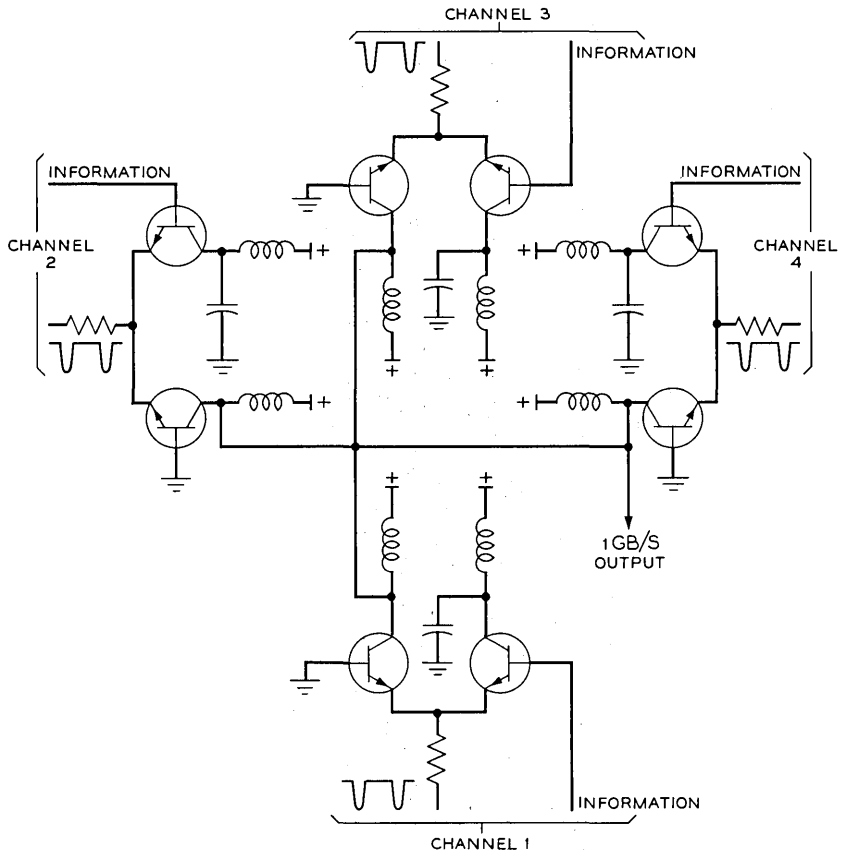


Fig. 22—Alternative realization of information gates and multiplexer functions.

collector output capacitance but present-day devices are available with values of capacitance significantly below a picofarad and time constants of less than 100 ps should easily be obtained.

With the previously described circuit elements, multilevel coding is readily implemented. This form of coding is achieved merely by supplying spike waveforms with the necessary quantization in amplitude levels. The system depicted in Fig. 23 is such a realization for a four-level coding scheme. Not shown is the conversion logic which is required at the inputs to the information gates but since this is at a relatively slow data rate (250 Mb/s) it is easily accomplished with standard integrated circuits. Multilevel PCM coding schemes are indeed attractive in view of the low system noise as will be discussed later.

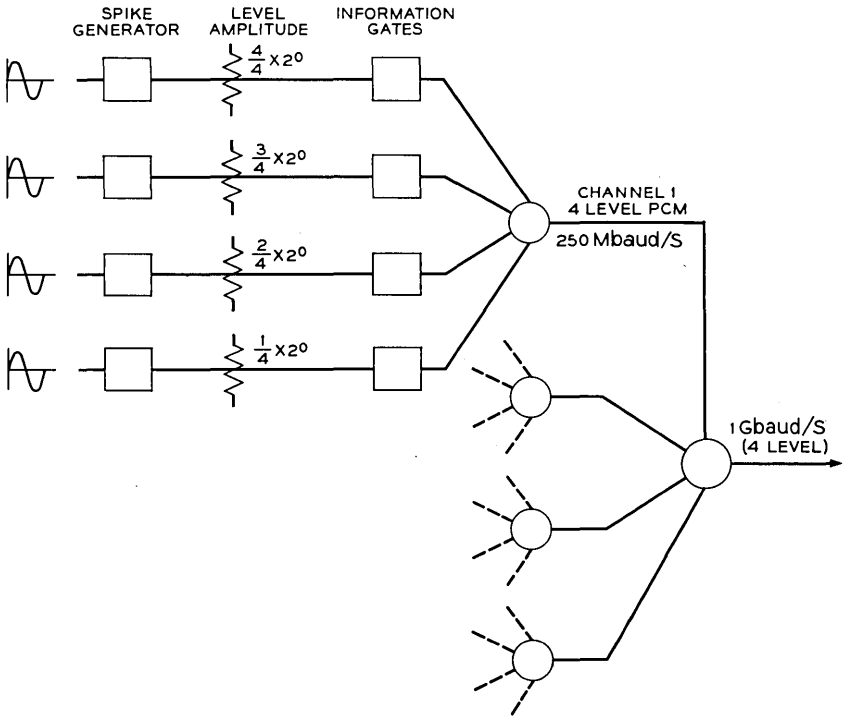


Fig. 23—Multilevel PCM coding schema.

VII. SYSTEM PERFORMANCE

The complete system was established in accordance with the basic schemas of Figs. 2 and 16. Optical loss and beam parameters are important factors in such systems where efficient use of beam power and optical component matching are required. Losses through passive optical components are relatively unimportant (typically 1 percent or less per surface). However, modulator losses are significant because of optical absorption and beam blockage due to large aspect ratio modulators. To achieve efficient optical matching, exact knowledge of the beam parameters is required; this was obtained by two experimental techniques, viz., slit measurements and edge measurements similar to those of Arnaud, et al.[50] Both techniques made possible reliable estimates of the beam intensity profile. Typical of these measurements is the experimentally obtained diffraction-limited waist pattern shown in Fig. 24 which compares well with the calculated waist pattern [from equation (14)] also presented. Such exact knowledge of beam parameters enabled modulator losses as low as 0.5 dB to be obtained. In practice, however, low-loss transmission does not coincide with high-efficiency modulation depths but, nevertheless, losses as low as a few decibels are obtainable.

Practically realizable system transfer function coefficients of approximately 0.17/W were obtained as opposed to 0.5/W predicted by equation (3). This lower value is accounted for primarily by the nonlinear modulation mode. Other factors, including reflection from the exiting face of the modulator crystal and analyzer losses, also reduce the coefficient. Nevertheless, output voltages of 20 mV result from applied modulation potentials of 4 V.

The total system performance at a bit rate of 1 Gb/s is illustrated in Fig. 25a. This output from the photodetector (lower trace) results from the impression of a 101101 word pattern on channel 1, channels 2 through 4 having all 1's impressed. This particular result is presented as typical of the many different word patterns that have been successfully transmitted over the laser beam. The detected electrical output from the photodetector after transmission via the beam is to be compared with the electrical modulation waveform exiting from the modulator structure (upper trace). Two points are to be noted, one the fidelity, and two, the low noise. The first point is indicative of

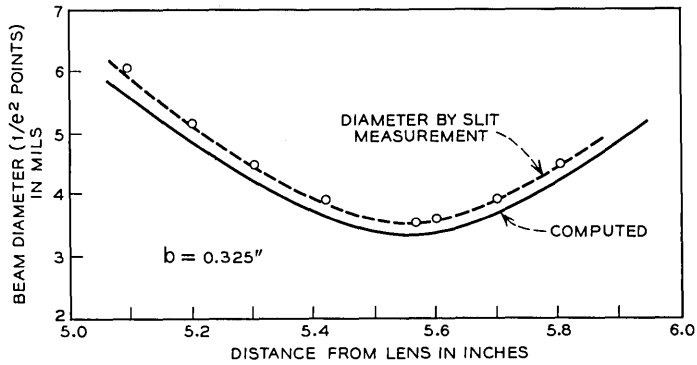


Fig. 24—Experimental and computed beam profiles at diffraction-limited waist.

the exceptionally wide bandwidth of the system while the latter augurs well for a multilevel type of coding.

The bandwidth of the system is mainly determined by the length of the modulator crystal and a length of 0.6 inch results in a frequency axis crossing point of approximately 5 GHz. This consequentially produces, in the time domain, a step response rise time of approximately 200 ps. Figure 25b is an experimentally observed response for a 250-ps rise time pulse excitation (upper trace); the detected pulse (lower trace) exhibits a rise time of approximately 400 ps which is commensurate with a system response time of 200 ps. The system bandwidth, of course, can be made arbitrarily large, up

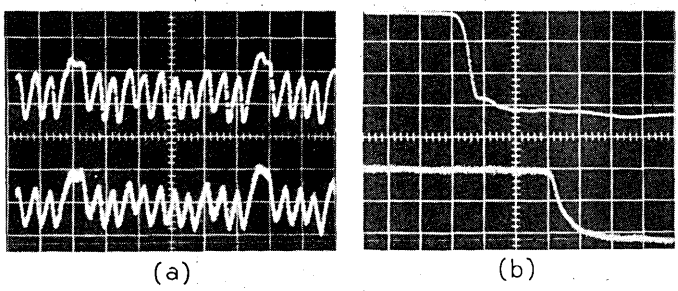


Fig. 25—(a) Upper trace: electrical modulation signal exiting from modulator (2 ns/cm, 2 V/cm); lower trace: detected optical signal after transmission over laser beam (2 ns/cm, 10 mV/cm). (b) Upper trace: applied electrical modulation waveform; lower trace: detected optical signal after transmission over laser beam (500 ps/cm).

to the limit of the photodetector, by reducing the length of the modulator crystal. This increase in bandwidth is paid for at the expense of smaller depths of modulation.

Some small deviations in waveform shape are observable between that exiting from the modulator crystal (Fig. 25a upper trace) and that shown in Fig. 21b for which there is a pure resistive driving point impedance. This is attributable to the slight impedance mismatch between the modulator structure and the transmission line structure in which it is embedded. Its primary cause is reflections in the traveling-wave modulator rather than coupling paths to the electronic circuits since no significant reflections are observed in this path for test waveforms. The effect, however, is not great and the overall result is that of widening the PCM eye pattern, which is a desirable feature.

The noise level observed in the waveform of Fig. 25a is attributable to a combination of the noise in the photodetector, that due to the laser beam, and that due to the sampling noise in the instrument used to observe the waveform. The latter source, which is intrinsically unrelated to the actual system, provides the dominant contribution. Nevertheless, taking the signal-to-noise ratio as the observed 20 dB, results in a possible system information rate in excess of 3 Gb/s in accordance with the Shannon-Hartley law. This information rate would require, of course, a more efficient coding scheme.

Long-term stability of the system is measured in hours. Loss of stability is due to the onset of mode competition resulting from etalon adjustment wandering, although it is not difficult to envisage a correction mechanism via a monitoring and servomechanism scheme to extend this time to arbitrarily long periods. The modulator crystal is not temperature controlled and ambient temperature variations will cause a loss of system bias point. As discussed earlier, this is readily rectified using a cascaded rod structure to achieve bias stabilization.

Once detected, the received signal may be demultiplexed conveniently by existing techniques. A suitable optical technique has been described by Kinsel[51] for demultiplexing two channels and a related technique for an arbitrarily large number of channels has been proposed by Chen.[52] Electronic techniques employing a sampling scheme are also applicable without placing severe constraints on the demultiplexer circuit. This relaxation in requirements results from the relatively small amplitudes of the detected waveform. In repeater applications, the demultiplexed pulses can be conveniently amplified in relatively

low-frequency amplifiers (250 Mb/s) in order to provide drive signals to subsequent information gates.

VIII. CONCLUSIONS

An optical communications system has been described which has a demonstrated information capacity of 1 Gb/s. The major difficulties to achieving high information rates have been outlined as lying in high-speed circuit realization and the interfacing of high-speed electronic circuits and optics. These difficulties have been alleviated by a system architecture and subsystem design which not only results in the high information rate described but shows promise for future systems of even higher rates. These higher rates could be achieved by a multiplexing of quantized electrical pulse amplitudes, and the possibility also exists for a further increase in rate with optical multiplexing.

The realization of the high-speed performance of the electronic circuits relies not only on the design configuration but heavily on the fabrication technology employed. In the case of the 1-Gb/s system a beam lead chip on alumina substrate technology has provided sufficiently low parasitic reactance values to achieve the desired speed performance. In a wholly electronic communications system, the circuits evolved in this study could find advantageous application in achieving data rates which are closer to the most efficient bandwidth utilization rate for coaxial systems.[53]

Transmission channels have not been considered in this paper; these channels could be of an open atmospheric, gas lens guide, or fiber optic type. Recent advances in low-loss fibers[54,55] make these look particularly attractive for optical transmission and efficient utilization of these fibers could call for wide bandwidth schemes of the type described.

Practical laser communications systems already are in use[14,56] and the exponential growth in bandwidth requirements will undoubtedly force a greater exploitation of the laser portions of the electromagnetic spectrum. Even the high information rates described nowhere near approach the theoretical bandwidth capabilities of optical communications systems. This ultimate limit will probably only be realized with improved optical processing of logic so that the potential of the extremely short pulses now available from such lasers as the neodymium YAG and the solid state GaAs forms can be realized.

IX. ACKNOWLEDGEMENTS

It is with pleasure that the author acknowledges all those who have contributed to this work and, in particular, H. J. Schulte and P. K. Runge for stimulating discussions, B. G. King and D. A. Hodges for their encouragement during the course of this work, M. R. Biazzo who mounted and tested the modulator crystals, H. Melchior who supplied the photodiodes, and last but not least G. M. Chin for his excellent technical assistance.

REFERENCES

1. Standley, R. D., "Performance of an 11 GHz Optical Modulator Using $LiTaO_3$," J. Appl. Opt., May 1971, pp. 1022–1023.
2. Melchior, H., and Lynch, W. T., "Signal and Noise Response of High Speed Germanium Avalanche Photodiodes," IEEE Trans. Elec. Devices, *ED-13,* No. 12 (December 1966), pp. 829–838.
3. Denton, R. T., Kinsel, T. S., and Chen, F. S., "224 Mc/s Optical Pulse Code Modulator," Proc. IEEE (Letters), *54,* October 1966, pp. 1472–1473.
4. Gordon, E. I., "A Review of Acoustic-optic Deflection in Modulation Devices," Appl. Opt., *5,* No. 10 (October 1966), pp. 1629–1639.
5. Zeeman, P., *Researches in Magneto-optic,* London: Macmillan and Co., Ltd., 1913.
6. Kaminow, I. P., and Turner, E. H., "Electrooptic Light Modulators," Appl. Opt., *5,* No. 10 (October 1966), pp. 1612–1628.
7. Pockels, F., *Lehrbuck Der Kristalloptik,* Leipzig: B. Teubner, 1906.
8. Kaminow, I. P., and Johnston, W. D., Jr., "Quantitative Determination of Sources of the Electro-optic Effect in $LiNbO_3$ and $LiTaO_3$," Phys. Rev., *160,* No. 3.
9. Ashkin, A., et al., "Optically-Induced Refractive Index Inhomogeneities in $LiNbO_3$ and $LiTaO_3$," Appl. Phys. Letters, *9,* No. 1 (July 1966), pp. 72–74.
10. Denton, R. T., Chen, F. S., and Ballman, A. A., "Lithium Tantalate Light Modulators," J. Appl. Phys., *38,* No. 4 (March 1967), pp. 1611–1617.
11. Smith, P. W., "Mode Locking of Lasers," Proc. IEEE, *58,* No. 9 (September 1970), pp. 1342–1357.
12. Kogelnik, H., and Patel, C. K. N., "Mode Suppression and Single Frequency Operation in Gaseous Optical Masers," Proc. IRE, *50,* No. 11 (November 1962).
13. Denton, R. T., and Kinsel, T. S., "Terminals for a High-Speed Optical Pulse Code Modulation Communications System: I. 224-Mbit/s Single Channel," Proc. IEEE, *56,* No. 2 (February 1968), pp. 140–145.
14. Masuda, T., Uchida, T., Ueno, Y., and Shimamura, T., "An Experimental High Speed PCM/AM Optical Communications System Using Mode-Locked HeNe Gas Lasers," IEEE Int. Conf. Commun., San Francisco, June 8–10, 1970.
15. Di Domenico, M., Jr., Geusic, J. E., Marcos, H. M., and Smith, R. G., "Generation of Ultrashort Optical Pulses by Mode Locking the YAlG:Nd Laser," Appl. Phys. Letters, *8,* No. 7 (April 1966), pp. 180–183.
16. Kinsel, T. S., and Denton, R. T., "Terminals for a High-Speed Optical Pulse Code Modulation Communications System: II. Optical Multiplexing and Demultiplexing," Proc. IEEE, *56,* No. 2 (February 1968), pp. 146–154.
17. Hirano, J., and Kimura, T., "Multiple Mode Locking of Lasers," IEEE J. Quantum Elec., *QE-5,* No. 5 (May 1969), pp. 219–225.

18. Runge, P., "Ein lm langer He-Ne-Laser für 0.63 μm mit nur einer einzigen Frequenz," Archiv der Elektrischen Übertragung, *19,* 1965, pp. 573–574.
19. Kinsel, T. S., "Wide-Band Optical Communication Systems: Part I—Time Division Multiplexing," Proc. IEEE, *58,* No. 10 (October 1970), pp. 1666–1683.
20. Hance, H. V., Peterson, D. G., Ward, R. B., and Ohlman, R. C., "A Laboratory Ultra-Wideband FM/IM Optical Data Relay System," IEEE Int. Conf. Commun. San Francisco, June 8–10, 1970.
21. Gabler, F., and Sokob, P., "Senarmont Compensator," Z. Instrumentenk., *58,* 1938, p. 301.
22. Poincaré, H., *Théorie Mathématique de la Lumiére,* Paris: Gauthiers-Villars, 1892.
23. Kogelnik, H., and Li, T., "Laser Beams and Resonators," Appl. Opt., *5,* No. 10 (October 1966), pp. 1550–1567.
24. Kogelnik, H., "Modes in Optical Resonators," *Lasers,* Vol. 1, Ed. A. K. Levine, New York: Marcel Dekker, Inc., 1966.
25. Bhawalkar, D. D., Goncharenko, A. M., and Smith, R. C., "Propagation of Gaussian Beams in Anisotropic Media," Brit. J. App. Phys., *18,* pp. 1431–1441.
26. Levinstein, H. J., et al., "Reduction of the Susceptability of Optically Induced Index Inhomogeneities in $LiTaO_3$ and $LiNbO_3$," J. Appl. Phys., *38,* No. 8 (July 1967), pp. 3101–3102.
27. Riesz, R. P., and Biazzo, M. R., "Gigahertz Optical Modulation," Appl. Opt., *8,* 1969, pp. 1393–1396.
28. Chow, K. K., Comstock, R. L., and Leonard, W. B., "1.5 GHz Bandwidth Light Modulator," IEEE J. Quantum Elec., *QE-5,* No. 12 (December 1969), pp. 618–620.
29. Rigrod, W. W., and Kaminow, I. P., "Wideband Light Modulation," Proc. IEEE, *51,* No. 1 (January 1963), pp. 1–4.
30. Bicknell, W. E., Yap, B. K., and Peters, C. J., "0 to 3 GHz Traveling-Wave Electro-optic Modulator," Proc. IEEE, *55,* No. 2 (February 1967), pp. 225–226.
31. Di Domenico, M., Jr., and Anderson, L. K., "Broadband Electro-optic Traveling-Wave Light Modulators," B.S.T.J., *42,* No. 6 (November 1963), pp. 2621–2678.
32. Grover, F. W., *Inductance Calculations, Working Formulas and Tables,* New York: Dover Publications, Inc., 1946.
33. Spencer, E. G., Lenzo, P. V., and Ballman, A. A., "Dielectric Materials for Electro-optic, Elasto-optic and Ultrasonic Device Applications," Proc. IEEE, *55,* No. 12 (December 1967), pp. 2074–2108.
34. Kaminow, I. P., and Sharpless, W. M., "Performance of $LiTaO_3$ and $LiNbO_3$ Light Modulators at 4 GHz," Appl. Opt. *6,* 1967, p. 351.
35. Kaminow, I. P., Kompfner, R., and Louisell, W. H., "Improvements in Light Modulators of the Traveling-Wave Type," IRE Trans. on Microwave Theory and Techniques, *MTT-10,* No. 5 (September 1962).
36. Kaminow, I. P., and Liu, J., "Propagation Characteristics of Partially Loaded Two-Conductor Transmission Line for Broadband Light Modulators," Proc. IEEE, *51,* No. 1 (January 1963), pp. 132–136.
37. Peters, C. J., "Gigacycle Bandwidth Coherent Light Traveling-Wave Phase Modulator," Proc. IEEE, *51,* No. 1 (January 1963), pp. 147–152.
38. Nelson, D. F., and Turner, E. H., "Electro-optic and Piezo-electric Coefficients and Refractive Index of Gallium Phosphide," J. Appl. Phys., *39,* No. 7 (June 1968), pp. 3337–3343.
39. Bond, W. L., "Measurement of the Refractive Indices of Several Crystals," J. Appl. Phys., *36,* No. 2 (May 1965), pp. 1674–1677.
40. Kaminow, I. P., "Microwave Modulation of the Electro-optic Effect in KH_2PO_4," Phys. Rev. Letters *6,* No. 10 (May 1961), pp. 528–530.
41. Peters, C. J., "Traveling Wave Amplitude Modulator," NEREM Record 1964, November 1964, pp. 70–71.
42. Peters, C. J., "Gigacycle-Bandwidth Coherent-Light Traveling Wave Amplitude Modulator," Proc. IEEE, *53,* No. 5 (May 1965), pp. 455–460.

43. Biazzo, M. R., "Fabrication of a Lithium Tantalate Temperature Stabilized Optical Modulator," Appl. Opt., *10,* No. 5 (May 1971).
44. Friis, H., "Analysis of Harmonic Generator Circuits for Step Recovery Diodes," Proc. IEEE, *55,* No. 7 (July 1967), pp. 1192–1194.
45. Dietrich, A. F., and Goodall, W. M., "Solid State Generator for 2 × 10^{-11} Second Pulses," Proc. IRE (Correspondence), *48,* 1960, pp. 791–792.
46. Gray, P. E., et al., *Physical Electronics and Circuit Properties of Transistors,* New York: John Wiley and Sons, 1964.
47. Kirk, C. T., Jr., "A Theory of Transistor Cut-off Frequency (f_T) Fall-off at High Current Density," IEEE Trans. Elec. Devices, *ED-9,* No. 2 (March 1962), pp. 164–173.
48. Lepselter, M. P., and Sze, S. M., "Silicon Schottky Barrier Diode with Near-Ideal I-V Characteristics," B.S.T.J., *47,* No. 2 (February 1968), pp. 195–208.
49. Mead, C. A., "Metal-Semiconductor Surface Barriers," Solid State Elec., *9,* 1966, pp. 1023–1033.
50. Arnaud, J. A., Franke, E. A., Hubbard, W. M., Mandeville, G D., de la Claviere, B., and Franke, J. M., "Technique for Fast Measurement of Gaussian-Beam Parameters," J. Opt. Soc. Am., *60,* No. 5 (May 1970), p. 739.
51. Kinsel, T. S., "Lightwave of the Future: Optical PCM," Electronics, September 16, 1968, pp. 123–128.
52. Chen, F. S., "Demultiplexers for High-Speed Optical PCM," IEEE J. Quantum Elec., *QE-7,* No. 1 (January 1971), pp. 24–29.
53. Freeny, S. L., and Chang, R. W., "Hybrid Digital Transmission Systems—Part 2: Information Rate of Hybrid Coaxial Cable Systems," B.S.T.J., *47,* No. 8 (October 1968), pp. 1687–1711.
54. Kapron, F. P., Keck, D. B., and Maurer, R. D., "Radiation Losses in Glass Optical waveguides," IEE Conf. Trunk Telecommunications by Guided Waves, September 20, 1970, London, p. 29.
55. Kapron, F. P., Keck, D. B., and Maurer, R. D., "Radiation Losses in Glass Optical Waveguides," Appl. Phys. Letters, *77,* November 15, 1970, p. 423.
56. IEEE Spectrum, 7, No. 6 (June 1970), p. 120.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Analysis of Dependence Effects in Telephone Trunking Networks

By J. M. HOLTZMAN

(Manuscript received March 10, 1971)

Since theoretical and computational difficulties often preclude exact solution of telephone trunking network problems, approximate methods are naturally used. A typical approach is to determine link blocking probabilities and from them calculate point-to-point blocking probabilities by invoking independence assumptions. Although the link blocking probabilities may be quite accurate, the point-to-point blocking probability calculations will, in some cases, suffer from the independence assumptions. This paper presents a method of taking dependence into account for certain networks by approximating conditional probabilities which reflect the dependence. The approximations avoid the problems of dealing with the large sums associated with this problem.*

I. INTRODUCTION

The exact analysis of telephone trunking networks often leads to severe computational problems due, e.g., to the large number of possible states. Approximate methods are thus naturally used. There has been much success in approximately calculating link blocking probabilities but less in determining point-to-point blocking probabilities. Errors in the point-to-point blocking probabilities can be caused by independence assumptions.

The purpose of this paper is to take advantage of existing techniques for approximating link blocking probabilities (which are quite accurate) and develop an approach for taking link dependences into account. In particular, we present a method for approximating the appropriate conditional probabilities for cases where the traffics are Poisson (or close to Poisson). The extension of the approach to the case of distinctly non-Poisson processes, such as arise in overflows, will be reported in Ref. 1. To obtain point-point blocking probabilities for non-Poisson processes,

* We use link to denote trunk group.

one must also solve for the blocking seen by individual traffics when more than one traffic is offered to a link (the equivalent random method gives the blocking seen by the combined traffic). This is also treated in Ref. 1.

To clarify the role of independence assumptions, suppose there is common traffic on links 1 and 2. Let A_1 and A_2 be the events that the common traffic is blocked on links 1 and 2, respectively. Then the probability of being blocked on either link 1 or on link 2, $P\{A_1 \cup A_2\} = P\{A_1\} + P\{A_2\} - P\{A_1 \cap A_2\}$, is often approximated by $P\{A_1 \cup A_2\} \cong P\{A_1\} + P\{A_2\} - P\{A_1\}P\{A_2\}$ or even by $P\{A_1 \cup A_2\} \cong P\{A_1\} + P\{A_2\}$. The last approximation is clearly accurate if $P\{A_1\} + P\{A_2\} \gg P\{A_1 \cap A_2\}$ and the first is accurate if $P\{A_1 \mid A_2\} \cong P\{A_1\}$ or if both $P\{A_1 \cap A_2\}$ and $P\{A_1\}P\{A_2\}$ are relatively small. On the other hand, it is easy to give examples where the neglect of dependence leads to non-negligible errors.

The errors due to neglecting dependence naturally depend strongly on mutual traffic. To see this in a transparent case, let us reconsider the two-link network mentioned above with each link having the same number of trunks. $P\{A_1 \mid A_2\}$ could vary from $P\{A_1\}$ (when the mutual traffic is zero) to unity (when there is no traffic on link 2 not shared with link 1). In the latter extreme case, assuming independence would give $P\{A_1\} + P\{A_2\}\ (1 - P\{A_1\})$ for $P\{A_1 \cup A_2\}$ compared to the correct answer $P\{A_1\}$. Thus, assuming independence could, in this case, conceivably overestimate the point-to-point blocking probabilities by something approaching 100 percent. In fact, it is possible to overestimate a point-to-point blocking probability on m tandem links by almost as much as m times by assuming independence.

II. BASIC APPROACH

For simplicity of explanation, first consider the situation shown in Fig. 1. λ_1, λ_2, and λ_{12} are the parameters of mutually independent Poisson processes. Holding times are mutually independent exponential random variables with unity mean (or the mean is the time unit) here and throughout this paper. Also, throughout the paper we shall assume that lost calls are cleared and that the system is in equilibrium. Links 1 and 2 have N_1 and N_2 trunks, respectively. Assume $N_1 \geqq N_2$.

Let A_1 and A_2 denote the events of link 1 and link 2 being blocked, respectively (i.e., there are N_1 calls up on link 1 and N_2 calls up on link 2).

All of the equilibrium state probabilities may actually be expressed in

closed form using, e.g., the form of solution given in Ref. 2, Section 7. For example,

$$P\{A_1 \cup A_2\} = \frac{\sum_{i_{12}=0}^{N_2} \frac{\lambda_{12}^{i_{12}}}{i_{12}!} \sum_{i_1=0}^{N_1-i_{12}} \frac{\lambda_1^{i_1}}{i_1!} \sum_{i_2=0}^{N_2-i_{12}} \frac{\lambda_2^{i_2}}{i_2!}}{\sum_{\substack{i_1+i_{12}\leqq N_1 \\ i_2+i_{12}\leqq N_2}} \frac{\lambda_1^{i_1}}{i_1!} \frac{\lambda_2^{i_2}}{i_2!} \frac{\lambda_{12}^{i_{12}}}{i_{12}!}}, \tag{1}$$

where the indices are nonnegative integers (throughout the paper). However, the use of such exact results very quickly becomes impractical for all but the smallest problems even for large computers because of the number of computations required.

The link blocking probabilities can usually be quite well approximated. For example, the following reduced load equations,

$$P_1 = B(N_1 , \lambda_1 + \lambda_{12}(1 - P_2)), \tag{2}$$

$$P_2 = B(N_2 , \lambda_2 + \lambda_{12}(1 - P_1)), \tag{3}$$

$$P\{A_1\} \cong P_1 , \tag{4}$$

$$P\{A_2\} \cong P_2 , \tag{5}$$

($B(N, \lambda)$ is the Erlang B formula) are often sufficiently accurate as they stand.* See Ref. 3 for much more discussion of link blocking probabilities.

Once the link blocking probabilities are determined (however they are determined) all that is left to calculate is $P\{A_1 \mid A_2\}$ in order to determine $P\{A_1 \cup A_2\}$. We first write down this probability exactly and then show a simple practical approximation. (Approximating the sums becomes even more important when considering more complicated dependences as in Section III.) To this end, note that

$$P\{A_1 \mid A_2\} = \frac{\sum_{i_{12}=0}^{N_2} \frac{\lambda_{12}^{i_{12}}}{i_{12}!} \frac{\lambda_2^{N_2-i_{12}}}{(N_2 - i_{12})!} \frac{\lambda_1^{N_1-i_{12}}}{(N_1 - i_{12})!}}{\sum_{i_{12}=0}^{N_2} \frac{\lambda_{12}^{i_{12}}}{i_{12}!} \frac{\lambda_2^{N_2-i_{12}}}{(N_2 - i_{12})!} \sum_{i_1=0}^{N_1-i_{12}} \frac{\lambda_1^{i_1}}{i_1!}}$$

$$= \frac{\sum_{i=0}^{N_2} \binom{N_2}{i} p^i q^{N_2-i} \frac{\lambda_1^{N_1-i}}{(N_1 - i)!}}{\sum_{i=0}^{N_2} \binom{N_2}{i} p^i q^{N_2-i} \sum_{i_1=0}^{N_1-i} \frac{\lambda_1^{i_1}}{i_1!}} \tag{6}$$

* It may easily be shown that these P_1 and P_2 agree exactly with $P\{A_1\}$ and $P\{A_2\}$ to first-order terms in λ_{12}. Although the P_i usually overestimate the $P\{A_i\}$, they can underestimate (seriously, in extreme cases) when λ_{12} is very large.

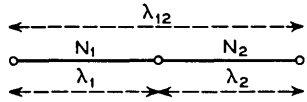


Fig. 1—Two-link network.

where

$$p = \frac{\lambda_{12}}{\lambda_{12} + \lambda_2}, \tag{7}$$

$$q = 1 - p. \tag{8}$$

Observe that (6) is the ratio of two sums each of which is essentially a Bernstein polynomial (see Ref. 4). However, a Bernstein polynomial is of the form $\sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} f(i/n)$ but our f is actually a function of i or of $n(i/n)$. Each sum is also the expectation of a function of a binomially distributed random variable. Using the following interpolation formula (see Ref. 5, p. 178),

$$f(i) \cong f(\bar{\imath}) + (i - \bar{\imath})\tfrac{1}{2}[f(\bar{\imath} + 1) - f(\bar{\imath} - 1)] + (i - \bar{\imath})^2\tfrac{1}{2}[f(\bar{\imath} + 1) - 2f(\bar{\imath}) + f(\bar{\imath} - 1)], \tag{9}$$

and taking expectations yields

$$\sum_{i=1}^{n} \binom{n}{i} p^i q^{n-i} f(i) \cong f(np) + \frac{npq}{2}\,[f(np + 1) - 2f(np) + f(np - 1)]. \tag{10}$$

(10) is the usual expression for the mean and variance approximation to an expectation of a function of a random variable but with the second derivative replaced by a central difference. Using (10) on the numerator and denominator of (6) yields

$$P\{A_1 \mid A_2\} \cong P_a\{A_1 \mid A_2\} = \frac{1 + \dfrac{N_2pq}{2}\left[\dfrac{\lambda_1}{N_1 - N_2p + 1} - 2 + \dfrac{N_1 - N_2p}{\lambda_1}\right]}{\dfrac{1}{B(N_1 - N_2p,\, \lambda_1)} + \dfrac{N_2pq}{2}\left[\dfrac{\lambda_1}{N_1 - N_2p + 1} - 1\right]}. \tag{11}$$

Interpolation, such as given on page 571 of Ref. 3, may be used to evaluate the Erlang B formula for a nonintegral number of trunks or an approximation for Erlang B in integral form can be used.

Note that we use the same type of approximation on the numerator

and denominator, which are closely related, so that the ratio can be more accurate than either taken separately.

(11) shows quantitatively what we expect qualitatively, namely, that the dependence effect gets small as the common traffic gets small. In particular, when $\lambda_{12} = 0$, $P\{A_1 \mid A_2\} = B(N_1, \lambda_1) = P\{A_1\}$ and when $\lambda_2 = 0$, $P\{A_1 \mid A_2\} = B(N_1 - N_2, \lambda_1)$(recall that we assumed $N_1 \geqq N_2$).

To examine the approximation, we find it convenient to first consider $P_a\{A_1^c \mid A_2\}$ where superscript c denotes complement. Actually, since $P\{A_1 \cup A_2\} = P\{A_1\} + P\{A_1^c \mid A_2\}P\{A_2\}$, $P\{A_1^c \mid A_2\}$ is really the crucial quantity. With $N_1 \geqq N_2$, we have

$$P_a\{A_1^c \mid A_2\} = 1 - P_a\{A_1 \mid A_2\}$$

$$= \frac{\dfrac{1}{B(N_1 - N_2p, \lambda_1)} - 1 + \dfrac{N_2pq}{2}\left[1 - \dfrac{2(N_1 - N_2p)}{\lambda_1}\right]}{\dfrac{1}{B(N_1 - N_2p, \lambda_1)} + \dfrac{N_2pq}{2}\left[\dfrac{\lambda_1}{N_1 - N_2p + 1} - 2\right]}. \tag{12}$$

If we let $N_1 = N_2 = N$, then

$$P_a\{A_1^c \mid A_2\} = \frac{\dfrac{1}{B(Nq, \lambda_1)} - 1 + \dfrac{N_2pq}{2}\left[1 - \dfrac{2Nq}{\lambda_1}\right]}{\dfrac{1}{B(Nq, \lambda_1)} + \dfrac{Npq}{2}\left[\dfrac{\lambda_1}{Nq + 1} - 2\right]}. \tag{13}$$

(12) could have been derived directly using a mean and second central difference approximation just as (6) was approximated by (11). The terms multiplied by $N_2pq/2$ in both the numerator and denominator are then the second central difference terms. If we hold λ_1 and p fixed and let N get large, it can be shown that the second central difference terms get small compared to the mean terms suggesting that the approximation is accurate for large N. However, since $P_a\{A_1 \mid A_2\} \to 0$ as $N \to \infty$ (with λ_1 and p fixed), investigation of this type of convergence is of limited practical value.

It is probably of more interest to examine the approximation for fixed p and to let λ_1 get large with N. To this end, let $\lambda_1 = kN$ with $k > q$. For sufficiently large n and $a > n$, $B(n, a) \cong 1 - n/a$ so that

$$B(Nq, kN) \cong 1 - \frac{q}{k}. \tag{14}$$

Hence, we obtain for large N,

$$P_a\{A_1 \mid A_2\} = \frac{1 + \frac{Npq}{2}\left[\frac{kN}{Nq+1} - 2 + \frac{Nq}{kN}\right]}{\frac{1}{1-\frac{q}{k}} + \frac{Npq}{2}\left[\frac{kN}{Nq+1} - 1\right]}$$

$$\cong B(Nq, kN) \tag{15}$$

which has the interesting interpretation that with link 2 full, the number of trunks on link 1 to handle the λ_1 traffic is reduced by the average number of calls on link 2 which are common to link 1 (we shall elaborate on this below). Although (15) is intuitively appealing and obviously correct for $p = 0$ or 1, its accuracy should be examined for $p \in (0, 1)$. Observe that for large N, the variance terms dominate.

By elaborating on the interpretation alluded to above, we can see why $B(Nq, \lambda_1)$ should tend to overestimate $P\{A_1 \mid A_2\}$. With D_i the event that i calls from the λ_{12} traffic are up,

$$D_i = \{i\ \lambda_{12} \text{ calls up}\}, \tag{16}$$

observe that

$$P\{A_1 \mid A_2\} = \sum_{i=0}^{N_2} P\{A_1 \mid A_2 \cap D_i\} P\{D_i \mid A_2\} \tag{17}$$

$$= \sum_{i=0}^{N_2} B(N_1 - i, \lambda_1) P\{D_i \mid A_2\}. \tag{18}$$

Now, *if* the λ_{12} calls were never blocked on link 1, then $P\{D_i \mid A_2\} = \binom{n}{i} p^i q^{n-i}$. With $N_1 = N_2 = N$, $\lambda_1 = kN$, it may be shown that $\sum_{i=0}^{N_2} B(N_1 - i, \lambda_1)\binom{n}{i} p^i q^{n-i} \to B(N_1 - N_2 p, \lambda_1) = B(Nq, \lambda)$ as $N \to \infty$. $B(N_1 - N_2 p, \lambda_1)$ is the blocking probability on link 1 when the number of trunks is reduced by the conditional mean of trunks occupied by λ_{12} calls. Hence, under these conditions, $B(Nq, \lambda_1)$ approximates $P\{A_1 \mid A_2\}$ by ignoring the blocking of λ_{12} calls on link 1. If we take this blocking into account, we would expect that for large N, $P\{A_1 \mid A_2\} < B(Nq, \lambda_1)$. This will be seen to be the case below.

Another approximation for $P\{A_1 \mid A_2\}$ is

$$P_b\{A_1 \mid A_2\} = \frac{r_1^{-N_1}(q + pr_1)^{N_2} e^{\lambda_1 r_1} / \sqrt{2\pi\beta_1}}{r_2^{-N_1}(q + pr_2)^{N_2} e^{\lambda_1 r_2} / (1 - r_2)\sqrt{2\pi\beta_2}} \tag{19}$$

where

$$r_1 = \frac{\sqrt{(\lambda_1 q - N_1 p + N_2 p)^2 + 4\lambda_1 N_1 pq} - \lambda_1 q + N_1 p - N_2 p}{2\lambda_1 p}, \tag{20}$$

$$\beta_1 = N_1 - N_2\left(\frac{pr_1}{q + pr_1}\right)^2, \tag{21}$$

$$\beta_2 = N_1 - N_2\left(\frac{pr_2}{q + pr_2}\right)^2 + \left(\frac{r_2}{1 - r_2}\right)^2, \tag{22}$$

and r_2 is the solution between 0 and 1 of the cubic

$$\lambda_1 p r_2^3 + (N_2 p - p - N_1 p + \lambda_1 q - \lambda_1 p) r_2^2 + (-\lambda_1 q + N_1 p - N_1 q - N_2 p - q) r_2 + N_1 q = 0. \tag{23}$$

This approximation is due to D. L. Jagerman (see the derivation in the Appendix).

For $N_1 = N_2 = N$, $\lambda_1 = kN$, and N large,

$$P_b\{A_1 \mid A_2\} \cong 1 - \frac{\sqrt{k^2q^2 + 4kpq} - kq}{2kp}. \tag{24}$$

Table I shows some numerical values of the exact and approximate conditional probabilities along with the values of the approximations for large N. It is seen that, although $P_b\{A_1 \mid A_2\}$ does not always surpass the simpler $P_a\{A_1 \mid A_2\}$, it is, on the whole, superior and its behavior with increasing N is clearly better. We dwelled on $P_a\{A_1 \mid A_2\}$ because it is useful in many cases and the mean and variance approach is easily

TABLE I—NUMERICAL VALUES

p	$\lambda_1 = N$	$P\{A_1 \mid A_2\}$	$P_a\{A_1 \mid A_2\}$	$P_b\{A_1 \mid A_2\}$	% error in $P_a\{A_1 \mid A_2\}$	% error in $P_b\{A_1 \mid A_2\}$
0.1	5	0.3256	0.3246	0.3404	−0.3	4.5
0.1	10	0.26087	0.26091	0.2732	0.015	4.7
0.1	15	0.2296	0.2287	0.2402	−0.4	4.6
0.1	20	0.2101	0.2096	0.2197	−0.2	4.5
0.1	25	0.1966	0.1957	0.2052	−0.4	4.4
0.1	∞		0.1	0.092		
0.5	5	0.5017	0.4926	0.5125	−1.8	2.15
0.5	10	0.4579	0.4635	0.4645	1.2	1.4
0.5	15	0.4386	0.4587	0.4431	4.6	1.0
0.5	20	0.4274	0.4599	0.4307	7.6	0.8
0.5	25	0.4200	0.4621	0.4226	10.0	0.6
0.5	∞		0.5	0.382		
0.9	5	0.7806	0.8161	0.7827	4.5	0.25
0.9	10	0.7521	0.8258	0.7535	9.8	0.19
0.9	15	0.7416	0.8412	0.7421	13.4	0.11
0.9	20	0.7360	0.8525	0.7366	15.8	0.07
0.9	25	0.7326	0.8604	0.7330	17.4	0.05
0.9	∞		0.9	0.718		

extendable to cases where extreme accuracy is not required (as in Section III). It can be easily seen that a percentage error in the conditional probability typically causes a smaller percent error in the point-to-point blocking probability.

III. MORE COMPLICATED DEPENDENCES

The extension of the basic idea of Section II is often rather straightforward. For example, consider the network in Fig. 2 which is complicated enough to illustrate the ingredients of a general approach.

Assume (for the sake of being concrete) that $N_1 \geqq N_3 \geqq N_2$. All of the traffics are mutually independent Poisson. Let A_1, A_2, A_3 represent the events of links, 1, 2, 3 being blocked, respectively. Then the following are approximations to conditional probabilities:

$$P\{A_1 \mid A_2\} \cong \frac{\sum_{i=0}^{N_2} \binom{N_2}{i} p_{12}^i q_{12}^{N_2-i} \dfrac{\lambda_1^{N_1-i}}{(N_1-i)!}}{\sum_{i=0}^{N_2} \binom{N_2}{i} p_{12}^i q_{12}^{N_2-i} \sum_{j=0}^{N_1-i} \dfrac{\lambda_1^j}{j!}}, \tag{25}$$

$$P\{A_3 \mid A_2\} \cong \frac{\sum_{i=0}^{N_2} \binom{N_2}{i} p_{32}^i q_{32}^{N_2-i} \dfrac{\lambda_3^{N_3-i}}{(N_3-i)!}}{\sum_{i=0}^{N_2} \binom{N_2}{i} p_{32}^i q_{32}^{N_2-i} \sum_{j=0}^{N_3-i} \dfrac{\lambda_3^j}{j!}}, \tag{26}$$

$$P\{A_1 \mid A_3\} \cong \frac{\sum_{i=0}^{N_3} \binom{N_3}{i} p_{13}^i q_{13}^{N_3-i} \dfrac{(\bar{\lambda}_1)^{N_1-i}}{(N_1-i)!}}{\sum_{i=0}^{N_3} \binom{N_3}{i} p_{13}^i q_{13}^{N_3-i} \sum_{j=0}^{N_1-i} \dfrac{(\bar{\lambda}_1)^j}{j!}}, \tag{27}$$

where

$$p_{12} = \frac{\lambda_{12} + \lambda_{13}(1 - P_3)}{\lambda_{12} + \lambda_{13}(1 - P_3) + \lambda_2 + \lambda_{23}(1 - P_3)}, \qquad q_{12} = 1 - p_{12}, \tag{28}$$

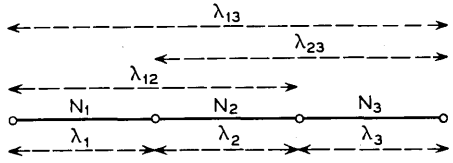


Fig. 2—Three-link network.

$$p_{32} = \frac{\lambda_{23} + \lambda_{13}(1 - P_1)}{\lambda_{23} + \lambda_{13}(1 - P_1) + \lambda_2 + \lambda_{12}(1 - P_1)}, \qquad q_{32} = 1 - p_{32}, \quad (29)$$

$$p_{13} = \frac{\lambda_{13}(1 - P_2)}{\lambda_{13}(1 - P_2) + \lambda_3 + \lambda_{23}(1 - P_2)}, \qquad q_{13} = 1 - p_{13}, \quad (30)$$

$$\overline{\lambda_1} = \lambda_1 + \lambda_{12}\left(1 - \frac{P\{A_3 \mid A_2\}P_2}{P_3}\right), \quad (31)$$

and the P_i are approximations to the link blocking probabilities. In (31), the approximation to $P\{A_3 \mid A_2\}$ is used (and we are obviously assuming $P_3 > 0$).

The rationale behind (25)–(27) is based on (6) at least in the case where the P_i are small so that the carried traffics are not too far from Poisson. We have already discussed approximating the sums of (25)–(27) in Section II.

But for the network on Fig. 2 we should also calculate $P\{A_1 \cup A_2 \cup A_3\}$, the blocking probability seen by the λ_{13} traffic. Since

$$\begin{aligned} P\{A_1 \cup A_2 \cup A_3\} &= P\{A_1\} + P\{A_2\} + P\{A_3\} \\ &\quad - P\{A_1 \mid A_2\}P\{A_2\} - P\{A_1 \mid A_3\}P\{A_3\} \\ &\quad - P\{A_3 \mid A_2\}P\{A_2\} + P\{A_1 \cap A_3 \mid A_2\}P\{A_2\}, \end{aligned} \quad (32)$$

we see that we need only discuss the evaluation of $P\{A_1 \cap A_3 \mid A_2\}$, the other quantities already having been treated.

Just as we derived (6), we can show that (exactly)

$$P\{A_1 \cap A_3 \mid A_2\} = \frac{\displaystyle\sum_{i_{12}+i_{23}+i_{13}+i_2=N_2} \frac{N_2!\, p_{12}^{i_{12}} p_{23}^{i_{23}} p_{13}^{i_{13}} p_2^{i_2}}{i_{12}!\, i_{23}!\, i_{13}!\, i_2!} \frac{\lambda_3^{N_3-i_{23}-i_{13}}}{(N_3 - i_{23} - i_{13})!} \frac{\lambda_1^{N_1-i_{12}-i_{13}}}{(N_1 - i_{12} - i_{13})!}}{\displaystyle\sum_{i_{12}+i_{23}+i_{13}+i_2=N_2} \frac{N_2!\, p_{12}^{i_{12}} p_{23}^{i_{23}} p_{13}^{i_{13}} p_2^{i_2}}{i_{12}!\, i_{23}!\, i_{13}!\, i_2!} \sum_{i_3=0}^{N_3-i_{23}-i_{13}} \frac{\lambda_3^{i_3}}{i_3!} \sum_{i_1=0}^{N_1-i_{12}-i_{13}} \frac{\lambda_1^{i_1}}{i_1!}} \quad (33)$$

where now

$$p_{12} = \frac{\lambda_{12}}{\lambda_{12} + \lambda_{23} + \lambda_{13} + \lambda_2}, \quad (34)$$

$$p_{23} = \frac{\lambda_{23}}{\lambda_{12} + \lambda_{23} + \lambda_{13} + \lambda_2}, \quad (35)$$

$$p_{13} = \frac{\lambda_{13}}{\lambda_{12} + \lambda_{23} + \lambda_{13} + \lambda_2}, \tag{36}$$

$$p_2 = \frac{\lambda_2}{\lambda_{12} + \lambda_{23} + \lambda_{13} + \lambda_2}. \tag{37}$$

These sums may be approximated by recognizing that they are both expectations of functions of multinominal random variables. Thus,

$$\begin{aligned} E[f(i_{23}, i_{13}, i_{12})] &\cong f(\bar{\imath}_{23}, \bar{\imath}_{13}, \bar{\imath}_{12}) \\ &+ \tfrac{1}{2}\{[E(i_{23} - \bar{\imath}_{23})^2 \delta_{23}^2 + E(i_{13} - \bar{\imath}_{13})^2 \delta_{13}^2 + E(i_{12} - \bar{\imath}_{12})^2 \delta_{12}^2] \\ &+ 2E[(i_{23} - \bar{\imath}_{23})(i_{13} - \bar{\imath}_{13})]\mu\delta_{23}\mu\delta_{13} \\ &+ 2E[(i_{23} - \bar{\imath}_{23})(i_{12} - \bar{\imath}_{12})]\mu\delta_{23}\mu\delta_{12} \\ &+ 2E[(i_{13} - \bar{\imath}_{13})(i_{12} - \bar{\imath}_{12})\mu\delta_{13}\mu\delta_{12}]\}f(\bar{\imath}_{23}, \bar{\imath}_{13}, \bar{\imath}_{12}) \end{aligned} \tag{38}$$

where $\bar{\imath}_{nm}$ here indicates expected value and where δ^2 is a second central difference, e.g.,

$$\begin{aligned} \delta_{23}^2 f(\bar{\imath}_{23}, \bar{\imath}_{13}, \bar{\imath}_{12}) &= f(\bar{\imath}_{23} + 1, \bar{\imath}_{13}, \bar{\imath}_{12}) - 2f(\bar{\imath}_{23}, \bar{\imath}_{13}, \bar{\imath}_{12}) \\ &\quad + f(\bar{\imath}_{23} - 1, \bar{\imath}_{13}, \bar{\imath}_{12}) \end{aligned} \tag{39}$$

and the operator $\mu\delta$ is given, e.g., by

$$\mu\delta_{13} f(\bar{\imath}_{23}, \bar{\imath}_{13}, \bar{\imath}_{12}) = \tfrac{1}{2}[f(\bar{\imath}_{23}, \bar{\imath}_{13} + 1, \bar{\imath}_{12}) - f(\bar{\imath}_{23}, \bar{\imath}_{13} - 1, \bar{\imath}_{12})] \tag{40}$$

so that, e.g.,

$$\begin{aligned} &\mu\delta_{23}\mu\delta_{13} f(\bar{\imath}_{23}, \bar{\imath}_{13}, \bar{\imath}_{12}) \\ &\quad = \tfrac{1}{4}[f(\bar{\imath}_{23} + 1, \bar{\imath}_{13} + 1, \bar{\imath}_{12}) + f(\bar{\imath}_{23} - 1, \bar{\imath}_{13} - 1, \bar{\imath}_{12}) \\ &\qquad - f(\bar{\imath}_{23} - 1, \bar{\imath}_{13} + 1, \bar{\imath}_{12}) - f(\bar{\imath}_{23} + 1, \bar{\imath}_{13} - 1, \bar{\imath}_{12})]. \end{aligned} \tag{41}$$

The variances and covariances for the multinominal distribution are given on page 164 of Ref. 6. Thus,

$$\bar{\imath}_{jk} = N_2 p_{jk}, \qquad jk = 23, 13, 12, \tag{42}$$

$$\sigma_{jk}^2 \equiv E(i_{jk} - \bar{\imath}_{jk})^2 = N_2 p_{jk}(1 - p_{jk}), \qquad jk = 23, 13, 12, \tag{43}$$

$$\sigma_{jk,lm}^2 = E[(i_{jk} - \bar{\imath}_{jk})(i_{lm} - \bar{\imath}_{lm})] = -N_2 p_{jk} p_{lm}, \qquad jk, lm = 23, 13, 12 (jk \neq lm). \tag{44}$$

One obtains

$$P\{A_1 \cap A_3 \mid A_2\} \cong \Big\{1 + \frac{1}{2}\sigma_{23}^2\Big(\frac{\bar{N}_3}{\lambda_3} - 2 + \frac{\lambda_3}{\bar{N}_3 + 1}\Big)$$

$$+ \frac{1}{2}\sigma_{13}^2\Big(\frac{\bar{N}_3\bar{N}_1}{\lambda_3\lambda_1} - 2 + \frac{\lambda_3\lambda_1}{(\bar{N}_3 + 1)(\bar{N}_1 + 1)}\Big) + \frac{1}{2}\sigma_{12}^2\Big(\frac{\bar{N}_1}{\lambda_1} - 2 + \frac{\lambda_1}{\bar{N}_1 + 1}\Big)$$

$$+ \frac{1}{4}\Big[\sigma_{23,13}^2\Big(\frac{\bar{N}_3(\bar{N}_3 - 1)}{\lambda_3^2}\frac{\bar{N}_1}{\lambda_1} - 2 + \frac{\lambda_3^2\lambda_1}{(\bar{N}_3 + 2)(\bar{N}_3 + 1)(\bar{N}_1 + 1)}\Big)$$

$$+ \sigma_{23,12}^2\Big(\frac{\bar{N}_3\bar{N}_1}{\lambda_3\lambda_1} - 2 + \frac{\lambda_3\lambda_1}{(\bar{N}_3 + 1)(\bar{N}_1 + 1)}\Big)$$

$$+ \sigma_{13,12}^2\Big(\frac{\bar{N}_3\bar{N}_1(\bar{N}_1 - 1)}{\lambda_3\lambda_1^2} - 2 + \frac{\lambda_3\lambda_1^2}{(\bar{N}_3 + 1)(\bar{N}_1 + 2)(\bar{N}_1 + 1)}\Big)\Big]\Big\} \Big/$$

$$\{B_3^{-1}B_1^{-1} + \tfrac{1}{2}\sigma_{23}^2(B_3^{-1}(-1)B_1^{-1} - 2B_3^{-1}B_1^{-1} + B_3^{-1}(1)B_1^{-1})$$

$$+ \tfrac{1}{2}\sigma_{13}^2(B_3^{-1}(-1)B_1^{-1}(-1) - 2B_3^{-1}B_1^{-1} + B_3^{-1}(1)B_1^{-1}(1))$$

$$+ \tfrac{1}{2}\sigma_{12}^2(B_3^{-1}B_1^{-1}(-1) - 2B_3^{-1}B_1^{-1} + B_3^{-1}B_1^{-1}(1))$$

$$+ \tfrac{1}{4}[\sigma_{23,13}^2(B_3^{-1}(-2)B_1^{-1}(-1) - 2B_3^{-1}B_1^{-1} + B_3^{-2}(2)B_1^{-1}(1))$$

$$+ \sigma_{23,12}^2(B_3^{-1}(-1)B_1^{-1}(-1) - 2B_3^{-1}B_1^{-1} + B_3^{-1}(1)B_1^{-1}(1))$$

$$+ \sigma_{13,12}^2(B_3^{-1}(-1)B_1^{-1}(-2) - 2B_3^{-1}B_1^{-1} + B_3^{-1}(1)B_1^{-1}(2))]\} \tag{45}$$

where

$$\bar{N}_3 = N_3 - \bar{\imath}_{23} - \bar{\imath}_{13}\,, \tag{46}$$

$$\bar{N}_1 = N_1 - \bar{\imath}_{12} - \bar{\imath}_{13}\,, \tag{47}$$

and for $j = 3$ or 1,

$$B_i^{-1} = 1/B(\bar{N}_i\,,\ \lambda_i), \tag{48}$$

$$B_i^{-1}(-1) = B_i^{-1} - 1, \tag{49}$$

$$B_i^{-1}(-2) = B_i^{-1} - 1 - \frac{\bar{N}_i}{\lambda_i}\,, \tag{50}$$

$$B_i^{-1}(1) = B_i^{-1} + \frac{\lambda_i}{\bar{N}_i + 1}\,, \tag{51}$$

$$B_i^{-1}(2) = B_i^{-1} + \frac{\lambda_i}{\bar{N}_i + 1} + \frac{\lambda_i^2}{(\bar{N}_i + 2)(\bar{N}_i + 1)}. \tag{52}$$

In many cases, $P\{A_1 \cup A_2 \cup A_3\}$ may be sufficiently well approximated by using $B(\bar{N}_3\,,\ \lambda_3)B(\bar{N}_1\,,\ \lambda_1)$ for $P\{A_1 \cap A_3 \mid A_2\}$ since the

last term in (32) is typically smaller than the preceding three which are, in turn, typically smaller than the first three.

IV. DISCUSSION

We tried the method out in a number of examples and found the dependence analyses to improve accuracy considerably when dependence is indeed important and to be useful in indicating when the dependence effect can be safely ignored. This investigation was actually motivated by a study of the use of a satellite to carry telephone traffic. (This study will be described more fully elsewhere.) One mode of such use, called "variable destination," allocates up-channels (each channel consists of a number of trunks) from ground stations to the satellite while all ground stations can receive all that is transmitted downward from the satellite. An equivalent network for five ground stations (the outer nodes) is shown in Fig. 3. Between each pair of ground stations there is an offered traffic; thus, there are 10 (2-way) traffics. For various choices of the traffics and allocations, the dependence effect becomes significant and it was important to know when it needed to be taken into account and when it could safely be ignored (which leads to simplification in the study, particularly the optimization algorithms—an aspect of the study was to optimize the channel allocations to minimize the traffic blocked). The method of this paper was found quite useful in this regard. It should be noted that, although the N_i are very large and results of the form of (1) lead to huge sums, the method was computationally simple.

The approach to dependence given here is one of approximations motivated by exact results in simple contexts. The results obtained thus far are promising and encourage further work in investigating the approximations and in extending the approach. Further experience is needed to determine the best way to apply the results to large net-

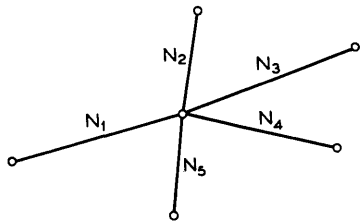


Fig. 3—Equivalent network for satellite variable destination mode.

works (see the last comment in Section III). Observe that improved accuracy (and, possibly, satisfactory accuracy from a practical viewpoint) in large networks can be achieved by considering only two-link or only two- and three-link dependences (already discussed in this paper). An inherent assumption that has been used is that the traffics are Poisson (or close to Poisson). As mentioned in Section I, the analysis of point-point blocking probabilities in conjunction with the use of the equivalent random method will be reported in Ref. 1.

V. ACKNOWLEDGMENT

I am grateful to D. L. Jagerman for stimulating discussions about approximations.

APPENDIX

In Section II, two approximations for $P\{A_1 \mid A_2\}$ were given. This Appendix derives the sum approximations for $P_b\{A_1 \mid A_2\}$. In particular, approximations for the following sums will be obtained:*

$$J_1 = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \frac{\lambda^{N-i}}{(N-i)!} = E\left[\frac{\lambda^{N-\xi}}{(N-\xi)!}\right], \tag{53}$$

$$J_2 = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \sum_{j=0}^{N-i} \frac{\lambda^j}{j!} = E\left[\sum_{j=0}^{N-\xi} \frac{\lambda^j}{j!}\right]. \tag{54}$$

The random variable ξ is distributed according to the binomial law. One has

$$\frac{\lambda^{N-j}}{(N-j)!} = \frac{1}{2\pi i} \oint z^{-N+j-1} e^{\lambda z}\, dz, \tag{55}$$

$$\sum_{l=0}^{N-j} \frac{\lambda^l}{l!} = \frac{1}{2\pi i} \oint z^{-N+j-1} \frac{e^{\lambda z}}{1-z}\, dz \tag{56}$$

in which the integration is on a circle around the origin as center.

Let $z = re^{i\theta}$, then

$$\frac{\lambda^{N-j}}{(N-j)!} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\lambda r e^{i\theta} - (N-j) i\theta} r^{-(N-j)}\, d\theta, \tag{57}$$

$$\sum_{l=0}^{N-j} \frac{\lambda^l}{l!} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\lambda r e^{i\theta} - (N-j) i\theta} \frac{r^{-(N-j)}}{1 - re^{i\theta}}\, d\theta. \tag{58}$$

* The derivations to follow are due to D. L. Jagerman who prepared this Appendix.

Computing the required expectations from (57) and (58), one obtains

$$J_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\lambda r e^{i\theta} - Ni\theta - N\ln r}\phi(\theta - i\ln r)\,d\theta, \tag{59}$$

$$J_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\lambda r e^{i\theta} - Ni\theta - N\ln r}\left(\frac{\phi(\theta - i\ln r)}{1 - re^{i\theta}}\right)d\theta \tag{60}$$

in which the characteristic function $\phi(\tau)$ is given by

$$\phi(\tau) = E[e^{i\tau\xi}] = (q + pe^{i\tau})^n. \tag{61}$$

Thus

$$J_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{h_1(\theta)}\,d\theta, \tag{62}$$

$$J_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{h_2(\theta)}\,d\theta, \tag{63}$$

in which

$$h_1(\theta) = \lambda r e^{i\theta} - Ni\theta - N\ln r + n\ln(q + pre^{i\theta}), \tag{64}$$

$$h_2(\theta) = \lambda r e^{i\theta} - Ni\theta - N\ln r + n\ln(q + pre^{i\theta}) - \ln(1 - re^{i\theta}). \tag{65}$$

The expansions of $h_1(\theta)$ and $h_2(\theta)$ in powers of θ are

$$\begin{aligned} h_1(\theta) &= \lambda r - N\ln r + n\ln(q + pr) \\ &\quad + \left(\lambda r - N + \frac{npr}{q + pr}\right)i\theta \\ &\quad - \frac{1}{2}\left\{\lambda r + n\frac{pr}{q + pr} - n\left(\frac{pr}{q + pr}\right)^2\right\}\theta^2 + \cdots, \end{aligned} \tag{66}$$

$$\begin{aligned} h_2(\theta) &= \lambda r - N\ln r + n\ln(q + pr) - \ln(1 - r) \\ &\quad + \left(\lambda r - N + \frac{npr}{q + pr} + \frac{r}{1 - r}\right)i\theta \\ &\quad - \frac{1}{2}\left\{\lambda r + n\frac{pr}{q + pr} - n\left(\frac{pr}{q + pr}\right)^2 + \frac{r}{1 - r} + \left(\frac{r}{1 - r}\right)^2\right\}\theta^2 + \cdots. \end{aligned} \tag{67}$$

Using the method of Hayman (see Ref. 7) the functions $h_1(\theta)$, $h_2(\theta)$ will be made stationary at $\theta = 0$ by defining r_1, r_2 as follows:

$$\lambda r_1 - N + \frac{npr_1}{q + pr_1} = 0, \tag{68}$$

$$\lambda r_2 - N + \frac{npr_2}{q + pr_2} + \frac{r_2}{1 - r_2} = 0. \tag{69}$$

Thus the values of r satisfy

$$\lambda p r_1^2 + (\lambda q - Np + np) r_1 - Nq = 0, \tag{70}$$

$$\lambda p r_2^3 + (np - p - Np + \lambda q - \lambda p) r_2^2 + (-\lambda q + Np - Nq - np - q) r_2 + Nq = 0. \tag{71}$$

For r_1, the following explicit formula is obtained:

$$r_1 = \frac{\sqrt{(\lambda q - Np + np)^2 + 4Npq} - \lambda q + Np - np}{2\lambda p}. \tag{72}$$

For r_2, the root between zero and one must be used. Let

$$\beta_1 = \lambda r_1 + \frac{npr_1}{q + pr_1} - n\left(\frac{pr_1}{q + pr_1}\right)^2, \tag{73}$$

$$\beta_2 = \lambda r_2 + n\frac{pr_2}{q + pr_2} - n\left(\frac{pr_2}{q + pr_2}\right)^2 + \frac{r_2}{1 - r_2} + \left(\frac{r_2}{1 - r_2}\right)^2, \tag{74}$$

then

$$h_1(\theta) = \lambda r_1 - N \ln r_1 + n \ln (q + pr_1) - \tfrac{1}{2}\beta_1\theta^2 + \cdots, \tag{75}$$

$$h_2(\theta) = \lambda r_2 - N \ln r_2 + n \ln (q + pr_2) - \ln (1 - r_2) - \tfrac{1}{2}\beta_2\theta^2 + \cdots. \tag{76}$$

Since

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-\frac{1}{2}\beta\theta^2}\, d\theta \cong \frac{1}{\sqrt{2\pi\beta}} \tag{77}$$

in which the approximation becomes the more accurate the larger β is, one now obtains

$$J_1 \cong r_1^{-N}(q + pr_1)^n \frac{e^{\lambda r_1}}{\sqrt{2\pi\beta_1}}, \tag{78}$$

$$J_2 \cong r_2^{-N}(q + pr_2)^n \frac{e^{\lambda r_2}}{(1 - r_2)\sqrt{2\pi\beta_2}}. \tag{79}$$

REFERENCES

1. Holtzman, J. M., "Point-Point Blocking Probabilities for Non-Poisson Processes," to appear.

2. Jensen, A., "An Elucidation of Erlang's Statistical Works Through the Theory of Stochastic Processes" in "The Life and Works of A. K. Erlang," Trans. Acad. Sci., No. 2, pp. 23–100 (1948).
3. Katz, S. S., "Statistical Performance Analysis of a Switched Communications Network," 5th Int. Teletraffic Congress, New York, Rockefeller Univ., pp. 566–575 (1967).
4. Lorentz, G. G., *Bernstein Polynomials*, Toronto: University of Toronto Press, 1953.
5. Froberg, C., *Introduction to Numerical Analysis*, Second Edition, Reading, Mass.: Addison-Wesley Publishing Co., 1969.
6. Fisz, M., *Probability Theory and Statistics*, Third Edition, New York: John Wiley and Sons, Inc., 1963.
7. Wyman, M., "The Asymptotic Behavior of the Laurent Coefficients," J. SIAM, 7, No. 3 (September 1959), pp. 325–342.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

A Laboratory System for Measuring Loudness Loss of Telephone Connections

By J. L. SULLIVAN

(Manuscript received August 5, 1970)

Transmission performance of the telephone message network has improved steadily over the years. Coincidental with this improvement has been the evolution of rating plans which provide the basis for transmission planning, design, and evaluation. Loudness of telephone speech was an important consideration in this evolution.

Telephone speech loudness continues to be one of the major factors which need to be taken into account in telephone transmission engineering. This paper covers a laboratory system, called EARS (*Electro-Acoustic Rating System*), *devised to make objective measurements of partial and overall telephone connections (including electro-acoustic transducer efficiencies) in a manner which reflects subjective loudness loss. Topics covered include a historical review of rating plans, computation of speech loudness, evolution of the* EARS, *and description of the system and its capabilities.*

The EARS *essentially comprises a sound source and a meter for measuring either acoustical pressure or electrical voltage. Design of the source and meter are based on telephone speech loudness considerations, and measurements made with the system approximate subjective loudness judgments with an accuracy which is sufficient for telephone engineering purposes.*

The EARS *may be used in implementing* any *telephone transmission rating plan incorporating speech loudness loss as an element. In such application, the system can be used for specifying connection losses, thus eliminating the need for extensive subjective tests of loudness. However, subjective tests will still be required to evaluate effects of other elements, e.g., noise, important in any given plan, and to evaluate the range of loudness losses acceptable for that plan.*

I. INTRODUCTION

A basic Bell System objective—to provide our customers with the best possible telephone message transmission consistent with the

state of the art and the economic climate—has remained essentially unchanged since the early days of telephony. However, continuing review of the telephone system in terms of this objective has resulted in steady improvement in transmission performance of the system over the years. Such improvement has been made possible by growth in our technical skills; it has been made necessary by evolving customer needs for improved transmission. Indeed, it has been postulated that as our customers use the telephone, they become accustomed to current performance and come to expect further improvement.[1,2]

What do we mean by telephone message transmission performance? In the broadest sense, this refers to the effect of the system on speech signals when these signals are transmitted over telephone connections. Customers conversing over telephone connections want to hear reasonably faithful, undistorted reproductions of each others' voices with a minimum of effort. Connections for which these conditions pertain can be thought of as providing satisfactory transmission performance. Connections exhibiting severe distortion would thus provide something less than satisfactory performance; customers might be able to converse but only with extreme difficulty.

Speech transmission capabilities of the telephone network are often considered in terms of individual transmission parameters, the combination of which determines overall transmission performance. Some of the more important parameters are loss, amplitude distortion, and unwanted interferences such as noise, crosstalk, and echo. Improvement in performance over the years has been achieved by design to control these parameters, singly and in combination, as dictated by the technology, economics, and needs of the times.

Coincidental with this improvement in transmission performance has been the evolution of telephone transmission rating. The problem of rating, that is evaluation and measurement, has been the subject of much thought and work over many years, and a number of different rating plans evolved to meet the transmission design needs for an improving and expanding telephone message network.

For present purposes, a transmission rating plan comprises (*i*) a rating criterion which represents the basis for rating telephone connections, (*ii*) a reference system (may be a physical simulation of an overall telephone connection, a set of definitions, or both) with some adjustable feature in terms of which connection ratings are established based on the selected criterion, and (*iii*) a rating scale which is essentially represented by the variable feature of the reference

system. Thus, a rating plan provides a framework for the design and evaluation of telephone connections.

Historically, the rating criterion was subjective in nature, and extensive subjective testing was required to evaluate telephone connections in terms of the reference. Moreover, application of the various rating plans required subjective testing to evaluate scale values, i.e., determine what scale values (and connections) represented acceptable performance.

Transmission performance of the telephone message network has improved to a point that loudness loss is a major variable which has to be taken into account in transmission planning. This paper describes a laboratory rating system, called the EARS (*E*lectro-*A*coustic *R*ating *S*ystem), which can be used to objectively measure loudness loss of telephone connections in a manner which closely approximates subjective loudness judgments. Thus, the system supplants subjective tests which would ordinarily be required to determine loudness ratings of connections. However, tests will still be needed both to determine subjective reaction to various amounts of loudness loss, and to determine interrelationships between loudness loss and other transmission parameters important in any given rating plan.

The EARS was devised to measure acoustic pressures and electric voltage as required by loudness rating definitions for partial and overall telephone connections.[3] The EARS has been used extensively over the past several years to characterize the loudness performance of telephone sets and connections, and to evaluate design plans such as unigauge design of the customer loop plant.[4] Moreover, loudness loss as measured with the EARS is an important element in current studies of transmission planning based on a multiparameter approach in which other transmission factors, e.g., noise and echo, are included with loudness loss.[5] Also, the EARS concept is one of several candidates for adoption as a standard method of specifying loudness loss.*

The EARS essentially comprises a sound source which is used to energize the talking end of a telephone connection and an indicating meter which is used to measure acoustic pressure or electrical voltage, and provides a simple means of measuring input and output signal levels for partial and overall telephone connections. Loudness losses

* Methods of measuring the loudness loss of telephone connections are currently being studied by (*i*) Study Group XII of the CCITT (Comite Consultatif International Telegraphique et Telephonique—International Telegraph and Telephone Consultative Committee) as outlined in Question 15/XII of Ref. 6 and (*ii*) the Task Force on Telephone Instrument Testing of the Institute of Electrical and Electronics Engineers.

of connections and connection components, including electro-acoustic transducers, i.e., telephone transmitters and receivers, are then the differences between input and output signal levels expressed in dB-like terms relative to appropriate reference signal levels.

The approach followed in this paper is first to consider the history of transmission rating plans and, second, to discuss the evolution of the EARS. Discussion of rating plans in Section II demonstrates the interrelationships of rating plans, rating systems, and network transmission performance improvements. Concluding remarks outline the relation between loudness rating definitions and the rating system, the EARS, which is used to determine parameter values as required by the rating definitions. Also mentioned are current studies concerned with multiparameter network design and evaluation.

Discussion of the evolution of the EARS involves several steps. We begin in Section III by addressing ourselves to a review of various techniques for computing the loudness of tones, noise, and speech, then discuss in some detail the derivation of a particular speech loudness computation method. This discussion indicates the manner in which frequency response characteristics of partial and overall connections can be measured and shows how, from the measured response for a given connection, we can compute a number which we will call loudness. Also included is a comparison of computed and experimental results.

Section IV covers derivation of the EARS from the speech loudness computation method referred to above and describes the EARS in its present form. Also described is a graphical computation method based on the design concepts leading to the EARS.

In Section V, we discuss the accuracy of the EARS in predicting subjective test results and consider the effects of simplifying assumptions employed in deriving the EARS. The vehicle for this is the graphical method referred to above. Results computed using this method are compared to observed results for the subjective tests used in validating the speech loudness computation method. This approach is necessary since the subjective test systems are no longer available, and hence could not be measured directly using the EARS. Therefore, the comparison of computed and observed results reflects the accuracy of the concepts on which the EARS is based, and not of the EARS itself.

As will become evident, the EARS in its present form is only one of several ways in which the computational method could be implemented as a laboratory measuring system, and is not an exact realiza-

tion of the computational method. This form was selected because it provides a satisfactory combination of simple equipment arrangements and suitable calibration and measurement procedures.

II. EVOLUTION OF TRANSMISSION RATING PLANS

Historically, transmission rating plans depended on (*i*) the use of reference circuits with which commercial circuits could be compared and (*ii*) systems of units for expressing the relative ratings thus determined. A great deal of subjective testing was necessary to determine ratings of commercial telephone equipments and facilities, and to correlate these ratings with measurable transmission properties. These ratings, in terms of the appropriate prevalent unit for expressing them, were used in design and evaluation of the telephone message network.

Considerable work was also necessary to determine transmission objectives for the network. That is, the reference system and related units provided the means of expressing transmission performance; it was then necessary to determine the rating, or range of ratings, to which the plant should be designed.

2.1 *Transmission Equivalent Plan*

For a time following the turn of the century, telephone circuits in commercial use had quite similar characteristics, and conditions were such that essentially only loudness, or volume, capability was under control of the transmission engineer. The performance of circuits was determined by comparing them, on a loudness basis, with a reference circuit which was adjustable in attenuation, but whose other transmission characteristics were typical of commercial circuits.

The reference circuit in use at that time was known as the Standard Cable Reference System (SCRS).[7-9] It consisted of telephone sets (including transmitters and receivers), cord circuits (for supplying the telephone set carbon transmitters with operating current), and an adjustable artificial line for interconnecting the cord circuits. These connection components were representative of types then used commercially.

The rating of telephone circuits by comparing them to the reference circuit involved a talker speaking alternately over the test circuit and the reference circuit, and a listener switching similarly at the receiving ends. The reference circuit artificial line, calibrated in terms of "miles of standard cable," was adjusted until the listener judged the volume

or loudness of the speech sounds reproduced by the two circuits to be equal. The number of miles of artificial line in the reference circuit was then used as the "transmission equivalent" of the circuit under test.[7] The effect of any change in a test circuit on the efficiency of that circuit could then be measured by determining the "transmission equivalent" before and after the change; the number of miles of artificial line required to compensate for the change was used as an index of this effect.

Performance ratings assigned in terms of loss of speech volume or loudness based on the comparison procedure outlined above constituted a practicable and effective means of assessing transmission performance. The adverse effects of sidetone, distortion, and noise on transmission quality were recognized, but no way was known of incorporating them, together with loudness loss, into a single figure of merit for rating transmission performance.[3] Moreover, the similarity between the reference circuit and commercial circuits at that time rendered such incorporation unnecessary.

As the state of the telephone art developed, modifications in the reference circuit became desirable in order for the circuit to more satisfactorily fulfill its purpose. Three factors leading to the modifications were (*i*) improved designs of telephone instruments and circuits, (*ii*) improved measuring techniques and instruments, and (*iii*) the need for a more suitable unit than the "miles of standard cable."

The effect of improved design was to introduce into the plant telephone instruments and circuits which had less distortion than corresponding parts of the Standard Cable Reference System. For this reason, it became desirable to have a new reference system with which transmission over the most perfect telephone circuit, or over circuits less perfect, could be simulated at will.[8,9]

Secondly, the performance of the Standard Cable Reference System was specified by stating the kinds of apparatus and circuits used. The electrical portion of the system could be checked by voltage, current, and impedance measurements. However, for the transmitters and receivers, reliance for constancy of performance was placed primarily upon the careful maintenance and frequent cross comparisons (subjective) of a group of transmitters and receivers which were specially constructed to reduce some of the sources of variation in the regular product instruments.[8,9] Improvements in measuring techniques and instruments made it possible to measure objectively characteristics of electro-acoustic transducers. This, in turn, permitted selection and maintenance of transducers to provide long-term stability.

Finally, a change in rating units became desirable for two reasons. First, circuits were being designed which had less distortion than the artificial cable of the Standard Cable Reference System. Since a mile of cable with this system corresponded not only to a certain volume change, but also to a distortion change, characterizing new circuits in terms of miles of standard cable implied a distortion degradation not necessarily attributable to the circuit under test. Secondly, two different reference systems were in common use, one in the United States, the other in some other countries.[10] These systems used artificial cable with different characteristics. That used in the United States had a loop resistance of 88 ohms and a capacitance of 0.054 microfarad per mile; the other used artificial cable having the same resistance and capacitance but had, in addition, an inductance of 1 millihenry and a conductance of 1 micromho per mile. Thus, a test circuit compared with each of these references would have two different ratings assigned to it. The first of the above reasons suggested the need for a distortionless unit, the second for a common unit.

A new unit, called the Transmission Unit (TU), was devised by the Bell System.[10*] This was a distortionless, logarithmic unit so chosen as to make use of common logarithms convenient in transmission computations. Its magnitude was very nearly the same as the loss of a mile of standard cable and, thus, existing experience learned in terms of miles of standard cable could be transferred to the new system with a minimum of difficulty. The Transmission Unit later became the decibel (dB).[11]

The three factors discussed in preceding paragraphs resulted in design of the Master Reference System (MRS).[8,9] This system utilized transducers with very low distortion, amplifiers to compensate for the lower efficiency of these transducers as compared to commercial instruments, and an artificial line (consisting essentially of a 600-ohm attenuator) for interconnecting the transmitting and receiving elements. The system included provision for inserting distorting networks into the transmitting and receiving elements so that performance of

* An intervening step between the "Mile of Standard Cable" and the "Transmission Unit" was the "800 Cycle Mile." This also was a distortionless logarithmic unit, equal in magnitude to the loss of a mile of standard cable at 800 Hz (actually, the loss at 796 Hz was used). There were thus two "800 Cycle Mile" units because of the two different standard cable specifications. The "800 Cycle Mile" and its successor, the "Transmission Unit," differed in that (*i*) the former represented a current ratio while the latter represented a power ratio and (*ii*) although both were logarithmic, the former was in units of log 1.115 (one 800 Cycle Mile corresponded to a current ratio of 1.115) while the latter was in units of 0.1 log 10 (one Transmission Unit corresponded to a power ratio of $10^{0.1}$).

commercial transducers, characterized by amplitude response curves with pronounced resonances, could be simulated.

Transmission engineering was still done on a loudness basis, and the MRS was used to obtain ratings in the same manner as that previously described for the SCRS. That is, (*i*) speakers talked alternately over the reference and test systems, (*ii*) observers listened over the systems, switching with the talkers, and (*iii*) the line (600-ohm attenuator) of the MRS was adjusted to obtain equal loudness. The average setting of the attenuator (in dB) at balance was then the rating of the system under test.

Adoption of the MRS by the Bell System had important long-term effects on international standardization of telephone reference systems. In 1926, the CCI* invited representatives of the Bell System to meet with a committee appointed by the CCI to consider adoption of a transmission reference system.[9] At the recommendation of this committee, the CCI adopted the MRS. Two of these systems were built by the Bell System. One was retained at Bell Telephone Laboratories in New York, the other was sent to the laboratory of the CCI in Paris in 1928. The CCI System was designated the SFERT.†

In the late 1950s, the SFERT was replaced by a new reference system, the NOSFER.‡ The NOSFER is so constructed and calibrated that it is essentially equivalent to the SFERT. Telephone connection ratings obtained with either the SFERT or the NOSFER are designated as Reference Equivalents (RE) in dB, and are numerically equal to the setting (in dB) of the reference system line attenuator required for loudness balance.

2.2 *Effective Loss Plan*

In the early 1930s, the Bell System adopted a new transmission plan. The change became necessary because of technological advances. Telephone sets incorporating antisidetone circuitry and improved transducers began to be used in quantity in the Bell System. These sets had poorer loudness performance than their predecessors, but

---

* CCI = Comité Consultatif International des Communications Telephoniques a Grande Distance (International Consultative Committee for Long Distance Telephone Communication). This organization later became the CCIF and is now the CCITT.

† SFERT = Systeme Foudamental European de Reference pour la Transmission Telephonique (European Fundamental Reference System for Telephone Transmission). See Ref. 12.

‡ NOSFER = Nouveau Systeme Foudamental pour la Determination des Equivalents de Reference (New Fundamental System for the Determination of Reference Equivalents). See Ref. 6.

provided marked improvement in such characteristics as sidetone, amplitude distortion, and nonlinear distortion, all of which have a marked effect on transmission. Evaluation of these telephone sets in terms of the old system tended to emphasize loudness differences at the expense of these other factors. Thus, there was need for a rating plan which properly recognized the improvements and, at the same time, retained applicability for the older telephone set types. The effective loss plan was devised to meet this need.[13,14]

Ratings under the new plan were in terms of dB of "effective" loss to distinguish them from "loudness" or "volume" losses of the old plan. Effective loss represented a figure of merit for evaluating the effectiveness of the transmission over telephone circuits and, as such, was a measure of the ability of telephone listeners to understand as well as to hear transmitted telephone speech.

Effective transmission data were determined in terms of the Working Reference System (WRS).[14] This system consisted of representative customer loops and telephone sets and a variable, distortionless trunk, i.e., an attenuator. Ratings for commercial connections were obtained in effect by comparing these, using live talkers and listeners, with the reference system, adjusting the distortionless trunk until the reference system and test connection were judged to provide equivalent transmission. (Basic data for the plan were obtained from two-way conversation tests, in contrast to the earlier plan for which tests were on a listening-only basis.) The criterion for such equality was repetition rate, that is, the rate of occurrence of repetitions requested by test subjects.[13,14] Thus, balance was achieved when the WRS (with a particular trunk setting) and the test connection provided the same repetition rate. The WRS trunk setting, in dB, was then a measure of the effectiveness of the circuit under test. The reference trunk setting for the Working Reference System was selected to provide numerical equivalence of ratings obtained with the new plan and its predecessor. With reference trunk setting, the Working Reference System had an effective loss of 18 dB which was also its transmission equivalent in terms of the earlier plan.

During the early 1930s when the effective loss plan was under development, in-plant telephone set transducers were predominantly of the resonant type, characterized by large amounts of amplitude and nonlinear distortion, telephone sets were largely of the sidetone variety, and a number of low cutoff cable loading systems were in use. However, improved transducers and antisidetone sets were being introduced into the plant, and additional improvements were under

development. (The transitional nature of the telephone set plant at that time and additional planned improvements were important considerations leading to adoption of the effective loss plan.) This resulted in introduction of the 300-type telephone set into the plant in 1937.[15–17] Subsequent study similarly resulted in introduction of the 500-type telephone set into plant in 1950.[18,19] The effective loss plan encouraged these improvements and, also, the adoption of higher cutoff cable loading systems.

By the mid 1950s, telephone sets provided sufficient antisidetone and freedom from amplitude and nonlinear distortion that little could be gained by further improvements. Thus, loudness appeared to be once again a major variable factor in plant design, likely to be the most important consideration in subsequent telephone set design.

The uniformly high level of transmission performance characterizing the plant suggested revision of telephone transmission planning to reflect more emphasis of loudness effects. Impetus for such revision was provided by (*i*) the availability of a speech loudness computation technique and a laboratory measuring system based on this technique and (*ii*) the need for planning in simpler terms than was possible with the effective loss plan which suffered from a number of practical disadvantages.[3]

2.3 *Loudness Rating Plan*[3]

For present purposes, the loudness rating plan comprises definitions for loudness ratings of partial and overall telephone connections.* Loudness is a subjective quantity and, therefore, loudness rating is defined as the ratio of the loudness of the speech into the listener's ear to the loudness of the speech out of the talker's mouth. As used here, however, loudness rating is considered to be determined by objective measurements.

The rating definitions involve both acoustic pressures and electric voltages. The EARS provides a means of measuring these quantities. The EARS consists of a source of acoustic energy, comprising a complex voice-frequency test tone which simulates certain properties of human speech and an artificial mouth, and an indicating meter for measuring voltages or sound pressures in a manner which simulates the loudness perception of an average listener.

Ratings obtained using the EARS do not precisely equal ratings

* Loudness ratings as specified by the definitions are expressed in decibels. Such usage does not strictly conform to the definition of the decibel, but this should have no effect on general use of the term "decibel."

obtained using human beings, in part because the artificial mouth and the 6-cm^3 coupler do not precisely duplicate the characteristics of their human counterparts. However, EARS ratings are sufficiently accurate to be highly useful in telephone transmission engineering although some problems associated with using the EARS to measure nonlinear transducers, e.g., carbon microphones, remain to be solved. Continuing research aimed at characterizing human mouths and ears may result in artificial counterparts the use of which in the EARS will render the differentiation between EARS and subjectively determined loudness ratings unnecessary.

The EARS is thus an instrument used to measure the loudness performance of telephone instruments and facilities in much the same manner that the 3A Noise Measuring Set is used to measure the subjective magnitude of telephone message circuit noise.[20] Loudness performance is expressed as constants and/or families of curves which are used in transmission planning. Examples of such use are the unigauge plan already cited[4] and current studies which are contemplating transmission design based on (*i*) noise and loss and on (*ii*) noise, loss, and echo.[5]

Earlier rating plans utilized physical reference systems which simulated overall telephone connections. Loudness rating as considered here is based on measurement of physical quantities, and does not require specification of a standard simulated connection.

III. COMPUTATION OF LOUDNESS OF SPEECH

A number of tests have been conducted to determine the loudness of telephone speech signals. Study of these test results, and of various methods devised for computing the loudness of speech and tones, resulted in a particular speech loudness computational procedure. This procedure, developed some time ago but not previously reported in the literature, is discussed in some detail in the present section.

We begin with a summary of the essential attributes of the procedure with appropriate reference to functions required in the computation of loudness ratings. Following this, several methods for computing the loudness of continuous spectrum sounds are referred to, and the background for the specific computation method of concern herein is reviewed. We then derive the functions required for the computational procedure. Finally, the procedure is used to compute the loudness performance of several laboratory test systems, and the computed values are compared to experimental results.

The computational method is not based on fundamental properties of speech sounds and of the hearing mechanism and, therefore, is probably not generally applicable in determining the loudness of other types of sounds. Rather, the method is based on a "black box" approach in which a defined "speech" signal is applied at the input, and the signal so processed that the output of the black box correlates with subjective experience of speech loudness. Internal operation of the black box is specified, but such is not intended to reflect exactly the specific manner in which the human being processes speech stimuli to obtain loudness although similarities are noted. This approach is used because, although there are a number of theories covering loudness perception, the specific operation of the hearing mechanism in determining loudness is not known.

3.1 *Description of Speech Loudness Computational Method*

The method, depicted on Fig. 1, is based on performing certain operations on the pressure spectrum delivered by a telephone connection to the ear of a listener. In essence, these operations comprise (*i*) dividing the received speech spectrum into a number of different frequency bands, (*ii*) determining the loudness due to each band, and (*iii*) summing across all bands to obtain the total loudness.

The received pressure spectrum consists of (*i*) a reference speech spectrum applied at the transmitting end of a connection modified by (*ii*) the amplitude transfer characteristic of the connection. The compromise spectrum and system response definitions are given on Figs. 2 and 3 respectively.*

Not all of the received spectrum contributes to loudness, i.e., that portion of the spectrum lower in level than the threshold of hearing does not contribute. Account is taken of this by defining a quantity termed effective spectrum which is the received spectrum minus X, the threshold of audibility for continuous spectrum sounds. The X function is shown on Fig. 4.

The effective spectrum is divided into 50 frequency bands selected such that each contributes equally to the total loudness produced by a flat effective spectrum. The frequency limits for the "2 percent" loudness bands can be derived from the function of Fig. 5.

The effective level in each band is then converted to loudness, in loudness units (LU), using the function of Fig. 6, and the loudness

* The ordinate of Fig. 2 is in terms of dBt. For purposes of this paper, dBt = dB relative to 2×10^{-5} newton/meter^2; 0 dBt = $20 \log_{10}(2 \times 10^{-5})$.

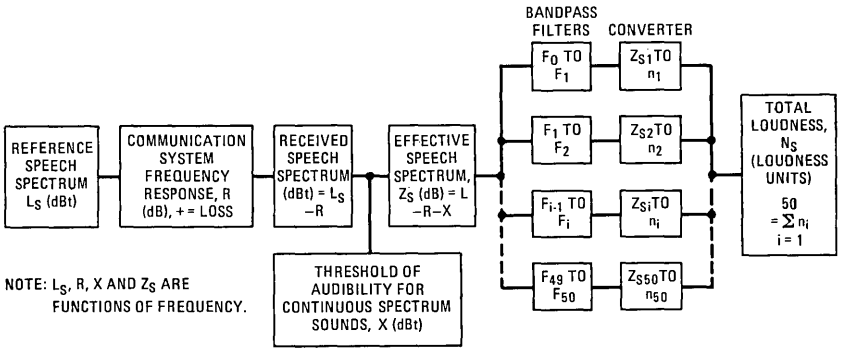


Fig. 1—Computation of speech loudness.

units summed for the 50 bands. The resulting sum, N_s, represents the total loudness contained in the particular received spectrum.

Once the total loudness, N_s, is obtained, the problem of its interpretation arises. One useful way of expressing the loudness is in terms of loudness level.* Functions which may be used to convert N_s to loudness level are shown on Fig. 7.

A more significant way of expressing loudness for speech communication problems is in terms of the level of a reference speech spectrum. The function relating total loudness to the level of the selected reference speech spectrum is given on Fig. 8.

3.2 *Survey of Loudness Studies*

Loudness has been studied extensively as demonstrated by the numerous publications relating to this subject. Studies reported have been largely concerned with single-frequency tones, tone complexes, and continuous spectrum steady sounds. Several different methods for computing the loudness of these different sounds have evolved.† Some of the methods, specifically those covering continuous spectrum

* The loudness level of a sound, in phons, is numerically equal to the median sound pressure level, in dBt, of a free progressive wave of frequency 1000 Hz presented to listeners facing the source, which in a number of trials is judged by the listeners to be equally loud. The unit of loudness is the sone. By definition, a 1000-Hz tone 40 dB above a listener's threshold produces a loudness of 1 sone; the loudness of any sound that is judged by the listener to be n times that of the 1-sone tone is n sones.[21] In this paper, we will use the term loudness units (LU) when considering loudness of speech, and the term sones in discussing other sounds. However, speech of N LU is equal in loudness to any other sound of N sones.

† References 22 through 27 describe some of these computational methods.

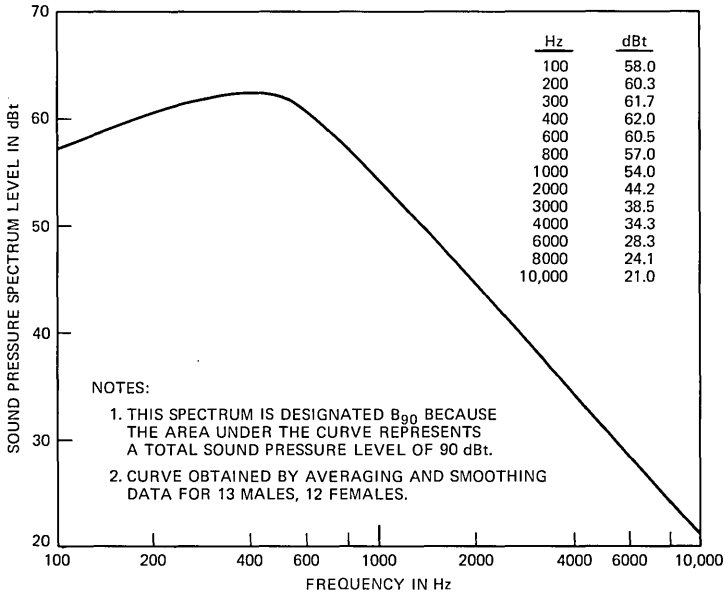


Fig. 2—Sound pressure spectrum of continuous speech at 2 inches from the lips of a talker.

steady sounds, are sufficiently general that with appropriate specification of the speech signal they may be suitable for computing the loudness of speech.* Such suitability has not, to the author's knowledge, been demonstrated and these methods are not considered in this report.

The study of speech loudness has received somewhat less attention than the general subject of loudness. Publications covering the study of speech loudness are limited in number, and of these, only a few consider speech loudness computational methods.†

In 1924, H. Fletcher and J. C. Steinberg devised a method for computing the loudness loss, due to changing the transmission system response, of a sound being transmitted to the ear.[28] Results obtained using the method were in close agreement with experimental results for the specific sounds—speech and a complex test tone—considered in the study.

* See, for example, Ref. 26.

† References 28, 30, 31, 32, and 33 report results of some speech loudness tests. References 28, 29, and 31 describe speech loudness computation methods developed.

In 1925, Steinberg developed a more general method of computing the loudness of any complex sound.[29] (The method referred to in the preceding paragraph was a special case.) Results computed using this method agreed with experimental results then available, including the data reported in the 1924 paper.

In 1938, W. A. Munson developed a method for computing the

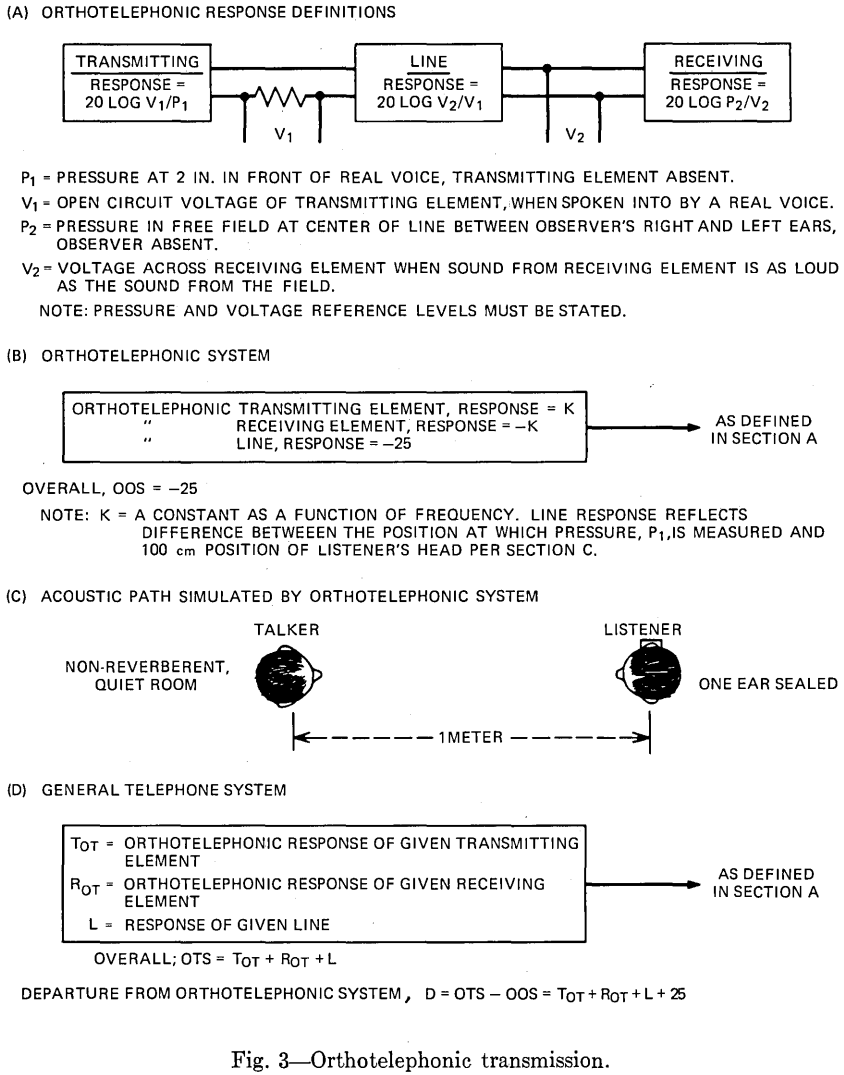


Fig. 3—Orthotelephonic transmission.

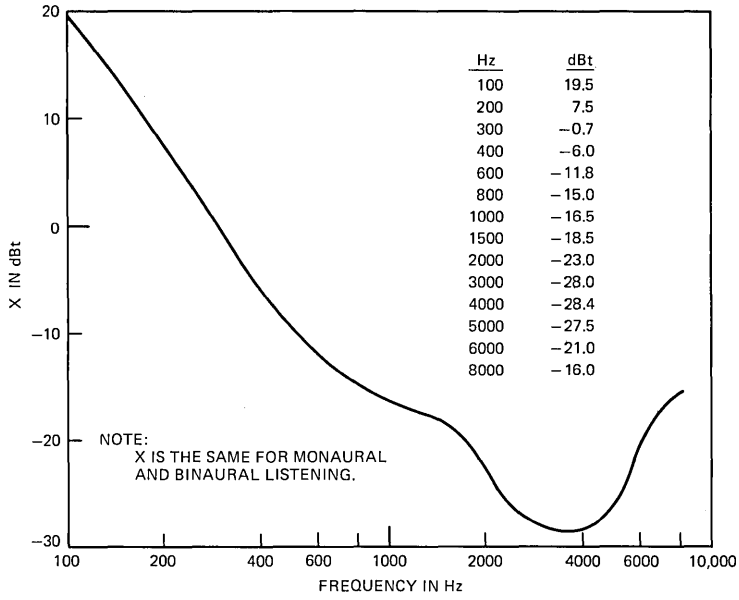


Fig. 4—X, the threshold of audibility for continuous spectrum sounds.

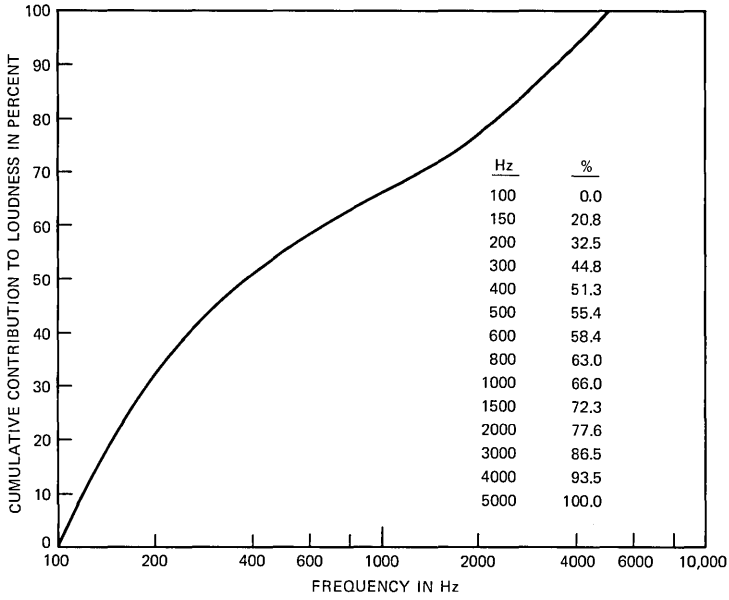


Fig. 5—Cumulative contribution to loudness of speech for a flat effective spectrum.

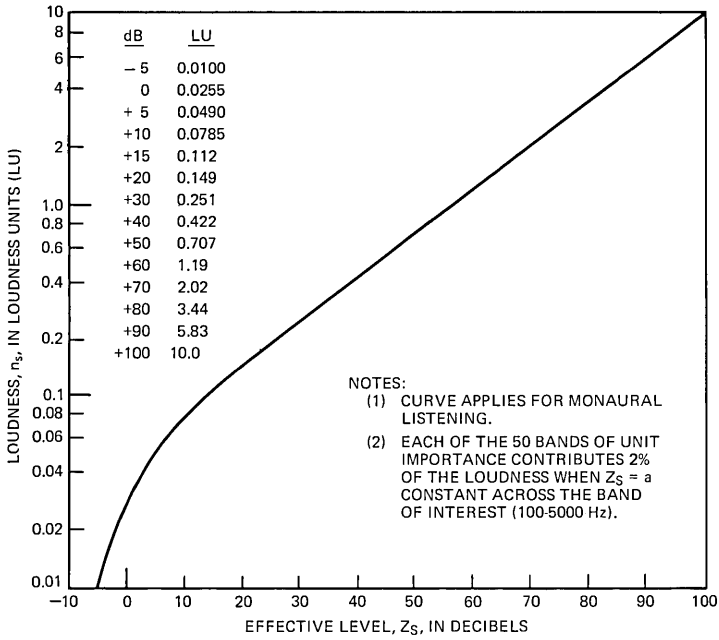


Fig. 6—Loudness (n_s) in a band of unit importance vs effective level (Z_s) in that band.

loudness of speech. The method was based on speech loudness tests conducted in 1935 and 1936.*

Thus, there were available at least two methods, those of Steinberg and Munson, for computing the effects of changes in the response characteristics of telephone circuits on speech loudness. These methods had been found to provide computed results which were reasonably consistent with observed results from experiments on which the methods were based. A review of these experiments revealed substantial inconsistencies between the different sets of data, leading to doubt concerning the generality of the procedures. Moreover, a substantial amount of additional speech loudness data became available in 1939 and 1940. These considerations led to a review of the available speech loudness computation procedures and the various sets of subjective test data for purposes of arriving at a simple procedure, suitable for telephone engineering purposes, which would be reasonably consistent with all of the data.

* Neither the computational procedure nor the tests were reported in the literature.

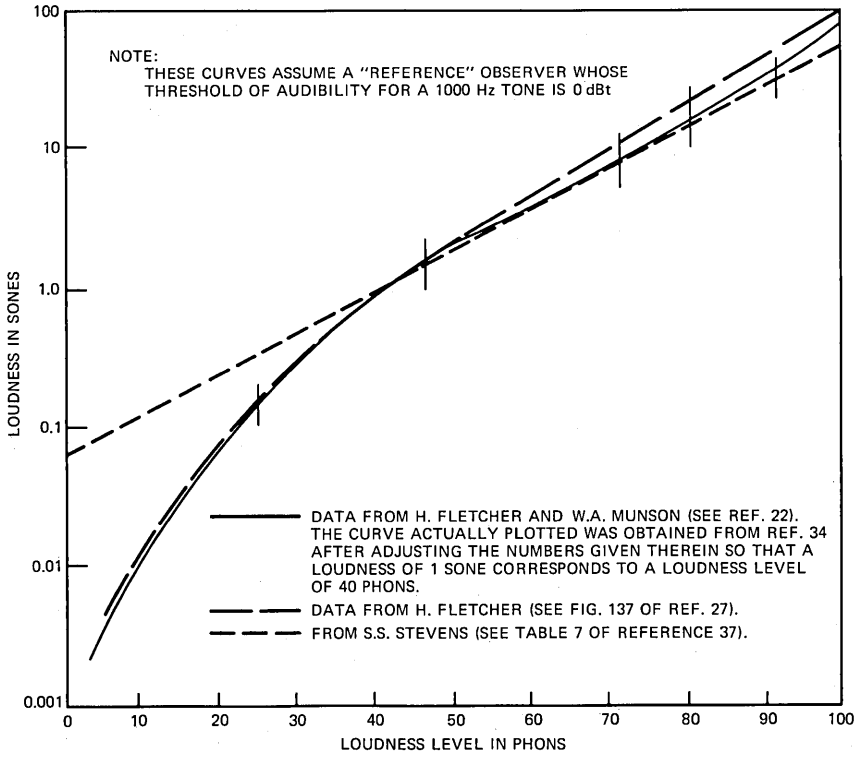


Fig. 7—Relation between loudness and loudness level.

The speech loudness computation procedure resulting from the review referred to above is described in subsequent paragraphs. This procedure, developed in 1941, and the 1939–40 speech loudness data have not been previously reported, primarily for two reasons. First, the group concerned with the speech loudness project was essentially disbanded in late 1941 and assigned other work in support of the war effort. This group was not re-established so that its accomplishments could be reduced to a form suitable for publication in the literature. Secondly, there was no pressing need for the procedure until the loudness rating plan was proposed.[3] Review initiated at that time has verified applicability of the procedure to telephone speech loudness problems.

3.3 *Development of the Speech Loudness Computation Method*

The speech loudness computation method evolved from a specific method devised by Fletcher and Munson for treatment of continuous

spectrum sounds, and uses as many of the steady-state definitions as are applicable.[23] Review of the Fletcher-Munson method therefore serves as a convenient introduction.

3.3.1 *Steady-State Sounds—Definitions and Formulae*

Loudness, N, is defined as the intensive attribute of an auditory sensation.[21] Loudness can be expressed in terms of a scale whose units have been found to agree with the common experience of observers estimating the intensive attribute of sounds. Ordinarily, the loudness of a sound is not specified in terms of units of a loudness scale, but, rather, in terms of its loudness level. (See Section 3.1 footnote.)

Results of some experiments to determine the relation between loudness and loudness level are shown on Fig. 7. The solid curve, due to Fletcher and Munson, was first proposed in 1933.[22]

This curve was later adopted in an American Standard.[35] Further study by Fletcher and Munson resulted in a slightly different relation shown by the long dashed line. This curve, first reported in 1937,[23] was later modified slightly.* The short-dashed curve, due to S. S. Stevens, resulted from effort directed at obtaining the best fit for all available data with the simplest possible relation.[26,36,37] This relation is included in a USA Standard.[38]

The loudness, N, for steady-state sounds can be computed from the masking spectrum, M, of the sound.† The loudness equation may be written as

$$N = \int F(M)\, dx. \tag{1}$$

$F(M)$, expressed in loudness units, is a function of the masking, and may be interpreted as expressing the intensity of the nerve stimulation at a particular position, x, along the basilar membrane. The quantity, x, is expressed in terms of the percent of total nerve endings that the maximum stimulation passes over as a stimulating tone is changed from lowest audible frequency to a frequency f, corresponding to x. The product $F(M)$ and dx represents the loudness contributed by a sound within the differential length dx and, consequently, the integration represents the total nerve stimulation, or loudness, over the length of the basilar membrane.

For continuous spectrum sounds, the masking, M,‡ at any fre-

* Figure 137, Ref. 27.

† See Ref. 23 and 39 for more detailed discussion of the concepts considered in this and subsequent paragraphs.

‡ $M = \beta - \beta_0$ where β = the level of a single frequency tone which is just audible in the presence of a specified noise and β_0 = the level of the same tone which is just audible when the noise is absent. (See Ref. 39.)

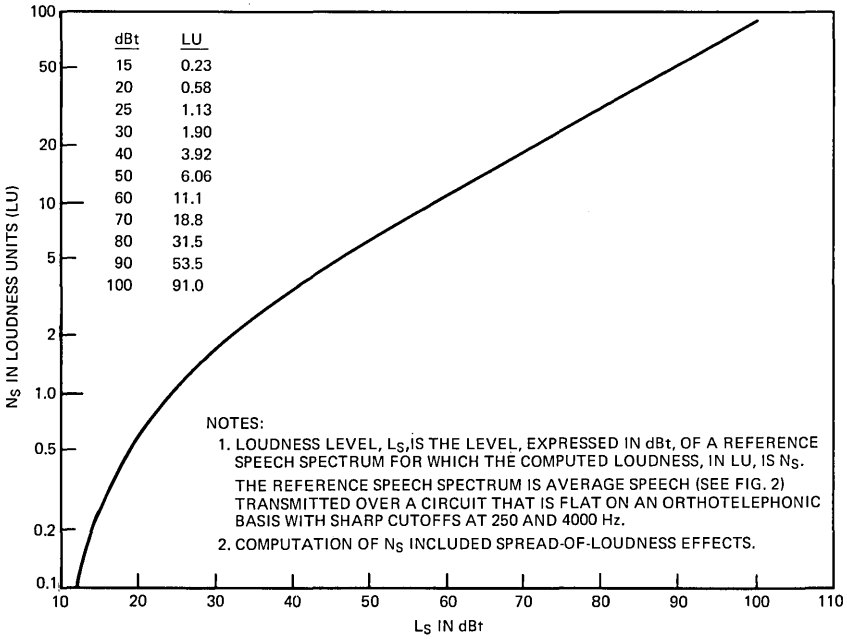


Fig. 8—Loudness of reference speech.

quency is related to the effective level, Z, of the masking sound at the same frequency. Thus, the loudness formula (1) can also be expressed as follows:

$$N = \int Q(Z)\, dx. \tag{2}$$

Interpretation of this formula is the same as that for (1) in that it represents the integrated nerve stimulation over the length of the basilar membrane. However, the intensity of the stimulation at any frequency is determined from the effective level of the continuous spectrum sound, Z, defined as follows:

$$Z = B - X \tag{3}$$

where

Z = effective spectrum level (dB)
B = pressure spectrum level (dBt)
X = a function (dBt), empirically determined from listening tests,

which reflects the fact that not all of the objectively measurable spectrum (B) is effective in producing loudness.

The X function plotted on Fig. 4 may be considered to be the threshold of audibility for continuous spectrum sounds. Although this is not strictly correct, it lends physical interpretation to the function and, as such, facilitates understanding of the loudness computation technique.*

The loudness, N, of a continuous spectrum sound can be computed by (*i*) dividing the audible frequency range into a number of successive frequency bands conveniently selected for computational purposes, (*ii*) determining from the effective level and the importance of each band the number of loudness units contributed by the band, and (*iii*) summing the loudness units over the audible range. This procedure, replacing the integration of (2), may be written as

$$N = \sum^{x} n_x \, \Delta x \tag{4}$$

where

N = total number of loudness units,
n_x = loudness, $Q(Z)$ from equation (2), integrated over a unit length of the basilar membrane, and
Δx = the number of unit lengths along the basilar membrane included in the computation band.†

The quantity, n_x, is a function of the effective level, Z, of that portion of the sound lying in the computation band.

3.3.2 *Speech Sounds—Definitions and Formulae*‡

Development of the speech loudness formula was based on the assumption that a formula of the same general type as (4) could be used, that is, the loudness of speech can be computed by dividing the speech spectrum into a number of successive bands selected for

* Consider the case in which no masking occurs, i.e., $M = 0$. Assuming $M = Z$, then $Z = B - X = 0$, and $X = B$, the pressure spectrum level of a noise which produces zero masking, hence zero loudness. X can thus be defined as the threshold of audibility for continuous spectrum sounds, useful for the purpose indicated, but not strictly correct since it has been found that at low values of M (e.g., $M = 0$), the relationship $M = Z$ is not valid. (See equation (10-5) of Ref. 27.)

† In Ref. 23, the unit length was taken as 1 percent of the basilar membrane length.

‡ Many of the steady-state definitions of Section 3.3.1 apply also for speech. However, since some of the functions differ, the subscript s is used where appropriate to denote those specifically applicable to the speech case.

convenience of computation, determining the loudness contributed by each band, and summing the loudness over the audible range. This can be expressed as follows:

$$N_s = \sum^{s} n_s \, \Delta S \tag{5}$$

where

n_s = total loudness produced by speech in a frequency band of unit importance

ΔS = the number of bands of unit importance included in a computation band.

The quantity, n_s, is a function of Z_s, the effective level of the received speech. Following the definition used in the formulation for continuous spectrum steady sounds, Z_s is defined as follows:

$$Z_s = B_s - X \tag{6}$$

where

B_s = sound pressure spectrum level of received speech (dBt)*

and X is defined in equation (3).

The bands of unit importance for equation (5) are so chosen that no matter where they are located on the basilar membrane, each contributes an equal amount to the total loudness when the effective level of the speech is independent of frequency, i.e., Z_s = k (a constant) across the audible range.† When the speech band is divided into a number of bands suitable for computational purposes, each of these may contain several bands of unit importance. Thus, the loudness, n_s, contributed by a band of unit importance must be multiplied by the number of such bands, ΔS, in the computation band. The value of ΔS used for any particular computation band depends on the width of that band (in Hz) and its location on the frequency scale.

The summation of ΔS along the frequency scale represents the cumulative number of bands of unit importance. The rate of change

---

*Speech varies rapidly in level with time. Accounting for this variability would have entailed development of complex formulae from the limited amount of applicable data available on speech. The approach followed was to develop the simplest possible method of computing speech loudness from the long-term average sound pressure spectrum of speech where long term is understood to comprise a time interval which is long compared to the average syllabic interval.

† Selection of the bands is based on a flat effective spectrum for purposes of simplicity. Any other shape, i.e., $Z_s \neq$ k, would require a different loudness versus Z_s function for each band of unit importance.

of S with frequency may, therefore, be interpreted as showing the relative importance to loudness of equal effective levels of speech in different frequency regions. This interpretation of ΔS and the interpretation of n_s as the loudness of a band of unit importance, based on the similarity of the quantities n_x and Δx in equation (4) and n_s and ΔS in equation (5), are probably not exact because of the radically different nature of speech and steady-state sounds. These interpretations should, therefore, be applied with caution, and regarded primarily as a point of view helpful in systematizing the basic data and in understanding the loudness formulations.

3.3.3 *Received Speech Spectrum*

For any given telephone connection, the sound pressure level of the received speech [B_s of equation (6)] and, hence, the loudness of the speech depends on (*i*) the speech spectrum applied at the talking end of the connection and (*ii*) the loss of the connection as a function of frequency. These factors are taken into account in the speech loudness computation method by assuming a compromise speech spectrum and expressing the connection loss in orthotelephonic terms.

3.3.3.1 *Compromise Speech Spectrum.* The compromise spectrum adopted is shown on Fig. 2 and is designated B_{90} because the area under the curve integrates to a total sound pressure level of 90 dBt.* This spectrum was obtained by averaging and smoothing long average power measurements of continuous speech (including pauses between words) from 13 males and 12 females.† Although some differences in spectral content were found between individual voices, the shapes of the spectrum curves were sufficiently alike to justify averaging across all voices. Moreover, measured spectra from other studies exhibit similar shapes.[41–43]

Several experimenters have reported measurements of the long-term average sound pressure level of speech. Results of these experiments, referred to a point 2 inches in front of a talker's lips, are given in Table I. (Selection of the 2-inch reference point is discussed in the next section.)

The values are in fair agreement excepting the last row. Benson and Hirsh considered the apparent discrepancies at some length, concluding that values reported by Dunn and White and by Rudmose,

* This spectrum is the same as that used in computing articulation index. See Fig. 2 of Ref. 39.

† Spectrum curves reported in Ref. 40 provided some of the data used in deriving this compromise spectrum.

TABLE I—LONG-TERM AVERAGE SOUND PRESSURE LEVEL OF SPEECH AT 2 INCHES FROM TALKERS' LIPS

Source	Males		Females	
	No.	Pressure (dBt)	No.	Pressure (dBt)
Sivian—Ref. 45	5	91.8	3	90.0
Dunn and Farnsworth—Ref. 44	1	90.4	—	—
Dunn and White—Ref. 40	6	90.8*	5	88.0*
Rudmose, et al.—Ref. 42	7	92.6*	—	—
Benson and Hirsh—Ref. 43	5	81.7*	5	79.7*

* Values reported were converted to apply at 2 inches by assuming an equivalent point source 0.6 cm behind the talker's lips. See Ref. 44.

et al., "approximate more nearly monitoring levels (peaks of energy probably corresponding to vowel sounds) than to an actual average."[43]

There is thus some uncertainty regarding the sound pressure level at 2 inches from a real voice. For present purposes, the level is assumed to be 90 dBt, an approximate average for the majority of the data (specifically excludes the Benson-Hirsh data). As will become evident, the exact value assumed for the level at 2 inches does not significantly affect the speech loudness computational procedure as long as the spectrum shape (see Fig. 2) is not changed.

3.3.3.2 *Orthotelephonic Transmission.* The concept of orthotelephonic transmission is based on relating telephone and face-to-face conversation.[16] Specifically, orthotelephonic transmission for present purposes implies reproduction by a telephone system of speech sounds which are indistinguishable from those received with an air transmission system. Conditions applicable for the latter case are negligible noise and acoustic reflections when listening monaurally at a distance 1 meter from the lips of the talker, the listener facing the talker. An orthotelephonic system is thus one that provides the required orthotelephonic reproduction of the speech sounds.

Application of the orthotelephonic concept to telephone connections is detailed in Fig. 3. The definitions are given here as a matter of convenience, and will be considered where appropriate in later discussion. Typical orthotelephonic amplitude responses of overall telephone connections employing specific telephone set types are given in Ref. 16. Such responses vary widely depending on the telephone set and transducer types used.

One point worth noting is that the orthotelephonic transmitting

response definition of Fig. 3 is in terms of the sound pressure level at 2 inches from a talker's lips whereby strict conformance with the orthotelephonic definition would require expression in terms of the 1-meter sound pressure level. This change was made for two reasons: (*i*) the 1-meter reference involves a long acoustic path which is ascribed to the transmitter; (*ii*) measurements at 2 inches are more reproducible since they are less likely to be affected by characteristics of the surrounding environment, e.g., reverberation. However, studies show that under free-field conditions the sound pressure spectra at 2 inches (approximately 5 cm) and at greater distances have about the same shape,[44,45] but differ in absolute level by about 25 dB for the 1-meter and 2-inch positions.[44]

3.3.4 *Derivation of n_s and S Functions*

The speech loudness formula [equation (5)] involves the two quantities n_s and S. As noted earlier, n_s is a function of Z_s and is the loudness due to that portion of speech contained in a frequency band of unit importance; S is a function of frequency which shows the relative importance to total loudness of equal effective levels of speech in different frequency regions. These functions, though somewhat interrelated, may be thought of as providing level weighting and frequency weighting respectively.

Data from three tests were used in deriving these functions. The n_s function was based on tests by Munson the results of which show how loudness varies as a function of received speech level. Data obtained by Steinberg and by Van Wynen were used to derive the S function. These data show the manner in which loudness depends on frequency content of the received speech signal. Other data were available, but were not used because (*i*) there were gross inconsistencies between the data and data from other tests and/or (*ii*) the test system and conditions were not reported in sufficient detail to permit utilization of the test results.

An important factor in deriving the n_s and S functions is Z_s, the effective level of the received speech. [See equation (6).] Effective levels for the test results used in the derivation are shown on Fig. 9. The reference effective spectrum curve, consisting of the compromise speech spectrum, B_{90}, of Fig. 2 minus the threshold function, X, of Fig. 4, is also plotted on Fig. 9. This function, utilized in the speech loudness computation technique later, is assumed to represent the average case for a large number of talker-listener pairs conversing over a telephone system with (*i*) orthotelephonic response and (*ii*) 25

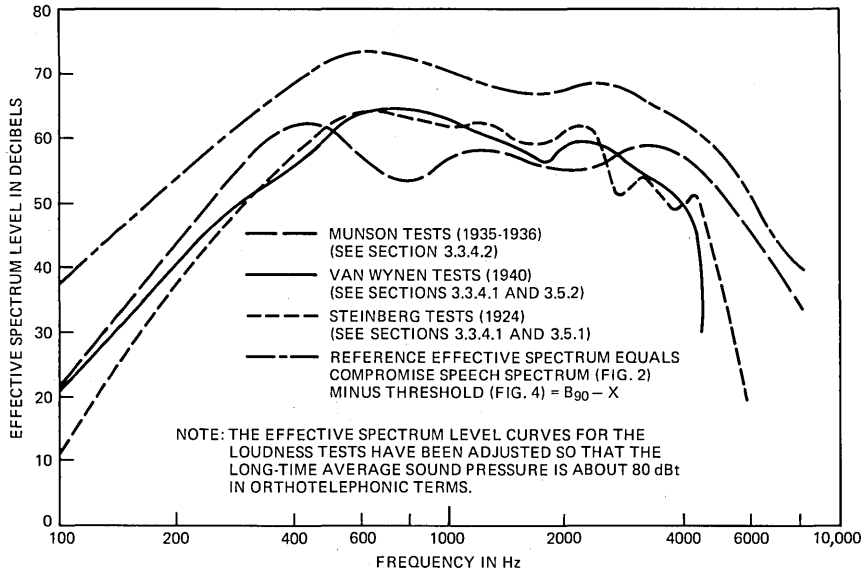


Fig. 9—Effective spectrum level curves.

dB of flat gain relative to a true orthotelephonic system. The function is given here to illustrate the difference between the effective spectrum level curve for a system providing orthotelephonic response and the effective spectrum level curves for the Steinberg and Van Wynen test systems.

Because n_s and S functions were interrelated, their derivations were carried out concurrently by means of a series of successive approximations. For simplicity, the two functions are treated separately in succeeding sections, and the derivation is described in greatly simplified terms. The actual derivation involved obtaining a first estimate of the S function, using the Steinberg and the Van Wynen data (including the respective loudness law exponents—see later discussion), for correcting the actual Z_s to a constant Z_s. This S function was then applied to the Munson data in order to obtain a curve of total loudness versus speech sound pressure level for a 100–5000 Hz band, the band of interest in the speech loudness computation method. The ordinate and abscissa of this curve were then corrected to reflect respectively the loudness in a band of unit importance and the Z_s in a band of unit importance.

These first estimates of the S and n_s functions were used to compute the loudness for a number of conditions. Computed results were compared to experimental results, and the functions modified as indicated by this comparison. This process was continued until computed values reflected adequate accuracy for engineering purposes.

3.3.4.1 *The S Function.* The S function was derived from results of tests by Steinberg (see Section 3.5.1) and by Van Wynen* (see Section 3.5.2). In these tests, speech received through high- and low-pass filters was balanced against undistorted speech. Test results important in the derivation are given in Fig. 10. The intersection of smooth curves drawn through the test results provides two items of information: (*i*) f_{50}, the frequency above and below which the loudness contributions are equal;† (*ii*) LL_{50}, the number of dB by which undistorted speech must be reduced for a 50-percent reduction in loudness.

By assuming a logarithmic relationship between loudness and effective level (see Section 3.3.4.2), we can determine the percentage of loudness contribution as a function of frequency. Thus,

$$Z_{s2} - Z_{s1} = \mathrm{k} \log \frac{N_2}{N_1} \tag{7}$$

where Z_{s1} and Z_{s2} are effective levels before and after a flat change in undistorted speech level, N_1 and N_2 are corresponding loudness numerics, and k is a constant. For the Steinberg tests, we find, substituting -9.1 dB for $Z_{s2} - Z_{s1}$ and 1/2 for N_2/N_1 (50-percent loudness reduction), that

$$\mathrm{k} = \frac{Z_{s2} - Z_{s1}}{\log \frac{N_2}{N_1}} = \frac{-9.1}{\log \frac{1}{2}} = 30.2\ dB$$

while for the Van Wynen tests, k = 41.2 dB.

We can express the relationship of equation (7) in the power function form frequently encountered in current literature[33,36] by letting P represent the stimulus magnitude and setting it equal to Antilog $Z/20$. Then

$$N = P^n \tag{8}$$

* Steinberg reports results for sensation levels = 100 dB and 39 dB.[29] (Sensation level is the level above threshold. See Ref. 21.) Only the 100-dB results were used in the derivation. Results for the 39-dB level were near the knee of the loudness curve (see Fig. 8 for example) and their inclusion would have substantially complicated the derivation.

† This assumes ideal filters. The filters used in the tests appear, for speech signals, to be sufficiently close to ideal to warrant making this assumption.

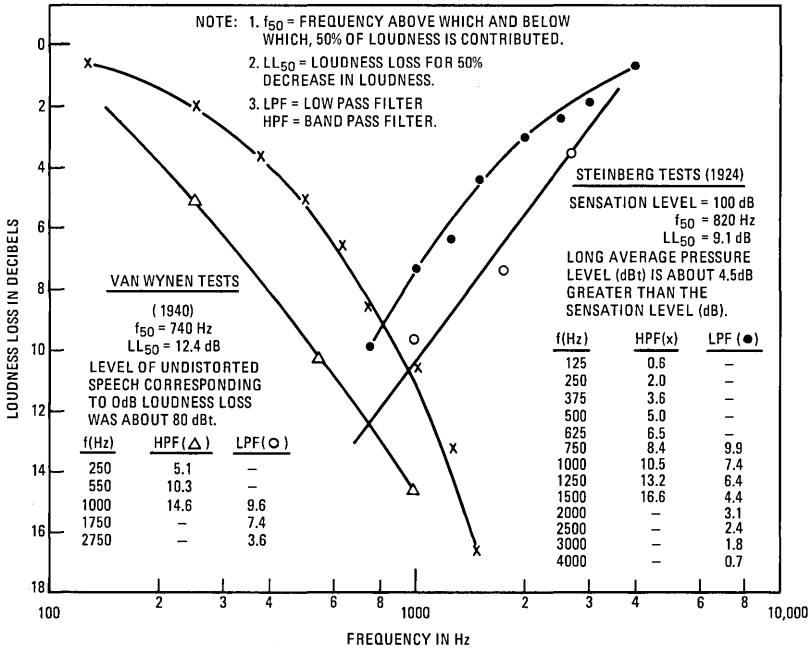


Fig. 10—Filter test results.

where N and P are as defined above and $n = 20/k$. The values of n are then 0.662 and 0.486 for the Steinberg and Van Wynen tests respectively.

By applying equation (7) to the curves of Fig. 10, we can determine the loudness reduction corresponding to various values of loudness loss. For example, the Van Wynen tests show that for a high-pass filter with cutoff frequency = 550 Hz, $Z_{s2} - Z_{s1} = -10.3$ dB. Using the equation, we find that the 550-Hz high-pass filter passes 56.2 percent of the loudness, suppressing 43.8 percent. Proceeding in a similar manner for all of the data, the points shown on Fig. 11 were obtained. It is of interest to note that there appear to be no systematic differences between the Steinberg and Van Wynen test results even though the exponents referred to in the preceding paragraph are somewhat different.

Because the effective spectrum level curves for tests were almost identical (see Fig. 9), the smooth curve of Fig. 11 was drawn to represent all of the data points. This curve shows the percent of total

loudness contributed by frequencies below the frequency of the abscissa for the effective spectrum level used in these tests.

The S function required, showing the cumulative contribution to loudness when the effective spectrum level is flat with frequency, is plotted on Fig. 5. The function was obtained from the curve of Fig. 11 by successive approximations as outlined in the introductory paragraphs of Section 3.3.4.

3.3.4.2 *The n_s Function.* The n_s function, relating loudness and effective speech level in a band of unit importance, was derived using test results obtained by Munson in 1935–1936. The test system and general methodology are described in Ref. 46. However, the test results have not previously been reported in the literature.

In the Munson tests, undistorted speech was loudness balanced against a 1000-Hz test tone. Also, thresholds were determined for the 1000-Hz and speech signals. Detailed results of the tests are given in Tables II and III. Average results are plotted on Fig. 12.

The ordinate of Fig. 12 was converted to a loudness scale using the solid curve of Fig. 7. The resulting relation between loudness and

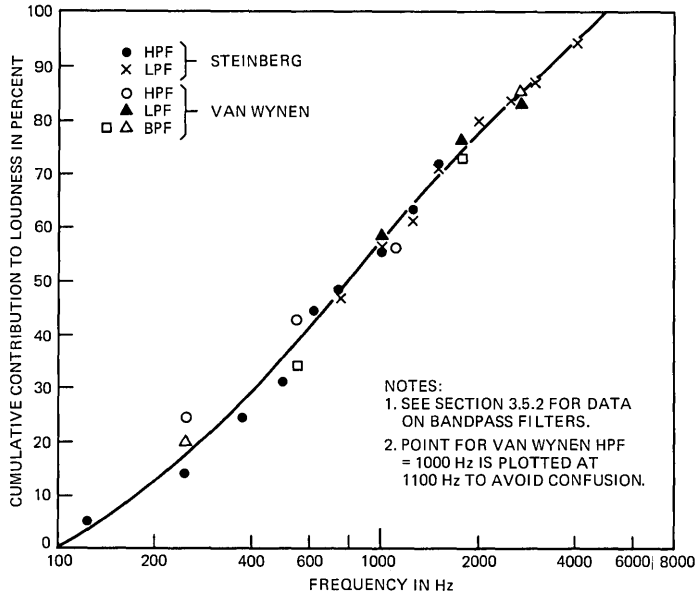


Fig. 11—Cumulative contribution to loudness of speech for the Steinberg and Van Wynen filter tests.

TABLE II—RESULTS OF MUNSON (1935–1936) SPEECH LOUDNESS TESTS

Speech Pressure Level-dBt	Parameter*	Subject Designation											Group†	
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	$\bar{\bar{X}}$	S
20.2	$\bar{X}$	29.5	39.1	20.6	29.5	29.1	19.2	24.2	20	20.9	29.6	18.7	25.5	6.4
	σ	4.2	9.2	4.8	12.8	3.4	2.8	8.5	3.3	3.5	5.6	5.1		
	N	6	4	6	6	3	5	4	6	6	4	5		
30.2	$\bar{X}$	54.2	56.3	50.4	52.7	57.1	49.0	41.4	35.9	39.5	43.0	36.9	46.9	7.8
	σ	5.8	5.2	11.4	9.1	9.2	11.8	11.1	3.1	6.7	8.0	8.0		
	N	8	6	8	8	5	7	4	7	8	6	7		
45.2	$\bar{X}$	78.6	76.3	76.5	76.1	76.3	72.3	66.6	61.2	62.9	56.0	60.7	69.4	7.9
	σ	5.8	3.8	5.1	6.8	5.0	6.2	8.7	3.9	3.2	9.0	7.0		
	N	13	12	13	13	7	11	5	12	13	7	11		
54.3	$\bar{X}$	91.5	84.0	86.6	76.5	82.5	86.5		76.5	71.5	55.5	74.5	78.6	10.3
	σ													
	N	1	1	1	1	1	1		1	1	1	1		
60.2	$\bar{X}$	91.5	88.5	90.9	84.6	83.7	85.1	78.8	77.0	75.6	66.2	71.9	81.3	8.1
	σ	3.6	3.8	5.7	4.9	4.0	3.8	5.5	3.7	3.5	9.5	5.6		
	N	10	9	10	9	6	8	5	9	10	8	8		
80	$\bar{X}$	98.2	98.5	103.6	95.9	93.2	100.2	93.3	94.0	91.5	76.9	82.7	93.5	7.7
	σ	4.7	3.9	5.4	3.0	3.2	3.5	1.8	3.8	7.0	6.8	5.9		
	N	8	7	8	7	5	6	5	8	8	6	6		
90.2	$\bar{X}$	107.2	105.2	108.5	99.6	101.5	107.7	100.3	105.0	97.2	78.3	91.9	100.2	9.7
	σ	4.7	3.4	6.5	4.2	9.0	4.5	0.2	2.5	4.0	10.0	4.8		
	N	5	3	5	4	4	5	2	5	5	3	5		

* Each observer participated in N tests, providing $X_1, X_2, \cdots, X_N$ estimates of the 1000-Hz tone level (in dBt) judged to be equal in loudness to the designated speech pressure level.

$$\bar{X} = \frac{\sum^N X_i}{N} \qquad \sigma = \sqrt{\frac{\sum^N (\bar{X} - X_i)^2}{N}}$$

$$\bar{\bar{X}} = \frac{\sum^S \bar{X}_i}{11} \qquad S = \sqrt{\frac{\sum^S (\bar{\bar{X}} - \bar{X}_i)^2}{10}}$$

TABLE III—RESULTS OF MUNSON (1935–1936) THRESHOLD TESTS

Subject Designation	Speech			1000-Hz Tone		
	$\bar{X}$-dBt	σ-dB	N	$\bar{X}$-dBt	σ-dB	N
S_1	12.7	0.8	4	2.2	0	1
S_2	10.3	1.9	4	7.9	0	1
S_3	16.0	3.1	4	5.6	0	1
S_4	16.9	2.0	4	5.9	0	1
S_5	12.4	0.6	3	—	—	—
S_6	12.7	1.2	3	4.3	0	1
S_7	10.4	1.3	2	4.5	0	1
S_8	8.0	2.5	4	2.6	0	1
S_9	12.7	0.3	4	5.7	0	1
S_{10}	11.9	3.6	4	1.2	0	1
S_{11}	11.9	0.9	3	1.6	0	1
	$\bar{\bar{X}}$-dBt 12.4	S-dB 2.5		$\bar{\bar{X}}$-dBt 5.1	S-dB 3.2	

speech sound pressure level is shown on Fig. 13. (The ordinate of Fig. 13 is in terms of Loudness Units—LU rather than sones. See footnote of Section 3.1.)

In terms of equation (8), the curve of Fig. 13 at higher sound pressure levels represents an exponent of 0.455. This is in close agreement with the Van Wynen test results with exponent = 0.486 (see

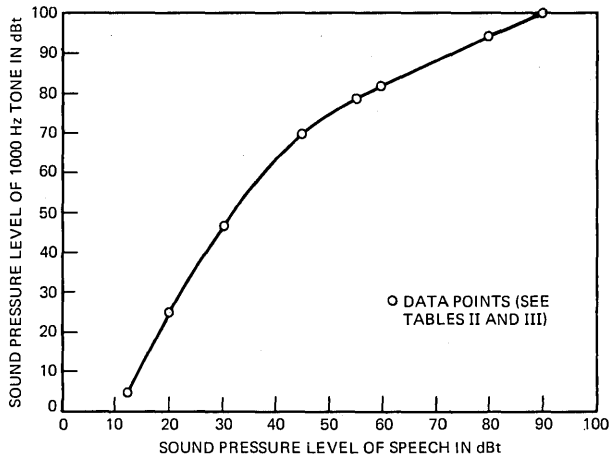


Fig. 12—Loudness balance test results for a 1000-Hz tone and speech (Munson, 1935–1936).

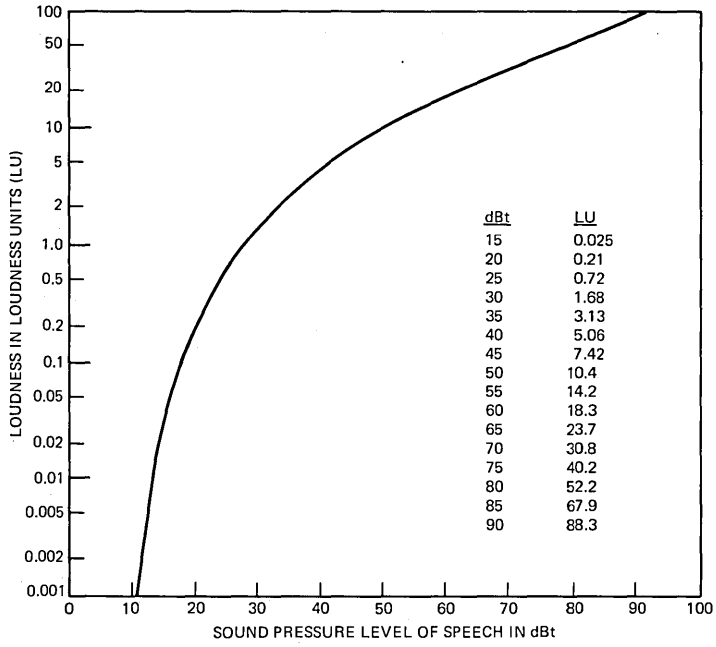


Fig. 13—Loudness (N_s) vs speech sound pressure level for Munson (1935–1936) tests.

previous section), with the "Derived" (from balancing speech against a 1-kHz tone) and "Direct" (from adjusting a speech signal to sound "half as loud" and "twice as loud" as a standard speech signal) from Ref. 32 with approximate exponents of 0.50 and 0.46 respectively, and with the derived curve of Fig. 52, Ref. 31, with an exponent of about 0.51. The exponent does not agree with that from the Steinberg tests which was 0.662 (see previous section), with the exponent of about 0.65 determined from the monaural versus binaural loudness matches of Ref. 32, with the exponent of 0.7 determined from the single phoneme magnitude estimation study of Ref. 33, and with the exponent of 0.6 determined from the limited magnitude estimation work reported in Ref. 47.

The relation of Fig. 13, applicable for wideband speech having the effective spectrum shown by the long-dashed curve of Fig. 9, was then modified to apply for narrow bands of speech, (bands of unit importance) and a flat effective spectrum. To do this, effective spectra were plotted for several of the loudness balance tests, e.g., see Fig. 9. For each of these, the frequency scale was divided into various num-

bers of bands of equal importance as determined from Fig. 5. It was then assumed that Z_s was constant across each band of unit importance, and was equal in magnitude to the Z_s at the center frequency of the unit band. This resulted in several "staircase" approximations to the smooth curve, each exhibiting a granularity depending on the number of bands selected. Study of the various spectra indicated that using 50 bands of unit importance, referred to as 2-percent loudness bands, represented a reasonable compromise between the conflicting requirements of (*i*) accuracy, improved by increasing the number of bands and thus obtaining a closer match between the "staircase" approximations and the smooth curves, and (*ii*) simplicity, obtained by reducing the number of bands.

The required function of n_s versus Z_s, shown on Fig. 6, was obtained from Fig. 13 by a series of successive approximations as outlined in the introductory paragraphs of Section 3.3.4. Specific steps involved included (*i*) adjusting the ordinate and abscissa scales of Fig. 13 to represent a band of 100–5000 Hz (required making allowance for the fact that the effective level of the speech spectrum used in the Munson tests was not flat with frequency) and (*ii*) dividing the ordinate by 50 (to reflect contribution by each 2-percent band) and by 2 (since the Munson test results and the solid curve of Fig. 7 are both in terms of binaural listening while the case of usual interest in telephony is monaural listening).

3.4 *Computational Procedure*

Mechanics of the computation for any given telephone connection consists of two parts, (*i*) the computation of the loudness, N_s, of the received speech, and (*ii*) interpretation of the computed loudness numbers. The former is carried out utilizing a computational form based on Figs. 5 and 9.

The computational method to be described applies specifically to the monaural case, of primary interest in telephone problems. If speech is received binaurally, we assume that the "average" listener is symmetrical and compute the loudness separately for each ear, adding the resulting loudness numbers to obtain the total loudness. [See equation (5).] Thus, in cases where the speech signals delivered to the two ears are identical, the resulting loudness is twice that obtained with one ear alone.

3.4.1 *Computation of Speech Loudness, N_s*

The procedure followed in computing speech loudness is most read-

ily described in terms of the computation form shown in Fig. 14. The computation involves four steps:

(1) Calculation of the received speech signal effective level, Z_s, at a number of selected frequencies;
(2) Determination of n_s, the loudness in a band of unit importance, from Z_s;
(3) Determination of the loudness in a computation band, the product of n_s and ΔS;
(4) Summation of the loudness numerics across the band of interest.

Column 1 of Fig. 14 lists the frequencies at which the computations are made and columns 2 through 7 are used for determining Z_s. The values entered in column 2, obtained from Fig. 9 (reference effective spectrum), apply for an acoustic talking level of 90 dBt.

Columns 3, 4, and 6 are respectively: T_{ot}, the orthotelephonic response of the transmitting element; R_{ot}, the orthotelephonic response of the receiving element; and L, the response of the circuit interconnecting the transmitting and receiving elements. (Refer to Fig. 3 for appropriate orthotelephonic definitions.) In many cases of interest in telephony, the same transmitting and receiving elements are likely to be used in many computations. Column 5 of the form provides for combining responses of these elements with $B_{90} - X$.

The effective level of the received speech spectrum, Z_s, resulting from combining entries of columns 5 and 6, is entered in column 7. The effective level in each computation band is converted to a loudness number for a band of unit importance using the curve of Fig. 6. The loudness numbers, entered in column 8, apply for frequency bands having 2-percent importance. The bands of Fig. 14 with midfrequencies of 600 Hz and above have been selected to give 2-percent importance. Below this frequency, however, the 2-percent bands become very narrow (see Fig. 5), and circuit measurements at the required frequencies provide a fineness in the response curves not required for telephone rating purposes. Thus, wider bands are used and the greater importance of these bands allowed for by means of column 9 which shows the number of 2-percent bands in each computation band. The loudness for a computation band is thus the product of column 8 and column 9 entries, and is entered in column 10. The total loudness, N_s, is obtained by summing column 10.

In some cases, the computation frequency is not the midfrequency of the computation band according to Fig. 5. The midfrequencies

COMPUTATION NO.

DATE

COMPUTED BY

TRANSMITTER

RECEIVER

CIRCUIT

(1) f	(2) B_{90}-X *	(3) T_{OT}	(4) R_{OT}	(5) 2+3+4	(6) L	(7) Z_S (5) + (6)	(8) n_s	(9) Δs	(10) $n_s\Delta s$
150	46.5							13	
200	53							7	
300	62.5							4	
400	68							3	
500	71.5							2	
600	73							1	
700	72.5							1	
800	72							1	
900	71.5							1	
1000	70							1	
1200	69							1	
1400	67.5							1	
1600	66.5							1	
1800	67							1	
2000	67.5							1	
2200	68							1	
2400	68							1	
2600	68							1	
2800	67							1	
3000	67							1	
3300	66							1	
3600	64.5							1	
3900	63.5							1	
4200	62							1	
4500	60.5							1	
4800	59							1	

* APPLIES ONLY FOR A LONG-TERM SPEECH SOUND PRESSURE LEVEL OF 90 dBt AT 2" FROM THE TALKER'S LIPS. FOR OTHER TALKING LEVELS, CHANGE ENTRIES OF COLUMN 2 BY THE DIFFERENCE BETWEEN 90 dBt AND THE TALKING LEVEL OF INTEREST.

$N_S = \Sigma n_s \Delta s =$ _____

Fig. 14—Speech loudness computation form.

were rounded off as shown for purposes of convenience. For the lowest computation band, this rounding off overemphasizes that band's contribution to the loudness, but this generally affects the total loudness number, N_s, only slightly.

The computation bands and the computation frequencies of Fig. 14 were selected for convenience. These can be changed, e.g., to consider 1/3-octave bands, by first locating the band limits on Fig. 5, then

determining (*i*) the center frequency for, and (*ii*) the number of bands of unit importance contained in, the selected computation band. Appropriate values of $B_{90} - X$ can be obtained from Fig. 9.

The procedure described above applies for cases where the loudness of a band of speech is determined by the effective level of the speech within that band. Masking experiments have shown that when the effective level of a continuous spectrum sound changes abruptly with frequency, e.g., a sharp cutoff filter applied to thermal noise, there may be masking in the suppressed frequency band due to energy in the passed frequency band. Since masking has been identified with loudness, it is reasonable to assume that the loudness of a band-limited speech signal will depend not only on the passed region but the suppressed region as well. This effect, termed spread-of-loudness, is considered in the computation by assuming that Z_s does not decrease by more than 10 dB between adjacent 2-percent loudness bands.* If Z_s (column 7 of Fig. 14) shows a more rapid change, a new Z_s is entered which is exactly 10 dB below the Z_s of the preceding or succeeding 2-percent band.

3.4.2 *Interpretation of Total Loudness, N_s*

Loudness in loudness units (LU), although a ratio scale reflecting observers assessment of sound magnitude, has little physical significance. It is thus desirable to express loudness in such a way as to convey a physical interpretation. This can be done by expressing the loudness of a sound in terms of the level of a reference sound which has the same number of loudness units. Thus, speech loudness can be expressed in terms of loudness level, i.e., the level of a 1000-Hz tone having the same loudness. Loudness level is an appropriate way to express speech loudness when comparing computed values with results of tests which involved loudness balancing speech and a 1000-Hz tone. Speech loudness (N_s) can be computed using the form of Fig. 14, then with $N_s = N_{1000}$, the loudness level found from Fig. 7.

Often speech through a test circuit is balanced against speech through some relatively distortionless reference circuit. In these cases, the loudness of the speech transmitted over the test circuit should probably be expressed in terms of the setting of the reference circuit.

* The spread-of-loudness effect was determined from masking data obtained by Fletcher and Munson (Figs. 11 and 12 of Ref. 23 or Figs. 141 and 142 of Ref. 27) using narrow bands of noise. The masking data were plotted in terms of number of 2-percent bands below and above the nominal bandlimits of the noise. A straight line with a slope of 10 dB/band fitted the data with reasonable accuracy, thus permitting use of a simple rule for taking spread-of-loudness effects into account.

Thus, speech loudness can be expressed in terms of several different reference sounds. A single reference would be desirable to enable comparison of results from different tests. Loudness level would be suitable for such purposes. However, considering the different nature of speech and a 1000-Hz tone, a speech signal of specified characteristics might be a more suitable reference, one that would be more intuitively satisfying.

The reference selected for the speech loudness computation procedure is that due to transmitting average speech (Fig. 2) having the reference effective spectrum (Fig. 9) over a circuit that is flat on an orthotelephonic basis (Fig. 3) except for sharp cutoffs at 250 and 4000 Hz. (These frequency limits, selected somewhat arbitrarily, were wide enough to include existing and planned commercial telephone circuits.) The loudness of speech from any telephone circuit can then be specified by the level of the reference that gives an equivalent N_s ; the level, in turn, can be conveniently specified by its integrated rms pressure in dBt. The level of the reference speech spectrum, designated as L_s , is related to its loudness, N_s , as shown on Fig. 8. This curve was derived using the form of Fig. 14, successively changing entries per column 2 thereon, and computing N_s for each such change. It should be noted that (*i*) L_s for the entries of column 2 is 90 dBt for the band 100–5000 Hz, 89.4 dBt for the band 250–4000 Hz and (*ii*) spread-of-loudness was taken into account in the computation of N_s .

3.5 *Comparison of Computed and Observed Results*

Observed results from six different subjective tests were compared to computed results obtained using the computation form of Fig. 14 and the frequency response characteristics of the subjective test systems. (Observed results from two of these tests were used in developing the computational method.) Observed results comprised reference circuit attenuator settings or trunk settings obtained by observers when they loudness-balanced reference circuits and test circuits.

Computed reference circuit settings were determined by first computing the loudness (N_s) for each test condition and its corresponding reference condition (a reference circuit with its attenuator or trunk setting at the value to which observers adjusted to obtain loudness balance) using the test and reference circuit frequency response characteristics.* The computed loudness values (N_s) were then converted to levels of the reference spectrum (L_s). By comparing the L_s of each

* Frequency response characteristics used for the loudness computations discussed in this paper are not reported.

test circuit with the L_s of corresponding reference circuit at observed balance, computed settings were obtained.

Computed and observed results are discussed and tabulated in Sections 3.5.1 thru 3.5.6 and are plotted on Fig. 15. (For the latter, the L_s for the reference circuit at balance is considered to be the observed value, the L_s for the test circuit is the computed value.) Comparisons of computed and observed results are summarized in Table IV in terms of error distributions. These, together with Fig. 15 and tabulated values of ensuing sections, indicate that the computational method provides a high degree of accuracy.

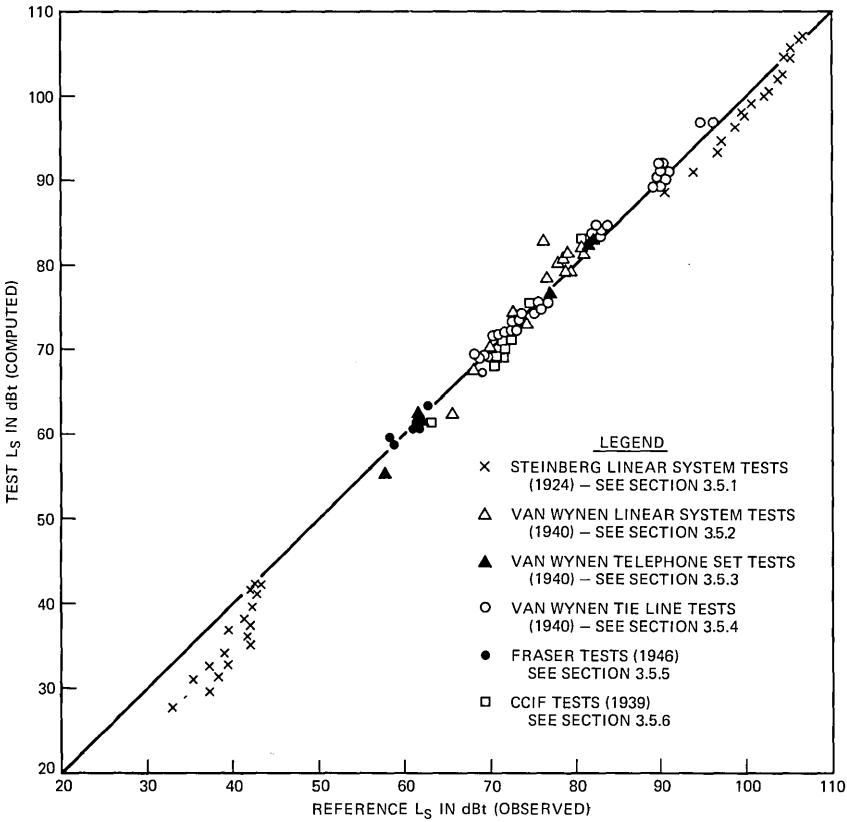


Fig. 15—Comparison of reference (observed) and test (computed) speech loudness.

TABLE IV—SUMMARY OF COMPARISON OF COMPUTED AND OBSERVED RESULTS

Test Designation	Detailed Results In	Computed Minus Observed	
		Average-dB	σ-dB
Steinberg tests (1924)			
Sensation level = 100 dB	3.5.1	1.4	1.1
Sensation level = 39 dB	3.5.1	4.0	2.2
Van Wynen linear system tests (1940)	3.5.2	−0.8	2.1
Van Wynen telephone set tests (1940)	3.5.3	0.3	1.1
Van Wynen tie line tests (1940)	3.5.4	−0.4	0.9
Fraser tests (1940)	3.5.5	0.2	0.9
CCIF tests (1939)	3.5.6	0.8	1.1

3.5.1 *Steinberg Linear System Tests* (1924)

In 1924, J. C. Steinberg conducted loudness balance tests with a system which utilized linear transducers interconnected by either of two channels, a test channel into which various filters were inserted and a reference channel which provided essentially distortionless transmission.[28,29] Observers adjusted an attenuator in the reference channel to obtain loudness balance for each of a number of the test networks.

Computed and observed loudness losses are given in Table V. For purposes of the table, loudness loss is equal to the amount of flat loss by which the reference circuit was varied from its reference setting (for the particular sensation level) so that speech through the reference channel had the same loudness as speech through the test channel. Observed values given for the 100-dB sensation level were used in deriving the S function. (See Section 3.3.4.1.)

For the reference channel, L_s with the 100-dB sensation level reference setting was 107.1 dBt; with the 39-dB sensation level reference setting, L_s = 42.6 dBt. Observed and computed loudness losses per the table may be converted to L_s values for the reference channel by subtracting entries in the "observed" column from the above reference values; L_s for the test channel may be obtained by subtracting entries in the "computed" column, for the appropriate test network, from the reference values given above. Thus, the ref-

erence channel L_s = 102.1 dBt and the test channel L_s = 100.1 dBt for a 500-Hz high-pass filter and a 100-dB sensation level.

Agreement between computed and observed values is reasonably good for the most part. In general, the agreement becomes poorer with increasing distortion and is poorer at the lower sensation levels. Neither of these affects is particularly restrictive from a telephone engineering standpoint since they represent extremes which a well-engineered telephone plant will seldom approach.

3.5.2 *Van Wynen Linear System Tests (1940)*

In 1940, K. G. Van Wynen conducted loudness balance tests using a system equipped with linear transducers. (These tests were not

TABLE V—COMPUTED AND OBSERVED LOUDNESS LOSS VALUES FOR THE STEINBERG LINEAR SYSTEM TESTS (1924)

	Sensation Level = 100 dB			Sensation Level = 39 dB		
	Loudness Loss (dB)			Loudness Loss (dB)		
Test Network*	A Observed	B Computed	C = B − A (dB)	D Observed	E Computed	F = E − D (dB)
125 Hz HPF	0.6	0.2	−0.4	0	0.2	0.2
250	2.0	2.7	0.7	0.4	1.0	0.6
375	3.6	5.0	1.4	−0.2	1.5	1.7
500	5.0	7.0	2.0	1.0	4.8	3.8
625	7.6	9.0	1.4	3.0	5.8	2.8
750	8.5	10.8	2.3	0.6	7.4	6.8
1000	10.5	13.8	3.3	3.2	9.8	6.6
1250	13.2	16.2	3.0	4.5	11.3	6.8
1500	16.6	18.5	1.9	5.3	13.0	7.7
			Avg. = 1.7			Avg. = 4.1
			σ = 1.1			σ = 2.8
750 Hz LPF	9.9	12.6	2.7	9.7	14.4	4.7
1000	7.4	9.4	2.0	7.3	11.6	4.3
1250	6.4	8.0	1.6	5.3	10.0	4.7
1500	4.4	6.5	2.1	3.7	8.6	4.9
2000	3.0	4.3	1.3	0.8	6.6	5.8
2500	2.4	2.5	0.1	1.4	4.3	2.9
3000	1.9	1.5	−0.4	0.2	3.0	2.8
4000	0.7	0.4	−0.3	−0.5	0.4	0.9
			Avg. = 1.1			Avg. = 3.9
			σ = 1.1			σ = 1.5

* HPF = high-pass filter; LPF = low-pass filter. Entries designate nominal cutoff frequencies of the filters. See Ref. 48 for description of the filters which were probably used in the Steinberg tests.

TABLE VI—COMPUTED AND OBSERVED RESULTS FOR THE VAN WYNEN LINEAR SYSTEM TESTS (1940)

Distorting Network*	Observed Reference Circuit Attenuator Setting-dB	L_s – dBt		Computed Reference Circuit Attenuator Setting-dB	Computed Minus Observed-dB
		Test Circuit	Reference Circuit (at Balance)		
2700 Hz LPF	24.2	81.4	80.8	23.6	−0.6
1750 Hz LPF	28.3	78.3	76.6	26.6	−1.7
1000 Hz LPF	31.3	74.1	73.4	30.6	−0.7
250 Hz HPF	25.6	79.1	79.2	25.7	0.1
550 Hz HPF	30.8	73.1	74.0	31.7	0.9
1000 Hz HPF	36.4	67.6	68.1	36.9	0.5
250–2700 Hz BPF	29.0	75.0	76.0	30.0	1.0
550–1750 Hz BPF	38.9	62.5	65.6	42.0	3.1
1000 Hz Peak					
15 dB	26.8	80.2	78.1	24.7	−2.1
30	24.2	81.9	80.8	23.1	−1.1
56	34.5	70.3	70.1	34.3	−0.2
700 Hz Peak					
30 dB	25.8	81.4	79.0	23.4	−2.4
2000 Hz Peak					
30 dB	28.7	82.8	76.2	22.1	−6.6
Falling Loss	25.6	79.2	79.2	25.6	0
Rising Loss	26.4	80.6	78.5	24.3	−2.1
				Avg. =	−0.8
				σ =	2.1

* LPF = low-pass filter, HPF = high-pass filter, BPF = bandpass filter; numbers denote nominal filter cutoff frequencies. "Peak" denotes a resonant circuit defined by the resonant frequency (minimum loss of circuit occurred at resonance) and a dB value indicative of the Q, e.g., 56 dB indicates a high Q, 15 dB indicates a low Q. "Falling Loss" designates a network whose loss decreased monotonically with increasing frequency while for the "Rising Loss" network, the loss increases monotonically with frequency.

reported in the literature.) The transducers were interconnected via one of two channels, the test channel into which various test networks were introduced and a reference channel which was essentially distortionless. Switching between channels was controlled by the observers who, for each network tested, adjusted an attenuator in the reference channel to produce equal loudness. Twelve talker-listener pairs participated in the tests.

Computed and observed attenuator settings are given in Table VI. The computed setting was obtained by noting the difference in L_s for the test circuit and the reference circuit at balance, and modifying the observed setting by this amount.*

* The high- and low-pass filter results, expressed in different terms, were used in deriving the speech loudness computational method. See Section 3.3.4.1 and Fig. 10.

The agreement between observed and computed reference circuit settings is quite good, the error being less than 2 dB in ten of the fifteen cases. Those cases showing the larger errors represent severe distortion conditions which would be approached rarely, if ever, in a well-engineered telephone plant.

3.5.3 *Van Wynen Telephone Set Tests (1940)*

In 1940, K. G. Van Wynen conducted tests in which different types of telephone sets were compared on a loudness basis. (These tests were not reported in the literature.) The laboratory system used consisted essentially of two separate acoustic-to-acoustic channels. One of these was equipped with reference transmitting and receiving telephone sets interconnected by a test network which included an adjustable attenuator. (The telephone sets were of the type described in Ref. 16.) The other channel used each of two different types of telephone sets, designated Test Telephone Sets A and B, with fixed connecting circuitry. (Test Telephone Sets A were of the type described in Ref. 15; Test Telephone Sets B were of a design similar to the telephone sets used with the Working Reference System described in Section 2.2 of this paper.) These sets differed from each other and from the reference telephone set in that they had transducers exhibiting different frequency response characteristics.

Results of the tests are expressed in terms of trunk loss, variable and controlled by the observer, required in the reference channel to deliver speech equal in loudness to that from the test channel with a fixed trunk loss. (Twelve talker-listener pairs participated in the tests.) Computed and observed trunk losses shown in Table VII are in close agreement. The computed trunk loss was obtained by noting the difference in L_s for the test channel and the reference channel, and modifying the observed trunk loss by this difference.

3.5.4 *Van Wynen Tie Line Tests (1940)*

In 1940, K. G. Van Wynen conducted loudness balance tests with laboratory systems which simulated selected circuit conditions characteristic of telephone connections between Bell Telephone Laboratories and the American Telephone and Telegraph Company. These connections involved circuits which were called "tie lines," and thus the designation "tie line tests."

TABLE VII—COMPUTED AND OBSERVED LOUDNESS LEVELS FOR THE VAN WYNEN TELEPHONE SET TESTS (1940)

Test Telephone Set	Network*	Observed Trunk Loss at Loudness Balance-dB†		L_s — dBt		Computed Reference Trunk Loss-dB	Computed Minus Observed Reference Trunk Loss-dB
		Test Channel	Reference Channel	Test Channel	Reference Channel		
A	4000 Hz LPF	5	16.8	76.3	76.8	17.3	0.5
	4000 Hz LPF	25	35.2	55.3	57.7	37.6	2.4
B	4000 Hz LPF	5	11.4	82.7	82.2	10.9	−0.5
	4000 Hz LPF	25	31.5	62.2	61.7	31.0	−0.5
	Rising Loss	5	11.7	82.5	81.8	11.0	−0.7
	Rising Loss	25	31.2	61.5	62.2	31.9	0.7
							Avg. = 0.3
							σ = 1.1

* 4000 Hz LPF = low-pass filter with a nominal cutoff frequency of 4000 Hz; Rising Loss = increasing loss with increasing frequency. The network was switched (by the observer) between reference and test channels and loudness balance achieved by adjusting an attenuator in the reference channel connecting circuit.

† Values are the 1-kHz losses (900-ohm source and termination) of the reference and test channel connecting circuits.

Two test systems, designated "overall" and "sidetone," were used.* Each of these permitted comparison between test networks and either of two reference networks. (The transmitter and receiver used in these tests were of a design described in Ref. 17.) The observers' task was to adjust an attenuator in the reference channel so that speech via this channel and the test channel were equally loud. Twelve talker-listener pairs participated in the tests.

Results of the tests expressed in terms of reference circuit settings are shown in Table VIII. Corresponding values of L_s (dBt) are given as parenthetical entries; those in the "Observed" columns apply for the reference channel while those in the "Computed" columns apply for the test channel. Computed settings were obtained by noting the difference between L_s for the test channel and L_s for the reference channel, and modifying the observed circuit setting by this difference. Computed and observed values are in close agreement.

3.5.5 *Fraser Tests (1946)*

In 1946, J. M. Fraser conducted tests in which two different acoustic-to-acoustic channels were compared on a loudness basis. (These tests were not reported in the literature.) The test channel comprised test telephone sets of the type which were the Bell System Standard at that time,[16,17] and connecting circuitry which simulated selected telephone connections. These simulations, arbitrarily designated in column 1 of Table IX, exhibited different amounts of increasing loss with increasing frequency. The reference channel was the Master Reference System discussed in Section 2.1 of this paper.

Results of the tests are expressed in terms of the Master Reference System trunk (600-ohm attenuator) setting, variable and controlled by the observer, required to achieve loudness balance between the test channel and the reference channel. Twelve talker-listener pairs participated in the tests. Computed and observed reference trunk settings shown in Table IX are in close agreement.

3.5.6 *CCIF Tests (1939)*

In 1939, the CCIF† conducted loudness balance tests on selected simulated telephone connections sent them by the American Telephone

---

* These designations reflect the frequency response characteristics simulated and not testing conditions. Specifically, those with the "sidetone" system were on a listening only basis, and did not involve a talker listening to himself via the system. Sidetone is discussed in Refs. 33, 49, and 50.

† See Section 2.1 of this paper.

TABLE VIII—COMPUTED AND OBSERVED LOUDNESS LOSSES FOR THE VAN WYNEN TIE LINE TESTS (1940)

Test Network*	DL Reference			Peaked Reference		
	Circuit Setting-dB		Computed Minus Observed-dB	Circuit Setting-dB		Computed Minus Observed-dB
	Observed	Computed		Observed	Computed	
Overall						
20 DL	—	—	—	6.3(76.0)	7.3(75.0)	1.0
20 DL	—	—	—	6.8(75.5)	7.3(75.0)	0.5
20–3000 LPF	21.2(73.8)	20.8(74.2)	−0.4	6.9(75.4)	8.1(74.2)	1.2
20–2400 LPF	22.5(72.5)	22.7(72.3)	0.2	9.6(72.7)	10.0(72.3)	0.4
20–1900 LPF	24.7(70.3)	25.7(69.3)	1.0	13.8(68.5)	13.0(69.3)	−0.8
20–1900 LPF	26.0(69.0)	25.7(69.3)	−0.3	12.9(69.4)	13.0(69.3)	0.1
20–RL1	22.0(75.0)	21.6(73.4)	−0.4	8.7(73.6)	8.9(73.4)	0.2
20–RL2	23.3(71.7)	22.9(72.1)	−0.4	10.4(71.9)	10.2(72.1)	−0.2
20–RL3	24.4(70.6)	23.3(71.7)	−1.1	11.3(71.0)	10.6(71.7)	−0.7
			$\bar{x}$ = −0.2			$\bar{x}$ = 0.1
			σ = 0.6			σ = 0.6
Sidetone						
Ref. DL	0.7(94.8)	−1.3(96.8)	−2.0	−13.7(96.2)	−14.3(96.8)	−0.6
−10 DL	11.5(83.0)	10.7(83.8)	−0.8	−1.6(83.3)	−2.1(83.8)	−0.5
−20 DL	19.2(75.3)	19.0(75.5)	−0.2	5.4(77.0)	6.9(75.5)	1.5
R900	5.4(89.8)	6.2(89.0)	0.8	−6.9(89.1)	−7.0(89.2)	−0.1
R1300	5.4(90.0)	4.4(91.0)	−1.0	−8.6(90.9)	−8.7(91.0)	−0.1
R1800	5.2(89.9)	3.1(92.0)	−2.1	−8.2(89.7)	−10.5(92.0)	−2.3
R3200-A	12.2(82.4)	11.1(83.5)	−1.1	−0.8(82.7)	−1.6(83.5)	−0.8
R3200-B	12.2(82.5)	10.1(84.6)	−2.1	−2.1(83.9)	−2.8(84.6)	−0.7
RL4	5.7(89.6)	5.0(90.3)	−0.7	−8.2(90.7)	−7.8(90.3)	+0.4
			$\bar{x}$ = 1			$\bar{x}$ = −0.4
			σ = 0.9			σ = 1.0
			Overall: $\bar{x}$ = −0.4			
			σ = 0.9			

* The number preceding the overall network designation defines the value of an attenuator associated with that network. For the DL network, this number (20) was the setting of an attenuator associated with the network when it was balanced against the peaked reference. The DL reference and the "20 DL" overall were the same circuit (incorporating in essence a variable attenuator) and differed from the DL circuits for the sidetone tests. The response of all DL networks showed decreasing loss with increasing frequency. LPF = low-pass filters with a nominal cutoff frequency (in Hz) noted. The RL networks had differing amounts of increasing loss with increasing frequency. Networks designated R were resonant circuits with high values of loss at the indicated resonant frequencies. Networks R3200-A and R3200-B differed in that the former had a slightly greater Q.

TABLE IX—COMPUTED AND OBSERVED LOUDNESS LEVELS FOR THE TEST AND REFERENCE CHANNELS OF THE FRASER TESTS (1946)

Test Network	Observed Reference Trunk Setting (dB)	L_s – dBt		Computed Reference Trunk Setting (dB)	Computed Minus Observed (dB)
		Test System	Reference System		
A	33.3	67.5	69.2	35	1.7
B	39.9	63.3	62.8	39.4	−0.5
C	41.5	60.8	61.2	41.9	0.4
D	43.0	58.8	58.9	43.1	0.1
E	40.3	61.6	61.9	40.6	0.3
F	43.8	59.7	58.4	42.5	−1.3
G	40.6	60.9	61.6	41.3	0.7
					Avg. = 0.2
					σ = 0.9

and Telegraph Company. (These tests were not reported in the literature.) These simulations were a subset of those discussed in Section 3.5.4 of this paper.

The CCIF tests involved two steps. First, the simulated connection designated "DL Reference" with 20 dB of added loss and the SFERT* were loudness balanced. Twenty talker-listener pairs, formed from a trained test crew of five persons, each provided seven balances. Results comprised settings (in dB) of the SFERT trunk (600-ohm attenuator) required to achieve balance; the average value was 30.9 dB.†

The L_s for the DL Reference (with 20 dB added loss) was computed using frequency response characteristics for the DL Reference of Section 3.5.4 while the L_s for the SFERT (with an attenuator setting of 30.9 dB) was computed from response characteristics for the reference system of Section 3.5.5. The respective values were 74 dBt and 71.6 dBt. That is, the CCIF loudness balance tests showed these circuits to be equal while the speech loudness computational method indicates that they differ by 2.4 dB. Probably this difference is due to the system response characteristics (primarily transducers) used in the computations, and not to the computational method itself being in error. In the computations, corrections reflecting differences in specific transducers of the same type were applied where known,

* See Section 2.1 of this paper.

† As noted in Section 2.1, the SFERT line attenuator setting (in dB) required for loudness balance with a given system under test is, by definition, the Reference Equivalent (in dB) of that system. Thus, the Reference Equivalent of the 20 DL Reference connection is 30.9 dB.

but these were not known in all cases, and average responses were necessarily used in these cases. This argument is supported to some extent by the internal consistency for the tests of Section 3.5.4, the tests of Section 3.5.5, and the CCIF tests.

The DL Reference was then used as a secondary reference against which ten simulated telephone connections were loudness balanced. Computed and observed results given in Table X are in close agreement.

3.5.7 *Comparison of Computed Loudness Losses and Reference Equivalents*

Reference Equivalents are, by definition, loudness ratings of telephone connections obtained using the SFERT (and, by inference, the MRS and NOSFER). The speech loudness computational method provides loudness ratings in orthotelephonic terms. Since both are based on speech loudness, information from the preceding sections can be used to derive a correction factor which would permit translating ratings between the two frames of reference. (Such translation depends on bandwidth characteristics of telephone connections and therefore will depend on, among other things, the particular transducer types involved.) Tables XI and XII summarize the appropriate information.

The average conversion factor from Table XI is 12.7 dB while that from Table XII is 15.0 dB, a difference of 2.3 dB. Some possible reasons for this difference were discussed in Section 3.5.6. In addition, the observer groups differed for the Fraser and CCIF tests, and there were probably some minor differences in test procedure.

IV. LABORATORY MEASURING SYSTEM

The speech loudness computational procedure could be realized as a laboratory measuring system in any one of several forms. The "ideal" system would implement exactly the various functions of the computational procedure. The system sound source would energize the transmit end of a telephone connection with an acoustic signal having the amplitude characteristic of Fig. 2 and a time rate-of-change based on Fig. 5. The signal received at the other end of the connection would be applied to a measuring subsystem containing a network with a response according to Fig. 4. System operation so far corresponds to computing the effective spectrum, Z_s. (See column 7 of the computation form on Fig. 14.)

TABLE X—LOUDNESS BALANCE RESULTS WITH DL REFERENCE CONNECTION AS REFERENCE

Test System		Observed DL Reference Trunk Setting (dB)	L_s − dBt		Computed DL Reference Trunk Setting (dB)	Computed Minus Observed (dB)
Type	Designation*		Test System	DL Reference System		
Overall	20–1900 LPF	23.5	68.3	70.5	25.7	2.2
	20–2400 LPF	21.7	71.3	72.3	22.7	1.0
	20–3000 LPF	20.8	73.2	73.2	20.8	0.0
	20–RL2	22.3	71.1	71.7	22.9	0.6
	20–RL3	22.2	70.7	71.8	23.3	1.1
						Avg. = 1.0
						σ = 0.7
Sidetone	Ref DL	13.5	82.8	81	11.7	−1.8
	−10 DL	23.7	69.8	70.8	24.7	1.0
	−20 DL	31.8	61.5	62.7	33.0	1.2
	R900	19.0	75.2	75.5	19.3	0.3
	R3200-A	22.9	69.5	71.6	25.0	2.1
						Avg. = 0.6
						σ = 1.3
					Overall: $\bar{x}$ = 0.8	
					σ = 1.1	

* See footnote of Table VIII.

TABLE XI—RELATION BETWEEN REFERENCE EQUIVALENTS (RE) AND LOUDNESS LOSS (LL) FROM THE FRASER TESTS (SECTION 3.5.5)

Test System	RE* (dB)	L_s* (dBt)	LL† (dB)	RE Minus LL (dB)
A	33.3	67.5	21.9	11.4
B	39.9	63.3	26.1	13.8
C	41.5	60.8	28.6	12.7
D	43.0	58.8	30.6	12.4
E	40.3	61.6	27.8	12.5
F	43.8	59.7	29.7	14.1
G	40.6	60.9	28.5	12.1
			Avg. =	12.7
			σ =	0.9

* Columns 2 and 3 of Table IX.
† 89.4 minus entries of column 3. See last paragraph of Section 3.4.2.

The remainder of the measuring subsystem would consist of a device to reflect the spread-of-loudness effect, a device having the law of Fig. 6, an integrating circuit, and, finally, a display mechanism reflecting the law of Fig. 8.

The "ideal" system could be modified, without changing its essen-

TABLE XII—RELATION BETWEEN REFERENCE EQUIVALENTS (RE) AND LOUDNESS LOSS (LL) FROM CCIF TESTS (SECTION 3.5.6)

Test System	RE* (dB)	L_s† (dBt)	LL‡ (dB)	RE Minus LL (dB)
20-DL REF	30.9	74	15.4	15.5
20-1900 LPF	34.4	68.3	21.1	13.3
20-2400 LPF	32.6	71.3	18.1	14.5
20-3000 LPF	31.7	73.2	16.2	15.5
20-RL2	33.2	71.1	18.3	14.9
20-RL3	33.1	70.7	18.7	14.4
REF DL	24.4	82.8	6.6	17.8
−10 DL	34.6	69.8	19.6	15.0
−20 DL	42.7	61.5	27.9	14.8
R900	29.9	75.2	14.2	15.7
R3200-A	33.8	69.5	19.9	13.9
			Avg. =	15.0
			σ =	1.1

* Derived from column 3 entries of Table X and relation that DL Reference with 20-dB trunk setting had a reference equivalent of 30.9 dB.
† From column 4 of Table X and text of Section 3.5.6.
‡ 89.4 minus entries of column 3. See last paragraph of Section 3.4.2.

tial attributes, by combining the "X" network response with that of the signal source. The source output would then have an amplitude characteristic as shown by the dash-dot curve of Fig. 9. (See column 2 of the computation form on Fig. 14.)

Realization of the system design outlined above is limited by a number of practical considerations. Exact implementation would require using an artificial mouth and an artificial ear which duplicated their human counterparts within the context of definitions given in Fig. 3. Such are not available. However, simpler devices are available which are probably adequate for purposes of many telephone engineering problems.

Secondly, the source signal does not incorporate the dynamic properties of speech. This is probably not important as long as the telephone connection under test is composed of linear elements. However, measurements on connections incorporating nonlinear elements, particularly carbon transmitters, may be in error. Carbon transmitters in real use are activated by real speech; testing such with an applied signal of the type referred to above might well result in different device operating characteristics. This matter is under study.

In this paper, we consider a system which is somewhat simpler than the "ideal" and which probably reflects adequate accuracy for telephone rating purposes. This system, derived from the computational procedure, is designated the EARS (for *E*lectro-*A*coustic *R*ating System), and is dealt with in ensuing sections.* A graphical computation procedure is also described. Finally, a comparison is made between ratings based on the EARS and ratings obtained from subjective tests.

The EARS utilizes an artificial voice and a 6-cm^3 coupler while the computational procedure utilizes the concept of orthotelephonic transmission. Response definitions appropriate to artificial mouth and 6-cm^3 coupler measurements are shown on Fig. 16 in terms similar to the orthotelephonic response definitions of Fig. 3.

Typical artificial voice/6-cm^3 coupler amplitude responses of connections employing specific telephone set types are given in Ref. 18. As with orthotelephonic responses, artificial voice/6-cm^3 coupler responses vary widely depending on telephone set and transducer types used. Conversion curves for translating between orthotelephonic and artificial voice/6-cm^3 coupler responses are not as variable, but are still highly dependent on telephone set and transducer type.

* Reference 51 describes derivation of a measuring system which is similar to the EARS.

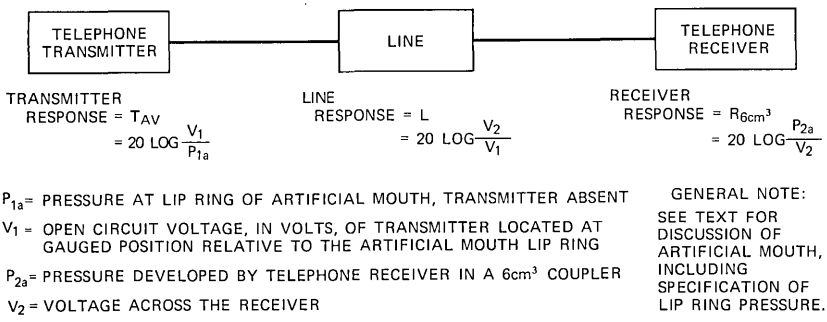


Fig. 16—Telephone circuit responses in artificial voice — 6-cm³ coupler terms.

4.1 *Derivation of EARS*

We begin by considering the computation form of Fig. 14. For the reference effective spectrum (column 2), we compute the cumulative contribution to loudness as a function of computation band, i.e., we assume an orthotelephonically flat transmission system with 0 dB of gain for the sum of columns 3, 4, and 6, and compute the entries for column 10. Cumulating these and dividing each cumulative entry by the total loudness (N_s) results in a relation between cumulative percentage contribution to loudness and frequency. This relation is shown by the small circles plotted on Fig. 17; the ordinate is cumulative and the abscissa is upper frequency limit of the computation band, obtained from Fig. 5. The relation of Fig. 17 differs from that of Fig. 5 because the latter is for a flat effective spectrum (Z_s = a constant) whereas the former is for the effective spectrum of speech ($Z_s = B_{90} - X$) shown by the dot-dash curve of Fig. 9.

The points plotted on Fig. 17 can be reasonably approximated by a straight line drawn from 100 Hz to 5000 Hz on the logarithmic frequency scale. Thus, 2-percent loudness bands, or any bands of equal loudness derived from this curve, have a logarithmic relation to frequency and may be found by laying off equal lengths along the logarithmic frequency scale.

In terms of a measuring system, the above suggests using a source signal which has a flat amplitude versus frequency characteristic but which sweeps across the band at a logarithmic rate. Such a signal would cover equal loudness bands of the effective spectrum ($B_{90} - X$) in equal time divisions, corresponding to equal length divisions along the abscissa of Fig. 17. Integration of the sweep on a time basis

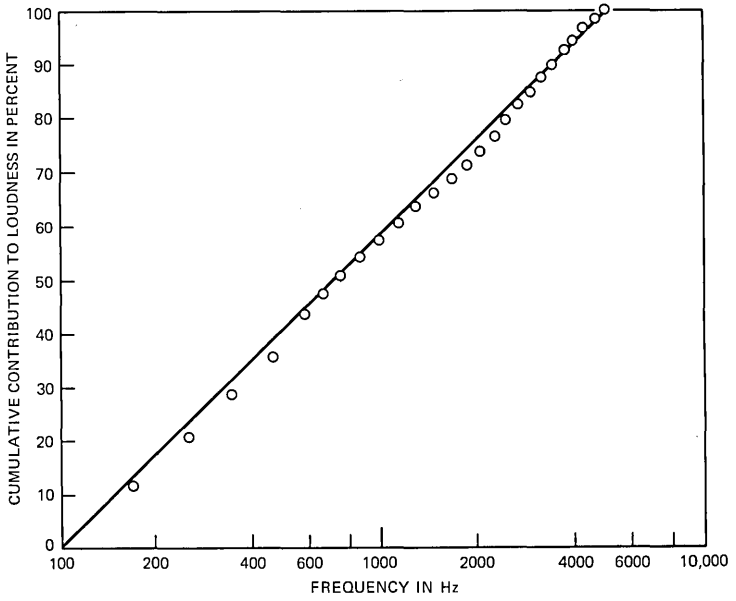


Fig. 17—Cumulative contribution to loudness of speech for the reference effective spectrum (see Figs. 9 and 14).

would result in a value proportional to the total loudness of the effective spectrum since the signal amplitude is constant with frequency.

In essence, the operation described above translates the amplitude weighting of $Z_s = B_{90} - X$ to frequency weighting. The computation form of Fig. 14 could be modified to reflect this by changing columns 1 and 9 to conform with the straight line of Fig. 17 and entering a constant value of Z_s in column 2. The value of Z_s can be determined by dividing the computed N_s for $B_{90} - X$ by the number of computation bands (e.g., 50) to find n_s (the loudness per band of unit importance), then entering the curve of Fig. 6 at that value of n_s.

The source signal described above, in combination with the acoustic-to-acoustic response of a telephone connection, provides the effective spectrum of the received speech for that connection. The next step, then, is to apply the loudness scale of Fig. 6 in order to convert to loudness units.

The loudness scale of Fig. 6 is linear on logarithmic coordinates, i.e., $Z_s = \text{k} \log n_s$, above the knee of the curve, and the number of loudness units increases tenfold for a 44-dB change in effective level,

or doubles for a 13.2-dB change in effective level. (This is reasonably checked by subjective loudness tests with filters as plotted on Fig. 10. In the Steinberg tests, a 9.1-dB reduction in effective level corresponded to a 50-percent reduction in loudness, while in Van Wynen tests the level reduction required to produce half loudness was 12.4 dB). This relation is assumed to hold over the entire range of effective levels. The effect of this assumption on accuracy will be considered in a later section.

At the receiving end of a telephone connection activated by the logarithmic signal source at the transmitting end, the received acoustic signal is applied to a measuring circuit and ultimately appears as a voltage across a resistance. At any instant, this voltage is proportional to $10^{C/20}$ where C is the response, in dB, of the telephone connection at a particular frequency. However, because of the sweep rate of the signal source, voltage elements appearing across the resistance in time sequence are proportional to $10^{Z_s/20}$, where Z_s is the effective level of the received speech spectrum. We have already postulated that loudness doubles for a 13.2-dB change in Z_s ; the voltage elements we would like to see across the resistance should therefore be proportional to $10^{Z_s/44}$.* Thus, the desired voltage elements, V_1 , are proportional to the 2.2 root of the voltage elements, V_{1x} , which actually appear, i.e.,

$$V_1 \propto 2.2 \sqrt{V_{1x}} \tag{9}$$

or

$$(V_1)^{2.2} \propto V_{1x} . \tag{10}$$

The above relationship suggests that voltage elements proportional to $10^{Z_s/20}$ be applied to the input of a 2.2-to-1 compressor. With such a device, changes at the input proportional to

$$10^{(Z_{s1}-Z_{s2}/20)}$$

appear at the output proportional to

$$10^{(Z_{s1}-Z_{s2}/44)},$$

---

* If loudness N_2 is twice as large as loudness N_1,

$$\frac{N_2}{N_1} = 10^{(Z_{s2}-Z_{s1}/x)} = 2,$$

$$Z_{s2} - Z_{s1} = x \log_{10} 2 = 13.2,$$

and

$$x = 44.$$

Thus

$$N \propto 10^{Z_s/44}.$$

which is the 2.2 root of the input change. Thus, the action of the compressor approximates the loudness scale of Fig. 6 in converting elements of the received effective spectrum to loudness units.

Integration of the voltage with respect to time may be either on a square-law or linear basis. Without the compressor, square-law integration adds voltage elements proportional to $10^{Z_s/10}$ while linear integration adds voltage elements proportional to $10^{Z_s/20}$. By including the compressor with linear integration, elements proportional to $10^{Z_s/44}$ are added, these in turn being proportional to loudness units. The combination of the two gives, in effect, 2.2 root law addition.

System operation to this point corresponds to use of the computation form of Fig. 14 in that the integrated voltage, V_s, is proportional to N_s. A means of expressing this voltage in dB is needed for the same reasons that the curve of Fig. 13 was derived to permit interpretation of N_s. This can be accomplished using an indicating meter with a dB scale. By making the meter circuit an "averaging" one with a time constant long compared to the sweep time of the source signal, deflections of the needle will be approximately proportional to the total number of loudness units. A dB scale corresponding to the conversion specified by the curve of Fig. 13 could be provided. This might be difficult since the constant of proportionality of the needle deflections to total number of loudness units is not readily obtainable. Instead, advantage was taken of another approximation.

Comparison of the curves of Figs. 6 and 13 shows that above their respective knees the shapes of these curves are the same. If we assume a straight-line relationship between L_s and N_s as was done earlier for the relation between n_s and Z_s, the addition of distortionless attenuation in any speech spectrum will produce a dB loudness level change equal to the change in attenuation. Thus, the meter scale can be so calibrated that it obeys the same law of addition as that provided by the compressor and linear integrator in converting elements of effective level to loudness, namely a 13.2-dB change in effective level would correspond to doubling the total number of loudness units. The scale on the meter would show a difference of 13.2 dB between half-scale and full-scale deflection of the needle, or between any two deflection points, the greater of which is twice the smaller. Consequently, the meter reading would reflect a change in distortionless loss in the telephone connection under test, dB for dB.

4.2 *Description of EARS*

A block diagram of EARS is shown on Fig. 18. The system consists of two parts: a signal source and a measuring subsystem. The latter includes provision for measuring both voltage and pressure.[3]

4.2.1 *Signal Source*

The logarithmic oscillator provides at its output terminals a signal which sweeps logarithmically with time from 300 Hz to 3300 Hz to 300 Hz and has a flat amplitude versus frequency characteristic. The reason for limiting the measuring band of the EARS to 300–3300 Hz is a practical one. The use of partial connection ratings as an engineering tool implicitly requires that, for any given connection, the sum of the partial ratings should approximately equal the overall rating. Thus, the bandwidth used to obtain these ratings should approximate the bandwidth of the most restrictive element(s) in the connection in order to avoid cumulating bandwidth penalties when summing partial ratings. The specific limits of 300 Hz and 3300 Hz were selected by reviewing bandwidth characteristics of various telephone equipments and facilities.

The oscillator sweep rate is 6 times per second where a sweep is defined in terms of the 300–3300-Hz band. Criteria leading to selec-

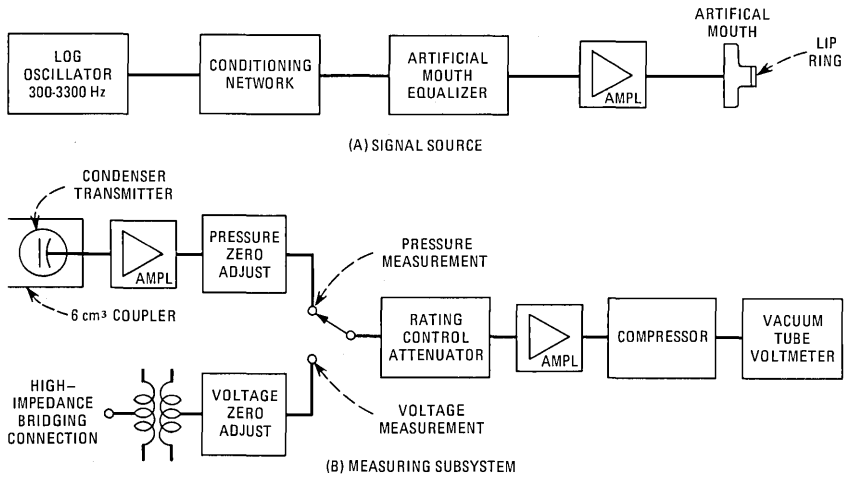


Fig. 18—Block diagram of the EARS.

tion of this sweep rate were (*i*) the sweep interval (1/sweep rate) should be small as compared to an easily realizable integrating time constant for the indicating meter and (*ii*) the rate should approximate the syllabic rate of speech. The latter characteristic appears desirable when measuring carbon transmitters in order to ensure that these operated at an efficiency comparable to that obtained under actual use conditions, i.e., with real speech applied. (Recent measurements indicate that the sweep rate can be changed over the range 2 to 10 sweeps per second without significantly changing the ratings of many telephone sets of modern design.)

The conditioning network (a 6-dB attenuator) is momentarily switched out of the source circuit prior to measuring carbon transmitters thus increasing the source signal level by 6 dB. The higher level is intended to condition the transmitters to operate at the proper level. Conditioning is not required when measuring linear transmitters.

The artificial mouth equalizer comprises a passive network whose frequency response is the inverse of the electric-to-acoustic response of the artificial mouth. The loss of this network is compensated for by the amplifier.

The artificial mouth used is a permanent magnet, moving coil loudspeaking unit, and is, for all practical purposes, the equivalent of an earlier proposal.[52] The mouth includes, as an integral part, a lip ring which is used as a reference for obtaining the proper spatial relationship between the artificial mouth and telephone instruments under test. The location of the lip ring has been empirically determined so as to correspond approximately to the plane of the lips of a human mouth.[52]

Ideally, the source arrangement should, with an oscillator output which is flat with frequency over the band of interest, provide an output pressure at the artificial mouth lip ring which is also flat with frequency.* Practically, control of the overall response to within ±1 dB of the 1000-Hz value provides acceptable operation, introducing less than about 0.2 dB error in ratings.

4.2.2 *Measuring Subsystem*

The measuring subsystem is arranged to permit both voltage and pressure measurements. For voltage measurements, the input is a

* The pressure is measured with a Type L microphone.[53] The microphone is located in a carefully gauged position so selected that the pressure measured at that point corresponds closely to the pressure at the center of the lip ring opening.[54]

high-impedance transformer, and is bridged across the selected impedance (usually a 900-ohm resistor) terminating the telephone connection at the point where voltage measurement is desired. The transformer is connected to an attenuator, used in system calibration to compensate for gain drift of amplifiers, thence to a switch node.

The pressure measuring circuit consists of a 6-cm^3 test coupler equipped with a Type L pressure microphone.* The microphone is connected to a condenser microphone amplifier which provides bias voltage to the microphone and which produces at its output a voltage proportional to the pressure developed in the 6-cm^3 cavity by the telephone receiver under test. The amplifier is connected through a calibrating attenuator to a switch node.

The measurement mode is selected by operating the switch to "pressure measurement" or "voltage measurement." The switch swinger connects to an attenuator (designated Rating Control), amplifier, compressor, and vacuum tube voltmeter. The compressor design employed provides a 2.2-to-1 characteristic over a limited range. Operation is confined to this range by holding the compressor output voltage at a constant value, indicated on the vacuum tube voltmeter as the "reference" or "zero" point, and adjusting for rating changes using the Rating Control. This takes advantage of the fact that, as previously noted, flat attenuation changes ahead of the compressor equate, on a dB-for-dB basis, to loudness level changes.

4.2.3 *System Calibration*

System calibration consists of first removing the condenser microphone from the coupler and locating it at gauged position relative to the artificial mouth lip ring, then adjusting the source to deliver reference test pressure at that point. This adjustment is based on the condenser microphone calibration, and does not involve the EARS measuring subsystem. When the proper test pressure has been obtained, the EARS measuring system is switched to the pressure measuring mode, Rating Control is set at the "0" dB, and the "Pressure Adjust Attenuator" set to obtain reference reading on the vacuum tube voltmeter. Since the pressure spectrum being measured is flat with frequency, pressure level equates to loudness pressure level.

Calibration of the voltage measuring mode proceeds in a similar

* The simple 6-cm^3 test coupler used, conforming in design to present standards (Fig. 3 of Ref. 55), has been found suitable for measurements of the type considered in this report. However, handsets with ear caps of unusual shape may require a different coupler configuration.

manner. A voltage signal, derived from the log oscillator, is applied to a 900-ohm resistor. The voltage developed is first read directly with a voltmeter. Then the high-impedance bridging connection of the EARS measuring subsystem is connected across the resistor, the system switched to the voltage measuring mode, and the Rating Control set to read the voltage measured with the voltmeter above, i.e., if the latter was −2 dB relative to 1 volt, then the rating control is set to read −2 dB relative to the "reference" or "zero" setting. The "Voltage Zero Adjust" is then set to obtain "reference" or "zero" reading on the measuring subsystem vacuum tube voltmeter. Since the voltage spectrum being measured is flat with frequency, voltage level equates to loudness voltage level.

4.2.4 *Rating Measurements*

Ratings of partial or overall telephone connections are established by (*i*) the reading of the Rating Control and (*ii*) the reference pressure and/or voltage levels employed. Examples of rating measurements, including signal levels employed with the present EARS, are given on Figs. 19, 20, and 21. The transmitting and receiving loops of Figs. 19 and 20 respectively are the same as those for the overall connection of Fig. 21. Sum of the component ratings indicates a loudness loss of 17.3 dB (= −18.3 + 25.6 + 10) while the overall rating from Fig. 16 is 17.1, a discrepancy of 0.2 dB. Had the measurement band been increased from 300–3300 Hz to, for example, 100–5000 Hz, the discrepancy would have been somewhat larger because of the bandpass response of both the transmitter and receiver.

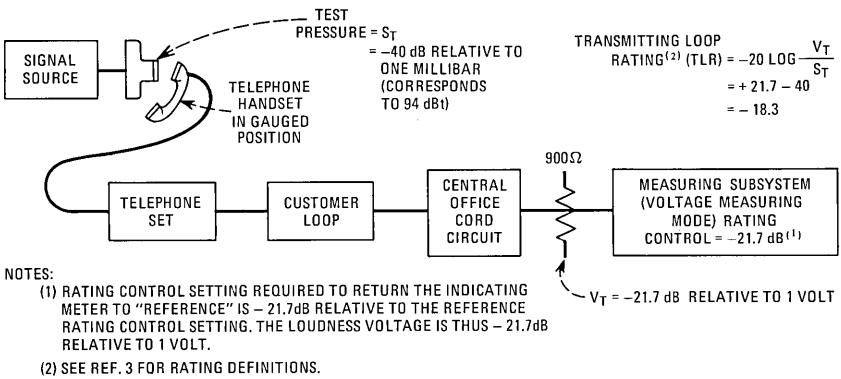
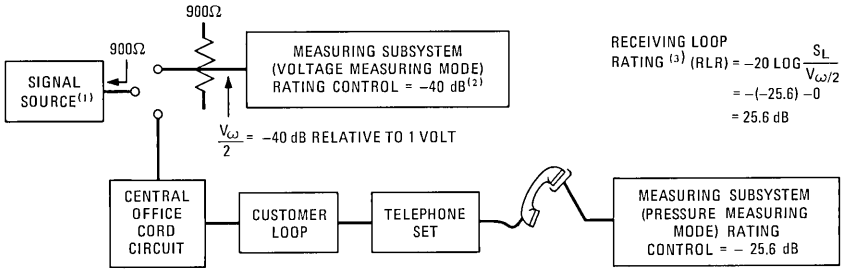


Fig. 19—Transmitting loop rating.



NOTES:

(1) SIGNAL SOURCE PROVIDES AN OUTPUT SIGNAL WHICH CONSISTS OF THE TEST SPECTRUM (FLAT WITH FREQUENCY) MODIFIED BY THE FREQUENCY RESPONSE OF COMPONENTS WHICH NORMALLY PRECEDE THE RECEIVING COMPONENT IN AN OVERALL CONNECTION. THIS IS NECESSARY IN ORDER TO PRESERVE ADDITIVITY OF COMPONENT RATINGS.

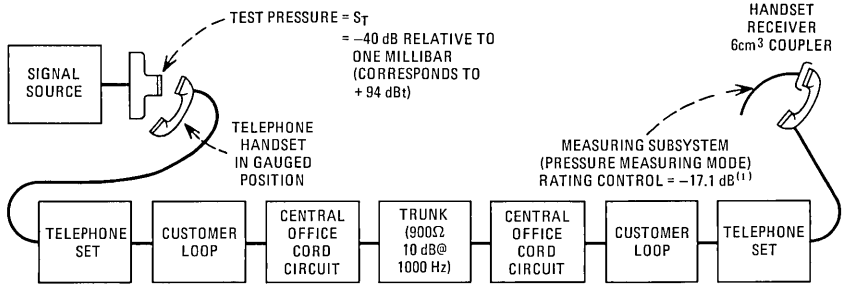
(2) RATING CONTROL SETTING RELATIVE TO REFERENCE SETTING IN ORDER THAT THE INDICATING METER NEEDLE IS AT REFERENCE FOR THE SELECTED TEST VOLTAGE. THE LOUDNESS TEST VOLTAGE IS THUS −40 dB RELATIVE TO 1 VOLT.

(3) SEE REF. 3 FOR RATING DEFINITIONS. THE TEST REFERENCE VOLTAGE IS SO SELECTED THAT IT IS NUMERICALLY EQUAL, IN dBre 1 VOLT, TO THE SELECTED TEST PRESSURE LEVEL IN dBt (SEE FIG. 19.) FOR ANY OTHER TEST VOLTAGE, THE EQUATION FOR RLR MUST CONTAIN A CONSTANT.

Fig. 20—Receiving loop rating.

4.3 *Graphical Computation of Loudness*

A form for graphical determination of loudness ratings from the amplitude response characteristics of telephone connections is shown on Fig. 22. This form is based on the same considerations as those leading to the EARS, and is discussed in this paper because it is the



NOTES:

(1) RATING CONTROL SETTING REQUIRED TO RETURN THE INDICATING METER TO "REFERENCE" IS 17.1 dB BELOW THE REFERENCE SETTING OF THE CONTROL. THUS, THE RECEIVED LOUDNESS PRESSURE IS 17.1 dB BELOW THE TEST PRESSURE LEVEL APPLIED AT THE TRANSMITTING END OF THE CONNECTION.

(2) SEE REF. 3 FOR RATING DEFINITIONS. NOTE THAT FOR OVERALL MEASUREMENTS, A POSITIVE RATING REPRESENTS A LOUDNESS LOSS.

OVERALL RATING (2) (OR) = −20 LOG S2 / ST
= −(−17.1) + 0
= 17.1 dB

Fig. 21—Overall rating.

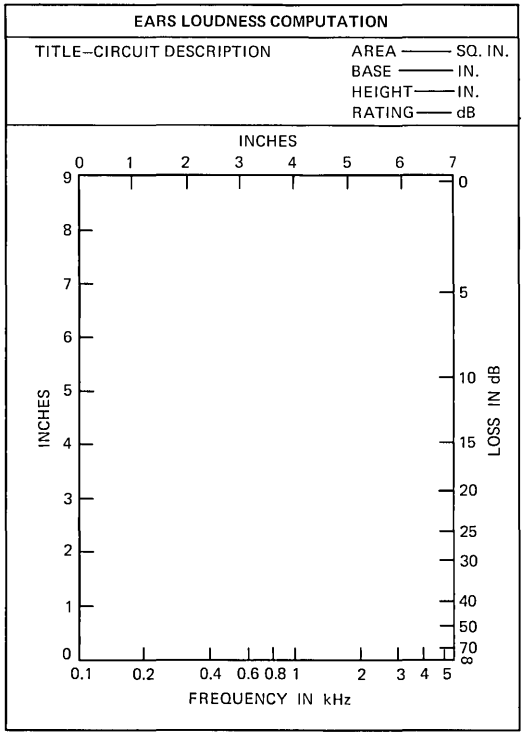


Fig. 22—Graph paper for computing loudness ratings.

vehicle used to study the effects of differences between the loudness computation method and the EARS.

The lower abscissa scale is frequency in Hz, corresponding on a logarithmic basis to the upper abscissa scale in inches. (Use of inches is arbitrary; any length units could be used.) Thus, equal increments on the upper scale correspond to equal increments of log f_2/f_1 where $f_2 > f_1$. (This reflects the straight-line approximation shown on Fig. 17). The equal distance increments closely approximate bands of the $B_{90} - X$ spectrum which are interpreted to be of equal loudness when listened to by the human ear.

The right-hand ordinate scale is proportional to the 2.2 root of a voltage, current, or pressure. This scale and grid are constructed according to the equation

$$44 \log 2X = 55 - L \tag{11}$$

where

X = inches (left-hand ordinate scale)
L = loss of circuit in dB (right-hand ordinate scale).

(Equation (11) assumes that the reference input to the connection is flat with frequency.) The ordinate is measured in inches from the bottom line of the graph which corresponds to zero output voltage, current, or pressure. Correspondence between ordinate scales is demonstrated in Table XIII. Boxes at the top of the sheet are provided for recording graph measurements.

In order to use the graph paper, the loss (in dB) of the partial or overall telephone connection must be known over the band of interest. This loss data may be in terms of

$$\frac{\text{input pressure}}{\text{output pressure}}, \qquad \frac{\text{input voltage}}{\text{output voltage}}, \qquad \frac{\text{input pressure (millibars)}}{\text{output voltage (volts)}},$$

or

$$\frac{\text{input voltage (volts)}}{\text{output pressure (millibars)}}.*$$

The loss frequency characteristic is plotted on the graph paper. The right-hand ordinate scale may be adjusted by a constant for negative losses, i.e., gains. Where large losses are encountered, greater accuracy can be obtained by similarly applying an adjustment constant. To illustrate, if the lowest loss across the band of interest is -15 dB, the values along the right-hand ordinate scale should have 15 subtracted from them. If, on the other hand, the lowest loss is $+15$ dB, the right-hand ordinate values should be increased by 15.

The average height of the response as plotted on the graph paper is determined by measuring the area (in square inches) under the curve and dividing by the base width (in inches). The average height is then located on the left-hand scale and the corresponding dB value read from the right-hand scale. This dB value is the rating based on reference input level.

For cases where the input spectrum is not flat with frequency, the graphical computation of a circuit rating involves determining two areas. The first of these is obtained from a plot of the actual input spectrum on the graph paper, the second from a plot of the input

* See Ref. 3.

TABLE XIII— RELATION BETWEEN ORDINATE DISTANCE AND LOSS SCALES

Left-Hand Ordinate Scale-Distance (Inches)	Input/Output	Right-Hand Ordinate Scale-Loss (dB)
0	∞	∞
0.5	561	55
5	3.55	11
8.89	1	0

spectrum modified by the circuit response. The loudness rating is then the dB difference between the losses computed for these areas.

V. COMPARISON OF GRAPHICALLY DETERMINED RATINGS AND OBSERVED RESULTS

The speech loudness computation method accurately predicts loudness performance of telephone connections (see Section 3.5). Laboratory realization of this method, the EARS, involved a number of simplifying assumptions. Thus, validation of the EARS approach requires considering (*i*) the accuracy of the EARS in predicting loudness performance and (*ii*) the effects of the various simplifying assumptions. We do this by comparing computed ratings and observed results for the subjective tests of Section 3.5. (Numerous other tests have been performed which support the EARS concept, but these are not considered because either the tests were limited, i.e., incomplete test designs, small observer groups, etc., or the test system amplitude response characteristics are not known.)

TABLE XIV—FEATURES OF THE SPEECH LOUDNESS COMPUTATION METHOD AND THE EARS METHOD

Feature	Computation Method	EARS Method
Loudness Law	Figs. 6 and 8	Linear Portions of Figs. 6 and 8
Spread-of-Loudness Correction	Yes	No
Analysis Bandwidth	100–5000 Hz	300–3300 Hz
Reference Bandwidth	250–4000 Hz	300–3300 Hz
Rating Definition	Orthotelephonic Terms (Fig. 3)	Artificial Voice/6-cm^3 Coupler Terms (Fig. 16)

Validation of the EARS utilized the graphical computation method described in Section 4.3. This approach was necessary because the systems used in the tests were not available for direct measurement using the EARS. However, the frequency response characteristics of these systems were known and, thus, their loudness ratings could be computed.

In theory, the graphical form (Fig. 22) and the computation form (Fig. 14) should provide about the same results. The reference of the former (the horizontal line at 0 dB of loss) corresponds to that of the latter (entries of column 2) because of the relation between frequency weightings of the two references. There is, however, one important difference between the two forms. For the graphical form, the loudness versus effective level relation has a constant slope (straight-line portion of Fig. 6) while for the computation form, this relation has a pronounced change in slope at low effective levels (see Fig. 6). The effects of this difference should be most apparent at low received speech levels where the graphical method would indicate a higher received speech loudness than would the computational method.

With the above in mind, we can now consider the ratings for the various tests under a number of different conditions. These conditions, listed in Table XIV, reflect the differences in features of the speech loudness computation method and the EARS.

Computed and observed ratings are given in Tables XV through XIX in terms of loudness loss (positive entries) or loudness gain (negative entries). The arrows and associated numbers below the tabular entries refer to distributions of differences between the various conditions considered.

We will consider Table XV in some detail, noting that since the other tables have an identical arrangement, such discussion will in general similarly apply to these tables. Referring to Table XV, column 1 designates the network tested, and is repeated from Table V. Column 2 designates the setting of the reference network for which the test observers judged the test and reference networks to provide equally loud speech. Columns 3 through 9 each contain 3 subcolumns; "a" and "b" are the loudness ratings of the test and reference channels respectively and "c" represents the error, equal to "a" entries minus "b" entries.

Column 3 entries are essentially a repeat of the information contained in Table V, obtained by converting the Table V entries to L_s values, then subtracting these from 89.4 dBt, the level of the reference speech spectrum of Fig. 2 transmitted over a system that is flat

TABLE XV—COMPUTED LOUDNESS LOSSES (dB) FOR STEINBERG LINEAR SYSTEM TESTS (1924)
(See also Table V)

1	2	3			4			5			6			7			8			9		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
					GRAPHICAL COMPUTATION (Fig. 22)																		
					Orthotelephonic Responses												Artificial Voice/6-cm³ Coupler Responses					
					Same as Computation Form Except																		
	Observed Ref. Ckt. Setting (dB)	Computation Form (Fig. 14)			Linear Loudness Law			Linear Loudness Law, No Spread-of-Loudness			Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz			Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz		
Test System		Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ
Sensation Level = 100 dB																						
125 Hz HPF	−22.0	−17.5	−17.1	−0.4	−17.3	−17.7	0.4	−18.9	−19.3	0.4	−12.3	−12.8	0.5	−15.4	−14.8	−0.6	−4.0	−4.1	0.1	−4.5	−3.9	−0.6
250 Hz HPF	−20.6	−15.0	−15.7	0.7	−14.5	−16.3	1.8	−15.9	−17.9	−2.0	−9.3	−11.4	2.1	−15.3	−13.4	−1.9	0.6	−2.7	3.3	−4.4	−2.5	−1.9
375 Hz HPF	−19.0	−12.7	−14.1	1.4	−12.5	−14.7	2.2	−13.7	−16.3	2.6	−6.7	−9.8	3.1	−13.4	−13.0	−0.4	2.6	−1.1	3.7	−2.7	−0.9	−1.8
500 Hz HPF	−17.6	−10.7	−12.7	2.0	−10.5	−13.3	2.8	−11.2	−14.9	3.7	−4.6	−8.4	3.8	−10.5	−10.6	0.1	4.8	0.3	4.5	−0.1	0.5	−0.6
625 Hz HPF	−15.0	−8.7	−10.1	1.4	−8.8	−10.7	1.9	−9.4	−12.3	2.9	−2.6	−5.8	3.2	−8.3	−8.0	−0.3	6.1	2.9	3.2	1.6	3.1	−1.5
750 Hz HPF	−14.1	−6.9	−9.2	2.3	−6.8	−9.8	3.0	−7.4	−11.4	4.0	−0.8	−4.0	4.1	−5.8	−7.1	1.3	7.1	3.8	3.3	3.0	4.0	−1.0
1000 Hz HPF	−12.1	−3.9	−7.2	3.3	−4.7	−7.8	3.1	−4.0	−9.4	5.4	2.6	−2.9	5.5	−1.5	−5.1	3.6	10.4	5.8	4.6	7.4	6.0	1.4
1250 Hz HPF	−9.4	−1.5	−4.5	3.0	−1.5	−5.1	3.6	−0.3	−6.7	6.4	6.2	−0.2	6.4	3.5	−2.4	5.9	13.2	8.5	4.7	11.2	8.7	2.5
1500 Hz HPF	−6.0	0.8	−1.1	1.9	1.7	−1.7	3.4	3.2	−3.3	6.5	9.8	3.2	6.6	8.8	1.0	7.8	16.1	11.9	4.2	15.0	12.1	2.9
750 Hz LPF	−12.7	−5.1	−7.8	2.7	−4.8	−8.4	3.6	−3.6	−10.0	6.4	3.0	−3.5	6.5	5.0	−5.7	10.7	11.2	5.2	6.0	18.7	5.4	13.3
1000 Hz LPF	−15.2	−8.3	−10.3	2.0	−7.3	−10.9	3.6	−7.1	−12.5	5.4	−0.3	−6.0	5.7	−0.6	−8.2	7.6	8.7	2.7	6.0	12.1	2.9	9.2
1250 Hz LPF	−16.2	−9.7	−11.3	1.6	−9.5	−11.9	2.4	−9.5	−13.5	4.0	−2.8	−7.0	4.2	−4.5	−9.2	4.7	7.7	1.7	6.0	11.0	1.9	9.1
1500 Hz LPF	−18.2	−11.2	−13.3	2.1	−11.1	−13.9	2.8	−11.3	−15.5	4.2	−4.7	−9.0	4.3	−7.2	−11.2	4.0	4.7	−0.3	5.0	5.7	−0.1	5.8
2000 Hz LPF	−19.6	−13.4	−14.7	1.3	−13.1	−15.3	2.2	−13.9	−16.9	3.0	−7.3	−10.4	3.1	−10.8	−12.6	1.8	2.6	−1.7	4.3	2.0	−1.5	3.5
2500 Hz LPF	−20.2	−15.2	−15.3	0.1	−14.3	−15.9	1.6	−15.7	−17.5	1.8	−9.1	−11.0	1.9	−13.3	−13.2	−0.1	0.2	−2.3	2.5	−1.5	−2.1	0.6
3000 Hz LPF	−20.7	−16.2	−15.8	−0.4	−15.3	−16.4	1.1	−16.4	−18.0	1.6	−9.9	−11.5	1.6	−14.3	−13.7	−0.6	−0.8	−2.8	2.0	−3.0	−2.6	−0.4
4000 Hz LPF	−21.9	−17.3	−17.0	−0.3	−16.8	−17.6	0.8	−18.1	−19.2	1.1	−11.5	−12.7	1.2	−15.5	−14.9	−0.6	−3.8	−4.0	0.2	−4.2	−3.8	−0.4
Error: Average				1.4			2.4			3.6			3.7			2.5			3.8			2.4
σ				1.1			1.0			1.9			1.9			3.6			1.7			4.3
Difference Distributions					0.2, 0.5			−0.5, 1.0			6.6			−3.3, 1.8					−1.0, 3.1			
														8.7, 1.0								
																+11.0, 2.1						
Sensation Level = 39 dB																						
125 Hz HPF	38.4	47.0	46.8	0.2	43.7	42.7	1.0	42.1	41.1	1.0	48.7	47.6	1.1	45.6	45.4	0.2	57.0	56.3	0.7	56.5	56.5	0
250 Hz HPF	38.8	47.8	47.2	0.6	45.5	43.1	2.4	45.1	41.5	3.6	51.7	48.0	3.7	45.7	45.8	−0.1	61.6	56.7	4.9	56.6	56.9	−0.3
375 Hz HPF	38.2	48.3	46.6	1.7	48.5	42.5	6.0	47.3	40.9	6.4	54.3	47.4	6.9	47.6	45.2	2.4	63.6	56.1	7.5	58.3	56.3	2.0
500 Hz HPF	39.4	51.6	47.8	3.8	50.5	43.7	6.8	49.8	42.1	7.7	56.4	48.6	7.8	50.5	46.4	4.1	65.8	57.3	8.5	60.4	57.5	2.9
625 Hz HPF	41.4	52.6	49.8	2.8	52.2	43.7	6.5	51.6	44.1	7.5	58.4	50.6	7.8	52.7	48.4	4.3	67.1	59.3	7.8	62.6	59.5	3.1
750 Hz HPF	39.0	54.2	47.4	6.8	54.2	43.3	10.9	53.6	41.7	11.9	60.2	48.2	12.0	55.2	46.0	9.2	68.1	56.9	11.2	64.0	57.1	6.9
1000 Hz HPF	41.6	56.6	50.0	6.6	56.3	45.9	10.4	57.0	44.3	12.7	63.6	50.8	12.8	59.5	48.6	10.9	71.4	59.5	11.9	68.4	59.7	8.7
1250 Hz HPF	42.9	58.1	51.3	6.8	59.5	47.2	12.3	60.7	45.6	15.1	67.2	52.1	15.1	64.5	49.9	14.6	74.2	60.8	13.4	72.2	61.0	11.2
1500 Hz HPF	43.7	59.8	52.1	7.7	62.7	48.0	14.7	64.2	46.4	17.8	70.8	52.9	17.9	69.8	50.7	19.1	77.1	61.6	15.5	76.0	61.8	14.2
750 Hz LPF	48.1	61.2	56.5	4.7	56.2	52.4	3.8	57.4	50.8	6.6	64.0	57.3	6.7	66.0	55.1	10.9	72.2	66.0	6.2	79.7	66.2	13.5
1000 Hz LPF	45.7	58.4	54.1	4.3	53.7	50.0	3.7	53.9	48.4	5.5	60.7	54.9	5.8	60.4	52.7	7.7	69.7	63.6	6.1	73.1	63.8	9.3
1250 Hz LPF	43.7	56.8	52.1	4.7	51.5	48.0	3.5	51.5	46.4	5.1	58.2	52.9	5.3	56.5	50.7	5.8	68.7	61.6	7.1	72.0	61.8	10.2
1500 Hz LPF	42.1	55.4	50.5	4.9	49.9	46.4	3.5	49.7	44.8	4.9	56.3	51.3	5.0	53.8	49.1	4.7	65.7	60.0	5.7	66.7	60.2	6.5
2000 Hz LPF	39.2	53.4	47.6	5.8	47.9	43.5	4.4	47.1	41.9	5.2	53.7	48.4	5.3	50.2	46.2	4.0	63.6	57.1	6.5	63.0	57.3	5.7
2500 Hz LPF	39.8	51.1	48.2	2.9	46.7	44.1	2.6	45.3	42.5	2.8	51.9	49.0	2.9	47.7	46.8	0.9	61.2	57.7	3.5	59.5	57.9	1.6
3000 Hz LPF	38.6	49.8	47.0	2.8	45.7	42.9	2.8	44.6	41.3	3.3	51.1	47.8	3.3	46.7	45.6	1.1	60.2	56.5	3.7	58.0	56.7	1.3
4000 Hz LPF	37.9	47.2	46.3	0.9	44.2	42.2	2.0	42.9	40.6	2.3	49.5	47.1	2.4	45.5	44.9	0.6	57.2	55.8	1.4	56.8	56.0	0.8
Error: Average				4.0			5.7			7.0			7.8			6.0			7.1			5.7
σ				2.2			3.9			4.6			4.6			5.3			3.9			4.6
Difference Distributions					−2.5, 2.5			−0.3, 1.0			6.6			−3.4, 2.2					−1.1, 3.3			
														8.7, 1.0								
																11, 2.1						

TABLE XVI—COMPUTED LOUDNESS LOSSES (DB) FOR VAN WYNEN LINEAR SYSTEM TESTS (See also Table VI)

1	2	3			4			5			6			7			8			9		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
		GRAPHICAL COMPUTATION (Fig. 22)																				
					Orthotelephonic Responses												Artificial Voice/6-cm³ Coupler Responses					
					Same as Computation Form Except						Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz			Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz		
Test System	Observed Ref. Ckt. Setting (dB)	Computation Form (Fig. 14)			Linear Loudness Law			Linear Loudness Law, No Spread-of-Loudness														
		Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ
2700 Hz LPF	24.2	8.0	8.6	−0.6	7.9	8.9	−1.0	6.9	7.6	−0.7	13.4	14.3	−0.9	10.6	12.5	−1.9	21.1	21.6	−0.5	20.2	21.7	−1.5
1750 Hz LPF	28.3	11.1	12.8	−1.7	10.9	13.0	−2.1	9.9	11.7	−1.8	16.5	18.4	−1.9	15.0	16.6	−1.6	24.6	25.7	−1.1	25.9	25.8	0.1
1000 Hz LPF	31.3	15.3	16.0	−0.7	16.3	16.0	0.3	15.9	14.7	1.2	22.5	21.4	1.1	23.8	19.6	4.2	29.7	28.7	1.0	34.2	28.8	5.4
250 Hz HPF	25.6	10.3	10.2	0.1	10.3	10.3	0	9.1	9.0	0.1	15.7	15.7	0	9.3	13.9	−4.6	25.5	23.0	2.5	19.4	23.1	−3.7
550 Hz HPF	30.8	16.3	15.4	0.9	14.9	15.5	−0.6	14.0	14.2	−0.2	20.7	20.9	−0.2	14.1	19.1	−5.0	29.2	28.2	1.0	22.8	28.3	−5.5
1000 Hz HPF	36.4	21.8	21.3	0.5	21.2	21.1	0.1	22.0	19.8	2.2	28.5	26.5	2.0	23.3	24.7	−1.4	36.0	33.8	2.2	31.2	33.9	−2.7
250–2700 Hz BPF	29.0	14.4	13.4	1.0	14.5	13.7	0.8	13.6	12.4	1.2	20.4	19.1	1.3	11.9	17.3	−5.4	30.0	26.4	3.6	21.4	26.5	−5.1
550–1750 Hz BPF	38.9	26.9	23.8	3.1	26.8	23.6	3.2	20.5	22.3	1.8	33.1	29.0	4.1	23.8	27.2	−3.4	43.0	36.3	6.7	33.6	36.4	−2.8
1000 Hz Peak																						
15 dB	26.8	9.2	11.3	−2.1	9.7	11.5	−1.8	8.0	10.2	−2.2	14.7	16.9	−2.2	11.0	15.1	−4.1	22.5	24.2	−1.7	20.3	24.3	−4.0
30 dB	24.2	7.5	8.6	−1.1	7.5	8.9	−1.4	5.6	7.6	−2.0	12.4	14.3	−1.9	7.0	12.5	−5.5	20.8	21.6	−0.8	16.2	21.7	−5.5
56 dB	34.5	19.1	19.3	−0.2	19.0	19.2	−0.2	17.2	17.9	−0.7	24.0	24.6	−0.6	17.0	22.8	−5.8	30.5	31.9	−1.4	26.3	32.0	−5.7
700 Hz Peak																						
30 dB	25.8	8.0	10.4	−2.4	8.4	10.5	−2.1	6.8	9.2	−2.4	13.4	15.9	−2.5	8.5	14.1	−5.6	21.7	23.2	−1.5	17.7	23.3	−5.6
2000 Hz Peak																						
30 dB	28.7	6.6	13.2	−6.6	7.3	13.4	−6.1	5.8	12.1	−6.3	12.3	18.8	−6.5	7.2	17.0	−9.8	19.8	26.1	−6.3	15.4	26.2	−10.8
Falling Loss	25.6	10.2	10.2	0	10.8	10.3	0.5	9.4	9.0	0.4	16.0	15.7	0.3	12.5	13.9	−1.4	23.8	23.0	0.8	21.5	23.1	−1.6
Rising Loss	26.4	8.8	10.9	−2.1	9.3	11.1	−1.8	7.5	9.8	−2.3	14.1	16.5	−2.4	14.8	14.7	0.1	21.4	23.8	−2.4	25.1	23.9	1.2
Error: Average				−0.8			−0.8			−0.6			−0.6			−3.5			0.1			−3.3
σ				2.1			2.0			2.6			2.4			3.1			2.9			3.6
Difference Distributions				0,0.5			−1.1,0.7			6.6			−4.4,3.0						−3.0,3.9			
													8.1,0.9									
																9.4,0.8						

TABLE XVII—COMPUTED LOUDNESS LOSSES (dB) FOR VAN WYNEN TELEPHONE SET TESTS (1940)
(See also Table VII)

1		2	3			4			5			6			7			8			9		
			a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
			GRAPHICAL COMPUTATION (Fig. 22)																				
						Orthotelephonic Responses												Artificial Voice/6-cm³ Coupler Responses					
						Same as Computation Form Except																	
Test System		Observed Ref. Ckt. Setting (dB)	Computation Form (Fig. 14)			Linear Loudness Law			Linear Loudness Law, No Spread-of-Loudness			Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz			Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz		
Test Tele. Set	Test Network		Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ
A	5dB-4000 Hz LPF	16.8	13.1	12.6	0.5	12.6	12.5	0.1	12.4	10.5	1.9	18.5	17.1	1.4	10.5	10.5	0	26.2	23.3	2.9	18.7	18.0	0.7
	25dB-4000 Hz LPF	35.2	34.1	31.7	2.4	32.6	30.9	1.7	32.4	28.9	3.5	38.5	35.5	3.0	30.5	28.9	1.6	46.2	41.7	4.5	38.7	36.4	2.3
B	5dB-4000 Hz LPF	11.4	6.7	7.2	−0.5	6.7	7.1	−0.4	7.7	5.1	2.6	14.7	11.7	3.0	7.2	5.1	2.1	23.4	17.9	5.5	16.3	12.6	3.7
	25dB-4000 Hz LPF	31.5	27.2	27.7	−0.5	26.7	27.2	−0.5	27.7	25.2	2.5	34.7	31.8	2.9	27.2	25.2	2.0	43.4	38.0	5.4	36.3	32.7	3.6
	5dB-Rising Loss	11.7	6.9	7.6	−0.7	6.5	6.9	−0.4	7.5	4.9	2.6	14.5	11.5	3.0	8.1	6.4	1.7	23.5	17.4	6.1	17.3	15.2	2.1
	25dB-Rising Loss	31.2	27.9	27.2	0.7	26.5	26.4	0.1	27.5	24.4	3.1	34.5	31.0	3.5	28.1	25.9	2.2	43.5	36.9	6.6	37.3	34.7	2.6
Error: Average					0.3			0.1			2.7			2.8			1.6			5.2			2.5
σ					1.1			0.8			0.5			0.7			0.7			1.2			1.0
Difference Distributions	Test Telephone Set A			−1.0,0.5			−0.2			6.1			−8.0			7.7			−7.5			
	Test Telephone Set B			−0.6,0.5			+1.0			7.0			−7.0,0.5			8.8			−6.6,0.5			
	Reference Telephone Set			−0.5,0.3			−2.0			6.6			−6.1,0.7			6.1			−4.3,1.5			

TABLE XVIII—COMPUTED LOUDNESS LOSSES (dB) FOR VAN WYNEN TIE LINE SYSTEM TESTS (See also Table VIII)

1	2	3			4			5			6			7			8			9		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
					GRAPHICAL COMPUTATION (Fig. 22)																		
					Orthotelephonic Responses												Artificial Voice/6-cm³ Coupler Responses					
					Same as Computation Form Except																		
		Computation Form (Fig. 14)			Linear Loudness Law			Linear Loudness Law, No Spread-of-Loudness			Analysis Band = Reference Band = 100-5000 Hz			Analysis Band = Reference Band = 300-3300 Hz			Analysis Band = Reference Band = 100-5000 Hz			Analysis Band = Reference Band = 300-3300 Hz		
Test System	Observed Ref. Ckt. Setting (dB)	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ
Overall-DL Reference																						
20-3000 LPF	21.2	15.2	15.6	−0.4	14.3	14.6	−0.3	13.4	13.4	0	19.9	20.0	−0.1	11.6	12.5	−0.9	25.9	26.5	−0.6	18.9	19.5	−0.6
20-2400 LPF	22.5	17.1	16.9	0.2	15.7	15.9	−0.2	15.1	14.7	0.4	21.7	21.3	0.4	13.4	13.8	−0.4	28.4	27.8	0.6	21.6	20.8	0.8
20-1900 LPF	24.7	20.1	19.1	1.0	17.6	18.1	−0.5	18.1	16.9	1.2	24.6	23.5	1.1	16.7	16.0	0.7	32.5	30.0	2.5	25.6	23.0	2.6
	26.0		20.4	−0.3		19.4	−1.8		18.2	−0.1		24.8	−0.2		17.3	−0.6		31.3	1.2		24.3	1.3
20-RL1	22.0	16.0	16.4	−0.4	15.5	15.4	0.1	13.9	14.2	−0.3	20.4	20.8	−0.4	12.9	13.3	−0.4	26.8	27.3	−0.5	20.2	20.3	−0.1
20-RL2	23.3	17.3	17.7	−0.4	17.0	16.7	0.3	15.1	15.5	−0.4	21.6	22.1	−0.5	14.3	14.6	−0.3	28.7	28.6	0.1	22.5	21.6	0.9
20-RL3	24.4	17.7	18.8	−1.1	17.3	17.8	−0.5	15.3	16.6	−1.3	21.9	23.2	−1.3	14.9	15.7	−0.8	28.9	29.7	−0.8	23.2	22.7	0.5
Overall-Peaked Reference																						
20-DL	6.3	14.4	13.4	1.0	13.4	12.1	1.3	12.2	10.7	1.5	18.8	17.2	1.6	11.3	9.0	2.3	25.3	24.4	0.9	18.3	16.5	1.8
	6.8		13.9	0.5		12.6	0.8		11.2	1.0		17.7	1.1		9.5	1.8		24.9	0.4		17.0	1.3
20-3000 LPF	6.9	15.2	14.0	1.2	14.3	12.7	1.6	13.4	11.3	2.1	19.9	17.8	2.1	11.6	9.6	2.0	25.9	25.0	0.9	18.9	17.1	1.8
20-2400 LPF	9.6	17.1	16.7	0.4	15.7	15.4	0.3	15.1	14.0	1.1	21.7	20.5	1.2	13.4	12.3	1.1	28.4	27.7	0.7	21.6	19.8	1.8
20-1900 LPF	13.8	20.1	20.9	−0.8	17.6	19.6	−2.0	18.1	18.2	−0.1	24.6	24.7	−0.1	16.7	16.5	0.2	32.5	31.9	0.6	25.6	24.0	1.6
	12.9		20.0	0.1		18.7	−1.1		17.3	0.8		23.8	0.8		15.6	1.1		31.0	1.5		23.1	2.5
20-RL1	8.7	16.0	15.8	0.2	15.5	14.5	1.0	13.9	13.1	0.8	20.4	19.6	0.8	12.9	11.4	1.5	26.8	26.8	0	20.2	18.9	1.3
20-RL2	10.4	17.3	17.5	−0.2	17.0	16.2	0.8	15.1	14.8	0.3	21.6	21.3	0.3	14.3	13.1	1.2	28.7	28.5	0.2	22.5	20.6	1.9
20-RL3	11.3	17.7	18.4	−0.7	17.3	17.1	0.2	15.3	15.7	−0.4	21.9	22.2	−0.3	14.9	14.0	0.9	28.9	29.4	−0.5	23.2	21.5	1.7
Sidetone-DL Reference																						
Ref. DL	0.7	−7.4	−5.4	−2.0	−7.3	−5.5	−1.8	−8.0	−7.1	−0.9	−2.2	−0.5	−1.7	−9.4	−8.2	−1.2	4.1	5.8	−1.7	−2.9	−1.1	−1.8
−10 DL	11.5	5.6	6.4	−0.8	5.1	5.7	−0.6	3.7	4.3	−0.6	10.2	10.8	−0.6	2.5	3.8	−1.3	16.6	17.1	−0.5	9.7	10.2	−0.5
−20 DL	19.2	13.9	14.1	−0.2	12.2	13.3	−1.1	10.8	11.8	−1.0	17.3	18.5	−0.8	9.9	10.9	−1.0	24.9	24.8	−0.1	17.8	18.0	−0.2
R 900	5.4	0.2	−0.6	0.8	0	−1.1	1.1	−1.4	−2.4	1.0	4.3	4.2	0.1	−1.7	−3.3	−1.6	11.3	10.7	0.6	5.1	3.8	1.3
R 1300	5.4	−1.6	−0.6	−1.0	−1.0	−0.9	−0.1	−2.4	−2.3	−0.1	4.2	4.3	−0.1	−2.9	−3.2	0.3	10.5	10.6	−0.1	4.0	3.9	0.1
R 1800	5.2	−2.6	−0.5	−2.1	−2.4	−1.0	−1.4	−3.7	−2.4	−1.3	2.8	4.3	−1.5	−4.5	−3.3	−1.2	9.6	10.5	−0.9	2.8	3.9	−1.1
R 3200-A	12.2	5.9	7.0	−1.1	4.8	6.3	−1.5	3.3	4.9	−1.6	10.0	11.5	−1.5	1.7	4.2	−2.5	17.4	18.1	−0.7	9.6	11.2	−1.6
R 3200-B	12.2	4.8	6.9	−2.1	4.2	6.3	−2.1	2.8	4.9	−2.1	9.3	11.5	−2.2	2.9	4.2	−1.3	16.4	17.8	−1.4	10.2	11.0	−0.8
RL4	5.7	−0.9	−0.2	−0.7	−1.6	−0.8	−0.8	−3.1	−2.3	−0.8	3.5	4.3	−0.8	−4.0	−3.2	−0.8	9.8	9.8	0	2.8	3.0	−0.2
Sidetone Peaked Reference																						
Ref. DL	−13.7	−7.4	−6.8	−0.6	−7.3	−7.2	−0.1	−8.0	−9.6	1.6	−2.2	−2.2	0	−9.4	−11.2	1.8	4.1	4.3	−0.2	−2.9	−3.8	0.9
−10 DL	−1.6	5.6	6.1	−0.5	5.1	4.7	+0.4	3.7	3.3	0.4	10.2	9.8	0.4	2.5	1.6	0.9	16.1	17.0	−0.9	9.7	9.2	0.5
−20 DL	5.4	13.9	12.4	1.5	12.2	11.0	1.2	10.8	9.7	1.1	17.3	16.3	1.0	9.9	7.9	2.0	24.9	23.3	1.6	17.8	15.4	2.4
R 900	−6.9	0.2	0.3	−0.1	0	−1.1	1.1	−1.4	−2.5	1.1	4.3	4.1	0.2	−1.7	−4.2	2.5	11.3	11.3	0	5.1	3.4	1.7
R 1300	−8.6	−1.6	−1.5	−0.1	−1.0	−3.0	2.0	−2.4	−4.4	2.0	4.2	2.1	2.1	−2.9	−6.0	3.1	10.5	9.6	0.9	4.0	1.7	2.3
R 1800	−8.2	−2.6	−0.3	−2.3	−2.4	−2.3	−0.1	−3.7	−3.7	0	2.8	2.9	−0.1	−4.5	−5.4	0.9	9.6	10.5	−0.9	2.8	2.2	0.6
R 3200-A	−0.8	5.9	6.7	−0.8	4.8	5.3	−0.5	3.3	3.8	−0.5	10.0	10.4	−0.4	1.7	2.2	−0.5	17.4	17.8	−0.4	9.6	9.9	−0.3
R 3200-B	−2.1	4.8	5.5	−0.7	4.2	4.1	0.1	2.8	2.7	0.1	9.3	9.3	0	2.9	1.1	1.8	16.4	16.7	−0.3	10.2	8.8	1.4
RL4	−8.2	−0.9	−1.3	0.4	−1.6	−2.7	1.1	−3.1	−4.0	0.9	3.5	2.5	1.0	−4.0	−5.5	1.5	9.8	9.8	0	2.8	1.8	1.0
Error: Average				−0.4			−0.1			0.2			0			0.3			0.1			0.8
σ				0.9			1.1			1.0			1.0			1.4			0.9			1.2
Difference Distributions				−0.8,0.6			−1.4,0.5			6.5			−7.7,0.6						−7.1,0.7			
													6.6,0.5									
																7.4,0.5						

TABLE XIX—COMPUTED LOUDNESS LOSSES (dB) FOR FRASER TESTS (1946) (See also Table IX)

1	2	3			4			5			6			7			8			9		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
		GRAPHICAL COMPUTATION (Fig. 22)																				
					Orthotelephonic Responses												Artificial Voice/6-cm³ Coupler Responses					
					Same as Computation Form Except																	
					Linear Loudness Law			Linear Loudness Law, No Spread-of-Loudness			Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz			Analysis Band = Reference Band = 100–5000 Hz			Analysis Band = Reference Band = 300–3300 Hz		
Test System	Observed Ref. Ckt. Setting (dB)	Computation Form (Fig. 14)																				
		Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ	Test	Ref.	Δ
A	33.3	21.9	20.2	1.7	21.6	20.1	1.5	20.3	18.4	1.9	26.9	25.0	1.9	21.0	22.3	−1.3	33.1	33.9	0.8	28.6	33.7	−5.1
B	39.9	26.1	26.6	−0.5	26.2	26.7	−0.5	25.0	25.0	0	31.5	31.6	−0.1	25.9	28.9	−3.0	37.1	40.5	−3.4	32.9	40.3	−7.4
C	41.5	28.6	28.2	0.4	28.1	28.3	−0.2	26.9	26.6	0.3	33.5	33.2	0.3	27.6	30.5	−2.9	39.6	42.1	−2.5	35.3	41.9	−6.6
D	43.0	30.6	30.5	0.1	29.7	29.8	−0.1	28.4	28.1	0.3	35.0	34.7	0.3	29.1	32.0	−2.9	41.8	43.6	−1.8	36.7	43.4	−6.7
E	40.3	27.8	27.5	0.3	27.3	27.1	0.2	26.1	25.4	0.7	32.7	32.0	0.7	26.8	29.3	−2.5	38.6	40.9	−2.3	34.3	40.7	−6.4
F	43.8	29.7	31.0	−1.3	29.7	30.6	−0.9	28.5	28.9	−0.4	34.9	35.5	−0.6	28.2	32.8	−4.6	40.8	44.4	−3.6	35.7	44.2	−8.5
G	40.6	28.5	27.8	0.7	29.2	27.4	1.8	28.0	25.7	2.3	34.6	32.3	2.3	28.8	29.6	−0.8	40.8	41.2	−0.4	36.5	41.0	−4.5
Error: Average				0.2			0.3			0.7			0.7			−2.6			−2.1			−6.5
σ				0.9			0.9			0.9			1.0			1.2			1.1			1.2
Test System				−0.2, 0.5			−1.2			6.6			−5.8, 0.3						−4.5, 0.4			
Difference Distributions													6.1, 0.3									
																7.5, 0.2						
Reference System				−0.2, 0.3			−1.7			6.6			−2.7						−0.2			
Difference Distributions													8.9									
																11.4						

on an orthotelephonic basis, is sharply bandlimited to 250–4000 Hz, and has 0 dB of loss within this band. For example, test network 125 Hz HPF (sensation level = 100 dB) of Table V has an L_s = 107.1 − 0.2 = 106.9 dBt. The loudness loss is then 89.4 − 106.9 = −17.5 dB.

Entries in columns 4 through 9 were computed using the graphical form of Fig. 22. Column 4 entries are essentially a repeat of the column 3 entries, the difference being, as noted earlier, that the former were computed using a linear loudness law while the latter were computed using the loudness law of Figs. 6 and 8. Column 5 entries were computed without applying spread-of-loudness corrections.

The reference spectrum for computing the values given in columns 3, 4, and 5 was bandlimited to 250–3000 Hz. For example, entries of column 5 were obtained by plotting each of the responses on the form of Fig. 22, then measuring the area (square inches) enclosed by this response curve, the base line (designated 0 inches), and the left-hand and right-hand boundaries (designated 100 Hz and 5000 Hz respectively). This area was then divided by the base (inches) corresponding to the 250–4000-Hz bandwidth, and the equivalent height (inches) converted to loudness rating (dB) using the left-hand and right-hand ordinate scales. Entries of column 6 were obtained in the same way as those of column 5 except that the area was divided by the base (inches) corresponding to a 100–5000-Hz bandwidth. Thus, the change between columns 5 and 6 was simply one of reference spectrum bandwidth.

Entries of column 7 were obtained in the same manner as those of column 6 except that the area measured was restricted to the 300–3300-Hz bandwidth characteristic of EARS, and the reference spectrum was limited to the 300–3300-Hz band. Entries of columns 8 and 9 were obtained in the same manner as were entries of columns 6 and 7 respectively except that artificial voice/6-cm^3 coupler responses were used for the former, orthotelephonic responses for the latter.

The entries of columns 9a and 9b were computed in a manner which reflects the essential features of the EARS and, therefore, these entries closely approximate (within the bounds of computational and measurement error) what would be measured with the EARS on those test and reference systems utilizing linear transducers. However, column 9a and 9b entries of Tables XVII and XVIII and column 9a entries of Table XIX do not necessarily represent what would be

measured with the EARS. The reason for this is that the systems for these cases utilized carbon transmitters which are nonlinear devices. Transmitter responses used in the computations, based on measurements made with real speech, would probably differ from the responses pertaining during an EARS measurement because of the highly different nature of speech and the EARS acoustic test signal. This matter is now under study.

The error distributions corresponding to each of the seven computation methods exemplified by the column 3 through 9 entries are summarized in Table XX. These distributions reflect the accuracy of the computed ratings in predicting reference channel setting for equal speech loudness. The entries of column 1 are repeated from Table IV. These show that the computational method provides an accurate means of predicting subjective loudness balances for all of the tests except the Steinberg linear system tests at a sensation level = 39 dB. As noted earlier, received speech levels in a well-engineered communications system will seldom be this low, and if they do occur, are likely to represent trouble conditions.

Comparing the entries of column 2 to those of column 1, we note that there is little change in the error distributions, indicating that the graphical method is a close approximation of the computational method. Exceptions to this are the results for the Steinberg tests, particularly at the lower sensation level where the average error and error standard deviation are somewhat greater than when using the computation form. The reason for this is the change from the loudness law of Figs. 6 and 8 to a linear loudness law.

Column 3 entries indicate a further reduction in accuracy due to neglecting spread-of-loudness. The greatest change occurs for the Steinberg tests which involved numerous filter conditions and the Van Wynen telephone set tests which involved transducers with pronounced resonances. Both of these represent instances where spread-of-loudness effects would be important. Somewhat less change occurs for the Van Wynen linear system tests which also involved filter conditions.

The remaining columns show errors which might be expected using the EARS as presently defined (column 7), and the EARS modified to incorporate a wider measuring band (column 6). For the most part, extending the band improves the accuracy appreciably with the exception of the Steinberg tests and the Van Wynen telephone set tests. As regards the former, examination of Table XV indicates that extend-

TABLE XX—ERROR DISTRIBUTIONS

Test Series Designation	(1)		(2)		(3)		(4)		(5)		(6)		(7)	
	Computation Form (Fig. 14)		GRAPHICAL COMPUTATION (Fig. 22)											
			Orthotelephonic Responses								Artificial Voice/6-cm³ Coupler Responses			
			Same as Computation Form Except											
			Linear Loudness Law		Linear Loudness Law, No Spread-of-Loudness		Analysis Band = Reference Band = 100–5000 Hz		Analysis Band = Reference Band = 300–3300 Hz		Analysis Band = Reference Band = 100–5000 Hz		Analysis Band = Reference Band = 300–3300 Hz	
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ
Steinberg Linear System														
Tests–Sensation Level = 100 dB	1.4	1.1	2.4	1.0	3.6	1.9	3.7	1.9	2.5	3.6	3.8	1.7	2.4	4.3
Sensation Level = 39 dB	4.0	2.2	5.7	3.9	7.0	4.6	7.8	4.6	6.0	5.3	7.1	3.9	5.7	4.6
Van Wynen Linear System Tests	−0.8	2.1	−0.8	2.0	−0.6	2.6	−0.6	2.4	−3.5	3.1	0.1	2.9	−3.3	3.6
Van Wynen Telephone Set Tests	0.3	1.1	0.1	0.8	2.7	0.5	2.8	0.7	1.6	0.7	5.2	1.2	2.5	1.0
Van Wynen Tie Line Tests	−0.4	0.9	−0.1	1.1	0.2	1.0	0	1.0	0.3	1.4	0.1	0.9	0.8	1.2
Fraser Tests	0.2	0.9	0.3	0.9	0.7	0.9	0.7	1.0	−2.6	1.2	−2.1	1.1	−6.5	1.2

ing the band substantially improves the accuracy for the low-pass filter conditions but decreases the accuracy for the high-pass filter conditions.

The entries of columns 4 and 5, in a sense counterparts of columns 6 and 7 respectively, indicate what might happen if an EARS were built to utilize an artificial voice and an artificial ear which were accurate simulations of the human voice and ear within the context of the orthotelephonic definitions (see Fig. 3), and the system were calibrated in conformance with these definitions. Comparing columns 4 and 6, and 5 and 7, we see that the accuracy improves for the Van Wynen telephone set tests and the Fraser tests, remaining about the same for the other tests.

Let us now direct our attention to the difference distributions of Tables XV through XIX. These are given at the bottom of the tables together with arrows showing the computational methods compared. To obtain the difference distributions, differences were obtained between numerical entries, condition by condition. The first number associated with an arrow represents the average difference, a positive number signifying that ratings entered in columns at the right-hand tip of the arrow are numerically larger than those at the left-hand tip of the arrow. The second number, if given, represents the standard deviation of the difference distribution. In many cases, a standard deviation is not given because it was found to be of the order of 0.1 dB, insignificant for present purposes.

The difference distributions referred to above are summarized in Table XXI. Note that there is a set of entries for each different transmitter-receiver combination excepting the Steinberg tests for which entries are given for each of the two sensation levels.

Column 1 of Table XXI indicates that the graphical method results and computation method results differ very little except, as was noted in discussion of Table XX, in the case of the Steinberg tests with the sensation level = 39 dB. Differences in graphical ratings with and without spread-of-loudness included are also seen to be relatively small from entries of column 2. Entries of column 3 show that the change in ratings due to changing the reference spectrum bandwidth from 250–4000 Hz to 100–5000 Hz is essentially a constant, although some slight variation from test to test is apparent.

The remaining columns, 4 through 7, show that (*i*) measuring system bandwidth differences and (*ii*) orthotelephonic versus artificial voice/6-cm^3 coupler differences are highly dependent on transmitter-receiver

TABLE XXI—DIFFERENCE DISTRIBUTIONS

Test Series Designation	(1) Graphical Loudness Rating Minus Computation Form Loudness Rating		(2) Linear Loudness Law Without Spread-of-Loudness Minus with Spread-of-Loudness		(3) Reference Spectrum 100–5000 Hz Minus Reference Spectrum 250–4000 Hz		(4) 300–3300 Hz Minus 100–5000 Hz		(5)		(6) Artificial Voice 6-cm³ Coupler Responses Minus Orthotelephonic Responses		(7)	
							Ortho-telephonic Responses		Artificial Voice/6-cm³ Coupler Responses		100–5000 Hz		300–3300 Hz	
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ
Steinberg Linear System Tests														
Sensation Level = 100 dB	0.2	0.5	−0.5	1.0	6.6	—	−3.3	1.8	−1.0	3.1	8.7	1.0	11.0	2.1
Sensation Level = 39 dB	−2.5	2.5	−0.3	1.0	6.6	—	−3.4	2.2	−1.1	3.3	8.7	1.0	11.0	2.1
Van Wynen Linear System Tests	0	0.5	−1.1	0.7	6.6	—	−4.4	3.0	−3.0	3.9	8.1	0.9	9.4	0.8
Van Wynen Telephone Set Tests														
Test Telephone Set A	−1.0	0.5	−0.2	—	6.1	—	−8.0	—	−7.5	—	7.7	—	8.2	—
Test Telephone Set B	−0.6	0.5	1.0	—	7.0	—	−7.0	0.5	−6.6	0.5	8.8	—	9.1	—
Reference Telephone Set	−0.5	0.3	−2.0	—	6.6	—	−6.1	0.7	−4.3	1.5	6.1	—	7.9	0.6
Van Wynen Tie Line Tests	−0.8	0.6	−1.4	0.5	6.5	—	−7.7	0.6	−7.1	0.7	6.6	0.5	7.4	0.5
Fraser Tests														
Test System	−0.2	0.5	−1.2	—	6.6	—	−5.8	0.3	−4.5	0.4	6.1	0.3	7.5	0.2
Reference System	−0.2	0.3	−1.7	—	6.6	—	−2.7	—	−0.2	—	8.9	—	11.4	—

combinations. These differences appear to fall into two categories: (*i*) the Steinberg and Van Wynen linear system tests and the reference system of the Fraser tests for all of which linear, but different type, transducers were used; (*ii*) the Van Wynen telephone set and tie line tests and the test system of the Fraser tests for all of which nonlinear transducers were used. The transmitter and receiver types were identical (although the specific transducers were different) for the reference telephone set of the Van Wynen telephone set tests, the Van Wynen tie line tests, and the test system from the Fraser tests.

That the difference distribution standard deviations are relatively small and there seems to be a strong dependence of average difference on transmitter-receiver combination, suggests the possibility of applying correction factors to ratings computed by one method, e.g., the EARS method, to obtain ratings for some other method, e.g., the computational method. The average differences summarized in Table XXI represent these correction factors.

Returning to Table XX, we note rather large errors in some cases with the EARS method (see column 7) while the errors for the computation method are rather small (see column 1). Obviously, application of correction factors to the Steinberg, Van Wynen linear system, and Van Wynen tie line tests will not improve the EARS method error distributions since in each of the tests, the particular transmitter-receiver combination was common to the test and reference channels.

Different transducers were used in the test and reference channels of the Van Wynen telephone set tests and the Fraser tests. Let us examine the effect of application of correction factors to column 9 entries of Tables XVII and XIX on error distributions. The correction factors can be obtained from Table XXI. When the appropriate entries of column 9 of Tables XVII and XIX are so converted, the error distribution for the Van Wynen telephone set tests is changed from one with average = 2.5 dB, σ = 1.0 dB to average = 0.3 dB, σ = 1.5 dB comparing favorably to the column 3 values (Table XVII). Similar values for the Fraser tests are average = −6.5 dB, σ = 1.2 dB before correction, average = 0 dB, σ = 1.3 dB after correction, comparing favorably to the column 3 values from Table XIX for which average = 0.2 dB, σ = 0.9 dB.

Review of preceding discussion and the information contained in Tables XV through XXI indicates that the EARS provides a simple and reasonably accurate method of measuring telephone connection

loudness loss. For a telephone plant utilizing telephone sets of a single design, or of several similar designs, loudness ratings determined following the EARS procedure can provide an effective tool in telephone transmission engineering. In situations where the plant utilizes somewhat different telephone set designs, it appears that reasonably good design can result from using the EARS method to determine performance with the individual set designs. Comparisons between designs may then be made by application of correction factors. Determination of these correction factors can be a time consuming and difficult task as is evident from consideration of the orthotelephonic and artificial voice/6-cm^3 coupler response definitions of Figs. 3 and 16 respectively, but it is worth noting that such correction factors need be determined only once for each transmitter-receiver combination mounted in a particular handset configuration.

VI. ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to the many people who have participated in the evolution of the loudness computation and measurements methods which have been described in this paper. Many of the contributors to the concepts discussed herein have already been named in the text, and many not so named appear as authors of references.

Particular mention should be made of P. V. Dimock, W. Koenig, and J. W. Emling who were largely responsible for the particular speech loudness computation method discussed in Section III and O. H. Coolidge and H. Kahl who translated the computation method into a laboratory measuring system. Mr. Koenig also contributed substantially to the design of the subjective tests reported in Sections 3.5.2, 3.5.3, and 3.5.4. In addition, the author received considerable encouragement and advice from F. T. Andrews, J. W. Emling, M. D. Fagen, R. W. Hatch, G. A. Ledbetter, R. C. Lummis, and P. D. Bricker in preparation of the present paper.

REFERENCES

1. Huntley, H. R., "Where We Are and Where We Are Going in Telephone Transmission," Trans. Amer. Inst. Elec. Eng., Part 1, Commun. Elec., *76,* March 1957, pp. 54–63.
2. MacAdam, W. A., "A Basis for Transmission Performance Objectives in a Telephone Communication System," Trans. Amer. Inst. Elec. Eng., Part 1, Commun. Elec., *77,* May 1958, pp. 205–209.
3. Blye, P. W., Coolidge, O. H., and Huntley, H. R., "A Revised Telephone Transmission Rating Plan," B.S.T.J., *34,* No. 3 (May 1955), pp. 453–472.

4. Pamm, L. R., "Unigauge—A New Subscriber Loop System," Bell Lab. Rec., *45,* No. 8 (September 1967), pp. 261–264.
5. Andrews, F. T., Jr., and Hatch, R. W., "National Telephone Network Transmission Planning in the American Telephone and Telegraph Company," IEEE Trans. Commun. Theor., *COM-19,* No. 3 (June 1971), pp. 302–315.
6. Volume V**, CCITT White Book, *Telephone Transmission Quality, Local Networks and Telephone Sets,* Published by the International Telecommunication Union, 1969.
7. Martin, W. H., "The Transmission Unit and Telephone Transmission Reference Systems," B.S.T.J., *3,* No. 3 (July 1924), pp. 400–408.
8. Sivian, L. J., "A Telephone Transmission Reference System," Elec. Commun., *3,* No. 2 (October 1924), pp. 114–126.
9. Martin, W. H., and Gray, C. H. G., "Master Reference System for Telephone Transmission," B.S.T.J., *8,* No. 3 (July 1929), pp. 536–559.
10. Hartley, R. V. L., "The Transmission Unit," Elec. Commun., *3,* No. 1 (July 1924), pp. 34–42.
11. Martin, W. H., "Decibel—The Name for the Transmission Unit," B.S.T.J., *8,* No. 1 (January 1929), pp. 1–2.
12. Volume IV, CCIF Green Book, *Recommendations and Measurements Concerning Transmission Quality of Telephone Apparatus,* Published by the International Telecommunication Union, Geneva, 1956.
13. Martin, W. H., "Rating the Transmission Performance of Telephone Circuits," B.S.T.J., *10,* No. 1 (January 1931), pp. 116–131.
14. McKown, F. W., and Emling, J. W., "A System of Effective Transmission Data for Rating Telephone Circuits," B.S.T.J., *12,* No. 3 (July 1933), pp. 331–346.
15. Jones, W. C., and Inglis, A. H., "Development of a Handset for Telephone Stations," B.S.T.J., *11,* No. 2 (April 1932), pp. 245–263.
16. Inglis, A. H., "Transmission Features of the New Telephone Sets," B.S.T.J., *17,* No. 3 (July 1938), pp. 358–380.
17. Jones, W. C., "Instruments for the New Telephone Sets," B.S.T.J., *17,* No. 3 (July 1938), pp. 338–357.
18. Inglis, A. H., and Tuffnell, W. L., "An Improved Telephone Set," B.S.T.J., *30,* No. 2 (April 1951), pp. 239–270.
19. Bennett, A. F., "An Improved Circuit for the Telephone Set," B.S.T.J., *32,* No. 3 (May 1953), pp. 611–626.
20. Cochran, W. T., and Lewinski, D. A., "A New Measuring Set for Message Circuit Noise," B.S.T.J., *39,* No. 4 (July 1960), pp. 911–931.
21. "Acoustical Terminology (including Mechanical Shock and Vibration)," American Standard S1.1, 1960.
22. Fletcher, H., and Munson, W. A., "Loudness, Its Definition, Measurement and Calculation," B.S.T.J., *12,* No. 4 (October 1933), pp. 337–430. Also see J. Acoust. Soc. Amer., *5,* No. 2 (October 1933), pp. 82–108.
23. Fletcher, H., and Munson, W. A., "Relation Between Loudness and Masking," J. Acoust. Soc. Amer., *9,* July 1937, pp. 1–10.
24. Beranek, L. L., Marshall, J. L., and Cudworth, A. L., "Calculation and Measurement of the Loudness of Sounds," J. Acoust. Soc. Amer., *23,* No. 3 (May 1951), pp. 261–269.
25. Garner, W. R., "On the Lambda Loudness Function, Masking, and the Loudness of Multicomponent Tones," J. Acoust. Soc. Amer., *26,* No. 5 (May 1954), pp. 602–607.
26. Stevens, S. S., "Procedure for Calculating Loudness: Mark VI," J. Acoust. Soc. Amer., *33,* No. 11 (November 1961), pp. 1577–1585.
27. Fletcher, H., *Speech and Hearing in Communication,* New York: D. Van Nostrand Company, Inc., 1953.
28. Fletcher, H., and Steinberg, J. C., "The Dependence of the Loudness of a Complex Sound Upon the Energy in the Various Frequency Regions of the Sounds," Phys. Rev., 2nd series, *24,* September 1924, pp. 306–317.
29. Steinberg, J. C., "The Relation Between the Loudness of a Sound and Its Physical Stimulus," Phys. Rev., 2nd series, *26,* October 1925, pp. 507–523.

30. Pollack, I., "The Effect of White Noise on the Loudness of Speech of Assigned Average Level," J. Acoust. Soc. Amer., *21,* No. 3 (May 1949), pp. 255–258.
31. Fletcher, H., and Galt, R. H., "The Perception of Speech and Its Relation to Telephony," J. Acoust. Soc. Amer., *22,* No. 2 (March 1950), pp. 89–151.
32. Pollack, I., "On the Measurement of the Loudness of Speech," J. Acoust. Soc. Amer., *24,* No. 3 (May 1952), pp. 323–324.
33. Lane, H. L., Catania, A. C., and Stevens, S. S., "Voice Level: Autophonic Scale, Perceived Loudness and Effects of Sidetone," J. Acoust. Soc. Amer., *33,* No. 2 (February 1961), pp. 160–168.
34. Geiger, P. H., "Loudness Level to Loudness Conversion Chart," J. Acoust. Soc. Amer., *11,* No. 3 (January 1940), pp. 308–310.
35. "American Standard for Noise Measurement," Z24.2, 1942, in J. Acoust. Soc. Amer., *14,* No. 1 (July 1942), pp. 102–110.
36. Stevens, S. S., "The Measurement of Loudness," J. Acoust. Soc. Amer., *27,* No. 5 (September 1955), pp. 815–829.
37. Stevens, S. S., "Calculation of the Loudness of Complex Noise," J. Acoust. Soc. Amer., *28,* No. 5 (September 1956), pp. 807–832.
38. "Procedure for the Computation of the Loudness of Noise," USA Standard, S3.4, 1968.
39. French, N. R., and Steinberg, J. C., "Factors Governing the Intelligibility of Speech Sounds," J. Acoust. Soc. Amer., *19,* January 1947, pp. 90–119.
40. Dunn, H. K., and White, S. D., "Statistical Measurements on Conversational Speech," J. Acoust. Soc. Amer., *11,* No. 3 (January 1940), pp. 278–288.
41. Stevens, S. S., Egan, J. P., and Miller, G. A., "Methods of Measuring Speech Spectra," J. Acoust. Soc. Amer., *19,* No. 5 (September 1947), pp. 771–780.
42. Rudmose, H. W., Clark, K. C., Carlson, F. D., Eisenstein, J. C., and Walker, R. A., "Voice Measurements with an Audio Spectometer," J. Acoust. Soc. Amer., *20,* No. 4 (July 1948), pp. 503–512.
43. Benson, R. W., and Hirsh, I. J., "Some Variables in Audio Spectrometry," J. Acoust. Soc. Amer., *25,* No. 3 (May 1953), pp. 499–505.
44. Dunn, H. K., and Farnsworth, D. W., "Exploration of Pressure Field Around the Human Head During Speech," J. Acoust. Soc. Amer., *10,* No. 3 (January 1939), pp. 184–199.
45. Sivian, L. J., "Speech Power and Its Measurement," B.S.T.J., *2,* No. 4 (October 1929), pp. 646–661.
46. Steinberg, J. C., and Munson, W. A., "Deviations in the Loudness Judgments of 1000 People," J. Acoust. Soc. Amer., *8,* No. 2 (October 1936), pp. 71–80.
47. Bricker, P. D., "Technique for Objective Measurement of Speech Levels," J. Acoust. Soc. Amer., *38,* No. 2 (June 1965), pp. 361–362.
48. Fletcher, H., "The Nature of Speech and Its Interpretation," J. Franklin Inst., *193,* No. 6 (June 1922), pp. 729–747.
49. Noll, A. M., "Subjective Effects of Sidetone During Telephone Conversation," IEEE Trans. Commun. Elec., *83,* No. 72 (May 1964), pp. 228–231.
50. Noll, A. M., "Effects of Head and Air-Leakage Sidetone During Monaural-Telephone Speaking," J. Acoust. Soc. Amer., *36,* No. 3 (March 1964), pp. 598–599.
51. Braun, K., and Koschel, H., "Reference Equivalent Measuring Set with Direct Indication and Its Importance for Improving Telephone Transmission," Fernmeldetech. Z., *5,* No. 10, 1952, pp. 447–455.
52. Inglis, A. H., Gray, C. H. G., and Jenkins, R. T., "A Voice and Ear for Telephone Measurements," B.S.T.J., *11,* No. 2 (April 1932), pp. 293–317.
53. "Specifications for Laboratory Standard Microphones," USA Standard S1.12, 1967.
54. "Proposed Method for Measuring Transmission Performance of Telephone Sets," IEEE No. 269, May 1966, (Published for Trial Use).
55. "American Standard Method for Coupler Calibration of Earphones," Z24.9, 1949.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

The Transmission Performance of Bell System Toll Connecting Trunks

By J. E. KESSLER

(Manuscript received April 19, 1971)

A systemwide survey of the performance of Bell System toll connecting trunks was undertaken in 1966. Various results of this survey have been used in other studies since that time. This paper collects the main results in a single writing. The sampling plan and measurement procedures are discussed briefly. Measurement results are presented in distributional form and as estimates for the transmission parameters relating to loss, message circuit noise, impulse noise, relative envelope delay, P/AR, harmonic distortion, and physical characteristics. Some results are presented separately for major facilities and facility categories.

I. INTRODUCTION

The Bell System telephone network consists of a hierarchy of transmission facilities available for connecting one subscriber to another. These include the loops from the subscribers to their local telephone office, toll connecting trunks from the local to the toll office, and intertoll trunks between these toll offices. Clearly, the transmission performance of a toll connection is affected by each part of this built-up connection. Overall measures of performance have been reported in previous connection surveys[1,2] and there has been a recently completed 1969–70 Connection Survey.[3–5] There is a concurrent need for detailed knowledge of the specific parts of toll connections. This information finds important applications in the setting of trunk objectives for the various transmission parameters, and in the effort to simulate, on a computer, a model of the Bell System transmission network. To this end, this paper reports various population estimates of the transmission performance of toll connecting trunks, based on a systemwide survey undertaken in 1966. Individual portions of this information have been used separately, and this paper serves to bring all the results together in a single document. To the author's knowledge, these data represent

the only published estimates of toll connecting trunk performance. Although the mixture of facilities has changed with time, estimates of parameters for individual types of facility are expected to be reasonably stable.

A toll connecting trunk is defined as any trunk connecting an end office (class 5) with a toll office (class 4, 3, 2, or 1). For purposes of this paper each toll connecting trunk consists of the trunk facility and any office equipment associated with the trunk. The office equipment may be attenuation pads, hybrid transformers, repeating coils, etc. Figure 1 shows the various types of toll connecting trunks used in the Bell System.

In general, a toll connecting trunk consists of switching equipment at both ends and transmission facilities in between. Consistent with their application to trunk transmission objectives studies, trunks are measured on a switch-through-switch basis (see Fig. 2). This enables one to use these data in constructing estimates for overall connections. In toll offices measurements are made from a toll testboard, master test frame, or an outgoing trunk (OGT) board. In the end offices the testing is done from a master test frame or OGT board. Thus, the trunk includes (besides the actual facility of two-wire or four-wire voice-frequency cable, or carrier channel) the trunk circuit, repeating coils, four-wire terminating sets, and switching equipment, etc.

The toll connecting trunks in the survey were measured for the following transmission parameters:

(*i*) 1000-Hz loss
(*ii*) Frequency response

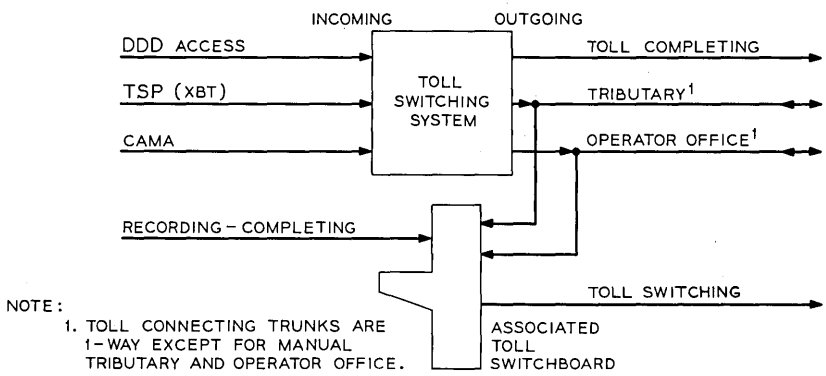


Fig. 1—Types of toll connecting trunks.

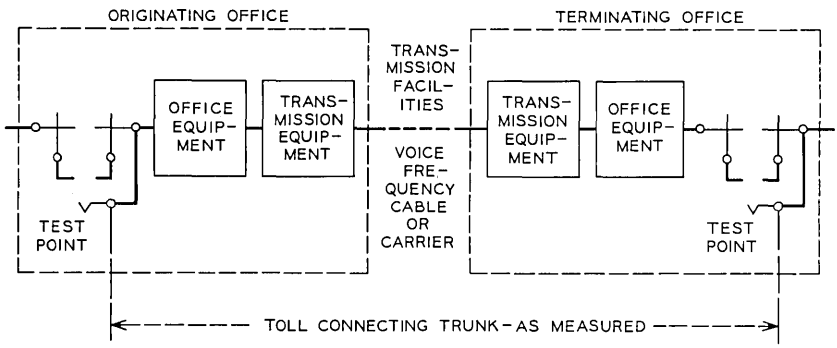


Fig. 2—Composition and testing arrangement.

(*iii*) Relative envelope delay
(*iv*) Message circuit noise
(*v*) Impulse noise
(*vi*) P/AR meter readings
(*vii*) Harmonic distortion
(*viii*) Level tracking on compandored facilities.

All measurements were made in both directions of transmission with the exception of impulse noise. It was measured only incoming to the class 5 office. The frequency response and relative envelope delay were measured at 17 frequencies from 200 to 3400 Hz. Bell Laboratories personnel carried out these measurements, with one person at each end of the toll connecting trunk.

In addition to actual measurements, information about trunk length, type of switching, and type of facility used for the toll connecting trunk was obtained from office records, such as circuit layout cards.

There are many different facilities in use as toll connecting trunks in the Bell System. This survey encountered several separably identifiable types of facilities and, where there was a significant sample, results are given for these facility subclasses.

II. SAMPLING PLAN

The sampled population consisted of all the Bell System toll connecting trunks in the continental United States and Canada. Although the exact number was not known, it was estimated that there were about 800,000 toll connecting trunks in service in 1966.

The sample of toll connecting trunks of this survey consisted of

150 trunks from 15 end offices, each having less than 400,000 annual outgoing toll messages (AOTM), and 242 trunks from 25 end offices each having more than 400,000 AOTM. These 392 trunks were measured for most parameters in both directions, thus yielding 784 population elements in the sample.

The sample design consisted of a two-stage plan with first-stage stratification and with the primary units selected with probabilities proportional to a measure of size. (See Refs. 6, 7, or 8 for a general discussion of sampling plans of this type.) The primary units were Bell System end office (class 5) buildings. These end offices were stratified into two disjoint subpopulations, according to the number of annual outgoing toll messages originating in the end office. This was done because a pilot survey in 1965 indicated that large end office buildings usually have shorter toll connecting trunks than small end office buildings. The large end offices also tend to use voice-frequency cable (VF) toll connecting trunks rather than carrier. By stratifying the population into two groups that tend to internal homogeneity one can more efficiently procure the necessary estimates with the desired precision. The confirmation of this rationale will be borne out in a later section discussing the physical characteristics of toll connecting trunks.

The frame for the first-stage sample was a list of the 9052 Bell System end office buildings in service on January 1, 1964. These end office buildings were divided into two subpopulations, according to whether they originated 400,000 AOTM or not. From stratum 1 of end office buildings with less than 400,000 AOTM, 15 end office buildings were selected. From stratum 2, twenty-five end office buildings were selected. Each end office building was assigned a probability of selection proportional to its AOTM. This measure of size had been used for the first-stage selection of offices in the concurrently conducted toll connection survey,[2] and practicality deemed its use again for the toll connecting trunk survey. There is a strong correlation between the AOTM and the number of toll connecting trunks in an end office building. The sampling of end office buildings in each primary stratum was performed independently.

For the second stage of sampling a listing of all toll connecting trunks in each primary unit (i.e., the selected end office building) was obtained. For controlling the complexity of the survey, a three-day visit was planned for each office. This practical constraint indicated the ability

to measure 8 to 12 trunks in each office. Therefore, where possible, 12 toll connecting trunks were chosen by simple random sampling, independently in each primary unit. The toll connecting trunks thus selected become the second-stage units of the sample design.

It is possible for an end office to have trunks to more than one toll office building, requiring the tester at the toll office end of the trunks to travel to more than one building. The goal of testing 8 to 12 trunks from each end office was met except in three offices. In one office, only seven toll connecting trunks existed, and in the other two offices practical difficulties arose to limit the measurements to seven and five trunks.

The data analysis consisted of using estimation formulas appropriate for the sample design described previously. These formulas yield estimates of population parameters (i.e., mean, variance, standard deviation, and 90-percent confidence intervals) based on the sample data. The estimates are weighted to account for the structure of the sampling plan, with its stratification and selection proportional to a measure of size. Probability sampling of this type has the distinct advantage of providing estimates of the important characteristics of the population from which we sample, and at the same time providing a quantitative measure of the sampling error.

III. PHYSICAL CHARACTERISTICS

For each of the toll connecting trunks of the sample a record was obtained of some of its physical characteristics. This included the route length of the trunk, the type of switching at each end of the trunk, and the type of facility comprising the trunk. From these data, estimates were made of trunk length distributions, facility composition of toll connecting trunks, and percentages of different switching machines used.

3.1 *Trunk Length*

The trunk length distribution is shown in Fig. 3. A substantial amount (21 percent) of the trunks are intrabuilding trunks resulting from the end office and toll office being in the same building. In these cases the toll connecting trunk usually consists of a 2-dB pad and switching equipment in tandem with a short length of cable.

Basically, toll connecting trunks span short distances. About half the trunks are less than 10 miles long, and only 10 percent of the trunks

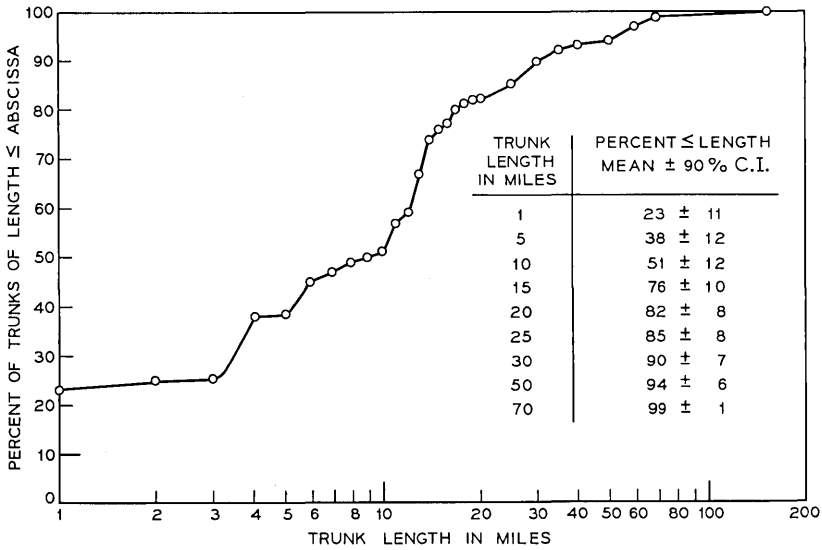


Fig. 3—Trunk length distribution.

are longer than 30 miles. However, there were two trunks in the sample that were 151 miles long. This resulted from indirect geographic routing which added about 100 miles to their actual length.

As previously mentioned, the survey sample had a primary stratification based on the belief that large and small offices were significantly different with regard to the types and length of the toll connecting trunks associated with them. These differences are corroborated in Table I which shows the estimated length and type of facility in large and small offices. The metropolitan offices (i.e., large, over 400,000 AOTMs) tend to have toll connecting trunks much shorter than those in small offices, and they utilize voice-frequency cable to a much greater degree than the small offices.

A finer relationship between trunk length and type of facility is shown in Table II, in terms of three specific length intervals, 0–15, 15–30, and over 30 miles. The estimated percentage of each facility type used within each interval is given.

3.2 *Facility Composition*

The sample of 392 trunks contained many different facilities. Those types appearing sufficiently often to be estimated reliably are shown

in Table III. The percentages of both categories of facilities and single types of facilities are estimated in Table III.

3.3 *Switching Machines*

The toll connecting trunk encounters some type of switching at both the end office (class 5) and the toll office (class 4 or higher). In this survey we noted the type of switching machine used at each end of each toll connecting trunk in the sample. The numbers of the type of switching machine used are classified by end office, toll office, and stratum, and are given in Table IV. There were 40 end office buildings and 66 toll office buildings in the survey, with some office buildings having more than one type of switching system.

It should be noted that these numbers of switching machines are only the types of switching that were found with the toll connecting trunks of this survey. They should not be taken to accurately represent the distribution of switching machines in the Bell System, since the basic population element that was sampled in this survey was a toll connecting trunk and not a switching machine. Also, even though the end office buildings of the primary strata selections may be related to switching machines, the same can not be said of the toll offices. In many cases, all the trunks from an end office home on the same switching machine in the toll office, thus biasing any estimates at the toll office.

At the toll offices there is a sizable amount of switching designated as "switchboard." This category is composed of those trunks for which the transmission measurements were made at a switchboard. Therefore, this category includes toll switching, recording-completing, and operator office trunks (see Fig. 2). This category does not constitute manual switching, where all of a customer's calls to the toll office have to be handled by an operator at a switchboard. In all the toll offices associated with this classification there also was a switching machine. This type of switching machine could have been used as an alternative designation to the switchboard classification. (In fact, operator office trunks may have a multipled switching machine appearance, as well as the switchboard appearance.) However, since the information recorded was to reflect the type of switching actually used by the sampled toll connecting trunks of the survey, the switchboard designation was chosen.

In the class 5 end offices encountered in this survey about 54 percent of the trunks utilized step-by-step switching systems, and 34 percent

used No. 5 crossbar. These two types account for the bulk of the switching machines in use at the class 5 end of the sampled toll connecting trunks. At the toll office, crossbar tandem systems were used most frequently, followed by step-by-step systems and switchboards (with the conditions mentioned above defining switchboards).

IV. TRANSMISSION LOSS

4.1 *Frequency Response*

The loss of toll connecting trunks was measured at 17 frequencies, from 200 to 3400 Hz. The measurements were made switch-through-switch as shown in Fig. 1. At each frequency a tone at 0 dBm was transmitted over the trunk and measured at the receiving end with a 25A Gain and Delay Measuring Set. This procedure was repeated for each direction of transmission over the trunk. Since the design of the measurement equipment precluded measuring losses greater than 31 dB, it was not always possible to measure some trunks at the higher frequencies. This was the situation in particular on carrier facilities, such as N1, ON, and Lenkurt. For losses greater than 31 dB, the value of 31 dB was used. This provides a conservative estimate of the mean loss, but is more realistic than simply eliminating these readings. Also, this effect occurred only at 3300 and 3400 Hz, and only on some carrier facilities.

The frequency response statistics for all the toll connecting trunks in the survey are shown in Table V. The results show the mean and its associated 90-percent confidence interval, and an estimate of the standard deviation for each distribution that is considered. Table V also shows the loss estimates for the two types of facilities, carrier and voice-frequency cable, used by toll connecting trunks. It appears that the loss responses for both types of facilities are similar throughout most of the voiceband, with the carrier facilities having a sharper cutoff at the high and low frequencies.

4.2 *Loss Relative to 1000 Hz*

For the individual types of facilities it is useful to present the loss relative to 1000 Hz. This enables one to see the loss characteristics of the facility itself more readily than examination of the absolute value of loss does, since different loss design procedures are encountered among the many trunks. Table VI shows the loss relative to 1000 Hz for the more commonly used carrier facilities. Table VII shows the loss relative to 1000 Hz for the major VF facilities.

V. NOISE

The term noise is applied to a variety of electrical phenomena, all of which are unwanted additive signals and tend to interfere with the transmission of information. There are numerous sources of noise, some man-made and some not. Man-made noise is generated by power supply hum, faulty contacts, electrical apparatus, quantizing errors, etc. The noise from non-man-made sources can originate within or without the systems under measurement. All systems are prone to thermal noise from the random motion of electrons in conductors, as well as to atmospheric electrical disturbances.

The effects of these forms of noise depend, in part, on the type of signal being transmitted over the telephone communication channel. For analog speech signals the average power of the noise is of primary concern. Interference from many noise sources contribute to this type of noise, referred to as "message circuit noise." However, data signals suffer more severely from noise voltage peaks (whether of short or long time duration), which are classified as "impulse noise." A different procedure is used to measure these two noise parameters and each is reported separately in the following sections.

5.1 *Message Circuit Noise*

The level of background noise was measured using the 3A Noise Measuring Set (hence, the name 3A noise) with both C-message and 3-kHz flat weighting networks.[9] The noise measurements were performed for both directions of transmission on the trunk. The measurements were made during the normal business day.

A summary of the noise measured with C-message weighting and 3-kHz flat weighting is shown in Table VIII in terms of the mean and its 90-percent confidence interval, and standard deviation. Estimates for specific facilities and groups of facilities are included where there were sufficient data for useful statistics.

It can be observed that in all categories the 3-kHz flat noise level is higher than with C-message weighting, since the 3-kHz flat weighting does not attenuate the low frequencies nearly as much as C-message weighting. The flat noise on voice-frequency facilities was slightly higher than on carrier, probably because carrier systems have a higher cutoff frequency at the low end of the frequency band, below which nothing is transmitted.

The estimates of flat noise were significantly affected by the direction of transmission being observed. A subclass analysis based on this characteristic showed that average flat noise measured incoming to

the end office (class 5) was considerably higher than that incoming to the toll office end of the trunk. Estimated statistics for the parameter, δ, defined as the end office flat noise minus the toll office flat noise, are given in Table IX. The difference in noise levels is more pronounced on voice-frequency cable facilities than on carrier facilities because there is less low-frequency energy transmitted on carrier facilities, in either direction, due to the action of channel filters.

Since the end office noise level is higher it suggests that its noise environment is worse than that of the toll offices. This could be due to extraneous low-frequency components (primarily 60 Hz and its harmonics) generated by office battery and other power plants. Reasoning that larger offices have more noise-producing loads one would expect the larger end offices (strata 2) to have higher average noise levels than the small end offices (strata 1). This is borne out in Fig. 4, where a difference of about 6 dB is observed. There is further support for these hypotheses in comparing the toll office flat noise of toll connecting trunks with the toll office flat noise of short intertoll trunks[10] (0–62.5 miles). The result, with 90-percent confidence limits, is shown in Table X.

The data for 3A noise with C-message weighting were also analyzed

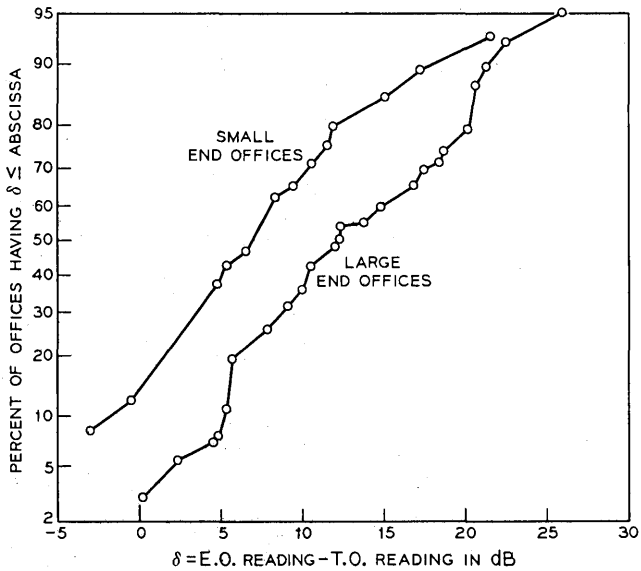


Fig. 4—3A noise with 3-kHz flat weighting.

with respect to trunk length. For this purpose six mileage bands were chosen, with each succeeding band being a doubling of the previous interval. The "double distance" criterion is chosen since previous studies[1,2,10] have shown a linear relationship between noise and double distance on toll connections. The toll connecting trunks of this survey are of relatively short length, but a similar linear relationship is evident, as shown in Table XI. Not only is there the effect of longer trunks accumulating more noise, but it should be recalled from Section III that the longer trunks are predominantly on carrier facilities. Both parameters, the length and facility type, contribute their effect to the noise level observed. The flat-weighted noise was independent of trunk length. This is still consistent with the past studies of intertoll trunks which showed only a very slight increase in flat noise with length for the shorter trunks.

5.2 *Background Noise on Compandored Trunks*

Since the syllabic compandor action on carrier facilities introduces a large loss to the noise in the absence of speech, a background noise reading with a quiet termination is not characteristic of a compandored channel in normal use (i.e., with speech present). In order to measure C-message weighted noise in a simulated speech condition, a holding tone was sent over the trunk in order to adjust the compandor gains, and then it was filtered out at the receiving end.

Measurements were made at four different levels of holding tone, in both directions of transmission on the trunk. The 1850-Hz tone was supplied by a Western Electric KS-19260 Pushbutton Oscillator, at levels of 0, −10, −20, and −30 dBm. After the tone was filtered out, the C-message weighted noise readings were made.

Among the compandored toll connecting trunks there were those using T1 carrier, a digital system.[11] This system uses instantaneous compandors which are active whether speech is present or not. The 1850-Hz holding tone introduced distortion products into the measured signal as the fundamental and harmonics of the 1850-Hz tone interacted with the sampling frequency of the T1 system. Contrary to analog systems, the noise on T1 is mostly quantizing noise introduced by the encoder in the terminal, there being relatively small amounts of line noise. For these reasons, the T1 data are treated separately from the analog data.

Table XII shows estimates of the mean, its 90-percent confidence interval, and the standard deviation of the C-message weighted noise as a function of input holding tone power level. A majority of the analog

compandored trunks were on N1 carrier, and their characteristics are also listed separately. For the analog carrier systems the noise components in the terminal equipment are low compared to the line noise. Hence, the measurements should reflect the expandor action at the near end where the measurements are made. Since a compression and expansion ratio of 2 : 1 is being used we would expect a 5-dB change in noise level for each 10-dB change in input signal level. For the analog facilities there is a ratio of about 10 : 7 between the 0- and −10-dBm level inputs, and a ratio of slightly less than 10 : 5 for the level inputs between −10 and −30 dBm. This compares to the theoretical 10 : 5 ratio for perfect companding action.

5.3 *Impulse Noise*

Impulse noise was measured on a switch-through-switch basis, incoming to the class 5 offices. The measurements were made with the 6F Voiceband Noise Measuring Set using C-message weighting. The four counters of the 6F Set were adjusted to cover the range of 4 to 45 counts in the 15-minute recording interval. Then, on each trunk measured, the recorded impulse counts and associated threshold levels are used to interpolate the levels that would correspond to 45, 15, and 4.5 counts per 15-minute interval. The difference in the two levels that yield 4.5 and 45 counts is defined as the slope (in dB per decade) of the peak amplitude distribution. (See Ref. 12 for a discussion of these parameters.) On compandored facilities an 1850-Hz tone at a level of −14 dBm was applied to the trunk in the toll office in order to simulate data transmission trunk conditions. At the end office the tone was suppressed by a 90-dB band-reject filter with a 50-Hz bandwidth.

The results presented in Table XIII are estimates of the mean and standard deviation of the 45- and 15-count level distributions, with a 90-percent confidence interval on the mean estimate. The same estimates are shown for the slope. These data all apply at the switch at the receive end of the trunk. The compandored carrier trunks have a higher mean impulse noise level, but lower average slope, than the voice-frequency cable facilities. A lower slope indicates a narrower distribution of peak amplitudes, so that a smaller change in the reference count level is required to produce a change by a factor of ten in the impulse count. Figure 5 illustrates these points. With the exception of the estimates for all toll connecting trunks, the data of Table XIII and Fig. 5 result from trunks of less than 60 miles length. There were 13 trunks longer than 60 miles, all of them on compandored carrier facilities.

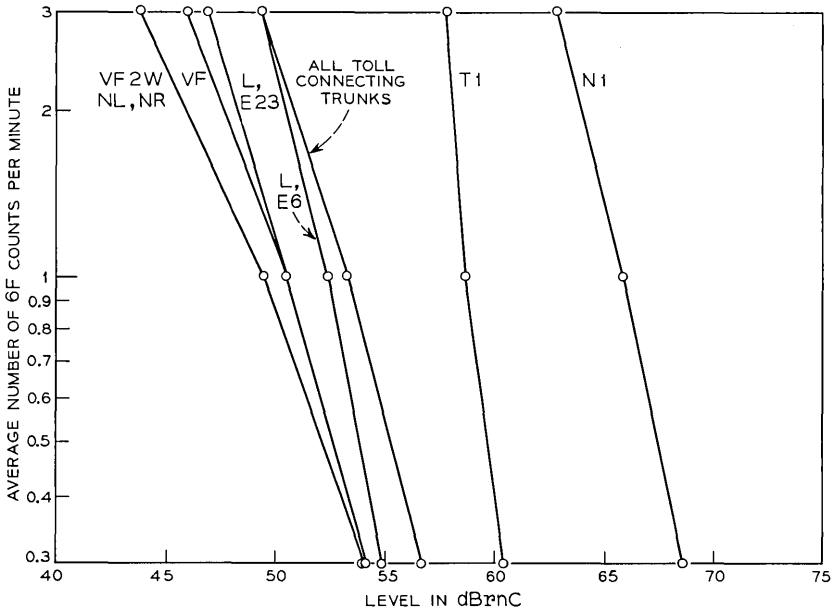


Fig. 5—Average impulse noise levels at receive switch.

VI. RELATIVE ENVELOPE DELAY

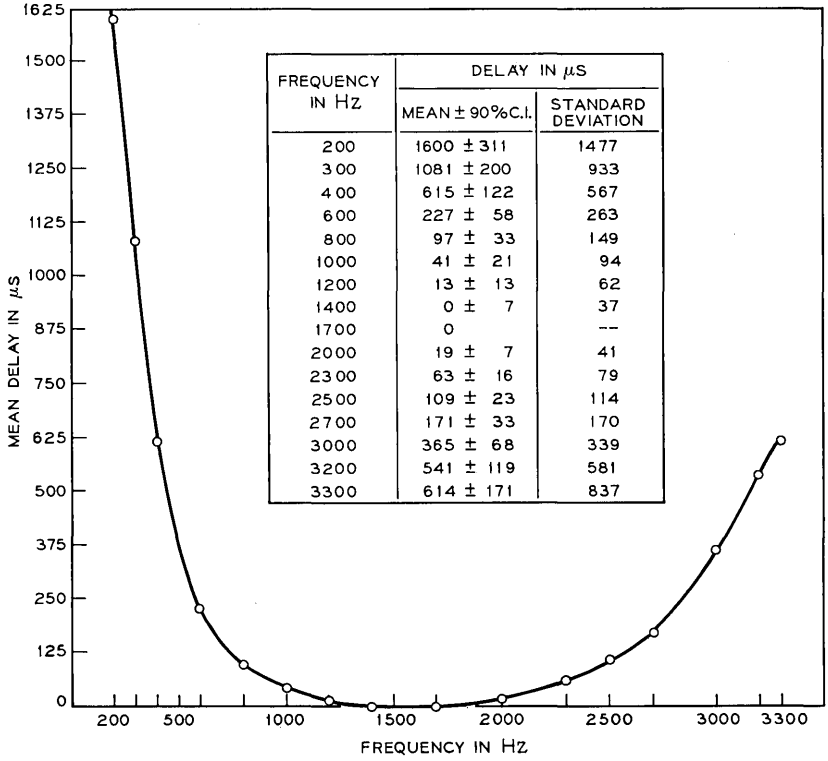
In the transmission of complex waveforms a nonlinear phase shift characteristic of the transmission medium results in the distortion of the signal due to the differing amounts of delay of the different frequencies present in the waveform. This delay distortion can be assessed by measuring the relative envelope delay, which is the delay of the envelope of a low-frequency amplitude-modulated voice-freqeuncy carrier relative to the envelope delay at a reference voice-frequency carrier. The relative envelope delay is an approximation to the derivative of the phase characteristic, so linear phase would yield constant envelope delay. In this survey all the envelope delay measurements are relative to 1700 Hz.

The measurements of relative envelope delay were made using the 25A Gain and Delay Measuring Set and KS-19260 Oscillator. A pair of these was required at each end of the toll connecting trunk, so as to measure both directions of transmission.

The relative envelope delay data were analyzed both for toll connecting trunks in general, and for various subclasses of interest. The

results show the estimated mean and its associated 90-percent confidence interval, and an estimate of the standard deviation of each distribution considered. Figure 6 presents the relative envelope delay characteristics based on the sample of all toll connecting trunks. This reflects the differing influences of different types of facilities. For example, the nonloaded voice-frequency cable facilities have essentially no relative delay at the higher frequencies, so the relative delay derives from the carrier and loaded voice-frequency cable facilities. Also, in all cases the office equipment associated with the transmission facility is a major contributor to the delay at low frequencies.[6]

The difference in the relative envelope delay between carrier and voice-frequency cable facilities is shown in Fig. 7. In general, the carrier facilities exhibit more relative envelope delay at the low and high frequency band-edges. The estimates for the carrier facilities are



FREQUENCY IN HZ	DELAY IN μS	
	MEAN ± 90%C.I.	STANDARD DEVIATION
200	1600 ± 311	1477
300	1081 ± 200	933
400	615 ± 122	567
600	227 ± 58	263
800	97 ± 33	149
1000	41 ± 21	94
1200	13 ± 13	62
1400	0 ± 7	37
1700	0	--
2000	19 ± 7	41
2300	63 ± 16	79
2500	109 ± 23	114
2700	171 ± 33	170
3000	365 ± 68	339
3200	541 ± 119	581
3300	614 ± 171	837

Fig. 6—Mean relative envelope delay (includes all types of facilities).

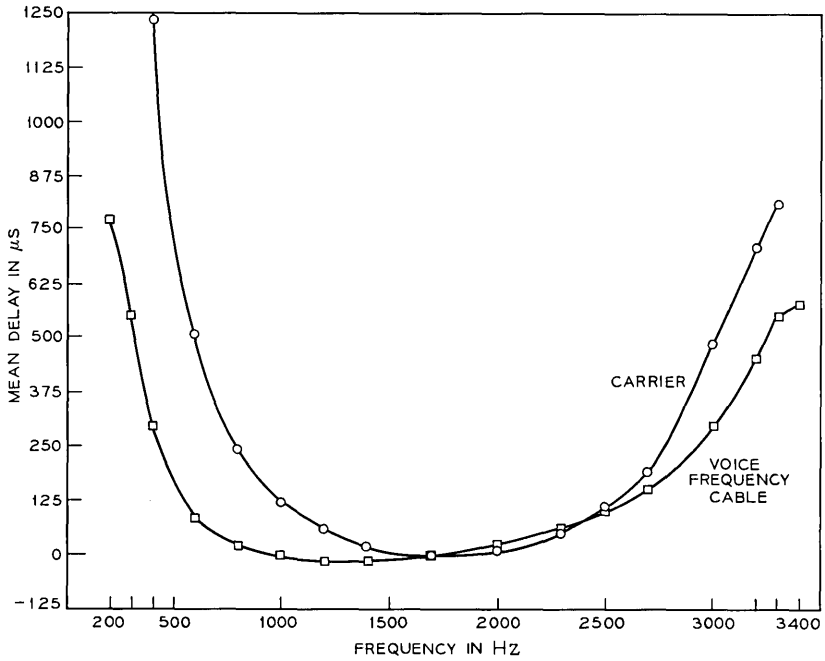


Fig. 7—Mean relative envelope delay (carrier and voice-frequency cable facilities).

only given up to 3200 Hz since excessive loss on the connection above this frequency precluded making the envelope delay measurements on the N1, ON, and O carrier facilities.

Tables XIV and XV show the estimates of relative envelope delay for specific facilities. These subclasses were considered since a sufficient number of trunks of each type was measured, and the trunks were from a number of different offices. The seven types of facilities tabulated represent the facilities more commonly used for toll connecting trunks and they comprised about 85 percent of the total sample for this survey. Among the carrier facilities the delay on T1 showed the least variability. The delay on N1 has been reported in the past[6] where N1 was being used as an intertoll trunk, and those findings are comparable to the values reported here, with N1 now used as a toll connecting trunk. Although all the voice-frequency cable facilities exhibit similar delay characteristics at low frequencies, there is a marked difference between loaded and nonloaded facilities at the higher frequencies. This accounts for the higher variances of the VF facilities in the higher frequencies,

with the loaded VF facilities contributing most of this variance. The estimates of the variance for VF trunks with E6 or E23 repeaters are larger than other categories due to the correlation between delay and trunk length—the longer trunks exhibiting more delay. For example, the E23 measurements came from eight different offices, of which four offices supplied shorter trunks (5.3, 5.7, 6.2, 6.7 miles) and four offices had longer trunks (10.2, 11.7, 12.8, 16.3 miles).

In considering toll connecting trunks in general, the correlation between trunk length and relative envelope delay is neither significant nor consistent. This is due to having many different facilities in the total sample and also to measuring the combined effects of office equipment and facility for each toll connecting trunk. Thus, there is only a slightly discernible relation between delay and trunk length at low frequencies (below 1700 Hz), since office equipment causes the predominant effect. Above 1700 Hz the "VF Loaded" facilities exhibited a significant length correlation. These facilities were for trunks of roughly 1 to 18 miles in length. A subclass analysis based on three mileage bands is shown in Fig. 8. The 90-percent confidence interval for each estimated mean value of delay (above 1700 Hz) is also shown.

VII. P/AR

The P/AR (Peak-to-Average Ratio) meter[13] measures the ratio of the peak and full-wave rectified average values of a low duty-cycle pulse train transmitted over a transmission path. Its purpose is to provide a measure of the overall transmission distortion that is present from many different causes. The P/AR meter is most sensitive to envelope delay distortion and is also affected by noise, bandwidth reduction, gain ripples, nonlinearities such as compression and clipping, and other impairments. The P/AR reading is a indication of the general transmission quality of the voiceband channel. In particular, it is a measure of the phase linearity of the toll connecting trunk. If the P/AR signal were received entirely undistorted the P/AR meter would read 100. Distortion normally causes readings lower than 100.

Four P/AR meter readings were made on each toll connecting trunk in the survey. Both directions of transmission were measured, using both polarities of the nonsymmetrical P/AR test signal. However, since there was no significant difference between readings using the normal and inverted polarity test signal, the average of the two was used in deriving estimates. P/AR generators now in production have eliminated this effect by inverting alternate signals. Table XVI shows

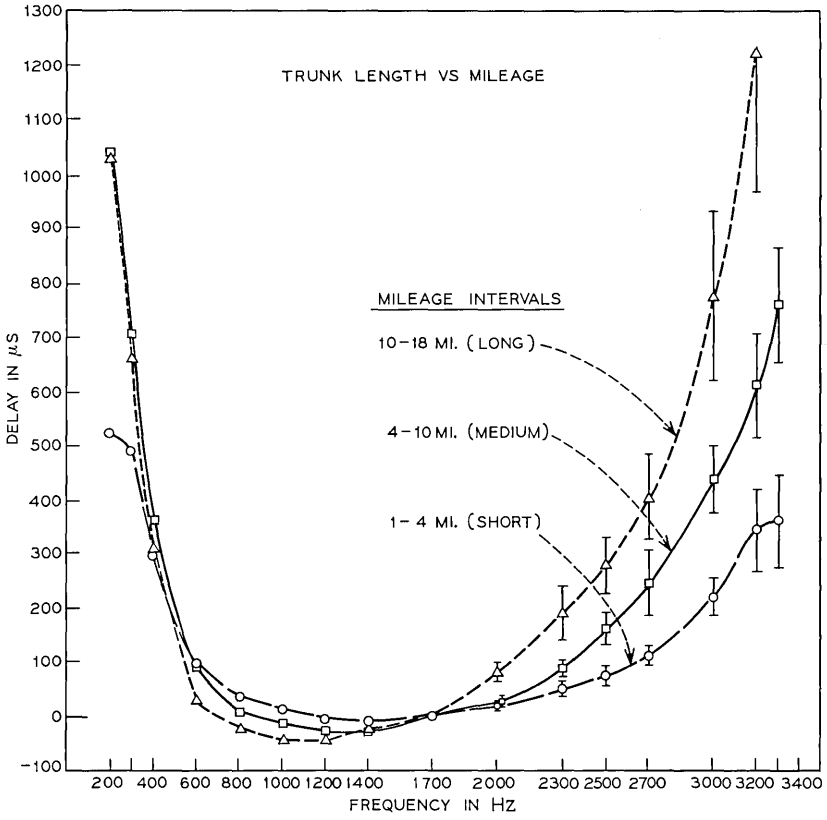


Fig. 8—Relative envelope delay (VF loaded facilities only).

the estimated values for the P/AR distributions of all toll connecting trunks and a variety of other subclasses based on types of facility.

It should be noted that different facilities have different distributions of P/AR readings, so that the same numerical value may indicate a lower than expected reading on one facility but be a higher than expected reading for a different facility. Hence, a comparison of subjective quality between trunks should be made only within a particular facility category, and not between categories. In general, comparing P/AR readings indicates the relative amount of intersymbol interference to be expected in each case.

Since the P/AR meter reading is affected by envelope delay and frequency response it is not surprising that there is a relationship

between these parameters and the P/AR readings. In fact, there is a high degree of correlation between these characteristics, which implies that a trunk's relative performance in any of the three measures is a good indication of its general transmission quality with relation to other trunks of the same type.

To show the relationship between P/AR, envelope delay distortion, and loss, consider the performance of the more commonly used facilities with respect to the three mentioned parameters. Besides the mean P/AR reading for a category, the mean loss relative to 1000 Hz (dB) and the relative envelope delay (μs) at 3000 Hz will be used. These values for the major identifiable categories are shown in Table XVII.

Since P/AR, loss, and envelope delay have different underlying scales of measurement and unknown distributions, one method of measuring their relationship is by a rank correlation test.[14] This measures the agreement between the performance with respect to the three parameters. The seven categories of Table XVII can be ranked,* as follows:

	Delay	P/AR	Loss
VF no load, no rptr	1	1	1
VF no rptr	2	2	3
T1	3	4	2
N2	4	3	4
VF loaded	5	5	5
VF repeatered	6	6	6
N1	7	7	7

This yields a coefficient of concordance[14] of 0.96, which is significant beyond the 0.01 point (if one feels that such small probabilities are significant). There evidently is a very high degree of correlation between these three transmission characteristics. In other words, if for a given facility one trunk exhibits a "better" P/AR reading than another trunk, it is also likely that it will have "better" loss and delay characteristics.

VIII. LEVEL TRACKING ON COMPANDORED FACILITIES

Of the 392 toll connecting trunks in the survey, 139 were on carrier facilities and these were compandored. These 139 trunks terminated in a total of 40 end offices.

* The number one (1) indicates lowest loss or delay, or highest P/AR reading.

Compandor tracking was measured by transmitting an 1850-Hz tone at each of four power levels (0, −10, −20, and −30 dBm) and then measuring the received level. Measurements were made in both directions of transmission, using a standard oscillator and the 25A Transmission Measuring Set (or 3A Noise Measuring Set for the −30-dBm level).

The channel net gain for any given input power level can be taken as the reference level, and the net gain for other input levels is then referred relative to the reference level. For example, Table XVIII shows the deviations from channel net gain, relative to a −10-dBm input, at the other input levels. So at each input power, the variable of interest is the net gain of that power level minus the net gain at −10 dBm. Estimated mean deviations relative to any other power level can also be obtained from Table XVIII.

IX. HARMONIC DISTORTION

The nonlinearities of the transmission path for the toll connecting trunks caused harmonic distortion of the signals being carried. The amount of harmonic distortion was determined by sending a 1000-Hz tone at each of four power levels (0, −10, −20, and −30 dBm) and measuring the power level of the received fundamental tone and the second and third harmonics. The fundamental was measured with a 25A Set if above −25 dBm. Below this level, a 3A Set with C-message weighting was used. The second and third harmonic measurements used the 3A Set with a narrow bandpass filter, centered at 2 or 3 kHz, in place of the C-message weighting network. The measurements were adjusted to reflect the loss of the bandpass filters. There was also a "measurement floor" to consider due to the sensitivity of the measuring equipment. This was determined to be 48 dB (i.e., a fundamental-to-harmonic ratio of 48 dB) so that harmonics more than 48 dB below the 1000-Hz fundamental were measured with increasing inaccuracy as the level decreased. Measurements were made in both directions of transmission.

The estimates of harmonic distortion are shown in Table XIX for all the toll connecting trunks measured in the survey. The results are presented as signal-to-harmonic distortion power ratios (S/D) showing means, confidence intervals, and standard deviations. An analysis of the data showed that the larger values of distortion (lower S/D ratios) almost always occurred on compandored systems (i.e., all the carrier facilities). The harmonic distortion on the noncompandored trunks

(i.e., voice-frequency cable facilities) was usually below the "measurement floor," resulting in only a few accurate measurements on these facilities. Hence, they were not analyzed in detail. The results for compandored toll connecting trunks (carrier facilities) are shown in Table XX. Among these carrier facilities a subclass analysis of the more commonly used facilities is shown in Table XXI. The estimates there are for the signal-to-distortion ratio, i.e., the total harmonic distortion below the fundamental.

In studying Table XIX and XX one notes that the average distortion power is roughly the same for the second and third harmonics. However, the second and third harmonic distortion power distribution are dissimilar, as can be seen from the plot of their distributions in Fig. 9. It shows clearly that for the high distortion levels, a greater percentage of the trunks have second harmonic power exceeding a chosen level than third harmonic power. About 20 percent of compandored facilities have a fundamental-to-second harmonic distortion ratio less than 30 dB, while only one percent have a fundamental-to-third ratio less than

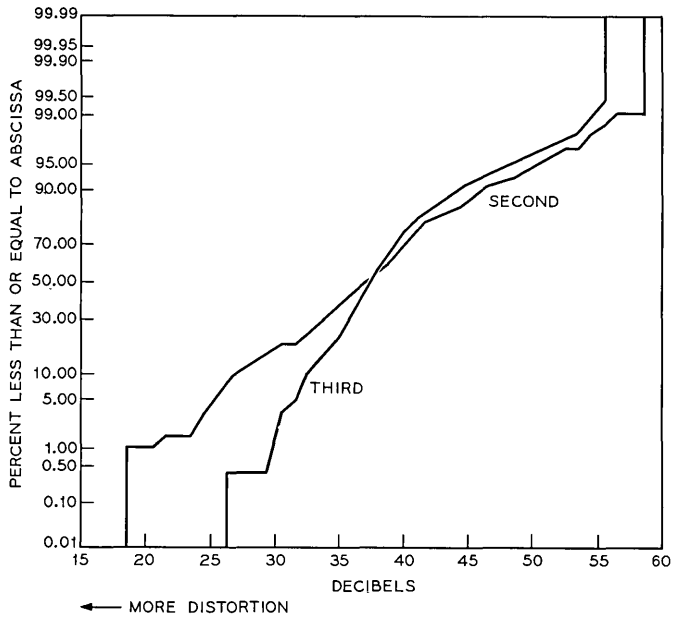


Fig. 9—Second and third harmonic levels in dB below first harmonic (received fundamental) for compandored systems—end and toll office data. Input level of fundamental = −10 dBm.

30 dB. Hence, when the total harmonic distortion is relatively large it is usually due to the second harmonic component. This feature is also evident in a distribution plot of the second and third harmonics for all toll connecting trunks, and implies a positive skewness[15,16] of the population of third harmonic measurements.

X. CONCLUDING REMARKS

The 1966 survey of toll connecting trunks has supplied information on the transmission characteristics of a key part of the built-up connection between subscribers. In addition, when combined with other such data, this information has been useful in generating a transmission model of the Bell System network.

Although carried out as a trunk survey, it was possible through subclass analysis to derive estimates of the transmission performance of toll connecting trunks using specific facilities. These included trunks using voice-frequency cable pairs, loaded and nonloaded, and those using voice-frequency repeaters, as well as the more commonly used short-haul carrier systems, N1 and T1. These trunk data reflect the effects of office equipment as well as those of the facilities.

Another use of the 1966 survey data has been in making comparisons with data from the 1964 Intertoll Trunk Survey.[10] Certain subclasses of each survey are similar enough to warrant investigation. For example, in both cases the trunks of 0 to 15 miles trunk length use voice-frequency facilities for about 80 percent of the trunks. Since their makeup is so similar their respective estimates for loss, noise, and delay can be compared. The agreement of these estimates is very good, and this lends mutual confidence to both estimates. The same is true if we consider the collection of trunks 15 to 30 miles long where N1 carrier was the predominantly used facility.

It should be remembered that although these results accurately picture the transmission characteristics of toll connecting trunks in 1966, with the passage of time changes can occur in the sampled population. Introduction of new facilities and the uneven growth of others are two obvious occurrences which create the need for continual updating of transmission data.

XI. ACKNOWLEDGMENTS

The 1966 toll connecting trunk survey represents the combined efforts of numerous people at Bell Laboratories, the American Telephone and Telegraph Company, and the Operating Companies. D. T. Osgood

of the American Telephone and Telegraph Company served as coordinator between Bell Laboratories and the Operating Companies, and the actual measurements were performed by more than 20 Bell Laboratories people. T. L. Bequette, G. P. McNamara, and J. T. Powers, Jr., as well as the author, supplied the data analysis associated with the large body of information gathered by the survey. This task was greatly facilitated by the use of computing programs and techniques created by F. P. Duffy. Only through a lengthy, collective effort has this survey been completed, and it is here duly acknowledged.

REFERENCES

1. Nåsell, I., "The 1962 Survey of Noise and Loss on Toll Connections," B.S.T.J. *43*, No. 2 (March 1964), pp. 697–718.
2. Nåsell, I., "Some Transmission Characteristics of Bell System Toll Connections," B.S.T.J., *47*, No. 6 (July–August 1968), pp. 1001–1018.
3. Balkovic, M. D., Klancer, H. W., Klare, S. W., and McGruther, W. G., "1969–70 Connection Survey: High-Speed Voiceband Data Transmission Performance on the Switched Telecommunications Network," B.S.T.J., *50*, No. 4 (April 1971), pp. 1349–1384.
4. Duffy, F. P., and Thatcher, T. W., Jr., "1969–70 Connection Survey: Analog Transmission Performance on the Switched Telecommunications Network," B.S.T.J., *50*, No. 4 (April 1971), pp. 1311–1347.
5. Fleming, H. C., and Hutchinson, R. M., Jr., "1969–70 Connection Survey: Low-Speed Data Transmission Performance on the Switched Telecommunications Network," B.S.T.J., *50*, No. 4 (April 1971), pp. 1385–1405.
6. Hansen, M. H., Hurwitz, W. N., and Madow, W. G., *Sample Survey Methods and Theory*, Vols. I and II, New York: J. Wiley and Sons, 1953.
7. Cochran, W. G., *Sampling Techniques*, New York: J. Wiley and Sons, 1963.
8. Kish, L., *Survey Sampling*, New York: J. Wiley and Sons, 1967.
9. Cochran, W. T., and Lewinski, D. A., "A New Measuring Set For Message Circuit Noise," B.S.T.J., *39*, No. 4 (July 1960), pp. 911–931.
10. Nåsell, I., Ellison, C. R., and Holmstrom, R., "The Transmission Performance of Bell System Intertoll Trunks," B.S.T.J., *47*, No. 8 (October 1968), pp. 1561–1613.
11. Fultz, K. E., and Penick, D. B., "The T1 Carrier System," B.S.T.J., *44*, No. 7 (September 1965), pp. 1405–1451.
12. Fennick, J. H., "Amplitude Distributions of Telephone Channel Noise and a Model for Impulse Noise," B.S.T.J., *48*, No. 10 (December 1969), pp. 3243–3263.
13. Cochran, W. T., "Measuring Waveform Distortion With a P/AR Meter," Bell Laboratories Record, *43*, No. 9 (October 1965), pp. 367–371.
14. Kendall, M. G., *Rank Correlation Methods*, 2nd ed., London: Charles Griffin and Co., 1955.
15. Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, 6th ed., New York: J. Wiley and Sons, 1967.
16. Mood, A. M., *Introduction to the Theory of Statistics*, New York: McGraw-Hill Co., 1950.

TABLE I—ESTIMATED TRUNK LENGTH AND FACILITY TYPE IN LARGE AND SMALL OFFICES

Office Size	Trunk Length (Miles)	Type of Facility (%)	
	Mean ± 90% C.I.	Carrier	Voice Frequency
Small (<400,000 AOTM*) Stratum 1	25.2 ± 11.2	63	37
Large (≧400,000 AOTM) Stratum 2	6.7 ± 2.4	19	81

* AOTM: Annual Outgoing Toll Messages

TABLE II—TRUNK LENGTH VS FACILITY TYPE

Facility Type	Percent of TCTs in Length Categories		
	0–15 Miles	15.1–30 Miles	Over 30 Miles
N1 Carrier	5.8	42.0	27.0
T1 Carrier	10.2	4.6	—
Other Carrier*	1.0	32.1	73.0
Voice-Frequency Cable	83.0	21.3	—
	100.0	100.0	100.0

* The facilities in this category consist of ON, ON/R, O, N2, and Lenkurt carrier types.

TABLE III—FACILITY COMPOSITION OF TOLL CONNECTING TRUNKS

Facility Type	Percent of all TCTs Mean ± 90% C.I.
1. Carrier:*	34
N1	13 ± 8
T1	8 ± 8
Other†	10
2. Voice-Frequency Cable:	66
Nonloaded	23 ± 11
Loaded:	43 ± 13
Loaded, nonrepeatered	6 ± 5
Loaded, E6 repeater	26 ± 10
Loaded, E23 repeater	7 ± 6

* 3% of these are carrier facilities in tandem with a section of loaded voice-frequency cable.

† The facilities in this category consist of ON, ON/R, O, N2, and Lenkurt carrier types in quantities individually too small for reliable estimates.

TABLE IV—SWITCHING USED BY TOLL CONNECTING TRUNKS OF SURVEY

	Stratum 1 (<400,000 AOTM*)—15 Small End Offices					
	At End Office		At Toll Office		Total Both Ends	
Type	Trunks	Machines	Trunks	Machines	Trunks	Machines
SXS	120	12	34	4	154	16
5XB	20	2	48	6	68	8
SWBD	10	1	24	5	34	6
XBT	—	—	32	4	32	4
4A	—	—	12	2	12	2
Total	150	15	150	21	300	36

	Stratum 2 (≧400,000 AOTM)—25 Large End Offices					
	At End Office		At Toll Office		Total Both Ends	
Type	Trunks	Machines	Trunks	Machines	Trunks	Machines
SXS	91	9	36	5	127	14
5XB	115	13	11	2	126	15
XBT	—	—	119	16	119	16
1XB	25	4	7	2	32	6
SWBD	—	—	44	13	44	13
4A	—	—	25	6	25	6
PANEL	11	3	—	—	11	3
Total	242	29	242	44	484	73
Grand Totals	392	44	392	65	784	109

Notes: SXS: Switchboard
5XB: No. 5 Crossbar
1XB: No. 1 Crossbar
SWBD: Switchboard
4A: 4A Crossbar
XBT: Crossbar Tandem

* AOTM: Annual Outgoing Toll Messages

TABLE V—FREQUENCY RESPONSE OF TOLL CONNECTING TRUNKS

Frequency (Hz)	Switch-to-Switch Loss (dB)					
	All Trunks		Carrier Facilities		Voice-Frequency Cable	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
200	8.1 ± 0.8	4.0	11.1 ± 1.3	4.5	6.6 ± 0.6	2.5
300	4.6 ± 0.3	1.9	5.4 ± 0.4	2.3	4.2 ± 0.3	1.6
400	3.7 ± 0.2	1.6	4.0 ± 0.4	2.1	3.5 ± 0.3	1.3
600	3.3 ± 0.2	1.5	3.5 ± 0.4	2.0	3.1 ± 0.2	1.2
800	3.0 ± 0.2	1.4	3.0 ± 0.4	1.8	2.9 ± 0.2	1.1
1000	2.8 ± 0.2	1.4	2.8 ± 0.3	1.7	2.8 ± 0.2	1.2
1200	2.8 ± 0.2	1.4	2.8 ± 0.2	1.6	2.9 ± 0.2	1.2
1400	2.9 ± 0.2	1.4	2.9 ± 0.3	1.7	2.9 ± 0.2	1.2
1700	3.1 ± 0.2	1.5	3.3 ± 0.4	1.8	3.0 ± 0.2	1.3
2000	3.3 ± 0.2	1.6	3.6 ± 0.4	1.9	3.1 ± 0.3	1.4
2300	3.5 ± 0.2	1.7	3.7 ± 0.4	2.0	3.4 ± 0.3	1.5
2500	3.8 ± 0.3	1.9	3.9 ± 0.5	2.2	3.8 ± 0.3	1.7
2700	4.3 ± 0.3	2.1	4.3 ± 0.5	2.3	4.2 ± 0.4	2.0
3000	5.6 ± 0.5	3.0	6.1 ± 0.9	2.9	5.3 ± 0.7	3.0
3200	8.9 ± 1.3	6.1	12.7 ± 3.0	6.8	7.0 ± 1.1	4.7
3300	12.4 ± 2.1	9.5	19.9 ± 5.2	10.3	8.5 ± 1.5	6.2
3400	15.2 ± 2.4	11.0	23.4 ± 6.1	10.8	10.9 ± 2.2	8.4

TABLE VI—LOSS RELATIVE TO 1000 Hz (dB) FOR CARRIER FACILITIES

Frequency (Hz)	N1		N2		T1	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
1000– 200	6.2 ± 0.5	0.4	4.5 ± 0.8	1.5	10.7 ± 0.6	1.6
– 300	2.2 ± 0.2	1.2	2.9 ± 1.0	1.7	3.0 ± 0.1	0.6
– 400	1.0 ± 0.2	0.8	1.5 ± 0.6	1.4	1.2 ± 0.1	0.2
– 600	0.6 ± 0.1	0.4	0.7 ± 0.3	0.6	0.3 ± 0.1	0.1
– 800	0.2 ± 0.1	0.2	0.5 ± 0.3	0.8	0.0 ± 0.1	0.1
–1000	—	—	—	—	—	—
–1200	0.0 ± 0.1	0.2	−0.2 ± 0.1	0.3	0.0 ± 0.1	0.1
–1400	0.1 ± 0.1	0.3	0.2 ± 0.1	0.7	0.1 ± 0.1	0.1
–1700	0.5 ± 0.1	0.5	0.5 ± 0.2	0.7	0.1 ± 0.1	0.1
–2000	0.9 ± 0.2	0.7	0.3 ± 0.1	0.5	0.3 ± 0.1	0.1
–2300	0.9 ± 0.3	1.0	0.9 ± 0.3	1.1	0.5 ± 0.1	0.2
–2500	1.2 ± 0.3	1.2	1.7 ± 1.0	2.7	0.7 ± 0.1	0.2
–2700	2.1 ± 0.4	1.4	1.7 ± 1.1	2.8	1.0 ± 0.1	0.2
–3000	4.8 ± 0.5	1.7	2.7 ± 2.1	5.3	1.9 ± 0.1	0.3
–3200	16.8 ± 0.8	2.5	3.7 ± 2.4	5.3	3.0 ± 0.1	0.3
–3300	—	—	3.4 ± 2.4	5.4	3.8 ± 0.2	0.3
–3400	—	—	5.3 ± 2.6	5.4	4.7 ± 0.2	0.3

TABLE VII—LOSS RELATIVE TO 1000 HZ (dB) FOR VOICE-FREQUENCY CABLE

Frequency (Hz)	VF2W Facilities							
	Nonloaded, No Rptr		Loaded, E23 Rptr		Loaded, E6 Rptr		Loaded, No Rptr	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
1000– 200	2.5 ± 0.5	1.2	4.1 ± 0.9	1.9	5.2 ± 0.9	2.2	2.2 ± 0.6	1.5
– 300	0.8 ± 0.3	0.7	1.1 ± 0.3	0.7	2.1 ± 0.5	1.2	0.5 ± 0.2	0.6
– 400	0.4 ± 0.2	0.5	0.6 ± 0.1	0.5	1.1 ± 0.3	0.8	0.1 ± 0.1	0.5
– 600	0.1 ± 0.1	0.3	0.4 ± 0.1	0.3	0.5 ± 0.1	0.5	0.0 ± 0.1	0.3
– 800	0.0 ± 0.1	0.1	0.0 ± 0.1	0.3	0.2 ± 0.1	0.3	0.0 ± 0.1	0.1
–1000	—	—	—	—	—	—	—	—
–1200	0.0 ± 0.1	0.1	0.1 ± 0.1	0.4	0.0 ± 0.1	0.3	0.1 ± 0.1	0.1
–1400	0.0 ± 0.1	0.3	0.2 ± 0.2	0.5	0.1 ± 0.1	0.3	0.1 ± 0.1	0.2
–1700	0.1 ± 0.1	0.5	0.3 ± 0.2	0.5	0.1 ± 0.1	0.4	0.3 ± 0.1	0.4
–2000	0.1 ± 0.2	0.6	0.5 ± 0.3	0.8	0.3 ± 0.1	0.5	0.5 ± 0.1	0.4
–2300	0.2 ± 0.2	0.8	0.8 ± 0.5	1.0	0.7 ± 0.1	0.5	0.9 ± 0.2	0.6
–2500	0.4 ± 0.2	0.9	1.2 ± 0.2	0.9	1.3 ± 0.2	0.8	1.2 ± 0.2	0.6
–2700	0.4 ± 0.2	0.9	1.7 ± 0.3	1.2	2.2 ± 0.4	1.1	1.4 ± 0.1	0.7
–3000	0.4 ± 0.2	1.0	3.0 ± 0.8	1.5	3.8 ± 0.7	1.9	2.3 ± 0.3	1.3
–3200	0.5 ± 0.2	1.0	5.0 ± 1.5	2.5	6.5 ± 1.2	2.9	3.5 ± 0.3	2.0
–3300	0.5 ± 0.2	1.0	6.8 ± 2.4	3.6	9.0 ± 1.7	4.2	4.6 ± 0.4	2.2
–3400	0.5 ± 0.2	1.0	11.4 ± 5.1	6.9	13.2 ± 2.6	6.1	6.2 ± 0.7	2.7

TABLE VIII—3A NOISE ON TOLL CONNECTING TRUNKS

Facility Type	C-Message Weighting			3-kHz Flat Weighting		
	Mean (dBrn)	90% C.I. (dB)	Std. Dev. (dB)	Mean (dBrn)	90% C.I. (dB)	Std. Dev. (dB)
Carrier						
N1	21.2	±0.9	4.8	34.7	±2.3	7.6
N2	17.5	—	8.1	38.8	—	7.9
T1	16.0	—	3.8	31.9	—	3.8
All Carrier	18.4	±1.7	6.4	33.3	±1.8	8.0
Voice-Frequency Cable						
Nonloaded, nonrepeatered	5.5	±1.4	4.4	33.5	±3.1	14.3
Loaded, nonrepeatered	9.5	±4.2	7.8	37.0	±3.2	13.2
Loaded, E6 repeater	11.3	±2.1	7.3	35.9	±3.3	11.1
Loaded, E23 repeater	13.2	±2.8	6.9	39.4	±4.4	11.4
All VF	9.5	±1.5	7.2	35.7	±1.9	12.7
All TCTs	12.5	±1.4	8.2	34.5	±1.5	11.4

TABLE IX—3A NOISE WITH 3-KHZ FLAT WEIGHTING
(δ = End Office Noise − Toll Office Noise)

Facility Type	Mean δ (dB)	90% C.I. (dB)	Std. Dev. (dB)
Carrier			
N1	6.1	±2.9	9.2
N2	9.0	—	10.9
T1	3.5	—	10.1
All Carrier	6.0	±2.7	9.9
Voice-Frequency Cable			
Nonloaded, nonrepeatered	18.7	±3.4	10.8
Loaded, nonrepeatered	19.5	±1.9	4.6
Loaded, E6 repeater	12.0	±2.9	9.2
Loaded, E23 repeater	13.9	±7.8	12.4
All VF	15.2	±2.3	10.4
All TCTs	12.1	±2.2	11.1

TABLE X—3A NOISE WITH 3-KHZ FLAT WEIGHTING

Survey	Mean (dBrn)	90% C.I. (dB)	Std. Dev. (dB)
1964 ITT Survey (0–62.5 miles)	29.4	±2.0	9.0
1966 TCT Survey (toll offices)	28.7	±1.4	9.0

TABLE XI—3A NOISE WITH C-MESSAGE WEIGHTING: TRUNK LENGTH ANALYSIS

Mileage Category	Mean (dBrnC)	90% C.I. (dB)	Std. Dev. (dB)
0.0– 2.0	5.5	±1.3	4.3
2.1– 4.0	9.5	±0.8	7.9
4.1– 8.0	12.3	±1.9	7.3
8.1–16.0	15.6	±2.0	6.6
16.1–32.0	17.3	±3.0	6.8
Over 32.0	19.0	±2.2	7.1

TABLE XII—C-MESSAGE NOISE AS A FUNCTION OF INPUT HOLDING TONE POWER

Transmit Level (dBm)	Received Noise Level (dBrnC)					
	Analog Systems (Syllabic Compandor)		N1 Carrier		T1 Carrier (Instantaneous Compandor)	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90%*	Std. Dev.
0	39.2 ± 1.8	6.3	41.9 ± 2.1	6.4	46.2 ±	2.0
−10	32.2 ± 2.0	6.5	35.1 ± 2.2	6.5	39.9 ±	2.1
−20	27.4 ± 2.0	6.4	30.5 ± 1.9	6.0	30.8 ±	2.1
−30	22.2 ± 2.1	6.3	25.7 ± 1.7	5.2	21.2 ±	2.0

* Confidence interval not shown because of small sample size.

TABLE XIII—IMPULSE COUNT LEVEL DISTRIBUTIONS

Facility Type	45-Count Level (dBrnC)		15-Count Level (dBrnC)		Slope (dB/decade)	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
N1 Carrier	62.8 ± 4.7	8.4	65.9 ± 4.7	8.4	5.8 ± 1.0	2.1
T1 Carrier	57.8 ± 4.3	7.6	58.7 ± 4.3	7.5	2.7 ± 0.2	1.8
VF2W:						
NL, NR	43.8 ± 4.3	10.1	49.5 ± 4.6	11.0	10.3 ± 1.3	3.8
L, NR	44.1 ± 4.2	10.2	54.1 ± 4.8	10.9	8.5 ± 2.6	4.3
L, E6	49.3 ± 1.9	7.0	52.4 ± 2.1	7.2	5.6 ± 0.6	2.0
L, E23	46.8 ± 1.5	6.7	50.4 ± 1.3	6.8	7.1 ± 0.9	3.4
All VF*	46.0 ± 1.9	8.7	50.5 ± 1.9	9.1	8.2 ± 0.7	3.9
All TCTs	49.4 ± 2.0	10.2	53.3 ± 1.9	10.0	7.3 ± 0.8	3.8

Notes: Measurement interval 15 minutes
L = loaded cable
NL = nonloaded cable
NR = nonrepeatered cable
E6, E23 = E6, E23 repeaters
* Includes 2-wire and 4-wire

TABLE XIV—RELATIVE ENVELOPE DELAY OF CARRIER FACILITIES*

Frequency (Hz)	Delay (μs)					
	N1		N2		T1	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
200	3390 ± 155	391	1602 ± 126	271	1707 ± 70	269
300	2055 ± 88	232	1245 ± 144	157	1113 ± 66	182
400	1134 ± 64	140	773 ± 101	115	683 ± 46	100
600	416 ± 36	75	324 ± 43	60	269 ± 12	36
800	175 ± 16	47	135 ± 13	47	117 ± 6	20
1000	63 ± 16	47	102 ± 27	69	50 ± 5	18
1200	27 ± 5	22	69 ± 34	86	15 ± 3	18
1400	6 ± 5	19	17 ± 10	28	0 ± 3	12
1700	0	—	0	—	0	—
2000	17 ± 7	18	41 ± 24	89	9 ± 1	7
2300	82 ± 6	24	30 ± 10	45	33 ± 2	10
2500	159 ± 13	40	40 ± 7	12	54 ± 3	19
2700	242 ± 13	49	108 ± 12	76	92 ± 3	9
3000	615 ± 19	91	170 ± 9	26	169 ± 4	15
3200	698 ± 23	99	201 ± 17	21	212 ± 2	18
3300	—	—	340 ± 15	53	238 ± 3	18
3400	—	—	429 ± 5	57	271 ± 3	16

* Includes office equipment.

TABLE XV—RELATIVE ENVELOPE DELAY OF VOICE-FREQUENCY CABLE FACILITIES*

Frequency (Hz)	Delay (μs)							
	VF2W {Nonloaded Nonrepeatered		VF2W-E6 Repeater		VF2W-E23 Repeater		VF2W {Loaded Nonrepeatered	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
200	604 ± 60	202	679 ± 227	505	973 ± 522	700	872 ± 139	280
300	442 ± 48	159	559 ± 105	304	587 ± 271	400	570 ± 87	219
400	251 ± 30	92	297 ± 54	186	305 ± 131	217	299 ± 47	121
600	102 ± 14	39	64 ± 29	83	42 ± 55	87	96 ± 17	47
800	50 ± 7	21	− 4 ± 26	55	5 ± 28	76	27 ± 11	25
1000	28 ± 4	12	−23 ± 21	45	−20 ± 24	38	5 ± 10	20
1200	13 ± 2	9	−29 ± 15	34	−41 ± 13	47	−11 ± 9	18
1400	6 ± 1	7	−22 ± 9	24	−29 ± 24	57	−10 ± 8	17
1700	0	—	0	—	0	—	0	—
2000	− 6 ± 2	7	46 ± 16	39	57 ± 26	63	21 ± 13	20
2300	−10 ± 3	9	116 ± 36	76	142 ± 80	109	36 ± 17	26
2500	−11 ± 4	11	177 ± 56	113	212 ± 44	84	70 ± 27	45
2700	−13 ± 4	12	255 ± 75	156	299 ± 122	181	121 ± 43	68
3000	−13 ± 5	13	501 ± 148	310	576 ± 188	265	217 ± 68	103
3200	− 9 ± 5	13	772 ± 229	470	899 ± 347	464	320 ± 126	167
3300	− 8 ± 5	13	957 ± 309	627	1097 ± 440	577	341 ± 155	224
3400	− 9 ± 5	13	1018 ± 363	711	1275 ± 601	802	398 ± 266	327

* Includes office equipment.

TABLE XVI—P/AR ESTIMATES

Category	Mean ± 90% C.I.	Std. Dev.
1. All toll connecting trunks	93.6 ± 1.3	7.4
2. All carrier facilities:	90.9 ± 2.2	6.9
N1 carrier	88.8 ± 1.7	6.9
N2 carrier	96.2 ± 2.6	6.4
T1 carrier	96.1 ± 0.7	4.7
3. All voice-frequency cable:	95.0 ± 1.7	7.1
VF-nonloaded	98.6 ± 0.6	2.6
VF-loaded	93.1 ± 2.5	8.0
VF2W-E23 repeater	92.4 ± 3.8	6.2
VF2W-E6 repeater	92.3 ± 3.8	9.2
VF2W-loaded, nonrepeatered	97.0 ± 1.8	2.8
VF4W	92.5	4.0

TABLE XVII—ESTIMATED MEANS OF RELATED TRANSMISSION PARAMETERS

Facility Type	Mean P/AR	Relative Envelope Delay (μs)	Loss* Relative To 1000 Hz (dB)
VF: no load, no repeater	98.2	−13	0.4
VF: no repeater	98.2	36	2.3
VF: loaded	93.1	475	3.6
VF: repeatered	92.5	522	3.8
Carrier: N1	88.8	615	4.8
N2	96.2	170	2.7
T1	96.1	169	1.9

* At 3000 Hz.

TABLE XVIII—COMPANDOR LEVEL TRACKING
(Estimated Deviations from Net Gain at −10-dBm Input)

Input Power (dBm)	Mean ± 90% C.I. (dB)	Std. Dev. (dB)
0	−0.3 ± 0.1	0.5
−10	0	—
−20	0.7 ± 0.1	0.5
−30	1.2 ± 0.3	1.2

TABLE XIX—STATISTICS FOR THE RATIO OF FUNDAMENTAL TO SECOND, THIRD, OR TOTAL HARMONIC POWER IN dB (ALL TOLL CONNECTING TRUNKS)

	Transmitted Level (dBm)							
	0		−10		−20		−30	
Harmonic	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
Second	47.1 ± 2.4	9.3	48.1 ± 2.8	9.3	48.9 ± 3.0	9.6	48.0 ± 2.9	9.4
Third	46.6 ± 2.0	9.0	47.9 ± 1.8	7.9	49.0 ± 2.0	8.2	47.2 ± 2.0	8.4
Total	43.3 ± 2.3	8.9	44.3 ± 2.5	8.6	45.3 ± 2.8	9.0	43.9 ± 2.7	9.1

TABLE XX—STATISTICS FOR THE RATIO OF FUNDAMENTAL TO SECOND, THIRD, OR TOTAL HARMONIC POWER IN dB (CARRIER FACILITIES)

	Transmitted Level (dBm)							
	0		−10		−20		−30	
Harmonic	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
Second	36.0 ± 1.2	6.5	37.0 ± 2.9	7.3	37.9 ± 3.9	8.2	38.9 ± 3.1	7.2
Third	35.6 ± 2.1	6.0	38.2 ± 0.9	5.1	39.4 ± 1.5	5.6	39.2 ± 1.3	5.1
Total	32.0 ± 1.3	5.2	33.5 ± 2.0	5.5	34.4 ± 3.0	6.6	34.9 ± 2.4	5.8

TABLE XXI—SIGNAL-TO-DISTORTION RATIO FOR CARRIER FACILITIES (IN dB)

Transmitted Level (dBm)	Carrier Facility							
	N1		N2		ON		T1	
	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.	Mean ± 90% C.I.	Std. Dev.
0	28.0 ± 0.5	3.3	38.0 ± 1.0	4.2	32.5 ± 1.0	2.8	34.1 ± 0.2	3.7
−10	33.6 ± 0.6	4.0	38.9 ± 1.6	3.2	35.1 ± 1.2	3.0	29.3 ± 0.5	4.5
−20	36.7 ± 0.6	3.9	38.2 ± 0.8	4.7	35.6 ± 1.5	3.2	28.0 ± 0.4	4.7
−30	37.2 ± 0.6	3.5	37.0 ± 1.1	4.6	36.5 ± 0.6	3.2	30.0 ± 0.5	4.0

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

An Adaptive Echo Canceller in a Nonideal Environment (Nonlinear or Time Variant)

By E. J. THOMAS

(Manuscript received October 19, 1970)

In this paper we calculate a lower bound on suppression provided by an adaptive echo canceller in either a nonlinear or time variant environment. Specifically, we examine the effects on performance of nonlinear echo paths described by a Volterra integral equation [equation (18)] or time variant echo paths caused by phase jitter on single-sideband suppressed carrier systems.

I. INTRODUCTION

In a previous paper[1] we have described the performance of an adaptive echo canceller in a linear time invariant environment. Qualitatively speaking, the environment in which an echo canceller will operate appears to be mainly linear and time invariant; however, some echo paths will be nonlinear and/or time variable. Some typical causes of nonlinearity are the volume dependent gain in compandored circuits† and harmonic distortion in amplifiers and repeaters. Time variability, on the other hand, may be caused by spurious modulation of the carrier of long-haul single-sideband suppressed carrier systems. This is commonly referred to as incidental FM or phase jitter. In certain cases these anomalies are of sufficient magnitude to degrade the performance of an adaptive echo canceller.

In this paper we examine the operation of an adaptive echo canceller studied previously in an ideal environment[1] in a nonideal environment (nonlinear, time variant). We restrict ourselves to nonlinearities which do not possess infinite memory and to time variability caused by phase jitter. In both cases we derive a lower bound on the suppression provided by the echo canceller and empirically verify

† This problem occurs when the compressor portion of the compandor is not perfectly compensated for by the expandor.

our results. For the sake of brevity we restrict our discussion to a digital implementation of the type shown in Fig. 1. However, one can investigate other implementations of the echo canceller and obtain similar results.

II. GENERAL CONSIDERATIONS AND NOTATIONAL FORMS

In Fig. 1 we show how an echo canceller would be connected in a typical connection. The input signal $x(t)$ produces an echo $y(t)$ corrupted by noise, $\zeta(t)$. An approximation of $y(t)$, $y_A(t)$, is subtracted from the actual echo and noise producing a cancelled echo $e(t)$. Examining the echo canceller in a little more detail we find that it is composed of M digital filters having the set of orthonormal impulse responses $\{\lambda_i(k)\}$, and the set of outputs $\{w_i(k)\}$ to input $x(k)$.† Every tap has associated with it the adaptive network shown for two taps in the figure. From Fig. 1 we conclude that the gain of the ith tap at the k + 1 sampling interval is given by

$$g_i(k+1) = g_i(k) + |K|\, e(k) w_i(k). \tag{1}$$

Let us assume that the response of the echo path to $x(t)$ may be given in the following form:

$$y(t) = x(t) * h(t) + \psi(t). \tag{2}$$

where $h(t)$ is any square integrable function, (*) denotes convolution, and $\psi(t)$‡ is any function which is required to make equation (2) correct. Qualitatively speaking we see from equation (2) that we are breaking $y(t)$ into a linear component $x(t) * h(t)$ and a distortion term $\psi(t)$. In subsequent sections of this paper we will demonstrate that the nonlinear or time variant echo paths that we choose to study yield $y(t)$'s of the form of equation (2). We will therefore first examine the effect of a system described by equation (2) on the performance of an adaptive echo canceller. We will then calculate $\psi(t)$ for a nonlinear or a time variant echo path.

2.1 *The Governing Equation*

Since we hypothesized that $h(t)$ was a square integrable function we may represent it by a generalized Fourier series

$$h(t) = \sum_{i=1}^{\infty} C_i \lambda_i(t)$$

† $x(k)$ denotes the value of $x(t)$ at the kth sampling interval; also, all sampling is assumed to be at the Nyquist rate.

‡ $\psi(t)$ may depend on the absolute time when $x(t)$ is applied.

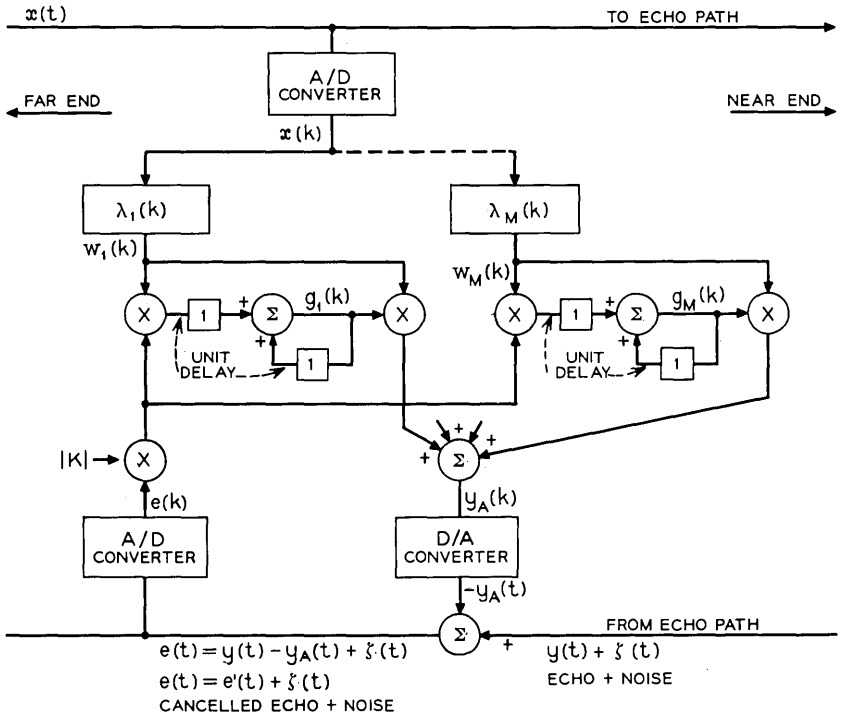


Fig. 1—An adaptive echo canceller.

where

$$C_i = \int_0^\infty h(t)\lambda_i(t)\,dt.$$

Substituting $\sum_{i=1}^{\infty} C_i\lambda_i(t)$ for $h(t)$ in equation (2) we obtain

$$y(t) = \sum_{i=1}^{\infty} C_i w_i(t) + \psi(t)$$

where $w_i(t) = x(t) * \lambda_i(t)$.
From Fig. 1 we see that

$$y_A(t) = \sum_{i=1}^{M} g_i(t)w_i(t).$$

Therefore,

$$e(t) = \sum_{i=1}^{M} r_i(t)w_i(t) + \sum_{i=M+1}^{\infty} C_i w_i(t) + \psi(t) + \zeta(t) \qquad (3)$$

where

$$r_i(t) = C_i - g_i(t). \tag{4}$$

Combining equations (1) and (4) we obtain

$$r_i(k+1) = r_i(k) - |K| w_i(k)e(k).$$

Squaring the above equation, summing over the M taps of the echo canceller, and using equation (3) we obtain

$$\begin{aligned}\sum_{i=1}^{M} (r_i^2(k+1) - r_i^2(k)) = -2|K| \sum_{i=1}^{M} r_i(k)w_i(k)\Big(\sum_{j=1}^{M} r_j(k)w_j(k) \\ + \sum_{j=M+1}^{\infty} C_j w_j(k) + \psi(k) + \zeta(k)\Big) \\ + |K|^2 \sum_{i=1}^{M} w_i^2(k)\Big(\sum_{j=1}^{M} r_j(k)w_j(k) \\ + \sum_{j=M+1}^{\infty} C_j w_j(k) + \psi(k) + \zeta(k)\Big)^2.\end{aligned} \tag{5}$$

In order to simplify our notation we make the following three definitions:

$$\theta(k) \equiv \sum_{i=M+1}^{\infty} C_i w_i(k) + \psi(k), \tag{6a}$$

$$u^2(k) \equiv \sum_{i=1}^{M} w_i^2(k), \tag{6b}$$

$$v(k) \equiv \sum_{i=1}^{M} r_i(k)w_i(k). \tag{6c}$$

As a result we obtain

$$\begin{aligned}\sum_{i=1}^{M} (r_i^2(k+1) - r_i^2(k)) \\ = -2|K|(v^2(k) + \theta(k)v(k) + \zeta(k)v(k)) \\ + |K|^2 u^2(k)\{v^2(k) + \theta^2(k) + \zeta^2(k) + 2\theta(k)v(k) \\ + 2\zeta(k)v(k) + 2\zeta(k)\theta(k)\}.\end{aligned} \tag{7}$$

The above equation describes the performance of the adaptive echo canceller shown in Fig. 1 in the environment described by equation (2). However, before we can apply equation (7) to the problem at hand one final definition is required. We define the average value of a sam-

pled function $f(k)$ in the interval $Qp + 1 \leqq k \leqq (p + 1)Q$ as

$$\overline{f(p)} = \frac{1}{Q} \sum_{k=Qp+1}^{(p+1)Q} f(k) \qquad p = 0, 1, 2, \cdots \tag{8}$$

where Q is the number of samples per interval and is constant. Therefore, $\overline{f(p)}$ may be interpreted as the average value of $f(k)$ averaged over the pth interval containing Q samples. Note that for all p the $p - 1$ and $p + 1$ intervals are adjacent to and do not overlap the pth interval.

2.2 *Choices of* $| K |$ *Which Allow Best Match to the Echo Path Fourier Coefficients*

Let us now apply our averaging technique defined by equation (8) to equation (7). We will assume that the circuit noise $\zeta(t)$ is zero mean, with variance σ_ζ^2. Also, we will assume that the noise is statistically independent of the other variables in equation (7). This is justified by virtue of the fact that we are assuming that K is small.[1] Furthermore, we will assume that the number of samples in the pth interval, Q, is large enough so that $\overline{\zeta(p)}$ and $\bar{\zeta}^2(p)$ are good estimates of the true mean and variance of the circuit noise. As a result of the above we obtain

$$\begin{aligned} 1/Q \sum_{i=1}^{M} \{r_i^2[(Qp + 1) + Q] - r_i^2(Qp + 1)\} \\ = -2 \mid K \mid \overline{v^2(p)} - 2 \mid K \mid \overline{\theta(p)v(p)} \\ + 2 \mid K \mid^2 \overline{\theta(p)v(p)u^2(p)} \\ + \mid K \mid^2 \overline{u^2(p)v^2(p)} + \mid K \mid^2 \sigma_\zeta^2 \overline{u^2(p)} \\ + \mid K \mid^2 \overline{\theta^2(p)u^2(p)}. \end{aligned} \tag{9}$$

From (9) it is clear that the sum of the r_i^2 is reduced when the right-hand side of the equation is negative.† In fact, it is continually reduced until the right-hand side of (9) can no longer be negative. Therefore a sufficient condition for convergence‡ can be written as

$$\begin{aligned} \Delta(p) = \mid K \mid^2 [\overline{(v(p) + \theta(p))^2 u^2(p)} + \sigma_n^2 \overline{u^2(p)}] \\ - 2 \mid K \mid (\overline{v^2(p)} + \overline{\theta(p)v(p)}) < 0 \end{aligned} \tag{10a}$$

† This may be more easily seen if $\Sigma_{i=1}^{M} r_i^2$ is considered to be the magnitude squared of a vector whose components are $\{r_i\}$. In this event the left-hand side of the equation is proportional to the difference of the magnitude square of a vector at some instant and Q samples earlier. When the right-hand side is negative it indicates that the magnitude of the vector is smaller than it was Q samples ago.

‡ In this paper convergence is taken to mean the reduction of the sum of the r_i^2.

or equivalently

$$\Delta'(p) = |K| [\overline{(v(p) + \theta(p)^2 u^2(p) + \sigma_n^2 u^2(p))}] - 2\overline{(v^2(p) + \theta(p)v(p))} < 0. \tag{10b}$$

By defining

$$B(p) \equiv \overline{(V(p) + \theta(p))^2 U^2(p) + \sigma_n^2 U^2(p)},$$
$$D(p) \equiv \overline{(V^2(p) + \theta(p) V(p))},$$

we find

$$\Delta(p) = |K|^2 B(p) - 2|K| D(p).$$

From the above equation it should be clear that

$$\Delta(p) \leqq 0$$

for

$$0 < |K| < 2\left|\frac{D(p)}{B(p)}\right|$$

and

$$D(p) \geqq 0.$$

Therefore we conclude that the echo canceller converges for $D(p) > 0$ and $0 \leqq |K| \leqq 2D(p)/B(p)$. However, since $|V(p)|$ is small in the neighborhood of equilibrium, $D(p)$ decreases quicker than $B(p)$ and as a result the upper bound on $|K|$ is reduced while convergence is taking place. Convergence ceases when

$$|K| = \frac{2D(p)}{B(p)}.$$

On the other hand for $D(p) < 0$ there is no positive choice of K which will allow convergence and therefore the canceller diverges until $D(p)$ becomes positive and $0 < |K| \leqq |2D(p)/B(p)|$. As a result the best choice of $|K|$ which will allow the most convergence (the smallest $|D(p)|$ and therefore the best match to the Fourier coefficients C_i) is the smallest $|K|$. Ideally $|K| \to 0$ will allow the best match to the echo path Fourier coefficients. However, the smaller $|K|$ is made the longer it takes for the canceller to converge, and in practice a compromise is required between amount and speed of convergence.

2.3 *A Lower Bound on Achievable Suppression*

We have shown previously (Section 2.2) that choosing $|K|$ very small will allow the best match of the tap gains $g_i(k)$ to the Fourier co-

efficients C_i. Taking the limit of equation (10b) as $|K|$ goes to zero we obtain

$$\lim_{|K|\to 0} \Delta'(p) = -\overline{2(v^2(p)} + \overline{\theta(p)v(p))} \leqq 0$$

or convergence takes place as long as

$$\overline{v^2(p)} > -\overline{\theta(p)v(p)} \tag{11a}$$

and equilibrium is reached when

$$\overline{v^2(p)} = -\overline{\theta(p)v(p)}. \tag{11b}$$

A sufficient condition for (11) to be true is

$$\overline{v^2(p)} \geqq |\ \overline{\theta(p)v(p)}\ |.$$

From the definition of $\overline{[\cdot]}$ [equation (8)] and applying the Schwarz inequality, we obtain

$$|\ \overline{\theta(p)v(p)}\ | \leqq \overline{|\ \theta(p)v(p)\ |} \leqq (\overline{v^2(p)})^{\frac{1}{2}}(\overline{\theta^2(p)})^{\frac{1}{2}}.$$

From the above two inequalities we see that a looser but still sufficient condition for (11a) and (11b) to be valid is

$$\overline{v^2(p)} \geqq (\overline{v^2(p)})^{\frac{1}{2}}[\overline{\theta^2(p)}]^{\frac{1}{2}}$$

or

$$\overline{v^2(p)} \geqq \overline{\theta^2(p)}. \tag{12}$$

Equation (12) is a sufficient condition to insure convergence in the pth interval. We therefore conclude that when $|K|$ is made infinitesimally small the echo canceller will reduce $\sum_{i=1}^{M} \overline{r_i^2(p)}$ at least until:

$$\overline{v^2(p)} \leqq \overline{\theta^2(p)}. \tag{13}$$

We will now use equations (13) and (11a) and (11b) to determine a lower bound on the achievable suppression. Referring to Fig. 1 we define the suppression S as

$$S \equiv 10 \log_{10} \frac{\overline{y^2(p)}}{\overline{(e'(p))^2}}. \tag{14}$$

From equations (3) and (6) it may be easily shown that

$$e'(k) = v(k) + \theta(k)$$

or

$$\overline{(e'(p))^2} = \overline{v^2(p)} + 2\overline{v(p)\theta(p)} + \overline{\theta^2(p)}.$$

Substituting equation (11b) into the above equation we obtain

$$\overline{(e'(p))^2} = \overline{v(p)\theta(p)} + \overline{\theta^2(p)}.$$

Also, since

$$\overline{(e'(p))^2} \leqq |\overline{v(p)\theta(p)}| + \overline{\theta^2(p)}, \tag{15}$$

we can apply Schwarz's inequality to the above and obtain

$$\overline{(e'(p))^2} \leqq (\overline{v^2(p)})^{\frac{1}{2}}(\overline{\theta^2(p)})^{\frac{1}{2}} + \overline{\theta^2(p)}.$$

Substituting (13) and the above into (14) we obtain

$$S \geqq 10 \log_{10} \frac{\overline{y^2(p)}}{2\overline{\theta^2(p)}} = S_1 . \tag{16}$$

On the other hand, if for a given input signal

$$\overline{v(p)\theta(p)} = 0,^{\dagger}$$

we see from (15) and (14) that

$$S \geqq 10 \log_{10} \frac{\overline{y^2(p)}}{\overline{\theta^2(p)}} = S_2 . \tag{17}$$

Equations (16) and (17) are lower bounds on the achievable suppression which occurs when $|K|$ is chosen to be infinitesimal. For any other than infinitesimal $|K|$ the suppression may or may not be less than the values given by the lower bound [(16) and (17)]. However, as will be seen in a later section of this paper, for the choices of $|K|$ which allow reasonable settling times, (16) and (17) do form lower bounds on suppression.

III. NONLINEAR ECHO PATHS

We will now become more specific and consider the effect of a nonlinear echo path. We will restrict our discussion to a class of nonlinear echo paths which possess the following four properties:‡

(*i*) They are time invariant.
(*ii*) They are deterministic.
(*iii*) They are "smooth." Qualitatively speaking by smooth we mean that the echo path cannot introduce any abrupt or switch-like changes in the output. If such a change is evident in the output then it must be due to a similar switch-like change in the input.
(*iv*) They possess noninfinite memory. That is the memory does not depend on the remote past.

† For example, this occurs when θ and v are independent and one is zero mean. As will be seen this is the case for systems exhibiting phase jitter.

‡ Most echo paths encountered in practice satisfy all the conditions except one and these are handled separately in Section IV.

Echo paths which possess the above four properties may be conveniently characterized by a Volterra integral equation.[2-6] As a result, $y(t)$ shown in Fig. 1 is related to $x(t)$ by the equation given below:

$$y(t) = \int_0^\infty h_1(\tau_1)x(t-\tau_1)\,d\tau_1 + \int_0^\infty \int_0^\infty h_2(\tau_1, \tau_2)x(t-\tau_1)x(t-\tau_2)\,d\tau_1\,d\tau_2 + \sum_{n=0}^{\infty} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{n} h_n(\tau_1 \cdots \tau_n) \prod_{i=1}^{n} x(t-\tau_i)\,d\tau_i \,. \tag{18}$$

We may easily place equation (18) in the form of equation (2):

$$y(t) = x(t) * h_1(t) + \sum_{n=2}^{\infty} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{n} h_n(\tau_1 \cdots \tau_n) \prod_{i=1}^{n} x(t-\tau_i)\,d\tau_i \,.$$

Comparing the above equation with equation (2) we see that

$$\psi(t) = \sum_{n=2}^{\infty} \underbrace{\int_0^\infty \cdots \int_0^\infty}_{n} h_n(\tau_1 \cdots \tau_n) \prod_{i=1}^{n} x(t-\tau_i)\,d\tau_i \,.$$

As a result we conclude that the bounds given by equations (16) and (17) are applicable to any nonlinear system possessing the properties outlined in the beginning of this section. We need only replace $\theta(k)$ by

$$\theta(k) = \sum_{i=M+1}^{\infty} C_i w_i(k) + \sum_{n=2}^{\infty} \underbrace{\int \cdots \int}_{n} h_n(\tau_1 \cdots \tau_n) \prod_{i=1}^{n} x(k-\tau_i)\,d\tau_i \,. \tag{19}$$

In this case, $\theta(k)$ is composed of that part of the linear portion of the echo path which the echo canceller cannot compensate for due to its limited number of taps and the nonlinear portion.

IV. TIME VARIABILITY

In this section we are specifically interested in time variability of the type predominantly found on single-sideband suppressed carrier systems and commonly referred to as incidental FM or phase jitter.†

† We use incidental FM and phase jitter interchangeably.

We will show that the effect of incidental FM on suppression can be calculated directly from the index of modulation, β, which is easily measured.

4.1 *Carrier System Model–Assumptions*

In Fig. 2 we show a block diagram of a typical single-sideband suppressed carrier system whose output will be corrupted by phase

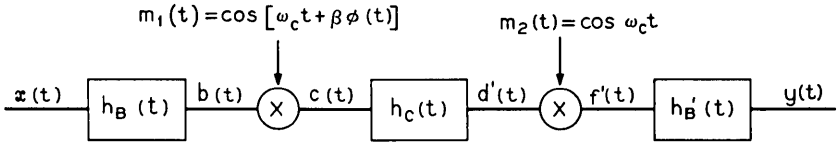


Fig 2—Single-sideband suppressed carrier system with phase jitter.

jitter. We will begin by making the following practical assumptions:

(*i*) The input signal $x(t)$ is bandlimited to ω_s. That is,

$$X(\omega) = \begin{cases} X(\omega) & \text{for} \quad |\omega| \leqq \omega_s \\ 0 & \text{otherwise.} \end{cases}^{\dagger}$$

(*ii*) The filter $h_B(t)$ is a low-pass filter bandlimited to ω_B,

$$H_B(\omega) = \begin{cases} H_B(\omega) & \text{for} \quad |\omega| < \omega_B \quad \text{where} \quad \omega_B > \omega_s + \omega_i \\ 0 & \text{otherwise.} \end{cases}$$

(*iii*) The incidental FM is narrow-band FM,

$$|\beta\phi(t)| \ll \pi/2 \quad \forall t$$

and

$$\phi(t) = \cos \omega_i t \qquad \omega_i \ll \omega_s.$$

(*iv*) The filter $h_c(t)$ is an ideal bandpass filter, at carrier frequencies, which passes the upper sideband,

$$H_c(\omega) = \begin{cases} 1 & \text{for} \quad \omega_c \leqq |\omega| \leqq \omega_c + \omega_s + \omega_i \,, \; \omega_c \gg \omega_B \\ 0 & \text{otherwise.} \end{cases}$$

---

† Capital letters are reserved for functions of frequency, small letters for functions of time.

$\therefore$ $x(t)$ has the Fourier Transform $X(\omega)$.

(v) The filter $h_{B'}(t)$ is a low-pass filter bandlimited to ω_B,

$$H_{B'}(\omega) = \begin{cases} H_{B'}(\omega) & \text{for } |\omega| < \omega_B \quad \text{where } \omega_B > \omega_s + \omega_i \\ 0 & \text{otherwise.} \end{cases}$$

4.2 *Derivation of the Relationship Between $x(t)$ and $y(t)$*

Consider Fig. 3 for a moment. $H_+(\omega)$ is an ideal, not physically realizable, filter which passes only positive frequencies. That is,

$$H_+(\omega) = \begin{cases} 1 & \omega \geqq 0 \\ 0 & \omega < 0. \end{cases} \tag{20}$$

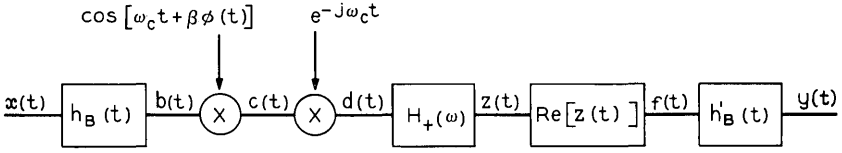


Fig. 3—Mathematically equivalent to Fig. 2.

and Re$(\cdot)$ signifies the real part of $(\cdot)$. It may be verified by the reader that Fig. 3 is mathematically equivalent to Fig. 2 as far as $x(t)$ and $y(t)$ are concerned. Therefore, we will use it for our analysis since it reduces the manipulations required. From Fig. 3,

$$c(t) = \frac{b(t)}{2}\left[e^{j(\omega_c t+\beta\phi(t))} + e^{-(j\omega_c t+\beta\phi(t))}\right] \tag{21}$$

and

$$d(t) = \frac{b(t)}{2} e^{j\beta\phi(t)} + \frac{b(t)}{2} e^{-j\beta\phi(t)} e^{-j2\omega_c t}. \tag{22}$$

Since the second term in equation (22) is a high-frequency term and its effect will be filtered out by the filter $h_{B'}(t)$ we will disregard it. Using the assumption $|\beta\phi(t)| \ll \pi/2$ we expand $e^{jB\phi(t)}$ obtaining

$$d(t) = \frac{b(t)}{2}(1 + j\beta\phi(t)). \tag{23}$$

Continuing, we obtain for $z(t)$ †

$$z(t) = \tfrac{1}{2}[d(t) + j\,\widehat{d(t)}] = \frac{b(t)}{4} - \frac{\beta\phi(t)\widehat{b(t)}}{4} + j\left(\frac{\widehat{b(t)}}{4} + \frac{\beta\phi(t)b(t)}{4}\right). \text{ ‡} \tag{24}$$

† The reader is referred to Ref. 7 for the proof of equation (24).

‡ $\widehat{[\cdot]}$ denotes Hilbert transform of $[\cdot]$.

Therefore,

$$\theta(t) = 1/4(b(t) - \beta\widehat{\phi(t)b(t)}) \tag{25}$$

and

$$y(t) = 1/4(b(t)*h_{B'}(t) - \beta(\widehat{\phi(t)b(t)})*h_{B'}(t)). \tag{26}$$

Comparing equations (2) and (26) we conclude that they are of the same form:

$$h(t) = 1/4h_B(t)*h_{B'}(t)$$

and

$$\psi(t) = 1/4\beta(\widehat{\phi(t)b(t)})*h_{B'}(t).$$

As a result the bounds given by equations (16) and (17) are applicable with $\theta(t)$ given by

$$\theta(t) = \sum_{i=m+1}^{\infty} C_i w_i(t) - 1/4\beta(\widehat{\phi(t)b(t)}) * h_{B'}(t). \tag{27}$$

However, unlike the nonlinear problem, we may simplify the above results considerably as shown below.

4.3 *Simplification of Bound for Systems Displaying Incidental FM*

If we assume that the basis set $\{\lambda_i(k)\}$ is complete we obtain from equation (27) that

$$\theta(t) = -1/4\,\beta(\widehat{\phi(t)b(t)})*h_{B'}(t).$$

For the situation where incidental FM is present we have found experimentally that quantity $\overline{v(p)\theta(p)}$ [see equation (6)] may be safely assumed to be zero. Therefore the bound S_2 given by equation (17) is applicable.

For bandlimited signals sampled at the Nyquist rate,

$$\frac{\overline{y^2(p)}}{\overline{\theta^2(p)}} \approx \frac{\lim_{\tau_1\to\infty} 1/\tau_1 \int_0^{\tau_1} y^2(t)\,dt}{\lim_{\tau_2\to\infty} 1/\tau_2 \int_0^{\tau_2} \theta^2(t)\,dt}.$$

Since the effect of incidental FM on the circuits that we are dealing with is a second-order effect and is quite small, it is reasonable to assume that

$$\overline{|\,b(t) * h_{B'}(t)\,|} \gg \overline{|\,\beta(\widehat{\phi(t)b(t)}) * h_{B'}(t)\,|}.$$

Therefore,

$$\lim_{\tau_1\to\infty} 1/\tau_1 \int_0^{\tau_1} y^2(t)\,dt \approx \frac{1}{16}\left(\lim_{\tau_1\to\infty} 1/\tau_1 \int_0^{\tau_1} (b(t) * h_{B'}(t))^2\,dt\right)$$

resulting in

$$\frac{\overline{y^2(p)}}{\overline{\theta^2(p)}} \approx \frac{1}{\beta^2} \frac{\lim_{\tau_1\to\infty} 1/\tau_1 \int_0^{\tau_1} [b(t) * h_{B'}(t)]^2}{\lim_{\tau_2\to\infty} 1/\tau_2 \int_0^{\tau_1} [\widehat{(\phi(t)b(t))} * h_{B'}(t)]^2\,dt} = \frac{1}{\beta^2} \frac{\int_{-\omega_s-\omega_j}^{\omega_s+\omega_j} S_y(\omega)\,d\omega}{\int_{-\omega_s-\omega_j}^{\omega_s+\omega_j} S_\theta(\omega)\,d\omega} \tag{28}$$

where $S_y(\omega)$ and $S_\theta(\omega)$ are proportional to the power spectral density of $y(t)$ and $\theta(t)$ respectively.

We will now show that the two integral expressions in (28) are approximately equal. In the frequency domain we have

$$\begin{aligned} S_y(\omega) &= |\, B(\omega)\,|^2\,|\, H_{B'}(\omega)\,|^2, \\ S_\theta(\omega) &= |\, \mathfrak{F}[\widehat{\phi(t)b(t)}]H_{B'}(\omega)\,|^2 \,^{\dagger} \\ &= |\, \mathfrak{F}[\widehat{\phi(t)b(t)}]\,|^2\,|\, H_{B'}(\omega)\,|^2. \end{aligned} \tag{29}$$

The Hilbert transform does not affect the magnitude of the Fourier transform. As a result we obtain

$$S_\theta(\omega) = |\, \mathfrak{F}[\phi(t)b(t)]\,|^2\,|\, H_{B'}(\omega)\,|^2. \tag{30}$$

We will now apply assumption *iii* and assume that the phase jitter function $\phi(t)$ is a cosine wave,‡

$$\phi(t) = \cos \omega_j t$$

where ω_j is typically either 60 Hz, 80 Hz, or 120 Hz. Equation (30) then becomes

$$S_\theta(\omega) = 1/4\,|\, B(\omega - \omega_j) + B(\omega + \omega_j)\,|^2\,|\, H_{B'}(\omega)\,|^2. \tag{31}$$

Since the effect of $\phi(t)$ is only to shift the spectrum $B(\omega)$ to the right and left a very small amount§ relative to its bandwidth,§ it should be clear from equations (29) and (31) that

$$\int_{-\omega_s-\omega_j}^{\omega_s+\omega_j} S_y(\omega)\,d\omega \approx \int_{-\omega_s-\omega_j}^{\omega_s+\omega_j} S_\theta(\omega)\,d\omega. \tag{32}$$

† $\mathfrak{F}[\phi(t)b(t)]$ indicates the Fourier transform of $\phi(t)b(t)$.
‡ In practice this is a good assumption.
§ Recall $\omega_s \gg \omega_j$. Typically ω_s is 3500 Hz and $\omega_j \leqq 120$ Hz.

Using equations (28) and (32) in (17) we obtain

$$S \geqq 20 \log_{10} \frac{1}{\beta}.^{\dagger} \tag{33}$$

Equation (33) is what we have been striving for; it relates achievable suppression to the modulation coefficient of the unwanted incidental FM for an echo canceller with a complete basis set.

V. EMPIRICAL RESULTS

5.1 *Time Variant Systems (Incidental FM)*

Since it is difficult to accurately create a desired amount of incidental FM on a working L carrier system, an analog simulation was used. For all practical purposes it is identical to a real system. The only difference is that it allows a controlled amount of incidental FM to be inserted.

We chose flatly weighted noise bandlimited to 4 kHz as our input signal. A digital tape was made of the input and output of the L terminal simulator and this tape was used as input to a computer simulation of the adaptive echo canceller shown in Fig. 1. The results are shown in Fig. 4.

From Fig. 4 we see that there is good agreement between the empirical results and the analytical results for peak-to-peak phase jitter angles above 4 degrees. Below 4 degrees the echo canceller performance is limited by the fact that it utilizes a finite number of taps. Recall in our derivation of equation (33) we assumed that the echo canceller employed a complete basis set (i.e., infinite number of taps). This assumption begins to fail in the vicinity of 4 degrees peak-to-peak phase jitter and the performance of the echo canceller becomes limited by the incompleteness rather than by the incidental FM.

5.2 *Nonlinear Systems*

In order to verify our results we simulated the nonlinear echo path shown in Fig. 5. We used various combinations of input signals and feedback gain constants $|K|$ and in all cases the results were comparable to those shown in Table I for a white noise input.

A word of explanation is needed to clarify Table I. Suppression was calculated according to equation (14). The sampling rate was 0.1 ms and the averaging interval, Q, consisted of 501 samples, i.e., a 50.1-ms interval.‡ Forty such averages were computed during the two-

† β must be in radians.
‡ See equation (8).

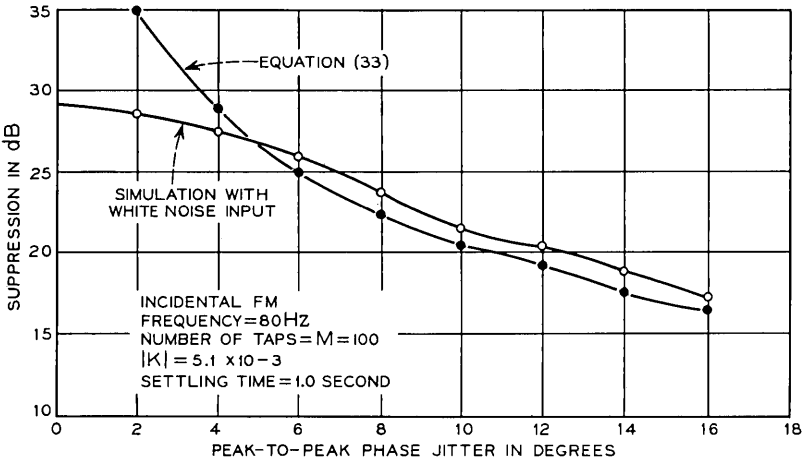


Fig. 4—Time variant simulation.

second time period after the settling time had elasped. The minimum and maximum suppression shown are those of the forty calculated suppressions, and the average suppression shown is the ensemble average of the forty calculated suppressions.

For all variation of input signals and feedback gain constants tested, we found that S_2 was 3 dB closer to the actual suppression than S_1 and within 5 dB of the actual suppression. However, as expected, the echo canceller always performed better than either bound.

VI. CONCLUDING REMARKS

We have obtained lower bounds on the suppression provided by an echo canceller in either a nonlinear or time invariant environment. Although theoretically the bounds are only valid when an infinitesimal feedback constant is used, we have found empirically that they apply for any practical choice of feedback gain constant. If one were able

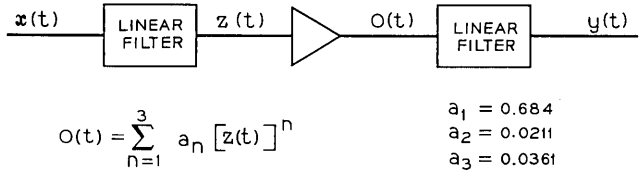


Fig. 5—Nonlinear simulation.

TABLE I—WHITE NOISE INPUT SIGNAL

Feedback gain $\|K\|$	5.1×10^{-3}
Settling time	1.0 second
Actual achieved suppression in dB	Min. = 18.2, max. = 22.1, ave. = 20.4
Bound S_1 [equation (16)]	Min. = 10.6, max. = 15.1, ave. = 13.7
Bound S_2 [equation (17)]	Min. = 13.6, max. = 18.1, ave. = 16.7

to measure the time variability of echo paths or to measure echo paths and then to separate the response into the linear and nonlinear portions [equation (18)], one could predict the minimum achievable suppression. Measurement of nonlinear echo paths has been the subject of study of others.[8,9]

For a nonlinear environment we have found that the nonlinear portion of the response and hence the lower bounds are dependent on the input signal. As a result, if this method is going to be used to predict the performance of an echo canceller for a speech input signal, it will be necessary to design an interrogation signal which possesses some of the same properties as does speech (syllabic rate, power level, peak-to-rms ratio, etc.)

The derived equations which relate the effect of incidental FM on suppression for the single-sideband suppressed carrier systems are sufficiently general to provide lower suppression bounds for different or more complex systems which exhibit incidental FM. One must merely know an effective index of modulation, β, measured between the input and the output. The calculations for more complex systems are rather involved and have been omitted.

Laboratory simulated echo cancellation for a nonlinear echo path and for an echo path with incidental FM produced cancellations somewhat better than the calculated lower bounds. Knowledge of the type of nonlinearity or time variability in the echo path environment may permit one to empirically increase the lower bounds derived in this paper.

VII. ACKNOWLEDGMENT

The author wishes to thank I. Jacobs and K. E. Fultz for their many helpful comments.

REFERENCES

1. Rosenberger, J. R., and Thomas, E. J., "Performance of an Adaptive Echo Canceller in a Noisy, Linear, Time Invariant Environment," B.S.T.J., *50*, No. 3. (March 1971), pp. 785–813.

2. Liou, Ming-Lei, "Nonlinear System Analysis and Synthesis," Technical Report No. 6554-6, October 1963, Systems Theory Laboratory, Stanford University, California.
3. Smits, H. B., "Analysis and Synthesis of Nonlinear Systems," IRE Trans. Circuit Theory, December 1960, pp. 459–469.
4. Van Trees, H. L., *Synthesis of Optimum Nonlinear Control Systems,* Cambridge, Massachusetts: M.I.T. Press, 1962.
5. Volterra, V., *Theory of Functionals,* London: Blackie and Son, 1930.
6. Parente, R. B., "Functional Analysis of Systems Characterized by Nonlinear Differential Equations," Technical Report 444, M.I.T. Research Laboratory of Electronics, July 15, 1966.
7. Rowe, H. E., *Signals and Noise in Communication Systems,* New York: D. Van Nostrand Co., 1965, p. 17.
8. Unrue, J. E., Jr., unpublished work.
9. Miller, G., unpublished work.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Some Considerations on the Application of the Volterra Representation of Nonlinear Networks to Adaptive Echo Cancellers

By E. J. THOMAS

(Manuscript received October 19,1970)

In this paper we apply the Volterra representation of nonlinear systems to the echo control problem and propose a generalized adaptive echo canceller which compensates for nonlinear echo paths. We prove that the proposed echo canceller converges and reduces the echo to zero and finally we suggest other applications for the system.

I. INTRODUCTION

The Volterra functional representation of a nonlinear system is a generalization of the well known convolution integral used for linear systems. The validity conditions for the Volterra model are sufficiently weak to be satisfied in many practical applications.

In this paper we apply the Volterra representation to the echo control problem and propose a generalized adaptive echo canceller which compensates for nonlinear echo paths. We prove that the proposed echo canceller converges and reduces the echo to zero in the absence of noise. We also suggest other applications for the adaptive echo canceller presented.

II. THE VOLTERRA FUNCTIONAL REPRESENTATION OF A NONLINEAR 2-PORT

For linear systems it is well known that the impulse response completely determines the input-output relationship. The output signal, $y(t)$, is functionally related to the input, $x(t)$, by the convolution integral,

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)\, d\tau, \tag{1}$$

where $h(t)$ is the system impulse response. It has been demonstrated by other investigators[1-3] that nonlinear systems, whose outputs do

not depend on the infinite past, obey a more general functional relationship,

$$y(t) = h_0(t) + \int_{-\infty}^{\infty} h_1(\tau_1)x(t - \tau_1)\, d\tau_1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1 , \tau_2)x(t - \tau_1)x(t - \tau_2)\, d\tau_1\, d\tau_2 + \sum_{n=3}^{\infty} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n} h_n(\tau_1 , \cdots , \tau_n) \prod_{i=1}^{n} x(t - \tau_i)\, d\tau_i . \quad (2)$$

This is an extension of the familiar power series representation of a memoryless nonlinear system and provides for the system to have memory. It is applicable to all nonlinear systems whose outputs depend on the remote past to a vanishingly small extent. For example, it is not applicable to hysterisis and/or switching systems. The terms of (2) were first studied by Volterra[4] and are called Volterra functionals. The kernels, $h_n(\tau_1 , \cdots , \tau_n)$, are generally called Volterra kernels.

For the systems that we are concerned with we will assume the following:

(*i*) The zero input response, $h_0(t)$, is identically zero.
(*ii*) The system is causal so that $h_n(\tau_1, \tau_2, \cdots , \tau_n) = 0$ for any $\tau_i < 0$.
(*iii*) The system is stable so that for all n,

$$\underbrace{\int_0^{\infty} \cdots \int_0^{\infty}}_{n} [h_n(\tau_1 , \cdots , \tau_n)]^2\, d\tau_1 , d\tau_2 , \cdots , d\tau_n < \infty .$$

Because of i we may rewrite (2) in the form

$$y(t) = \sum_{n=1}^{\infty} y_n(t), \quad (3)$$

where

$$y_n(t) = \underbrace{\int_0^{\infty} \cdots \int_0^{\infty}}_{n} h_n(\tau_1 , \cdots , \tau_n) \prod_{i=1}^{n} [x(t - \tau_i)\, d\tau_i]. \quad (4)$$

III. AN ADAPTIVE ECHO CANCELLER FOR NONLINEAR SYSTEMS

We will now generalize the adaptive echo canceller proposed by Sondhi[5] to cancel the echo of a nonlinear echo path for which the Volterra representation is applicable.

Since the Volterra kernels are square integrable we may represent them by an n dimensional generalized Fourier series,[6-9]

$$h_n(\tau_1 , \tau_2 , \cdots , \tau_n) = \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} C_{i_1, i_2, \cdots, i_n} \cdot \Gamma_{i_1}(\tau_1)\Gamma_{i_2}(\tau_2) \cdots \Gamma_{i_n}(\tau_n), \quad (5)$$

where $\{\Gamma_i(t)\}$* is an orthonormal set complete in the $L2$ space. The coefficients of (5) are given by

$$C_{i_1, i_2, \cdots, i_n} = \int_0^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} h_n(\tau_1 , \tau_2 , \cdots , \tau_n) \cdot \Gamma_{i_1}(\tau_1)\Gamma_{i_2}(\tau_2) \cdots \Gamma_{i_n}(\tau_n)\, d\tau_1 \cdots d\tau_n . \quad (6)$$

We will assume that the highest order nonlinearity is of order N. Then substituting (5) into (3) we obtain

$$y(t) = \sum_{n=1}^{N} \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} C_{i_1, i_2, \cdots, i_n} \int_0^{\infty} \Gamma_{i_1}(\tau_1)x(t - \tau_1)\, d\tau_1 \cdot \int_0^{\infty} \Gamma_{i_2}(\tau_2)x(t - \tau_2)\, d\tau_2 \cdots \int_0^{\infty} \Gamma_{i_n}(\tau_n)x(t - \tau_n)\, d\tau_n . \quad (7)$$

Before we proceed it will be convenient to adopt a shorthand notation. We will define

$$\sum_{I_n} (\cdot) \equiv \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} (\cdot), \quad (8)$$

$$C_{I_n} \equiv C_{i_1, i_2, \cdots, i_n} , \quad (9)$$

and

$$w_{i_n} \equiv \int_0^{\infty} \Gamma_{i_n}(\tau_n)x(t - \tau_n)\, d\tau_n .^{\dagger} \quad (10)$$

Thus, (7) becomes

$$y(t) = \sum_{n=1}^{N} \sum_{I_n} C_{I_n} w_{i_1} w_{i_2} \cdots w_{i_n} . \quad (11)$$

Now consider the system shown in Fig. 1. It contains N subsystems designated by the circled numbers. The filters having the set of mutually orthonormal impulse responses $\{\Gamma_{i_1}(t)\}$ and the set of outputs $\{w_{i_1}\}$

* Some typical sets $\{\Gamma_n(t)\}$ are the impulse responses of tapped delay lines or Laguerre networks.

† Note that w_{i_n} is a function of t although it is not explicitly shown.

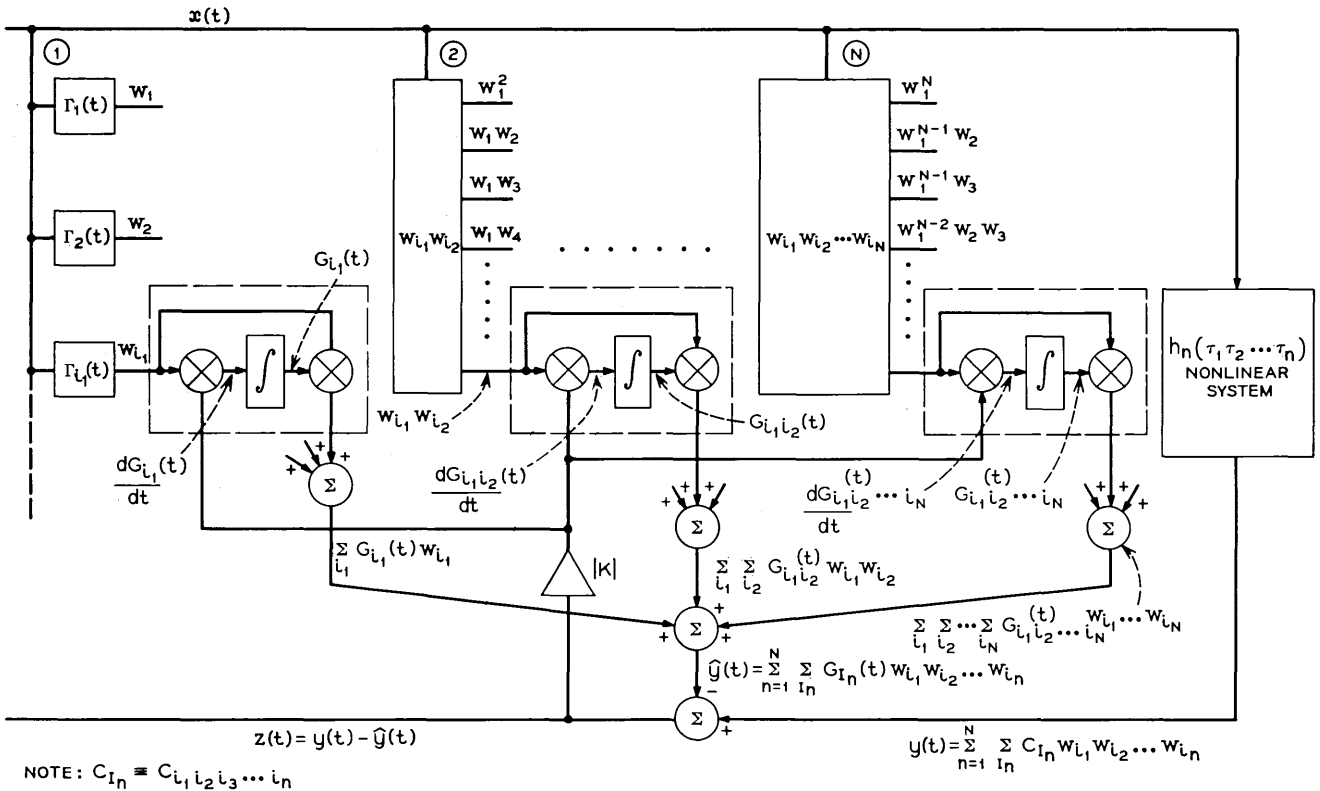


Fig. 1—An adaptive echo canceller for nonlinear systems.

due to the input $x(t)$ are common to all the subsystems. Subsystem 1 produces the output of the filters taken one at a time. Subsystem 2 produces all possible combinations of the product of the filter outputs taken two at a time (i.e., $w_{i_1}w_{i_2}$), and subsystem N produces all possible combinations of the filter outputs taken N at a time (i.e., $w_{i_1}w_{i_2}w_{i3}\cdots w_{i_n}$). Every output tap of every subsystem has associated with it two multipliers and an integrator connected in the configuration shown in Fig. 1 for only three taps. Other quantities pertinent to the discussion below are also defined in Fig. 1.

We will now show that the tap gains, $G_{I_n}(t)$, of the system converge to the generalized Fourier coefficients C_{I_n} of (6) and, consequently, that $\hat{y}(t)$ converges to $y(t)$ so that the residual echo, $z(t)$, vanishes.

We first define

$$R_{I_i}(t) \equiv C_{I_i} - G_{I_i}(t). \tag{12}$$

Then, by inspection, the equation governing the gain of each of the taps of the pth subsystem is

$$\frac{dG_{I_p}(t)}{dt} = |K| w_{i_1}w_{i_2}\cdots w_{i_p}\sum_{n=1}^{N}\sum_{I_n} R_{I_n}(t)w_{i_1}w_{i_2}\cdots w_{i_n}.^{*} \tag{13}$$

From (12) we see that

$$\frac{dR_{I_i}(t)}{dt} = -\frac{dG_{I_i}(t)}{dt}. \tag{14}$$

Also,

$$\frac{dR_{I_i}^2(t)}{dt} = 2R_{I_i}(t)\frac{dR_{I_i}(t)}{dt}. \tag{15}$$

Applying (14) and (15) to (13) we obtain

$$\frac{dR_{I_p}^2(t)}{dt} = -2|K|R_{I_p}(t)w_{i_1}w_{i_2}\cdots w_{i_p}\sum_{n=1}^{N}\sum_{I_n} R_{I_n}(t)w_{i_1}w_{i_2}\cdots w_{i_n}. \tag{16}$$

Summing over all the taps within the pth subsystem, and over the N subsystems, yields

$$\frac{d}{dt}\sum_{p=1}^{N}\sum_{I_p} R_{I_p}^2(t) = -2|K|\left(\sum_{n=1}^{N}\sum_{I_n} R_{I_p}(t)w_{i_1}w_{i_2}\cdots w_{i_n}\right)^2. \tag{17}$$

Note that the right-hand side of (17) is always negative or zero.

* Where the set $(i_1i_2\cdots i_p)$ are the components of I_p.

Also note that

$$\psi(t) \equiv \sum_{p=1}^{N} \sum_{I_p} R_{I_p}^2(t) \geqq 0. \tag{18}$$

Thus, we have a nonnegative function, whose derivative is always nonpositive, and conclude that $\psi(t)$ must be nonincreasing. It is strictly decreasing whenever

$$\sum_{n=1}^{N} \sum_{I_n} R_{I_n}(t) w_{i_1} w_{i_2} \cdots w_{i_n} \not\equiv 0.$$

Since, in general

$$w_{i_1} w_{i_2} \cdots w_{i_n} \not\equiv 0,$$

it is clear that

$$\lim_{t \to \infty} \psi(t) = 0. \tag{19}$$

Since every term in $\psi(t)$ is positive, we further conclude that

$$\lim_{t \to \infty} R_{I_p}(t) = 0 \quad \text{for} \quad p = 1, 2, \cdots, N. \tag{20}$$

Applying (20) to (12) we see that

$$\operatorname*{limit}_{t \to \infty} G_{I_p}(t) = C_{I_p}; \qquad p = 1, 2, \cdots, N. \tag{21}$$

Thus we have shown that the tap gains converge to the generalized Fourier coefficients of $h_n(\tau_1, \cdots, \tau_n)$, which is what we set out to prove.

The reader can no doubt appreciate that solving (17) for anything but the final value would be difficult and would only be valid for a specific $x(t)$. The solution for even a very simple input signal may not be tractable. We will not attempt such an analysis here. However, from (17) we see that the larger we make $|K|$ the quicker the echo vanishes. However, $|K|$ cannot be made arbitrarily large because doubling talking periods* will cause divergence problems.

IV. EFFECTS OF CIRCUIT NOISE AND OTHER CONSIDERATIONS

The previous analysis assumes a noiseless output signal $y(t)$. When circuit noise, $n(t)$, generated in the echo path is present equation (17) would be written as

* Doubling talking is said to occur when both speakers in a telephone conversation are speaking simultaneously.

$$\frac{d}{dt}\sum_{j=1}^{N}\sum_{I_p}R_{I_p}^2(t) = -2\mid K\mid\left(\sum_{n=1}^{N}\sum_{I_n}R_{I_n}(t)w_{i_1}w_{i_2}\cdots w_{i_n}\right)^2$$
$$-2\mid K\mid n(t)\left(\sum_{n=1}^{N}\sum_{I_n}R_{I_n}(t)w_{i_1}w_{i_2}\cdots w_{i_n}\right).$$

By choosing $|K|$ small so that the rate of change of the G_{I_p} is small, it is reasonable to assume that $n(t)$ is statistically independent of all other variables in the equation. Then, assuming $n(t)$ has zero mean, upon taking the ensemble average of both sides of the above equation we obtain

$$\frac{d}{dt}\left(\left\langle\sum_{j=1}^{N}\sum_{I_p}R_{I_p}^2(t)\right\rangle\right)$$
$$= -2\mid K\mid\left\langle\left(\sum_{n=1}^{N}\sum_{I_n}R_{I_n}(t)w_{i_1}w_{i_2}\cdots w_{i_n}\right)^2\right\rangle. \tag{22}$$

The same argument used before then yields

$$\lim_{t\to\infty}\left\langle\sum_{j=1}^{N}\sum_{I_p}R_{I_p}^2(t)\right\rangle = 0. \tag{23}$$

Of course, before this limit is reached our statistically independent assumption would no longer be valid, but at least we may conclude that convergence will take place until the noise power becomes comparable to the echo power.

There are several other difficulties associated with the echo canceller. Recall that we represented the nth-order Volterra kernel by an "n" fold infinite summation [equation (5)]. Obviously, if such a system is to be implemented it cannot include an infinite number of elements and (5) can only be approximated by

$$h_n(\tau_1, \cdots, \tau_n) \approx \sum_{i_1}^{J_1}\sum_{i_2}^{J_2}\cdots\sum_{i_n}^{J_3}G_{i_1,i_2,\cdots,i_n}\Gamma_{i_1}(\tau_1)\Gamma_{i_2}(\tau_2)\cdots\Gamma_{i_n}(\tau_n), \tag{24}$$

where the J's are finite and as small as possible for an adequate description.

Furthermore, even if the J's are small the system shown in Fig. 1 requires a large number of components. However, one point may not be obvious. It can be shown that the Volterra kernels are symmetrical.[2] As a result it may be verified that all the $C_{i_1 i_2 \cdots i_n}$ are equal for any permutation of a specific set of numbers $i_1, i_2, \cdots, i_n$. For example,

$$C_{123} = C_{231} = C_{312} = C_{213}, \text{ etc..}$$

Since the set of gains, $G_{i_1 \cdots i_n}$, for any combination of a specific set of integers $i_1 \cdots i_n$ are the result of the same physical operations, only one tap is required to account for all such terms. For example, consider the w_1w_2 tap and the w_2w_1 tap. From Fig. 1,

$$G_{12} = \int w_1w_2[y(t) - \hat{y}(t)]\, dt,$$

$$G_{21} = \int w_2w_1[y(t) - \hat{y}(t)]\, dt.$$

It is obvious that

$$G_{21} = G_{12}.$$

The contribution of G_{12} and G_{12} taps toward $\hat{y}(t)$ is given by

$$\hat{y}(t)\,|_{12} = G_{12}w_1w_2 + G_{21}w_2w_1 = (G_{12} + G_{21})w_{12} = 2G_{12}w_{12}.$$

Thus, the G_{12} and G_{21} taps may be replaced by a single tap. Any set of taps, $G_{i_1 \cdots i_n}$, for any permutation of a specific set of numbers, $i_1 \cdots i_n$, can be replaced by a single tap. This considerably reduces the total number of taps required.

V. OTHER APPLICATIONS

Although we have stressed the echo cancellation application, these ideas may also be useful in other areas. Two possible uses are described below.

5.1 *Compensator for a Nonlinear System*

Assume that we wish to linearize the nonlinear system shown in Fig. 1, such that the resulting output, $z(t)$, can be expressed by the linear convolution integral,

$$z(t) = \int_0^\infty h_1(\tau_1)x(t - \tau_1)\, d\tau_1 \,. \tag{25}$$

This can be done by first allowing the adaptive system of Fig. 1 to converge long enough so that the members of the set $\{G_{I_j}(t)\}$ can be considered to equal the corresponding members of the set $\{C_{I_j}\}$ [equation (21)]. After convergence, the members of the set $\{G_{I_1}\}$ are forced to zero while the members of the sets $\{G_{I_j}\}$, $j \neq 1$ are fixed at the values determined previously. As a result the compensated output will satisfy equation (25).

5.2 *Nonlinear System Synthesizer*

Suppose one wishes to study the electrical characteristics of a nonlinear system which cannot be brought into the laboratory. He could do this by making input/output tape recordings of the system, and use these as an input to a computer simulation of the adaptive system of Fig. 1. A good choice of an input signal would be white noise or any other easily generated broadband signal. By allowing the simulation to converge and then fixing the tap gains G_{I_n} at their final value, the nonlinear characteristic can be identified. Then it can be determined how the field system will behave for any arbitrary input by applying this input to the computer simulation with the tap gains fixed at the values determined previously.

VI. SUMMARY AND CONCLUSIONS

The Volterra representation is a concise method of characterizing nonlinear systems with memory when the output does not depend on the infinite past. It is a generalization of the convolution integral used in linear analysis, and many familiar concepts may be extended. Using the Volterra representation, we have proposed a generalized adaptive echo canceller capable of nonlinear compensation. A disadvantage of the proposed echo canceller is its complexity. The pursuit of these ideas would be greatly enhanced if an efficient means of measuring the Volterra kernels of an unknown network could be found. The author and other researchers are presently working toward this end and several ideas have been proposed. To date, unfortunately, they all suffer from the complexity problem.

REFERENCES

1. Brilliant, M. B., *Theory of the Analysis of Nonlinear Systems,* Tech. Report 345, March 3, 1958, M.I.T. Research Laboratory of Electronics, Cambridge, Mass.
2. George, D. A., *Continuous Nonlinear Systems,* Tech. Report 355, July 24, 1958, M.I.T. Research Laboratory of Electronics, Cambridge, Mass.
3. Narayanan, S., "Transistor Distortion Analysis Using Volterra Series Representation," B.S.T.J., *46,* No. 5 (May-June 1967), pp. 991–1024.
4. Volterra, V., *Theory of Functionals,* London: Blackie and Son, 1930.
5. Sondhi, M. M., "An Adaptive Echo Canceller," B.S.T.J., *46,* No. 3 (March 1967), pp. 497–511.
6. Tolstov, G. P., *Fourier Series,* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1965, pp. 173–196.
7. Sneddon, I. N., *Fourier Transforms,* New York: McGraw-Hill, 1951, pp. 79-82.
8. Bateman Manuscript Project, Vol. 2, *Higher Transcendental Functions,* New York: McGraw-Hill, 1953, pp. 264–293.
9. Lee, Y. W., and Schetzen, M., "Measure of the Wiener Kernels of a Nonlinear System by Crosscorrelation," Int. J. Control, September 1965, pp. 237–254.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

Some Comparisons of Load and Loss Data With Current Teletraffic Theory*

By R. I. WILKINSON

(Manuscript received February 26, 1971)

Data are presented and compared with current traffic loss theory in three trunking areas: high-usage groups; full (nonalternate routed) groups; and final groups. Both single hour and average busy season busy hour load-loss comparisons are made. Methods of estimating offered loads from carried loads, and from the proportion of calls blocked, are considered. The blocking which results from offering nonrandom traffic to overflow groups is discussed. Modifications in the theory commonly applied are indicated in each case when necessary to obtain satisfactory agreement with observed values.

I. INTRODUCTION

The ultimate test of theory is comparison with reality. The procurement of reliable and reproducible traffic data is relatively difficult since little control can be exercised over the character and level of the input to the load carrying system without invalidating the genuineness of the input. There is then a continuing need to compare the commonly used load-loss relations with observations made under a variety of trunking configurations and load characteristics. Data are presented here and compared with loss theory in three trunking areas: high-usage groups; full (nonalternate routed) groups; and final groups. The various findings are summarized at the end of the paper.

II. HIGH-USAGE GROUP STUDIES

2.1 *Loads and Losses on High-Usage Groups*

It is commonly assumed that the requirements of a "random" or Poisson input are met until originating traffic has passed through a restrictive group of paths or switches. It would be difficult and cer-

* Substantially as presented at the Sixth International Teletraffic Congress, Munich, 1970.

tainly impractical in most operational situations to check such an assumption by examining the call arrival instants or to analyze their interarrival times, as well as make a corresponding study of their service times. Since in any event exact conformity with the theoretical assumptions could not be found, the question would remain as to the relative adequacy with which they were met.

The traffic engineer's usual wish is to describe the blocking which will occur in real-life situations when a given average load is offered to a group of paths or switches. It will generally suffice then to compare observed load versus loss relationships with those theoretically derived, rather than attempt an assessment of the agreement of more basic requirements.

The simplest trunking situation occurs when a parcel of traffic arising from an indefinite (but potentially large) number of customers is offered to a full access high-usage trunk group. A high-usage group is presumed here to receive calls quite directly from customers without earlier significant constriction and serves them immediately if paths are available. If, however, all of the high-usage trunks are busy, arriving calls are alternate routed to other groups in the hierarchy; and negligibly few calls return to the high-usage group for retrial. Thus the conditions for the application of Erlang's loss formula are apparently well satisfied. How well do observations made in actual exchanges agree with this theory?

Data of the above character have been obtained on many high-usage groups, both interlocal and intertoll. For making comparisons, hourly readings of three recording registers are commonly taken:

(*i*) Number of calls offered, m
(*ii*) Number of calls blocked, n
(*iii*) Average number of calls seen on trunks during the period by switches-in-use counts taken every 100 seconds, giving an estimate of the average load carried in erlangs, ℓ.

Observations made by A. Descloux[1] indicate that a substantially unbiased estimate a of the hourly offered load is obtained from

$$a = \ell \Big/ \left(1 - \frac{n}{m}\right). \tag{1}$$

Typical results of plotting proportion blocked, n/m, versus load offered a are shown in Figs. 1, 2, and 3. Figures 1 and 2 are for interlocal groups studied at an Arlington, Massachusetts, No. 5 crossbar office, and at

Kildare (Chicago), a No. 1 crossbar office. Figure 3 data are on intertoll groups at Memphis, Tennessee, a sectional center in the toll hierarchy.

We know that over a long period an equilibrium load of a erlangs will show a significant hour-to-hour variability in offered loads.[2] Also, when such variable loads are offered to a trunk group, we should expect the individual hour losses to tend to fall on the convex side of an arithmetic plot of the long-run theoretical load-loss curve, since when aggregated their average loss should just meet the theory. To examine this phe-

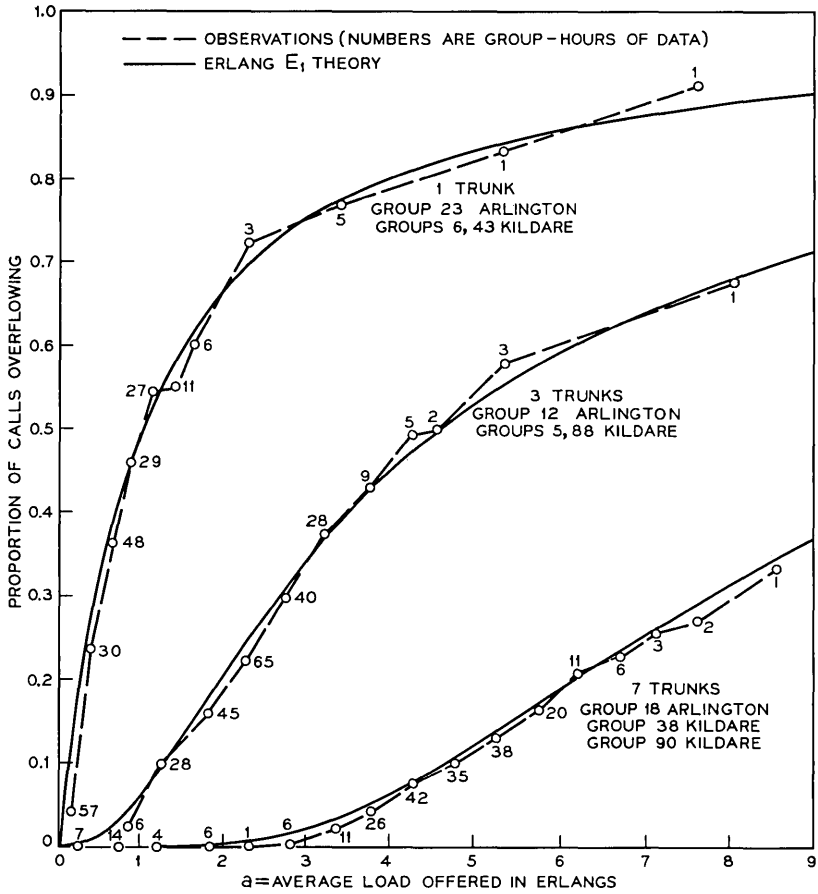


Fig. 1—Comparison of Erlang's loss theory with observations on interlocal high-usage trunk groups: busy hour, Arlington, 29 days; and busiest 3 hours, Kildare, 20 days, 1958.

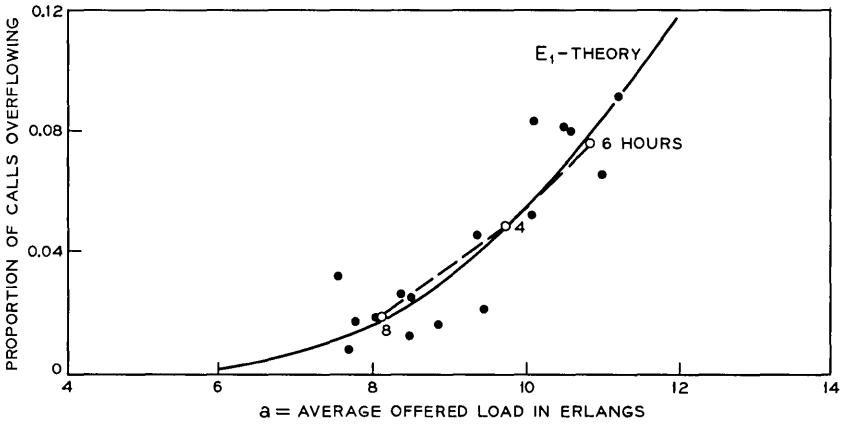


Fig. 2—Comparison of Erlang's loss theory with observations on interlocal high-usage trunks: Kildare group No. 12, 14 trunks, 11–12 A.M., 18 days, 1958. (• single hour. ° aggregated hourly values.)

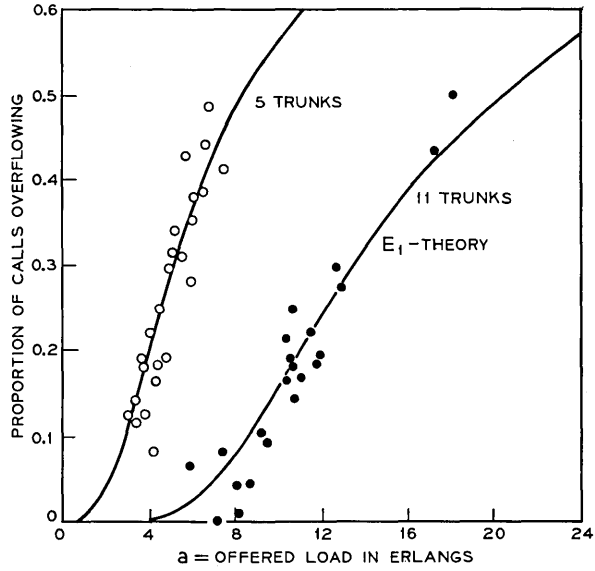


Fig. 3—Comparison of Erlang's loss theory with observations on intertoll high-usage trunk groups, Memphis, 1957–1958: 5 trunks to Milwaukee, 11 trunks to Cleveland.

nomenon under controlled conditions two simulations were run with a designed offer of 1.50 erlangs comprised of 180-second exponential holding time calls. The actual average offer observed was $\bar{a} = 1.485$ erlangs to 1 trunk for 400 hours, and 1.409 erlangs to 4 trunks for 100 hours. The results are displayed on Fig. 4. The individual hour losses, while conforming generally to the shape of the theoretical curve, appear to be slightly more linearly disposed. As longer summary periods are used, the observed values would be expected to approach the Erlang loss curve. This is seen clearly in the simulation results. A slight semblance of this tendency is seen on Figs. 1 thru 3, but it is considerably obscured by other fluctuations.

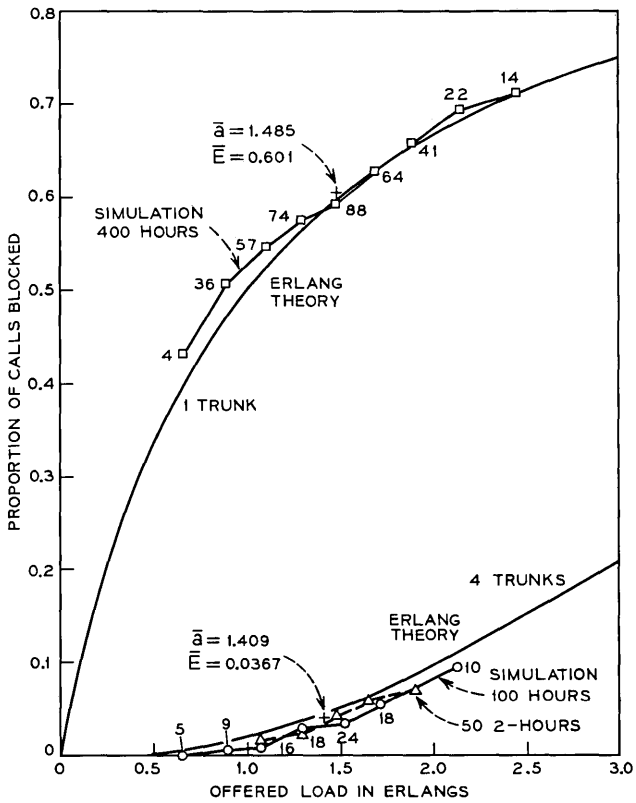


Fig. 4—Load vs loss simulation, with equilibrium offered loads, results summarized by 1- and 2-hour periods. (Numbers at points show hours of data averaged; $\bar{a}$ and $\bar{E}$ are mean values for all hours.)

We are directed to the conclusion that although Erlang's loss theory is properly applied only to a system operating in statistical equilibrium over a long period of time, for engineering purposes it is suitable for describing conversation type traffic load versus loss results summarized by periods as short as one hour (summary period $\geqq 20$ average holding times).

2.2 *Estimation of Hourly Offered Loads from Switch Counts*

It is customary to furnish only a switch counting device on high-usage trunk groups, and to depend on the accuracy of Erlang's load-loss relation,

$$E_{1,x}(a) = \frac{a^x/x!}{1 + a + a^2/2! + \cdots + a^x/x!},$$

to estimate the corresponding offered load by solving for a in

$$\ell = a[1 - E_{1,x}(a)] \tag{2}$$

where x is the number of trunks in the group. Descloux[1] has shown that a certain amount of bias occurs by this method. The reason for this is made clear by examining Fig. 5.

Figure 5a shows the loads offered, as estimated by equation (1), to three 1-trunk groups during the 3 busiest hours of the day for 20 business days. The load carried (= occupancy here), as determined by switch counting, is plotted as the ordinate. The Erlang relation is drawn as the solid line.

The points comprise a correlation scatter diagram. The classical regression line of "y on x," that is, the average value of carried load for a selected average offer, would be expected to agree with the Erlang loss relation, since the latter is derived on the basis of offered load as the "independent variable." The "y on x" regression line generated from the data is indeed seen on Fig. 5b (dashed broken line) to conform nicely with the Erlang theory.

Unless there were perfect correlation, the "x on y" regression line could not be expected to agree with Erlang's theory, and this is corroborated by the observed regression (solid broken) line on Fig. 5b. At lower than median occupancies (and offered loads) this regression line lies slightly below the Erlang theory, and for larger than median occupancies, it lies consistently above. Thus when estimating offered loads from carried loads by Erlang's loss formula [equation (2)] we should expect the values to be smaller than true for lower than average occupancies, and greater than true for larger than average occu-

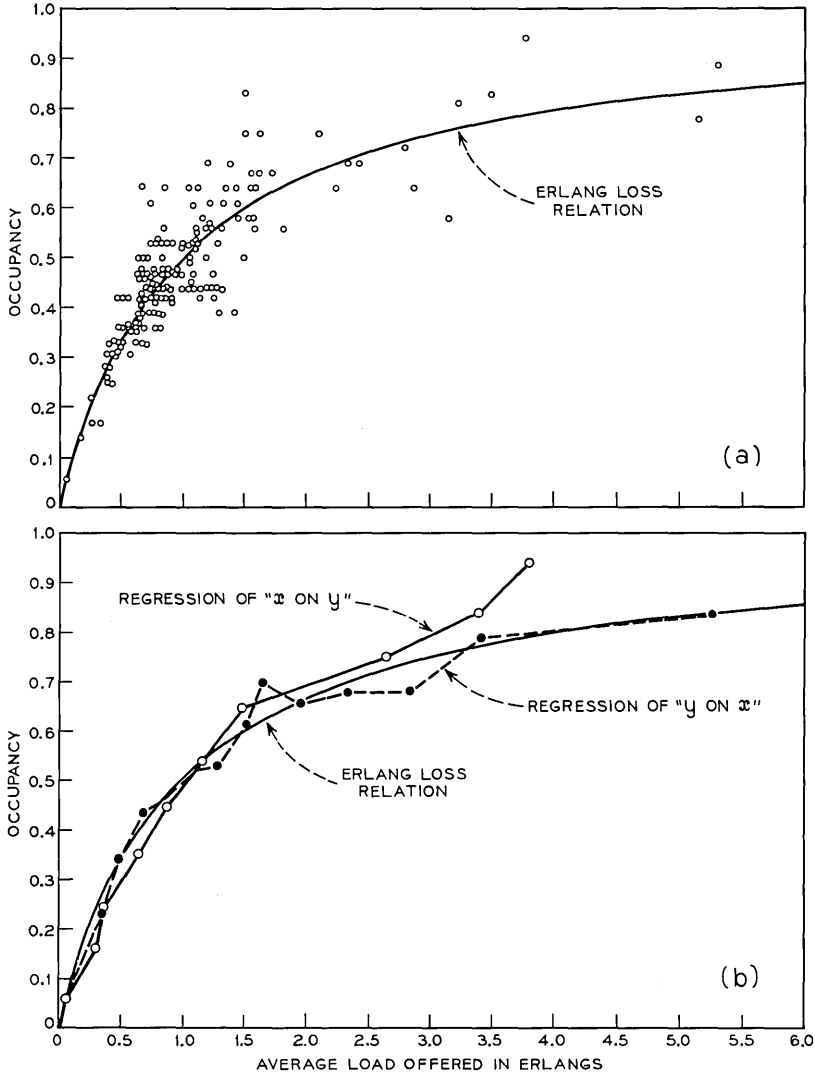


Fig. 5—Carried vs offered loads, data on three 1-trunk high-usage groups, Kildare, 1958, 3 hours per day, 20 business days. (a) scatter diagram, (b) observed regression lines.

pancies. For example, for one trunk at an observed load carried (occupancy) of 0.75 erlang, Erlang theory would predict an offered load of 3.0 erlangs, while the observed regression line would more nearly suggest an offer of 2.64 erlangs, a difference of 14 percent.

An appropriate theoretical expression for the "x on y" regression in this circumstance is difficult to generate, depending partially as it does on the day-to-day distribution of the offered loads, and this last becomes a matter of observation of customer characteristics. Since the regression lines may be decidedly nonlinear, the theoretical approaches of simple correlation theory are inadequate. In the face of such difficulties, one reverts to the reduction of data taken in the field for practical estimating results.

Figure 6 shows the ratio of two estimates of the offered load, a_2/a_1, versus observed occupancy of several 1-trunk groups. Here a_1 is the assumed unbiased estimate from equation (1), and a_2 is from equation (2), the latter making use only of the carried load measurement. In spite of the rather wide differences in loading among the five groups shown, the a_2/a_1 ratios are quite similar and one would not have too much difficulty in drawing a central ratio line through the field for general correction of a_2 values.

Figure 7 shows a similar chart for three trunks, indicating the ten-

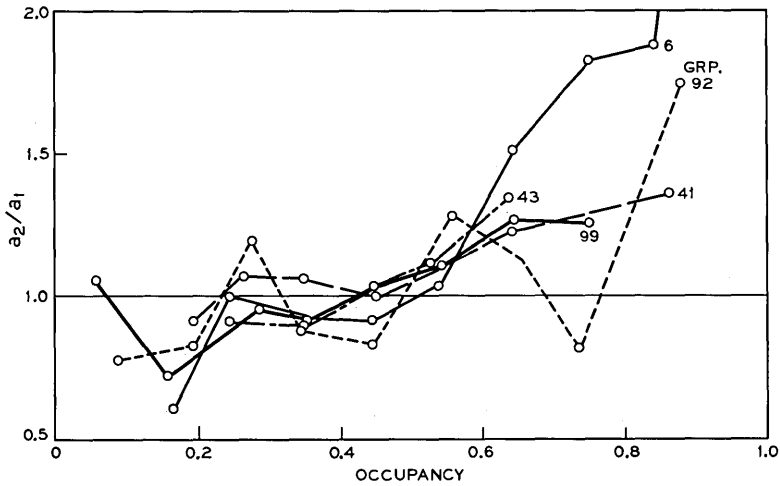


Fig. 6—Ratio of estimates of single hour loads offered to 1 trunk. Estimate a_1 is from switch and call counts, estimate a_2 from switch count and Erlang's loss formula.

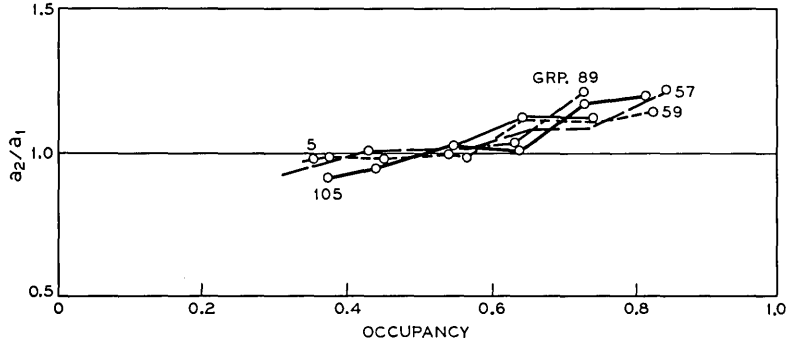


Fig. 7—Ratio of estimates of single hour loads offered to 3 trunks. Estimate a_1 is from switch and call counts, estimate a_2 from switch count and Erlang's loss formula.

dency for the ratios to approach 1.0 with increasing trunk group size. Figure 8 shows a rough summary of the occupancies at which the ratios would exceed 1.05 with various-sized trunk groups. Thus one might conclude, for example, that estimates by Erlang's loss formula of hourly offered loads on 12-trunk groups would not be acceptable

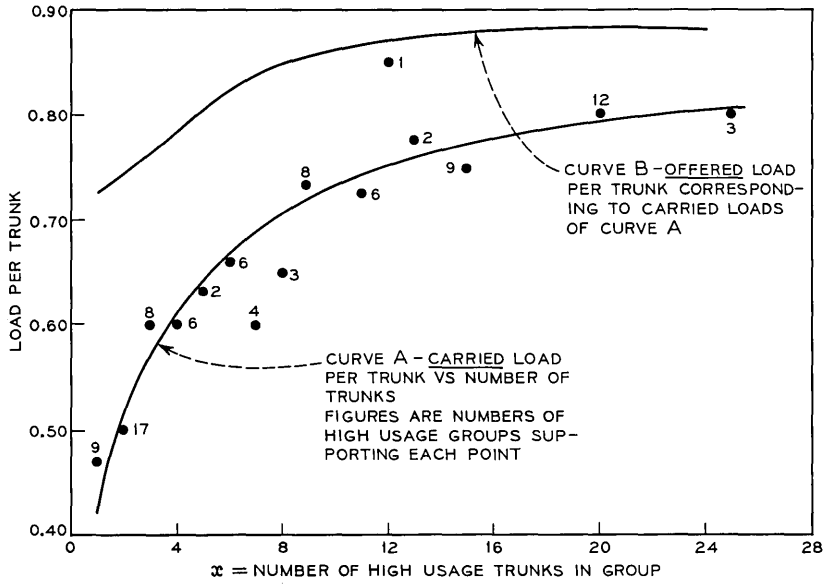


Fig. 8—Median occupancy (load per trunk) beyond which ratio a_2/a_1 for single hours exceeds 1.05 (approx.). Kildare data for 3 morning hours.

without correction (that is they would over-estimate by more than 5 percent) for occupancies greater than 0.75.

2.3 *Estimation of Average Overflow Load from Average Offered Load During a Busy Season*

When an offered load a varies from hour to hour in excess of the amount expected in 1-hour segments of a longtime load in statistical equilibrium,[2] the average overflow $\bar{\alpha}$ over a period of time will exceed that estimated by entering the Erlang loss relation with $\bar{a}$, that is, $\bar{\alpha} > \bar{a} \cdot E_{1,x}(\bar{a})$. This result is caused by the always concave upwards shape of the offered load, overflow load curves. The amount of such excess will depend strongly on the magnitude and character of the offered loads; in any event as a gets larger, the effect decreases for a given size of trunk group.

Numerous studies have shown that day-to-day busy hour variations in the busy season tend to follow a Type III Pearson distribution,

$$\theta(a) = Ka^h e^{-ca}, \tag{3}$$

in which the constants are determined from the mean $\bar{a}$ and variance v of the data, as

$$h = \bar{a}^2/v - 1$$

$$c = \bar{a}/v$$

and K is the normalizing coefficient, $c^{h+1}/\Gamma(h + 1)$. Typical is the example shown in Fig. 9 for a 9-trunk interlocal group in which $\bar{a} = 5.72$, $\text{Var}(a) = 1.61$, yielding $h = 19.20$, $c = 3.55$.

Again, there is found a considerable correlation between the day-to-day variance and the mean of such a distribution. Figure 10 shows on a log-log plot the field of variances versus means of loads offered during 3 hours each day for 20 days on 72 high-usage interlocal groups at Kildare office. Summaries at other exchanges in the United States confirm a similar association of variances and means. The general line of regression of variance on mean for the corresponding scatter diagrams is shown by the solid line, whose equation is approximately

$$\text{Var}\ (a) \doteq 0.31\bar{a}. \tag{4}$$

(A dashed line has also been drawn to approximate the major axis of the elliptical pattern of points. Its equation is

$$\text{Var}\ (a) = 0.13\bar{a}^{1.58}. \tag{5}$$

It will be referred to in a later section.)

If the Erlang loss formula can be used to estimate the proportion of calls which overflow a high-usage group during a single hour, the average load overflowing, $\bar{\alpha}$, over a series of hours is then calculable from

$$\bar{\alpha} = \int_{a=0}^{\infty} aE_{1,x}(a)\theta(a)\, da. \tag{6}$$

Similarly, the day-to-day variance of the overflow loads is determined from

$$\text{Var}\,(\alpha) = \int_{a=0}^{\infty} [aE_{1,x}(a)]^2\theta(a)\, da - \bar{\alpha}^2. \tag{7}$$

Curves have been calculated by numerical integration, using the regression relation of Fig. 10, which give the ratio of $\bar{\alpha}$ to $\alpha_{\bar{a}}$, the latter being the overflow corresponding to an offer of $\bar{a}$, found according to

$$\alpha_{\bar{a}} = \bar{a}E_{1,x}(\bar{a}). \tag{8}$$

Values of $\bar{\alpha}/\alpha_a$ are given on Fig. 11. At constant loads on the last trunk (shown by dashed lines for 100 call seconds per hour, or CCS) the ratios $\bar{\alpha}/\alpha_{\bar{a}}$ are relatively constant; hence a simple table, Table I,

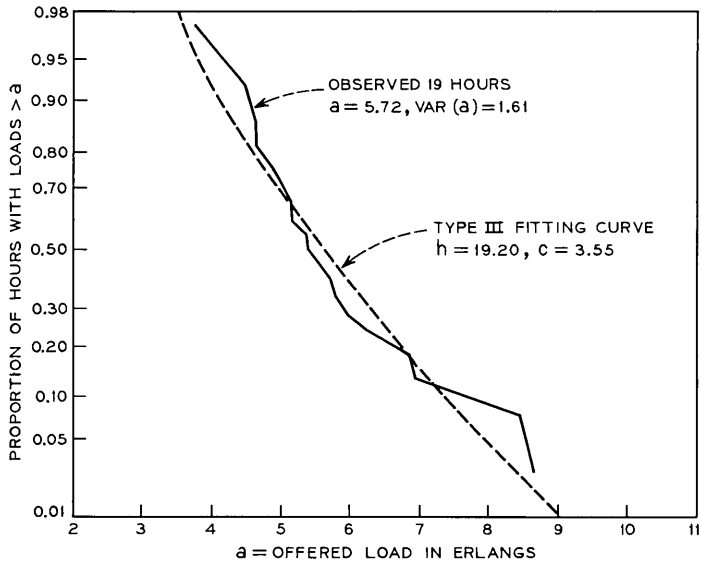


Fig. 9—Variations in day-to-day busy hour loads, Kildare group No. 67, 9 trunks, 20 days.

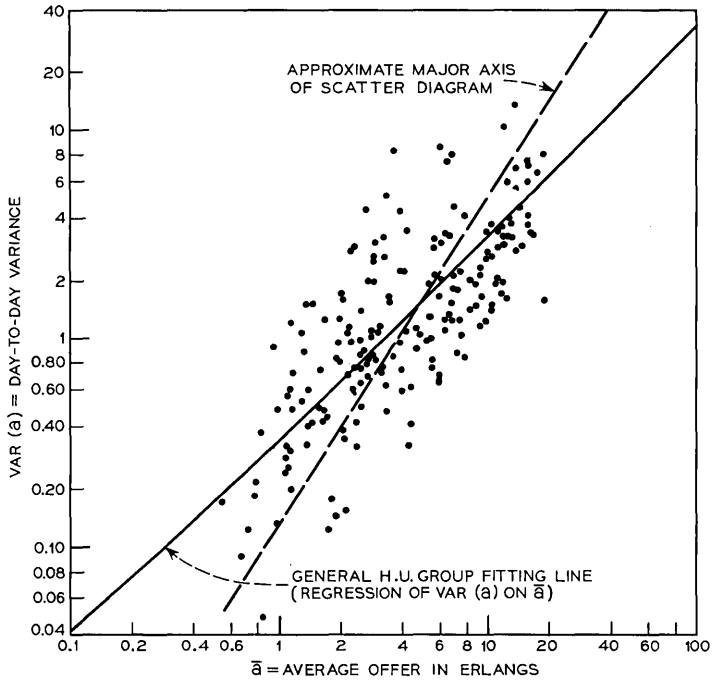


Fig. 10—Day-to-day variance vs average load in clock hours, 72 interlocal high-usage groups, 3 hours, Kildare, 1958.

using the last trunk load as the index is adequate for many working purposes. An example of the need for correcting by the factors of Table I is shown in Fig. 12. In the left diagram, uncorrected $\alpha_{\bar{a}}$ values are plotted against average overflow loads, $\bar{\alpha}_1$, calculated by the believed generally unbiased procedure

$$\bar{\alpha}_1 = \frac{1}{s} \sum_{i=1}^{s} \ell_i \frac{n_i/m_i}{1 - n_1/m_i} \quad (9)$$

where s is the number of hours summarized for each group. [Compare with equation (1).]

After correction the overflow estimates are shown in the right diagram to be in much better agreement with the $\bar{\alpha}_1$ values.

2.4 *Estimation of Average Offered Loads from Average Loads Carried*

To obviate the labor of calculating individual hourly estimates of the offered loads from observed hourly carried loads on high-usage

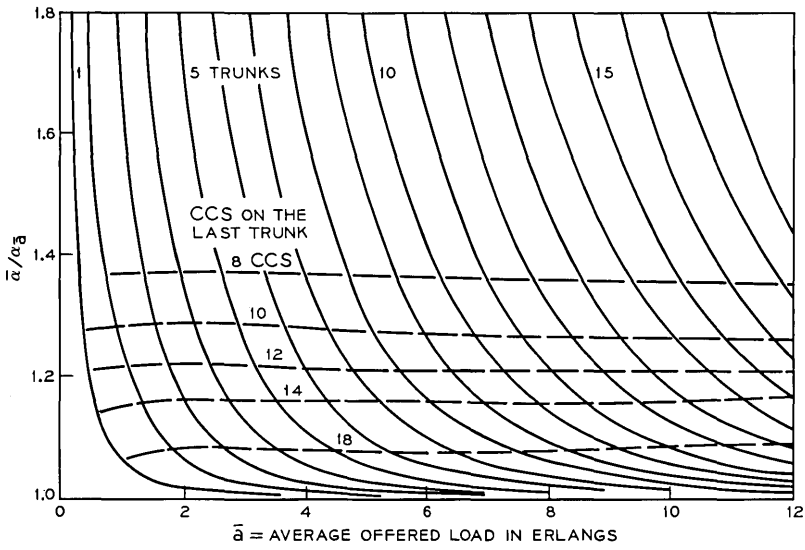


Fig. 11—Correction required in $\alpha_{\bar{a}}$ to estimate $\bar{\alpha}$.

groups (which would in turn require a correction as discussed in Section 2.2), the carried loads ℓ are commonly averaged first, and this average $\bar{\ell}$ is then entered in Erlang loss theory curves or tables to obtain $a_{\bar{\ell}}$ from

$$\bar{\ell} = a_{\bar{\ell}}[1 - E_{1,x}(a_{\bar{\ell}})]. \tag{10}$$

One needs then to compare $a_{\bar{\ell}}$ with the true offer $\bar{a}$. This is done as follows. Choosing a value of $\bar{a}$ as the offer to x trunks, $\bar{\alpha}$ is obtained by correcting $\alpha_{\bar{a}}$ employing the appropriate factor from Fig. 11. One then obtains $\bar{\ell}$ from

$$\bar{\ell} = \bar{a} - \bar{\alpha} \tag{11}$$

TABLE I—CORRECTIONS IN $\alpha_{\bar{a}}$ TO ESTIMATE $\bar{\alpha}$

Load on the Last Trunk		Range of Corrections Seen in Fig. 11	Corrections to $\alpha_{\bar{a}}$ for Practical Use
CCS	erlangs		
8	0.22	1.32–1.37	1.33
10	0.28	1.22–1.28	1.25
12	0.33	1.16–1.21	1.20
14	0.39	1.11–1.16	1.15
18	0.50	1.05–1.09	1.08

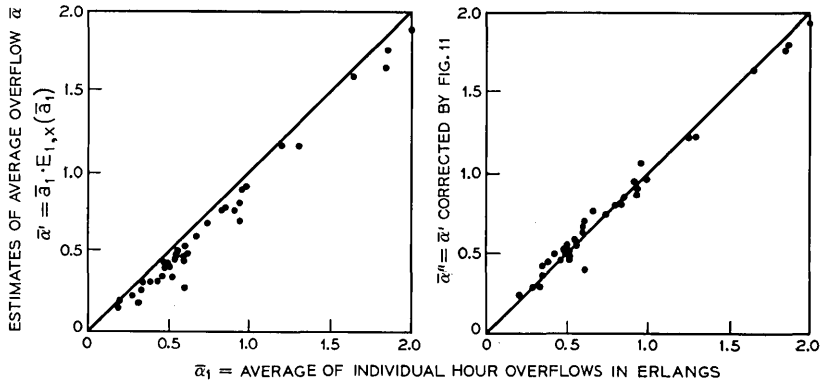


Fig. 12—Effect of corrections by Fig. 11 on estimates of overflows from high-usage groups of Kildare Tandem No. 1, 9–10 A.M.

since that load which does not overflow must naturally be carried. Relation (10) then gives $a_{\bar{t}}$. A field of values of $\bar{a}/a_{\bar{t}}$ have been calculated and form the curves of Fig. 13. We see that the ratio is sensitive both to proportion of overflow and to the number of trunks.

Corresponding data for Kildare groups are shown on Fig. 14. Similar forms of probability density distributions are observed.

It will be noted that there is a generally maximum ratio on Fig. 13 which tends to occur roughly at $E_{1,x}(a) = 0.2$, that is, at an expected overflow of 20 percent. It is interesting that this maximum correction required on $a_{\bar{t}}$ occurs squarely in the middle of the most common economic high-usage group operating levels.

For practical use we have constructed the traces of the 3-dimensional surface of Fig. 13 which correspond to several values of "economic CCS on the last trunk." (For a discussion of economic CCS, see Ref. 3, Section 7.6.) The theoretical $\bar{a}/a_{\bar{t}}$ ratios for 8, 14, 20, and 25 CCS on the last trunk are shown on Fig. 15. Comparison with data taken on a number of intertoll high-usage groups at Memphis, Tennessee, is shown in Fig. 16. Although the dispersion among individual groups is considerable, the grouped average values show reasonable agreement with theory. Table II shows the values of corrections to $a_{\bar{t}}$ required to best approximate the true offer $\bar{a}$ for 20 days of data, for the four selected last-trunk CCS. A brief study indicates that when the busy season includes fewer than 20 days, the required correction is somewhat smaller.

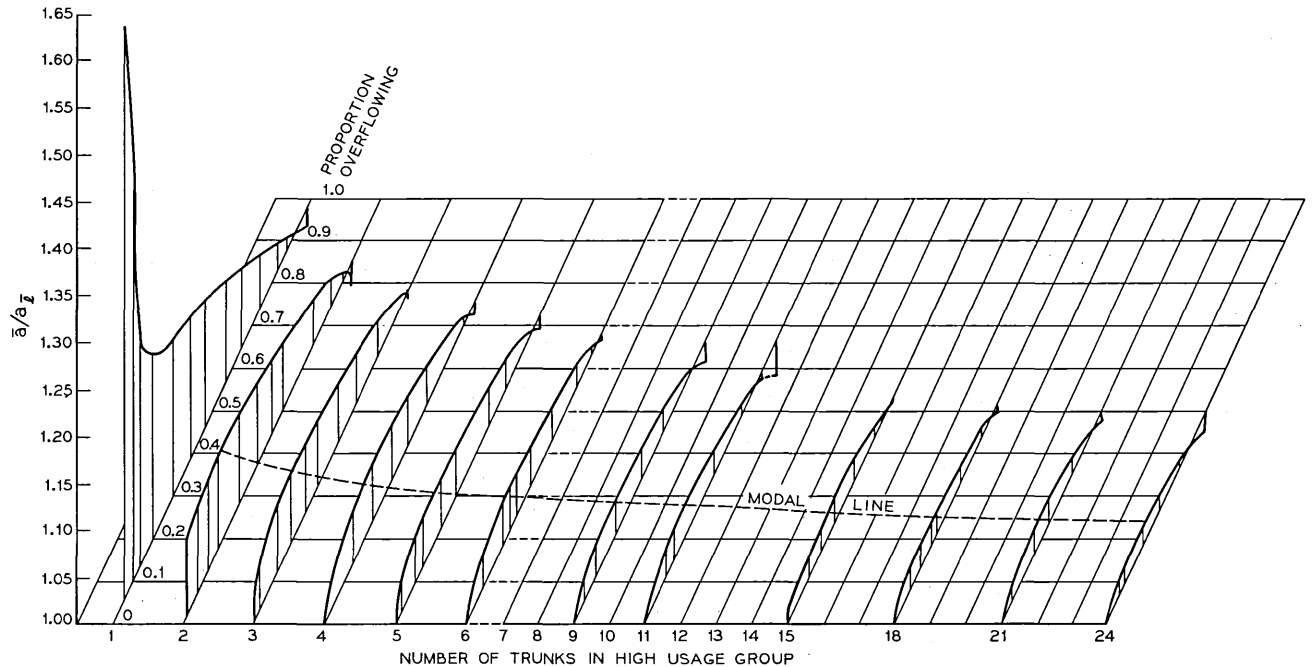


Fig. 13—Theoretical corrections required in a_I to obtain improved estimates of offered load $\bar{a}$.

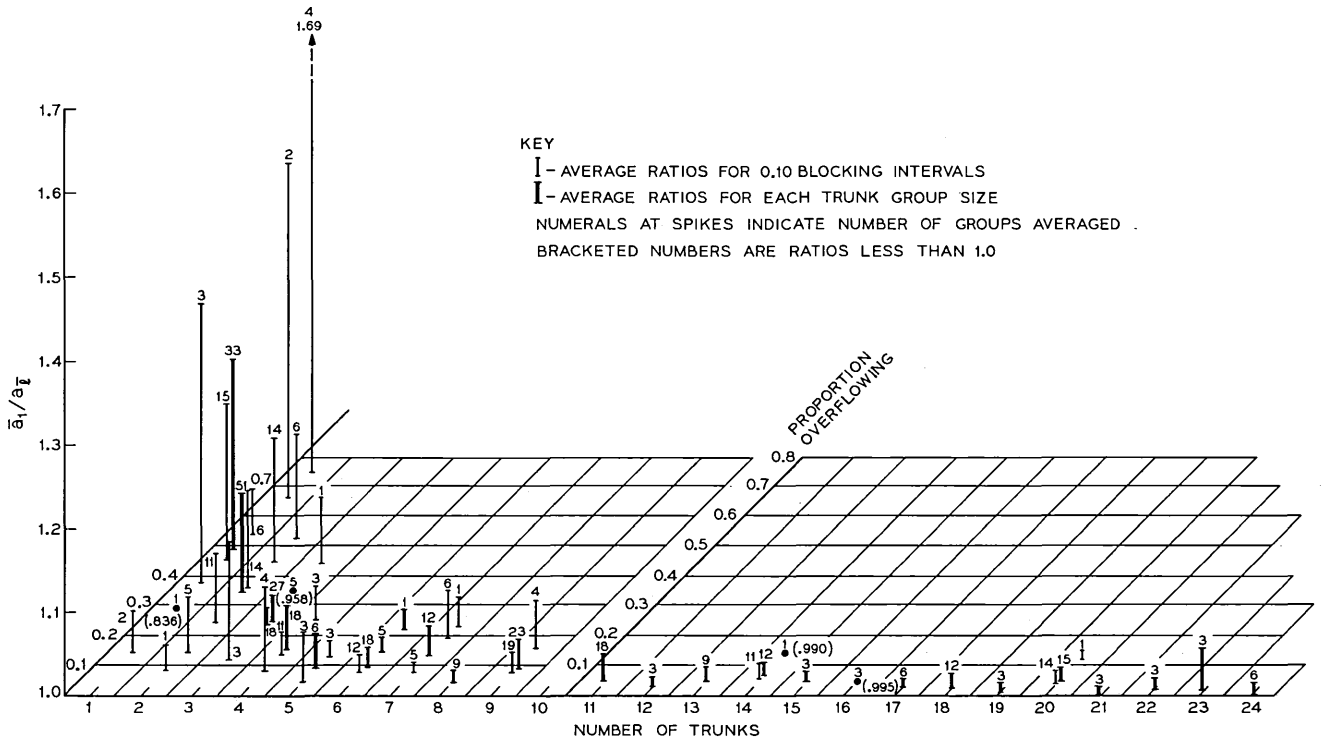


Fig. 14—Observed corrections required in $a_{\bar{t}}$ to obtain improved estimates of offered load $\bar{a}$, 100 high-usage groups in 3 tandems, Kildare, 1958, 3 hours for 20 days.

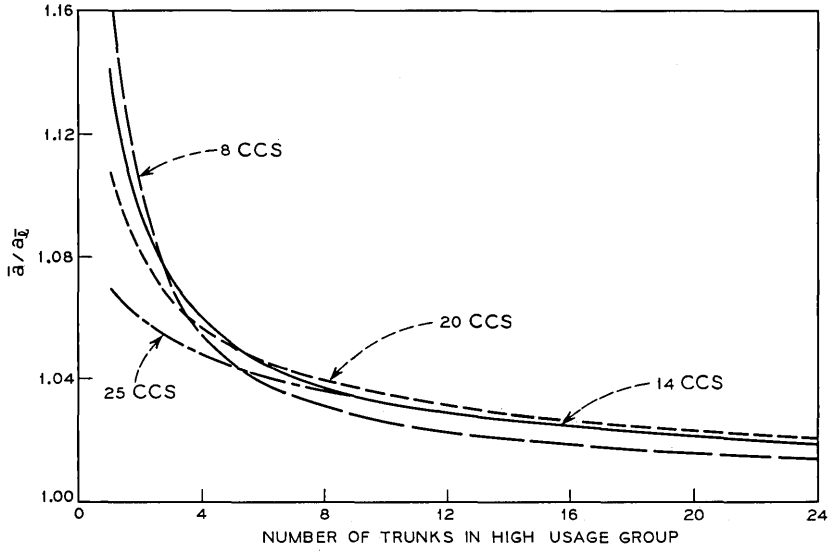


Fig. 15—Practical correction of $a_{\bar{I}}$ to estimate $\bar{a}$, with dependence on engineering level of load on "last trunk."

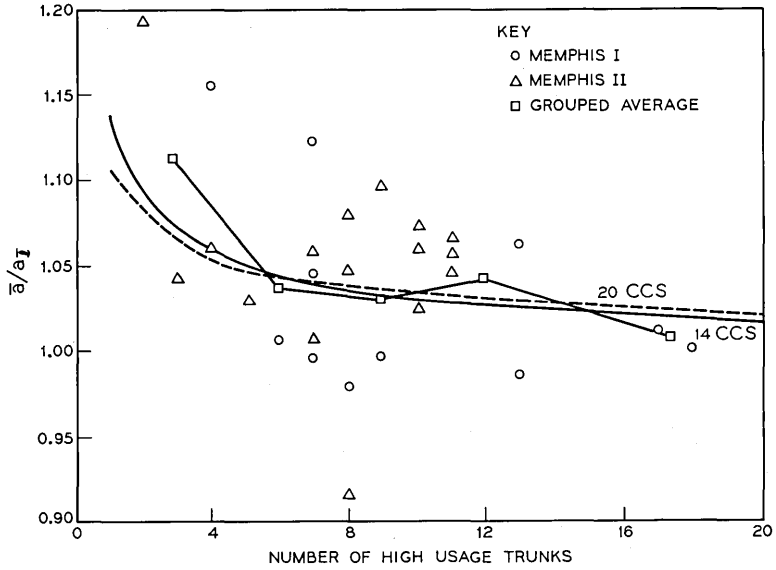


Fig. 16—Comparison of theory and data in corrections of $a_{\bar{I}}$ to estimate $\bar{a}$, intertoll data at Memphis, 1957–1958.

TABLE II—AVERAGE CORRECTIONS TO BE APPLIED TO $a_{\bar{\ell}}$ TO OBTAIN IMPROVED ESTIMATES OF $\bar{a}$
($\bar{\ell}$ determined as 20-day average)

No. Trunks	Correction Factors for y CCS on the Last Trunk			
	$y = 8$	$y = 14$	$y = 20$	$y = 25$
1	1.16	1.14	1.11	1.07
2	1.10	1.09	1.08	1.06
3	1.07	1.07	1.07	1.05
4	1.05	1.06	1.06	1.05
5	1.04	1.05	1.05	1.04
6	1.04	1.04	1.05	1.04
7	1.03	1.04	1.04	1.04
8	1.03	1.04	1.04	1.04
10	1.03	1.03	1.03	1.03
12	1.02	1.03	1.03	1.03
15	1.02	1.03	1.03	1.03
20	1.02	1.02	1.02	1.02
25 and up	1.01	1.02	1.02	1.02

2.5 *Estimation of Average Offered Loads from Average Overflow Ratios*

To monitor the level of traffic flow in some systems, a pair of registers which count numbers of calls offered m, and numbers of calls blocked n, are provided instead of the more common carried load meters. The value n/m is called the overflow ratio. Descloux[1] has shown that, under the condition that the observation period is long, say 20 holding times, little bias exists in estimates made of individual hourly offered loads to high-usage groups when a is determined from the Erlang loss relation

$$E_{1,x}(a) = n/m. \tag{12}$$

To increase the reliability of such estimates, the overflow ratios for a number of hours are commonly averaged and entered as for a single hour in equation (12). We inquire whether the estimate of average offer, $a_{\overline{n/m}}$, so obtained is unbiased, or requires correction. The expected loss probability when a load a, varying from hour to hour according to $\theta(a)$, is submitted to x trunks is

$$\bar{E}_{1,x}(a) = \int_0^\infty E_{1,x}(a)\theta(a)\, da. \tag{13}$$

When a constant load equal to the average load $\bar{a}$ is offered to x trunks, a loss probability E' will result,

$$E' = E_{1,x}(\bar{a}). \tag{14}$$

In general $\bar{E}$ will be different from E'.

Suppose that actual busy hour overflow ratios are observed over a period of days, and averaged giving $\overline{n/m}$. If $\bar{a}$ is estimated as $\bar{a}_1$ through equating $\overline{n/m}$ to $\bar{E}$ in equation (13) we should expect $\bar{a}_1$ to be unbiased since both $\overline{n/m}$ and $\bar{E}$ contemplate the presence of day-to-day variations, the first in reality, and the second by the inclusion of $\theta(a)$ in equation (13). However estimating $\bar{a}$ through substituting $\overline{n/m}$ for E' in equation (14) will yield $a_{\overline{n/m}}$, containing a certain amount of error since it merely determines the single hour load corresponding to the average loss.

A range of offered loads to trunk groups of size 1–24, covering the cases of 8 to 18 "CCS on the last trunk," has been studied and $\bar{a}_1/a_{\overline{n/m}}$ plotted against the numbers of trunks. This is shown in Fig. 17. For

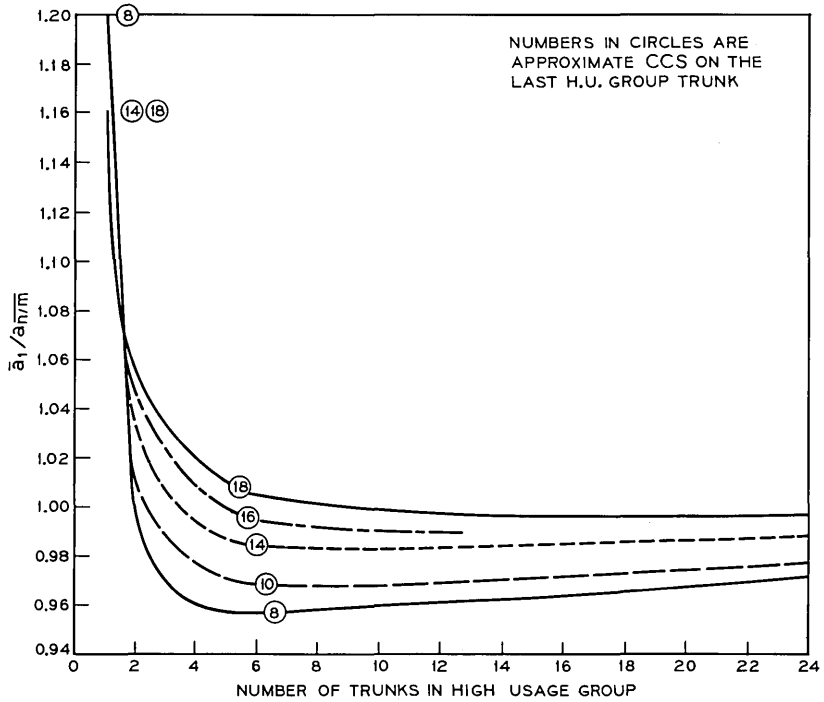


Fig. 17—Correction required in offered load estimated from averaged n/m readings.

one and two trunks the estimation procedure based on equation (14) considerably underestimates the average offered load; but beyond four trunks, overestimates of the true average offer are generally made. The cause of the inversion in the correction required lies in the shapes of the Erlang load versus loss curves in the ranges of interest. The one- and two-trunk curves are predominantly concave downwards (see for example Fig. 1), while for the larger trunk numbers the curves are concave upwards. Only for the one-trunk case are the corrections sufficient in most practical situations to warrant their making. Here a median correction is about 1.18.

Available data for groups having call and overflow counters, while showing considerable dispersion, confirm the shape and location of the correction curves of Fig. 17.

III. GROUPS WITHOUT ALTERNATE ROUTES

There has long been interest in the different theoretical formulas used by American and European administrations for engineering interlocal trunk groups. The former have relied on the cumulative Poisson while the latter have favored Erlang's loss formula. What do data taken on such groups indicate?

In 1959 data were taken on some 30 direct trunk groups terminating in the Arlington, Massachusetts, No. 5 crossbar office. Hourly observations comprised numbers of calls offered to each group, the number blocked, and, by switch counting, the average load carried. Illustrative of these is group No. 26 with 32 trunks observed for 4 hours a day for 29 business days. Offered loads were estimated for each hour by equation (1). The load versus blocking relationship observed for 116 hours is shown in Fig. 18. Superimposed is the Erlang loss relation. The agreement is seen to be very good.

On Fig. 19 are shown only the twenty-nine 10–11 A.M. busy hour loads of Fig. 18. The agreement with Erlang's theory is again understandably excellent. Other group data comparisons were nearly as satisfactory. We are led to conclude that Erlang's loss formula describes quite well the hourly blocking for conversation traffic on direct groups which do not have alternate routes. Apparently the return of blocked calls here was of a nature that caused little disturbance of the Poisson character of the offer—or perhaps they contributed to it!

In the lower section of Fig. 19 is shown a distribution of the busy hour loads offered to group No. 26 as they varied from day to day over the 6 weeks busy season. For the average offer of $\bar{a} = 21.37$

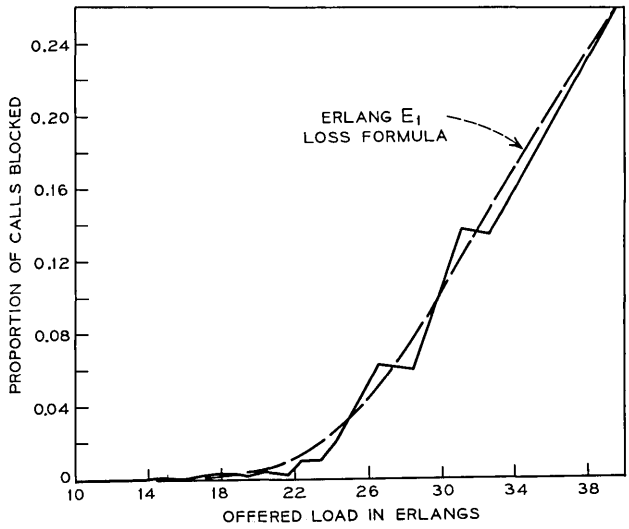


Fig. 18—Comparison of load vs loss data with Erlang's theory: full group No. 26 of 32 trunks, Arlington, 4 busiest hours of each day for 6 weeks, 1959.

erlangs, a variance $\mathrm{Var}(a) = 20.3$ was observed. Since the probability of loss curve in this range is concave upwards, the average loss for all the hours tends to exceed the loss at the average load. Thus the theoretical single hour loss at 21.37 erlangs offered is 0.0074. However the observed average loss was found to be 0.0197 when hourly losses were given equal weight (as in American practice). (If the losses had been weighted by the corresponding offered loads, the average loss would be still greater at 0.0267.)

Clearly, if Erlang's formula describes well the losses for individual hours, it will not usually give an adequate estimate of the average loss over a series of busy hours. However, when the more conservative Poisson summation values,

$$P(x, \bar{a}) = \sum_{r=x}^{\infty} \frac{\bar{a}^r e^{-\bar{a}}}{r!},$$

are laid on Fig. 19, they are seen to pass directly through the (unweighted) average loss point, indicating that at least for the amount of day-to-day variation present in this sample of hours, the Poisson provides just the right amount of "improvement" to Erlang's loss formula.

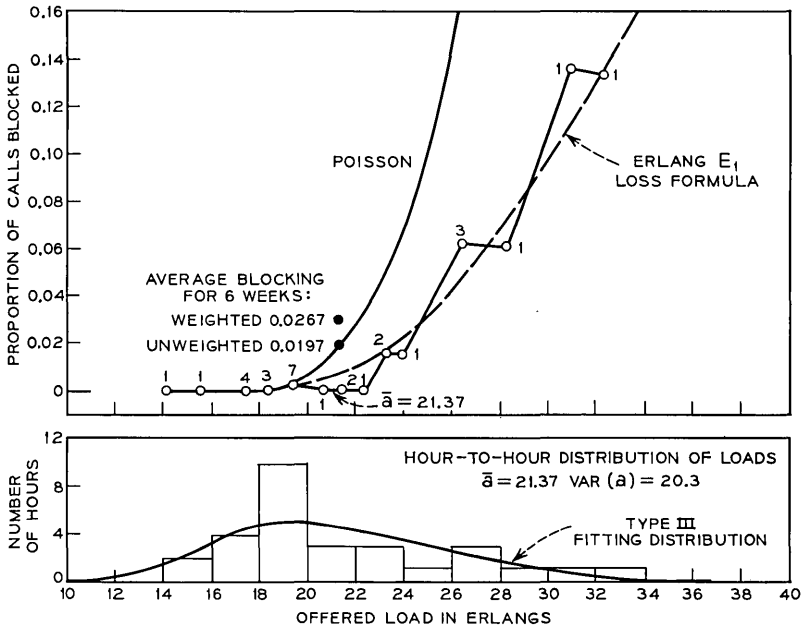


Fig. 19—Comparison of load vs loss data with Erlang's theory: full group No. 26 of 32 trunks, Arlington, 10–11 A.M. busy hour each day for 6 weeks, 1959. (Numbers at points are hours averaged.)

There is then a particular variance of day-to-day busy hour loads which when applied to Erlang's loss formula will just produce the average loss of the Poisson formula. Figure 20 shows a field of curves indicating the Var (a) required for average loads, $\bar{a}$, such that at the loss levels given, the Poisson summation will closely relate the average busy season busy hour loss to the average offer. On the same field are shown the actual variances versus averages observed for the trunk groups in the Arlington study mentioned above. It is seen that the variances found in practice are such that at commonly used interlocal average busy season busy hour grades of service, 0.005 to 0.03, the Poisson summation provides an excellent means of specifying group average capacity. It may be noted that in American practice, the objective grades of service are set to be met by the average (unweighted) blocking in the busy season busy hours.

The dashed line on Fig. 20 is a centrally located curve which describes the general relation seen between day-to-day variance and average load offered. It is interesting – and reasonable – that its equation is

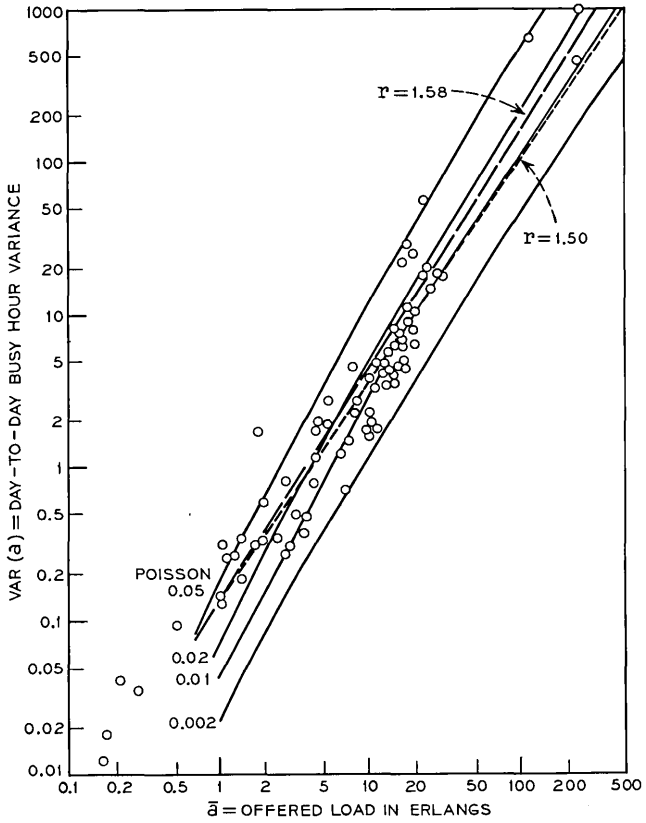


Fig. 20—Variations in time consistent hourly loads on 32 direct groups of trunks: Arlington, 4 busiest hours of each day for 6 weeks, 1959, compared with "Poisson assumption."

Var $(a) = 0.13\ \bar{a}^r$ where $r = 1.58$, identical with that of the central axis line drawn through the high-usage group observations on Fig. 10. The dotted line on Fig. 20 results from choosing $r = 1.50$; it corresponds nicely with Poisson blocking of 0.01 for loads of 10 to 100 erlangs.*

* It may be noted that Bell System traffic capacity tables for nongraded groups carrying random and nonrandom offers have been generated at several blocking levels, and contemplate the following hour-to-hour load variations: Single hour blocking; "low" hour-to-hour variation ($r = 1.50$); "medium" hour-to-hour variation ($r = 1.70$); "high" hour-to-hour variation ($r = 1.84$). At loads of one erlang, these several traffic tables contain the same allowance for day-to-day variations. Below one erlang the theoretical day-to-day variations, while inverted from their relative positions at loads above one erlang, are small compared with the momentary variations inherent in random traffic.

IV. FINAL GROUPS

Much has been written in the past on estimation of the effect of the nonrandomness (peakedness) of overflow traffic from high-usage groups upon the trunking requirements in final groups.[3] Studies have also been made (as in the previous section) showing the added trunks required to accommodate the increased demand on groups whose offered loads show significant day-to-day variations.

The two effects will usually appear simultaneously in the engineering of final groups. Moreover, loads offered to final groups may be expected to show generally larger day-to-day busy hour variations than do loads to high usage and direct groups not having alternate routes. Figure 21 shows a field of variance versus average loads offered to 28 interlocal final groups in which negligible first-routed traffic was present. A central fitting line, having the equation

$$\log_{10} \text{Var } (a) = 1.8404 \log_{10} \bar{a} - 0.8861, \tag{15}$$

has been drawn through the points, or equivalently,

$$\text{Var } (a) = 0.13\bar{a}^{1.84}. \tag{16}$$

When this variation is assumed, the carrying capacities of 6 to 15

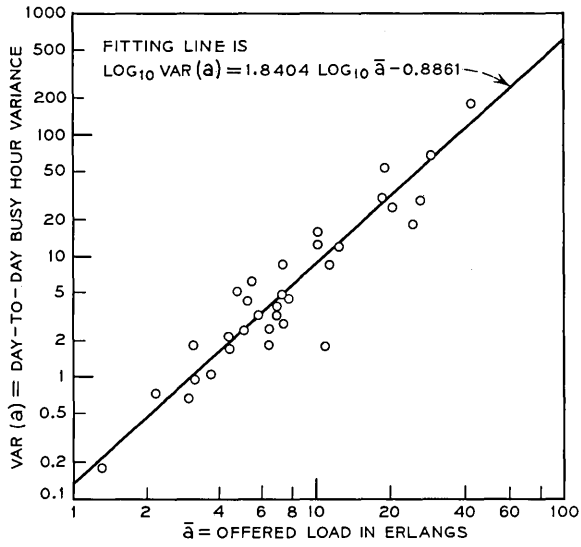


Fig. 21—Variations in busy hour loads offered to final groups in 28 interlocal alternate route systems.

trunks at an 0.02 average blocking are shown illustratively on Fig. 22. For 10 trunks, for example, the erlang capacity for random traffic is reduced from 5.08 to 4.47 erlangs, a drop of 12 percent resulting from the day-to-day variations.

Similarly, if day-to-day variations are *not* introduced, but instead a nonrandomness, characterized by a peakedness factor P.F. = (variance)/(mean) = 1.5, is assumed, the capacity of 10 trunks is reduced from 5.08 to 3.96 erlangs, a reduction of 22 percent.

When day-to-day variations and nonrandomness are jointly introduced, the capacity of a group is further reduced; thus the 10 trunks will now accommodate an offer of only 3.57 erlangs, a reduction of 30 percent from the basic Erlang loss formula value. It is clear that each cause can produce substantial reductions in trunking capacity.

When the offer to a final group contains a significant fraction of first-routed traffic, both the peakedness factor and the day-to-day variations will commonly be lowered; the trunking requirements will then be correspondingly reduced on both accounts.

An illustration of the importance of using theory which considers both day-to-day variations and the peakedness of the traffic is given in Fig. 23. The results are shown of observations on a Kildare final group of 59 trunks, with 62 subtending high-usage groups plus 4 first-routed traffic items. From high-usage group load and trunk configurations the estimated peakedness factor of the final group offered load is 1.93. The load-loss characteristics observed for the 20 days are:

Hour of Day	Average Offer (erlangs)	Day-to-Day Variance of the Offer	Average Blocking Observed
9–10 A.M.	26.5	28.5	0.001197
10–11 A.M.	44.0	180.3	0.04570

As seen on the figure, the simple Erlang loss values (dotted lines) are an order of magnitude below the observed losses. Although allowance for nonrandomness (dashed lines) makes a marked improvement, it is still far from describing the actual losses. Nor is the Poisson model (dots on Fig. 23), which might be expected to compensate partially for day-to-day variations in the 0.01 to 0.03 loss range, a nearly sufficient improvement. But employing the typical day-to-day variance magnitudes of Fig. 21 in conjunction with the estimated peakedness of the offer, the solid line load-loss curves for 58 and 60 trunks exactly bracket the average loss values seen for 20 days for the 9–10

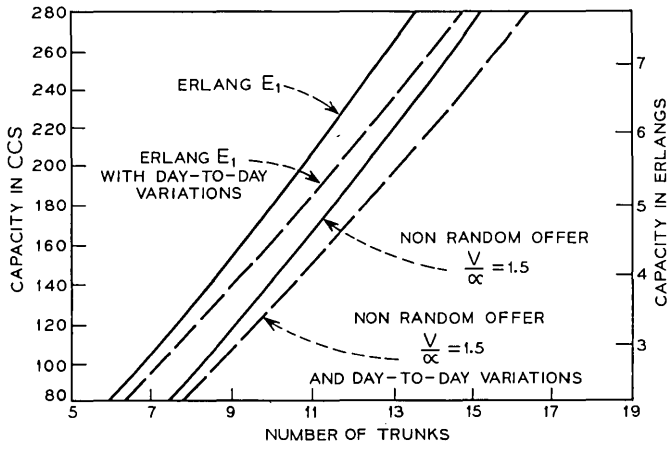


Fig. 22—Final trunk group engineering with various assumptions as to the character of offered loads. Average blocking = 0.02.

A.M. and the 10–11 A.M. hours on the 59-trunk final group. Construction of such load-loss relations is described in Ref. 4.

V. SUMMARY

Examples have been given comparing observation, simulation, and theory in various areas of traffic flow on trunks comprising direct and alternate routed plans. Particular attention has been drawn to the need for developing adequate relationships between offered, carried, and overflow loads, both single hour and average busy season, suitable to each operating condition. Where possible, a physical understanding of the principal factors is followed by statistical theory which may require approximate numerical calculations. Further insight may be gained from controlled simulations. In order for the relationships developed to be useful, they must finally be found to agree with the flow of traffic in real situations.

The following particular results have been determined:

(*i*) Single hour load-loss relationships on high-usage groups are found to be well described by Erlang's $E_{1,x}$ (*a*) loss formula. Comparisons are made with data and simulations (Figs. 1, 2, 3, 4).

(*ii*) Estimating single hour loads offered to high-usage groups from observed carried loads by use of Erlang's loss relationship yields too-

large values at the heavier occupancy levels, and slightly too-small values at low occupancies. This is explained and illustrated by recourse to regression theory (Fig. 5). The magnitude of the corrections needed are indicated in Figs. 6, 7, 8.

(*iii*) If busy season busy hour carried loads are first averaged and this value entered in Erlang's E_1-relation the average offer will be underestimated. Corrections required for a 20-day busy season are given in Table II and Figs. 13, 14, 15, 16.

(*iv*) Estimates of the average offer obtained from averaged hourly blocking proportions show that both positive and negative corrections may be required (Fig. 17). Except for the 1-trunk case, however, the corrections indicated are small.

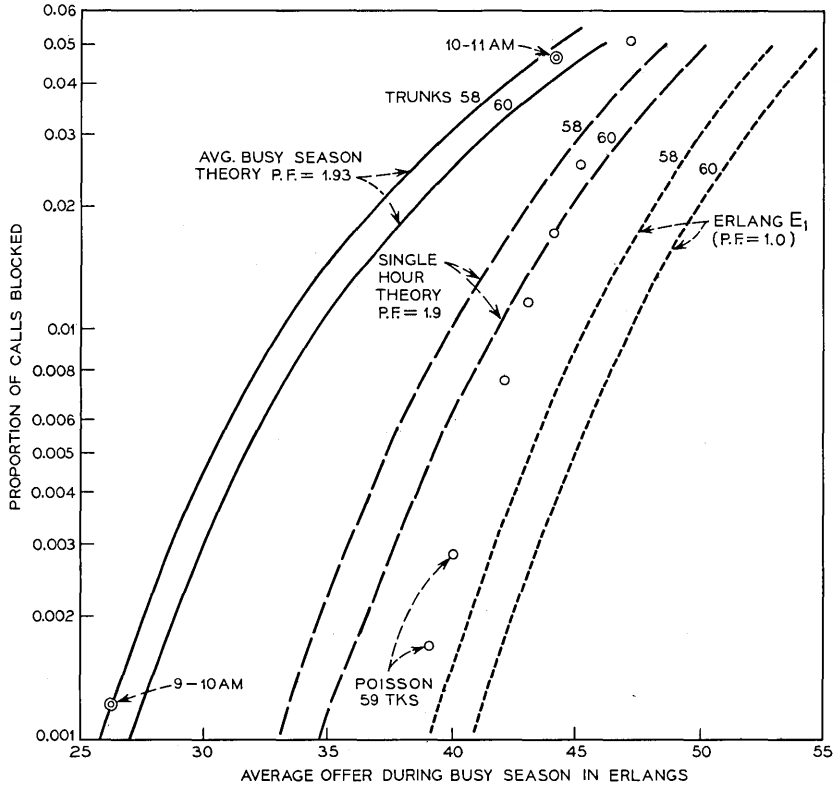


Fig. 23—Comparison of blocking theories with performance of Kildare final group No. 1, 62 high-usage groups plus 4 first-routed items, 59 trunks, P.F. = 1.93, 20 days, October–November 1958.

(*v*) When a series of busy hour loads is submitted to a high-usage group, the true average overflow load will exceed that estimated by entering Erlang's relation with the average offer. Corrections are given in Table I and Fig. 11.

(*vi*) For full groups (i.e., those without alternate routes) it is found that Erlang's formula describes well the load-loss relation for single hours (Figs. 18, 19).

(*vii*) When busy hour loads offered to full groups are averaged for the busy season, the corresponding average loss will usually be better estimated by the Poisson summation formula, if the loss values are in the common range from 0.005 to 0.03 blocking.

(*viii*) Loads offered to final groups will normally be nonrandom (peaked) and hence require special procedures for engineering. Examples are given in which both peakedness and day-to-day busy hour variations are present. Unless both influences are allowed for, the average load-average loss estimate may be far from that actually observed (Fig. 23).

REFERENCES

1. Descloux, A., "On the Variability of the Proportion of Unsuccessful Attempts in Loss Systems," Proc. Fourth Int. Teletraffic Congress, London, 1964.
2. Riordan, J., "Telephone Traffic Time Averages," B.S.T.J., *30,* No. 4 (October 1951), pp. 1129–1144.
3. Wilkinson, R. I., "Theories for Toll Traffic Engineering in the U.S.A.," B.S.T.J., *35,* No. 2 (March 1956), pp. 421–514.
4. Wilkinson, R. I., *Nonrandom Traffic Curves and Tables for Engineering and Administrative Purposes,* Traffic Studies Center, Bell Telephone Laboratories, 1970.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

When Are Transistors Passive?*

By B. GOPINATH† and D. MITRA†

(Manuscript received December 8, 1970)

The paper presents results on the stability and dynamic behavior under large signal conditions of networks consisting of transistors and sources connected to a linear, passive, memoryless subnetwork. The transistors' model incorporates various nonlinearities. A characteristic common to the main results of the paper is that they relate to properties of the transistors alone and, hence, are independent of the passive part of the network.

Sufficient conditions are obtained for asymptotic and bounded input-bounded output stability. The conditions impose restrictions on some of the physical constants of the transistors' model. These conditions have an interesting physical interpretation in terms of temperature differentials in the transistor junctions. In particular, any transistor with the exponential type of static diode characteristic is passive only if the ratio of the junction temperatures lies inside an interval determined by the α's.

In the state space of the network there exists a well-defined region R specified by the transistors' model with the property that constant terminal states in R are independent of initial conditions. The region R is in a certain sense maximal.

I. INTRODUCTION AND DERIVATION OF THE DIFFERENTIAL EQUATION

The network considered is shown in Fig. 1 and it consists of a number of transistors connected to a subnetwork composed of voltage sources, current sources, and linear, passive, memoryless devices such as resistors and transformers. Sandberg[1] has analyzed the dynamic behavior of such networks and has defined a class which exhibit various features of stability. The results of this paper are essentially different in that they relate to properties of the transistors alone and, hence, are independent of the passive part of the network.

The analysis is concerned with certain network-theoretic properties of

---

* This paper was presented at the 1970 IEEE International Symposium on Circuit Theory, Atlanta, December 1970.

† The sequence of names was chosen by flipping a coin.

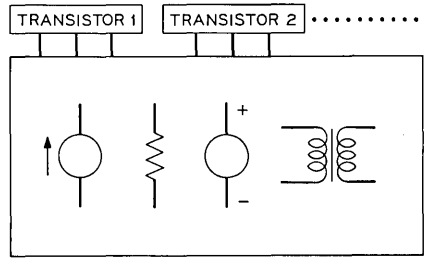


Fig. 1—General network containing transistors, sources, resistors, and transformers.

transistor models that are assumed in extensively used Network-Analysis Programs, such as NET 1, CIRPAK, CIRCUS, etc., (see Ref. 2). The intent of the paper is to study the large signal behavior of transistor networks using this model and certain generalizations of it. Each transistor is represented by a model of the type shown in Fig. 2 which takes into account the nonlinear dc properties as well as the presence of nonlinear junction capacitances. Six parameters are associated with the model: α_1, α_2, τ_1, τ_2, c_1, and c_2, all of which are positive; also α_1, $\alpha_2 < 1$. The two nonlinear static diode functions are denoted by $f_1(\cdot)$ and $f_2(\cdot)$. Initially it is assumed that $f_1(\cdot)$ and $f_2(\cdot)$ are monotone, strictly increasing mappings of the interval $(-\infty, \infty)$ into itself; $f_1(0) = f_2(0) = 0$, and $f_1(\cdot)$ and $f_2(\cdot)$ are continuously differentiable on $(-\infty, \infty)$. Further assumptions are made about $f_1(\cdot)$ and $f_2(\cdot)$ in the course of the paper.

Suppose the network has n transistors and with the polarity indicated in Fig. 2 let v_{2i-1} and v_{2i} respectively denote the emitter-to-base voltage and collector-to-base voltage of the ith transistor. Similarly, let i_{2i-1}

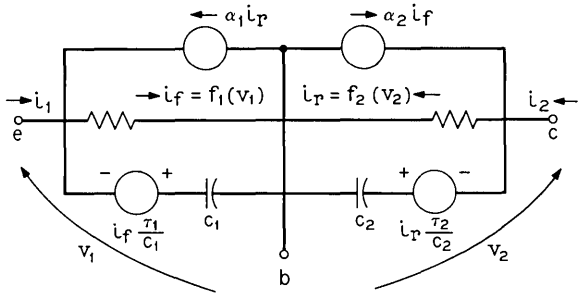


Fig. 2—Transistor model.

and i_{2i} respectively denote the emitter current and the collector current of the ith transistor. An identical scheme of subscripting is used to define $f_i(\cdot)$, c_i, and τ_i for $i = 1, 2, \cdots, 2n$. v, i, and $F(\cdot)$ are $2n$-column vectors formed by arranging $\{v_i\}$, $\{i_i\}$, and $\{f_i(\cdot)\}$ respectively. In applying the current law to the transistors' model it follows that

$$i = \frac{d}{dt} C(v) + TF(v) \tag{1}$$

where

$$[C(v)]_j = c_j v_j + \tau_j f_j(v_j)$$

and $T = T_1 \oplus T_2 \oplus \cdots \oplus T_n$ the direct sum of n 2 by 2 matrices T_j in which

$$T_j = \begin{bmatrix} 1 & -\alpha_1^j \\ -\alpha_2^j & 1 \end{bmatrix}$$

for $j = 1, \cdots, n$.

The subnetwork to which the transistors are connected is assumed to impose a constraint of the form

$$i = -Gv + b \tag{2}$$

in which G is the conductance matrix and hence $G \geqq 0$,* and b is an element of all real bounded continuous $2n$-vector-valued functions of t on $[0, \infty)$.

From (1) and (2),

$$\frac{d}{dt} C(v) + TF(v) + Gv = b. \tag{3}$$

The above equation is also derived in Ref. 1. Since all of the c_i and τ_i are positive and each of the $f_i(\cdot)$ is continuous and monotone increasing, $C_i^{-1}(\cdot)$ exists; obviously $C^{-1}(\cdot)$ is also strictly monotone increasing and $C^{-1}(0) = 0$.

II. ASYMPTOTIC STABILITY FOR THE UNFORCED SYSTEM

In this section we show how to prove an intuitively reasonable result concerning the asymptotic stability of the system. A Lyapunov function with a simple energy interpretation is introduced and used to prove

* $G \geqq 0$ indicates that G is a positive semidefinite matrix. Unless otherwise stated, G will only be assumed to be positive semidefinite.

stability. Let

$$L(v) \triangleq \sum_{i=1}^{2n} \int_0^{C_i(v_i)} C_i^{-1}(\sigma)\, d\sigma. \tag{4}$$

Since $C_i(\cdot)$ defines the charge-voltage relationship in the nonlinear junction capacitances, $L(v)$ is the total stored electrical energy in the network. It is easily verified that (*i*) $L(v) \geqq 0$, equality holding only when $v = 0$, and (*ii*) $L(v) \to \infty$ as $||v|| \to \infty$. Thus $L(v)$ is a Lyapunov function if $(d/dt)L(v) < 0$. For the unforced network

$$\frac{d}{dt} L(v) = -v^t TF(v) - v^t Gv. \tag{5}$$

Clearly if $v^t TF(v) \geqq 0$ with equality holding only if $v = 0$ then $L(v)$ is a Lyapunov function since G has been assumed to be positive semidefinite. It then follows from a well-known result in stability theory[3] that $v(t) \to 0$ as $t \to \infty$.

Note that for constant v the term $v^t TF(v)$ expresses the net power flow into the transistors and if all the transistors are passive then certainly $v^t TF(v) > 0$. However it is shown later that in certain abnormal conditions the transistors are not all passive and the judicious choice of the passive network allows one to extract energy from the transistors. To show this we first prove the following:

Lemma 1: *Given two real* $2n$*-vectors* x *and* y *such that* $x^t y > 0$, *there exists a* $G > 0$ *for which the following holds*

$$y = Gx. \tag{6}$$

Proof. Define a $2n$ by $2n - 1$ matrix Z such that the columns of Z span the $(2n - 1)$-dimensional subspace orthogonal to x. Then $Z^t x = 0$. Since by assumption y is not orthogonal to x, y is not an element of the range space of Z and, hence, the columns of Z together with y span E^{2n}. For these reasons the following construction of G,

$$G = \frac{1}{x^t y} yy^t + ZZ^t, \tag{7}$$

suffices to prove the lemma.

When for some v_1, $v_1^t TF(v_1) < 0$ then by Lemma 1 there exists a $G > 0$ such that

$$TF(v_1) + Gv_1 = 0. \tag{8}$$

Hence with initial condition $v(0) = v_1$ the solution of the network equation (3) with $b \equiv 0$ is,

$$v(t) \equiv v_1 \tag{9}$$

and the system is not globally asymptotically stable.

The above results may be summarized in the following:

Theorem 1: If $v^t TF(v) > 0 \ \forall v \neq 0$ *then the network is globally asymptotically stable. Furthermore, if there exists a* v_1 *such that* $v_1^t TF(v_1) < 0$ *then there exists a* $G > 0$ *for which the system is not globally asymptotically stable.*

2.1 *Sufficient Conditions for Asymptotic Stability*

Consideration is given to the positivity of $v^t TF(v)$. Due to the quasi-diagonal structure of T we need only consider in detail the behavior of

$$\varphi(v_{2i-1}, v_{2i}) \triangleq (v_{2i-1}, v_{2i}) \begin{bmatrix} 1 & -\alpha_1^i \\ -\alpha_2^i & 1 \end{bmatrix} \begin{bmatrix} f_{2i-1}(v_{2i-1}) \\ f_{2i}(v_{2i}) \end{bmatrix}. \tag{10}$$

For notational convenience α_1^i, α_2^i, v_{2i-1}, v_{2i}, $f_{2i-1}(\cdot)$, and $f_{2i}(\cdot)$ are respectively denoted by α_1, α_2, v_1, v_2, $f_1(\cdot)$, and $f_2(\cdot)$.

Note that when $v_1 v_2 \leqq 0$, $\varphi(v) \geqq 0$ and $\varphi(v) = 0$ only if $v_1 = v_2 = 0$. Let $v_1 \neq 0$, $v_2 = \rho v_1$, and $\rho \geqq 0$. Then

$$\varphi(v) = v_1 f_1(v_1) - \alpha_1 v_1 f_2(v_2) + v_2 f_2(v_2) - \alpha_2 v_2 f_1(v_1) \tag{11}$$

and

$$g(\rho) \triangleq \frac{\varphi(v)}{v_1 f_1(v_1)} = (1 - \rho\alpha_2) + \frac{f_2(\rho v_1)}{f_1(v_1)} (\rho - \alpha_1). \tag{12}$$

If $\rho = \alpha_1$ then $g(\rho) > 0$ since $1 - \alpha_1\alpha_2 > 0$. If $\rho \neq \alpha_1$, $g(\rho) = 0$ if and only if

$$\frac{1 - \rho\alpha_2}{\alpha_1 - \rho} = \frac{f_2(\rho v_1)}{f_1(v_1)}. \tag{13}$$

$(1 - \rho\alpha_2)/(\alpha_1 - \rho)$ is plotted as a function of ρ in Fig. 3. The salient features of the function for $\rho \geqq 0$ are:

$$\left.\begin{aligned} \frac{1 - \rho\alpha_2}{\alpha_1 - \rho} &\geqq \frac{1}{\alpha_1} \quad \text{for} \quad 0 \leqq \rho < \alpha_1 \\ &\leqq 0 \quad \text{for} \quad \alpha_1 < \rho \leqq \frac{1}{\alpha_2} \\ &< \alpha_2 \quad \text{for} \quad \frac{1}{\alpha_2} < \rho < \infty \end{aligned}\right\}. \tag{14}$$

Now $[f_2(\rho v_1)/f_1(v_1)]$ is a monotone, strictly increasing function of ρ and

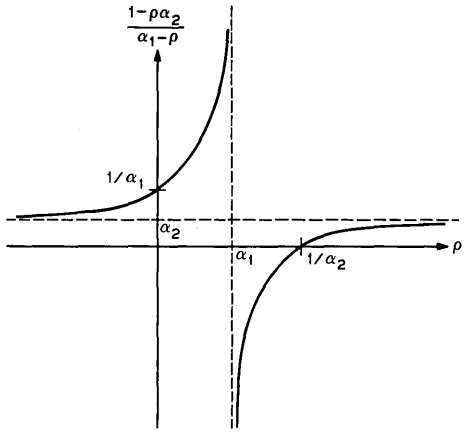


Fig. 3—Plot of $(1 - \rho\alpha_2)/(\alpha_1 - \rho)$.

$[f_2(\rho v_1)/f_1(v_1)]|_{\rho=0} = 0$. Clearly if

$$\text{and} \quad \left.\begin{aligned} \frac{f_2(\alpha_1 v_1)}{f_1(v_1)} &\leqq \frac{1}{\alpha_1} \\ \frac{f_2\left(\frac{1}{\alpha_2} v_1\right)}{f_1(v_1)} &\geqq \alpha_2 \end{aligned}\right\} \forall\ v_1 \neq 0 \tag{15}$$

then no solution of (13) exists. Since $g(\rho)$ is continuous in ρ, $g(0) > 0$, and $g(\rho) \neq 0\ \forall\ \rho > 0$, it follows that $g(\rho) > 0\ \forall\ \rho \geqq 0$. Hence sufficient conditions for the positivity of $v^t TF(v)$ are:

$$\text{and} \quad \left.\begin{aligned} \frac{f_{2i}(\alpha_1^i v)}{f_{2i-1}(v)} &\leqq \frac{1}{\alpha_1^i} \\ \frac{f_{2i}(v)}{f_{2i-1}(\alpha_2^i v)} &\geqq \alpha_2^i \end{aligned}\right\} \forall\ v \neq 0 \quad \text{and} \quad i = 1, 2, \cdots, n. \tag{16}$$

Remark:

$$\left.\begin{aligned} \frac{f_{2i}(\rho v)}{f_{2i-1}(v)} &< \frac{1 - \rho\alpha_2}{\alpha_1 - \rho} \quad \text{for} \quad 0 < \rho \leqq \alpha_1 \\ &> \frac{1 - \rho\alpha_2}{\alpha_1 - \rho} \quad \text{for} \quad \frac{1}{\alpha_2} \leqq \rho < \infty \end{aligned}\right\} \forall\ v \neq 0 \tag{17}$$

is both necessary and sufficient for $v^t TF(v) > 0$. However, the practical value of such a condition is limited.

2.2 *Necessary and Sufficient Conditions for the Exponential Type of Nonlinearities*

The functions $f_i(\cdot)$ in the conventional Ebers-Moll model and in the charge-control model[4] are of the following form:

$$f_i(v) = a_i (\exp \lambda_i v - 1). \tag{18}$$

Such exponential nonlinearities are subsumed in a class of nonlinearities with the following properties:

(*i*) $f_i(v_i) = a_i f(\lambda_i v_i)$, $i = 1, 2, \cdots, 2n$, where a_i and λ_i are positive constants.
(*ii*) $f(\cdot)$ is a monotone, strictly increasing function.
(*iii*) $f(0) = 0$.
(*iv*) $\lim_{v \to -\infty} f(v) = -1$.
(*v*) $\lim_{v \to \infty} [f(\rho v)/f(v)] = 0$ for all ρ where $0 < \rho < 1$.

When the transistors' static diode characteristics fall under this class then the following is true:

necessary and sufficient conditions for $v^t TF(v) \geqq 0$ ($v^t TF(v) = 0$ only when $v = 0$) are

$$\alpha_1^i \leqq \frac{a_{2i-1}}{a_{2i}}, \frac{\lambda_{2i-1}}{\lambda_{2i}} \leqq \frac{1}{\alpha_2^i} \tag{19}$$

for $i = 1, 2, \cdots, n$.

Sufficiency follows from the definitions and (16). For the necessity part consider the case when $v_{2i-1} \to -\infty$ with the remaining $2n - 1$ variables fixed. It is necessary that the coefficient of v_{2i-1} in the expansion of $v^t TF(v)$ approach a nonpositive value, i.e.,

$$\lim_{v \to -\infty} \{f_{2i-1}(v) - \alpha_1^i f_{2i}(v)\} \leqq 0 \tag{20}$$

for $i = 1, 2, \cdots, n$.
Similarly,

$$\lim_{v \to -\infty} \{f_{2i}(v) - \alpha_2^i f_{2i-1}(v)\} \leqq 0 \tag{21}$$

for $i = 1, 2, \cdots, n$. Hence

$$-a_{2i-1} + \alpha_1^i a_{2i} \leqq 0$$

and,

$$-a_{2i} + \alpha_2^i a_{2i-1} \leqq 0$$

or

$$\alpha_1^i \leqq \frac{a_{2i-1}}{a_{2i}} \leqq \frac{1}{\alpha_2^i} \tag{22}$$

for $i = 1, 2, \cdots, n$. To obtain the remaining conditions consider equation (12) in the notation of this section. By property (v),

$$\lim_{v_{2i-1}\to\infty} \frac{f_{2i}(\rho v_{2i-1})}{f_{2i-1}(v_{2i-1})} = 0$$

for all ρ such that $\rho < \frac{\lambda_{2i-1}}{\lambda_{2i}}$ and $i = 1, 2, \cdots, n$.

Hence, a necessary condition for the positivity of $v^t TF(v)$ is: $1 - \rho\alpha_2^i > 0$ is implied by $\rho < \lambda_{2i-1}/\lambda_{2i}$; i.e.,

$$\frac{\lambda_{2i-1}}{\lambda_{2i}} \leqq \frac{1}{\alpha_2^i} \tag{23}$$

for $i = 1, 2, \cdots, n$. Repeating the argument with the roles of v_{2i-1} and v_{2i} interchanged yields

$$\alpha_1^i \leqq \frac{\lambda_{2i-1}}{\lambda_{2i}} \tag{24}$$

for $i = 1, 2, \cdots, n$. This concludes the proof.

Remark 1: It is clear from the derivation that the conditions stated in equation (19) with the ratios replaced by the appropriate limits are necessary for the positivity of $v^t TF(v)$ when $\{f_i(\cdot)\}$ are general monotone, strictly increasing functions.

Remark 2: If condition (19) is violated then there exists a v_1, $\| v_1 \| < \infty$, such that $v_1^t TF(v_1) < 0$. The proof is as indicated below.

Consider equation (12) and say $\lambda_1/\lambda_2 > 1/\alpha_2$; then for ρ such that $1/\alpha_2 < \rho < \lambda_1/\lambda_2$ it follows from property (v) that there exists a V such that for $v_1 > V$, $g(\rho) < 0$. Likewise, for $a_1/a_2 > 1/\alpha_2$ and $\rho(\rho > 0)$ such that $(1 - \alpha_1(a_2/a_1)) + \rho(a_2/a_1 - \alpha_2) < 0$ it follows from property (iv) that there exists a V such that for $v_1 < V$, $g(\rho) < 0$. For the remaining possibilities the proof follows by interchanging the roles of v_1 and v_2 in the definition of $g(\rho)$ and proceeding as above.

III. PASSIVITY

This section is devoted to a discussion on condition (19) which has been shown to be both necessary and sufficient for the passivity of

transistors for which the static diode characteristics are of the form

$$f_i(v_i) = a_i(\exp \lambda_i v_i - 1) \tag{25}$$

in which a_i and λ_i are positive constants. When the temperature in the transistor is uniform, then from the well-known physical model,[5] $\lambda_i = (q/\mathrm{K}T)$ where q, K, and T are respectively the electron charge, Boltzman's constant, and temperature. However, when a temperature differential exists in the transistor junctions, for instance in a p-n-p transistor the temperature in the neighborhoods of the p-n and n-p junctions are respectively T_1 and T_2, then the same physical model holds[6] and

$$\lambda_{2i-1} = \frac{q}{\mathrm{K}T_1} \quad \text{and} \quad \lambda_{2i} = \frac{q}{\mathrm{K}T_2}$$

so that

$$\frac{\lambda_{2i-1}}{\lambda_{2i}} = \frac{T_2}{T_1}. \tag{26}$$

Hence when the ratio of the temperatures at the base-collector and emitter-base junctions of at least one transistor, say the ith, lies outside the interval $[\alpha_1^i, 1/\alpha_2^i]$ then there exists a $2n$-vector v_1 such that $v_1^t TF(v_1) < 0$. Furthermore, by Lemma 1, there exists a positive definite matrix G such that the solution of the network equation (3) with $b \equiv 0$ and initial condition v_1 is $v(t) \equiv v_1$. Then,

$$\left.\begin{matrix}\text{power delivered to the} \\ \text{passive subnetwork}\end{matrix}\right\} = v_1^t G v_1 = -v_1^t TF(v_1) > 0.$$

The temperature dependence of the a_i's may similarly be exploited to deliver power. The phenomenon of differential heating of transistors as described above is the basis of the thermocouple effect and thermoelectric generators.[5]

IV. BOUNDED INPUT—BOUNDED OUTPUT STABILITY

It is shown in this section that for a positive definite matrix G, bounded inputs, and passive transistors there exists a bounded neighborhood of the origin which, loosely speaking, is sure to contain the forced response of the network. The norm used is the Euclidean norm.

Now

$$v^t TF(v) > 0 \quad \forall v \neq 0 \tag{27}$$

and let

$$|| b(t) || \leqq \mathrm{K}. \tag{28}$$

$L(v)$ is as defined in equation (4). For the forced system, $v \neq 0$,

$$\begin{aligned} \frac{d}{dt} L(v) &= -v^t TF(v) - v^t Gv + v^t b \\ &< -v^t Gv + | v^t b | \\ &\leqq -v^t Gv + \mathrm{K} || v || \quad \text{by Schwarz's inequality} \\ &\leqq -\lambda_{\min} || v ||^2 + \mathrm{K} || v || \end{aligned} \tag{29}$$

where $\lambda_{\min}$ is the smallest eigenvalue of G and is positive by assumption. Hence when $||v|| \geqq \mathrm{K}/\lambda_{\min}$, $(d/dt)L(v) < 0$. Thus there exists a T such that for $t > T$,

$$|| v(t) || < \frac{\mathrm{K}}{\lambda_{\min}}.$$

V. A PROPERTY OF TRANSISTOR NETWORKS

Let $R = \{v' \mid (v' - v)^t T[F(v') - F(v)] > 0 \ \forall v \neq v'\}$. Then: (*i*) if G is positive definite, v' and b' are constant vectors such that $TF(v') + Gv' = b'$, and $v' \in R$, then for all inputs which approach b' as $t \to \infty$, the state vector approaches v' independent of initial conditions; (*ii*) there exists a passive subnetwork for which the overall network is such that if the terminal state corresponding to an input approaching a constant vector is not in $\bar{R}$, the closure of R, then there exists another input approaching the same constant vector for which the terminal state is different.

Remark: It has been pointed out by Sandberg[1] that the independence of the steady state from initial conditions is a basic property, and in Ref. 1 it has been proved that if (T, G) belongs to a certain class then the region R in statement *i* extends to the entire state space of the network. However, the region R defined here is independent of G.

In switching circuits with "memory" the dependence on initial conditions is a salient feature. Statement (*i*) observes that the design of such circuits be such that the steady states lie outside the region R, i.e., the steady state bias voltages violate the conditions which define the region R. Of course, the fact that the transistors are passive merely implies that R is non-empty, and their use in switching circuits with "memory" is not precluded. Statement (*ii*) states that in a certain sense

the region R is maximal.

Proof: (*i*)

$$TF(v') + Gv' = b' \tag{30}$$

and

$$\begin{aligned}\frac{d}{dt}C(v) + TF(v) + Gv - b' &= \frac{d}{dt}\{C(v) - C(v')\} \\ &\quad + T\{F(v) - F(v')\} + G(v - v') \\ &= b(t) - b'\end{aligned} \tag{31}$$

from (3) and (30). Define

$$\bar{v} \triangleq v - v', \tag{32a}$$

$$\bar{C}(\bar{v}) \triangleq C(v) - C(v'), \tag{32b}$$

and

$$\bar{F}(\bar{v}) \triangleq F(v) - F(v'). \tag{33}$$

From (31),

$$\frac{d}{dt}\bar{C}(\bar{v}) + T\bar{F}(\bar{v}) + G\bar{v} = b(t) - b'. \tag{34}$$

Note that $\bar{C}_i(\cdot)$ for $i = 1, 2, \cdots, 2n$ are monotone, strictly increasing mappings of the real interval $(-\infty, \infty)$ onto itself and $\bar{C}_i(0) = 0$; also there exists $\bar{C}_i^{-1}(\cdot)$. Since $v' \in R$, $\bar{v}^t TF(\bar{v}) > 0$. Also by assumption for any $\epsilon > 0$ there exists T such that

$$|| b(t) - b' || \leqq \lambda_{\min} \epsilon \quad \text{for} \quad t \geqq T \tag{35}$$

where $\lambda_{\min}$ is the smallest eigenvalue of G. On applying the results on bounded input-bounded output stability, as stated in Section IV, to the system described by equation (34) it follows that there exists a T' such that for $t > T'$,

$$|| \bar{v}(t) || = || v(t) - v' || < \frac{\lambda_{\min}\epsilon}{\lambda_{\min}} = \epsilon \tag{36}$$

and the proof is complete.

(*ii*) It will be shown that if $v_\infty \notin \bar{R}$ then there exists a $G > 0$ such that

$$TF(v_\infty) + Gv_\infty = b_\infty , \tag{37a}$$

$$TF(v_1) + Gv_1 = b_\infty , \tag{37b}$$

and

$$v_\infty \neq v_1 \,. \tag{37c}$$

In (37a) v_∞ denotes the terminal vector corresponding to an input approaching the constant vector b_∞. The solution of the network equation (3) with $b \equiv b_\infty$ and initial condition v_1 is $v(t) \equiv v_1$.

The proof is by simply observing from the definition of R and Lemma 1 that there exist v_1 and $G > 0$ such that

$$T\{F(v_\infty) - F(v')\} + G(v_\infty - v') = 0. \tag{38}$$

At least when the static diode characteristics are of the conventional exponential type the region R is easily obtained as shown below:

$$(v' - v)^t T\{F(v') - F(v)\} = \bar{v}^t T D_{v'} F(\bar{v}) \tag{39}$$

where

$$\bar{v} = v' - v$$

and

$$D_{v'} = \text{diag}\,(\exp \lambda_1 v_1' \,, \exp \lambda_2 v_2' \,, \cdots \,, \exp \lambda_{2n} v_{2n}').$$

The necessary and sufficient conditions for the positivity of $\bar{v}^t T D_{v'} F(\bar{v})$

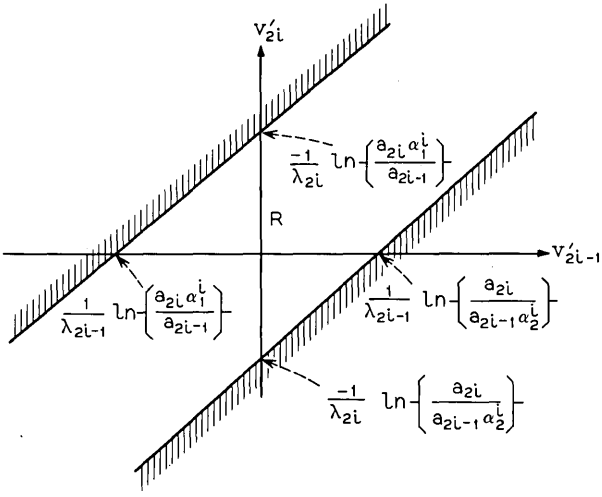


Fig. 4—Region R for the ith transistor (exponential diode characteristics).

are, from (19),

$$\alpha_1^i \leqq \frac{\lambda_{2i-1}}{\lambda_{2i}}, \qquad \frac{a_{2i-1} \exp(\lambda_{2i-1} v'_{2i-1})}{a_{2i} \exp(\lambda_{2i} v'_{2i})} \leqq \frac{1}{\alpha_2^i} \tag{40}$$

for $i = 1, 2, \cdots, n$. (40) is equivalent to

$$\left.\begin{aligned} \alpha_1^i \leqq \frac{\lambda_{2i-1}}{\lambda_{2i}} \leqq \frac{1}{\alpha_2^i} \\ \ln\left\{\frac{a_{2i}\alpha_1^i}{a_{2i-1}}\right\} \leqq (\lambda_{2i-1} v'_{2i-1} - \lambda_{2i} v'_{2i}) \leqq \ln\left\{\frac{a_{2i}}{a_{2i-1}\alpha_2^i}\right\} \end{aligned}\right\} \tag{41}$$

for $i = 1, 2, \cdots, n$. The region R is plotted in Fig. 4.

REFERENCES

1. Sandberg, I. W., "Some Theorems on the Dynamic Response of Nonlinear Transistor Networks," B.S.T.J., *48*, No. 1 (January 1969), p. 35.
2. Shichman, H., "Integration System of a Non-linear Network Analysis Program," IEEE Trans. Circuit Theory, August 1970, pp. 378–386.
3. LaSalle, J. P., and Lefschetz, S., *Stability by Liapunov's Direct Method*, New York: Academic Press, 1961.
4. Koehler, D., "The Charge-Control Concept in the Form of Equivalent Circuits, Representing a Link Between the Classic Large Signal Doide and Transistor Models," B.S.T.J., *46*, No. 3 (March 1967), pp. 523–576.
5. Feynman, R. P., Leighton, R. B., and Sands, M., *The Feynman Lectures on Physics, Quantum Mechanics*, Vol. III, Reading, Mass.: Addison-Wesley, 1965.
6. Gummel, H. K., private communication.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 8, October, 1971
Printed in U.S.A.

A Fast Amplitude Approximation For Quadrature Pairs

By G. H. ROBERTSON

(Manuscript received April 13, 1971)

This paper gives an algorithm from which an approximation to the amplitude of a quadrature pair can be obtained simply and without requiring more bits than are needed to represent the largest allowable amplitude, A.

If $A \cos \theta$ and $A \sin \theta$ are a quadrature pair, the amplitude is

$$A = (A^2 \cos^2 \theta + A^2 \sin^2 \theta)^{\frac{1}{2}}.$$

If AX is the larger magnitude of the pair and AY the smaller, an approximation to A is

$$A' = AX + \tfrac{1}{2}AY.$$

Computations using this algorithm with random values of θ showed that when the true amplitude was A, the mean obtained was $1.087A$, with standard deviation $0.031A$.

When A consists of a signal appreciably contaminated with noise so that several estimates are required before the signal can be detected reliably, no significant degradation ($\ll 0.1$ dB S/N) in detectability was found using the algorithm described above.

I. INTRODUCTION

When a discrete Fourier transformation is applied to samples from a data segment the results are usually desired in the form of phase and amplitude estimates of the components into which the data is resolved. When FFT[1] algorithms are used to perform the transformation the information from which these estimates are obtained is usually produced in the form of quadrature pairs of numbers $A \cos \theta$, and $A \sin \theta$, where A is the amplitude and θ the phase of the estimate. Obviously

$$A = (A^2 \cos^2 \theta + A^2 \sin^2 \theta)^{\frac{1}{2}}. \tag{1}$$

The intermediate step in this computation of squaring the components of the pair requires in general twice as many bits as are required to represent A. Computing the squares requires appreciable time as does the final square root computation. The situation is aggravated by the fact that there are as many pair members as the original data samples.

This paper describes the use of an algorithm which gives an approximation that in many if not all cases is adequately accurate and greatly reduces the computation time as well as promoting efficiency in the use of hardware by not requiring more bits than are needed to represent the final estimates.

II. DISCUSSION

Let the larger of $\cos^2\theta$, $\sin^2\theta$, be represented by X^2 and the smaller by Y^2. Then

$$\begin{aligned} A &= AX\left(1 + \frac{Y^2}{X^2}\right)^{\frac{1}{2}} \\ &= AX\left(1 + \frac{1}{2}\frac{Y^2}{X^2} - \frac{1}{8}\frac{Y^4}{X^4} + \cdots\right), \end{aligned} \tag{2}$$

where $1 \geqq Y/X$.
An approximation to A is thus

$$A' = AX + \tfrac{1}{2}AY\left(\frac{Y}{X}\right), \tag{3}$$

which suggests the computationally more attractive form

$$A' = AX + \tfrac{1}{2}AY. \tag{4}$$

Note that AX and AY are magnitudes and (4) implies that the amplitude of a Fourier component is approximately the sum of the larger and half the smaller of its quadrature pair magnitudes.

III. EVALUATION

Curve (a) in Fig. 1 shows a plot of (4) as θ ranges from zero to $\pi/4$ where the curve is reflected to $\theta = \pi/2$. The pattern is then repeated for the remaining three quadrants. The corresponding plot for (3) is shown in curve (b).

From the samples displayed on this figure the mean and standard deviations were calculated and found to be:

For (3), mean = 1.01, standard deviation = 0.0157, and
for (4), mean = 1.087, standard deviation = 0.031.

The algorithm described by (3) thus gives more accurate estimates (as expected) than that described by (4) which, however, is easier to implement and requires less computation.

When the Fourier analysis is carried out in an attempt to detect signal components contaminated by noise it is important to determine how processing techniques affect the detectability of such signals. In order to check this a program was prepared to compute performance estimates when the algorithm based on (4) was used. These estimates were plotted on standard ROC curves[2] for a linear detector and no significant degradation in detectability was observed. Extensive computations using the proposed algorithm have produced spectrograms of noise-contaminated signals that are undistinguishable from those obtained conventionally.

Use of the algorithm described in equation (4) results in a phase-dependent variation of the detector output. However, when the phase is uniformly distributed in the range $(0, 2\pi)$ the amplitudes are quite densely clustered around the mean because of the shape of the modulation characteristic. Eighty-five percent of the amplitudes lie within a range of about 0.4 dB. The accuracy obtained using the fast amplitude algorithm based on (4) may thus be adequate even when the estimates are not corrupted by noise.

IV. CONCLUSION

An algorithm has been described for computing the approximate amplitudes of Fourier harmonics from their quadrature pairs. The

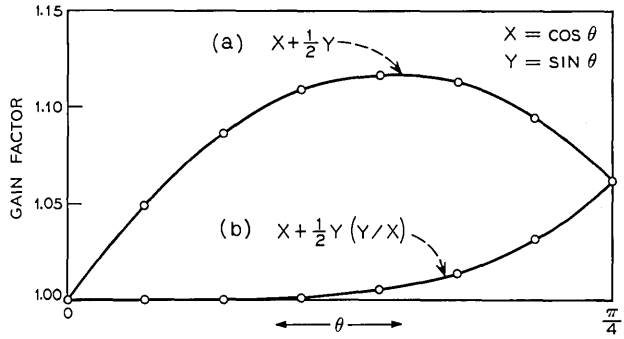


Fig. 1—Algorithm modulation characteristics.

algorithm is fast and does not require the use of more bits than are required to represent the largest permissible result. Use of the algorithm does not significantly degrade the detectability of signals contaminated by noise, compared with use of a perfect linear detector. Estimates obtained using the algorithm are amplitude modulated depending on their phase but the total modulation range is less than 1 dB. Representing the magnitude of the larger component of a quadrature pair by AX, and the smaller by AY, the algorithm is

$$A' = AX + \tfrac{1}{2}AY.$$

REFERENCES

1. Cooley, J. W., and Tukey, J. W., "An Algorithm for the Machine Calculation of Complex Fourier Series," Math. of Computation, *19,* No. 90 (April 1965), pp. 297–301.
2. Robertson, G. H., "Operating Charactertistics for a Linear Detector of CW Signals in Narrowband Gaussian Noise," B.S.T.J., *46,* No. 4 (April 1967), pp. 755–774.

Contributors to This Issue

BENJAMIN AVI-ITZHAK, B.S.M.E., 1955, Dipl. Ing., 1956, B.S.I.E., 1960, M.S., 1960, D.Sc., 1963, Technion-Israel Institute of Technology; Professor of Operations Research at the Faculty of Industrial and Management Engineering, Technion, Israel; Sabbatical leave at Bell Telephone Laboratories, 1970—. Mr. Avi-Itzhak is currently engaged in research on data traffic and processing systems. Member, ORSIS, ORSA, TIMS.

B. GOPINATH, M.S. (Mathematical Physics), 1964, University of Bombay, India; M.S.E.E. and Ph.D.(E.E.), 1968, Stanford University; Postdoctoral Research Associate, Stanford University, 1967–1968; Bell Telephone Laboratories, 1968—. Mr. Gopinath's primary interest, as a member of the Mathematics of Physics and Networks Department, is in the applications of mathematical methods to physical problems.

R. L. GRAHAM, B.S., 1958, University of Alaska; M.A., Ph.D., 1962, University of California (Berkeley); Bell Telephone Laboratories, 1962—. Mr. Graham's recent interests have been in the fields of combinatorial analysis, graph theory, finite structures and their applications to coding theory, and the analysis of algorithms. He is Head of the Discrete Systems and Control Department. Member, American Mathematical Society, Mathematical Association of America, Society for Industrial and Applied Mathematics, Sigma Xi, American Association for the Advancement of Science.

E. Y. Ho, B.S.E.E., 1964, The National Taiwan University; Ph.D., 1969, University of Pennsylvania; Bell Telephone Laboratories, 1969—. Mr. Ho has been engaged in developing and analyzing automatic equalizers for data transmission systems. Member, IEEE.

JACK M. HOLTZMAN, B.E.E., 1958, City College of New York; M.S., 1960, University of California (Los Angeles); Ph.D., 1967, Polytechnic Institute of Brooklyn; Hughes Aircraft Company, 1958–1963; Bell Telephone Laboratories, 1963—. Mr. Holtzman has worked in various aspects of systems and control theory and is the author of *Nonlinear System Theory—A Functional Analysis Approach* (Prentice-Hall, 1970). He is the supervisor of a group working on system theory and stochastic problems. Member, SIAM.

JAMES E. KESSLER, B.S.E.E., 1958, Purdue University; M.S.E.E., 1960, New York University; Bell Telephone Laboratories, 1958—. Mr. Kessler was involved initially with long-range studies of the waveguide transmission system, including cost studies and network configurational aspects. He has also participated in field trials and evaluation of the T1 carrier system. More recently, he has done data analysis of transmission parameters and customer interconnection planning.

DIETRICH MARCUSE, Diplom Vorpruefung, 1952, Dipl. Phys., 1954, Berlin Free University; D.E.E., 1962, Technische Hochschule, Karlsruhe, Germany; Siemens and Halske (Germany), 1954–57; Bell Telephone Laboratories, 1957—. At Siemens and Halske, Mr. Marcuse was engaged in transmission research, studying coaxial cable and circular waveguide transmission. At Bell Telephone Laboratories, he has been engaged in studies of circular electric waveguides and work on gaseous masers. He spent one year (1966–1967) on leave of absence from Bell Telephone Laboratories at the University of Utah where he wrote a book on quantum electronics. He is presently working on the transmission aspect of a light communications system. Member, IEEE, Optical Society of America.

D. MITRA, B.Sc.(E.E.), 1964, and Ph.D.(E.E.), 1967, University of London; United Kingdom Atomic Energy Authority Research Fellow 1965–1967; University of Manchester, U.K., 1967–1968; Bell Telephone Laboratories, 1968—. Mr. Mitra, a member of the Systems Theory Research Department, is interested in the application of mathematical methods to physical problems.

HENRY O. POLLAK, B.A., 1947, Yale University; Ph.D., 1951, Harvard University; Bell Telephone Laboratories, 1951—. Since joining Bell Laboratories, Mr. Pollak has been engaged in mathematical research in communications. He has written technical papers on analysis, function theory, and probability theory. At present he is Director, Mathematics and Statistics Research Center. Member, American Mathematical Society, Mathematical Association of America (formerly on its Committee on the Undergraduate Program), Advisory Board of the School Mathematics Study Group of the NSF Advisory Panel for Mathematical Sciences, International Commission on Mathematical Instruction.

GEORGE H. ROBERTSON, B.Sc., 1943, and Post Graduate Certificate (Natural Philosophy), 1948, University of Glasgow; Bell Telephone

Laboratories, 1948—. Until 1958 Mr. Robertson was engaged in electronics research and a variety of electron tube development projects. Since 1958 he has been working on signal propagation and processing studies in the Ocean Systems Research Department. Associate member, IEEE; member, AAAS.

R. A. SEMPLAK, B.S. (Physics), 1961, Monmouth College, N.J.; Bell Telephone Laboratories, 1955—. Mr. Semplak, a member of the Atmospheric Physics Research Department, has also been engaged in research on microwave antennas and propagation. He participated in the Project Echo and *Telstar*® communications satellite experiments. He currently is concerned with the attenuation effects of rain on propagation at 18.5 and 30.9 GHz. Member, Sigma Xi, Commission II of International Scientific Radio Union (URSI).

JOHN L. SULLIVAN, B.S.E.E., 1953, Iowa State University; M.S.E.E., 1959, Newark College of Engineering; Bell Telephone Laboratories, 1953—. Mr. Sullivan has worked on television test equipment and on establishing transmission objectives for exchange area trunks, wide-area data service, electronic telephone sets, and *Touch-Tone*® calling. He presently supervises a group involved in subjective testing to determine observer reaction to telephone message and television transmission impairments and in studying methods of objectively characterizing the transmission performance of telephone connections. In addition, he is involved in Bell System, national, and international standards activities. Member, Acoustical Society of America.

EDMOND J. THOMAS, B.S.E.E., 1964, and M.S., 1965, Rensselaer Polytechnic Institute; Rutgers University, 1968; Bell Telephone Laboratories, 1965—. Mr. Thomas has been engaged in studying the applicability of various adaptive control systems to the echo control problem. Specifically he has been characterizing the performance of these systems in nonlinear and time varying environments. He also has been involved in designing a computer testing system for measuring the electrical characteristics of telephone networks. Presently he is investigating various methods of transmitting data over microwave radio. Member, IEEE, Tau Beta Pi, Eta Kappa Nu, Sigma Xi.

GERARD WHITE, B.Sc., 1963, and Ph.D., 1966, University of Wales, Bangor; Bell Telephone Laboratories 1967—. Mr. White is a member of the Communication Principles Research Laboratory where his ac-

tivities have included time domain applications of Gunn effect devices, high-speed circuits, and optical communications systems. More recently his interests have been in the application of new device technologies to system studies. Member, IEEE Solid State Circuits Committee; senior member, IEEE.

ROGER I. WILKINSON, B.S.E.E., 1924, and Prof.E.E., 1950, Iowa State University; A.T.&T. Co. Dev. and Res. Dept., 1924–1934; Bell Telephone Laboratories, 1935–1968; Special Consultant to U. S. Air Force in Operations Analysis, 1943–45. Mr. Wilkinson has been engaged in the application of the theory of probability to telephone traffic problems since joining the Bell System. He was also concerned with the design of early sampling procedures for quality indexes in operating companies. He has contributed to studies of customer characteristics; the combination of probability distributions; and the extension of standard traffic theory to allow for day-to-day variations, repeated attempts, and nonrandom offered loads, by use of "equivalent random" procedures. At the time of his retirement in 1968 Mr. Wilkinson was Head of the Traffic Research Department. Fellow, IEEE; Member, American Statistical Association, Institute of Mathematical Statistics, Operations Research Society of America, American Society for Quality Control, Eta Kappa Nu, Tau Beta Pi, Pi Mu Epsilon, Phi Kappa Phi.

YU S. YEH, B.S.E.E., 1961, The National Taiwan University; M.S.E.E., 1964, and Ph.D., 1966, University of California, Berkeley; Harvard University, 1967; Bell Telephone Laboratories, 1967—. Mr. Yeh is a member of the Radio Transmission Research Department and is doing research work concerning mobile radio communication.

Bell System