THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 51 July-August 1972 Number 6

Fifty Years of BSTJ

It is appropriate, on this fiftieth anniversary of the founding of BSTJ, to recall the genesis and purpose of the publication, as set forth in the Foreword to Vol. 1, No. 1, of July 1922 (see opposite page).

This perceptive and prophetic note by the founders of the BSTJ states clearly their conviction that the art of telecommunication would require the ever-increasing application of scientific knowledge and the scientific method, and that this effort would, in turn, expand the boundaries of human knowledge in many fields. The BSTJ was planned to make these advances available to the world, treating a range of subjects as broad as the "science and technique of electrical communication itself."

How well has BSTJ performed this function? How best to catch the excitement and import of the technical papers that span these fifty years? One approach is to select a few out of the many outstanding papers to illustrate the forward thrust of telecommunications over the period. (See foldout.)

The early days

From the beginning, BSTJ authors saw their primary goal as extending the depth of understanding, and particularly of *quantitative* understanding, of the science and technology involved – to lay a solid foundation for the ever-growing nationwide telecommunication network.

This quantitative base extended from Harvey Fletcher's paper on "The Nature of Speech and Its Interpretation" to articles such as George Campbell's "Physical Theory of the Electric Wave-Filter," Harry Nyquist's "Certain Factors Affecting Telegraph Speed," and Clinton Davisson's "The Discovery of Electron Waves." With a systems view the fundamentals of a network combining efficient transmission and switching were worked out. Using the understanding

of telephony as a springboard, important contributions were made in a variety of related fields–movie-making, orthophonic recording, and television are examples.

In this era were laid the foundations for the great transmission advances: electric wave-filters, crystal filters and oscillators, Harold Black's epochal invention of stabilized feedback amplifiers, fundamental work on coaxial cable and waveguide systems, HF radio across oceans. All these were based on advances in understanding of the physics and mathematics involved. System applications were paced by the rapid sophistication of design of vacuum tubes. In the same period, panel and crossbar dial systems were developed to meet the needs for improved switching.

The middle period

The World War II years saw a great burst of application to military uses of the knowledge built up during the earlier period. Some of the basic electron tube and radio work reported in BSTJ in the late 30's paved the way for many of these advances. For example, this background made possible the development of the magnetron, newly invented in England, into a reliable generator of microwave pulses and a key to practical radar systems. "The Magnetron as a Generator of Centimeter Waves" by James Fisk, Homer Hagstrum, and Paul Hartman represents one important result of this work.

The mood of the post-World War II period was one of great confidence and expectation, fulfilled in a giant forward step – the discovery of the transistor. This was the key to unlocking the miracles of modern semiconductor electronics. Papers by William Shockley, John Bardeen, and Walter Brattain documented this advance. The transistor was destined to change radically all elements of the telephone network – transmission, switching, and customer systems. Outside the Bell System, it formed the cornerstone for a revolutionized electronics business and a huge new computer industry. Other major advances included Claude Shannon's "A Mathematical Theory of Communication," John Pierce's and Rudi Kompfner's traveling-wave tube work, Harald Friis' microwave antenna and repeater work, and Jack Morton's microwave triode.

The early 1950's saw also the beginning of the bold and massive effort to spread automatic switching throughout the nationwide network, as described in the paper "Automatic Switching for Nationwide Telephone Service" by A. B. Clark and Harold Osborne. And the modern approaches to materials science and engineering were firmly laid, leading to synthesis of a wide variety of new materials with properties especially tailored to meet communications needs – needs extending all the way from tough low-cost cable sheathing to exacting semiconductor properties.

In this period also was made the far-reaching decision that all new Members of Staff in the *development* areas at Bell Laboratories should receive advanced

training beyond the Bachelors level. (This was already the practice in the research areas.) This trend shows in BSTJ papers in later years, with many very scholarly and fundamental papers written by members of the development staff. The prophetic words of the "Foreword" of 1922 were to come true with a completeness that might have surprised those early contributors.

Since 1955

The most recent decade and a half shows a continuing evolution in system complexity to meet the telecommunications needs of a population whose telephone usage reached 170 billion calls in 1971. We start this period with the transistor finding its first truly widespread application, and in its wake, a rapidly maturing technology in solid state electronics and an expanding computer capability built firmly on this technology. The mood is one of excitement that "almost anything" in the way of new systems concepts is technologically achievable. Contributing to this mood are the discovery of the traveling-wave maser and the laser, of hard superconductors, and the ever-increasing sophistication in the understanding and use of materials. Parallel advancements in computer science pave the way to rapid development of languages and software systems which spark an explosion of computer applications to design, simulation, control, and manufacture.

In this period of BSTJ history, individual papers shine as before, but our evolving network complexity is suggested in a new way by the increasing number of special "systems" issues. The *Telstar*® Experiment issue documents man's first big step into satellite communications. The issue on No. 1 ESS describes the revolution in switching systems. And the wide range of technology and systems work to achieve a brand new two-way switched audio-video service is reviewed in the *Picturephone*® System issue.

Looking ahead

We enter the 70's with new building blocks such as charge-coupled devices, magnetic bubbles, miniature solid-state lasers, and minicomputers becoming available, and this reinforces our confidence that almost any technical challenge can be met. Within the field of switching, the burgeoning use of stored program electronic systems continues to point the way to new services. In transmission there is a strong feeling that we are on the threshold of another giant step - one which may have almost as widespread an effect on how we carry future telecommunication signals as the transistor has had on today's transmission network. That step is the emergence of low-loss optical fibers as a practical medium for information transmission.

But we also enter the 70's with increased awareness that our new system choices must show substantial economic and service margins over existing

systems if we are adequately to provide the new services that our advancing technological base promises. Topping past achievements is no small challenge, but the use of new technology to build on these accomplishments promises achievements yet unforeseen.

The sweeping observation of fifty years ago that "electrical communication touches upon almost every branch of science" seems equally applicable today. This observation also applies to the social sciences, as evidenced by increasing attention to the social impact of telecommunications and to the needs and aspirations of the individuals that comprise a large organization. Thus we can expect that in addition to continued "hard-science" advances in areas such as the basic understanding of materials, and of circuit, transmission, and switching theory, there will be increased contributions by BSTJ authors to solving relevant societal problems as well. A growing involvement in *operational* aspects of the nationwide network is one step in this direction.

As the BSTJ begins its second half-century one thing seems clear, even in this era of change. The flow of ideas, understanding, and concrete realization reported in its pages will continue to represent the main stream of progress foreshadowing the systems that will supply tomorrow's telecommunications demands efficiently, economically, and responsibly.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

Network for Block Switching of Data

By J. R. PIERCE

(Manuscript received December 4, 1970)

This paper describes a possible sort of data network for the transmission of addressed blocks of data between data terminals. The network consists of a number of closed rings or loops around which data blocks circulate. Data terminals on local loops can read data into empty blocks and read data addressed to them from full blocks. Buffered devices transfer data between local and regional loops and between regional loops and a national loop. Loops so interconnected need not be synchronized; they need not even operate at the same bit rate. Provisions are made for a "busy" signal, for special loops to carry heavy traffic between particular local loops, and for alternate routing.

I. GENERAL INTRODUCTION

1.1 *The Aim of This Paper*

This paper describes a data network in which addressed blocks of data find their way through a network of interconnected loops. The network does not use common control. Rather, simple customer terminals and simple devices which monitor and interconnect "loops" or "rings" would be added as such a network was brought into being. Thus, the cost and complexity would grow as the network grew.

1.2 *The Digital Communication Situation*

There is an increasing amount of digital traffic associated with computers and computerlike devices. Some of this traffic is highly intermittent but demands good transmission and quick response. Most networks at present available either offer quick response but monopolize a channel (private lines) or require an appreciable setup time (line switching).

One approach to handling intermittent, quick-response traffic is to transmit such traffic in blocks around loops or rings.[1-6] This paper describes a system in which simple customer terminals put blocks of

data onto or take blocks of data off of a loop or ring. Nonsynchronous loops or rings can be interconnected to form a widespread network.[6] This network takes digital circuits as they are; it does not impose special modulation formats as has been proposed.[5]

II. THE ENVIRONMENT OF A SWITCHED DATA NETWORK

The environment in which a switched data network would operate is different from the environment of telephone service. About all one can say for sure about "data" is that the input and output will appear in, or can be translated by the user into, binary form, and into serial binary form if need be. It is commonplace for the user or his machine to address messages, to look for addresses, to check for errors in various ways, and to perform other complex functions. Thus, user equipment can do, the writer believes, many things that the carrier might consider doing.

A data network would be different in another way. The line-switched telephone network is already there, in reach of wherever a data terminal or a data switching equipment might be. Hence, it is not at all necessary that messages concerning faults or the monitoring of operations be transmitted by the data network itself. The line-switched telephone network can be called in at any time.

Finally, in order to succeed, a data network must be able to grow gradually, gracefully, and economically, both geographically and in traffic capacity. Data channels for growth can readily be provided. Switching means will be needed at an increasing number of locations as the service grows. Switching means, which must of course be highly reliable, should be as simple and inexpensive as possible, consistent with providing satisfactory service to a community of sophisticated users.

III. SERVICE REQUIREMENTS

What service requirements should we put on an address-switched data communication network? What should be reserved for the carrier and what should be left to the user?

In addressing these questions we should remember that equipment or functions "left to" the user can as an option be supplied by the carrier through special "add on" equipment.

In the list below, several areas of requirements are discussed.

3.1 *Modulation and Timing*

The modulation and timing should be under control of the carrier;

otherwise there can be no assurance that transmission facilities will work or can be interconnected. The user should be provided with binary digits and a timing wave, which can be used in accepting digits or in putting digits into a terminal.

This argues strongly the desirability of performing all network functions other than transmission on binary streams, and providing modems to convert to binary at the ends of all transmission circuits which are used.

3.2 *Signal Format*

It is assumed that messages are transmitted by means of addressed blocks of binary digits, blocks of a common length. Certain positions within a block must be reserved for synchronizing and supervisory functions. Other positions should be accessible to the user for message digits and addresses. The sender's address might be put on by the carrier (to avoid misuse) or by the user. There seems no reason why the user should not supply the destination address.

3.3 *"Privacy"*

The carrier should supply in its switching system or data terminal means to assure that a message reaches only the customer to which it is addressed.

3.4 *Blocking*

A certain probability of blocking is inevitable.[7,8] However, provision should be made to prevent a single user from blocking a system for a prolonged time by transmitting continuously. This could be achieved by assuring that the user puts data into the system at an average rate considerably lower than the speed of the channel which serves him.[9] Thus, he would necessarily leave the channel idle after transmission of any block.

3.5 *Errors*

Nothing that the carrier can do can entirely prevent errors. The cost of error correction can be high. Different users may tolerate different error rates. The computer art is sophisticated. It seems best to leave error correction to the user. The carrier should endeavor to supply a low-error-rate service. If this could not be done over certain transmission links without error correction, the carrier might use error correction in these transmission links. The carrier can of course offer error correcting equipment as an option.

3.6 *Failure to Deliver*

Some messages are bound to fail to reach their destination, either because of blocking in transmission, errors in transmission, or because the user is out of operation or busy. The simplest way to deal with failure to deliver is acknowledgment of receipt. This can be left up to the user. If the user divides a long message into many blocks, he can, if he wishes, number the blocks and acknowledge only the first block, the end block, and loss of an intermediate block (when he receives a block whose number is not one greater than that of the preceding block which he received.) Such precautions could take care of the possible but unlikely time interchange of blocks—a block being delayed by alternate routing and arriving after a block which was sent later.

3.7 *User Busy*

In different systems, "user busy" can mean different things. It might merely mean that the user wasn't reading things out of a buffer fast enough. This would certainly happen if user equipment failed. The carrier could provide some notification to the sender if a terminal fails to accept a message intended for it.

3.8 *Failure of Service*

The system should somehow monitor its performance and request service when it fails. Alternate routing can allow service around a failed portion of the system. When the user believes that service has failed, the most sensible provision would seem to be a telephone number that would put him in touch with a computer or a person.

3.9 *Buffering at Data Terminals*

Some buffering will probably be necessary at data terminals to allow adequate notification of receipt of data and to allow notification of time when a block can be transmitted and correct timing of transmitted data. Buffering beyond this minimum required amount could be provided by the user, with an optional offering by the carrier.

IV. A PARTICULAR NETWORK

This section describes a particular data network. It makes use of closed digital loops, as in some other networks.[1–5] However, taken as a whole, it differs from other loop networks.[1–5] It is assumed that access to transmission facilities is through modems whose inputs and outputs are binary digits and a timing wave.

Figure 1 shows several interconnected loops which could form a small data network or a part of a large data network. Each loop is a data transmission channel. T1 lines look very attractive for local loops to which customers are connected. Other lines, such as 50-kilobit lines, could serve as trunk loops.

Three sorts of boxes appear in the loops:

Each loop has a box A which contains a clock and a buffer, so as to time and close the loop. A also performs other functions.

Unless a loop is a trunk loop, it also has boxes B which put blocks of data on and take blocks of data from the loop.

Loops are interconnected by boxes C, which transfer blocks of data from one loop to another and perform other functions.

Loops need not be synchronous, and speeds of transmission on different loops can be very different. Thus, transmission means can be fitted to the traffic on the loop. In transferring blocks from one loop to another, buffering will be provided to take care of differences in bit rate. Commonly, the buffer size will be one or more block lengths.

The operation of the network in Fig. 1 is best explained by considering particular features and blocks.

4.1 *Block Format*

Figure 2 shows a possible block format. The block is divided into several sections.

One section provides bits used for synchronization. These are written in by an A and used by B's and C's.

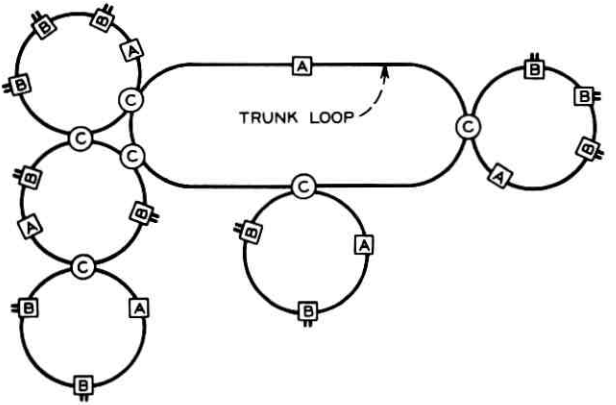


Fig. 1—Several interconnected loops in a data network.

Another section indicates whether the block is vacant or full; this reads vacant until a B or a C writes in *full* prior to writing address and message digits into the block.

Another part of the block is reserved for other supervisory marks; this space is used by the A's, B's, and C's, as will be explained later.

Other spaces in the block are reserved for destination address, sender's address, and message digits. These digits are accessible to the user on receipt of a message and, with the possible exception of sender's address, are supplied by the user.

4.2 *Box A*

An A contains a clock and circuitry for writing in the synch digits at the beginning of an outgoing block. In addition, it contains a buffer for storing the digits of an incoming block. The buffer can conveniently be a whole block long.

A block written by a B may have to pass an A in order to reach another B on the same loop. In this case, the A will simply read the bits of the block out from the buffer into the corresponding positions in the outgoing block. However, an A *marks* any full block that passes through it. When a marked block passes through an A, the A can interchange the sender's and destination addresses of any marked block. This returns the block to the sender and so provides a "busy signal" (the returned block) in case of any failure to deliver.

When the addresses are interchanged, the block should be marked so as not to be erased if it passes an A box once but to be erased if it passes *any* A box twice (rather than another interchange of address). This keeps undeliverable blocks from circulating endlessly through the network.

Rather than interchanging addresses in a marked block, an A could simply erase a marked block. In this case the system provides no "busy signal."

The A box should monitor the incoming signal and somehow send

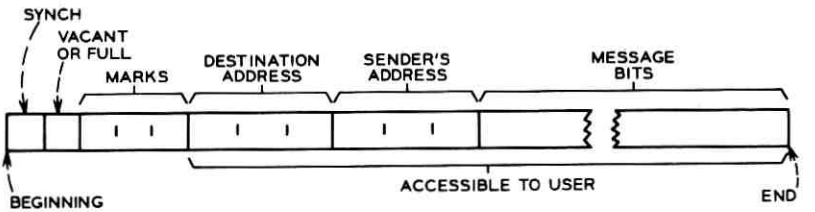


Fig. 2—Block format.

an alarm signal if there is none, or in the case of any other detectable malfunction.

4.3 *Box B*

The purpose of box B is to write data from a source into empty blocks and to read data from blocks.

B has to synchronize an internal counter or clock with the synchronizing pulses A writes in the blocks. This can be done by various methods.

In putting a message into a vacant block, B will have stored in its buffer part or all of the bits to go into the block. B will monitor the loop for a vacant block. When it finds one, it will (*i*) mark the block as full, (*ii*) signal its message source that it is transmitting so that more bits can be read in from the computer or other source, and (*iii*) write in address and message bits. B may write in the sender's address from the message source or from an internal store.

In taking a message from a block, B will examine the destination address of any full block. If this is the address of that B, the B will (*i*) signal the computer to accept bits, (*ii*) read address and message bits into buffer (if desired, message bits only can be outputted to the computer), and (*iii*) mark the block as empty and erase any harmful supervisory marks.

If desired, box B can signal the computer whenever a complete block has been received or transmitted, as well as when reception or transmission starts.

4.4 *Box C*

The purpose of box C is to transfer blocks from one loop to another. As the loops which box C interconnects may have different bit rates, buffering up to several block lengths may be necessary or desirable.

A C must decide whether to transfer a block or not. Thus, it must examine a part of the destination address of the message. The part it must examine is the address *of* the loop on which the destination lies, not the address *in* the loop on which the destination lies. The examination should certainly not involve a lengthy table lookup.

In Section 4.6 the writer describes one simple scheme. In this, the C need merely compare a part of the destination address with one particular address and determine whether the addresses are the same or different.

When a C transfers a block from one loop to another, it marks the block on the first loop as empty and makes any necessary erasures.

When a C effects a transfer, the block it puts onto the new loop is marked as not having passed an A; otherwise the block would be erased in passing an A in the new loop. The C, however, leaves intact a supervisory mark which shows that destination and sender's addresses have been interchanged.

An appropriate system layout will allow alternate routing; a message may be transferred by any of several C's on a loop, but not by others.

If the buffer of a C is full when a message which it might have transferred arrives, the C rejects the message. A message rejected by all C's on a loop could pass an A twice, and the A would either interchange sender's and destination addresses, so that the sender would get a "busy signal" (the returned block), or else erase the block. However, if the addresses had already been changed and the block so marked, the block would be erased on passing a particular A for the second time.

A C with a full buffer should signal for help. If the buffer is often full, there is danger of blocking. If the buffer stays full, something is wrong.

4.5 *Lost Blocks and Other Troubles*

Several things can cause the loss of a block.

Mutilation of address (through errors in transmission or in box function) can render a block undeliverable. Such a block will eventually be erased or returned.

Buffer overflow in a C may result in rejection of a block and subsequent destruction at an A. This is the *carrier's* fault; more buffering or more channel capacity could have prevented loss.

A block may reach the proper B but be rejected because of inadequate buffer capacity. If this bothers the recipient, he should add more buffer capacity.

A block may reach the proper B but the equipment for which it is intended may be down.

Errors introduced in going through the system may alter the address and send a block to the wrong destination. If the number of possible addresses is much larger than the number of addresses in use, this is not very likely to happen. Users can guard against false receipt of blocks by reserving some of their message bits for identification of desired messages, or by accepting messages only if the sender's address is in some restricted group.

4.6 *Systems Plan and Logic*

In the foregoing account the exact logic of the C's and the overall organization of the system have not been spelled out.

Figure 3 indicates the general features of the system plan. There are several hierarchies of loops—three in the figure. These have been labeled L (local), R (regional), and N (national). There is only one N loop. Various R loops connect to it by C's. L loops connect to the R loops by C's. L loops have B boxes on them to serve customers.

An address of a message destination (or a customer) consists of three parts:[10,11]

n_1 , Number *on* the local loop (L)

n_2 , Number *of* the local loop (L)

n_3 , Number of the regional loop (R).

Let us consider how a C should act in transferring blocks between loops:

(*i*) A block on an L loop is transferred to an R loop if its destination n_2 (number of the local loop) is *different from* the number of the L loop it is on.

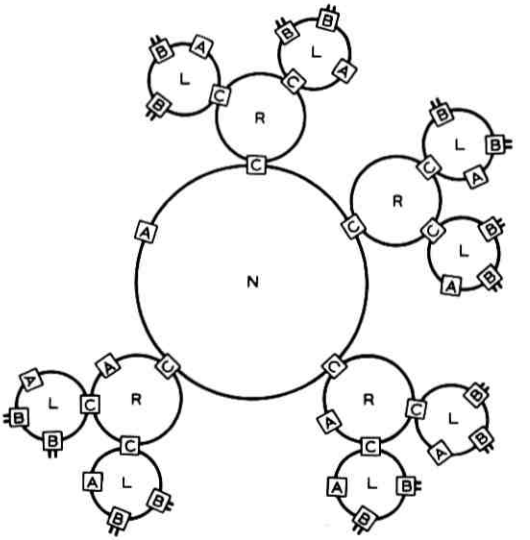


Fig. 3—General features of system plan.

(*ii*) A block on an R loop is transferred to the N loop if its destination n_3 is *different from* the number of the R loop it is on.

(*iii*) A block on the N loop is transferred to an R loop if its destination n_3 is *the same as* the n_3 of the R loop to which it can be transferred.

(*iv*) A block on an R loop is transferred to an L loop if its destination n_2 is *the same as* the n_2 of the L loop to which it can be transferred.

It is easy to see that this logic always gets a block to the destination address.

The above scheme does not provide alternate routing, and it channels all traffic between different regional loops or their local loops through the national loop. There may be heavy traffic between particular pairs of regional or local loops. To accommodate this, R or L loops can be interconnected by special TRUNK loops, as shown in Fig. 4. The C's for such interconnection should be located so that a block passes them just before it reaches the C to a higher-order (N or R) loop, as shown in the figure. The C's to the TRUNK transfer a block only if the destination n_3 (for a TRUNK connecting R loops) or the destination n_2 and n_3 (for a TRUNK connecting L loops) is the same as the n_3 or n_2 and n_3 of the loop to which the TRUNK can take the block.

Sometimes one might wish to connect a physically remote R or L loop to an N or R loop. This might be done by a physical extension of the remote loop, as shown at the left of Fig. 5. It might be more economical to provide a trunk loop, as at the right of Fig. 5. Thus, blocks on

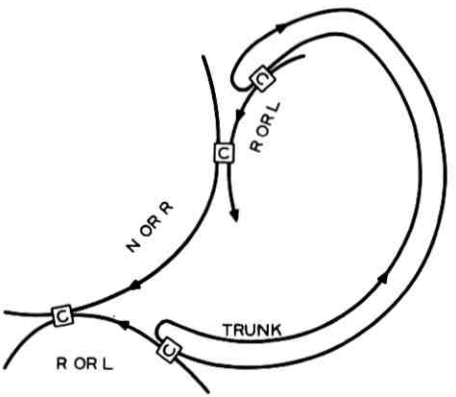


Fig. 4—Special TRUNK loop.

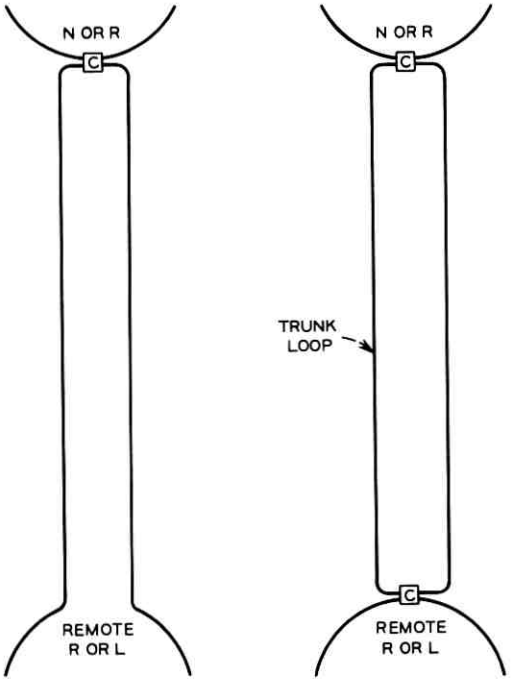


Fig. 5—Methods of connecting to a remote loop.

the remote loop would not have to travel the length of a long extension. The trunk loop might be a slower-speed loop than the remote loop. The two C's should have the same logic as the one C to the left.

What if the N loop fails? We can provide alternate routing by interconnecting regional loops with ALTERNATE loops as shown in Fig. 6. The C's connecting the R loops to the ALTERNATE loop should come *after* a C connecting an R loop to the N loop. Ordinarily, a block whose destination n_3 is not the same as the n_3 of the R loop it is on will be transferred to the N loop. If this fails, the following C will transfer the block to another R loop, as shown in the figure. Further ALTERNATE loops can be arranged so that a block which fails to reach the N loop will be passed from one R loop to another until it finally reaches the right R loop.

Reliability can also be increased by making the N loop and all R loops double and interconnecting them as shown in Fig. 7.

The L loops can be made double, too, perhaps by using the common two-way equipping of T1 lines. In the case of the L loops, one of the

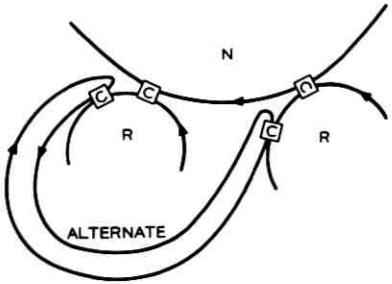


Fig. 6—ALTERNATE loop.

loops should be used in preference until it fails; then the A, the B's, and the C's should automatically shift to the other loop of the pair.

If a B loses power or fails, an automatic switch should bridge across it so as to keep the loop closed.

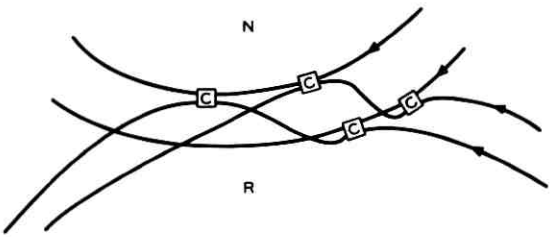


Fig. 7—Another method of increasing reliability.

4.7 *Blocking and Traffic Considerations*

The fact that material to be transmitted may be held for the appearance of a vacant block means that it is somewhat more likely that a full block will be followed by a full block than by an empty block. This must be taken into account in analyzing blocking.

It has already been mentioned that an increase in buffer size at a C should decrease blocking.

V. SOME CONCLUDING THOUGHTS

Unless the network is overloaded, the data network outlined in the preceding sections will get a block of data from source to destination very quickly.

Transmission of blocks of fixed length may seem restrictive, but it has a number of advantages. It allows easy synchronization at B's and C's and sure location of positions within the block. It lends itself

to acknowledgment and retransmission. Combined with a requirement that the average rate at which a user can transmit be considerably lower than the loop rate, it prevents one user from blocking a loop.

If the C's are to have a simple criterion for transferring blocks, the numbering of loops must satisfy some simple criterion. The freedom available through translation between directory and office numbers is not available. Translation would certainly run the cost up and might slow operation. Perhaps it is wise to forego translation in a data network, and stick to a carefully planned numbering scheme.

The data network has been deliberately kept very simple. It is multiprocessing with a vengeance. A few standard modules do everything. The logic and much of the circuitry of all A's is the same; adaptation is to the speed of the line. The same can be said for the B's and C's. Of course, for different types of channel, different modems will be needed in going to and from binary and in extracting and using the timing wave. Nonetheless, the network could be constructed of A, B, and C boxes and existing circuits and the modems for them.

REFERENCES

1. Company brochures on the Collins C-8500 Direct Digital Control System.
2. Grosser, H. K. M., and Schramel, F. J., "Data Transmission and Switching Equipment for the Seat Reservation System of United Air Lines (1)," Philips Telecommunication Review, *24*, Feb. 1963, pp. 13–24.
3. Martin, James, *Teleprocessing Network Organization*, Englewood Cliffs, N. J.: Prentice Hall, Inc., 1970, pp. 122–125.
4. Steward, Edgar H., "A Loop Transmission System," Proceedings of the 1970 International Conference on Communications, San Francisco, June 8, 9, 10, Institute of Electrical and Electronics Engineers, Inc. (70-CP-370-COM), pp. (36-1)–(36-9).
5. Farmer, W. D., and Newhall, E. E., "An Experimental Distributed Switching System to Handle High-Speed Aperiodic Computer Traffic," *Proceedings of ACM Symposium on Problems on the Optimization of Data Communication Systems*, Oct. 13–16, 1969.
6. Pierce, J. R., Coker, C. H., and Kropfl, W. J., "An Experiment in Addressed Block Data Transmission Around a Loop," IEEE International Convention Record, March 1971, pp. 222–223.
7. Hayes, J. F., and Sherman, D. N., "Traffic Analysis of a Ring Switched Data Transmission System," B.S.T.J., *50*, No. 9 (November 1971)., pp. 2947–2978.
8. Avi-Itzhak, B., "Heavy Traffic Characteristics of a Circular Data Network," B.S.T.J., *50*, No. 8 (October 1971), pp. 2521–2549.
9. Another solution is given in Kropfl, W. J., "An Experimental Data Block Switching System," B.S.T.J., this issue, pp. 1147–1165.
10. Kropfl proposes a slightly different way of handling n_2 and n_3 in Ref. 9 above.
11. Graham, R. L., and Pollak, H. O., in "On the Addressing Problem for Loop Switching," B.S.T.J., *50*, No. 8 (October 1971), pp. 2495–2519, propose a quite different addressing scheme.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

An Experimental Data Block Switching System

By W. J. KROPFL

(Manuscript received February 23, 1971)

In a companion paper, J. R. Pierce has described a novel data communication network which makes use of a hierarchy of interconnected rings or loops. The basic elements necessary to realize this network are called "A," "B," and "C" stations. Data blocks are circulated on closed loops formed by the interconnection of "A" and "B" stations. The "B" stations provide user access to the network while one "A" station on each loop performs supervisory functions. Isolated loops are interconnected by "C" stations. This paper describes an experimental hardware implementation of specific "A" and "B" station designs which operate via a T1 carrier system loop. A hog prevention technique is incorporated into the system which prevents any group of stations from monopolizing a loop. A likely "C" station realization and loop transferring algorithm is outlined. A bypass box which would automatically remove malfunctioning stations from the loop is also described.

I. GENERAL INFORMATION

A new type of data communication network has been described in a companion paper by J. R. Pierce.[1] It uses asynchronous multiplexing, buffered switching, and a distributed control system.

In this loop switching system, users are connected to the network by stations which are interconnected by a closed loop transmission line as shown in Fig. 1. Data is entered into and taken from the system in fixed-size blocks. Each data block consists of a unique synchronization word and a header which contains source and destination addresses as well as control information.

Each loop contains an "A" station which serves to close the loop, selectively repeating messages around the loop, and provides clocking and synchronizing information for all messages on the loop.

Another type of station, called a "B" station, utilizes the clock and

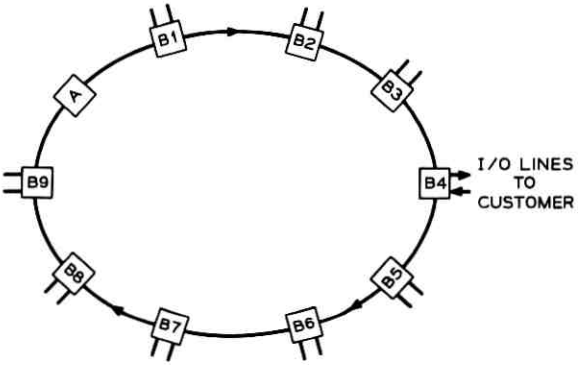


Fig. 1—Basic closed loop transmission network.

synchronizing information provided from the "A" station to write message blocks onto and read message blocks from the transmission loop.

A third type of station, called a "C" station, is used to transfer message blocks between loops. It also supervises the routing of messages through a maze of loops to their ultimate destination.

Laboratory models of the "A" and "B" stations have been designed and built to implement an experimental version of the data loop network described above. This paper describes the essential external features of this hardware and a probable "C" station configuration.

It should be noted the work reported on herein is solely a research project to test the feasibility of the concepts and to discover any unforeseen problems.

II. CONTROL FEATURES

2.1 *Message Format*

The message formats and headers shown in Fig. 2 were chosen to keep the "B" station as simple as possible, since it is the most numerous component in the system. The first three words after the start of the block SYNC word are used for supervisory and current (or local) loop addressing functions.

When traffic is confined to a single loop or to the same local loop, the local message format is used which consists of a local loop header immediately followed by N bytes of useful data. For any given network, N is fixed and was made equal to 54 for the loop described herein.

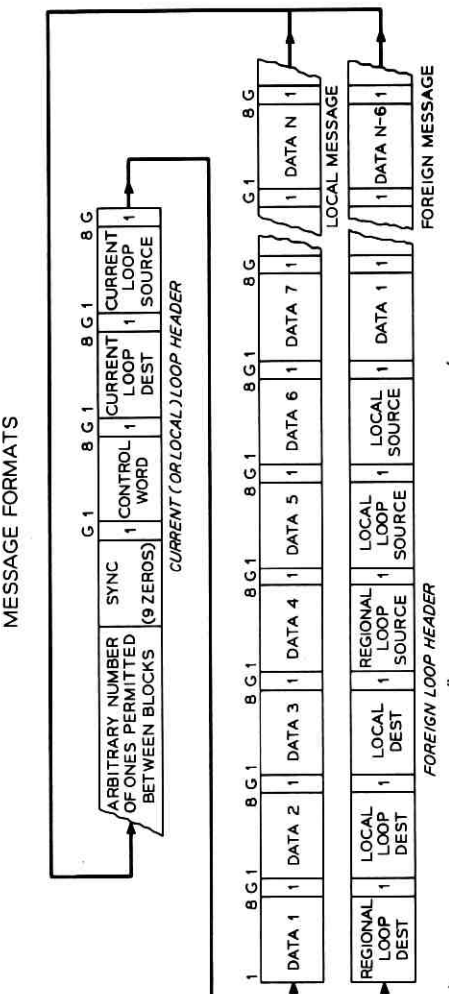


Fig. 2—Message formats.

When addressing a station on a different loop the foreign message format shown in Fig. 2 is used. It consists of a current loop header and a foreign loop header followed by $N - 6$ bytes of data. The contents of the foreign header is used by the "C" stations to pass the message from loop to loop to its ultimate destination.

The first word of each message block is a control word which is subdivided into fields that carry a coded representation of the status of the message block such as whether the block is vacant or full, private or common, and other conditions to be described later. The format of the Data Block Control Word (DBCW) and its control fields is summarized in Table I.

2.2 *Loop Synchronization*

The following synchronization scheme was chosen to make the system compatible with the Bell System's T1 carrier system.[2,3] As shown in Fig. 2, the message block is composed of 8-bit words which are always preceded by a guard bit. This prevents long strings of zeros which would cause the T1 carrier repeaters[4] to lose clock synchronism. The guard bit also allows one to construct a very simple loop synchronization scheme. A start-of-block sync pulse is generated whenever NINE consecutive zeros followed by a ONE are detected, with the result that an arbitrary number of ONEs are permitted between blocks. This allows the system to operate over large variations in loop length.

TABLE I—DATA BLOCK CONTROL WORD FORMAT

12	34	56	78		
TC					Type of Message Control Field (2 bits)
	LC				Loop Vacant-Full Control Field (2 bits)
		SB			Spare Bits (2 bits)
			HC		Hog Prevention Control Field (2 bits)
00				PM	Private Message
10				CM	*Common Message
01				UFM	Undeliverable Foreign Message, Foreign Source & Destination Interchanged
11				SCM	*Special Common Message, can be written only via an "A" or "C" Station
	00			VCC	Block Vacant
	10			FCC1	Block Full
	01			FCC2	Block Full and Passed an "A" Station Once
	11			FCC3	Block Full and Passed an "A" Station Twice, Current Loop S & D Interchanged

* Valid for within loop traffic only

III. "B" STATION FUNCTIONS

The main function of a "B" station is to provide an access port to the data loop network. It permits a user to read and write fixed-length message blocks under the supervision of a DBCW. Two general types of messages, namely private and common, are provided to facilitate message handling by the network. A private message, as the name implies, is used for personal or nonpublic communication between individual stations. The common message is used to broadcast the same message to a number of stations on the same loop. The main functional operations of a "B" station are described below and their corresponding logical equations are summarized in Table II.

3.1 *Reading Private Message Blocks*

If a "B" station detects a full block and recognizes the current loop destination address as its own, a message can be read from the data loop if the station's ReaD ReQuest line is enabled. The current loop source address and data is made available on the Parallel and Serial OUTput lines. A timing chart for both the Parallel and Serial Read STroBe lines and the terminal ReaDing gate is given in Fig. 3 together with a diagram summarizing the I/O signals needed to interface to a "B" station. The contents of the DBCW are stored in a register and made available during the entire read cycle on the Control Word OuTput lines. The station acknowledges reading a block by writing a block vacant mark into the DBCW's LC field.

3.2 *Reading Common Message Blocks*

A full block that is marked as being a common or special common message can be read by any station on the loop between the sender and the station nominally addressed, provided its Common Read ReQuest line is enabled. The nominally addressed station always marks the block empty to prevent its further propagation on the loop whether it is read or not. In other respects reading proceeds as described in the paragraph above. These messages can only be used for within-loop traffic. They will not be treated as foreign messages when received by a "C" station. A special common message is reserved for system use and can be written only by an "A" or "C" station. Its purpose is to provide a means by which system status information can be efficiently transmitted about a loop so that economical network management and supervision schemes can be implemented. For example, it would be useful in setting up loop testing, billing, and automated maintenance

TABLE II—SUMMARY OF A AND B STATION LOGIC

Text Key	For Station	If	Set DBCW To			Then
			TC	LC	H1-H2	
3.1	B	(VCCD−) (RDRQ+) (TDAD+)		VCC		Read DBCW, Source & Data
3.2	B	(VCCD−) (CRRQ+) (TC1D+)				Read DBCW, Source & Data
3.2	B	(VCCD−) (TDAD+) (TC1D+)		VCC		
3.3	B	(VCCD+) (WRRQ+) (ENWR+)	CWIN	FCC1		Write Dest, Source, Data & then set HPFF
3.5 (*i*)						
3.5 (*iii*)		Where (ENWR+) = (HCZD+) + (HPFF−)				
3.5 (*ii*)	B	(VCCD−) (WRRQ+) (HPFF−)			1	
3.5 (*v*)	B	(HCZD+)				Reset HPFF
3.5 (*iv*)	A				H2 0	
4.1 (*i*)	A	FCC1D+		FCC2	H2 0	
4.1 (*ii*)	A	FCC2D+		FCC3	H2 0	Interchange Local S & D
4.1 (*iii*)	A	FCC3D+		VCC	H2 0	

Mnemonics not used elsewhere in text:

ENWR	Enable Write
HCZD	HC Field Zero Detected
TC1D	TC Field Bit 1 Detected
TDAD	Terminal Destination Address Detected

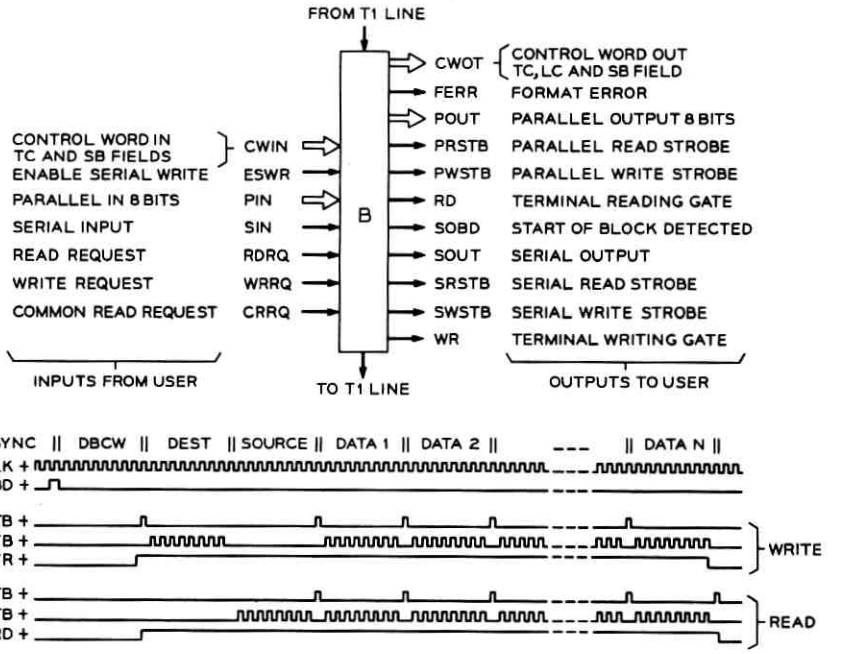


Fig. 3—"B" station I/O signals and timing diagram.

procedures. The common message feature can be useful on special purpose loops where a number of stations wish to share a common data base such as when driving slave-type display systems.

3.3 *Writing Message Blocks*

A message is written when a station, whose WRite ReQuest line is enabled, detects a vacant block. When this occurs a block-full marker is written into the DBCW's LC field. The type of message code to be written (see Table I) is determined by the status of the Control Word IN lines and is written into the DBCW's TC field. The current loop destination address and data are entered, from a user's buffer, through either the Parallel INput or Serial INput lines as determined by the logic level of the Enable Serial WRite line. The current loop source address is hard-wired within the station in order to provide positive sender identification. The timing for both the Parallel and Serial Write STroBe pulses and the terminal WRiting gate is given in Fig. 3.

3.4 *Format Error Detection*

The "B" station contains logic to continuously monitor the format

of the data loop. This is done by counting the number of words between consecutive start of block syncs. If the number of words is shorter than that needed to store a message block or longer than needed to store two blocks, a format error signal is generated. The "B" station also generates a format error if it detects a missing clock pulse. This error signal notifies the user as to the operational condition of the loop.

3.5 *Prevention of Data Loop Hogging*

Loop hogging can occur when certain send-receive patterns are established on a loop. For example, in Fig. 1, stations B2 and B5 through B8 cannot write messages onto the loop if B1 and B4 transmit continuously to B3 and B9 respectively. This problem can be solved if, after any "B" station sends a message, it is prevented from sending another message until all other write requests on the loop are acknowledged. This was implemented by manipulating a 2-bit Hog prevention Control (HC) field in the DBCW of each message block in the following manner;

(*i*) When a "B" station writes a message, a Hog Prevention Flip-Flop (HPFF) in that station is set to ONE. The block's HC field, however, is circulated on the loop unmodified.

(*ii*) If a "B" station on the loop has its WRite ReQuest line enabled when a full block is detected and if its HPFF is set to
  (a) ZERO, HC2 of the Hog prevention Control field is set to ONE;
  (b) ONE, nothing is done to HC2.

(*iii*) If a "B" station detects a vacant block and if the HC field is
  (a) ZERO, writing is independent of the state of HPFF;
  (b) NOT ZERO, the "B" station can write if and only if HPFF = 0.

(*iv*) When a data block passes an "A" station the contents of HC2 become the contents of HC1 and ZERO becomes the contents of HC2, i.e., (HC2) → (HC1) and 0 → (HC2).

(*v*) Whenever a "B" station detects an HC field whose contents are zero, its HPFF is reset; if it writes into this block, its HPFF will be set.

The scheme outlined above prevents any group of stations from monopolizing a loop. The service rendered is equitable in that the resources of the loop are divided equally among all of the users requesting service. If M users have their WRRQ lines enabled, each user is guaranteed that he will have to wait for no more than M vacant blocks to pass before he can write—provided no new WRRQ lines are enabled while the station is waiting. It should be noted at this time that vacant

blocks are generated, not only by a "B" station that reads a message addressed to it from the loop, but also by an "A" station (see Section 4.1) which detects three passes of the same message block. Considering this and the worst traffic situation that can occur on a loop, namely that all stations continually write messages to busy or nonexistent stations, we obtain, for an N-station loop, a maximum possible waiting time of $3N$ message block periods between consecutive write permits.

3.6 *Basic Delay Properties of Loop*

From a user's point of view, message string delay, D_S, is probably the most important variable by which system performance can be evaluated. For a given loop, D_S is the total time encountered in sending a message string consisting of K blocks between two stations on the same loop, and is given by

$$D_S = KD_A + D_P \tag{1}$$

where D_A is an average access delay for the writing interval and D_P is a propagation delay.

Propagation delay, D_P, is the time interval from when a message is written onto a loop until it arrives at its destination on the same loop and is simply equal to the number of delay elements between the sender and receiver. Each "B" station contains an 18-bit shift register which results in an 18-bit delay for each station on the loop.

Access delay, D_A, is defined as the time lapse between when a station requests and is subsequently granted permission to write a message block onto the loop. If a station requests service continuously, D_A is the time between two consecutive write permits for that station. It is a traffic-dependent random variable and as such requires statistical assumptions concerning user behavior to model it in a complete and rigorous manner. This random delay has been characterized in a study by J. F. Hayes and D. N. Sherman[5] for a data loop proposed by J. R. Pierce,[1] which did not include the above-mentioned anti-hogging scheme.

It has been demonstrated in the section above, however, that the constraint placed on loop traffic by the anti-hogging control scheme places a maximum limit on this delay of $3N$ blocks and is independent of user behavior. We thus have a worst case access delay of

$$(D_A)_{\max} < 3N * L/C \tag{2}$$

for a "B" station on a loop with the following characteristics:

N = number of stations on the loop,

L = block length in bits (including header, message body, guard bits, and filter bits),
C = loop bit rate capacity in bits/second.

The maximum propagation delay possible for a given loop is

$$(D_P)_{\max} < S * L/C \tag{3}$$

where S is the number of message block sectors circulating on the loop.

We now obtain a maximum message string delay of

$$(D_S)_{\max} < (3N * K + S) * L/C. \tag{4}$$

Although these results can be used to estimate worst case loop delays under saturated traffic conditions, they are much too conservative for use on local loops. Few stations on a local loop will send messages continuously. Equation (2) can therefore be multiplied by an average utilization factor,

$$\mu = 1/N \sum_{1}^{N} \mu_i \tag{5}$$

where μ_i is the probability that station i has its WRRQ enabled, and results in the following

$$D_A < 3\mu N * L/C \tag{6}$$

more realistic average access delay. Further analysis requires statistics concerning user behavior.

The worst case traffic assumption that led to the factor 3 in the above analysis is rather conservative; for example, it may be replaced by unity if circumstances are such that one can assume that all messages are properly addressed and encounter no busy stations.

IV. "A" STATION FUNCTIONS

4.1 *Supervision of Unclaimed Messages*

One of the primary functions of an "A" station is to dispose of undeliverable messages that occur due to being addressed to busy or nonexistent stations. This is done in the following manner:

(*i*) When a message passes an "A" station for the first time, its DBCW is marked to this effect as shown in Table I.

(*ii*) If the same message passes an "A" station twice, the destination and source addresses are interchanged and the DBCW marked to this effect. This sends the message back to the sender and thus serves as a station busy signal.

(*iii*) If the same message passes an "A" station for the third time, the block is marked vacant.

4.2 *Line Buffering*

In order to close the data loop, the "A" station must contain a loop closing buffer. The size of the buffer is dictated by the need to guarantee that the total loop bit capacity will always be large enough to store an integral number of data blocks. In the experimental model the buffer size is 512 bits and is fixed. The fixed buffer can result in a variable gap between messages, depending upon the number of stations on the loop. In an actual operating system it will probably be desirable to make this a variable length buffer in order to compensate for large variations in loop bit capacity due to changes in loop length. Loop length would be subject to temperature variations and changes due to taking malfunctioning stations off the loop. In any case, since the proposed synchronization scheme is insensitive to variations in loop length, this becomes a separate problem.

Another problem can present itself at an "A" station when closing the loop. Data errors will sometimes result due to phase difference variations between the "A" station's internal crystal clock and the received T1 carrier clock. In short loops this difference is nearly constant and can be compensated for by delaying the T1 clock by a fixed amount. This was done in the laboratory model. On long loops this phase difference will vary due to repeater-induced clock jitter as well as those variations discussed in the previous paragraph. This problem can be eliminated, using a four-bit elastic store[6] in the "A" station, by reading data into a buffer under the control of the incoming T1 repeater clock and reading out under control of the "A" station's crystal clock. The buffer must be initialized during a format loop cycle so that the buffer cell being read into is two bits removed from the cell being read out. This prevents data errors by eliminating the possibility that a buffer cell will be overwritten before its contents are used by the "A" station.

4.3 *Formatting*

When an "A" station detects a format error signal, it reformats the loop. In cases where it is possible to have only ONE data block circulating on the loop, reformatting is easily done by filling the line with all ONEs and then inserting a sequence of nine ZEROs. If the error condition persists, the loop is down and some maintenance procedures must be initiated.

In order to perform the various functions described above, an "A"

station needs a substantial portion of the logic contained in a "B" station. Therefore, the elements required to perform these operations are incorporated into one of the loop's "B" stations. A block diagram summarizing the additional logic needed is shown in Fig. 4.

V. "C" STATION FUNCTIONS

The primary purpose of a "C" station is to provide a means for interconnecting isolated loops. A likely realization of a "C" station is shown in Fig. 5. It consists of two "B" stations interconnected by a Buffer Memory and Controller. Messages destined for a station outside its own particular local loop are read and subsequently rewritten by "C" stations from loop to loop in the same manner as local interloop traffic. Buffering is needed because messages will often be delayed in going from one loop to another since messages already on the adjacent loop have the right of way. Buffering also has the desirable effect of allowing adjacent loops to operate at different bit rates.

If a network consists of a hierarchy of loops, a particularly simple foreign addressing scheme results. Such a network due to Pierce[1] is shown in Fig. 6. Individual subscribers are connected together by a local loop. The local loops are interconnected by the "LC" stations forming regional loops which are in turn connected by "RC" stations to form a national loop.

5.1 *Loop Transferring Algorithms*

To gain some insight into how a "C" station can be designed using the "B" station hardware described above, let us send a foreign message from X (R1, L2, B1) to Y (R4, L2, B2) in Fig. 6, examining the address portion of the message at each step.

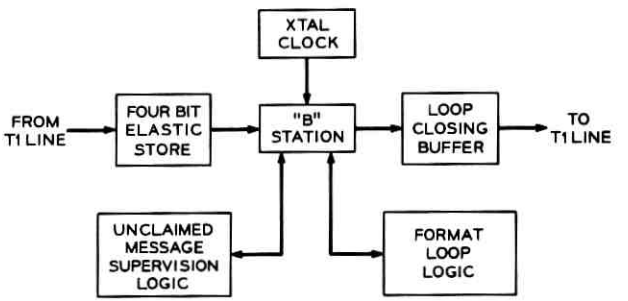


Fig. 4—"A" station.

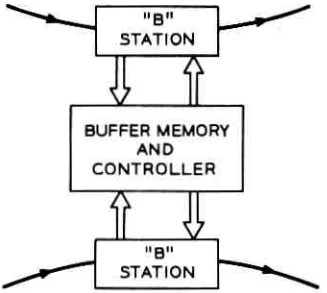
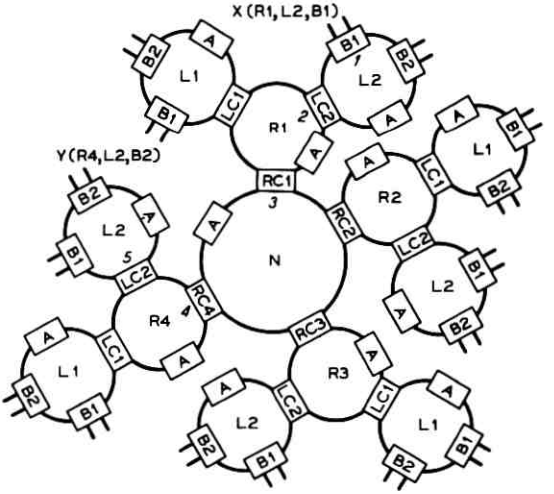


Fig. 5—"C" station.

Station X starts the process by writing a private message, with the header shown in *step 1* of Fig. 6, to its local loop "C" station. Note that all private messages sent and received by the local loop portion of a local "C" station are declared foreign.

HIERARCHICAL NETWORK OF LOOPS



FOREIGN MESSAGE HEADER

SYNC	DBCW	CLD	CLS	RLD	LLD	LD	RLS	LLS	LS
STEP									
1	X TO L	LC	.B1.	R4	L2	B2	---	---	---
2	L TO R	.RC.	.LC2.	R4	L2	B2	.RC1.	.LC2.	B1
3	R TO N	R4	.RC1.	R4	L2	B2	RC1	LC2	B1
4	N TO R	L2	.RC.	R4	L2	B2	RC1	LC2	B1
5	R TO L	B2	.LC.	R4	L2	B2	RC1	LC2	B1

Fig. 6—Foreign addressing scheme.

The message is then read by the addressed "C" station if its buffer is not full. The LC field of the DBCW is checked to make certain that the current loop source and destination addresses have not been interchanged, i.e., LC $\neq$ FCC3. The various control field codes are defined in Table I.

The message header is then operated upon by the appropriate procedure denoted by entry point SW shown in Fig. 7. The position of switch SW depends upon which loop is being traversed. The message with its new header is now written onto the adjacent loop. By following the process to its conclusion one can easily see that the algorithm given in Fig. 7 always gets the message to its designated destination address if no "C" station with a full buffer is encountered. The addresses bracketed by periods in Figs. 6 and 7 are defined and written at the step indicated by the network's "B" or "C" stations and cannot be altered by a customer. This is done to insure proper message source identification. They and the user defined addresses of *step 1* are subsequently manipulated by the network's "C" stations as shown in Fig. 6.

5.2 *Undeliverable Foreign Messages*

If a destination "B" station, or any "C" station, encountered by a message block is busy, the condition is detected by an "A" station which interchanges the current loop source and destination addresses and marks this fact by making LC = FCC3. As a result, the message is sent back to the last "C" station from which it came. This "C" station examines the contents of the DBCW and upon finding LC = FCC3 takes one of the following alternate corrective actions:

(*i*) If this is the message's first encounter with an obstacle, i.e., TC = PM, the message is read into the "C" station's buffer memory where its foreign source and destination addresses are interchanged. The message is also marked as being an undeliverable foreign message. The process then proceeds as directed by entry point SW in Fig. 7 and thus sends the message back to the sender.

(*ii*) If the message has already been marked as being an undeliverable foreign message, i.e., TC = UFM, the message is not read into the buffer memory. Therefore, any message that encounters two busy stations as it wanders through a maze of loops is destroyed.

5.3 *Buffer Status Controls the Relaying of Messages*

In the system above messages can be rejected and even destroyed by the system. A message should be rejected only due to encountering

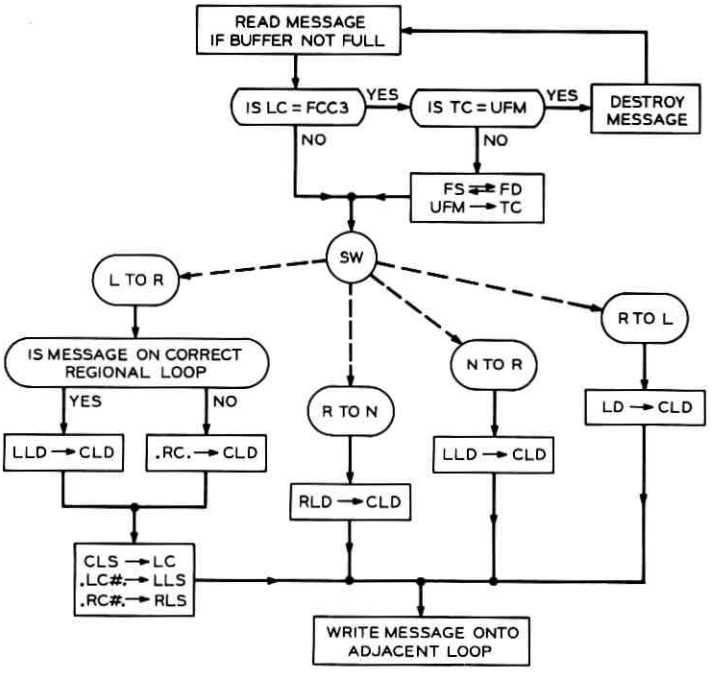


Fig. 7—"C" station loop transferring algorithms.

a busy destination "B" station. This ideal system will be approached as system "C" station buffer capacity is increased.

The following strategy can be employed to prevent message loss due to blockage by the system even in the presence of limited "C" station buffering.

Each "C" station buffer is divided into two sections—an upward buffer and a downward buffer. As their names imply these buffers pass messages up and down the hierarchy of loops. When one of these "C" station buffers becomes *nearly* full, its station sends a buffer full status signal, via a special common message, to all the other stations on the loop from which the buffer in question receives messages.

This buffer status information is used by the other "C" stations to prevent the relaying of messages to a station with a full buffer.

The buffer full status signal is sent before the buffer is completely full because messages destined for the station may still be in transition—the maximum number of messages in transit being the capacity of the loop.

The "C" station with the full buffer later sends a buffer clear status

signal to all stations on its loop when its buffer has room for additional messages.

5.4 *Redundant Loops*

The reliability and capacity of the hierarchical system just described would be substantially improved if redundant pathways were added. The alternate message routing and loop doubling schemes outlined in Pierce's paper[1] can be readily incorporated into the system for this purpose.

VI. BYPASS BOX

Loop integrity could be more readily insured and maintenance more easily performed if loop ByPass "BP" boxes were placed at strategic locations in the network. The logic of these boxes is given below. Such boxes would not only protect the loop against "B" station malfunctions but also protect the system against faulty repeaters and cables.

The "B" station in Fig. 8 would be bypassed by the "BP" box if any of the following conditions occur:

(*i*) Missing clock pulses are detected at IN2 but NOT detected at IN1 within a present period of time.

(*ii*) Varying data pulses are detected at IN1 but NOT detected at IN2 within a preset period of time.

(*iii*) Could be tripped manually or automatically from a central office to isolate some of the more subtle and unpredictable faults which will undoubtedly occur as in any system.

A "BP" box can be tripped only by a fault which occurs on the section of the loop it parallels. When it is tripped a delay equivalent to the bypassed section must be introduced into the loop. This could be done by having a fixed delay within the "BP" box. It may however be advantageous, especially from an installation cost viewpoint, to add some loop length measurement logic to an "A" station and have it

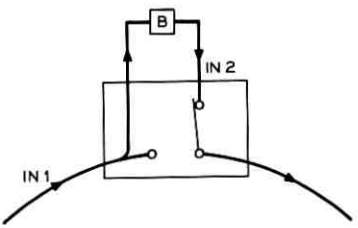


Fig. 8—Loop bypass box.

control a variable length loop closing buffer in order to compensate for the delay. Once a "BP" box has been tripped it can be reset only by the carrier. Normally the carrier would trip or reset a "BP" box only when the loop is being formatted. This is necessary to prevent the destruction of messages that may be currently circulating on the loop.

The notion of using the detection of missing clock pulses and the absence of missing data pulses to transmit information about major loop malfunctions are especially attractive because of the ease with which they can be implemented. Their use, however, requires careful consideration of T1 repeatered line performance when subjected to sparse or no input pulses.

6.1 *Repeater Performance with No Input Pulses*

If no input pulses are applied to a repeater, several possible consequences result. A number of these are summarized below.

A repeater may go into self-oscillation at approximately 10 KHz. This state would produce missing clock pulses which would be sensed by bypass condition *i* above.

An eight-out-of-eight pulse train may be generated due to cross talk from the clock circuit on the other side of the repeater. Line repeaters contain two complete regenerators in one case which share a common voltage regulator. This state would be sensed by bypass condition *ii*.

A regenerator which goes into self-oscillation may draw an excessive current from the voltage regulator thus adversely affecting the operation of its companion repeater. This condition must be eliminated or careful consideration must be given to how companion regenerators are used in the network.

A new T1 repeater[7] has recently been developed and is currently undergoing field trials. The availability of this repeater will eliminate many of the problems described above because it was designed not to oscillate during the absence of an input signal.

VII. HARDWARE PARTICULARS

Two "B" stations and an "A" station have been implemented and interconnected using T1 carrier system repeaters as a component part of the stations. The T1 system uses bipolar pulse transmission techniques to span up to 6000 feet between repeaters and has a bit rate of 1.544 MHz.

The "B" station was built with 56 chips of standard 7400 series TTL circuitry, 6 of which are MSI circuits. The "A" station contains all

the logic of a "B" station plus 7 additional chips needed for message supervision and reformatting control logic and a 512-bit MOS loop closing buffer. A photograph of an experimental "A" station is shown in Fig. 9.

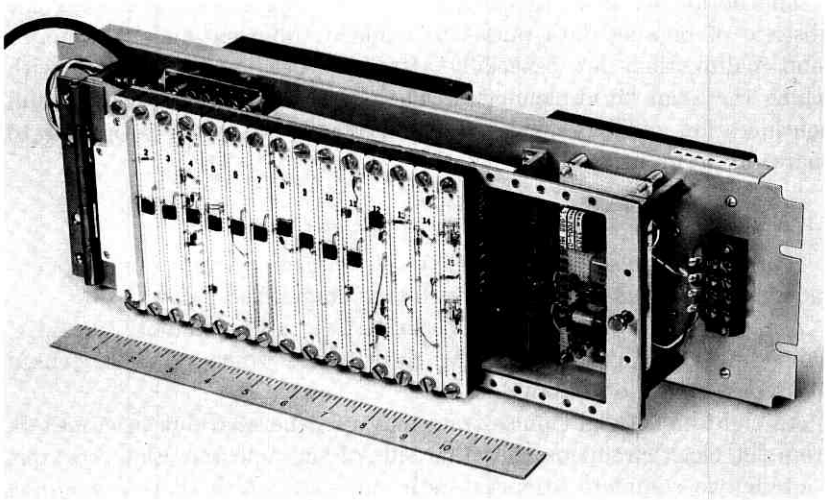


Fig. 9—Experimental "A" station.

A fixed message length of 522 bits was chosen because it allows a complete message to circulate on a loop made by closing an "A" station upon itself. The additional bits in excess of 512 come from the storage inherent within the "B" box portion of an "A" station. Therefore, each local message can accommodate 54 bytes of data while a foreign message contains 48 bytes of useful data. The bit rate and message block size can be easily altered if needed. The bit rate is dictated only by the characteristics of the T1 repeaters.

VIII. CONCLUSION

The system described above has been used to interconnect two DDP 516 laboratory computers. The computer interface and hardware for this system are described in a companion paper by C. H. Coker.[8] The equal sharing property of a loop and minimal constraints on the data format and simple addressing scheme allow the user a great deal of flexibility to structure the system to his needs.

REFERENCES

1. Pierce, J. R., "Network for Block Switching of Data," B.S.T.J., this issue, pp. 1133–1145.
2. Fultz, K. E., and Penick, D. B., "The T1 Carrier System," B.S.T.J., *44*, No. 7 (September 1965), pp. 1405–1451.
3. Mayo, J. S., "A Bipolar Repeater for Pulse Code Modulation Signals," B.S.T.J., *41*, No. 1 (January 1962), pp. 25–97.
4. Travis, L. F., and Yaeger, R. E., "Wideband Data on T1 Carrier," B.S.T.J., *44*, No. 8 (October 1965), pp. 1567–1604.
5. Hayes, J. F., and Sherman, D. N., "Traffic Analysis of a Ring Switched Data Transmission System," B.S.T.J., *50*, No. 9 (November 1971), pp. 2947–2978.
6. Members of the Technical Staff, Bell Telephone Laboratories, *Transmission Systems for Communications*, Fourth Edition, 1970, p. 616.
7. Cunningham, P. B., Durand, D. L., Lombardi, J. A., and Tarbox, R. A., "A New T-1 Regenerative Repeater," Bell Laboratories RECORD, *49*, No. 4 (April 1971), p. 109.
8. Coker, C. H., "An Experimental Interconnection of Computers Through a Loop Transmission System," B.S.T.J., this issue, pp. 1167–1175.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

An Experimental Interconnection of Computers Through a Loop Transmission System

By C. H. COKER

(Manuscript received October 13, 1971)

Two laboratory computers have been interconnected through an addressed-block data transmission system (ring) as described by J. R. Pierce and implemented by W. J. Kropfl. This paper gives an idea of the equipment, programming, and protocols of communication through that system.

I. INTRODUCTION

J. R. Pierce has described a digital communications system in which addressed messages are transmitted through a hierarchy of interconnecting loops or rings.[1] Components for one ring have been implemented by W. J. Kropfl.[2] We have used the ring to interconnect two laboratory computers.

1.1 *User's View of the Ring*

To the user, the system resembles a high-speed telegraph service. A message, headed by a destination address, can be "put on the wire," and a moment later, it will be delivered to the addressee. Inside the network, the message is multiplexed onto a loop of circulating message blocks. If the addressee is on the same loop as the sender, the message travels around until it reaches him. If the addressee is on another ring, the message is passed from ring to ring, up and down the hierarchy, until it reaches him. Thereupon it is removed and its place on the loop marked "empty."

But these internal details are invisible to the user. From his viewpoint, he can transmit directly to anyone he designates–without prior negotiation with the network to place a call, without worries of maintaining the connection for possible further communication, or of breaking it when finished.

II. HARDWARE: THE BASIC INTERFACE

Electrical communication with the network terminal ("B" station[1,2]) is simple. To send a message, the user signals the terminal that he is ready to transmit (A in Fig. 1). Milliseconds later, depending on network timing and traffic, the terminal begins sending back clock pulses (B), by which the sender is to shift into the terminal first an address and then data (C).

The message propagates through the network, ultimately to the addressee. If his receive ready line (D) is set, his terminal delivers the message, headed by the sender's address, serially (E) with clock pulses (B) for shifting.

The main constituents of a computer interface to this terminal are a shift register (F), for matching the parallel format of the computer to the serial ring; parallel full-computer-word buffers (G, H), for matching the narrow timing tolerances of the shift register to the more asynchronous responses of the computer; and logic (J), for controlling

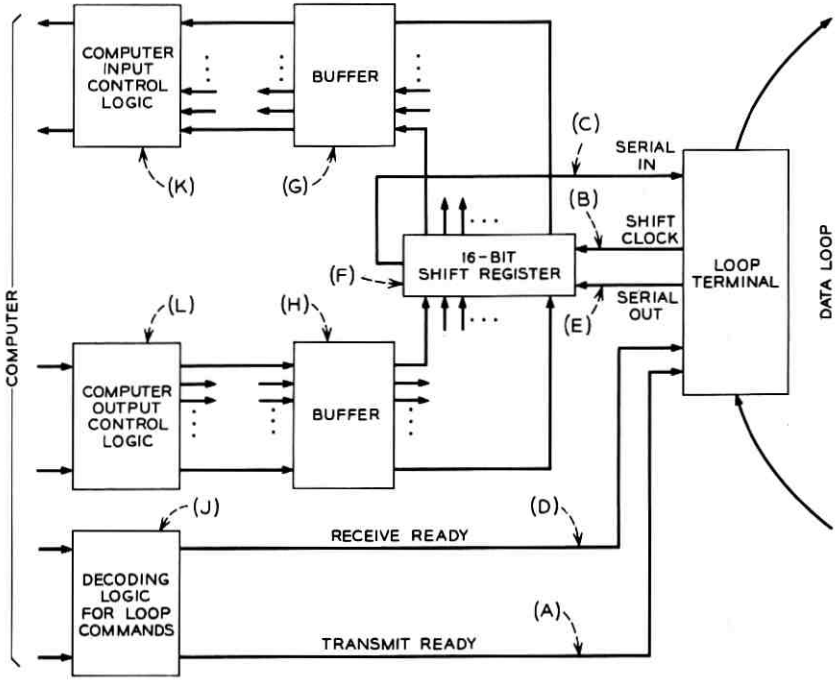


Fig. 1—Basic interface between a computer and the data loop.

read or write lines to the station, and for getting data into (K) and out of (L) the computer. Not shown in the figure are a counter, to segment serial bit streams into computer words, and assorted components to control buffer loading.

A basic interface for computer input-output at a fundamental level (through an "I/O bus") requires about 40 7400-series integrated circuit chips. Depending on the machine, some or most of these interface operations can be accomplished with less special design, using standard computer options.

The speed of the ring allows, in most cases, computer input-output to be done by block transfers through data channels, or to be programmed word-by-word. Some machines, however, are not fast enough to keep up with the T1 carrier loop system, transmitting word-by-word.

2.1 *Options: Buffering to Receive Unexpected Messages*

The simple interface above is adequate for most computers. However, for the very small and the very large, certain changes are desirable.

In a dedicated (one-user-at-a-time) computer, the situation can be controlled so that incoming messages are correctly anticipated and the computer is always prepared to receive them. In a time-shared multi-job computer, however, messages may arrive at any time–even when the computer is attempting to send a message.

One way to guarantee reception of ill-timed messages is to provide separate data channels of the computer for input and output (K and L in Fig. 1), so that the computer can prepare for both sending and receiving at the same time. A less expensive alternative is to provide means to rescind an output command upon appearance of an incoming message. When it recognizes an incoming address, the loop station produces a pulse that can be used for a computer interrupt. The interface must have enough buffer capacity to store incoming data until the interrupt request is honored and the channel or program readied for input. In the computer of this study (a DDP-516), worst-case delay for the sequence (without contemporaneous I/O in other data channels) is 30 μs. Three-word buffers (G in Fig. 1) are sufficient. Depending on the computer, buffers to receive messages by interrupt will require 10 to 20 7400-series chips.

2.2 *Hardware Addressing: Rejection of Intruding Messages*

A different type of unexpected message problem occurs when a machine that is prepared to receive a message from one sender gets a message from another. If the receiver is an unsophisticated device,

the device might be unable to detect the intrusion, or to recover from it in time to receive the desired message.

Hardware to discard messages from any but a designated sender is relatively simple using the 8-bit parallel output of Kropfl's station. Furthermore, the interface can be structured to demultiplex data and addresses onto separate lines, and to perform a similar multiplexing function for transmission.

The address of a correspondent can then be designated once, and will remain the same until redefined. The user has the option of treating the interfaced ring as an addressed-block system, by setting a new address for each transmission; or of treating it as a direct line, by leaving the address the same for as long as he wants. Data blocks can be any size–possibly much larger than a loop block. At the user's option the interface can reject intruding messages, or set off alarms or program interrupts.

Engineering for such features is straight-forward. About 10 integrated circuit chips are required if addresses of correspondents are to be set manually; 20 chips, if they are to be changeable under program control.

III. USE OF THE EXPERIMENTAL LOOP

Two identical Honeywell DDP-516 computers of the Bell Laboratories Acoustics Research facility have been connected to a local ring. The computers have 16 thousand 16-bit words of 0.96-μs memory; 800 thousand words of 3.3-megabit disk memory; hardware multiply, divide, double precision, and floating point; and printers, card readers, and analog-to-digital and digital-to-analog conversion equipment. One of the machines has 300- and 2000-baud *Dataphone*® data sets. The two machines are used for a variety of on-line applications in speech analysis, synthesis, and perception research.

3.1 *The Experimental Interface*

The two computer interfaces use 4-μs/word data channels and provide hardware multiplexing of data and addresses. In addition, the interface recognizes a special bit in the data block, and upon receipt of a block with that bit set, causes a program interrupt. These interrupt or "command blocks" simplify synchronization and provide more positive control of a remote computer in program debugging.

Figure 2 is a photograph of a loop station installed above its interface circuit. The interface is done in a card logic used for other devices on this computer. Engineering time was approximately 1.5 man-months.

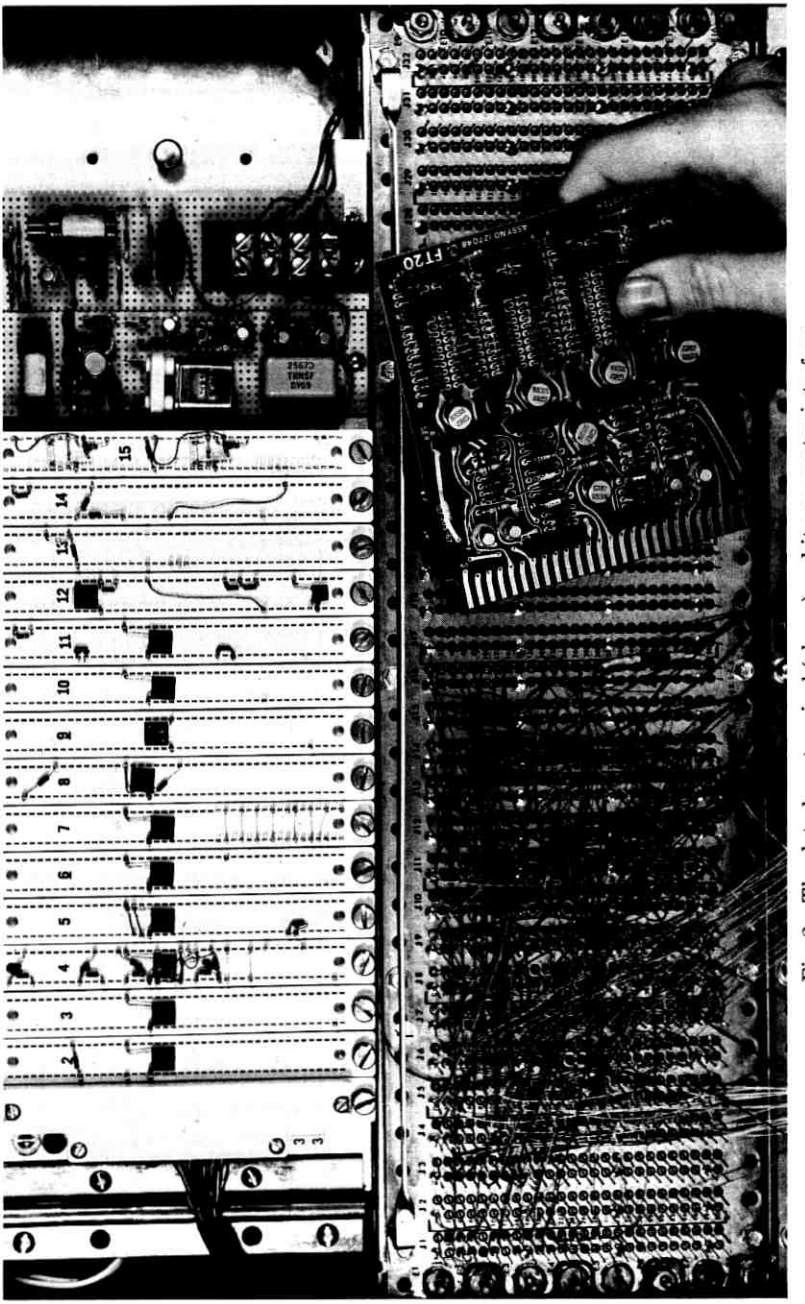


Fig. 2—The data loop terminal (above) and its computer interface.

IV. SOFTWARE

Programming for the ring system is essentially the same as for other means of transmission. Problems of error detection, correction, and retransmission; problems of format compatability between dissimilar computers; problems of distinguishing data from control information; etc., are in no way different for the ring. Users are free to transmit or receive as the network permits. There are no required self-disciplines, except perhaps to discourage nuisance calls or "junk mail."

4.1 *Program Synchronization*

A point that deserves mention is a result of high-speed transmission in general, not specifically of the ring system. Without programmed precautions, the speed of the T1 loop allows a slow receiver to be overrun by a fast sender, whereas with slower transmission, the channel itself might have been the limiting factor.

A trivial solution, appropriate for simple receiving terminals, is simple open-loop control of transmission rate. The sender, knowing the limitations of the receiver, waits for a fixed time after transmitting each block before sending the next. With more capable receivers, a basic "dialogue" procedure that is good practice for other transmission media also works well for the loop. In an initial "handshake," sender and receiver agree that a fixed amount of data will be sent, after which the sender will hold, awaiting a "go-ahead" message from the receiver. Normally, the sender will retain the transmitted data while waiting, in case the reply is a request for retransmission.

4.2 *Software Multiplexing of Addresses*

Hardware insertion and removal of addresses and rejection of intrusions is attractive for a computer as well as for a simple device–especially in an open-shop, real-time environment where computer users want direct, low-level control of I/O. Hardware addressing is not necessary, however, for program efficiency. Multiplexing can be done in software in several ways, depending on the particular interface. At the worst, it is no more complicated than copying data to or from a transmission buffer headed by address information. This processing is quite modest, compared to the translation, reformatting, packing, and unpacking frequently done for storage and input-output media.

4.3 *An Example Program*

A system utility program written for the loop provides a means to copy data or programs from the disk of one machine to that of the other,

and to perform several other functions. For transmission, the same program is used in both machines. Transmission can be controlled from either end.

The idling sequence of both machines is an attempt to input from the typewriter. The originator of a transaction gets a command from the keyboard and sends it through the network as a "command block," which interrupts the responding computer and takes it out of the idling sequence. The responder appends a coded acceptence or rejection and returns the message. The originator checks the reply and, if there is no error and the responder agrees on the amount, format, and disposal of data, then both computers enter the data sequence. The action proceeds by toggling between sending a block of data and sending back an acknowledgment.

In each computer, attempts to read control replies, data, and acknowledgments are subject to fixed time limits. Failure to receive the message in time is taken as an error. Errors of any kind are reported on the typewriter where the command was originated. Responsibility for requesting status, or restarting transmission, are left to the operator.

The program consists of 400 instructions in assembly language. Approximately 100 of these are ring I/O and error checking, 150 are typewriter I/O and command interpretation, and 150 are communications with the disk and printer, and the sequencing of subroutine calls to implement the commands.

4.4 *User Access to the Loop*

Commands of the loop utility program allow a user's program to be transported to a remote computer, loaded and placed into execution. Subsequently, a special interrupt-command block can cause the remote program to be aborted, dumped onto disk, and the general loop utility restored for continued remote operation. This allows both machines to be operated from one console, even in most program debugging.

There are no restrictions or special disciplines for use of the loop. It can be used in direct access by user programs, either in assembly language with a nine-instruction sequence, or in FORTRAN using existing library subroutines. Two calls define the address of a correspondent and transfer any amount of data up to 200 thousand bits, using data-loop blocks as they become available. For error checking, users may echo all transmissions, use simple checksums or use a burst-resistant multiple-error-detecting subroutine developed for magnetic tape.

Most of our uses of the loop involve the transmission of a program

or block of data 200 thousand bits or less in length. Transmissions are sufficiently infrequent to account for a very low average bit rate.

If necessary to reduce congestion, Kropfl's station can be made to impose a limit on peak transmission rate. All of our uses, thus far, are consistent with a peak rate limit of 50 thousand bits/second. Most transmissions would last only one or two seconds with that constraint.

Even a very demanding requirement–transmission of simple computer-generated motion pictures with our graphics system–can be done within

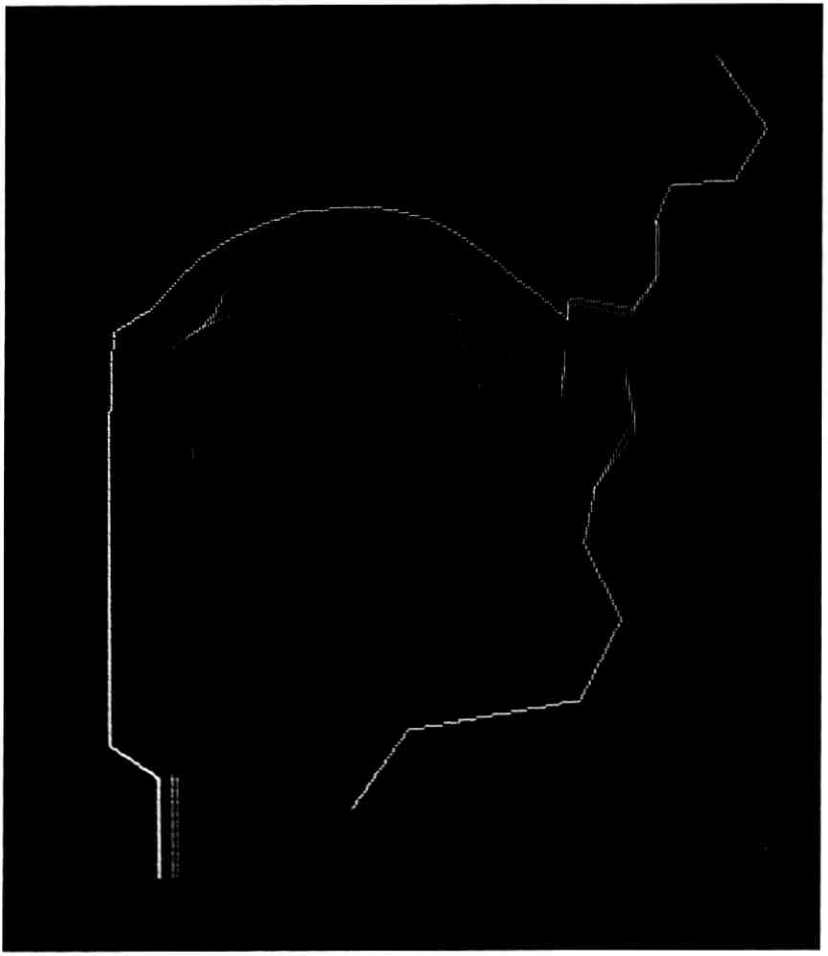


Fig. 3—The loop system supports rapid computer graphics. Using a dot-by-dot transmission code, real-time motion pictures of this complexity can be transmitted at 50 kilobits/second. Using a vector code, they would require only 10 kilobits.

this limit. The example in Fig. 3, drawn as a series of separately controlled dots, requires less than 3000 bits per frame for the part that moves. With a vector scope, it could be drawn with 600 bits per frame–less than 10 kilobits/second, 1/120 of the loop capacity.

V. OTHER APPLICATIONS

The above applications involve transmissions between essentially equal computers. Addressed-block transmission is potentially very useful between unequal correspondents also.

In remote-batch operation of a large computer, an interface from the remote terminal to the addressed-block system is the same as that for voice-grade and leased-line services. With conventional transmission the main computer has a separate modem and interface for each trunk. These are connected to a fairly large special processor whose job is to sort the simultaneously incoming streams of data into separate messages, and present them sequentially to the computer; and to perform an inverse function for output. These operations are inherent properties of addressed-block transmission! Messages are forced "into line" getting onto and passing through the network. They arrive at the computer and leave sequentially, through a single interface.

The availability of low-cost but powerful processors is making on-line computers desirable for every laboratory. But to be most useful, a machine should have access to a variety of expensive but infrequently used peripheral devices. A local data loop will allow a number of small and intermediate computers to share a pool of special equipment. Communication instead of duplication combines the economy and versitility of centralization with the on-line computing power of separate machines.

VI. SUMMARY

We have interfaced two laboratory computers to an experimental addressed-block data transmission system. The project went smoothly; there were no disappointments or surprises. Programming for the system is equally pleasant and uncomplicated. The general-purpose data transmission program for the system was written and debugged in a week. We are presently extending the system to other computers. We see addressed-block transmission as a simple but convenient solution to our computer communications needs.

REFERENCES

1. Pierce, J. R., "Network for Block Switching of Data," B.S.T.J., this issue, pp. 1133–1145.
2. Kropfl, W. J., "An Experimental Data Block Switching System," B.S.T.J., this issue, pp. 1147–1165.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

The Design of Finite Impulse Response Digital Filters Using Linear Programming Techniques

By L. R. RABINER

(Manuscript received October 18, 1971)

In this paper it is shown how standard linear programming techniques can be applied to designing finite impulse response digital filters. Attention is concentrated on designing filters having exactly linear phase, and arbitrary magnitude response. The design method is illustrated by examples of the design of frequency sampling filters with constraints on in-band ripple, optimal filters where the passband and stopband cutoff frequencies may be specified exactly, and filters with simultaneous constraints on the time response and frequency response.

I. INTRODUCTION

Many techniques exist for designing digital filters using optimization procedures. Herrmann and Schuessler have designed equiripple error approximations to finite impulse response (FIR) lowpass and bandpass filters using nonlinear programming procedures.[1,2] This work has been extended by Hofstetter, Oppenheim, and Siegel,[3] and by Parks and McClellan[4] to solve for the desired filters using polynomial interpolation techniques. Rabiner, Gold, and McGonegal[5] used a steepest descent technique to obtain FIR digital filters with minimax error in selected bands with the constraint that only a few of the filter coefficients were variable. Steiglitz,[6] and Athanasopoulos and Kaiser[7] have used nonlinear optimization techniques to obtain recursive filter approximations to arbitrary frequency response specifications.

Recently, attention has been focused on the use of linear programming techniques for the design of digital filters.[8-10] Many digital filter design problems are inherently linear in the design parameters, and hence are natural candidates for linear programming optimization. Further, linear programs are easy to implement and are generally guaranteed

to converge to a unique solution. The rate of convergence of the programs is moderately fast, thus making this technique practical for problems with 100 parameters or so.

There are many areas of FIR filter design where linear programming can be used conveniently. These include:

(*i*) design of filters with minimax ripple in the passband and/or stopband;

(*ii*) design of optimal (minimax) absolute or relative error approximations to arbitrary frequency response characteristics, where the passband and stopband edge frequencies of the filter may be specified exactly;

(*iii*) design of two-dimensional filters of the frequency sampling type, or with optimal error approximation;

(*iv*) and design of filters with simultaneous constraints on characteristics of both the time and frequency response of the filter.

Several of these design areas have been examined and examples will be presented showing how to apply linear programming techniques in specific cases. In the next section, the general framework of linear programming is presented and several practical aspects of linear programs are discussed. The following sections show how the general FIR, linear phase, filter design problem is linear in either the filter impulse response coefficients, or equivalently the Discrete Fourier Transform (DFT) coefficients, and how this problem is solved in specific cases.

II. LINEAR PROGRAMMING

The general linear programming problem can be mathematically stated in the form: find $\{X_j\}$, $j = 1, 2, \cdots, N$ subject to the constraints:

$$X_j \geqq 0, \qquad j = 1, 2, \cdots, N; \tag{1}$$

$$\sum_{j=1}^{N} c_{ij} X_j = b_i, \qquad i = 1, 2, \cdots, M (M < N); \tag{2}$$

such that:

$$\sum_{j=1}^{N} a_j X_j \quad \text{is minimized.} \tag{3}$$

The above problem is referred to as the "primal problem" and by a duality principle can be shown to be mathematically equivalent to

the "dual problem": find $\{Y_i\}$, $i = 1, 2, \cdots, M$ subject to the constraints:

$$\sum_{i=1}^{M} c_{ij} Y_i \leqq a_i , \qquad j = 1, 2, \cdots, N, \tag{4}$$

such that:

$$\sum_{i=1}^{M} b_i Y_i \quad \text{is maximized.} \tag{5}$$

The remainder of this paper refers to the dual problem as this is the most natural form for the digital filter design problems under consideration.

One characteristic of linear programs is that, given there is a solution, it is guaranteed to be a unique solution; and there are several well defined procedures for arriving at this solution within $(M + N)$ iterations. There are also straight-forward techniques for determining if the solution is unconstrained or poorly constrained.

The next section shows that linear phase FIR filters are linear in the design parameters and hence can be optimally designed using linear programming techniques.

III. LINEAR PHASE FIR FILTERS

Let $\{h_n\}$, $n = 0, 1, \cdots, N - 1$ be the impulse response of a causal FIR digital filter. The requirement of linear phase implies that

$$h_n = h_{N-n-1} . \tag{6}$$

The filter frequency response can be determined, in terms of the $\{h_n\}$, as:

$$H(e^{j\omega T}) = \sum_{n=0}^{N-1} h_n e^{-j\omega T n}. \tag{7}$$

For the case where N is odd, eq. (7) can be combined with eq. (6) to give:

$$H(e^{j\omega T}) = \underbrace{e^{-j\omega((N-1)/2)T}}_{\substack{\text{linear phase} \\ \text{term}}} \cdot \underbrace{\left[h_{(N-1)/2} + \sum_{n=0}^{(N-3)/2} 2h_n \cos\left[\left(\left(\frac{N-1}{2}\right) - n\right)\omega T\right]\right]}_{\substack{\text{purely real} \\ \text{linear in } \{h_n\}}}. \tag{8}$$

Equation (8) shows $H(e^{j\omega T})$ to consist of a purely linear phase term corresponding to a delay of $((N-1)/2)$ samples, and a term which is purely real and linear in the impulse response coefficients. It is the second term in eq. (8) which is used for approximating arbitrary magnitude response characteristics. Where N is even, the result of eq. (8) is modified to:

$$H(e^{j\omega T}) = \underbrace{e^{-j\omega((N-1)/2)T}}_{\substack{\text{linear phase} \\ \text{term}}} \underbrace{\left[\sum_{n=0}^{((N/2)-1)} 2h_n \cos\left(\frac{N-1}{2} - n\right)\omega T\right]}_{\substack{\text{purely real} \\ \text{linear in } \{h_n\}}}. \tag{9}$$

Equation (9) shows that for N even, the linear phase term corresponds to a delay of an (integer $+ \frac{1}{2}$) number of samples. The center of symmetry of $\{h_n\}$ is midway between samples $(N/2)$ and $(N/2 - 1)$. The remainder of eq. (9) is again a real term which is linear in the impulse response coefficients.

The DFT relation can be used to show that the filter frequency response is also a linear function of the DFT coefficients $\{H_k\}$. It is derived elsewhere[11] that the frequency response of linear phase FIR filters can be written:

$$H(e^{j\omega T}) = e^{-j\omega T((N-1)/2)} \frac{\sin\frac{\omega NT}{2}}{N} \cdot \left[\frac{H_0}{\sin\frac{\omega T}{2}} - \sum_{k=1}^{K} \frac{(-1)^k H_k 4 \cos\frac{\pi k}{N} \sin\frac{\omega T}{2}}{\left(\cos\omega T - \cos\frac{2\pi k}{N}\right)}\right], \tag{10}$$

when

$$K = \begin{cases} (N-1)/2 & \text{for } N \text{ odd} \\ \dfrac{N}{2} - 1 & \text{for } N \text{ even} \end{cases}.$$

The significance of eq. (10) is that the frequency response of a linear phase FIR filter is linear in the $\{H_k\}$ as well as in the $\{h_n\}$; hence linear programming techniques can be used to optimize the values of all or a selected set of DFT or impulse response coefficients.

IV. DESIGN OF FREQUENCY SAMPLING FILTERS

Previously, design of frequency sampling filters was accomplished using a steepest descent minimization.[5] This technique was capable only of minimizing the peak out-of-band ripple when several DFT coefficients in a transition band between passbands and stopbands were varied. Another limitation of the technique was that the amount of computation it took to optimally choose the variable DFT coefficients grew exponentially with the number of unconstrained variables. The largest problems attempted had four coefficients variable. This problem is readily solved in a much more general form using linear programming techniques. Furthermore, the computation required to calculate the more general solutions is considerably less than for the steepest descent algorithm used previously.

A typical specification for a lowpass filter to be approximated by a frequency sampling design is shown in Fig. 1. The heavy points show the DFT coefficients, and the solid curve shows the interpolated frequency response. The passband edge frequency is F_1 and the stopband edge frequency is F_2 . Since the length of the filter impulse response

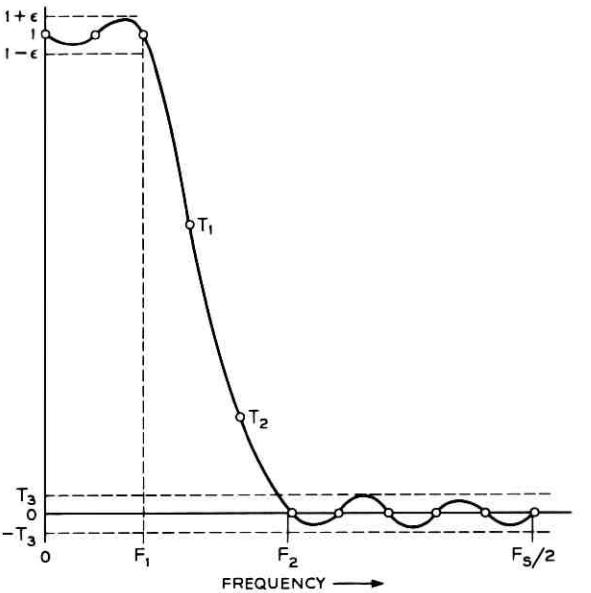


Fig. 1—Typical specification for a frequency sampling lowpass filter with transition coefficients T_1 , T_2 .

is N samples (assume N even), there are $(N/2 + 1)$ DFT coefficients (called frequency samples) to be specified. Those DFT coefficients which are in the passband are arbitrarily assigned the value 1.0, and those that fall in the stopband are assigned the value 0.0. The DFT coefficients in the transition band are free variables, and are labeled T_1 , T_2 in Fig. 1. The approximation problem can be set up as a linear program in the following manner. We let

$$T_3 = \text{peak stopband ripple.}$$

Then the design problem consists of finding values of (T_1 , T_2) to satisfy the constraints:

(*i*) The in-band ripple is less than or equal to some prescribed tolerance, ϵ.

(*ii*) The peak out-of-band ripple, T_3 is to be minimized.

Mathematically this problem can be stated as: find (T_1 , T_2 , T_3) subject to the constraints:

$$1 - \epsilon \leqq F(\omega) + \sum_{i=1}^{2} T_i D(\omega, i) \leqq 1 + \epsilon, \qquad 0 \leqq \omega \leqq 2\pi F_1 , \tag{11}$$

$$-T_3 \leqq F(\omega) + \sum_{i=1}^{2} T_i D(\omega, i) \leqq T_3 , \qquad 2\pi F_2 \leqq \omega \leqq \pi F_s , \tag{12}$$

where $F(\omega)$ is the contribution of the fixed DFT coefficients (the 1.0's in-band) and $D(\omega, i)$ is the contribution of the ith variable transition coefficient and is of the form shown in eq. (10), and F_s is the sampling frequency.

A suitable reshuffling of terms in eqs. (11) and (12) puts the set of equations in the form of the dual problem of linear programming. The final equations are of the form: find (T_1 , T_2 , T_3) subject to the constraints:

$$\left.\begin{aligned} \sum_{i=1}^{2} T_i D(\omega, i) &\leqq 1 + \epsilon - F(\omega) \\ -\sum_{i=1}^{2} T_i D(\omega, i) &\leqq -1 + \epsilon + F(\omega) \end{aligned}\right\} 0 \leqq \omega \leqq 2\pi F_1 , \tag{13}$$

$$\left.\begin{aligned} \sum_{i=1}^{2} T_i D(\omega, i) - T_3 &\leqq -F(\omega) \\ -\sum_{i=1}^{2} T_i D(\omega, i) - T_3 &\leqq F(\omega) \end{aligned}\right\} 2\pi F_2 \leqq \omega \leqq \pi F_s , \tag{14}$$

$(-T_3)$ is maximized.

The inequalities of eqs. (13) and (14) are evaluated at a dense set of frequencies in the appropriate range of interest (an 8-1 interpolation between DFT coefficients is sufficient) to yield the necessary set of equations for the linear program.

V. RESULTS ON FREQUENCY SAMPLING DESIGNS

A wide variety of frequency sampling filters has been designed using the results of eqs. (13) and (14). Previously, using the steepest descent algorithm, constraints on the in-band ripple, ϵ, could not be maintained.[5] With the linear programming design, tradeoff relations between in-band and out-of-band ripple can be obtained for a fixed number of transition samples, or equivalently a fixed width of transition band. Such tradeoff relations are illustrated in Figs. 2 and 3 for two and three transition samples.* In both these figures, the log

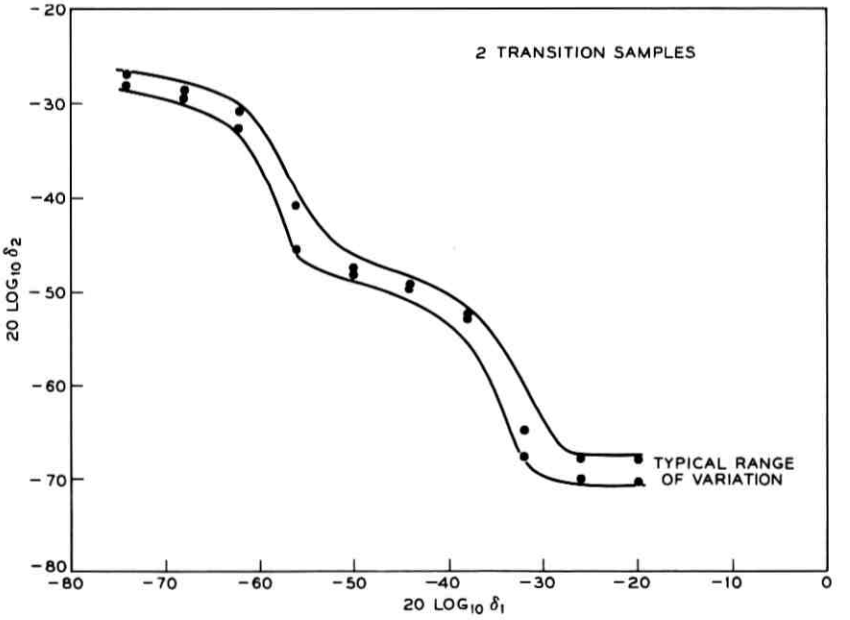


Fig. 2—Tradeoff relations between δ_1 and δ_2 for lowpass frequency sampling filters with two variable transition coefficients.

* The varying nature of the curves of Figs. 2 and 3 is due to the variance in the measured points (heavy dots) as a function of filter bandwidth. Solid curves are shown as an underbound and overbound on the typical behavior of the tradeoff relations.

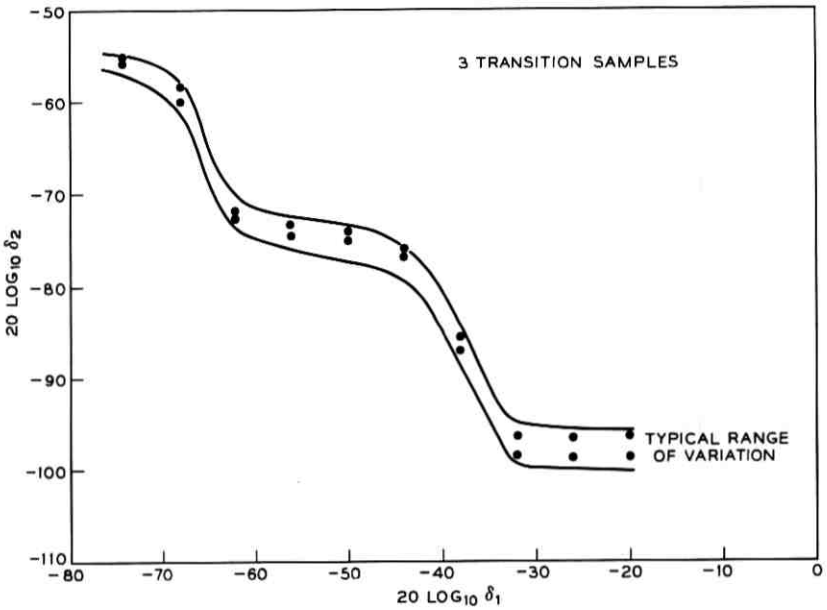


Fig. 3—Tradeoff relations between δ_1 and δ_2 for lowpass frequency sampling filters with three variable transition coefficients.

of out-of-band ripple, δ_2, versus the log of in-band ripple, δ_1, is plotted. Figure 2 shows that for in-band ripples larger than about 0.03 (i.e., 20 $\log_{10}$ δ_1 greater than -30 dB), the out-of-band ripple is in the range -66 to -71 dB. These figures correspond to the cases designed earlier[5] when no constraint on in-band ripple was in effect. At the other extreme of the curve, the out-of-band ripple flattens to between -25 and -30 dB with the in-band ripple, δ_1, in the range 0.0002 to 0.0005 (-74 to -66 dB). The midrange of the curve shows the tradeoff attainable between the two ripples. Figure 3 shows similar results for the case of three transition samples. No simple explanation is available for the general shape of these curves or the differences between the data in Figs. 2 and 3.

Figure 4 shows a comparison between equiripple filters and frequency sampling designs for the specialized case where in-band ripple and out-of-band ripple are equal. In this figure the normalized width of transition band* is plotted as a function of log δ, where δ is the ripple.

* The normalized width of transition band is defined as $D = N[(F_2 - F_1)/F_S]$ where N is the impulse response duration, F_S is the sampling frequency, and F_1 and F_2 are the passband and stopband cutoff frequencies in Hertz.

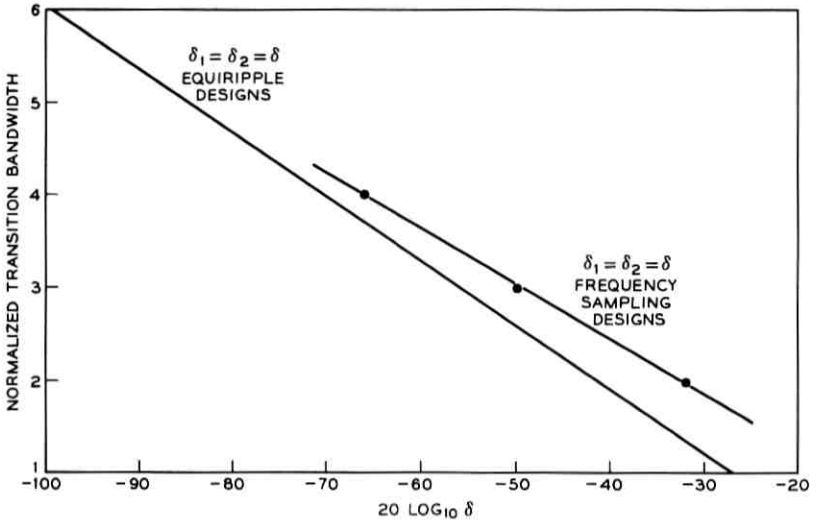


Fig. 4—A comparison between the curves of normalized transition bandwidth versus δ for equiripple filters and frequency sampling filters.

For the frequency sampling designs, the normalized transition widths are 4, 3, and 2 corresponding to 3, 2, and 1 transition samples. At these normalized transition bandwidths the ripple is −66, −50, and −32 dB respectively. The equiripple designs attain the same ripple values at normalized transition bandwidths of approximately 3.7, 2.6, and 1.4. The percentage difference in transition bandwidth for the 3 cases is 8.1, 15.4, and 42.9. Thus, except for the 1 transition point case, the transition bandwidths for frequency sampling designs are reasonably close to the bandwidths for equiripple filters.

VI. DESIGN OF OPTIMAL FILTERS

Just as a few of the DFT coefficients in a transition band could be varied to design reasonably efficient frequency sampling filters, all of the DFT coefficients, or equivalently all of the impulse response coefficients could be varied to give an optimal* approximation to any desired frequency response. Similar optimal approximations have been designed previously using nonlinear optimization procedures[1,2]

* The filters being discussed in this section are optimal in the sense of the theory of Chebyshev approximation on compact sets (i.e., the error of approximation exhibits at least $[(N+1)/2]+1$ alternations (of equal amplitude) over the frequency ranges of interest). In most cases, all the peaks of the error function are of the same amplitude, therefore, these filters are often referred to as equiripple filters.

and by polynomial interpolation methods.[3,4] However, the use of linear programming techniques, although significantly slower in running, offers many advantages over other existing design procedures. The design procedure is guaranteed to converge within a fixed number of iterations. Critical frequencies of the desired response can be specified exactly. The programs converge over a very wide range of parameter values. Finally, with the existence and increased understanding of integer linear programming techniques, the design problem can be combined with the coefficient quantization problem to design optimum filters with a prescribed wordlength.

To see how the design of optimal linear phase filters can be accomplished using linear programming techniques, consider the design of a lowpass filter to meet the following set of specifications:

Stopband magnitude ripple $\pm\delta_2$ Passband magnitude ripple $\pm\delta_1$	minimized or specified
Passband edge F_1	specified
Stopband edge F_2	specified

$F_1 < F_2$

(Phase response is to be linear.)

In this example either δ_1 , or δ_2 , or some linear combination is minimized. One can also consider the situation where δ_1 and δ_2 are proportionally related (i.e., $\delta_1 = k_1\delta$, $\delta_2 = k_2\delta$ where k_1 and k_2 are constants, and δ is minimized). In this manner a constant ratio between passband and stopband ripple is maintained. Consider the case where δ_1 is specified, and δ_2 is minimized. The linear program which realizes the above specifications can be stated as: find $\{h_n\}$, δ_2 subject to the constraints*:

$$\left.\begin{aligned} h_0 + 2\sum_{n=1}^{(N-1)/2} h_n \cos \omega nT &\leqq 1 + \delta_1 \\ -h_0 - 2\sum_{n=1}^{(N-1)/2} h_n \cos \omega nT &\leqq -1 + \delta_1 \end{aligned}\right\} 0 \leqq \omega \leqq 2\pi F_1 , \tag{15}$$

$$\left.\begin{aligned} h_0 + 2\sum_{n=1}^{(N-1)/2} h_n \cos \omega nT - \delta_2 &\leqq 0 \\ -h_0 - 2\sum_{n=1}^{(N-1)/2} h_n \cos \omega nT - \delta_2 &\leqq 0 \end{aligned}\right\} 2\pi F_2 \leqq \omega \leqq \pi F_s , \tag{16}$$

$(-\delta_2)$ maximized.

* From this point on, for convenience, we are assuming h_n is defined from $(-(N-1))/2 \leqq n \leqq (N-1)/2$, and is symmetric around $n = 0$. Since N is odd, eq. (8) can be simplified to the form: $H(e^{j\omega T}) = h_0 + \sum_{n=1}^{(N-1)/2} 2h_n \cos \omega nT$.

Before proceeding to typical designs, it is important to note some properties of linear programming problems, and show how they affect the optimal filter design problem. The solution to a linear programming problem of the type shown above with L variables, and M inequality constraints occurs when at least L of the M equations are solved with equality (instead of inequality); the remaining inequalities being met with inequality. For the optimal filter design problem this implies that there are at least L frequencies at which the ripple obtains a maximum. The practical implications of this result are best illustrated in Fig. 5 which shows the frequency response of an equiripple optimal filter with passband ripple δ_1 , stopband ripple δ_2 , passband edge frequency F_1 , and stopband edge frequency F_2 . The length of the filter impulse response is N samples. If

$$N_p = \text{number of ripples in the passband, and}$$
$$N_s = \text{number of ripples in the stopband,}$$

then

$$N_p + N_s \leqq \frac{(N+1)}{2} \quad (N \quad \text{odd}), \tag{17}$$

since an Nth degree polynomial (the z-transform of the filter impulse response) has at most $(N + 1)/2$ points of zero derivative in the frequency range from 0 to $F_s/2$ Hz. In addition to attaining a maximum

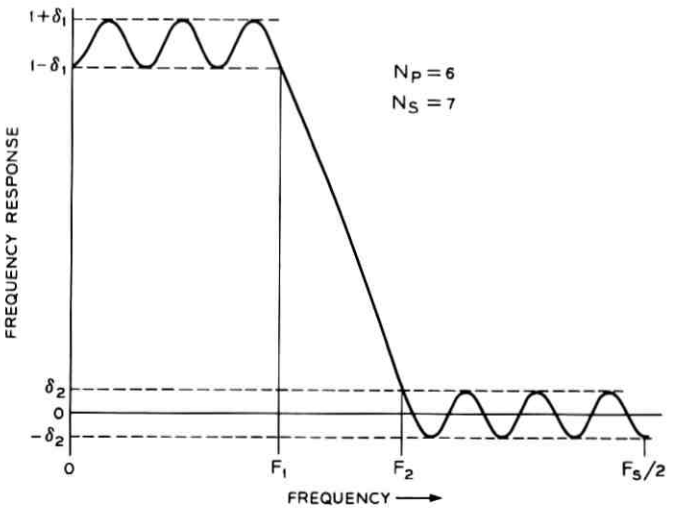


Fig. 5—A typical frequency response for an optimal filter, defining N_p , the number of passband maxima, and N_s the number of stopband maxima.

value at each of the ripple frequencies, the error attains a maximum value at $f = F_1$ and at $f = F_2$ (i.e., at the edges of the transition band). In fact this is how the transition band edges are defined. Thus the number of error maxima, N_e, satisfies the inequality

$$N_e \leqq \frac{(N+1)}{2} + 2. \tag{18}$$

The number of variables N_v in the linear program of eqs. (15) and (16) is

$$N_v = \frac{(N+1)}{2} + 1, \tag{19}$$

where $(N+1)/2$ coefficients of the impulse response are variable, and one ripple coefficient is variable. Thus eq. (19) shows that the minimum number of error maxima from the linear program solution, although optimal, is one less than the maximum number of error maxima obtainable.* A discussion of the effects of the extra ripple peak on the

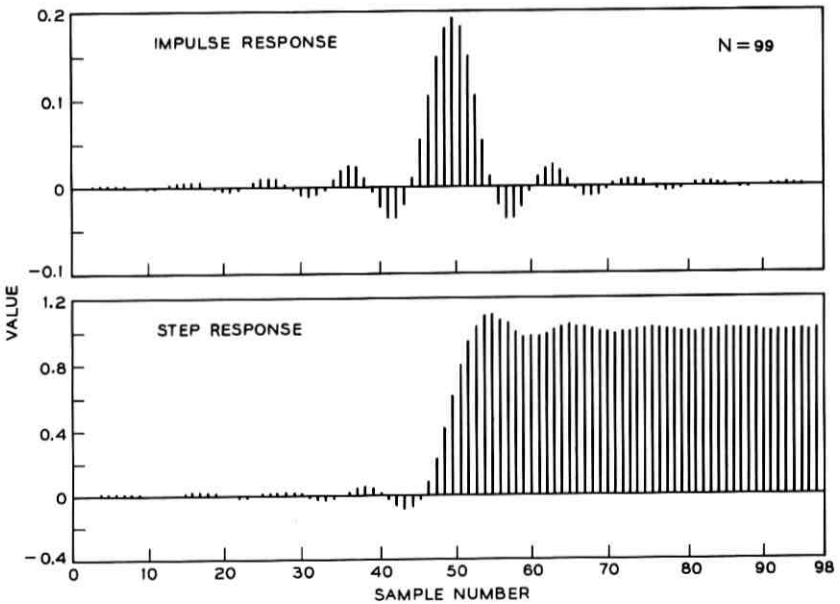


Fig. 6—The impulse and step response for an optimal digital filter with a 99-point impulse response.

* Parks and McClellan[4] have labeled the cases where all the ripples are present as "extra ripple" designs.

width of the transition band is given by Hofstetter, et al.[12] For all practical purposes the loss of the extra ripple is negligible in terms of normalized transition bandwidth, etc. At this point it is worthwhile showing some results of the design procedure.

VII. OPTIMAL FILTER DESIGNS–LOWPASS FILTER EXAMPLES

Using the linear program of eqs. (15) and (16), filters were designed with impulse response durations of up to 99 samples. Figures 6 and 7 show plots of impulse and step responses, and the log magnitude response of a lowpass filter designed from the specifications:

In-band ripple	δ
Out-of-band ripple	δ
Passband edge frequency	808 Hz
Stopband edge frequency	1111 Hz
Sampling frequency	10000 Hz.

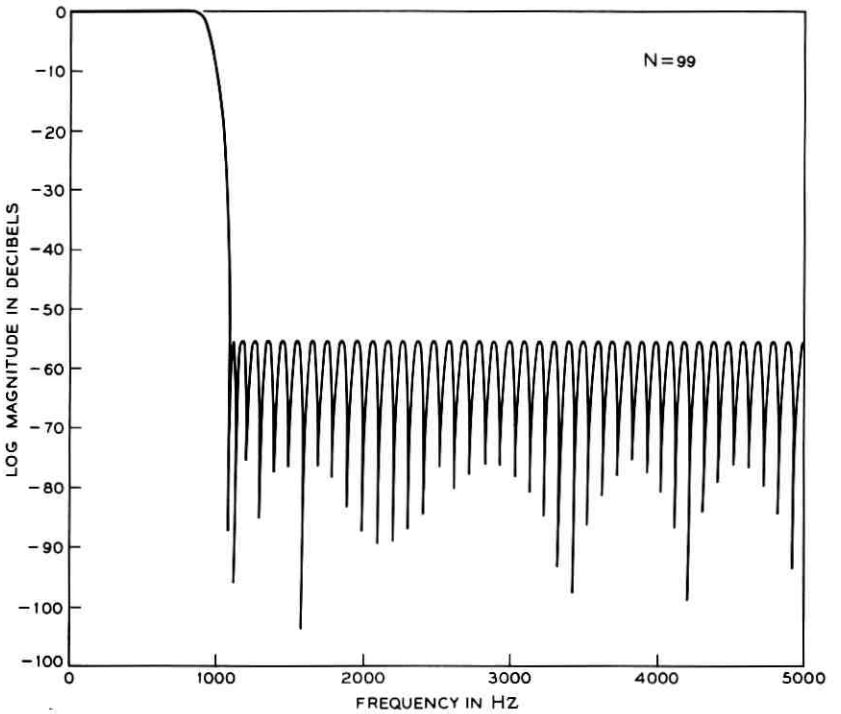


Fig. 7—The frequency response of an optimal digital filter with a 99-point impulse response.

The minimum value of δ, as chosen by the linear program, was $\delta = 0.001724$ or -55.3 dB.

Figure 8 shows a comparison of the normalized transition bandwidth versus log δ_2 for Herrmann–Schuessler equiripple filters with the maximum number of ripples, and the optimal linear program filters. The solid line in this figure shows the Herrmann–Schuessler data for $\delta_1 = \delta_2$, and the data points show the linear program data for several values of N, the impulse response duration. Clearly the differences between the data are insignificant as stated earlier. (The data points which fall below the solid line in Fig. 8 are due to the error in representing the equiripple data by a straight line on these coordinates.)

VIII. OPTIMAL FILTER DESIGNS–OTHER EXAMPLES

As stated earlier, the linear programming technique can design optimal approximations to any desired frequency response. To illustrate this feature several full band differentiators[13] and several filters for use in a digitized version of the A-channel bank[14] (a frequency transmission system in use in the Bell System) were designed.

To design a full band differentiator $H(e^{j\omega T})$ must approximate the normalized response,

$$\hat{H}(e^{j\omega T}) = j\frac{\omega}{\left(\frac{\omega_s}{2}\right)}, \tag{20}$$

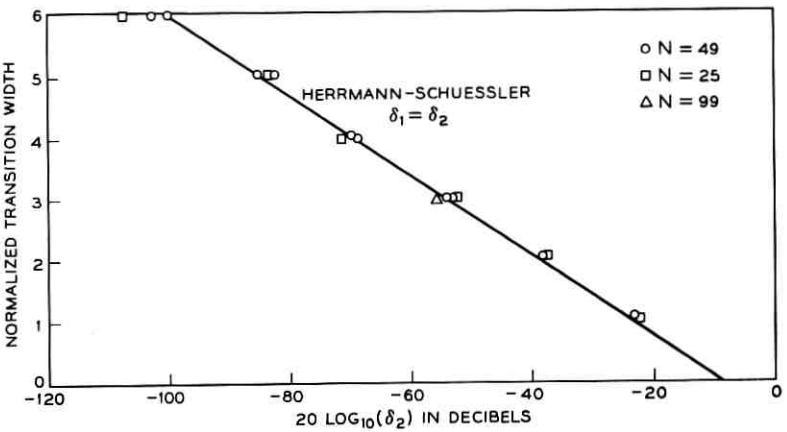


Fig. 8—A comparison between the curves of normalized transition bandwidth versus δ for equiripple filters with the maximum number of ripples, and optimal filters with one ripple omitted. Normalized bandwidth is defined as $D = N(F_2 - F_1)/F_s$.

where $(\omega_s/2)$ is half the radian sampling frequency. To get an optimal error approximation requires,

$$-\delta \leqq | H(e^{j\omega T}) - \hat{H}(e^{j\omega T}) | \leqq \delta, \tag{21}$$

where δ is minimized. To get a purely imaginary frequency response as in eq. (20), the impulse response is required to satisfy the symmetry condition,

$$h_n = -h_{N-n-1}, \qquad n = 0, 1, \cdots, \frac{N}{2} - 1, \tag{22}$$

where N is even to take advantage of the noninteger delay.[13] An illustrative example of an $N = 32$ sample differentiator is shown in Fig. 9. This figure shows the impulse response, magnitude response, and the error curve. The peak error, δ, is approximately 0.0057.

One could also consider designing optimal relative error filters by

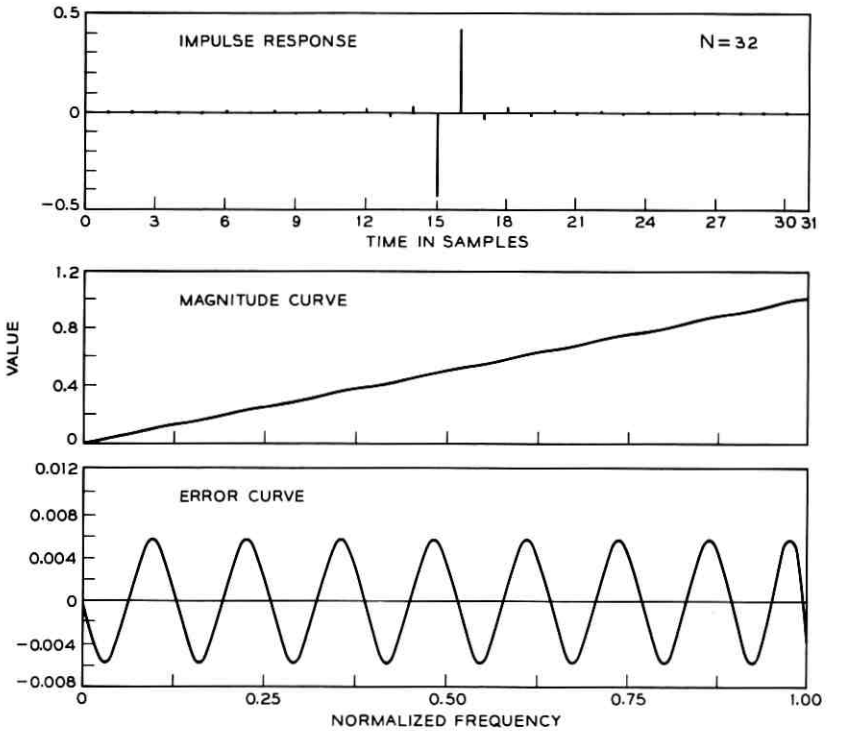


Fig. 9—The impulse response, frequency response, and error curve for a 32-point differentiator with optimal equiripple error.

changing the design equations slightly. For example, to design an optimal relative error differentiator requires,

$$-\delta\omega \leqq | H(e^{i\omega T}) - \hat{H}(e^{i\omega T}) | \leqq \delta\omega \tag{23}$$

(i.e., the envelope of the error in approximation is linear with frequency because the desired frequency response is linear in frequency). An example of an $N = 32$-point differentiator designed in this manner is shown in Fig. 10. The peak error, δ, is now 0.0062, only slightly higher

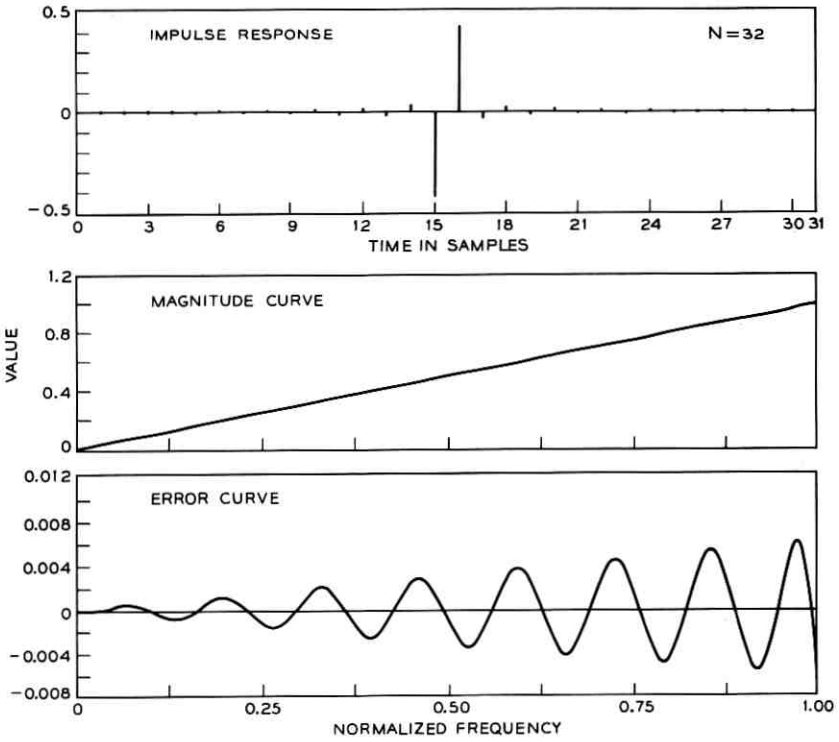


Fig. 10—The impulse response, frequency response, and error curve for a 32-point differentiator with optimal equiripple relative error.

than δ in the optimal error solution. The linearity of the error envelope is evident in Fig. 10.

To illustrate further the versatility of the linear programming approach, a special purpose filter for use in a digital transmission system was designed.[14] The specifications of the filter were:

$F_s = 112$ kHz

Passband ripple,	$0 \leqq f \leqq 2$ kHz,	$20 \log_{10} (1 + \delta_1) \leqq 0.25$ dB
Stopband ripple,	$14 \leqq f \leqq 18$ kHz,	$20 \log_{10} \delta_2 \leqq -63$ dB
	$30 \leqq f \leqq 34$ kHz,	$20 \log_{10} \delta_2 \leqq -63$ dB
	$46 \leqq f \leqq 50$ kHz,	$20 \log_{10} \delta_2 \leqq -63$ dB

In all other frequency bands, the frequency response was not specified. As an additional constraint on the impulse response, N was chosen arbitrarily to be 21 samples.

Since the filter was completely constrained by the above specifications, it was of interest to see how close the designs could get to the desired specifications. A linear program was written which allowed δ_1 and δ_2 to vary. The results of this program are plotted in Fig. 11. This figure shows a plot of $20 \log_{10} (\delta_2)$ versus $20 \log_{10} [(1 + \delta_1)/(1 - \delta_1)]$ obtained from the program. It also shows a triangle for the desired specifications, and a square for the filter obtained from a manual optimization by S. Freeny. Although none of the filters meets the specifications, the computer optimized designs come much closer than the manual optimization. Figure 12 shows a plot of the filter that comes closest to

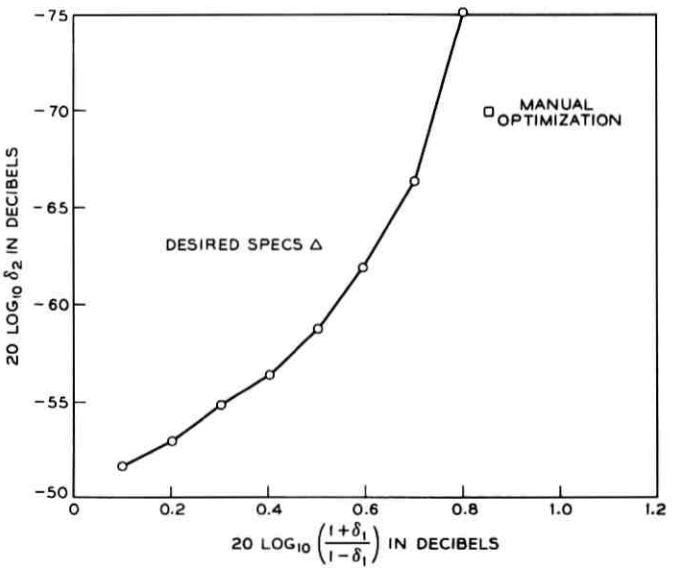


Fig. 11—A plot of the tradeoff relations between $20 \log_{10} ((1 + \delta_1)/(1 - \delta_1))$ and $20 \log_{10} \delta_2$ for a lowpass filter for a digital transmission system.

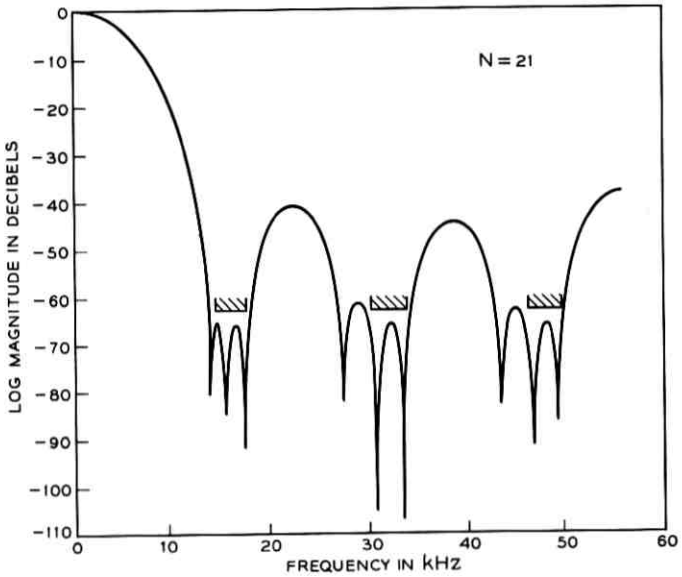


Fig. 12—The frequency response for the closest approximation to the desired specifications for the lowpass filter for the digital transmission system.

the desired specifications. The in-band response differs from specifications by about 0.1 dB, and the out-of-band response meets specifications by over 1 dB. The error is equiripple in each of the out-of-band regions.

IX. DESIGN OF FILTERS WITH SIMULTANEOUS CONSTRAINTS ON THE TIME AND FREQUENCY RESPONSE

The design of digital filters which approximate characteristics of a specified frequency response only has been discussed. Quite often one would like to impose simultaneous restrictions on both the time and frequency response of the filter. For example, in the design of lowpass filters, one would often like to limit the step response overshoot or ripple; at the same time maintaining some reasonable control over the frequency response of the filter. Since the step response is a linear function of the impulse response coefficients, a linear program is capable of setting up constraints of the type discussed above. Consider the design of a lowpass filter (N odd) with specifications:

Passband

$$1 - \delta_1 \leqq h_0 + \sum_{n=1}^{(N-1)/2} 2h_n \cos \omega nT \leqq 1 + \delta_1 , \tag{24}$$

Stopband

$$-\delta_2 \leqq h_0 + \sum_{n=1}^{(N-1)/2} 2h_n \cos \omega nT \leqq \delta_2 , \tag{25}$$

Step Response

$$-\delta_3 \leqq g_n \leqq \delta_3 \qquad n = -\left(\frac{N-1}{2}\right), \cdots, -\left(\frac{N-1}{2}\right) + N_1 , \tag{26}$$

where h_n is the symmetric impulse response of the filter ($h_n = h_{-n}$, $n = 0, 1, \cdots, (N-1)/2$), g_n is the filter step response defined by

$$g_n = \begin{cases} \sum_{m=-(N-1)/2}^{n} h_m & -\left(\frac{N-1}{2}\right) \leqq n \leqq \infty \\ 0 & n < -\left(\frac{N-1}{2}\right) \end{cases} , \tag{27}$$

and N_1 is the number of samples of the step response being constrained. For optimization there are several alternatives which are possible. One could fix any one or two of the parameters δ_1, δ_2, or δ_3 and minimize the other(s). Alternatively one could set $\delta_1 = \alpha_1\delta$, $\delta_2 = \alpha_2\delta$, and $\delta_3 = \alpha_3\delta$ where α_1, α_2, and α_3 are constants, and simultaneously minimize all three deltas.

To demonstrate this technique, a lowpass filter with $N = 25$ and no constraint on δ_3 was designed. This design is an optimal filter as discussed earlier, and is shown in Fig. 13. In this case δ_1 is set to $25\delta_2$ and the optimization gives $\delta_3 = 0.12$, $\delta_1 = 0.06$, and $\delta_2 = 0.00237$. The results of setting $\delta_3 = 0.03$ and then minimizing the frequency ripple are shown in Fig. 14. The equiripple character of the frequency response has been sacrificed in order to constrain the peak step response ripple. The ripple values for this new design are $\delta_1 = 0.145$ and $\delta_2 = 0.00582$. Using this linear programming technique, one can obtain tradeoffs between any of the deltas to get a design best suited to the particular application. The filter of Fig. 14 was designed for smoothing characteristic speech parameters where step response overshoot is a very important perceptual parameter.

X. DESIGN OF TWO-DIMENSIONAL FIR FILTERS

The techniques of FIR filter design using linear programming are readily extendable to two or more dimensions.[15] Filters of both the frequency sampling type and optimal type have been designed in this manner.

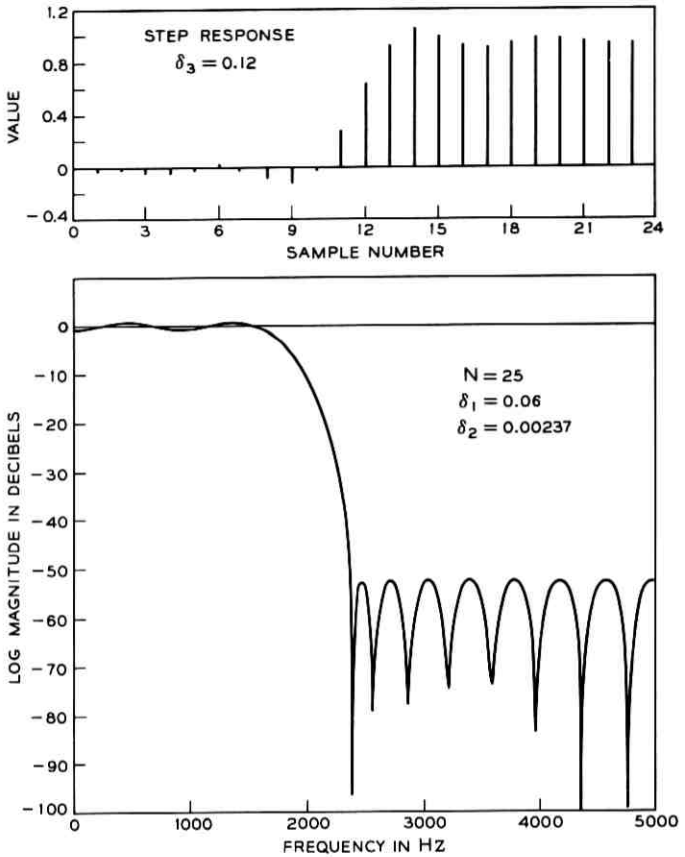


Fig. 13—A plot of the step response and frequency response of an optimal equiripple lowpass filter.

XI. COMPUTATIONAL CONSIDERATIONS

Since one of the major aspects of digital filter design by optimization procedures is the amount of computation necessary to produce a desired result, it is worthwhile discussing some of the details of our simultations.

The programs used throughout this study are APMM,[16] an IBM scientific subroutine which computes a Chebyshev approximation of a given real function over a discrete range, and MINLIN, a program written at Bell Laboratories by Mrs. Wanda Mammel. The running time of these programs is highly dependent on the number of variables, L, the number of inequalities, P, and the "complexity" of the results which determines the number of iterations required to attain a solution.

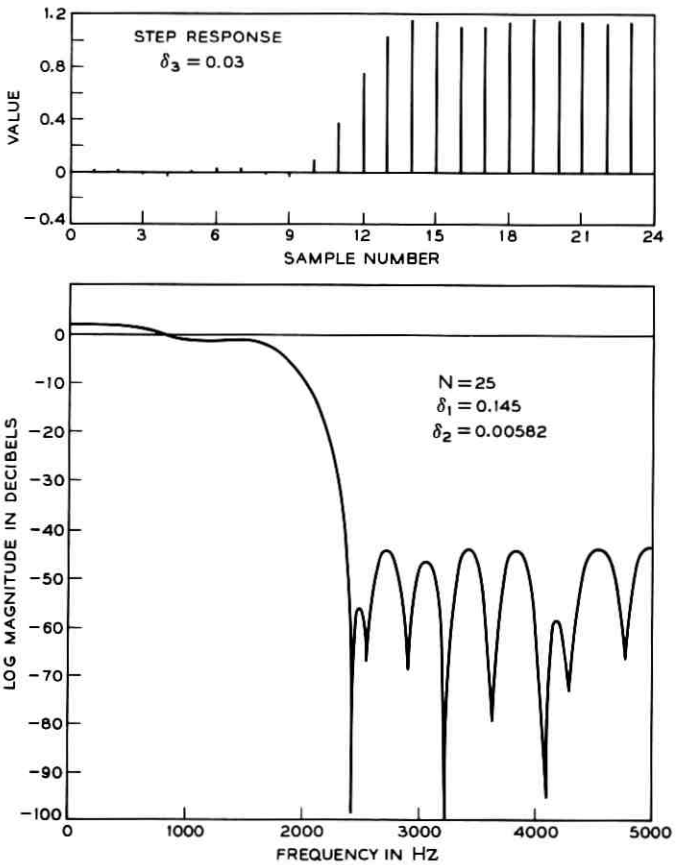


Fig. 14—A plot of the step response and frequency response for a lowpass filter with simultaneous constraints on both the time and frequency response.

The time per iteration is approximately proportional to L^2P. Typical experience indicates that it takes on the order of ten seconds to design the frequency sampling filters discussed earlier (i.e., $L \leqq 3$, P on the order of 1000). The total range of times to design optimal filters using APMM on the Honeywell 645 computer is shown below.

N	No. Iterations	Total Time
25	27 to 58	14 to 47 seconds
49	53 to 82	117 to 194 seconds
99	128	1200 seconds

Although the computation time is reasonably high, it is not impractical

to design high order filters with this technique. The argument can also be made that the most important application of these techniques is in the designs of FIR filters with small values of N (i.e., $N \leqq 50$) in which case the computation time starts becoming more reasonable.

XII. CONCLUSIONS

The design of linear phase FIR digital filters is shown to be a linear programming problem, and many appropriate problems can be solved using this technique. Examples have illustrated several filter areas which are reasonable candidates for linear program designs.

REFERENCES

1. Herrmann, O., and Schuessler, H. W., "On the Design of Selective Nonrecursive Digital Filters," IEEE Arden House Workshop, Harriman, N. Y., Jan. 1970.
2. Herrmann, O., "On the Design of Nonrecursive Digital Filters with Linear Phase," Electronics Letters, *6*, No. 11, 1970, pp. 328–329.
3. Hofstetter, E., Oppenheim, A., and Siegel, J., "A New Technique for the Design of Nonrecursive Digital Filters," Proceedings of the Fifth Annual Princeton Conference on Information Sciences and Systems, 1971, pp. 64–72.
4. Parks, T. W., and McClellan, J. H., "Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase," to be published in IEEE Trans. on Circuit Theory.
5. Rabiner, L. R., Gold, B., and McGonegal, C. A., "An Approach to the Approximation Problem for Nonrecursive Digital Filters," IEEE Trans. on Audio and Electroacoustics, *AU-18*, (June 1970), pp. 83–106.
6. Steiglitz, K., "Computer-aided Design of Recursive Digital Filters," IEEE Trans. on Audio and Electroacoustics, *AU-18*, (June 1970), pp. 123–129.
7. Athanasopoulos, J., and Kaiser, J. F., "Constrained Optimization Problems in Digital Filter Design," IEEE Arden House Workshop, Harriman, N. Y., Jan. 1970.
8. Helms, H., "Digital Filters with Equiripple or Minimax Responses," IEEE Trans. on Audio and Electroacoustics, *AU-19*, (March 1971), pp. 87–93.
9. Tufts, D. W., Rorabacher, D. W., and Mosier, W. E., "Designing Simple, Effective Digital Filters," IEEE Trans. on Audio and Electroacoustics, *AU-18*, (June 1970), pp. 142–158.
10. Cavin, R. K., Ray, C. H., and Rhyne, V. T., "The Design of Optimal Convolution Filters via Nonlinear Programming," IEEE Trans. on Geoscience Electronics, *GE-7*, (July 1969), pp. 142–145.
11. Rabiner, L. R., and Schafer, R. W., "Recursive and Nonrecursive Realizations of Digital Filters Designed by Frequency Sampling Techniques," IEEE Trans. on Audio and Electroacoustics, *AU-19*, (Sept. 1971), pp. 200–207.
12. Hofstetter, E., and Oppenheim, A., "On Optimum Nonrecursive Digital Filters," presented at the 9th Allerton Conference on Circuit and System Theory, Oct. 1971.
13. Rabiner, L. R., and Steiglitz, K., "The Design of Wideband Recursive and Nonrecursive Digital Differentiators," IEEE Trans. on Audio and Electroacoustics, *AU-18*, (June 1970), pp. 204–209.
14. Freeny, S. L., Kieburtz, R. B., Mina, K. V., and Tewksbury, S. K., "Design of Digital Filters for an All Digital Frequency Division Multiplex–Time Division Multiplex Translator," IEEE Trans. on Circuit Theory, *CT-18*, No. 6 (November 1971), pp. 702–711.
15. Hu, J. V. "Frequency Domain Design of Two-dimensional Finite Impulse Response Digital Filters," Proceedings of the Two-Dimensional Digital Signal Processing Conference, Univ. of Missouri, Columbia, Oct. 1971.
16. *IBM System/360 Programming Manual*, Scientific Subroutine Package (360 A-CM-03 X), Fortran program APMM, pp. 283–288.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

Pulse Propagation in Multimode Dielectric Waveguides

By D. MARCUSE

(Manuscript received November 4, 1971)

Using coupled power equations to describe the average performance of a multimode waveguide with random coupling, it is shown that a Gaussian input pulse remains approximately Gaussian with a pulse width that increases proportionally to the square root of the length of the waveguide. The proportionality factor is determined for the model of a slab waveguide. Since coupling between guided modes of necessity causes coupling of some of the guided modes to radiation modes, radiation losses are unavoidable. A desired improvement in pulse distortion that is accomplished by coupling the guided modes intentionally to each other must be paid for by a certain loss penalty. This loss penalty is also evaluated for the special case of the slab waveguide model. Pulse dispersion improvement can be achieved by providing intentional roughness of the core-cladding interface of the dielectric waveguide. The "power spectrum" of the core-cladding interface function must be designed very carefully in order to minimize the radiation loss penalty that accompanies any attempt to reduce pulse dispersion. The dependence of the loss penalty on the shape of the "power spectrum" of the core-cladding interface function is studied in this paper. Design criteria for the improvement of multimode pulse dispersion are given based on the slab waveguide model. The connection between the slab waveguide model and the round optical fiber is pointed out.

I. INTRODUCTION

S. D. Personick[1] was the first to realize that coupling between the guided modes of a multimode waveguide is capable of reducing the pulse dispersion that is caused by the fact that modes with higher group velocity arrive at the receiver earlier than modes with lower group velocity. Multimode pulse dispersion can, of course, be avoided by designing the waveguide to operate with only a single mode. However, single-mode waveguides cannot be excited efficiently by incoherent

light sources such as luminescent diodes. A simple communications system using luminescent diodes instead of more expensive lasers as light sources needs multimode optical waveguides as the transmission medium. Unless multimode pulse dispersion can be reduced by some means, the information-carrying capacity of a multimode optical fiber is severely limited. Even though Personick[1] pointed the way for achieving an improvement in the multimode pulse dispersion he did not give design criteria for their construction nor did he discuss the loss penalty that inevitably must be paid for any improvement in pulse dispersion. Furthermore, Personick's paper deals primarily with two modes even though some thought is given to the multimode case. H. E. Rowe and D. T. Young[2] rederived Personick's results using a more rigorous analysis but also limited themselves to the two-mode case. Patent applications by E. A. J. Marcatili, S. E. Miller, and S. D. Personick are pending. These patents describe the geometry of a fiber designed to improve multimode pulse distortion by means of mode coupling.

The theory presented in this paper is applicable to an arbitrary number of modes. Utilizing coupled equations (derived in an earlier paper[3]) for the average power carried by the modes of the guide and extending the discussion of the steady state multimode waveguide to the time varying case, a complete description of pulse propagation in multimode waveguides is formally set forth. This complete theory can be evaluated only approximately by means of perturbation theory. Using a second-order perturbation approach a solution of the pulse problem is presented with the assumption that the input pulse has a Gaussian shape (in time). Numerical evaluation of the theory requires matrix diagonalization that can be accomplished on a high-speed electronic computer. The theory is applied to the dielectric slab waveguide; and design criteria for this case are obtained. However, we also extrapolate the slab waveguide results to the more interesting case of the round optical fiber.

The coupling coefficients used in this paper are derived from a first-order perturbation theory. Therefore, they hold only for weak coupling. In case of strong coupling, the actual radiation losses are expected to be larger than predicted here.

II. DISCUSSION OF THE PRINCIPLE OF PULSE DISTORTION REDUCTION

S. D. Personick[1] discovered that coupling between the guided modes of a multimode waveguide with a random coupling function reduces the spread of a pulse whose power is shared by a large number of modes

traveling with slightly different group velocities. Dielectric waveguides have two types of modes. The guided modes that are capable of transporting power through the waveguide and radiation modes that allow the description of the radiation field around the waveguide. Imperfections in the waveguide that couple the guided modes among each other also tend to couple the guided modes to the radiation field. The coupling of guided modes to each other and to radiation modes is well understood.[4,5] In particular it is known that–to first order–two guided modes couple only by means of one component of the Fourier spectrum of the coupling function.[4,6] Figure 1 shows a schematic plot of the possible propagation constants β of the modes of a dielectric slab waveguide. Also shown is a bracket connecting two guided modes. The separation of these modes is

$$\Delta\beta = \beta_\nu - \beta_\mu . \tag{1}$$

The coupling function can be represented as the product of a constant term times a function of z, the distance along the waveguide axis.[7]

$$c_{\nu\mu}(z) = K_{\nu\mu} f(z). \tag{2}$$

The function f can be expanded in a Fourier series[4]

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{i\phi_n x} \tag{3}$$

with

$$\phi_n = \frac{2\pi}{z} n. \tag{4}$$

The two guided modes ν and μ are coupled only by the Fourier com-

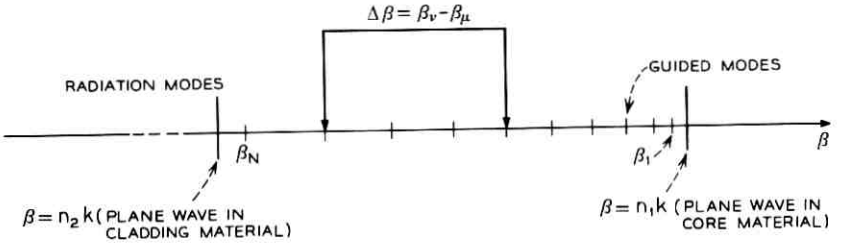


Fig. 1—Schematic representation of the propagation constants β_ν of the guided modes. $n_1 k$ is the propagation constant of plane waves in the core material, $n_2 k$ is the plane wave propagation constant in the cladding material. The line labeled $\Delta\beta$ indicates two guided modes that are coupled by a sinusoidal core-cladding interface irregularity of mechanical frequency $\Delta\beta$.

ponent whose mechanical frequency is given by

$$\phi_n = \beta_\nu - \beta_\mu . \tag{5}$$

If no Fourier component at this mechanical frequency exists the modes remain uncoupled. In order to couple all the guided modes shown in Fig. 1 to each other, we need a Fourier spectrum that has components at all those mechanical frequencies that correspond to existing differences $\beta_\nu - \beta_\mu$. However, in addition to coupling the guided modes among each other, the coupling mechanism also couples guided modes to radiation modes. The coupling law remains the same. Any Fourier component that has a mechanical frequency in the range[6]

$$\beta_\nu - n_2 k \leqq \phi_n \leqq \beta_\nu + n_2 k \tag{6}$$

($n_2 k$ is the propagation constant of plane waves in the medium of the cladding material) couples the mode ν to the radiation field. Coupling between guided modes and radiation modes results in radiation loss. It is thus apparent that we must avoid coupling guided modes to radiation modes or at least try to couple as few of the guided modes as possible to the radiation field without destroying the coupling between the guided modes. It is apparent from Fig. 1 that it is possible to couple all the guided modes among each other and couple only the highest-order guided mode to the radiation field. This selective coupling is made possible by the fact that the spacing between guided modes in β space decreases with decreasing mode number. Ideally we would want a "power spectrum" $F(\phi)$ of the function $f(z)$ as shown in Fig. 2. This spectrum is flat from zero mechanical frequencies to the maximum frequency $\phi_{max} = \beta_{N-2} - \beta_{N-1}$ that is chosen to be equal to the separation between the modes $N - 2$ and $N - 1$. The last guided mode, N, is close to the radiation field so that the Fourier spectrum of the func-

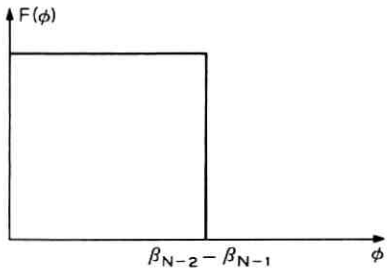


Fig. 2—Ideal shape of the "power spectrum" of the core-cladding interface irregularities.

tion $f(z)$ shown in Fig. 2 couples this mode strongly to the radiation field without coupling it to mode $N - 1$ or any other of the guided modes. The high-order modes are coupled only to their next neighbor while more than two low-order guided modes couple directly to each other because of their close spacing. It would be more efficient for the purposes of pulse distortion reduction to couple each guided mode individually to all the other guided modes. However, since it is impossible to accomplish this without simultaneously coupling all the guided modes directly to radiation modes we must be content to try to couple each guided mode at least to its nearest neighbor. The mechanical "power spectrum" of Fig. 2 would accomplish pulse distortion reduction by means of coupling between the guided modes without any radiation loss penalty. The reduction in the pulse length compared to the uncoupled case comes about because some of the power traveling in a fast mode is eventually transferred to a slow mode while power starting out in a slow mode finds itself at least partly in a fast mode so that the extremes of the group velocity spread are partly equalized causing the center of gravity of the pulse distribution to travel at an average velocity.

The perfect pulse distortion reduction scheme just outlined cannot be realized in practice since it is impossible to build filters with infinitely steep slopes. We can imagine that it is possible to produce a mechanical spectrum of core-cladding interface irregularities by changing the pulling speed of the fiber as it is drawn from the melt or from a preform. If the speed modulation is derived from an electrical noise signal that is filtered by a low-pass filter the problem is reduced to designing an electrical filter with as steep a slope as possible. More details of the required slope will be discussed later when we study the results of the numerical analysis of pulse propagation in slab waveguides.

III. THEORY OF PULSE PROPAGATION IN MULTIMODE WAVEGUIDES

It was shown in an earlier paper[3] how, starting from coupled wave equations, it is possible to obtain stochastic equations for the average power P_ν carried by N modes. In the steady state case discussed in Refs. 3 and 7 the coupled power equations assume the form

$$\frac{dP_\nu}{dz} = -(\alpha_\nu + b_\nu)P_\nu + \sum_{\mu=1}^{N} h_{\nu\mu}P_\mu \tag{7}$$

with

$$b_\nu = \sum_{\mu=1}^{N} h_{\nu\mu} \, . \tag{8}$$

The form of the symmetric matrix elements (power coupling coefficients) $h_{\nu\mu}$ depends on the type of waveguide and the particular coupling mechanism that is being considered, α_ν is the power loss coefficient that results from the coupling of mode ν directly to the radiation field.

Our first task is to generalize the steady state equation (7) to the time-dependent case. This generalization is achieved by considering a moving observer traveling at the group velocity v_ν of one of the guided modes. Whereas a stationary observer sees the average power in this mode grow or diminish as a function of z without noticing any change in time (in the steady state), the moving observer sees the mode power grow in time. The derivative dP_ν/dz noticed by the stationary observer corresponds to the derivative $(1/v_\nu)dP_\nu/dt$ observed in a coordinate system traveling at velocity v_ν . We can thus write for the steady-state case

$$\frac{dP_\nu}{dz} = \frac{1}{v_\nu}\frac{dP_\nu}{dt}. \tag{9}$$

The extension to the time-varying case consists in using the right-hand side of (9) even if the stationary observer sees the mode field change in time. This extension is certainly plausible if the time variations are not too rapid. Keeping in mind that the total time derivative corresponds to the change seen by the moving observer, we introduce the space- and time-dependent functions $P_\nu(z, t)$ and write

$$\frac{dP_\nu}{dt} = \frac{\partial P_\nu}{\partial t} + \frac{\partial P_\nu}{\partial z}\frac{dz}{dt} = v_\nu\frac{\partial P_\nu}{\partial z} + \frac{\partial P_\nu}{\partial t}. \tag{10}$$

The partial derivatives on the right-hand side of (10) are again the changes that are seen by a stationary observer. The time-dependent coupled power equations can thus be written as

$$\frac{\partial P_\nu}{\partial z} + \frac{1}{v_\nu}\frac{\partial P_\nu}{\partial t} = -(\alpha_\nu + b_\nu)P_\nu + \sum_{\mu=1}^{N} h_{\nu\mu}P_\mu . \tag{11}$$

Equation (11) is the starting point for the study of pulse propagation in multimode dielectric waveguides.

We obtain a formal solution of the time-dependent problem by substitution of the trial solution

$$P_\nu(z, t) = B_\nu e^{-\alpha z + i\omega t}. \tag{12}$$

The parameter ω would usually be considered to be the frequency of the time-dependent process (12). However, the average power P_ν is not sinusoidally time-varying so that it cannot be associated with a frequency. The parameter ω is thus simply a variable of integration for a Fourier integral expansion of the function $P_\nu(z, t)$. Substitution of (12)

into (11) leads to the following algebraic eigenvalue problem for the eigenvalue α and the eigenvector B_ν (B_ν is the νth element of an N dimensional vector):

$$\left[-\alpha + i\frac{\omega}{v_o} + i\omega\left(\frac{1}{v_\nu} - \frac{1}{v_o}\right)\right]B_\nu = -(\alpha_\nu + b_\nu)B_\nu + \sum_{\mu=1}^{N} h_{\nu\mu}B_\mu \,. \tag{13}$$

The term with the average group velocity v_o was added for reasons that will become clear shortly. This eigenvalue problem has N solutions. The complete solution of (11) is obtained as a linear superposition of the N eigensolutions plus an integration over the parameter ω:

$$P_\nu(z, t) = \sum_{j=1}^{N} \int_{-\infty}^{\infty} c_j(\omega)B_\nu^{(j)}(\omega)e^{-\alpha^{(j)}(\omega)z}e^{i\omega t}\, d\omega. \tag{14}$$

The superscript j was attached to label the eigenvalues $\alpha(\omega)$ and the eigenvectors $B_\nu(\omega)$. Using the orthogonality of the eigenvectors $B_\nu(\omega)$ of the symmetric real matrix defined by (13),

$$\sum_{\nu=1}^{N} B_\nu^{(i)}B_\nu^{(j)} = \delta_{ij} \,, \tag{15}$$

and the inversion of the Fourier integral allows us immediately to express the expansion coefficient in terms of the power distribution $P_\nu(0, t)$ at the input of the waveguide.

$$c_i(\omega) = \frac{1}{2\pi}\sum_{\nu=1}^{N} \int_{-\infty}^{\infty} B_\nu^{(i)}(\omega)P_\nu(0, t)e^{-i\omega t}\, dt. \tag{16}$$

Equations (14) and (16) represent the complete solution of the time-dependent multimode waveguide problem. In its complete form this formal solution is of little practical value. Thus we proceed to the perturbation solution of a particular problem.

Equation (13) was written in such a way as to suggest a perturbation problem. We added and subtracted the average group velocity v_o in order to obtain the small quantity

$$V = \frac{1}{v_\nu} - \frac{1}{v_o}. \tag{17}$$

V is small since the group velocities of the N modes are only slightly different from each other. Provided we need not include very large values of ω in the analysis ωV can be regarded as a perturbation term in (13). In the spirit of second-order perturbation theory we write the eigenvalue as follows:*

* To first order of perturbation theory only a change in the average group velocity appears. The change of the pulse width depends on the second-order term in (18).

$$\alpha^{(j)}(\omega) = \alpha_0^{(j)} + i\frac{\omega}{v_o} + i\omega\alpha_1^{(j)} + \omega^2\alpha_2^{(j)}. \tag{18}$$

The first term, $\alpha_0^{(j)}$, is the zero-order approximation that corresponds to the solution of the time-independent stationary problem for $\omega = 0$. The terms α_1 and α_2 are first- and second-order perturbation terms. Since the eigenvalue appears in the exponent of an exponential function multiplied by the large quantity z, the perturbation terms can influence the solution (14) very much. The eigenvector $B_\nu(\omega)$ must also be expanded in a similar fashion. However, the zero-order term, $B_{\nu o}$, is by far the most important term in the perturbation series of the eigenvector. The first- and second-order terms of $B_\nu(\omega)$ are of the same relative importance for all values of z and t and can never change very much the zero-order approximation consisting of $B_{\nu o}$ alone. It is therefore sufficient if we approximate $B_\nu(\omega)$ by $B_{\nu o}$. The second-order approximation α_2 is obtained by the well known methods of perturbation theory.

$$\alpha_2^{(i)} = \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{\left[\sum_{\nu=1}^{N}\left(\frac{1}{v_\nu} - \frac{1}{v_o}\right)B_{\nu o}^{(i)}B_{\nu o}^{(j)}\right]^2}{\alpha_0^{(j)} - \alpha_0^{(i)}}. \tag{19}$$

The range of applicability of the second-order perturbation theory is discussed in the Appendix.

We assume that the input power at $z = 0$ is given by

$$P_\nu(0, t) = G_\nu \exp\left[-(t/\tau)^2\right]. \tag{20}$$

From (16) we obtain

$$c_i(\omega) = \frac{\tau}{2\sqrt{\pi}} k_i \exp\left[-\left(\frac{\tau}{2}\omega\right)^2\right] \tag{21}$$

with

$$k_i = \sum_{\nu=1}^{N} G_\nu B_{\nu o}^{(i)}. \tag{22}$$

The integral appearing in (14) is now of the form

$$I = \int_{-\infty}^{\infty} \exp\left[i\omega\left(t - \frac{z}{v_o} - \alpha_1^{(i)}z\right)\right] \exp\left[-\omega^2\left(\frac{\tau^2}{4} + \alpha_2^{(i)}z\right)\right] d\omega$$

$$= \frac{2\sqrt{\pi}}{(\tau^2 + 4\alpha_2^{(i)}z)^{\frac{1}{2}}} \exp\left\{-\frac{\left[t - \left(\frac{1}{v_o} + \alpha_1^{(i)}\right)z\right]^2}{\tau^2 + 4\alpha_2^{(i)}z}\right\}. \tag{23}$$

The second-order perturbation solution of the multimode waveguide problem with a Gaussian input pulse is thus

$$P_\nu(z, t) = \sum_{i=1}^{N} \frac{\tau}{(\tau^2 + 4\alpha_2^{(i)} z)^{\frac{1}{2}}} k_i B_\nu^{(i)} \cdot \exp(-\alpha_0^{(i)} z) \exp\left\{ -\frac{\left[t - \left(\frac{1}{v_o} + \alpha_1^{(i)} \right) z \right]^2}{\tau^2 + 4\alpha_2^{(i)} z} \right\}. \quad (24)$$

IV. DISCUSSION OF THE RESULT OF PERTURBATION THEORY

Equation (24) contains a description of the propagation of a Gaussian input pulse in a multimode waveguide in case of coupling between the guided modes including radiation losses. However, the solution (24) holds also in the absence of coupling. If we assume, for a moment, that there is no coupling between the guided modes, $h_{\nu\mu} = 0$, and that the only losses are heat losses, $\alpha_\nu = \alpha_h$, we immediately have the following solution of (13):

$$\alpha_0^{(i)} = \alpha_h \quad (25)$$

$$\alpha_1^{(i)} = \frac{1}{v_\nu} - \frac{1}{v_o} \quad (26)$$

$$\alpha_2^{(i)} = 0 \quad (27)$$

$$B_{\nu o}^{(i)} = \delta_{i\nu}. \quad (28)$$

This solution means that the modes are uncoupled, each traveling independently of the others with its own group velocity and with the common attenuation constant α_h. Equation (24) can be written in the absence of coupling

$$P_\nu(z, t) = G_\nu e^{-\alpha_h z} \exp\left\{ -\frac{\left(t - \frac{z}{v_\nu} \right)^2}{\tau^2} \right\}. \quad (29)$$

The input pulse, if spread out over all the modes, arrives at the detector at $z = L$ as a succession of pulses. The total spread in the arrival time of the different pulses is

$$T = \left(\frac{1}{v_N} - \frac{1}{v_1} \right) L. \quad (30)$$

Next, we consider the case that all the modes are coupled among each other, $h_{\nu\mu} \neq 0$. Now the eigenvalues can no longer be written down

explicitly. However, we know from earlier work[7] that there are N eigenvalues with their associated eigenvectors. The eigenvalues can be ordered in sequence such that the zero-order solution $\alpha_0^{(1)}$ has the smallest value while all other zero-order eigenvalues assume increasingly larger values. The lowest-order eigenvalue $\alpha_0^{(1)}$ assumes the meaning of the steady-state loss of the multimode waveguide.[7] The steady-state distribution of power over all the modes is proportional to $B_{\nu o}^{(1)}$. By definition, the steady-state distribution is obtained when the sum in (24) can be approximated by its first term since the exponents $\alpha_0^{(i)}z$ have grown so large that all but the first term in the sum are negligible. In this limit, which always exists provided that the waveguide is long enough, we obtain from (24)

$$P_\nu(z, t) = \frac{\tau}{(\tau^2 + 4\alpha_2^{(1)}z)^{\frac{1}{2}}} k_1 B_{\nu o}^{(1)} \cdot \exp(-\alpha_0^{(1)}z) \exp\left\{-\frac{\left[t - \left(\frac{1}{v_o} + \alpha_1^{(1)}\right)z\right]^2}{\tau^2 + 4\alpha_2^{(1)}z}\right\}. \quad (31)$$

The expression (31) is very different from the expression (29) for the uncoupled modes. Whereas each mode traveled independently of all the others with its own group velocity in the absence of coupling, we see from (31) that all the modes travel with the group velocity

$$v_o = \frac{1}{\frac{1}{v_o} + \alpha_1^{(1)}} \quad (32)$$

once the steady-state distribution is established. The term $\alpha_1^{(1)}$ can always be made to vanish by suitable choice of v_o. Furthermore we see that, to the approximation implicit in (31), the pulse remains Gaussian. All the modes suffer identical attenuation according to the attenuation constant $\alpha_0^{(1)}$. The distribution of power over all the modes is determined by the lowest-order eigenvector $B_{\nu o}^{(1)}$. The other eigenvectors have no physical meaning. In fact, these higher eigenvectors can have negative elements whereas the power in each mode must be a positive quantity. Most important for our present discussion is the width of the Gaussian pulse. We see that at $z = L$ it is given by the expression

$$\Delta t = 2(\tau^2 + 4\alpha_2^{(1)}L)^{\frac{1}{2}}. \quad (33)$$

The pulse width increases as the multimode package travels along

the waveguide. For sufficiently large values of L we have $\tau \ll 4\alpha_2^{(1)}L$ and the pulse width becomes proportional to the square root of L. We can define an improvement factor R by the relation

$$R = \frac{\Delta t}{T} = \frac{2(\tau^2 + 4\alpha_2^{(1)}L)^{\frac{1}{2}}}{\left(\frac{1}{v_N} - \frac{1}{v_1}\right)L}. \tag{34}$$

If the width of the input pulse is much less than $4\alpha_2^{(1)}L$, (34) simplifies:

$$R = \frac{4\sqrt{\alpha_2^{(1)}}}{\left(\frac{1}{v_N} - \frac{1}{v_1}\right)\sqrt{L}}. \tag{35}$$

It is apparent that for sufficiently large values of L the improvement factor R, describing the shortening of the pulse as a result of mode coupling, is less than unity.

It is possible to express the loss suffered by the pulse directly as a function of the improvement factor R. The power loss of the multimode signal is given by $\alpha_0^{(1)}L$. If we are interested only in the loss penalty that must be paid for a certain improvement R we can express L in terms of R with the help of (35) and obtain for the loss per improvement R

$$\alpha_0^{(1)}L_R = 16\,\frac{\alpha_0^{(1)}\alpha_2^{(1)}}{\left(\frac{1}{v_N} - \frac{1}{v_1}\right)^2 R^2}. \tag{36}$$

The actual length of waveguide required to incur the loss (36) does no longer appear on the right-hand side. The loss penalty is thus expressed in terms of the zero-order eigenvalue $\alpha_0^{(1)}$ and its second-order perturbation $\alpha_2^{(1)}$, the difference in the group velocities of the first and Nth mode and the desired improvement R. Methods of minimizing this loss penalty occupy most of the remainder of this paper. The length required to obtain a certain pulse width or a certain improvement in the spreading of the pulse is determined by the second-order perturbation $\alpha_2^{(1)}$ of the eigenvalue of the lowest-order mode.

V. APPLICATION TO SLAB WAVEGUIDES

The coupling coefficients $h_{\nu\mu}$ were obtained in Ref. 7 for the case of a slab waveguide. The coupling mechanism is assumed to be the irregular boundary between the core and the cladding of the waveguide. The cladding is assumed to be infinitely extended. In Ref. 7 we assumed

that the function describing the core-cladding interface irregularity had a Gaussian correlation function. Dropping this restriction and expressing the coupling coefficient in terms of the "power spectrum" $F(\beta_\nu - \beta_\mu)$ of the core-cladding interface function $f(z)$ we obtain for the slab waveguide[7]

$$h_{\nu\mu} = \frac{n_1^2 k^2 \sin^2 \theta_\nu \sin^2 \theta_\mu}{2d^2\left(1 + \frac{1}{\gamma_\nu d}\right)\left(1 + \frac{1}{\gamma_\mu d}\right) \cos \theta_\nu \cos \theta_\mu} F(\beta_\nu - \beta_\mu). \tag{37}$$

The parameters appearing in this equation have the following meaning:

n_1 = refractive index of core material
n_2 = refractive index of cladding material
$k = 2\pi/\lambda$ = free-space propagation constant
d = half-width of core
β_ν = propagation constant of mode ν
θ_ν = characteristic angle of mode ν

$$\gamma_\nu = (\beta_\nu^2 - n_2^2 k^2)^{\frac{1}{2}} \tag{38}$$

$$\cos \theta_\nu = \frac{\beta_\nu}{n_1 k}. \tag{39}$$

Values for the radiation losses α_ν were given in Ref. 7. In this paper we consider only relatively narrow "power spectra" $F(\phi)$ coupling each mode only to its nearest neighbor or at the most to a few of its neighbors. In this case only the highest-order mode is coupled to the radiation field. It has been determined by trying out different numerical examples that the radiation losses of the coupled mode system do not depend critically on the loss value α_N, provided that it is large enough. We thus use

$$\left.\begin{aligned} \alpha_\nu &= 0 \qquad \nu \neq N \\ \alpha_N &\to \infty \end{aligned}\right\}. \tag{40}$$

The reason for this insensitivity of the multimode losses on the value of α_N can be explained if we consider that power coupled from any of the guided modes to mode N is quickly lost to radiation. The actual rate of loss from mode N to the radiation modes is not important as long as this rate is high. The actual losses of the multimode guide are determined by the rate at which power flows from mode $N - 1$ to mode N. This rate is determined correctly by the coupling coefficient $h_{N,N-1}$. It is thus not necessary to know the exact values of α_ν. For

actual computations (40) was used with values for α_N that were several orders of magnitude larger than $\alpha_0^{(1)}$.

In Ref. 7 we described the propagation constant of the guided slab modes by using the approximation that is valid far from cutoff. This approximation is based on using

$$\kappa_\nu d = \nu \frac{\pi}{2} \tag{41}$$

with

$$\kappa_\nu = n_1 k \sin \theta_\nu \tag{42}$$

or

$$\beta_\nu = (n_1^2 k^2 - \kappa_\nu^2)^{1/2}. \tag{43}$$

The mode number ν assumes all values from 1 through N with odd numbers indicating even modes while even numbers indicate odd modes. For our present purpose this approximation is not accurate enough since we are interested in accurate values for $\alpha_0^{(1)}$ and $\alpha_2^{(1)}$, particularly in the region where the guided modes are coupled only by the tail of the "power spectrum," so that the spacing between the modes has a critical influence on the actual values of the parameters. Therefore, we chose to use the exact values for κ_ν, β_ν, and γ_ν which are obtained as solutions of the eigenvalue equation

$$\tan \kappa_\nu d = \frac{\gamma_\nu}{\kappa_\nu} \tag{44}$$

for even guided TE modes and from

$$\tan \kappa_\nu d = -\frac{\kappa_\nu}{\gamma_\nu} \tag{45}$$

for odd TE modes. We restrict ourselves to those modes of the slab waveguide that have no variation in y direction and can thus be classified as TE and TM modes. However, only TE modes are being considered. The solutions of (44) and (45) with $n_1 = 1.5$ and $n_1/n_2 = 1.01$ are given in Table I for three particular cases of 5-, 10-, 20-, and 39-mode operation. Also shown in this table are the differences between adjacent values of β_ν. These numbers make it apparent how the spacing between the guided modes increases with increasing mode number. It is also interesting to compare the values of Table I with the approximation (41). It is apparent that (41) approximates the actual values better for the modes of low order in waveguides that support many modes.

TABLE I—PARAMETERS FOR THE TE MODES OF THE SLAB WAVEGUIDE FOR $n_1/n_2 = 1.01$

ν	$\kappa_\nu d$	$\beta_\nu d$	$(\beta_\nu - \beta_{\nu+1})d$
$kd = 35$, 5-mode case ($n_1 = 1.5$, $n_1/n_2 = 1.01$)			
1	1.3821	52.48180	0.05429
2	2.7580	52.42751	0.08936
3	4.1193	52.33815	0.12187
4	5.4507	52.21628	0.14679
5	6.7096	52.06949	———
$kd = 70$, 10-mode case			
1	1.4708	104.98970	0.03089
2	2.9407	104.95881	0.05140
3	4.4086	104.90741	0.07281
4	5.8733	104.83560	0.09199
5	7.3332	104.74361	0.11186
6	8.7861	104.63175	0.13115
7	10.2286	104.50060	0.14939
8	11.6544	104.35121	0.16531
9	13.0498	104.18590	0.17295
10	14.3634	104.01295	———
$kd = 145$, 20-mode case			
1	1.5210	217.49468	0.01595
2	3.0418	217.47873	0.02658
3	4.5624	217.45214	0.03721
4	6.0826	217.41493	0.04783
5	7.6023	217.36710	0.05844
6	9.1214	217.30865	0.06904
7	10.6396	271.23961	0.07962
8	12.1568	217.15999	0.09018
9	13.6728	217.06981	0.10070
10	15.1873	216.96911	0.11119
11	16.7000	216.85793	0.12161
12	18.2105	216.73631	0.13197
13	19.7182	216.60434	0.14222
14	21.2226	216.46213	0.15231
15	22.7225	216.30981	0.16218
16	24.2167	216.14764	0.17168
17	25.7028	215.97596	0.18051
18	27.1767	215.79545	0.18789
19	28.6291	215.60757	0.19023
20	30.0270	215.41733	———

VI. THE DEPENDENCE OF THE LOSS PENALTY ON THE SHAPE OF THE "POWER SPECTRUM"

We observed earlier that a "power spectrum" of the core-cladding interface function of the form shown in Fig. 2 would couple all the guided modes (except mode N) without causing radiation losses. Unfortunately, it is not possible to design dielectric waveguides with core-cladding interfaces whose power spectrum cuts off abruptly at a given specified mechanical frequency. It is thus necessary to study the loss penalty (36) for different "power spectra" in order to determine its dependence on the slope of the "power spectrum."

TABLE I—PARAMETERS FOR THE TE MODES OF THE SLAB WAVEGUIDE FOR $n_1/n_2 = 1.01$ (Continued)

ν	$\kappa_\nu d$	$\beta_\nu d$	$(\beta_\nu - \beta_{\nu+1})d$
kd = 290, 39-mode case			
1	1.5455	434.99725	0.00824
2	3.0909	434.98901	0.01373
3	4.6364	434.97529	0.01922
4	6.1818	434.95607	0.02471
5	7.7271	434.93136	0.03020
6	9.2723	434.90117	0.03569
7	10.8175	434.86547	0.04118
8	12.3625	434.82430	0.04667
9	13.9074	434.77762	0.05216
10	15.4521	434.72546	0.05765
11	16.9967	434.66782	0.06314
12	18.5410	434.60468	0.06862
13	20.0852	434.53605	0.07411
14	21.6291	434.46194	0.07960
15	23.1727	434.38235	0.08508
16	24.6160	434.29727	0.09056
17	26.2590	434.20671	0.09605
18	27.8016	434.11066	0.10152
19	29.3438	434.00914	0.10699
20	30.8856	433.90215	0.11246
21	32.4269	433.78969	0.11792
22	33.9676	433.67176	0.12339
23	35.5077	433.54838	0.12883
24	37.0472	433.41955	0.13428
25	38.5859	433.28572	0.13970
26	40.1237	433.14557	0.14512
27	41.6606	433.00045	0.15052
28	43.1964	432.84993	0.15589
29	44.7310	432.69404	0.16124
30	46.2642	432.53280	0.16655
31	47.7956	432.36625	0.17181
32	49.3251	432.19444	0.17701
33	50.8521	432.01743	0.18212
34	52.3762	431.83531	0.18710
35	53.8964	431.64821	0.19188
36	55.4114	431.45633	0.19630
37	56.9191	431.26003	0.20007
38	58.4149	431.05996	0.20201
39	59.8867	430.85796	

We begin by considering a "power spectrum" that couples all the guided modes equally except for the last two modes that are coupled by the tail of the "power spectrum" distribution. By equally coupled modes we mean that the "power spectrum" remains flat in the region that contributes to coupling between the guided modes. The actual amount of coupling depends, in addition to the "power spectrum," on the mode angles θ_ν [see equation (37)] so that a constant power spectrum couples higher-order modes more strongly than lower-order modes.

The "power spectrum" that is used as a model for a more realistic case is shown in Fig. 3. The flat portion of this spectrum couples mode $N-1$ to mode $N-2$, mode $N-2$ to mode $N-3$, etc., finally coupling several of the low-order modes among each other. Mode N is coupled to mode $N-1$ via the tail of the power spectrum distribution. The factor of importance in this discussion is the level to which the "power spectrum" has decayed from its maximum value at the point that is instrumental in coupling the last two modes to each other. Strong coupling of the last two modes leads to high radiation losses since it causes more power to flow from the guided modes to the radiation modes. The last mode, mode N, couples strongly to the radiation modes since the flat part of the spectrum connects this mode with the continuum of radiation modes. As explained earlier, we do not bother to compute the exact value of the coupling coefficient of mode N to the radiation field since it depends critically on the exact shape of the tail of the "power spectrum," complicating the discussion. Furthermore, it has been established that the actual amount of coupling of mode N to the radiation modes has no influence on the radiation losses of the multimode waveguide as long as the coupling exceeds a certain threshold value. At worst we may overestimate the radiation losses by assigning a large but arbitrary value to the parameter α_N.

Using the "power spectrum" of Fig. 3, we determine the loss penalty of equation (36). Analytically, we express the power spectrum as follows:

$$F(\phi) = \begin{cases} \dfrac{\pi\sigma^2}{\beta_{N-2} - \beta_{N-1}} & \text{for} \quad |\phi| \leqq \beta_{N-2} - \beta_{N-1} \\ K\dfrac{\pi\sigma^2}{\beta_{N-2} - \beta_{N-1}} & \text{for} \quad |\phi| = \beta_{N-1} - \beta_N \end{cases} . \tag{46}$$

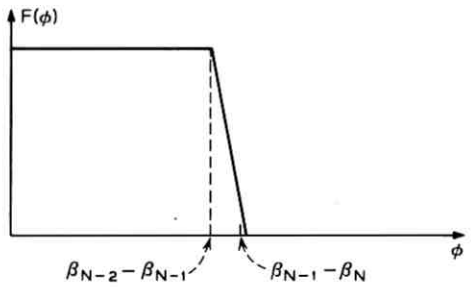


Fig. 3—A more realistic model of the desirable core-cladding interface "power spectrum."

The factor K determines the fraction of the maximum value to which the "power spectrum" has decayed at the point $\phi = \beta_{N-1} - \beta_N$ that is instrumental in coupling mode $N - 1$ to mode N.

The actual calculation is based on the slab waveguide model. According to (13) we must solve the zero-order eigenvalue problem

$$HB^{(i)} = -\alpha_0^{(i)} B^{(i)}. \tag{47}$$

$B^{(i)}$ is a column vector whose elements are $B_{\nu o}^{(i)}$ and H is a matrix with the elements

$$H_{\nu\mu} = h_{\nu\mu} - (\alpha_\mu + b_\mu)\delta_{\nu\mu} \tag{48}$$

that are determined by (8), (37), and (40). The eigenvalues $\alpha_0^{(i)}$ and eigenvectors $B_{\nu o}^{(i)}$ are determined numerically with the help of a computer. Using the result of the zero-order calculation we calculate $\alpha_2^{(1)}$ from equation (19). Instead of $\alpha_0^{(1)}$ and $\alpha_2^{(1)}$, the normalized dimensionless quantity

$$\bar{\alpha}_0^{(1)} = \frac{d}{\sigma^2 k^2} \alpha_0^{(1)} \tag{49}$$

was used in Ref. 7. In this paper the second-order perturbation of the eigenvalues is plotted in the following normalized form:

$$\bar{\alpha}_2^{(1)} = v_o^2 \frac{\sigma^2 k^2}{d} \alpha_2^{(1)}. \tag{50}$$

These normalizations have the advantage of removing the rms deviation σ of the core-cladding interface from the equations so that we need not specify any particular value for this parameter. Figure 4 shows a plot of the normalized* loss penalty $R^2\alpha_0^{(1)}L_R$ as a function of the factor K defined by the second line of (46). The four curves appearing in Fig. 4 were calculated for $kd = 35$ resulting in 5 guided modes, $kd = 70$ corresponding to 10 guided modes, $kd = 145$ or the 20-mode case, and finally for $kd = 290$ which gives 39 guided modes. The curve for 39 modes seems to deviate from the straight line beyond the region for which it appears drawn out in Fig. 4. Since the numerical diagonalization of a 39 by 39 element matrix is quite expensive no attempt was made to explore the precise shape of this curve.

Let us study the 10-mode case in more detail. For $K = 10^{-4}$ we find from Fig. 4 the value $R^2\alpha_0^{(1)}L_R = 0.01$ dB. This means that if we want to obtain a reduction of the uncoupled pulse distortion by a factor of

* This normalization involves only the factor R^2 and is unrelated to (49).

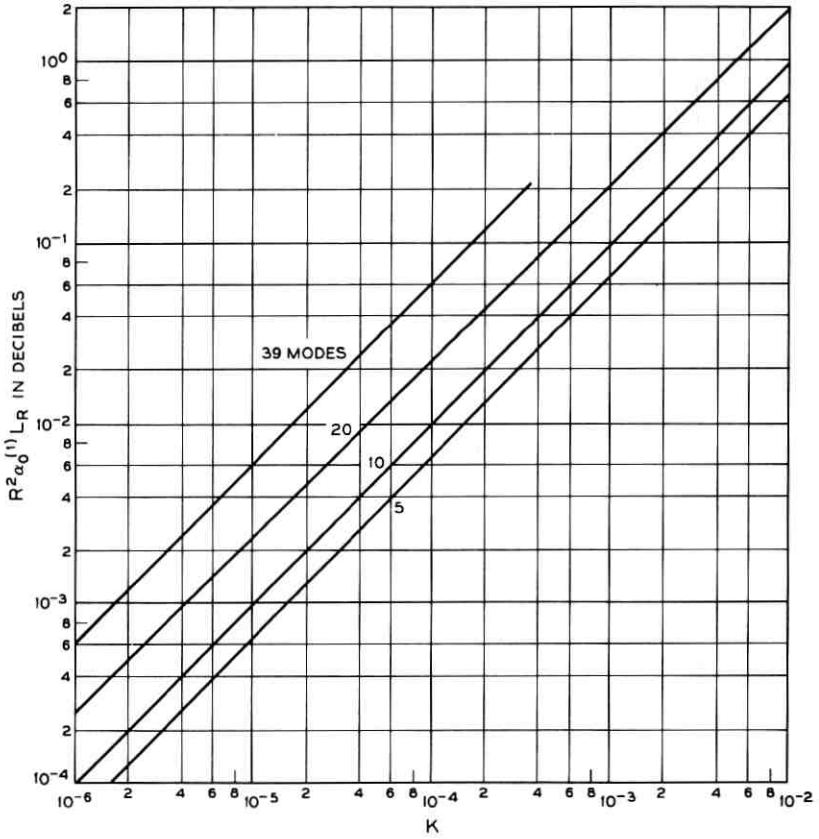


Fig. 4—The loss penalty $R^2\alpha_0{}^{(1)}L_R$ (in dB) that must be paid for a pulse distortion reduction R, ($R < 1$). The variable K is the fraction to which the "power spectrum" has decayed from its flat region at the point where it couples mode $N - 1$ to mode N.

$R = 0.1$, we would have to pay a loss penalty of 1 dB. A reduction of $R = 0.033$ would have to be paid for with 10 dB radiation loss. Let us next consider the slope of the "power spectrum". For the 10-mode case we have according to Table I $(\beta_8 - \beta_9)d = 0.165$. The width of the flat portion of Fig. 3 is thus $0.165/d$. Since $(\beta_9 - \beta_{10})d = 0.173$ we must require that the spectrum drop from unity to 10^{-4} in the "distance" $0.008/d$. The ratio of the region of the slope to the width of the flat region is thus $0.008/0.165 = 0.05$ or the slope extends over 5 percent of the flat portion of the spectrum. For the 39-mode case we obtain correspondingly $(\beta_{37} - \beta_{38})d = 0.2$ and $(\beta_{38} - \beta_{39})d = 0.202$;

the ratio of slope width to the width of the flat part of the spectrum is thus 0.002/0.2 = 0.01 or 1 percent. This comparison shows that the relative width of the slope region must be smaller if more modes can propagate on the waveguide. The "power spectrum" considered for these two examples results in a steady-state power distribution that puts nearly equal power into all the modes with the exception of the last mode which carries essentially no power because of its strong coupling with the radiation modes.

The pulse distortion expected for these examples can be obtained from the data in Table II. In the absence of mode coupling the pulse distortion is given by (30). The inverse of the group velocity of the slab waveguide modes can be calculated with the help of the eigenvalue equations (44) and (45),

$$\frac{1}{v_\nu} = \frac{\partial \beta_\nu}{\partial \omega} = \frac{\beta_\nu^2 + n_1^2 k^2 \gamma_\nu d}{\omega \beta_\nu (1 + \gamma_\nu d)}. \tag{51}$$

The center column of Table II provides the values of T/L. The last column contains the normalized second-order perturbation of the eigenvalues of the matrix eigenvalue problem (13) which is almost independent of K. Given a desired improvement factor, we can calculate the required length L for given rms deviation σ of the core-cladding interface or, vice versa, obtain the rms deviation σ from the given guide length L. Let us consider the following example. We assume that the waveguide length is given as $L = 1$ km, the wavelength of the light signal is $\lambda = 1$ μm. The numerical values in our figures and tables apply to the case $n_1 = 1.5$ and $n_1/n_2 = 1.01$. We want to obtain a pulse distortion improvement resulting in $R = 0.1$. Table III shows

TABLE II—NUMERICAL VALUES FOR THE PULSE DISTORTION IN THE ABSENCE OF COUPLING (THIRD COLUMN) AND THE NORMALIZED SECOND-ORDER PERTURBATION OF THE EIGENVALUE (FOURTH COLUMN) IN THE PRESENCE OF COUPLING FOR THE "POWER SPECTRUM" OF EQUATION (46)

(The normalized second-order perturbation is nearly independent of K)

Number of Modes	kd	$\frac{c}{n_1}\left(\frac{1}{v_N} - \frac{1}{v_2}\right) = \frac{cT}{n_1 L}$	$v_o^2 \frac{\sigma^2 k^2}{d} \alpha_2^{(1)}$
5	35	3.94×10^{-3}	0.02
10	70	5.02×10^{-3}	0.07
20	145	6.69×10^{-3}	0.4
39	290	8.13×10^{-3}	2.0

TABLE III—NUMERICAL VALUES FOR THE EXAMPLE LISTED IN TABLE II CORRESPONDING TO THE "POWER SPECTRUM" OF EQUATION (46) OR FIG. 3

($R = 0.1$, $\lambda = 1$ μm, $K = 10^{-4}$, $L_R = 1$ km)

Number of Modes	d(μm)	T(s)	σ(μm)	$\alpha_o{}^{(1)}L_R$(dB)
5	5.57	1.97×10^{-8}	1.71×10^{-2}	0.63
10	11.1	2.51×10^{-8}	3.56×10^{-2}	1
20	23.1	3.35×10^{-8}	9.16×10^{-2}	2.3
39	46.2	4.07×10^{-8}	2.39×10^{-1}	6

the numerical values of several interesting quantities that follow for this example from Fig. 4 and Table II with the ordinate value of Fig. 3, $K = 10^{-4}$. Listed in Table III are the slab half-width d, the time T to which an infinitely narrow input pulse is stretched out in the absence of coupling (pulse distortion caused by frequency dispersion in the material and due to waveguide effects is not being considered), the rms core-cladding interface irregularity that is required to provide the proper amount of coupling to achieve a pulse distortion improvement of $R = 0.1$ over a distance of $L = 1$ km, and finally, in the last column of the table, the loss penalty that results from $R = 0.1$. The most remarkable result of this example is the fact that such a slight core-cladding interface irregularity is so effective in coupling the guided modes. There may be problems in designing an optical fiber with such a slight core-cladding irregularity. In particular it might be expected that random core-cladding interface irregularities exist whose spectrum is very different from the desired shape shown in Fig. 3. Such unwanted core-cladding interface irregularities would be detrimental since they provide unwanted radiation losses. If the desired core-cladding interface irregularity is made larger, the multimode pulse dispersion is improved more than the factor $R = 0.1$ assumed here. However, such an improved mode mixing must be paid for with a higher loss penalty that can be reduced only by decreasing the factor K [the ordinate in Fig. 4; see also equation (46)].

The "power spectrum" of Fig. 3 is not realizable in practice and was used only to gain insight into the relation between pulse distortion improvement and loss penalty. It is interesting to pursue this question further and study other "power spectra." An obvious choice is a "power spectrum" of the form

$$F(\phi) = \frac{m\sigma^2 \sin \frac{\pi}{m}}{\Delta\beta\left[1 + \left|\frac{\phi}{\Delta\beta}\right|^m\right]}. \tag{52}$$

It can be seen in Fig. 5 that (52) approximates the idealized power spectrum of Fig. 3 for large values of m. The "power spectrum" (52) has the advantage of making it possible to study the loss penalty not only as a function of the slope of the power spectrum but also as a function of its width $\Delta\beta$. Figure 6 shows a number of loss penalty curves for the 10-mode case as a function of the width parameter $\Delta\beta d$ of the "power spectrum" (52). The curve parameter is the exponent m. Figure 6 reveals several interesting properties of the loss penalty in relation to the shape of the "power spectrum". The curves shown in Fig. 6 have a maximum. To the right of this maximum the loss penalty is decreasing very rapidly making it appear as though this were a good region of operation. However, a more detailed investigation of the lowest-order eigenvector $B_{\nu o}^{(1)}$ reveals that the steady-state power distribution in the region to the right of the maximum in Fig. 6 permits

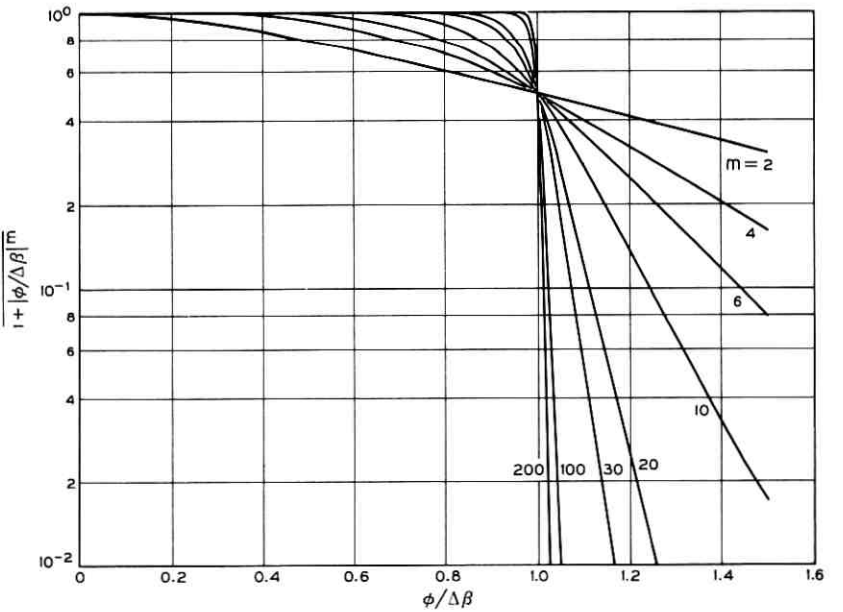


Fig. 5—A simple power law model for the "power spectrum."

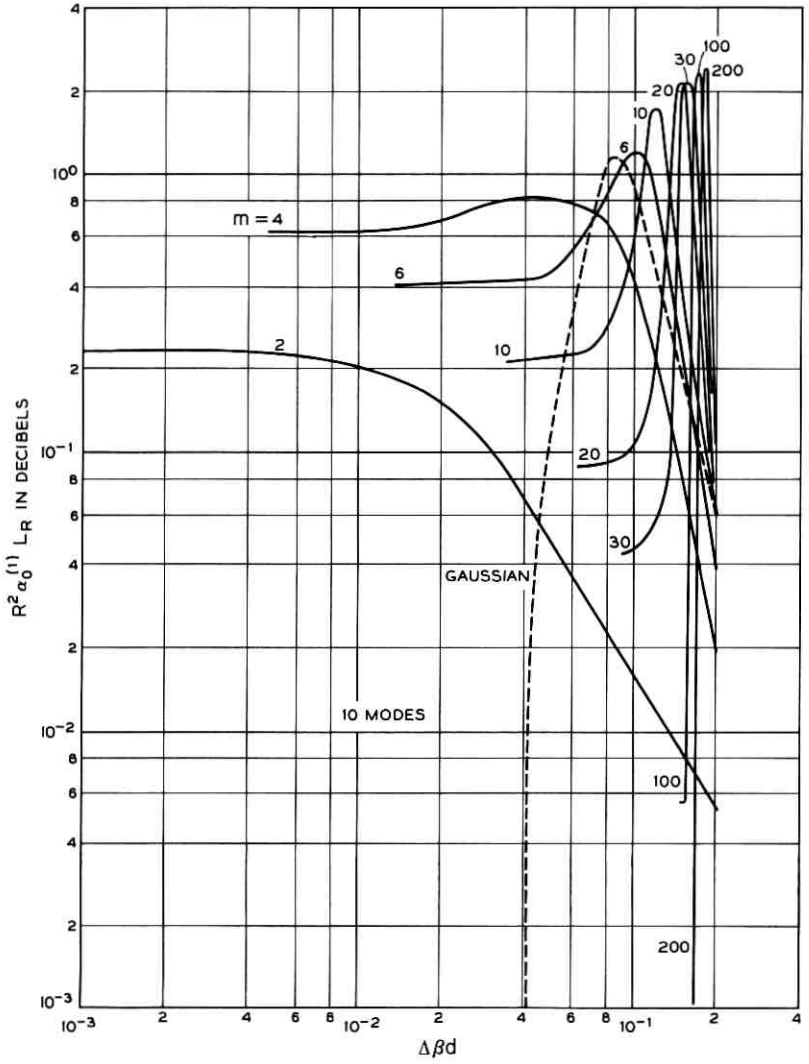


Fig. 6—The loss penalty for the "power spectrum" of Fig. 5 for $n_1/n_2 = 1.01$.

power to remain essentially only in the lower-order modes. This mode of operation is, of course, favorable from the point of view of pulse distortion. But if the fiber is to be excited with a light-emitting diode feeding power equally into all the modes, most of the power is lost in the transient before the steady-state power distribution establishes

itself. The region to the right of the maxima is thus considered to be unsuitable for efficient power transmission in a multimode waveguide. To the left of the maxima we find a sharp drop off ending in a plateau. On the left slope of the curve we enter a region in which all the modes of the waveguide share more evenly the total amount of power carried by the guide. On the plateaus the power distribution versus mode number is flat. The plateau regions are thus desirable from the point of view of efficient multimode operation with reduced pulse dispersion. Unfortunately the level of the loss penalty is rather high for low values of the exponent. We see also that the plateau regions move to the left for decreasing values of the exponent m. This indicates that low loss can be achieved only if the high-order modes are coupled only through the tails of the "power spectrum" curve. The last column of Table I gives for $\nu = 8$ the value of $(\beta_8 - \beta_9)d = 0.165$ for the 10-mode case. Since $(\beta_9 - \beta_{10})d = 0.173$, this means that the ideal region of operation would be $0.173 > \Delta\beta d > 0.165$ and $m \to \infty$. Figure 6 shows clearly the trend in this direction. For lower values of m the point of operation must be shifted to much smaller values of $\Delta\beta d$ and is accompanied by an increase in the loss penalty in the plateau region. Simultaneously with decreasing loss we obtain a decrease in the coupling between the guided modes. This behavior is very apparent in Fig. 7 which shows a plot of the normalized second-order perturbation of the eigenvalue which, as we know, determines the width of the steady state Gaussian pulse. In the plateau regions of Fig. 6 the normalized value of $\alpha_2^{(1)}$ becomes very large indicating a rapidly decreasing efficiency of pulse delay distortion reduction. The actual value of the length of the Gaussian pulse or of the improvement factor R is, of course, dependent on the rms deviation σ of the core-cladding interface. The flattening out of the loss penalty curves in the plateau region can be attributed to the fact that the loss of coupling efficiency among the guided modes with decreasing width of the "power spectrum" is accompanied by a corresponding reduction in power transfer to the radiation modes.

The region immediately to the left of the plateaus shown in Fig. 6 is interesting also. We terminated the curves since the numerical matrix diagonalization routine failed to function for values of $\Delta\beta d$ to the left of the end of each curve. This mathematical phenomenon has an important physical reason. As the coupling between the guided modes decreases, we encounter a regime of instability where the steady-state power distribution depends on the initial power distribution. In the absence of coupling each arbitrary power distribution is a steady-state distribution since power is no longer exchanged between the

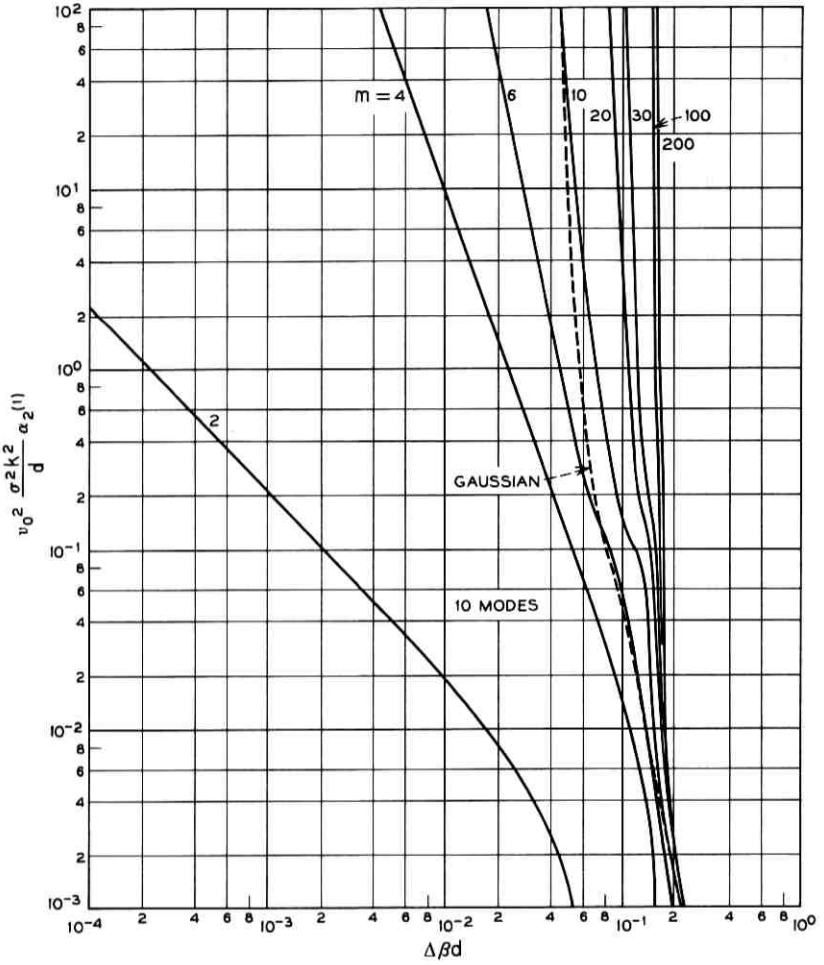


Fig. 7—The normalized second-order perturbation of the eigenvalue (this parameter determines the width of the pulse) $\alpha^{(1)}$ as a function of the width parameter of the "power spectrum" of Fig. 5. $n_1/n_2 = 1.01$.

modes. The onset of this instability causes the matrix diagonalization routine to fail. In fact, it is interesting to observe how different matrix diagonalization programs return different eigenvectors and eigenvalues in the region to the left of the plateaus. This phenomenon does not indicate errors in these programs but shows that the final solution depends on random fluctuations and is no longer uniquely determined. It is clear that the regions to the left of the plateaus are unsuitable for

purposes of pulse distortion reduction by means of mode coupling. Figure 6 thus shows that significant pulse distortion reduction with low loss penalty is possible at least in principle by using a very large value of the exponent m and operating at the lowest point of the steep slopes to the left of the maxima. The required tolerances for $\Delta\beta d$ become increasingly more critical as the value of the exponent m is being increased.

The curve for $m = 2$ in Fig. 6 remains entirely in the region, corresponding to the right of the maxima of the remaining curves, where most of the power is carried by the lower-order modes. The power spectrum (52) with $m = 2$ is thus unsuitable for our purposes. For comparison purposes Fig. 6 shows the loss penalty for the Gaussian "power spectrum,"

$$F(\phi) = \frac{2\sqrt{\pi}\,\sigma^2}{\Delta\beta} \exp\left[-\left(\frac{\phi}{\Delta\beta}\right)^2\right], \tag{53}$$

as a dotted line.

For large values of m the curves of Fig. 7 are not very suitable for computing the required rms deviation σ for given values of R and L because of their extremely steep slopes. It is advisable to use Fig. 5 and the values for the propagation constant differences of Table I to define an equivalent factor K [compare (46)] and use Fig. 4 for obtaining the loss penalty and Table II for obtaining the value of the normalized second-order perturbation of the lowest-order eigenvalue. (This value is very nearly independent of K. The steep slopes of the curves of Fig. 7 result from their sensitive dependence on $\Delta\beta d$.) Instead, we consider as an example a moderately large value of m. Let us assume that $m = 20$ and let us use $R = 0.1$, $L = 1$ km, $\lambda = 1$ μm, and the 10-mode case, $kd = 70$. Choosing as the operating point $\Delta\beta d = 0.1$ we obtain from Fig. 6 the loss penalty $\alpha_0^{(1)} L_R = 10$ dB. From Fig. 7 we find

$$v_0^2 \frac{\sigma^2 k^2}{d} \alpha_2^{(1)} = 4.2. \tag{54}$$

The difference of the group velocities of mode 1 and mode 10 follows from Table II. We thus obtain, with the help of (35), $\sigma = 0.27$ μm. This example results in ten times higher loss than the 10-mode case listed in Table III but the required rms deviation is more easily realizable than that of Table III.

Figures 6 and 7 apply to the 10-mode case. In order to explore the dependence of the loss penalty and the pulse length factor $\alpha_2^{(1)}$ on the

number of modes, Fig. 8 shows a comparison of the loss penalty for the exponent $m = 10$ [see (52)] for 5, 10, 20, and 39 modes as a function of the "power spectrum" width parameter $\Delta\beta d$. The curve for $N = 39$ was not extended out of the plateau region because of the cost involved in the diagonalization of the large matrix. The curves of Fig. 8 show that the loss penalty decreases slowly with increasing mode number. It appears that the dependence of the loss penalty in the plateau region on the mode number is approximately given by $N^{-1/2}$ for large values of N. The curve for $N = 5$ does not obey this law, possibly because this number is still too small.

The corresponding dependence of the normalized second-order perturbation of the eigenvalue on mode number is shown in Fig. 9. The two points for the curve of $N = 39$ in the region shown in Fig. 8 lie above the range of the figure.

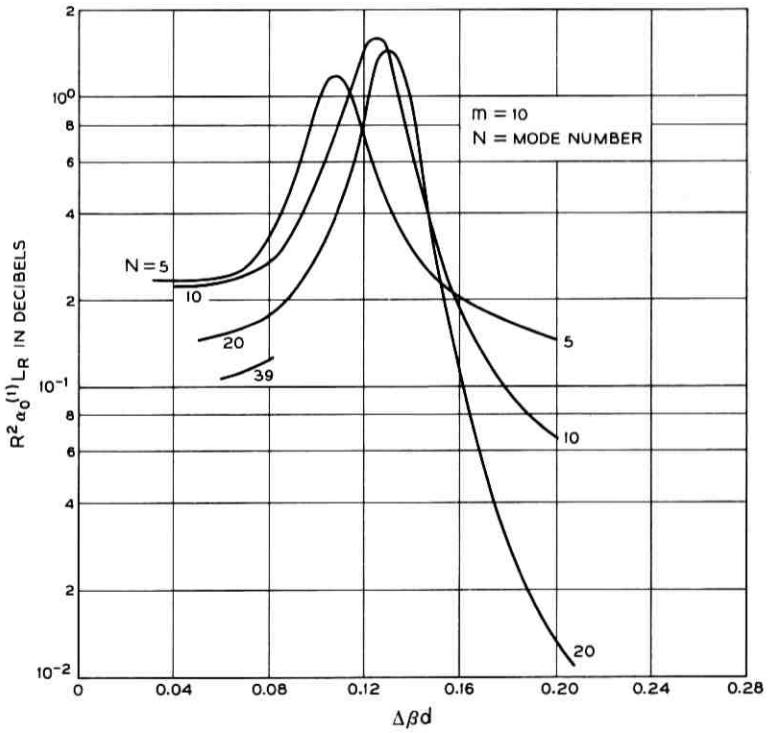


Fig. 8—Mode dependence of the loss penalty for $m = 10$ is shown as a function of the width parameter of the "power spectrum."

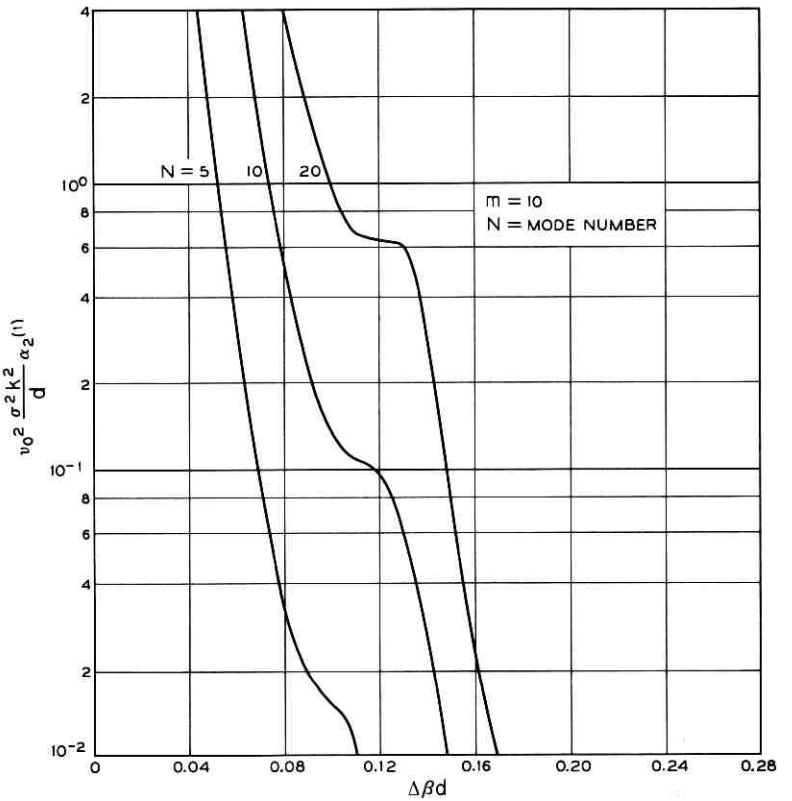


Fig. 9—The second-order perturbation of the eigenvalue $\alpha^{(1)}$ for three different mode numbers as a function of the width parameter of the power spectrum for $m = 10$.

In order to obtain an insight into the meaning of these curves (Figs. 8 and 9), Table IV presents the rms deviation that is required to achieve a pulse distortion improvement of $R = 0.1$ at $\Delta\beta d = 0.08$ for $\lambda = 1$ μm. The required values of the rms deviation are only slightly larger than those listed in Table III but the loss penalty for a pulse distortion reduction of $R = 0.1$ is far higher since the numbers in Table IV pertain to a relatively small value of the exponent, $m = 10$.

VII. DEPENDENCE ON THE REFRACTIVE INDEX DIFFERENCE

So far, all our examples applied to dielectric slab waveguides with a core-to-cladding-index ratio of $n_1/n_2 = 1.01$. In order to explore the

TABLE IV—NUMERICAL VALUES FOR AN EXAMPLE BASED ON THE "POWER SPECTRUM" OF EQUATION (52) OR FIG. 5

($R = 0.1$, $\lambda = 1$ μm, $\Delta\beta d = 0.08$, $m = 10$, $L_R = 1$ km)

Number of Modes	$\sigma(\mu m)$	$\alpha_o^{(1)} L_R$(dB)
5	2.16×10^{-2}	33
10	9.44×10^{-2}	27
20	2.97×10^{-1}	17
39	7.96×10^{-1}	12

dependence of our results on the index differences we give a few data for the case $n_1/n_2 = 1.005$. Table V shows results similar to those of Table I for the 10-mode case. The most important differences of these two examples are the increase of the value of kd required to obtain 10 guided modes from 70 for $n_1/n_2 = 1.01$ to $kd = 97$ for $n_1/n_2 = 1.005$. In addition, we find that the values for $(\beta_\nu - \beta_{\nu+1})d$ have become smaller. The "power spectrum" of the core-cladding interface irregularities must thus become narrower and have steeper slopes in order to yield the same loss penalty in both cases. The task of designing this "power spectrum" is thus more difficult for smaller index differences between core and cladding. The loss penalty that must be paid in this case is shown in Fig. 10 which is similar to Fig. 6 except that the abscissa is now represented on a linear scale. A few of the curves of Fig. 6 are also shown in Fig. 10 as dotted lines for comparison purposes. It is apparent that the smaller index difference causes the curves to shift

TABLE V—PARAMETERS FOR THE TE MODES OF THE SLAB WAVEGUIDE FOR $n_1/n_2 = 1.005$

[$kd = 97$, 10-mode case ($n_1 = 1.5$, $n_1/n_2 = 1.005$)]

ν	$\kappa_\nu d$	$\beta_\nu d$	$(\beta_\nu - \beta_{\nu+1})d$
1	1.4693	145.49258	0.02224
2	2.9375	145.47034	0.03700
3	4.4037	145.43334	0.05166
4	5.8665	145.38168	0.06615
5	7.3243	145.31553	0.08036
6	8.7746	145.23517	0.09410
7	10.2137	145.14107	0.10699
8	11.6347	145.03408	0.11791
9	13.0213	144.91617	0.12071
10	14.3012	144.79546	—

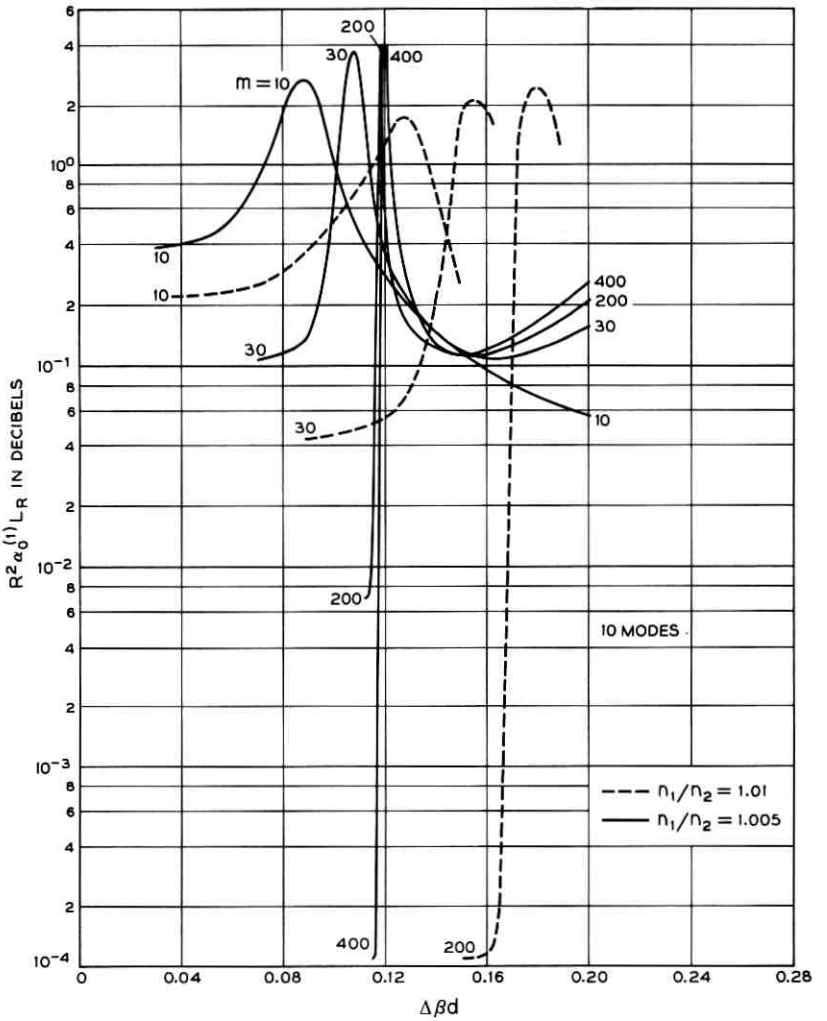


Fig. 10—The loss penalty for the "power spectrum" of Fig. 5. The solid lines apply to the case $n_1/n_2 = 1.005$, the dotted lines are drawn in for comparison with the case $n_1/n_2 = 1.01$.

to the left. Furthermore, we see that the loss penalty is increased for a given value of m.

The normalized second-order perturbation of the lowest-order eigenvalue is shown in Fig. 11 for $n_1/n_2 = 1.005$. To obtain a performance for $n_1/n_2 = 1.005$ comparable to the case $n_1/n_2 = 1.01$ requires a higher-

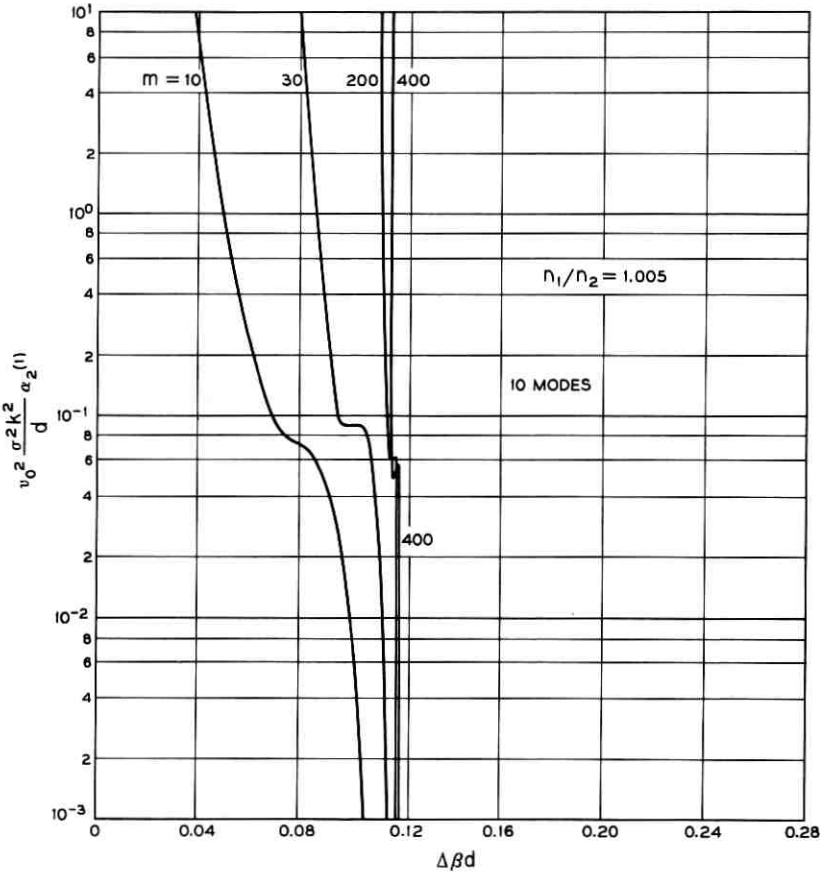


Fig. 11—The normalized second-order perturbation of the eigenvalue as a function of the width parameter of the "power spectrum" for the case $n_1/n_2 = 1.005$.

power m of the "power spectrum" curve (52) and a narrower spectrum.

The discussion of the loss penalty was based on the relative improvement factor R that determines how much the pulse is shortened compared to the pulse distortion in the absence of coupling. For an absolute comparison of pulse distortion it is important to keep in mind that the pulse distortion in the absence of coupling is less severe in a waveguide with smaller core-cladding index difference. Whereas we obtain

$$\frac{c}{n_1}\left(\frac{1}{v_N} - \frac{1}{v_1}\right) = 5.02 \times 10^{-3}$$

for the 10-mode case and $n_1/n_2 = 1.01$ (from Table II), we now have

$$\frac{c}{n_1}\left(\frac{1}{v_N} - \frac{1}{v_1}\right) = 1.94 \times 10^{-3}$$

for ten modes and $n_1/n_2 = 1.005$. The pulse distortion is already improved by a factor of 2.6 even in the absence of mode coupling simply by the reduction in the core-cladding index difference. The higher loss penalty that must now be paid for a given relative improvement R (provided we assume the same power factor m in both cases) is thus not quite as serious in absolute terms since the pulse distortion is already smaller for the smaller index difference.

VIII. APPLICATION TO ROUND OPTICAL FIBERS

Our discussion of numerical results of the pulse distortion reduction by means of coupling between guided modes was limited to the example of the slab waveguide. However, the slab waveguide results are meaningful for predicting the behavior of round optical fibers. In order to find the connection between round fibers and the slab waveguide let us use one intermediate step and consider a dielectric waveguide with square cross section. The behavior of dielectric waveguides with square cross section is very similar to the behavior of dielectric waveguides with circular cross section if the cross-sectional areas of the two waveguides are comparable. The modes of the slab waveguide can be visualized as plane waves (traveling in the core medium) that are reflected at the core-cladding interface. In the case of the slab waveguide we were concerned only with plane waves whose propagation vectors all lie in a plane that is positioned perpendicular to the core-cladding interface. The complete set of modes of the square waveguide is also made up of plane waves, except that now, for each plane wave propagating at a certain angle with respect to one parallel pair of interfaces, we have N waves whose propagation angle with one set of interfaces is fixed but whose angles with respect to the perpendicular set of interfaces are varying. Instead of the original N slab waveguide modes we find N^2 modes in the dielectric waveguide with square cross section. This discussion becomes more accurate with increasing values of N.

Turning now to the round optical fiber we can deduce from its similarity with the waveguide of square cross section that its total number of guided modes is also approximately given by N^2, where N is the number of slab waveguide modes. We assume that the slab half-width d corresponds approximately to the radius r of the round fiber. The 10-mode slab waveguide thus corresponds to a round optical fiber

supporting 100 guided modes. The largest number of modes considered in this paper, $N = 40$, thus corresponds to a round optical fiber supporting 1600 guided modes.

We assume that the coupling mechanism was caused by core-cladding interface irregularities of the slab waveguide with no variation in y direction. Extending our results to round optical fibers we consider a fiber with random core radius fluctuations with a carefully designed "power spectrum" but with no dependence of the core-cladding interface on the azimuth. Core radius variations of a round optical fiber couple only modes with the same azimuthal symmetry. This means that whole families of modes would remain uncoupled from each other. However, this defect does not prevent the pulse distortion reduction scheme from working. We must keep in mind that the modes within each family (all having the same azimuthal symmetry) cover the entire range in β space between n_2k and n_1k. The modes in each family would thus give rise to the spreading of the pulse described by (30). The random radius changes couple all the modes within each family reducing the spreading of the pulse to the amount given by (33). We thus can expect that instead of many pulses each traveling with its own group velocity, we now have families of pulses each traveling with an average velocity but each pulse being shortened by the coupling mechanism. Only if the pulses composed of each family of modes traveled with different average velocities would the pulse distortion reduction in the round fiber work less efficiently than predicted by our slab waveguide model. However, it can be expected that the average velocity of each pulse lies half way between the group velocities of the fastest and the slowest pulse in each family of modes. These average velocities must be very nearly the same. We thus expect that the pulse distortion reduction described in this paper is applicable to the round optical fiber. In designing the "power spectrum" of the random radius variation function, we must consider the spacing in β space between the modes of each family and must try to shape the "power spectrum" such that the guided modes within each family are coupled to each other with the exception of the highest-order mode whose strong coupling to the radiation field would cause excessive radiation losses. If this condition cannot be fulfilled for all mode families we may have to pay a higher loss penalty, losing certain families of modes more rapidly than others.

IX. CONCLUSIONS

We have found that pulse distortion resulting from the different group velocities of the many guided modes of a multimode waveguide

can be effectively reduced by providing coupling between the guided modes. We have seen, moreover, that we must be very careful to limit the coupling to those guided modes that are not coupled directly to the radiation field. Designing the power spectrum of the core-cladding interface irregularities responsible for the coupling mechanism in such a way, that the last guided mode becomes effectively uncoupled from the other guided modes, reduces radiation losses. The detailed discussion of the loss penalty that must be paid for a certain pulse distortion reduction showed that the slope of the core-cladding interface "power spectrum" must be extremely steep. Ideally, an infinitely steep slope would be desirable. The flat region of the power spectrum, shown in Figs. 3 and 5, is not a critical requirement. The delay distortion reduction would not be impaired if the power spectrum has ripples in this region.

In closing, it appears prudent to repeat the warning that our predictions are based on first-order perturbation theory. They hold only for weak coupling. For strong coupling the radiation losses are expected to be larger than predicted here. Furthermore, our theory predicts only average power values. Higher-order effects and the fluctuation problem will be discussed in future publications.

APPENDIX

The second-order perturbation theory is valid only provided that certain requirements are met. It is apparent from (18) that perturbation theory can be applied only if

$$\alpha_0^{(i)} \gg \omega^2 \alpha_2^{(i)} \tag{55}$$

with a suitably chosen value of ω. Instead of $\alpha_0^{(i)}$ and $\alpha_2^{(i)}$ the normalized quantities

$$\bar{\alpha}_0^{(i)} = \frac{d}{\sigma^2 k^2} \alpha_0^{(i)} \tag{56}$$

and

$$\bar{\alpha}_2^{(i)} = \frac{v_o^2 \sigma^2 k^2}{d} \alpha_2^{(i)} \tag{57}$$

are used for the actual numerical calculations. In terms of the numerical values used, (55) can thus be written as

$$\frac{v_o^2 \sigma^4 k^4}{\omega^2 d^2} \bar{\alpha}_0^{(i)} \gg \bar{\alpha}_2^{(i)}. \tag{58}$$

The effective value for ω to be used in these inequalities is related to the width τ of the input pulse,

$$\omega \approx \frac{1}{\tau}. \tag{59}$$

The required condition for the applicability of the perturbation theory is thus

$$\frac{v_o^2 \sigma^4 k^4 \tau^2}{d^2} \bar{\alpha}_0^{(i)} \gg \bar{\alpha}_2^{(i)}. \tag{60}$$

REFERENCES

1. Personick, S.D., "Time Dispersion in Dielectric Waveguides," B.S.T.J., *50*, No. 3 (March 1971), pp. 843–859.
2. Rowe, H. E., and Young, D. T., "Transmission Distortion in Multimode Random Waveguides," IEEE Transactions on Microwave Theory and Techniques, *MTT-20*, No. 6 (June 1972), pp. 349–365.
3. Marcuse, D., "Derivation of Coupled Power Equations," B.S.T.J., *51*, No. 1 (January 1972), pp. 229–237.
4. Marcuse, D., "Mode Conversion Caused by Surface Imperfections of a Dielectric Slab Waveguide," B.S.T.J., *48*, No. 10 (December 1969), pp. 3187–3215.
5. Marcuse, D., *Light Transmission Optics*, New York: Van Nostrand Reinhold Comp., 1972.
6. Marcuse, D., "Radiation Losses of Waveguides in Terms of Power Spectrum of the Wall Distortion Function," B.S.T.J., *48*, No. 10 (December 1969), pp. 3233–3242.
7. Marcuse, D., "Power Distribution and Radiation Losses in Multimode Dielectric Slab Waveguides," B.S.T.J., *51*, No. 2 (February 1972), pp. 429–454.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

Synthesis of Voiced Sounds From a Two-Mass Model of the Vocal Cords

By K. ISHIZAKA and J. L. FLANAGAN

(Manuscript received January 13, 1972)

A model of voiced-sound generation is derived in which the detailed acoustic behavior of the human vocal cords and the vocal tract is computed. The vocal cords are approximated by a self-oscillating source composed of two stiffness-coupled masses. The vocal tract is represented as a bilateral transmission line. One-dimensional Bernoulli flow through the vocal cords and plane-wave propagation in the tract are used to establish acoustic factors dominant in the generation of voiced speech. A difference-equation description of the continuous system is derived, and the cord-tract system is programmed for interactive study on a DDP-516 computer. Sampled waveforms are calculated for: acoustic volume velocity through the cord opening (glottis); glottal area; and mouth-output sound pressure. Functional relations between fundamental voice frequency, subglottal (lung) pressure, cord tension, glottal area, and duty ratio of cord vibration are also determined.

Results show that the two-mass model duplicates principal features of cord behavior in the human. The variation of fundamental frequency with subglottal pressure is found to be 2 to 3 Hz/cm H_2O, and is essentially independent of vowel configuration in the programmed tract. Acoustic interaction between tract eigenfrequencies and glottal volume flow is strong. Phase difference in motion of the cord edges is in the range of 0 to 60 degrees, and control of cord tension leads to behavior analogous to chest/falsetto conditions in the human. Phonation-neutral, or rest area of cord opening, is shown to be a critical factor in establishing self-oscillation. Finally, the complete synthesis system suggests an efficient, physiological description of the speech signal, namely, in terms of subglottal pressure, cord tension, rest area of cord opening, and vocal-tract shape.

I. GENERATION OF VOICED SOUNDS IN SPEECH

The vocal cords constitute the sound source for all voiced sounds in speech. The cords consist of opposing ligaments which form a con-

striction at the top of the trachea where it joins to the lower vocal tract. When air is expelled at sufficient velocity through this orifice (the glottis), the cords vibrate and act as an oscillating valve which interrupts the air flow into a series of pulses. These pulses of volume flow serve as the excitation source for the vocal tract in all voiced sounds, and the passive resonances of the vocal tract are excited by the glottal pulses. Voice quality and prosodic features of speech are therefore strongly dependent upon the properties of cord vibration.

In the synthesis of speech by machines (for automatic voice response from computers, for example) it is desirable to make the synthetic voice as natural sounding as possible. Toward this end, it is necessary to understand the fundamental acoustic principles of voiced-sound generation and how these factors might be incorporated into a machine voice. Further, in a rather different area, the successful medical diagnosis (and correction) of voice disorders depends upon an understanding of the critical parameters of vocal-cord behavior. As in the case of computer synthesis, medical diagnosis can be facilitated through an accurate and viable model of the human vocal cords. Applications such as these, together with fundamental interests in the acoustics of speech, provide a motivation for modeling the acoustic behavior of the vocal cords.

II. SELF-OSCILLATING MODELS OF THE VOCAL CORDS

The first quantitative self-oscillating model of the vocal cords was devised by one of the authors and implemented with a vocal-tract synthesizer on a digital computer.[1,2] This model was subsequently elaborated to include the mechanism of voiceless sound generation as well,[3] and was used for the synthesis of simple speech samples.

In this early work, the vocal cords were approximated as a simple mechanical oscillator, composed of single opposing masses, springs, and nonlinear damping–that is, a so-called one-mass approximation of each vibrating cord. The oscillating masses were permitted only lateral displacement and were driven by a function of the subglottal pressure and the Bernoulli pressure in the glottal orifice. The heretofore much-used assumption of linear separability of sound source and vocal tract was not made, and acoustic factors such as voice pitch, waveform of glottal flow, and glottal duty factor were derived as self-determined functions of physiological parameters, namely, subglottal (lung) pressure, vocal-cord tension (or natural frequency), "neutral" glottal area, and vocal-tract shape.

The waveforms of glottal area and volume velocity obtained in this first study were similar to those observed in high-speed motion pictures of the human vocal cords and in inverse filtering of natural speech. Further, the results revealed how the acoustic interaction between the vocal cords and the vocal-tract shape (through its driving-point impedance) could influence the waveform and period of the glottal flow. Control of the physiological parameters, subglottal pressure, cord tension, neutral area, and vocal-tract shape, were shown to be sufficient for the synthesis of voiced and voiceless sounds.[3]

Although the one-mass model could produce acceptable voiced-sound synthesis and simulate many of the properties of glottal flow, it was inadequate to produce other physiological detail in vocal cord behavior. For example, the amount of acoustic interaction displayed between source and tract was greater than observed in human speech.* The one-mass model was congenitally incapable of sustained oscillation for a capacitive input load of the vocal tract–corresponding to oscillation at a frequency just above a formant (or eigen) frequency of the tract. Also, a physiologically-natural correlate of chest and falsetto registers and a phase-difference in the motion of the cord edges were lacking.

To incorporate more physiological properties, multiple-mass representations of the cords were therefore considered.[4–6] The cord ligaments can be mechanically represented with as distributed a system as desired, i.e., periodic structures of masses, springs, and losses. However, theoretical work has shown that a two-mass approximation[6,7] can account for most of the relevant glottal detail, including phase differences of upper and lower edges and oscillation for a capacitive input impedance of the vocal tract. An initial effort at computer simulation of these factors[4] produced realistic phase differences and chest-falsetto dichotomy, but nonrealistic dependence on acoustic load. The difficulty lay in the equivalent circuit of the glottal orifice, the manner of its control, and the physiological data available for the simulation.

The present work seeks a comprehensive and definitive treatment of the relevant acoustic theory and the existing physiological data. As in the earlier study,[2] simulation on an interactive DDP-516 laboratory computer is the tool by which the model is assessed and the unknown constants are estimated. In the sequel, the level of understanding and the realism attained by the two-mass model will be discussed.

---

* The amount of interaction is critically dependent upon the trans-glottal pressure distribution. In the first work, van den Berg's measurements of glottal pressure were used.

III. MECHANICAL RELATIONS FOR THE TWO-MASS MODEL

The vocal cords are assumed to be bilaterally symmetric. The properties of only one cord are therefore discussed, the same being implied for the opposing cord. A schematic diagram of the glottal system is shown in Fig. 1. The trachea, leading to the lungs, is represented by the pipe to the left. The larynx tube, leading to the vocal tract, is to the right. These tubes are assumed to be cylindrical in shape and are fixed in size. The glottis constitutes a constriction between these tubes, and the size of the constriction depends upon the cord displacement. The inlet to the glottal constriction occurs over the contraction distance l_c. Expansion back to the vocal-tract cross section occurs over the distance l_e. Aerodynamic pressures relevant to the following discussion are indicated in Fig. 1.

In the two-mass model, the vocal cord is divided in depth (thickness) into an upper and a lower part. Each part consists of a simple mechanical oscillator having a mass, spring, and damping (m, s, and r). The two masses of a cord, m_1 and m_2 are permitted only lateral motion, x_1 and x_2, and the masses are coupled by a linear spring, of stiffness k_c, as shown in Fig. 1. Other factors shown in Fig. 1 are:

l_g	the effective length of the vocal cords (or of the glottal slit),
d_1 and d_2	the thickness of m_1 and m_2, respectively,
s_1 and s_2	the equivalent springs,
r_1 and r_2	the equivalent viscous resistances,
A_{g01} and A_{g02}	the cross-sectional areas of the glottal, slit when m_1 and m_2 are at rest (i.e., the "phonation neutral" areas),
U_g	the average volume velocity across the glottal area.

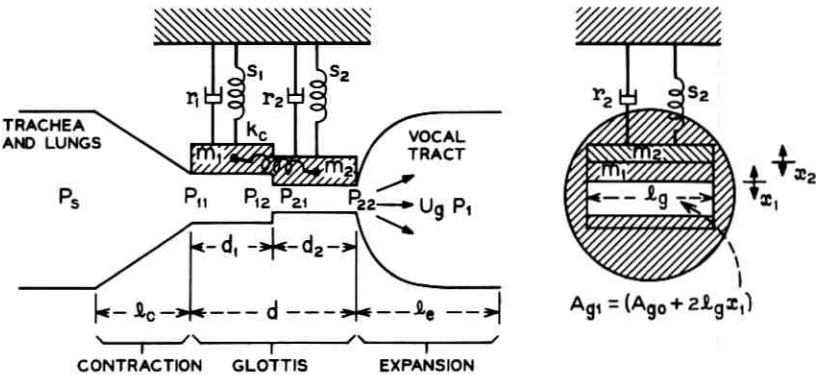


Fig. 1—Schematic diagram of the two-mass approximation of the vocal cords.

Owing to the assumption of bilateral symmetry, variations in cross-sectional areas due to the lateral displacements x_1 and x_2 are doubled to give the total area variation; that is, the cross-sectional areas at the two masses are:

$$A_{g1} = A_{g01} + 2l_g x_1$$
$$A_{g2} = A_{g02} + 2l_g x_2.$$

3.1 *Nature of the Vocal-Cord Springs*

The function of the linear coupling spring, k_c, is to represent an effect of flexural stiffness in the lateral direction of the vocal cords. This variable flexural stiffness results from varying the thickness and stiffness of the cords by action of the thyroarytenoid muscle (vocalis).

The springs, s_1 and s_2, are an equivalent representation of the tension of the vocal cords, which becomes firmer due to contraction of the anteriol cricothyroid muscle and other muscles. The springs, s_1 and s_2, are given a nonlinear characteristic to conform to the stiffness as measured on fresh, excised human vocal cords.[8] The nonlinear relation between the deflection from the equilibrium position and the force required to produce this deflection is given by

$$f_{sj} = k_j x_j (1 + \eta_{kj} x_j^2), \qquad j = 1, 2, \tag{1}$$

where f_{sj} is the force required to produce x_j, k_j is the linear stiffness, and η_{kj} is the coefficient describing the nonlinearity of the spring, s_j, being positive in this case.

During closure of the glottis, the model should satisfy realistic conditions at the colliding surfaces of the vibrating masses, m_1 and m_2 with their opposing counterparts. A contact force at collision will cause some deformation in the flesh of the vocal cords. The restoring force at this deformation can be represented by an equivalent spring s_{hj} ($j = 1, 2$). For simplicity, a nonlinear characteristic, similar to eq. (1), is assumed for the spring s_{hj}, that is,

$$f_{hj} = h_j\left(x_j + \frac{A_{g0j}}{2l_g}\right)\left\{1 + \eta_{hj}\left(x_j + \frac{A_{g0j}}{2l_g}\right)^2\right\} \tag{2}$$

for

$$x_j + A_{g0j}/2l_g \leqq 0 \qquad j = 1, 2,$$

where f_{hj} is the force required to produce the deformation to mass, m_j, during collision, h_j is the linear stiffness, and η_{hj} is a positive coefficient representing the nonlinearity of the contacting vocal cords. The resultant restoring force acting on m_j during closure is, therefore, the sum of f_{sj} and f_{hj}, that is, eq. (1) and eq. (2). This change in spring stiffness at closure is schematically illustrated in Fig. 2.

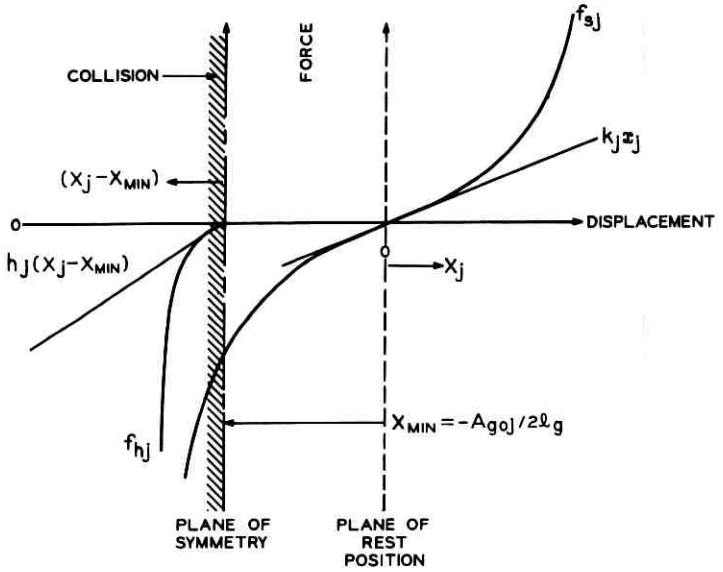


Fig. 2—Characteristics of the nonlinear stiffnesses.

3.2 *Nature of the Vocal-Cord Losses*

As in the earlier formulation,[1] the viscous loss of the vibrating cords is assumed piece-wise linear. The loss is caused to increase step-wise on closure of the cords to represent the "stickiness" of the soft, moist contacting surfaces as they form together.

It is convenient to express the equivalent viscous resistances in terms of damping ratios, ζ_1 and ζ_2, for the uncoupled oscillators, where

$$r_1 = 2\zeta_1\sqrt{m_1 k_1} \quad \text{and} \quad r_2 = 2\zeta_2\sqrt{m_2 k_2}, \tag{3}$$

and where, as shown in eq. (1), k_1 and k_2 are the linear components of stiffness for the springs s_1 and s_2. During the open-glottis condition, the loss is taken as $\zeta_1 = 0.1$ and $\zeta_2 = 0.6$ for a typical condition of the cord model. As in the earlier work, the loss during the closed-glottis condition is taken essentially as critical damping, namely

$$\zeta_1 = (1.0 + 0.1) \quad \text{and} \quad \zeta_2 = (1.0 + 0.6). \tag{4}$$

IV. PRESSURE DISTRIBUTION ALONG THE GLOTTIS

Because of the small dimensions of the glottis (compared to a wavelength at the frequencies of interest), and because of the high velocity

of the glottal flow (compared to the vocal-cord velocity), we can assume the glottal flow to be quasi-steady.[7] We shall use the Bernoulli equation for one-dimensional flow to obtain the pressure distribution along the glottal flow.

The abrupt contraction in cross-sectional area at the inlet to the glottis produces a *vena contracta* surrounded by stagnant air. The *vena contracta* makes the inlet area A_{g1} appear smaller than it actually is and the pressure drop greater than that dictated by an ideal area change. The loss factor for such a contraction has been studied in fluid flow experiments[9] and found to be on the order of 0.4 to 0.5. Flow measurements by van den Berg, et al.,[10] on plaster cast models of the larynx set the loss figure at 0.37. This latter figure is therefore taken to estimate the pressure drop at the inlet, and we fix this drop at

$$P_{B1}(1.00 + 0.37), \quad \text{or} \quad 0.69\rho(U_g^2/A_{g1}^2),$$

where $P_{B1} = \frac{1}{2}\rho u_{g1}^2$ is the Bernoulli pressure, ρ the air density, and u_{g1} the particle velocity at the lower cord-edge.

Within the constriction formed by the lower part of the cord, the pressure drop is assumed to be governed by viscous loss, which is also consistent with van den Berg's measurements. In this region the pressure falls linearly with distance according to a resistance to the volume flow equal to $12\mu d_1 l_g^2/A_{g1}^3$, where μ is the shear viscosity coefficient.

At the junction between the masses m_1 and m_2, the volume flow U_g is continuous, but the particle velocity changes. There is a corresponding abrupt change in pressure equal to the change in kinetic energy per unit volume of the fluid. This pressure change at the junction is

$$\begin{aligned} \Delta p &= 1/2\rho(u_{g1}^2 - u_{g2}^2) \\ &= 1/2\rho U_g^2(1/A_{g2}^2 - 1/A_{g1}^2). \end{aligned} \tag{5}$$

Throughout the constriction formed by the upper cord-edge, m_2, viscous loss is assumed to govern the pressure drop and, like the lower cord portion, the resistance is taken as $(12\mu d_2 l_g^2/A_{g2}^3)$.

At the abrupt expansion of the glottal outlet, the pressure recovers toward the atmospheric value (assuming no constrictions in the relatively large vocal tract). Van den Berg, in his model flow measurements, found the pressure recovery to be about 0.5 P_B. However, for small constrictions this measurement is difficult and uncertain. It seems preferable to base an estimate of the pressure recovery on momentum considerations, which hold in the theory of fluid flow.

Consider at the sudden expansion the relations for Newton's law,

$f = (d/dt)(mv)$. Then, because U_g is continuous,

$$\rho U_g(u_1 - u_{g2}) = A_1(P_{22} - P_1)$$

or

$$\begin{aligned}(P_1 - P_{22}) &= 1/2\rho u_{g2}^2[2N(1 - N)] \\ &= 1/2\rho \frac{U_g^2}{A_{g2}^2}[2N(1 - N)] \\ &= P_{B2}[2N(1 - N)], \end{aligned} \tag{6}$$

where $N = A_{g2}/A_1$, $P_{B2} = 1/2\rho u_{g2}^2$, and A_1 is the input area to the vocal tract. The value of $2N(1 - N)$ is typically in the order of 0.05 to 0.40, which is somewhat smaller than van den Berg's value. This difference is significant to the acoustic interaction between the vocal tract and the cord source.[1] The pressure distribution along the steady flow through the glottis is indicated in Fig. 3.

In the time-varying condition of the cords, the inertance of the air masses involved should also be taken into account. When combined with the loss terms just discussed, the pressure distribution along the glottis is described by

$$\begin{aligned}P_s - P_{11} &= 1.37 \frac{\rho}{2}\left(\frac{U_g}{A_{g1}}\right)^2 + \int_0^{l_c} \frac{\rho}{A_c(x)}\, dx \cdot \frac{dU_g}{dt} \,* \\ P_{11} - P_{12} &= 12 \frac{\mu l_g^2 d_1}{A_{g1}^3} \cdot U_g + \frac{\rho d_1}{A_{g1}} \cdot \frac{dU_g}{dt} \\ P_{12} - P_{21} &= \frac{\rho}{2} U_g^2\left(\frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2}\right) \\ P_{21} - P_{22} &= 12 \frac{\mu l_g^2 d_2}{A_{g2}^3} U_g + \frac{\rho d_2}{A_{g2}} \cdot \frac{dU_g}{dt} \\ P_{22} - P_1 &= -\frac{\rho}{2}\left(\frac{U_g}{A_{g2}}\right)^2 \cdot 2 \frac{A_{g2}}{A_1}\left(1 - \frac{A_{g2}}{A_1}\right). \end{aligned} \tag{7}$$

V. EQUIVALENT CIRCUIT FOR THE GLOTTIS

On the basis of the pressure difference relations of eq. (7), the acoustic impedance elements of the glottal orifice constitute the equivalent circuit shown in Fig. 4, where the U_g current is continuous. The elements

* The $(U_g(dL/dt))$ term in $(d/dt)(LU_g)$ is negligible, where $L = (\rho d/A)$.

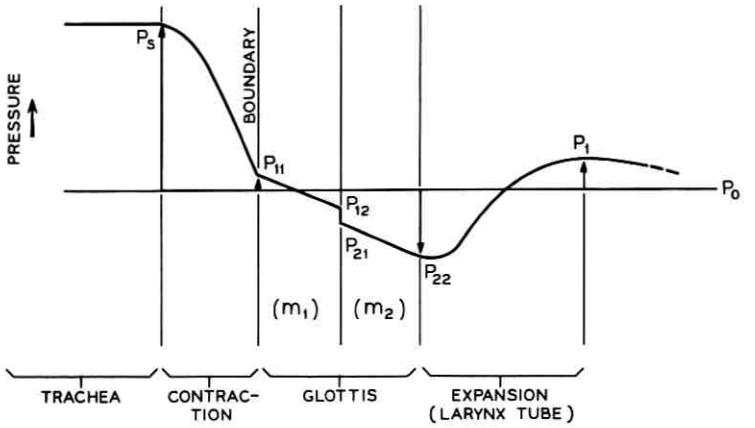


Fig. 3—Pressure distribution along the glottal flow.

of the acoustic circuit are given by

$$R_c = 1.37 \frac{\rho}{2} \frac{|U_g|}{A_{g1}^2}, \qquad L_c = \int_0^{l_c} \frac{dx}{A_c(x)}$$

$$R_{v1} = 12 \frac{\mu l_g^2 d_1}{A_{g1}^3}, \qquad L_{g1} = \frac{\rho d_1}{A_{g1}}$$

$$R_{12} = \frac{\rho}{2}\left(\frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2}\right)|U_g|$$

$$R_{v2} = 12 \frac{\mu l_g^2 d_2}{A_{g2}^3}, \qquad L_{g2} = \frac{\rho d_2}{A_{g2}}$$

$$R_e = -\frac{\rho}{2} \cdot \frac{2}{A_{g2} A_1}\left(1 - \frac{A_{g2}}{A_1}\right)|U_g|. \tag{8}$$

The total acoustic impedance of the glottis, Z_g, is therefore

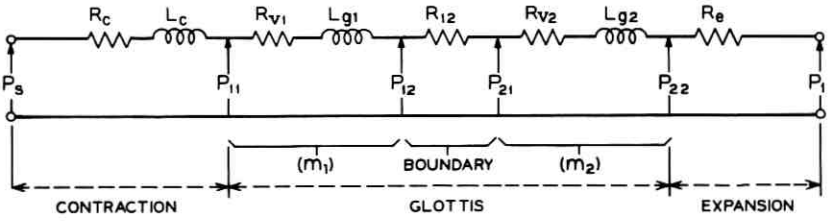


Fig. 4—Equivalent circuit for the glottis.

$$Z_g = \frac{\rho}{2} |U_g| \left\{ \frac{0.37}{A_{g1}^2} + \frac{1 - 2\frac{A_{g2}}{A_1}\left(1 - \frac{A_{g2}}{A_1}\right)}{A_{g2}^2} \right\} + (R_{v1} + R_{v2}) + j\omega(L_{g1} + L_{g2} + L_c) \quad (9)$$

or

$$Z_g = (R_{k1} + R_{k2}) \mid U_g \mid + (R_{v1} + R_{v2}) + j\omega(L_{g1} + K_{g2} + L_c),$$

$$Z_g = (R_{k1} + R_{k2}) \mid U_g \mid + (R_{v1} + R_{v2}) + j\omega(L_{g1} + L_{g2} + L_c), \quad (10)$$

where

$$R_{k1} = \frac{0.19\rho}{A_{g1}^2}, \qquad R_{k2} = \frac{\rho\left[0.5 - \frac{A_{g2}}{A_1}\left(1 - \frac{A_{g2}}{A_1}\right)\right]}{A_{g2}^2}.$$

In general, L_c can be neglected in comparison to $(L_{g1} + L_{g2})$.

The glottal impedance relation of eq. (10) can be linked to that obtained in flow measurements by van den Berg et al.[10] Using the pressure recovery found by van den Berg for the glottal outlet, namely $1/2\ P_{B2}$, [instead of the momentum relations in eq. (6)] gives a value $R_e = -(\rho/4) \mid U_g \mid /A_{g2}^2$. For the case of $A_{g1} = A_{g2} = A_g$, the total glottal impedance becomes

$$z_g = -0.87\frac{\rho}{2}\frac{|U_g|}{A_g^2} + 12\frac{\mu l_g^2 d}{A_g^3} + j\omega L_g . \quad (11)$$

The real part of this impedance is identical with that given by van den Berg.

VI. MODEL SYSTEM FOR VOICED SOUNDS

A network representation of the vocal system for voiced sounds is shown in Fig. 5. Beginning at the left, the subglottal system–comprised

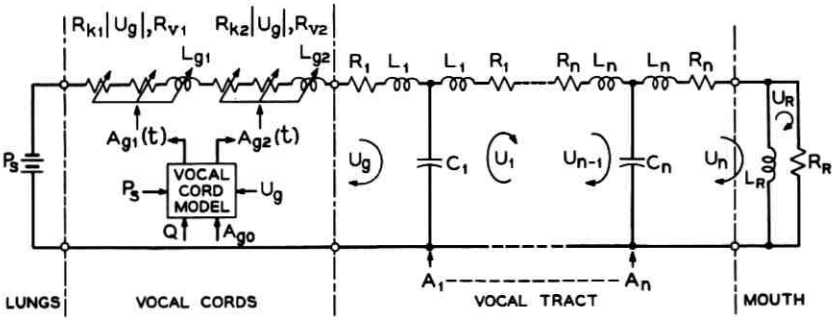


Fig. 5—Network model for the synthesis of voiced sounds.

of the trachea, bronchi, and lungs–is neglected and, as in the earlier study,[1] the subglottal pressure is approximated by a constant excess pressure in the lungs. Neglecting the subglottal system is also based on the finding that its first resonance is relatively high, with a mean value of 650 Hz and a bandwidth of 250 Hz. These values were determined from direct measurements of the subglottal driving-point impedance made on five laryngectomized subjects.[11] The 650 Hz figure is considerably higher than the value of 300 Hz reported by van den Berg.

The vocal tract is represented in Fig. 5 as a transmission line of n cylindrical, hard-walled sections, the element values of which are determined by the cross-sectional areas $A_1 \cdots A_n$, and the cylinder lengths $l_1 \cdots l_n$.[12] The inductances are $L_n = \rho l_n/2A_n$ and the capacitances are $C_n = (l_n \cdot A_n/\rho c^2)$, where c is the sound velocity. In the present work $n = 4$.

To account in part for tract losses, serial resistances R_n are taken to have the form of a viscous loss at the pipe wall, namely $R_n = (S_n/A_n^2)\sqrt{\rho\mu\omega/2}$, where S_n is the circumference of the nth section and ω is the radian frequency. The frequency for evaluation of this loss is fixed at the natural frequency of the lower oscillator, $f = (1/2\pi)\sqrt{k_1/m_1}$, and a multiplicative coefficient (ATT) is applied to increase the loss beyond that contributed by viscous loss at the walls and to produce formant bandwidths appropriate to a closed-glottis condition.* (The typical range for ATT is 20 to 25.)

The transmission line is terminated in a radiation load equal to that for a circular piston in an infinite baffle, namely $L_R = 8\rho/3\pi\sqrt{\pi A_n}$ and $R_R = (128\rho c/9\pi^2 A_n)$, where A_n is the final (mouth) area.[12]

From Fig. 5, the differential equations which relate the volume velocities of the system are:

$$(g\text{–loop}) \quad (R_{k1} + R_{k2})\,|U_g|\,U_g + (R_{v1} + R_{v2})U_g + (L_{g1} + L_{g2})\frac{dU_g}{dt} + L_1\frac{dU_g}{dt} + R_1 U_g + \frac{1}{C_1}\int_0^t (U_g - U_1)\,dt - P_s = 0$$

$$(1\text{–loop}) \quad (L_1 + L_2)\frac{dU_1}{dt} + (R_1 + R_2)U_1 + \frac{1}{C_2}\int_0^t (U_1 - U_2)\,dt + \frac{1}{C_1}\int_0^t (U_1 - U_g)\,dt = 0$$

* Other vocal tract losses not included *per se* are those arising from non-rigid walls and from heat conduction losses at the wall. The former is quite significant in lower-formant damping. The latter is essentially negligible. See Ref. 12.

$$\text{(2-loop)}\quad (L_2 + L_3)\frac{dU_2}{dt} + (R_2 + R_3)U_2 + \frac{1}{C_3}\int_0^t (U_2 - U_3)\,dt + \frac{1}{C_2}\int_0^t (U_2 - U_1)\,dt = 0$$

$$\text{(3-loop)}\quad (L_3 + L_4)\frac{dU_3}{dt} + (R_3 + R_4)U_3 + \frac{1}{C_4}\int_0^t (U_3 - U_L)\,dt + \frac{1}{C_3}\int_0^t (U_3 - U_2)\,dt = 0$$

$$\text{(4-loop)}\quad (L_4 + L_R)\frac{d(U_L)}{dt} + R_4 U_L - L_R\frac{d(U_R)}{dt} + \frac{1}{C_4}\int_0^t (U_L - U_3)\,dt = 0$$

$$\text{(5-loop)}\quad L_R\frac{d}{dt}(U_R - U_L) + R_R \cdot U_R = 0. \tag{12}$$

VII. FORCING RELATIONS FOR THE VOCAL-CORD OSCILLATOR

The masses of the cord oscillator are driven by mean pressures acting on their exposed faces, namely,

$$P_{m1} = \tfrac{1}{2}(P_{11} + P_{12}) = P_s - 1.37\,\frac{\rho}{2}\left(\frac{U_g}{A_{g1}}\right)^2 - \frac{1}{2}\left(R_{v1}U_g + L_{g1}\frac{dU_g}{dt}\right)$$

and

$$P_{m2} = \tfrac{1}{2}(P_{21} + P_{22}) = P_{m1} - \tfrac{1}{2}\left\{(R_{v1} + R_{v2})U_g + (L_{g1} + L_{g2})\frac{dU_g}{dt}\right\} - \frac{\rho}{2}U_g^2\left(\frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2}\right). \tag{13}$$

The exposed areas are $l_g d_1$ and $l_g d_2$, respectively. A shape of the cords is assumed such that the forces F_1 and F_2 acting on m_1 and m_2 over their displacements x_1 and x_2 are:

x_1	x_2	$F_1/l_g\,d_1$	$F_2/l_g\,d_2$
$x_1 > x_{1\min}$	$x_2 > x_{2\min}$	P_{m1}	P_{m2}
$x_1 \leqq x_{1\min}$	$x_2 > x_{2\min}$	P_s	0
$x_1 > x_{1\min}$	$x_2 \leqq x_{2\min}$	P_s	P_s
$x_1 \leqq x_{1\min}$	$x_2 \leqq x_{2\min}$	P_s	0,

(14)

where $x_{1\,\min} = -(A_{g01}/2l_g)$, $x_{2\,\min} = -(A_{g02}/2l_g)$, and A_{g01}, A_{g02} are the "phonation neutral" values of the glottal area. The equations of motion for the two masses are therefore:

$$m_1 \frac{d^2x_1}{dt^2} + r_1 \frac{dx_1}{dt} + s_1(x_1) + k_c(x_1 - x_2) = F_1$$

$$m_2 \frac{d^2x_2}{dt^2} + r_2 \frac{dx_2}{dt} + s_2(x_2) + k_c(x_2 - x_1) = F_2 ,$$

where

$$A_{g1} = (A_{g01} + 2l_g x_1), \qquad A_{g2} = (A_{g02} + 2l_g x_2),$$

$$s_j(x_j) = k_j(x_j + \eta_{kj} \cdot x_j^3), \quad j = 1, 2, \quad \text{for} \quad x_j > -\frac{A_{g0j}}{2l_g} ,$$

and

$$s_j(x_j) = k_j(x_j + \eta_{kj} \cdot x_j^3) + h_j\left\{\left(x_j + \frac{A_{g0j}}{2l_g}\right) + \eta_{hj}\left(x_j + \frac{A_{g0j}}{2l_g}\right)^3\right\} ,$$

$$\text{for} \quad x_j \leqq -\frac{A_{g0j}}{2l_g} , \qquad (15)$$

and F_1 and F_2 are given by the force table of eq. (14). These equations are coupled to the flow equations through the fact that x_1 and x_2 determine A_{g1} and A_{g2}. Also note that the coupling between the masses, which are permitted only lateral motion, has been linearized to be proportional to $(x_2 - x_1)$. [A more detailed consideration of the elongation produced in the coupling spring by a displacement difference $(x_2 - x_1)$, and of the lateral component of restoring force, leads to modifying the coupling term to $2k_c(x_2 - x_1)^3/(d_1 + d_2)^2$.]

VIII. DIGITAL SIMULATION

The differential equations are approximated by difference equations in which

$$\frac{df(t)}{dt} \cong \frac{f(t_i) - f(t_{i-1})}{(t_i - t_{i-1})} = \frac{f_i - f_{i-1}}{T}$$

$$\int_0^t f(t)\, dt \cong (t_i - t_{i-1}) \sum_{j=0}^{i-1} f(t_j) = T \sum_{j=0}^{i-1} f_j . \qquad (16)$$

These discrete approximations applied to eqs. (12) and (15) yield:

$$(g\text{–loop}) \quad (R_{k1i} + R_{k2i})\,|U_{gi}|\,U_{gi} + (R_{v1i} + R_{v2i})U_{gi}$$
$$+ (L_{g1i} + L_{g2i})\frac{U_{gi} - U_{gi-1}}{T} + L_1\frac{U_{gi} - U_{gi-1}}{T}$$
$$+ R_1 \cdot U_{gi} + \frac{T}{C_1}\sum_{j=0}^{i-1}(U_{gj} - U_{1j}) - P_s = 0$$

$$(1\text{–loop}) \quad \left\{\frac{L_1 + L_2}{T} + (R_1 + R_2)\right\}U_{1i} - \frac{L_1 + L_2}{T}U_{1i-1}$$
$$+ \frac{T}{C_2}\sum_{j=0}^{i-1}(U_{1j} - U_{2j}) + \frac{T}{C_1}\sum_{j=0}^{i-1}(U_{1j} - U_{gj}) = 0$$

$$(2\text{–loop}) \quad \left\{\frac{L_2 + L_3}{T} + (R_2 + R_3)\right\}U_{2i} - \frac{L_2 + L_3}{T}U_{2i-1}$$
$$+ \frac{T}{C_3}\sum_{j=0}^{i-1}(U_{2j} - U_{3j}) + \frac{T}{C_2}\sum_{j=0}^{i-1}(U_{2j} - U_{1j}) = 0$$

$$(3\text{–loop}) \quad \left\{\frac{L_3 + L_4}{T} + (R_3 + R_4)\right\}U_{3i} - \frac{L_3 + L_4}{T}U_{3i-1}$$
$$+ \frac{T}{C_4}\sum_{j=0}^{i-1}(U_{3j} - U_{Lj}) + \frac{T}{C_3}\sum_{j=0}^{i-1}(U_{3j} - U_{2j}) = 0$$

$$(4\text{–loop}) \quad \left\{\frac{L_4 + L_R}{T} + R_4\right\}U_{Li} - \frac{L_4 + L_R}{T}U_{Li-1}$$
$$- \frac{L_R}{T}(U_{Ri} - U_{Ri-1}) + \frac{T}{C_4}\sum_{j=0}^{i-1}(U_{Lj} - U_{3j}) = 0$$

$$(5\text{–loop}) \quad \frac{L_R}{T}\left\{(U_{Ri} - U_{Li}) - (U_{Ri-1} - U_{Li-1})\right\} + R_{Ri}U_{Ri} = 0,$$

where

$$R_{k1i} = \frac{0.19\rho}{A_{g1i-1}^2}, \quad R_{k2i} = \left[\frac{0.5 - \dfrac{A_{g2i-1}}{A(1)}\left(1 - \dfrac{A_{g2i-1}}{A(1)}\right)}{A_{g2i-1}^2}\right]\rho,$$

$$L_{g1i} = \frac{\rho d_1}{A_{g1i-1}}, \quad L_{g2i} = \frac{\rho d_2}{A_{g2i-1}}, \quad R_{v1i} = 12\mu l_g^2\frac{d_1}{A_{g1i-1}^3},$$

$$R_{v2i} = 12 l_g^2\frac{d_2}{A_{g2i-1}^3}, \tag{17}$$

and

$$\frac{m_1}{T^2}(x_{1i} - 2x_{1i-1} + x_{1i-2}) + \frac{r_1}{T}(x_{1i} - x_{1i-1}) + s_1(x_{1i}) + k_c(x_{1i-1} - x_{2i-1}) = F_{1i}$$

$$\frac{m_2}{T^2}(x_{2i} - 2x_{2i-1} + x_{2i-2}) + \frac{r_2}{T}(x_{2i} - x_{2i-1}) + s_2(x_{2i}) + k_c(x_{2i-1} - x_{1i-1}) = F_{2i},$$

where

$$A_{g1i} = A_{g01} + 2l_g \cdot x_{1i},$$

$$A_{g2i} = A_{g02} + 2l_g \cdot x_{2i},$$

$$s_1(x_{1i}) = k_1 \cdot (x_{1i} + \eta_{k1} \cdot x_{1i-1}^3), \quad \text{for} \quad x_{1i} > -\frac{A_{g01}}{2l_g},$$

$$s_1(x_{1i}) = k_1 \cdot (x_{1i} + \eta_{k1} \cdot x_{1i-1}^3) + h_1 \cdot \left\{\left(x_{1i} + \frac{A_{g01}}{2l_g}\right) + \eta_{h1} \cdot \left(x_{1i-1} + \frac{A_{g01}}{2l_g}\right)^3\right\}, \quad \text{for} \quad x_{1i} \leqq -\frac{A_{g01}}{2l_g},$$

$$s_2(x_{2i}) = k_2 \cdot (x_{2i} + \eta_{k2} \cdot x_{2i-1}^3), \quad \text{for} \quad x_{2i} > -\frac{A_{g02}}{2l_g},$$

$$s_2(x_{2i}) = k_2 \cdot (x_{2i} + \eta_{k2} \cdot x_{2i-1}^3) + h_2 \cdot \left\{\left(x_{2i} + \frac{A_{g02}}{2l_g}\right) + \eta_{h2} \cdot \left(x_{2i-1} + \frac{A_{g02}}{2l_g}\right)^3\right\}, \quad \text{for} \quad x_{2i} \leqq -\frac{A_{g02}}{2l_g},$$

$$F_{1i}/d_1 l_g = P_{m1i} = P_s - 1.37\,\frac{\rho}{2}\left(\frac{U_{gi}}{A_{g1i-1}}\right)^2 - \frac{1}{2}\left\{R_{v1i}U_{gi} + \frac{L_{g1i}}{T}(U_{gi} - U_{gi-1})\right\},$$

$$F_{2i}/d_2 l_g = P_{m2i} = P_{m1i} - \left\{\frac{1}{2}(R_{v1i} + R_{v2i})U_{gi} + (L_{g1i} + L_{g2i})\,\frac{U_{gi} - U_{gi-1}}{T}\right\} - \frac{\rho}{2}\,U_{gi}^2\left(\frac{1}{A_{g2i-1}^2} - \frac{1}{A_{g1i-1}^2}\right). \tag{18}$$

These difference equations were programmed in Fortran IV and compiled for experiment on one of the DDP-516 laboratory computers

of the Acoustics Research Department at Bell Laboratories.[13] Simultaneous solution of eqs. (17) and (18) yields all relevant volume velocities, glottal areas, and displacement. The time derivative of the mouth volume velocity (i.e., through the radiation load) is a good approximation to the radiated sound pressure.[12] Time samples of all dependent variables are obtained by iterating the solution for as many samples as are desired.

The sampling interval T is chosen as the longest interval that yields a stable solution to the difference equations. This interval is determined primarily by the time required for sound to transit the shortest length of the vocal tube. Because the distributed vocal tract is approximated as lumped constant T-sections, and because the behavior of these elements is further approximated by finite differences, the sampling interval T must be considerably shorter than the sound transit time through the shortest tube element. In the absence of appropriate sampling theory for this situation, the broad range of stable solutions was determined interactively on the DDP-516 computer and the longest stable interval used. In the present work, sampling rates in the range of 10 kHz to 30 kHz were used.

The iteration "loop" of the equations can be closed owing to the manner in which the glottal impedance elements and the forcing functions are taken to involve samples of glottal area; for example, current values of impedance and forcing function involve only past values of glottal areas. The iteration, therefore, proceeds as follows.

The cords and tract are initially assumed at rest, and initial currents are zero. The first sample of U_{gi} is calculated from loop-g using $A_{gi-1} = A_{g0}$ (i.e., $x_{i-1} = 0$). The initial samples of all other loop currents are likewise calculated, out to the radiation load. The first sample of U_{gi} is then used to calculate the first samples of the forcing functions and, from the mechanical equations, the first samples of the displacements x_{1i} and x_{2i}. The latter dictate new values of A_{g1} and A_{g2} which are entered back into the glottal impedance elements for the calculation of the next sample of U_g and all other currents. The process is continued until as much of the solution as desired is obtained.

For synthesis of continuous speech, the vocal-tract area values change as do the values of P_s, A_{g0}, and cord constants.* These changes are slow by comparison to the sample variations in volume velocities, displacements, and pressure. The solutions for the continuously changing vocal system can therefore be considered as quasi-steady solutions of

* As indicated in Fig. 5, a cord-tension parameter, Q, constitutes an input to the vocal-cord model. This parameter determines the mechanical constants of the oscillator.

eqs. (17) and (18), and the mouth output samples taken as the synthetic speech signal.

IX. PHYSIOLOGICAL CONSTANTS FOR THE VOCAL-CORD MODEL

Few numerical values are available for the physiological parameters of the vocal cords. Using the sparse data available, the simulation on the DDP-516 computer was used to establish relevant ranges for the parameters.

First, the range of parameters which allows self-oscillation of the model was studied for a uniform vocal tract, 16 cm long, 5 cm^2 in cross-section, and terminated in the radiation load. The DDP-516 computer was used interactively to establish the self-oscillation region. The allowed oscillation range as a function of k_2 and k_c is shown in Fig. 6. In this plot, the axes are normalized by the factor m_1/m_2k_1. The parameters in the figure are the damping coefficients of the mechanical oscillators, ζ_1 and ζ_2. For these cases, all other glottal parameters

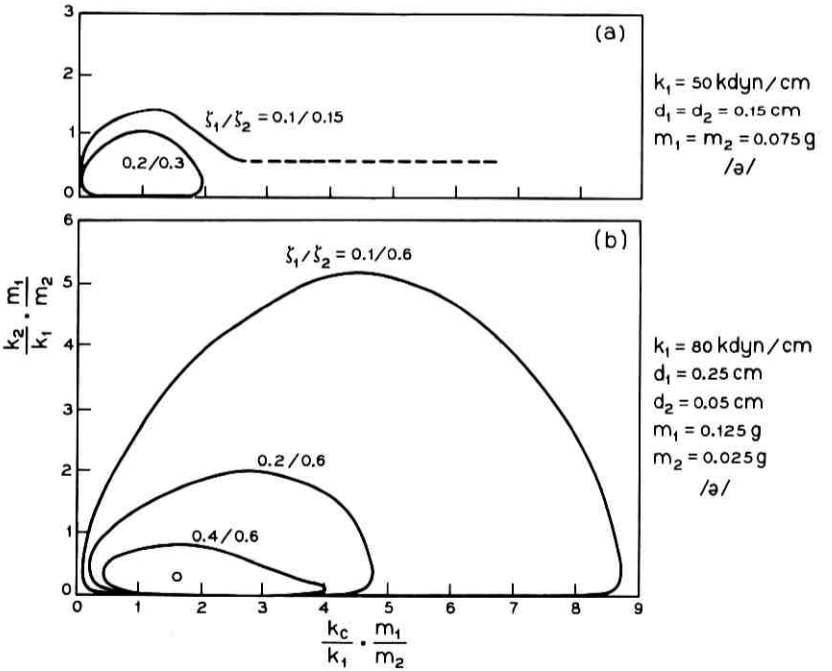


Fig. 6—Allowed regions of oscillation for the two-mass model. The parameter is the open-glottis damping ratio. The vocal-tract shape is for the vowel /ə/.

are held constant at physiologically realistic values: namely, $P_s = 8$ cm H_2O, $l_g = 1.4$ cm, $A_{g01} = A_{g02} = 0.05$ cm^2, thickness of the vocal cords $d_1 + d_2 = 0.3$ cm, total mass $m_1 + m_2 = 0.15$ g, nonlinear coefficient of the springs $\eta_{k1} = \eta_{k2} = 100$ and $\eta_{h1} = \eta_{h2} = 500$, and $h_1 = 3k_1$, $h_2 = 3k_2$. In particular, values for the spring constants are based upon measurements of static tensile stress versus displacement for an excised human larynx.[8] From these measurements, for example, η_k is deduced to be on the order of 50 to 100.

For Fig. 6a, the vocal cords are divided into equal parts, with thickness and mass 0.15 cm and 0.075 g, respectively, and with $k_1 = 50$ kdyn/cm. For Fig. 6b, the lower part of the vocal-cord model is thicker than the upper part, that is, $d_1 = 0.25$ cm and $d_2 = 0.05$ cm, and the masses, $m_1 = 0.125$ g and $m_2 = 0.025$ g, are chosen proportional to the thicknesses, keeping the same total mass of 0.15 g as in Fig. 6a.

Kaneko[11] has measured the damped oscillations of a fresh excised human larynx when excited by a mechanical impulse and with no air flow through the glottis. From this data, the damping ratio for the excised human cords can be estimated to be of the order of 0.1 to 0.2 (which, incidentally, is the same order as deduced in the earlier simulations[1]). This range of damping seems particularly appropriate for the bulk of the cords, that is, for m_1 of the model.*

An acoustic load of the vocal tract, whose driving-point impedance has an inductive reactance at the fundamental frequency of the vocal-cord vibration, acts to enhance oscillation of the model. An increase in damping (loss) of the vocal tract at lower frequencies, as would be caused by wall vibration in the vocal tract, however, acts to oppose oscillation. Also, the tendency to oscillate is suppressed by an increase in the mechanical damping of m_1 and especially of m_2.

The behavior of the vocal-cord model, calculated for values of k_2 and k_c specified by the small circle in Fig. 6b, will now be considered. This glottal condition is chosen as the "typical" one throughout the present study; namely, $k_1 = 80$ kdyn/cm, $k_2 = 8$ dyn/cm, and $k_c = 25$ kdyn/cm.

* An equivalent damping ratio for the bulk of the cords can be estimated as follows:

$$(r_1 + r_2) = 2\,\zeta_1\sqrt{m_1k_1} + 2\,\zeta_2\sqrt{m_2k_2}.$$

For $k_c \to \infty$,

$$(r_1 + r_2) = 2\,\zeta_{\text{equi}}\sqrt{(m_1 + m_2)(k_1 + k_2)}.$$

Substituting (for the "typical" conditions) $m_2 = m_1/5$, $k_2 = k_1/10$, $\zeta_1 = 0.1$, and $\zeta_2 = 0.6$ gives

$$\zeta_{\text{equi}} = \frac{1}{\sqrt{66}}\,(\sqrt{50}\,\zeta_1 + \zeta_2) = 0.16,$$

which corresponds favorably with Kaneko's measurements.

X. RESULTS OF THE DIGITAL SIMULATION

The vocal-cord and vocal-tract program, specified by eqs. (17) and (18), was used interactively on the DDP-516 computer to calculate waveforms of glottal flow, glottal area, and mouth sound pressure.

10.1 *Waveforms for Typical Glottal Conditions*

Measurements made at the typical glottal condition and for a uniform vocal tract are illustrated in Fig. 7. Waveforms of A_{g1}, A_{g2}, U_g, and mouth sound pressure are shown for the initial 30 ms of voicing. The negative values of A_{g1} and A_{g2} indicate glottal closure. (One can imagine the cords forming into one another upon contact, and the negative areas correspond to the continued displacement of the center of mass of the cords.)

The results show that the phase difference between m_1 and m_2 is about 55 degrees, and the duty ratio (glottis open time to total period) is about 0.6. These values compare well with observations which have been made on human vocal cords by high-speed motion picture techniques[14] and by inverse filtering.[15] One notices the differences between the glottal area wave and the corresponding glottal flow wave, as pointed out in the earlier work.[1] The glottal flow wave is typically characterized by some temporal detail, asymmetry, and steep falling slope, while the area wave shows little temporal detail, is less steep,

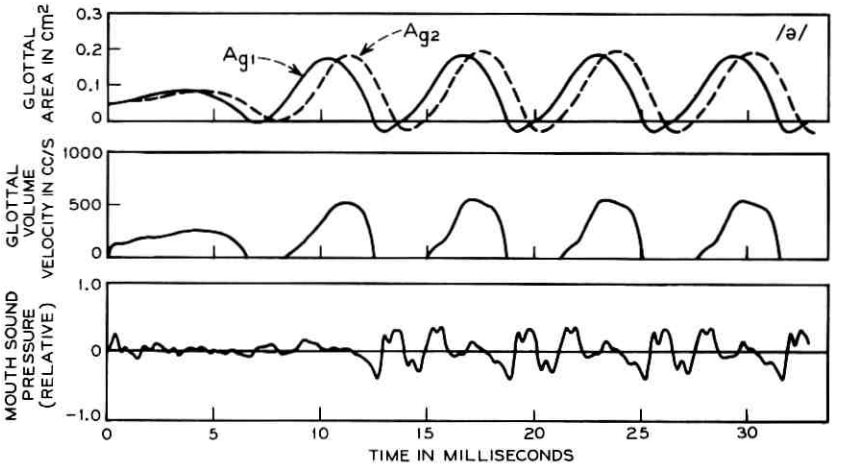


Fig. 7—Vocal-cord and vocal-tract functions computed from the DDP-516 simulation. Glottal areas, A_{g1} and A_{g2}, glottal volume velocity, U_g, and mouth-output sound pressure are shown for the initial 30 ms of voicing for the neutral vowel /ə/.

and tends to be more symmetrical. Because the cords are massive and are generally forced at a frequency above their natural frequency, their mechanical displacement does not reflect the detail of acoustic interaction which the glottal flow displays. The sound output wave reflects the periodicity established by the cord oscillator, and the greatest formant oscillation (or excitation) typically occurs (with about 0.5 ms delay) at the closing phase of the U_g wave. This effect has been observed previously.[2]

10.2 *Effect of Cord Stiffnesses*

The normalized k_2 versus k_c plane of Fig. 6 is a convenient medium for demonstrating the effects of spring constants. Using this plane, waveforms of U_g, A_{g1}, and A_{g2} are sketched for corresponding stiffnesses in Fig. 8. As before, the vocal tract is a uniform pipe (/ə/) and other glottal conditions are kept at their typical values.

An increase in k_c above the typical value reduces the phase difference between A_{g1} and A_{g2}. It also diminishes the steep falling slope of the flow waveform, and tends to make the wave more symmetrical and triangular. An increase in k_c also produces an increase in the build-up time required for the oscillation to settle to a steady state. For still larger values of k_c, close to the bounds of the oscillation range, both

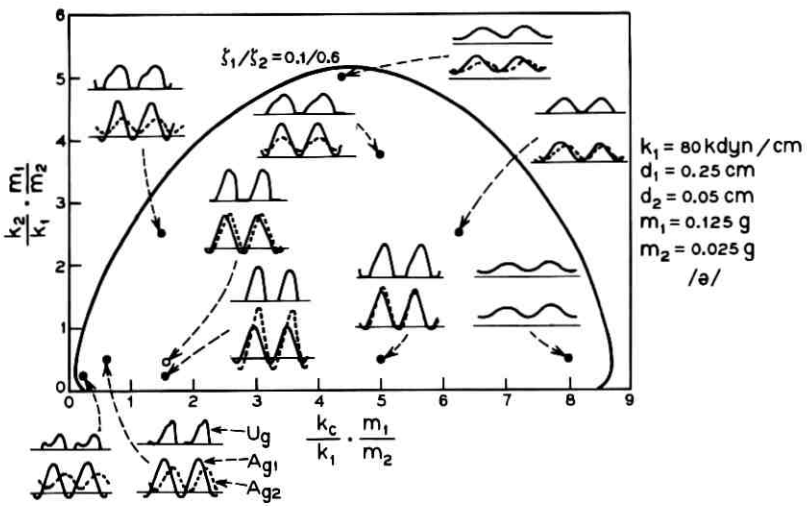


Fig. 8—Sketches of cord-tract functions for points on the k_2-k_c plane. The axes are normalized by the function (m_1/k_1m_2). The vowel is /ə/. Compare with Fig. 6b.

the glottal flow and area waveforms become sinusoidal on a dc component, and the glottis does not close.

The range of the sinusoidal behavior is expanded if the damping coefficients are made smaller. This special case is shown in Fig. 6a for the damping coefficients $\eta_{k1} = 0.1$ and $\eta_{k2} = 0.15$. Here, k_c has no limitation for the oscillation when k_2 is less than 20 kdyn/cm. Owing to the large k_c, the two-mass model behaves just as the one-mass model in the extended region in which the oscillation is sustained by the inductive reactance of the vocal tract and glottis. This projecting tail disappears with an increase in the losses, either of the vocal tract or of the vocal cords.

In contrast, an increase in k_2, with other conditions constant, decreases the amplitude of A_{g2} without a change of the phase difference. Further increase of k_2 leads to no closure of A_{g2} while A_{g1} can close completely during the cycle. Owing to the small amplitude of A_{g2} and its dc component, the glottal flow increases in upward roundness and also increases in duty ratio. A small, broad hump appears on the rising slope of the glottal flow wave, at which point the area A_{g1} is equal to A_{g2}.

By comparison, a decrease in k_2 increases the amplitude of A_{g2} and the glottal waves tend to a symmetrical form. This same dependence on k_2 and k_c also occurs for the case of equal thicknesses, $d_1 = d_2 = 0.15$ cm. A change in proportion of the damping coefficients, ζ_1 and ζ_2, also influences the relations between A_{g1} and A_{g2}. For example, the typical condition $\zeta_1 = 0.1$ and $\zeta_2 = 0.6$ produces an amplitude of A_{g1} slightly larger than that of A_{g2} for /ə/, as seen in Fig. 7. A smaller value of ζ_2 for the same values of ζ_1 and other parameters produces an amplitude of A_{g2} larger than A_{g1} without a change in phase difference. A steeper rising slope of the glottal area wave also results, but the falling slope remains unchanged.

10.3 *Effect of Neutral Area*

The behavior of the vocal-cord model with respect to the "phonation-neutral" area, or the equilibrium value A_{g0}, is another case where we can find correspondence between the complex behavior of the human vocal cords and the vibrations of the vocal-cord model. In human phonation the neutral area is maintained by laryngeal adjustment. Typical results from the simulation for different values of A_{g0} are illustrated in Fig. 9. These data were measured for the typical glottal conditions with $\zeta_1 = 0.1$ and $\zeta_2 = 0.6$ and for the vowel /i/. One sees that the build-up time required for the oscillation to reach a steady state increases as A_{g0} gets larger. The value $A_{g0} = 0.30$ cm^2 surpasses

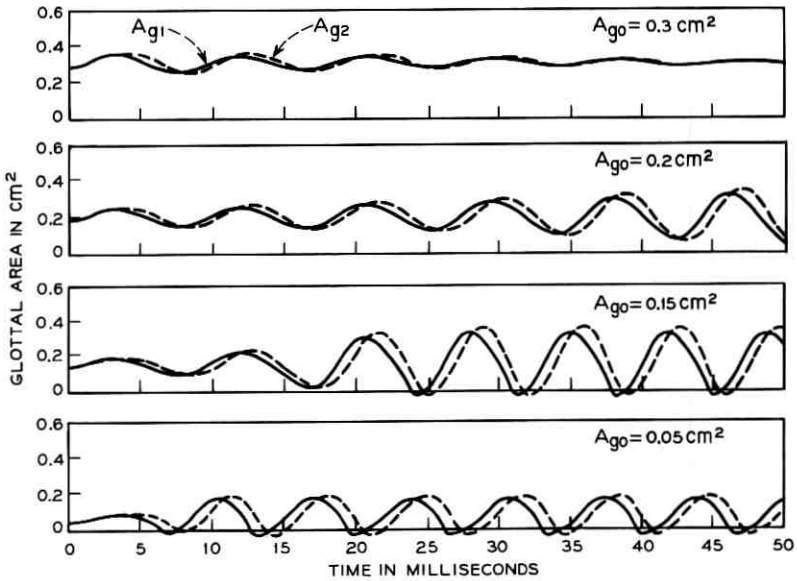


Fig. 9—Effect of the "phonation-neutral" or rest area, A_{g0}, upon the glottal area. The vowel is /i/.

a critical limit (about 0.25 cm^2) beyond which the model does not oscillate for these conditions.

During the voicing build-up time the pitch period is much longer than that of the steady-state oscillation. The change in pitch at the onset of voicing is similar to the starting motion of the human cords when they are brought to the phonation position from an open position. In this instance, unestablished low subglottal pressure also contributes to the reduction of the fundamental frequency. The oscillation period before cord closure is a value between the damped natural frequencies of the two mechanical oscillators. This is consistent with the value calculated from the acoustic theory of the two-mass model neglecting the collision and the nonlinearity of the springs.

Although the model, in this case, does not self-oscillate for $A_{g0} > 0.25$ cm^2, the maximum glottal area for phonation depends on the damping of the mechanical oscillators and of the vocal tract and on the subglottal pressure. For $\zeta_1 = 0.2$ and $\zeta_2 = 0.6$, and with $P_s =$ 8 cm H_2O, the maximum glottal area reduces to about 0.20 cm^2. An increase in the phonation-neutral area also causes an increase in the amplitude of the vibration with no significant change in the period of the steady-state oscillation.

10.4 *Effect of Tract Shape*

Excitation of the vocal tract by the cord model was studied for several vowels. Area waves, glottal flow, and mouth-output sound pressure are shown for the vowels /i/, /u/, and /a/ in Figs. 10a, b, and c. For all these cases, the typical glottal conditions hold (same as for /ə/ in Fig. 7).

One notices that the waveforms of glottal area and the fundamental frequency are almost independent of the vocal-tract shape, while the shape can substantially influence the waveform of the glottal flow, similar to the results obtained from the one-mass model in the earlier work.[1] The acoustic interaction between the glottal flow and the acoustic load depends on the resonance characteristics of the vocal tract. Vowels having high resonant Q for the first formant show noticeable interaction in the glottal flow wave, as is seen for /a/. Also a low first formant can affect the glottal flow wave to a considerable extent, for example in /i/. However, the relatively large dissipation of the vocal tract in the frequency range of low first formants (such as for /i/ and /u/) caused primarily by vibration of the vocal-tract walls acts to reduce the interaction, but the glottal flow waveforms still differ markedly from each other. In all these cases, the tract losses are set to give bandwidths for the first formant equal to values measured on the human tract for the closed-glottis condition.[16]

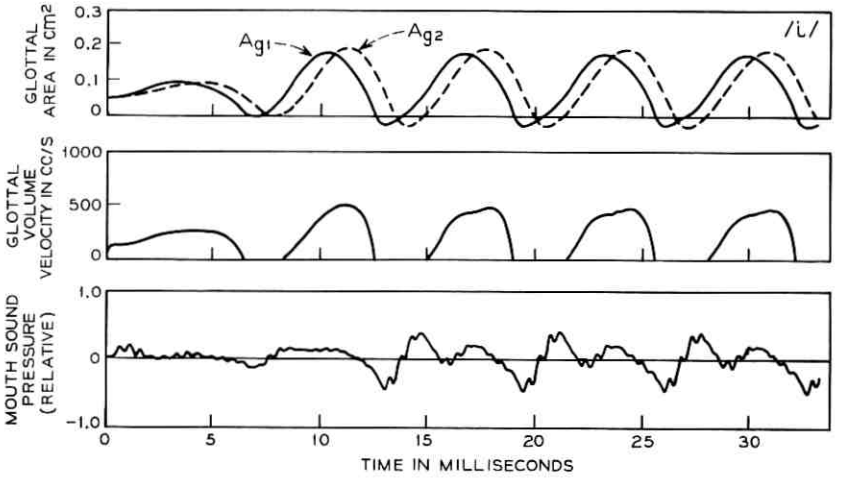


Fig. 10a—Results of the DDP-516 simulation for the vowel /i/ showing area waves, glottal flow, and mouth-output sound pressure.

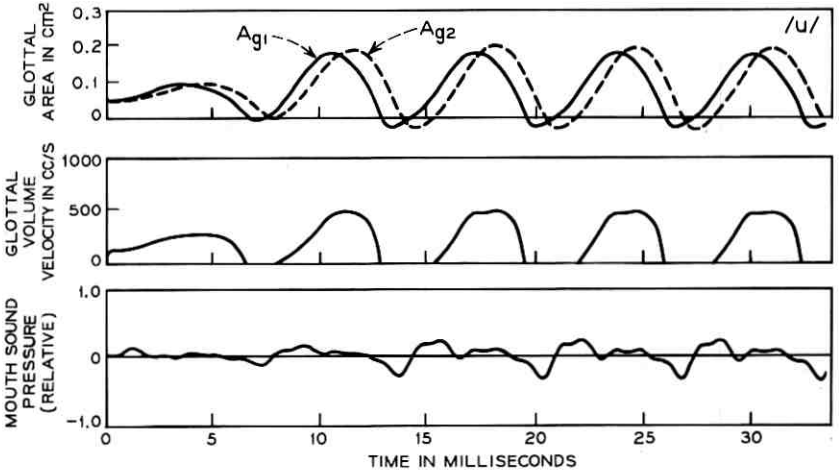


Fig. 10b—Same as Fig. 10a for the vowel /u/.

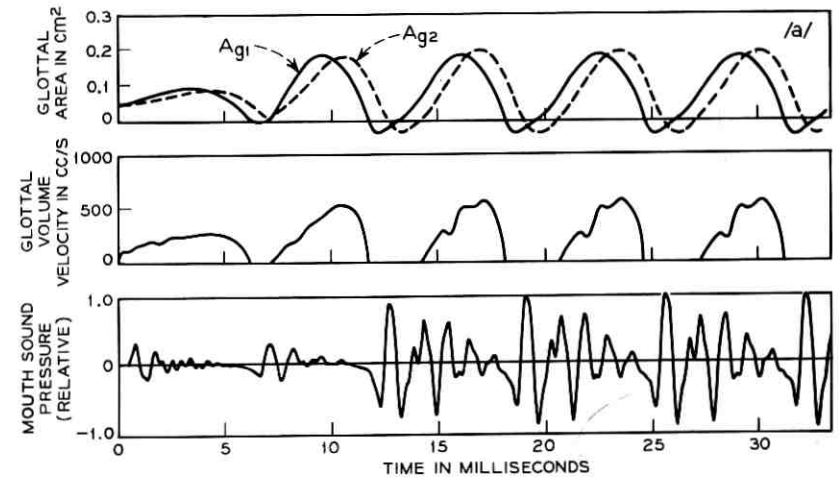


Fig. 10c—Same as Fig. 10a for the vowel /a/.

The data of Fig. 10* permit a comparison between the glottal waveform and the speech pressure wave. The comparison is familiar from the results of inverse filtering.[15,17] There is a delay time difference of about 0.5 ms between the waves, corresponding to the time required for the

* Sound spectrograms of the computed mouth-output sound pressure are shown for several vowels in Fig. 10d.

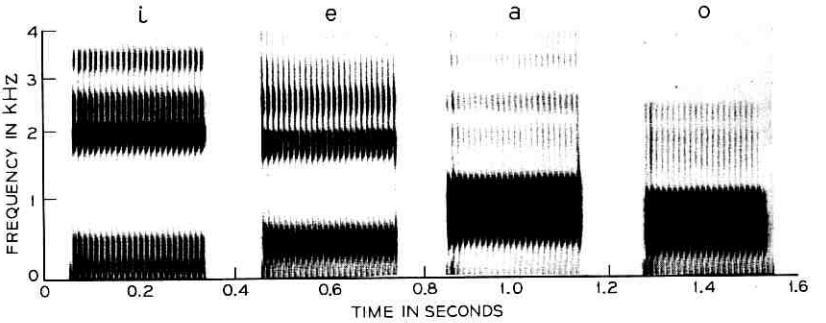


Fig. 10d—Sound spectrograms of the computed mouth-output sound pressure for the vowels /i, e, a, o/.

sound to travel from the glottis to the lips. The waveforms for /a/, /i/, and /u/ show that the formants are excited largely at the closure of the cords. The output pressure waves attenuate rapidly with increasing glottal area during the opening phase of the glottal cycle.

10.5 *Effect of Subglottal Pressure*

The influence of the subglottal pressure on the fundamental frequency of the vocal-cord vibration is another important aspect of voice production. Typical behavior of the model for these factors is shown in Fig. 11. The nonlinear coefficient of the spring, η_k, is shown as the parameter for the vowel /ə/. The data for the vowels /i/ and /a/ correspond to $\eta_k = 100$ solely. For all these cases, the coefficient describing the nonlinearity in the deformation of the vocal cords at closure is taken as $\eta_h = 5\eta_k$.

The slope of the fundamental frequency as a function of subglottal pressure is seen to be about 2.5 Hz/cm H_2O for $\eta_k = 100$, independent of the vowel configuration. This represents good agreement with measurements which have been made on human speech in the chest register by Hixon, et al.[18] The curve for $\eta_k = 0$, that is, linear springs, shows a saturation characteristic for subglottal pressures greater than 8 cm H_2O. These results suggest that pitch variations with subglottal pressure might be ascribed to two causes. One is the collision of the vocal cords at closure when the amplitude of vibration is not too large and the subglottal pressure is less than several cm H_2O. Another is the nonlinearity of the deflection of the muscles and ligaments at large amplitudes of vibration and at subglottal pressures more than several cm H_2O. In the latter case, the nonlinearity becomes dominant when large

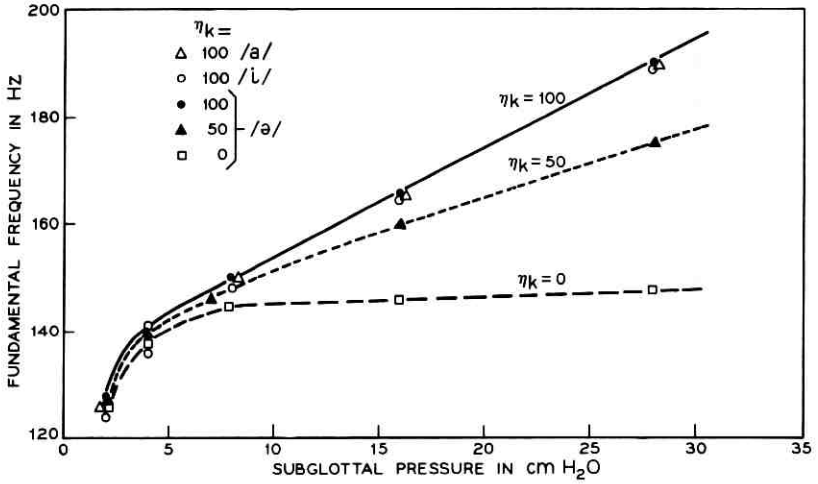


Fig. 11—Variation of fundamental frequency with subglottal pressure. The parameter is the nonlinear coefficient of the stiffnesses, η_k.

displacement amplitude increases the effective stiffness of the springs. This tends to increase fundamental frequency. The minimum subglottal pressure for vowel production is about 2 cm H_2O.

In the earlier work with the one-mass model, significant influences were found on fundamental frequency as a function of tract configuration. This influence was due in large part to the pressure recovery assumed at the glottal outlet, namely $1/2\ P_B$ according to van den Berg's data. When the intraglottal pressure distribution derived here is used in the one-mass model, the interaction across vowels and with subglottal pressure is much less.

The two-mass model becomes a one-mass model if k_c is increased to a large value. For this condition, the variation in fundamental frequency with subglottal pressure is shown for several vowels in Fig. 12. The behavior is similar to the two-mass model. Under these conditions, the duty ratio of the former tends to be slightly greater than the latter.

Duty ratio is another aspect of the model that can be compared to human phonation. An increase in subglottal pressure produces an increase in glottal flow and in glottal amplitude. Duty ratio (open time to total period) decreases for this increase in subglottal pressure and is asymptotic to a value around 0.5, as shown in Fig. 13. This value compares well with measurements on natural speech.[12]

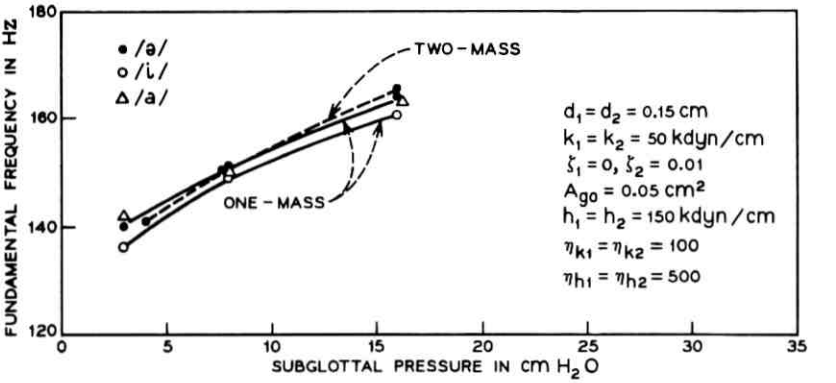


Fig. 12—Variations of fundamental frequency with subglottal pressure for the one-mass and two-mass models. The parameter is vowel configuration.

10.6 *Effect of Cord Tension*

As in the previous studies,[1-3] it is convenient to apply a "tension parameter," Q, to control fundamental frequency. This can be achieved by causing the masses and thicknesses to be scaled down and the springs scaled up by the factor Q, causing the fundamental frequency to vary proportionally with Q. Phase difference, duty ratio, and glottal area waveforms are essentially uninfluenced by Q, and the amplitudes of glottal area and glottal flow decrease gradually with increasing Q. The glottal flow waveform also varies in detail depending on the fundamental frequency, because the formants contributing to the temporal detail of the glottal

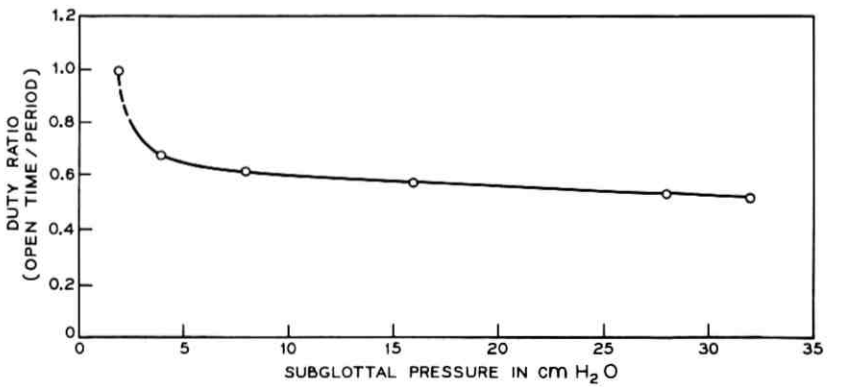


Fig. 13—Variation of duty ratio with subglottal pressure.

flow varies as a function of Q. Changes in flow waveform with pitch variation are greatest in cases where acoustic interaction is especially pronounced (such as for /a/).

In human speech the duty ratio has a tendency to increase with the fundamental frequency.[19] This feature can also be given to the vocal-cord model by modifying the coupling-tension parameter k_c to increase more than in linear proportion to Q. A variation as Q^2 appears more realistic. Physiologically this corresponds to the considerable decrease in compliance and thickness of the vocal cords when they are stretched by contraction of the cricothyroid muscle and other muscles associated with contracting of the vocalis. The increase of k_c more than proportional to Q is equivalent to shifting the glottal operation condition on a line parallel to the abscissa in Fig. 6. As indicated in Fig. 8, a shift to the right reduces the phase difference and increases the duty ratio without changing other features of the cord vibration, except near the boundaries of the oscillation range.

Behavior of the cord model with the Q parameter so defined is shown in Figs. 14 and 15. Variations in waveforms with Q are shown for the vowel /a/ in Fig. 14. The relations between fundamental frequency, duty ratio, and amplitude of glottal area with Q are plotted in Fig. 15. Variation of the duty ratio with frequency falls into the range measured in inverse filtering experiments.[19]

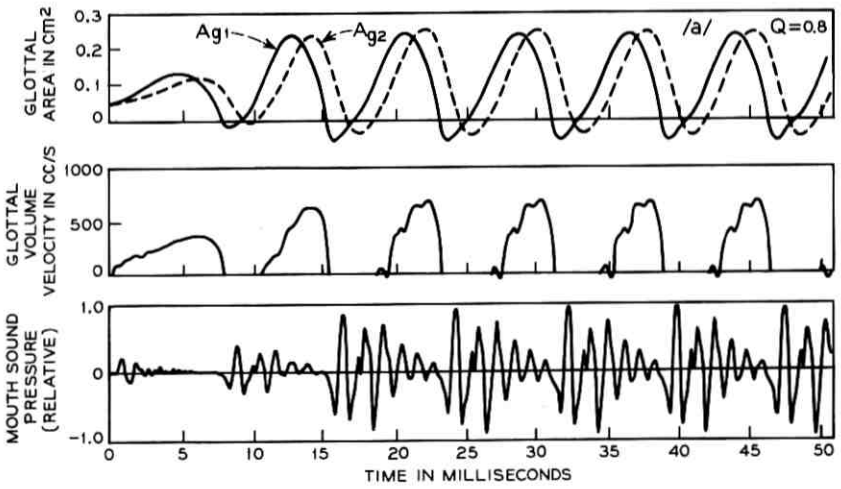


Fig. 14a—Effect of tension parameter, Q, on cord-tract output for the vowel /a/. $Q = 0.8$.

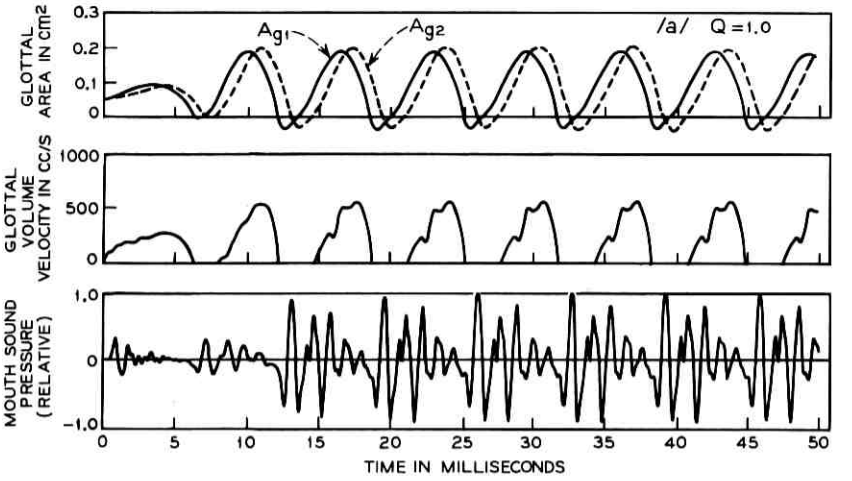


Fig. 14b—Same as Fig. 14a with $Q = 1.0$.

XI. INTERACTION EFFECTS WITH LARGE ACOUSTIC LOADS

11.1 *Differences Between Two-Mass and One-Mass Models*

The measurements discussed previously show that the fundamental frequency and the area waveforms of the cord model are not strongly influenced by tract geometry. The interaction with glottal flow, however,

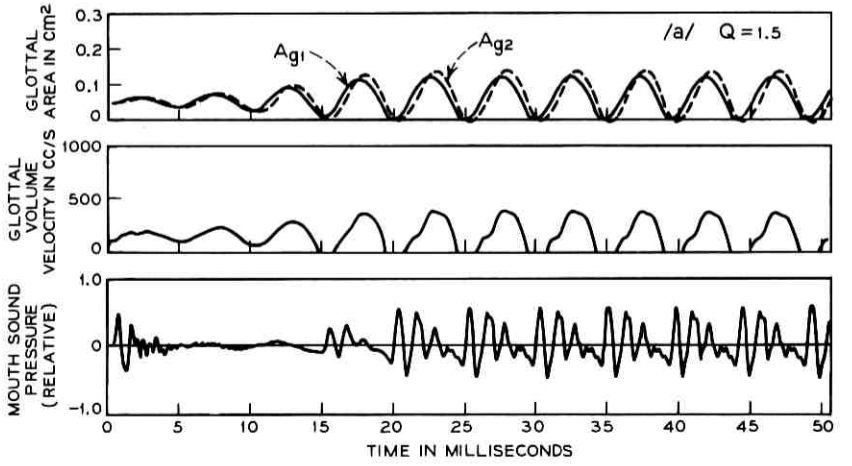


Fig. 14c—Same as Fig. 14a with $Q = 1.5$.

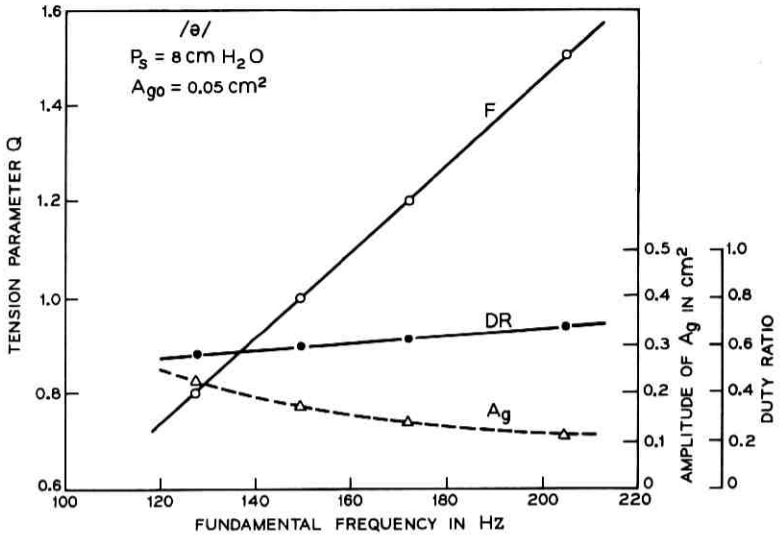


Fig. 15—Effect of tension parameter (Q) on fundamental frequency (F), duty ratio (DR), and glottal area (A_g).

is marked. We have further investigated the effect of acoustic load by lowering the frequency of the lowest resonance of the acoustic load (the first formant) into the range of the fundamental frequency. This increases the driving-point impedance at the fundamental frequency and strong coupling between source and load is expected.

The formant frequencies are lowered by lengthening the simulated vocal tract. Measurements of the fundamental frequency are shown in Fig. 16 as a function of the length of a uniform vocal-tract tube, 5 cm^2 in cross-section. Data are shown for both the two-mass cord model and an equivalent one-mass model ($k_c \rightarrow \infty$). The measurements are for the typical glottal conditions. The shunt impedance of the vocal-tract wall (wall vibration) is not taken into account *per se*, and this effect is only approximated by an increase in damping for the first 16-cm section of the tube (as was used for the /ə/ configuration). The remaining tube is regarded as an ideal hard-wall tube. The first resonance frequency of the vocal-tract tube, F_{01}, is shown by the solid line.

The frequency of the two-mass model decreases more gradually than that of the one-mass model with increasing the tube length. When the oscillation frequency of the former meets the first formant frequency of the vocal-tract tube, a sharp increase of the fundamental frequency occurs for further increase in tube length. The frequency returns to

almost the same value as for a short tube. The frequency jump occurs at the resonant frequency of the vocal-tract tube, independent of dissipation and of glottal conditions. For example, an increase in acoustic dissipation of the vocal tract and a decrease in mechanical damping of m_1 and m_2 raises the onset frequency of the jump, but the frequency where the jump occurs is still the first resonant frequency of the tube. The variation of frequency with vocal-tube length is shown for two conditions of damping in Fig. 16.

The curve of F_{01} as a function of tube length marks the boundary between an inductive driving-point impedance (to the left) and a capacitive driving-point impedance (to the right). The frequency jump for the two-mass model, which occurs at F_{01} regardless of the glottal conditions, places its new oscillation in the capacitive region, that is, between the first pole and second zero of the driving-point impedance.

A frequency jump also occurs in the one-mass model. In this case, however, the jump is to the original frequency for which the driving-point impedance is still an inductive impedance, that is, between the

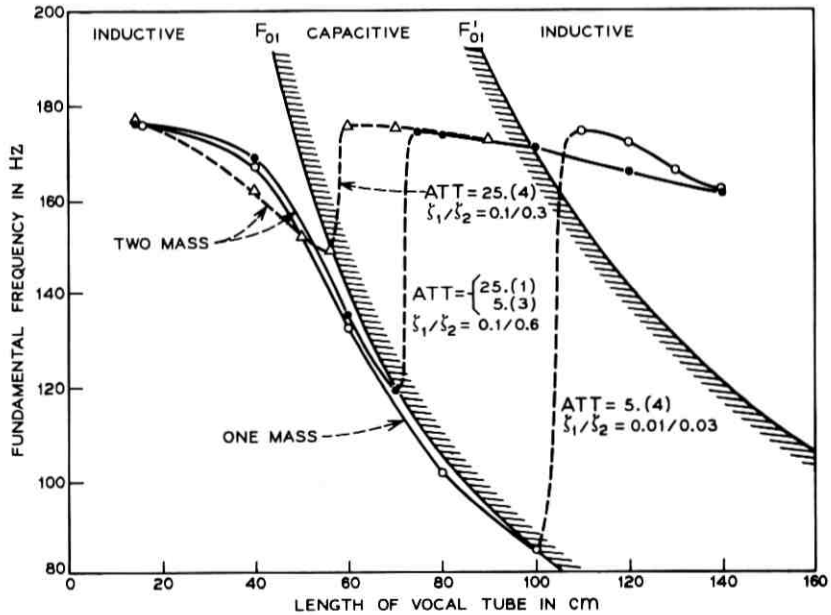


Fig. 16—Variation of fundamental frequency with acoustic load for the two-mass and one-mass models. F_{01} shows the frequency of the first pole of the driving-point impedance, and F_{01}' shows the first zero.

second zero and second pole of the driving-point impedance. This behavior can be predicted by an analysis of the oscillator with a uniform transmission line as a load.

11.2 *Effects of Acoustic Load on Human Voicing*

For comparison with the model behavior, we have measured similar loading effects in human voicing. To bring a first formant resonance into the range of the voice pitch, subjects phonated into a long metal tube the length of which was periodically changed from 39 cm to 73 cm by a motor (i.e., a bazooka-like sliding pipe). The subjects were instructed to pronounce the sustained vowel /ə/ at medium sound level and with constant glottal adjustment regardless of the change in tube length. Fundamental frequency (pitch) measurements were made at several frequencies in the chest register. Typical results for one subject are shown in Fig. 17.

The voice pitch was measured at 10-ms intervals by a pitch-extracting program.[20] The length of the metal tube (exclusive of the subject's vocal tract) is also indicated on the abscissa along with the corresponding time scale for the length change. Adjacent open and closed points (circles or triangles) pertain to different cycles of the pipe in one set of measurements. One sees frequency jumps similar to those in the two-mass model. However, the observed onset frequencies of the jumps are generally higher than the resonant frequency of the compound tube consisting of the metal tube and the subject's vocal tract (neglecting the shunt impedance of the vocal-tract wall). The deviation from the resonant frequency becomes especially noticeable for lower frequencies.

Toward an interpretation, it is known that the shunting impedance caused by vibration of the walls of the vocal tract produces a "cutoff frequency" of the sound transmission and constrains the lowest first formant frequency of the vocal tract.[21] This effect will contribute to raising the resonant frequency of the compound tube in a frequency range near the cutoff frequency. In the present instance, one could conceive of the walls of the cheeks, pharynx, and soft velum to yield to vibration because of the vocal-tract geometry for /ə/ and because of the long wavelength. At the cutoff frequency of the vocal tract, the first resonance frequency of the combined vocal tract and metal pipe is essentially that of the metal pipe alone. The latter is shown in Fig. 17 by the broken line.

From Fig. 17, we can presume the cutoff frequency of the vocal tract for /ə/ to be a little lower than 200 Hz. The effect of the wall vibration could thus account for the rightward shift of the observed pitch jumps. The rightward shift is most noticeable at the lower frequencies as this

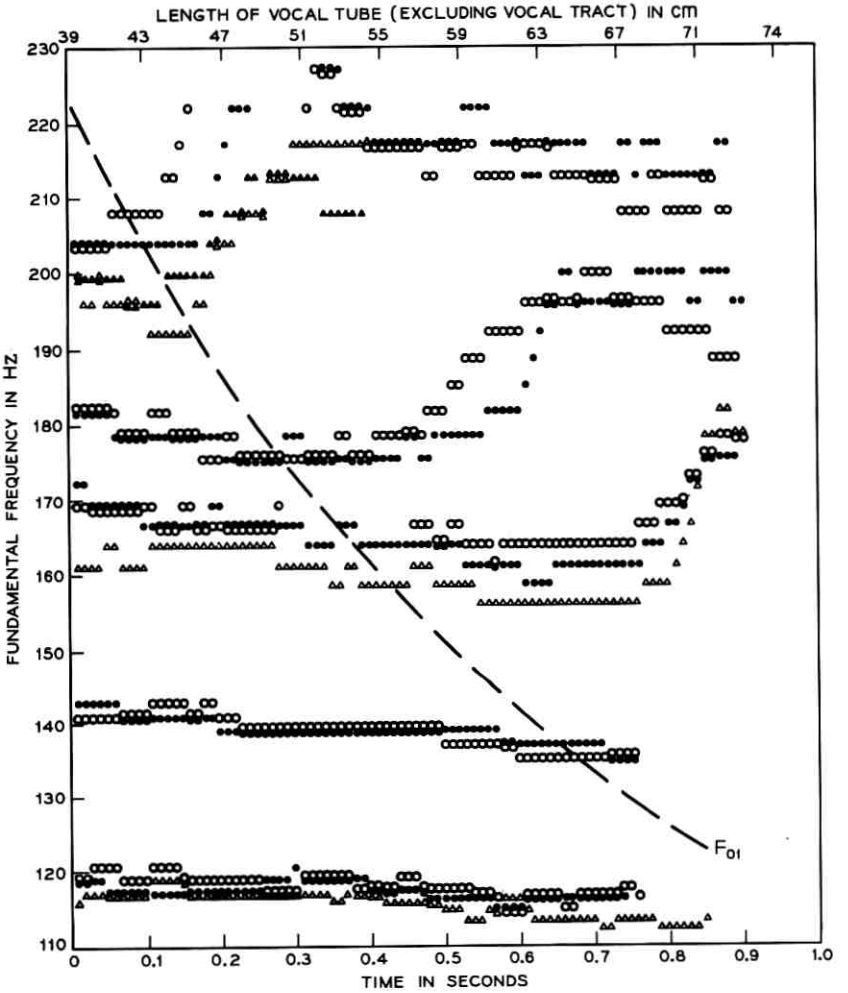


Fig. 17—Fundamental frequency measurements made for a human subject when the acoustic load on his vocal cords is varied. The acoustic load is varied by periodically changing the length of a uniform tube fitted to the subjects' mouth. The broken line shows the first resonant (pole) frequency of the uniform tube.

argument would predict. Even with these uncertainties, we see the close similarity in the dependence of fundamental frequency on acoustic load between the human larynx and the two-mass model.* It is further of interest that the vocal cords can self-oscillate without regenerative

* *Note added in proof:* After this paper was written, we measured the "cutoff frequency" for the vocal tract and tube combination. We found its value to be 195 Hz.

feedback from the subglottal and supraglottal system. In addition, the vibration of the soft walls of the vocal tract acts as a buffer to aid stable operation in the presence of coupling between the vocal cords and the vocal tract as the latter takes on a wide variety of shapes.

XII. CONCLUSION

The two-mass formulation of the vocal-cord model is seen to yield physiologically realistic behavior. In particular, the phase differences between upper and lower cord-edges corresponds well with motion observed in high-speed photography. The two-mass formulation also leads to a natural correlate to chest and falsetto register with coupling stiffness (lax in chest and tense in falsetto) being an important factor along with mass and thickness of the cords.

The computer measurements show that the two-mass model is capable of oscillation just above the resonant frequencies of the acoustic load (i.e., the formant frequencies of the vocal tract), duplicating a capability of the human cords. The one-mass model cannot oscillate in this frequency range, where the driving-point reactance is capacitive. Further, the intra-glottal pressure distribution derived for use with the two-mass model yields cord-tract interaction similar to human speech. Fundamental frequency varies with subglottal pressure approximately as 2 to 3 Hz/cm H_2O, and changes in vowel configuration do not markedly influence the fundamental frequency. Closures tighter than those which occur in vowel shapes (for example, at consonant-vowel boundaries) can of course influence the fundamental frequency. The improved intra-glottal pressure distribution is also applicable to a one-mass formulation, and it produces physiologically realistic cord-tract interactions with a one-mass model.

The programmed cord oscillator and the digitally simulated vocal tract constitute a complete synthesizer for voiced sounds. The system so implemented has potential for speech synthesis applications such as computer voice response. Especially for techniques such as text synthesis,[22] the cord model and vocal tract offer means for natural control of tract and larynx parameters, i.e., subglottal pressure, cord tension, neutral area, and tract shape. These parameters appear sufficient for describing both voiced and voiceless sounds in continuous speech.[3] In some synthesis applications, the complexity of the two-mass model may not be needed and a simpler one-mass formulation may serve. In normal voice production, phonation occurs at a fundamental frequency always below the first vocal resonance (formant). Here, the driving-point imped-

ance is inductive and the one-mass oscillator performs acceptably, particularly with the improved intra-glottal pressure distribution.

The two-mass model, because of its physiological detail, also provides a potential tool for medical analyses of voice disorder. Although the present simulation assumes bilateral symmetry of the opposing cords, asymmetric configurations can be implemented. The effects of deficiencies such as unilateral cord paralysis can therefore be investigated and quantified. Biomedical engineering is making increased use of digital simulations of physiological behavior. The simulation technique described here not only permits acoustic analysis of voice functions but of human respiration as well.

XIII. ACKNOWLEDGMENTS

We wish to thank several members of the Acoustics Research Department for their substantial contributions to this study. A. E. Rosenberg collaborated on the design and measurements with the impedance tube, D. E. Dudgeon made early simulations with the two-mass program, D. E. Bock assisted in an interactive implementation of the program on the DDP-516 computer, and K. Shipley programmed the pitch extractor used for the real voice measurements.

REFERENCES

1. Flanagan, J. L., and Landgraf, L. L., "Self-Oscillating Source for Vocal-Tract Synthesizers," Proc. IEEE-AFCRL Symposium on Speech Commun. and Process., Boston, Mass., (Nov. 1967), also published in IEEE Trans. Audio and Electroacoustics, *AU-16*, (March 1968), pp. 57–64.
2. Flanagan, J. L., "Use of an Interactive Laboratory Computer to Study an Acoustic-Oscillator Model of the Vocal Cords," IEEE Trans. Audio and Electroacoustics, *AU-17*, (March 1969), pp. 2–6.
3. Flanagan, J. L., and Cherry, L., "Excitation of Vocal-Tract Synthesizers," J. Acoust. Soc. Amer., *45*, (March 1969), pp. 764–769.
4. Dudgeon, D. E., "Two-Mass Model of the Vocal Cords," J. Acoust. Soc. Amer., *48*, (July 1970), p. 118A.
5. Ishizaka, K., "On Models of the Larynx," J. Acoust. Soc. Japan, *22*, 1966, pp. 293–294.
6. Ishizaka, K., and Matsudaira, M., "What Makes the Vocal Cords Vibrate," 6th Int. Congr. Acoust., Tokyo, (Aug. 1968), pp. B1–3.
7. Ishizaka, K., and Matsudaira, M., "Acoustic Theory of a Two-Mass Model of the Vocal Cords," unpublished work.
8. Ishizaka, K., and Kaneko, T., "On Equivalent Mechanical Constants of the Vocal Cords," J. Acoust. Soc. Japan, *24*, No. 5, 1968, pp. 312–313.
9. Kaufmann, W., *Fluid Mechanics*, New York: McGraw-Hill Co., 1963, p. 111.
10. van den Berg, J. W., Zantema, J. T., and Doornenbal, Jr., P., "On the Air Resistance and the Bernoulli Effect of the Human Larynx," J. Acoust. Soc. Amer., *29*, 1957, pp. 626–631.
11. Ishizaka, K., Kaneko, T., and Matsudaira, M., unpublished work.
12. Flanagan, J. L., *Speech Analysis, Synthesis and Perception*, 2nd Edition, New York, Berlin: Springer Verlag, 1972.

13. Flanagan, J. L., "Focal Points in Speech Communication Research," Proc. 7th Int. Cong. Acoust., Budapest, (August 1971). Also, IEEE Trans. Commun. Tech., (December 1971).
14. Farnsworth, D. W., "High-Speed Motion Pictures of the Human Vocal Cords," Bell Labs Record, *18*, No. 7 (March 1940), pp. 203–208.
15. Miller, R. L., "Nature of the Vocal Cord Wave," J. Acoust. Soc. Amer., *31*, 1969, pp. 667–677.
16. Fujimura, O., and Lindqvist, J., "Sweep Tone Measurements of Vocal-Tract Characteristics," J. Acoust. Soc. Amer., *49*, (Feb. 1971), pp. 554–558.
17. Holmes, J. N., "An Investigation of the Volume Velocity Waveform at the Larynx During Speech by Means of an Inverse Filter," Proc. Speech Commun. Seminar, Stockholm, 1962.
18. Hixon, T. J., Mead, J., and Klatt, D. H., "Influence of Forced Transglottal Pressure Changes on Vocal Fundamental Frequency," J. Acoust. Soc. Amer., *49*, (Jan. 1971), p. 105A.
19. Lindqvist, J., "The Voice Source Studied by Means of Inverse Filtering," Quarterly Report, Speech Transmicro Laboratory, Stockholm, Sweden, (Jan. 1970).
20. Gold, B., and Rabiner L., "Parallel Processing Techniques for Estimating Pitch Periods of Speech in the Time Domain," J. Acoust. Soc. Amer., *46*, (Aug. 1969), pp. 442–448.
21. Fant, G., "Speech at High Ambient Air-Pressure," STL-QPSR, (Feb. 1964), pp. 9–21.
22. Flanagan, J. L., Coker, C. H., Rabiner, L. R., Schafer, R. W., and Uneda, N., "Synthetic Voices for Computers," IEEE SPECTRUM, *7*, (October 1970), pp. 22–45.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

Exact Solutions to Some Deterministic and Random Transmission Line Problems

By J. McKENNA and J. A. MORRISON

(Manuscript received January 18, 1972)

A special class of transmission lines is considered, in which the modes decompose into two noninteracting sets. Both a single transmission line with constant characteristic impedance and variable propagation factor, and two transmission lines with equal propagation factors and variable coupling, in which the forward modes do not interact with the backward modes, are investigated. Exact expressions are obtained for the reflection and transmission coefficients when a section of such a transmission system connects two semi-infinite transmission systems consisting of constant impedance and admittance lines. These results hold for arbitrarily varying propagation factors and coupling; and while they are of independent interest in the case of deterministic variations, we make an application of them here in the case of stochastic variations.

Exact results are obtained for the ensemble averages of the transmission coefficient and transmitted power, and their variances, for the inserted section of single line, when the variable propagation factor is a random function involving either a Gaussian process or the random telegraph process. Asymptotic results are also obtained in the general case of weak fluctuations and long inserted sections. Analogous results may be obtained for the inserted section of two lines when they are randomly coupled, and the results are given in the case of matched lines, for which no reflections occur. Finally, some of the time domain statistics for lossless lines are considered, and expressions are derived for the ensemble averages of the transmitted pulse, due to pulses incident on the inserted section.

I. INTRODUCTION

This paper deals with a special class of the generalized equations of telegraphy. The starting point is the following simple observation. Consider the telegraphist equations in the frequency domain for a

single transmission line

$$\frac{d\mathcal{V}}{dx} = -\mathcal{Z}(x)\mathcal{I}(x),$$

$$\frac{d\mathcal{I}}{dx} = -\mathcal{Y}(x)\mathcal{V}(x), \tag{1}$$

where $\mathcal{V}(x)$ and $\mathcal{I}(x)$ are the time Fourier transforms of the voltage and current in the line, and $\mathcal{Z}(x)$ and $\mathcal{Y}(x)$ are the impedance and admittance, respectively. Then, if the characteristic impedance

$$K = \sqrt{\mathcal{Z}(x)/\mathcal{Y}(x)} \tag{2}$$

is a constant independent of x, it is simple to show that (1) has the two fundamental solutions

$$\mathcal{V}_+(x) = \exp\left\{-\int_0^x \Gamma(\xi)\, d\xi\right\}, \qquad \mathcal{I}_+(x) = \frac{1}{K}\mathcal{V}_+(x),$$

$$\mathcal{V}_-(x) = \exp\left\{\int_0^x \Gamma(\xi)\, d\xi\right\}, \qquad \mathcal{I}_-(x) = -\frac{1}{K}\mathcal{V}_-(x), \tag{3}$$

where

$$\Gamma(x) = \sqrt{\mathcal{Z}(x)\mathcal{Y}(x)}. \tag{4}$$

This simple result has several interesting consequences. Note that $\mathcal{V}_+(x)$ and $\mathcal{I}_+(x)$ describe a wave moving to the right and $\mathcal{V}_-(x)$ and $\mathcal{I}_-(x)$ describe a wave moving to the left (the time factor is assumed to be $e^{i\omega t}$). Furthermore, the wave moving to the right does not induce a reflected wave moving to the left, and vice versa, except possibly at the beginning or termination of the line.

This decomposition into noninteracting right and left moving waves suggests that a similar decomposition may exist in the case of n coupled transmission lines. This is the subject of Section II. It is shown there, that when a condition analogous to (2) is satisfied by the impedance and admittance matrices of the system, the system of $2n$ equations can be decomposed into two noninteracting sets of n equations. Under some circumstances, the n fundamental solutions of one system correspond to waves moving to the right; and the n fundamental solutions of the other system correspond to the left moving waves. In particular, we show that the model of two interacting waves recently studied by Rowe and Young[1,2] corresponds to just such a decomposition.

In addition, the form of the solution (3) for the single transmission line makes these models particularly convenient to study when $\mathcal{Z}(x)$

and $\mathcal{Y}(x)$ are stochastic processes. In Section III, we consider a finite section of length l of lossless transmission line with variable impedance and admittance satisfying condition (2). This section of line connects two semi-infinite lossless transmission lines having constant inductances and capacitances. We derive various expressions for the reflection and transmission coefficients of the inserted section. Using these expressions, in Section IV, we calculate exact expressions for the ensemble average of $T(\omega, l)$, the transmission coefficient, and of $| T(\omega, l) |^2$ and $| T(\omega, l) |^4$, when $\mathcal{Z}(x)$ and $\mathcal{Y}(x)$ are particular stochastic processes. Asymptotic results are obtained in the general case of weak fluctuations in $\mathcal{Z}(x)$ and $\mathcal{Y}(x)$, and long inserted sections. These results are based on a limit theorem of Khas'minskii,[3] and the details are given in the Appendix.

In Section V, we study a more complicated model consisting of two coupled transmission lines. Here the model involves a finite length, l, of two transmission lines having equal propagation constants, with variable inductive and capacitive coupling, connecting two semi-infinite, constant impedance and admittance lines. The self-impedance and admittance of the inserted section are also constant, and the semi-infinite lines are uncoupled. We derive expressions for the reflection and transmission matrices of the inserted section. Exact results may be obtained for the ensemble averages of the elements of the reflection and transmission matrices in the case of random coupling, for particular stochastic processes, using the results of Section IV. Since the results are quite lengthy in the general case, we give them only in the case of matched lines, so that no reflections occur.

Finally, in Section VI, we consider some of the time domain statistics of our models. Exact expressions are derived for the ensemble average of the transmitted pulse, due to a pulse incident on the inserted section of single transmission line with random inductance and capacitance. It is of interest to note that if the fluctuations in the propagation factor are described by a Gaussian process, the transmitted wave violates causality. This is not the case when the fluctuations are described by the random telegraph process. Analogous results are obtained for the transmitted pulses, due to pulses incident on the inserted section of two randomly coupled transmission lines, in the case of matched, lossless lines.

II. CLASS OF TRANSMISSION LINES

The generalized equations of telegraphy are the starting point of this paper:

$$\frac{\partial}{\partial x}\mathbf{V}(x, t) = -\mathbf{R}(x)\mathbf{I}(x, t) - \mathbf{L}(x)\frac{\partial}{\partial t}\mathbf{I}(x, t), \tag{5a}$$

$$\frac{\partial}{\partial x}\mathbf{I}(x, t) = -\mathbf{G}(x)\mathbf{V}(x, t) - \mathbf{C}(x)\frac{\partial}{\partial t}\mathbf{V}(x, t). \tag{5b}$$

These equations are typically used to describe the time and space variations of the current and voltage along n coupled transmission lines.[4] In this case, $\mathbf{V}$ and $\mathbf{I}$ are column vectors whose elements $V_p(x, t)$ and $I_p(x, t)$, $p = 1, 2, \cdots, n$, are, respectively, the voltage of the pth line relative to some fixed voltage, and the current in the pth line. $\mathbf{R}$, $\mathbf{L}$, $\mathbf{G}$, and $\mathbf{C}$ are $n \times n$ matrices, the resistance, inductance, conductance, and capacitance respectively, which typically are functions of the distance x along the lines.

We will, for the most part, find it convenient to work in the frequency plane, and so we introduce the Fourier transforms

$$\mathbf{V}(x, \omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbf{V}(x, t)e^{-j\omega t}\,d\omega, \tag{6a}$$

$$\mathbf{I}(x, \omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbf{I}(x, t)e^{-j\omega t}\,d\omega. \tag{6b}$$

Then $\mathbf{V}$ and $\mathbf{I}$ satisfy the equations

$$\frac{d\mathbf{V}}{dx} = -\mathbf{Z}(x)\mathbf{I}, \qquad \frac{d\mathbf{I}}{dx} = -\mathbf{Y}(x)\mathbf{V}, \tag{7}$$

where

$$\mathbf{Z}(x) = \mathbf{R}(x) + j\omega\mathbf{L}(x), \qquad \mathbf{Y}(x) = \mathbf{G}(x) + j\omega\mathbf{C}(x) \tag{8}$$

are the impedance and admittance matrices, respectively. It should be noted that the frequency domain equations appear in other contexts,[5] but there, the frequency dependance of $\mathbf{Z}(x)$ and $\mathbf{Y}(x)$ is generally more complicated. The remainder of this section is devoted to some general properties of the frequency domain eqs. (7).

It follows from (7) that

$$\frac{d}{dx}(\mathbf{V}^t\mathbf{I}^* + \mathbf{V}^{*t}\mathbf{I}) = -\mathbf{I}^t[\mathbf{Z}^t(x) + \mathbf{Z}^*(x)]\mathbf{I}^* - \mathbf{V}^t[\mathbf{Y}^t(x) + \mathbf{Y}^*(x)]\mathbf{V}^*, \tag{9}$$

where t denotes transpose and $*$ denotes complex conjugate. Hence, for lossless lines

$$\mathbf{Z}^t(x) + \mathbf{Z}^*(x) = 0, \qquad \mathbf{Y}^t(x) + \mathbf{Y}^*(x) = 0. \tag{10}$$

We wish to consider the class of coupled transmission lines for which

$$\mathbf{Z}(x) = \mathbf{K}\mathbf{Y}(x)\mathbf{K}, \tag{11}$$

where $\mathbf{K}$ is a nonsingular constant matrix. Note that, in the lossless case, the first condition in (10) follows from the second if $\mathbf{K}^t = \mathbf{K}^*$. When (11) is satisfied, the solutions of (7) may be split into two groups, namely

$$\mathbf{V} = \mathbf{K}\mathbf{I}, \qquad \frac{d\mathbf{V}}{dx} = -\mathbf{K}\mathbf{Y}(x)\mathbf{V}, \tag{12}$$

and

$$\mathbf{V} = -\mathbf{K}\mathbf{I}, \qquad \frac{d\mathbf{V}}{dx} = \mathbf{K}\mathbf{Y}(x)\mathbf{V}. \tag{13}$$

If the lines are appropriately matched at both ends, then either one set of solutions or the other occurs, and reflections are avoided. Since $\mathbf{K}$ is constant, this matching is independent of the length of the lines.

As a particular example, let $n = 2$ and

$$\mathbf{K} = \begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} \dfrac{\Gamma_1}{K_1} & \dfrac{-jc(x)}{(K_1K_2)^{\frac{1}{2}}} \\ \dfrac{-jc(x)}{(K_1K_2)^{\frac{1}{2}}} & \dfrac{\Gamma_2}{K_2} \end{bmatrix}. \tag{14}$$

Thus, from (11),

$$\mathbf{Z} = \begin{bmatrix} \Gamma_1K_1 & -j(K_1K_2)^{\frac{1}{2}}c(x) \\ -j(K_1K_2)^{\frac{1}{2}}c(x) & \Gamma_2K_2 \end{bmatrix}. \tag{15}$$

Corresponding to (12), we have

$$\begin{aligned} \frac{d\mathcal{V}_1}{dx} + \Gamma_1\mathcal{V}_1 &= jc(x)\left(\frac{K_1}{K_2}\right)^{\frac{1}{2}}\mathcal{V}_2 , \\ \frac{d\mathcal{V}_2}{dx} + \Gamma_2\mathcal{V}_2 &= jc(x)\left(\frac{K_2}{K_1}\right)^{\frac{1}{2}}\mathcal{V}_1 . \end{aligned} \tag{16}$$

The substitutions $\mathcal{V}_1 = K_1^{\frac{1}{2}}I_1$, $\mathcal{V}_2 = K_2^{\frac{1}{2}}I_2$ lead to the equations for two coupled modes traveling in the same direction, which have been considered previously.[1,2] We remark that if we choose $\mathbf{K} = \left[\begin{smallmatrix} K_1 & 0 \\ 0 & -K_2 \end{smallmatrix}\right]$ instead, then we are led to equations for two modes traveling in opposite directions.

Next, we consider a particular class of transmission lines, satisfying (11), for which

$$\mathbf{K}\mathbf{Y}(x) = \Gamma(x)\mathbf{I} - jc(x)\mathbf{A}, \tag{17}$$

where $\mathbf{I}$ is the unit matrix of order n, and $\mathbf{A}$ is a constant matrix. Then there are solutions of (12) and (13) of the form

$$\mathbf{V} = \mathbf{b} \exp\left\{\mp \int [\Gamma(x) - j\lambda c(x)]\, dx\right\} = \pm \mathbf{K}\mathbf{I}, \tag{18}$$

where $\mathbf{b}$ is a constant vector satisfying

$$(\lambda \mathbf{I} - \mathbf{A})\mathbf{b} = 0. \tag{19}$$

The eigenvalues λ are given by $|\lambda \mathbf{I} - \mathbf{A}| = 0$.

We will assume that $c(x)\mathbf{A}$ describes only the coupling between lines, so that $\mathbf{A}$ has diagonal elements equal to zero. The case $n = 1$ (for which $\mathbf{A} = 0$, and $\lambda = 0$ is the only eigenvalue) corresponds to the well-known case of a single line with constant characteristic impedance and variable propagation factor. This case is considered further in subsequent sections. The case $n \geqq 2$ corresponds to n transmission lines with identical propagation factors and variable coupling. Such a situation might arise in the consideration of n twisted pairs in a cable, although the relationship (11) is not too realistic. In Section V, we consider the case $n = 2$ corresponding to $\Gamma_1 = \Gamma_2$ in (14), so that (17) holds.

III. SINGLE TRANSMISSION LINE

In this and the following section, we study in some detail the following example. Consider an infinitely long, lossless, single transmission line ($n = 1$ in the classification of Section II) which for $x < 0$ has the constant impedance and admittance $\mathfrak{z}_0 = j\omega L_0$, $\mathfrak{y}_0 = j\omega C_0$, for $x > l$ has the constant impedance and admittance $\mathfrak{z}_l = j\omega L_l$ and $\mathfrak{y}_l = j\omega C_l$, while the central section $0 < x < l$ has the variable impedance and admittance $\mathfrak{z}(x) = j\omega L(x)$, $\mathfrak{y}(x) = j\omega C(x)$. A wave traveling to the right in the region $x < 0$ will be partially reflected and partially transmitted on striking the central region in $0 < x < l$. We study the transmitted wave under the assumption that $\mathfrak{z}(x)$ and $\mathfrak{y}(x)$ satisfy condition (11), i.e.,

$$K = \sqrt{\mathfrak{z}(x)/\mathfrak{y}(x)} = \sqrt{L(x)/C(x)} \tag{20}$$

is a positive constant independent of x.

Although this case is probably hard to realize physically, it is nevertheless of considerable interest, since it is mathematically simple enough so that many interesting questions about it can be answered.

Let

$$\begin{aligned} \Gamma_\alpha &= \sqrt{\mathfrak{z}_\alpha \mathfrak{y}_\alpha} = j\omega \sqrt{L_\alpha C_\alpha} = j\omega\gamma_\alpha , \\ K_\alpha &= \sqrt{\mathfrak{z}_\alpha/\mathfrak{y}_\alpha} = \sqrt{L_\alpha/C_\alpha} , \qquad \alpha = 0, l. \end{aligned} \tag{21}$$

Then in $x \leqq 0$ and $x \geqq l$, we can solve eqs. (7) simply. In $x \leqq 0$, we have a solution

$$\mathcal{V}(x) = e^{-\Gamma_0 x} + R(\omega, l)e^{\Gamma_0 x},$$

$$\mathcal{I}(x) = \frac{1}{K_0}(e^{-\Gamma_0 x} - R(\omega, l)e^{\Gamma_0 x}). \tag{22}$$

This represents a plane wave $e^{-\Gamma_0 x}$ moving to the right and a reflected wave $R(\omega, l)e^{\Gamma_0 x}$ moving to the left, where $R(\omega, l)$, the reflection coefficient, is a function of ω and l. Similarly, in $x \geqq l$, the solution can be written

$$\mathcal{V}(x) = T(\omega, l)e^{-\Gamma_l(x-l)} = K_l \mathcal{I}(x), \tag{23}$$

representing a transmitted plane wave moving to the right, where $T(\omega, l)$ is the transmission coefficient.

In $0 \leqq x \leqq l$ we define the propagation factor

$$\Gamma(x) = \sqrt{\mathcal{Z}(x)\mathcal{Y}(x)} = j\omega K C(x) = j\omega \frac{L(x)}{K}. \tag{24}$$

Then, from (18), we can write the general solution of (7) in $0 \leqq x \leqq l$ as

$$\mathcal{V}(x) = \mathrm{A}\mathcal{V}_+(x) + \mathrm{B}\mathcal{V}_-(x),$$

$$\mathcal{I}(x) = \frac{1}{K}(\mathrm{A}\mathcal{V}_+(x) - \mathrm{B}\mathcal{V}_-(x)), \tag{25}$$

when A and B are constants and

$$\mathcal{V}_\pm(x) = \exp\left\{\mp \int_0^x \Gamma(\xi)\, d\xi\right\}. \tag{26}$$

We now have a solution depending on four unknown constants A, B, R and T which can be determined from the condition that $\mathcal{V}(x)$ and $\mathcal{I}(x)$ must be continuous at $x = 0$ and $x = l$. The resulting four linear equations are easily solved and yield for the reflection and transmission coefficients

$$R(\omega, l) = \frac{(K - K_0)(K + K_l)\mathcal{V}_-(l) - (K + K_0)(K - K_l)\mathcal{V}_+(l)}{(K + K_0)(K + K_l)\mathcal{V}_-(l) - (K - K_0)(K - K_l)\mathcal{V}_+(l)}, \tag{27}$$

$$T(\omega, l) = \frac{4KK_l}{(K + K_0)(K + K_l)\mathcal{V}_-(l) - (K - K_0)(K - K_l)\mathcal{V}_+(l)}. \tag{28}$$

Notice that if $K = K_0 = K_l$, then $R(\omega, l) = 0$.

We confine our study to the transmitted wave, although the reflected

wave can be studied equally well by the techniques we employ. In particular, there is the easily proved energy conservation relationship:

$$1 = |R(\omega, l)|^2 + \frac{K_0}{K_l} |T(\omega, l)|^2. \tag{29}$$

Before proceeding, let us assume instead that a section of line $0 \leqq x \leqq l$ is driven by a voltage source in series with an impedance Z_0 and that the line is terminated in an impedance Z_l, as shown in Fig. 1. Then, $\mathcal{V}(x)$ and $\mathcal{I}(x)$ are still given by (25), but the boundary conditions, from which A and B are determined, are now

$$\mathcal{V}(0) + Z_0 \mathcal{I}(0) = E_0,$$
$$\mathcal{V}(l) = Z_l \mathcal{I}(l). \tag{30}$$

Then, it is easily shown that

$$\mathcal{V}(l) = \frac{2KZ_l E_0}{(K + Z_0)(K + Z_l)\mathcal{V}_-(l) - (K - Z_0)(K - Z_l)\mathcal{V}_+(l)}, \tag{31}$$

so the transfer impedance in this formulation is essentially identical with the transmission coefficient in the first formulation. We shall continue to use the first formulation.

We now further specialize the model, and let

$$L(x) = L(1 + \epsilon N(x)), \tag{32}$$

where we assume that $0 \leqq \epsilon \leqq 1$ is a dimensionless constant, $L > 0$ is a constant with the dimension of inductance, and $N(x)$ is a (dimensionless) stochastic process with zero mean. It follows that L is the stochastic mean of $L(x)$,

$$L = \langle L(x) \rangle. \tag{33}$$

The symbol $\langle\ \rangle$ will be used throughout to denote the stochastic mean.

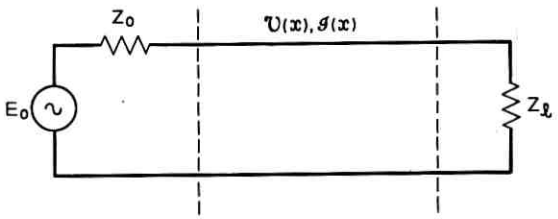


Fig. 1—Diagram of transmission line circuit driven by source E_0 at $x = 0$ with internal impedance Z_0 and terminated at $x = l$ by impedance Z_l.

It follows from (20) that

$$C(x) = C(1 + \epsilon N(x)), \qquad K = \sqrt{\frac{L}{C}}. \tag{34}$$

Further define

$$\gamma = \sqrt{CL}, \tag{35}$$

so that

$$\Gamma(x) = j\omega\gamma(1 + \epsilon N(x)). \tag{36}$$

In addition, we let

$$\theta(l) = \int_0^l N(\xi)\, d\xi, \tag{37}$$

and

$$\rho = \frac{(K - K_0)(K - K_l)}{(K + K_0)(K + K_l)}, \qquad \lambda = \frac{4KK_l}{(K + K_0)(K + K_l)}. \tag{38}$$

Then we can write expression (28) for $T(\omega, l)$ as

$$T(\omega, l) = \lambda e^{-j\omega\gamma(l+\epsilon\theta(l))}[1 - \rho e^{-2j\omega\gamma(l+\epsilon\theta(l))}]^{-1}. \tag{39}$$

We can now derive some series expansions of $T(\omega, l)$ and some of its powers which will prove useful in the next section. Since K, K_0 and K_l are positive,

$$|\rho| < 1, \tag{40}$$

hence we have the geometric series expansion

$$T(\omega, l) = \lambda \sum_{r=0}^{\infty} \rho^r e^{-(2r+1)j\omega\gamma(l+\epsilon\theta(l))}. \tag{41}$$

Next let

$$f(\varphi) = \lambda^2/[(1 - \rho e^{-i\varphi})(1 - \rho e^{i\varphi})]. \tag{42}$$

Then $f(\varphi)$ has the partial fraction expansion

$$f(\varphi) = \frac{\lambda^2}{1 - \rho^2}\left[\frac{1}{1 - \rho e^{i\varphi}} + \frac{\rho e^{-i\varphi}}{1 - \rho e^{-i\varphi}}\right], \tag{43}$$

and hence $f(\varphi)$ has the series expansion

$$f(\varphi) = \frac{\lambda^2}{1 - \rho^2} \sum_{n=-\infty}^{\infty} \rho^{|n|} e^{in\varphi}. \tag{44}$$

Then, it follows that

$$|T(\omega, l)|^2 = \frac{\lambda^2}{1 - \rho^2} \sum_{r=-\infty}^{\infty} \rho^{|r|} e^{2rj\omega\gamma(l+\epsilon\theta(l))}. \tag{45}$$

We finally need to examine double sums of the form

$$\mathrm{S} = \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} \rho^{|r|+|s|} h(r+s). \tag{46}$$

If we make the change of summation variables $u = r + s$, $t = r$, we can write S as

$$\mathrm{S} = \sum_{u=-\infty}^{\infty} h(u)\xi(u), \tag{47}$$

where

$$\xi(u) = \sum_{t=-\infty}^{\infty} \rho^{|t|+|t-u|}. \tag{48}$$

The series $\xi(u)$ can be evaluated simply, and so we obtain

$$\mathrm{S} = \sum_{u=-\infty}^{\infty} \left\{ |u| + \frac{1+\rho^2}{1-\rho^2} \right\} \rho^{|u|} h(u). \tag{49}$$

If we now square expression (45) for $|T(\omega, l)|^2$, we obtain a series of the form (46), and so it follows that

$$|T(\omega, l)|^4 = \frac{\lambda^4}{(1-\rho^2)^2} \sum_{r=-\infty}^{\infty} \left\{ |r| + \frac{1+\rho^2}{1-\rho^2} \right\} \rho^{|r|} e^{2rj\omega\gamma(l+\epsilon\theta(l))}. \tag{50}$$

IV. FREQUENCY DOMAIN STATISTICS

In this section we study some of the frequency domain statistics of the model described in Section III. In particular, we obtain expressions for the stochastic average of $T(\omega, l)$ and $|T(\omega, l)|^2$ and examine the standard deviation of these quantities.

It is clear from (41), and (45) and (50) that the problem of calculating $\langle T(\omega, l) \rangle$, $\langle |T(\omega, l)|^2 \rangle$ and $\langle |T(\omega, l)|^4 \rangle$ has been reduced to the problem of calculating $\langle e^{j\alpha\theta(l)} \rangle$.

Consider first the case where $N(x)$ is a zero-mean, Gaussian random process. Then $\theta(l)$ is a zero-mean, Gaussian random variable with variance[6]

$$\sigma^2(l) = \int_0^l \int_0^l \langle N(x)N(y) \rangle \, dx \, dy, \tag{51}$$

and

$$\langle e^{j\alpha\theta(l)}\rangle = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{j\alpha\theta-\theta^2/2\sigma^2}\,d\theta = e^{-\frac{1}{2}\alpha^2\sigma^2}. \tag{52}$$

Consequently,

$$M_1(\omega, l) = \langle T(\omega, l)\rangle = \lambda\sum_{r=0}^{\infty}\rho^{|r|}e(2r+1, \omega, l), \tag{53}$$

$$M_2(\omega, l) = \langle |T(\omega, l)|^2\rangle = \frac{\lambda^2}{1-\rho^2}\sum_{r=-\infty}^{\infty}\rho^{|r|}e(2r, \omega, l), \tag{54}$$

and

$$M_4(\omega, l) = \langle |T(\omega, l)|^4\rangle = \frac{\lambda^4}{(1-\rho^2)^2}\sum_{r=-\infty}^{\infty}\left\{|r| + \frac{1+\rho^2}{1-\rho^2}\right\}\rho^{|r|}e(2r, \omega, l), \tag{55}$$

where

$$e(r, \omega, l) = \exp\{-jr\omega\gamma l - \tfrac{1}{2}r^2\omega^2\gamma^2\epsilon^2\sigma^2(l)\}. \tag{56}$$

If $N(x)$ is white noise, then $\langle N(x)N(y)\rangle = D_0\,\delta(x-y)$ and

$$\sigma^2(l) = D_0 l\,, \tag{57}$$

where D_0 is a constant having dimensions of length. If $N(x)$ is a wide-sense stationary Gaussian process,

$$\langle N(x)N(y)\rangle = g(|\,x-y\,|), \tag{58}$$

with continuous $g(\xi)$, then

$$\sigma^2(l) = 2\int_0^l (l-\xi)g(\xi)\,d\xi. \tag{59}$$

In particular, if $g(\xi) = e^{-2b\xi}$, then

$$\sigma^2(l) = \frac{1}{b}\left[l - \frac{1-e^{-2bl}}{2b}\right]. \tag{60}$$

Since $\sigma(0) = 0$ in all these cases, $e(r, \omega, 0) = 1$, and so

$$M_n(\omega, 0) = \left(\frac{\lambda}{1-\rho}\right)^n, \qquad n = 1, 2, 4. \tag{61}$$

In many cases, such as (57) and (60), $\sigma^2(l)$ tends monotonely to ∞ as $l \to \infty$. In these cases we obtain the asymptotics of the moments as $l \to \infty$ simply. For $\omega\gamma\epsilon\sigma(l) > 1$, we have

$$M_1(\omega, l) \approx \lambda e^{-j\omega\gamma l - \frac{1}{2}\omega^2\gamma^2\epsilon^2\sigma^2(l)}, \tag{62}$$

$$M_2(\omega, l) \approx \frac{\lambda^2}{1-\rho^2}\{1 + 2\rho\cos(2\omega\gamma l)e^{-2\omega^2\gamma^2\epsilon^2\sigma^2(l)}\}, \tag{63}$$

$$M_4(\omega, l) \approx \frac{\lambda^4}{(1-\rho^2)^3}\{1 + \rho^2 + 4\rho\cos(2\omega\gamma l)e^{-2\omega^2\gamma^2\epsilon^2\sigma^2(l)}\}. \tag{64}$$

The mean-amplitude transmission coefficient $M_1(\omega, l)$ decays exponentially to zero as $l \to \infty$, while the mean-power transmission coefficient tends to the limit

$$\frac{\lambda^2}{1-\rho^2} = \frac{4KK_l^2}{(K_0 + K_l)(K^2 + K_0K_l)}. \tag{65}$$

This can be explained qualitatively by noting that, when the transmitted amplitudes are averaged over the ensemble, cancellation can take place, while the transmitted powers all have the same sign and so no cancellation can take place on averaging.

It is easily seen from (62) through (64) that the ratio of the standard deviation to the mean of $T(\omega, l)$ is

$$\Sigma_1(\omega, l) \approx \frac{1}{\sqrt{1-\rho^2}} e^{\frac{1}{2}\omega^2\gamma^2\epsilon^2\sigma^2(l)}, \tag{66}$$

while the ratio of the standard deviation to the mean of $|T(\omega, l)|^2$ is

$$\Sigma_2(\omega, l) \approx \sqrt{\frac{2\rho^2}{1-\rho^2}}[1 - \rho\cos(2\omega\gamma l)e^{-2\omega^2\gamma^2\epsilon^2\sigma^2(l)}]. \tag{67}$$

For the examples (57) and (60), $\Sigma_1(\omega, l) \to \infty$ exponentially as $l \to \infty$, while $\Sigma_2(\omega, l)$ tends to the limit

$$\sqrt{\frac{2\rho^2}{1-\rho^2}} = \frac{|K - K_0||K - K_l|}{\sqrt{2K(K_0 + K_l)(K^2 + K_0K_l)}}. \tag{68}$$

To get some feel for the numbers, we note that for

$$\frac{1}{2} \leqq \frac{K_0}{K} = \frac{K_l}{K} \leqq 2, \quad 0 \leqq \sqrt{\frac{2\rho^2}{1-\rho^2}} \leqq 0.158.$$

However, if K_0 and K_l differ too much from K, this ratio becomes much larger than 1.

As a second example, consider the case where $N(x)$ is the random telegraph process.[7] It is an ensemble of square wave functions $\{N(x)\}$, such that each sample function $N(x)$ can assume only the values ± 1. For fixed x, a sample function chosen at random will equal $+1$ or -1

with equal probability. The probability $p(n, x)$ of a given sample function changing sign n times in an interval of length x is given by the Poisson process

$$p(n, x) = \frac{(bx)^n}{n!} e^{-bx}, \qquad (n = 0, 1, 2, \cdots). \tag{69}$$

This process has zero mean and correlation function

$$\langle N(x)N(y)\rangle = e^{-2b|x-y|}. \tag{70}$$

The probability density function for the integral of the random telegraph process has been derived by McFadden,[8] and in our notation

$$P(\theta, l) = \tfrac{1}{2}e^{-bl}\Bigg[\delta(l-\theta) + \delta(l+\theta) + b\Bigg\{I_0[b(l^2-\theta^2)^{\frac{1}{2}}] + \frac{lI_1[b(l^2-\theta^2)^{\frac{1}{2}}]}{(l^2-\theta^2)^{\frac{1}{2}}}\Bigg\}H(l-\theta)H(l+\theta)\Bigg], \tag{71}$$

where $H(u)$ is the Heaviside function

$$H(u) = 1, \qquad u > 0,$$
$$H(u) = 0, \qquad u < 0. \tag{72}$$

It follows that[9]

$$\langle e^{j\alpha\theta(l)}\rangle = e^{-bl}\Bigg[\cosh\{(b^2-\alpha^2)^{\frac{1}{2}}l\} + \frac{b}{(b^2-\alpha^2)^{\frac{1}{2}}}\sinh\{(b^2-\alpha^2)^{\frac{1}{2}}l\}\Bigg]. \tag{73}$$

If we define

$$f(r, \omega, l) = \exp\{-jr\omega\gamma l - bl\} \cdot \Bigg[\cosh\{(b^2 - r^2\omega^2\gamma^2\epsilon^2)^{\frac{1}{2}}l\} + \frac{b\sinh\{(b^2-r^2\omega^2\gamma^2\epsilon^2)^{\frac{1}{2}}l\}}{(b^2-r^2\omega^2\gamma^2\epsilon^2)^{\frac{1}{2}}}\Bigg], \tag{74}$$

then the expressions for $\langle T(\omega, l)\rangle$, $\langle|T(\omega, l)|^2\rangle$ and $\langle|T(\omega, l)|^4\rangle$ for the random telegraph case can be obtained from (53) through (55) on on replacing $e(r, \omega, l)$ by $f(r, \omega, l)$.

If $\epsilon > 0$ is small enough so that for a given positive integer r_0

$$\eta \equiv (\epsilon r_0\omega\gamma/b)^2 \ll 1, \qquad \eta^2 bl \ll 1,$$

then for $0 \leqq r \leqq r_0$,

$$f(r, \omega, l) = \exp\{-jr\omega\gamma l - \tfrac{1}{2}r^2\omega^2\gamma^2\epsilon^2 l/b\}[1 + 0(\eta) + 0(\eta^2 bl)]. \tag{75}$$

If furthermore $bl \gg 1$, then from (60) for the Gaussian with correlation function (70), we have

$$\sigma^2(l) = \frac{l}{b}\left[1 + 0\left(\frac{1}{bl}\right)\right].$$

Hence, for these two cases, for $0 \leqq r \leqq r_0$,

$$\frac{f(r, \omega, l)}{e(r, \omega, l)} = [1 + 0(\eta) + 0(\eta^2 bl)]. \tag{76}$$

Note that both $e(r, \omega, l)$ and $f(r, \omega, l)$ are exponentially small for $r > r_0$ if ηbl is moderately large. Also, from (74), as $l \to \infty$ the first-order moment of $T(\omega, l)$ tends to zero, and the second moment and its standard deviation tend to the same limits as in the Gaussian case.

We now consider a general case of weak fluctuations in the inductance and capacitance, and long sections of line, so that $0 < \epsilon \ll 1$ in (32), and $l = \Lambda/\epsilon^2$. It is assumed that $N(x)$ is a bounded, zero-mean, wide-sense stationary stochastic process, with correlation function given by (58). An application is made in the Appendix of a limit theorem due to Khas'minskii,[3] in order to determine the behavior of

$$\epsilon\theta(\Lambda/\epsilon^2) = \epsilon \int_0^{\Lambda/\epsilon^2} N(x)\, dx, \tag{77}$$

for Λ bounded and $\epsilon \to 0$.

If the stochastic process $N(x)$ satisfies a certain strong mixing condition,[3] then it is found that the process $\epsilon\theta(\Lambda/\epsilon^2)$ converges weakly to a Markov diffusion process $\Theta(\Lambda)$, with probability density function

$$p(\Theta, \Lambda) = \frac{1}{(2\pi\bar{a}\Lambda)^{\frac{1}{2}}} \exp\left[\frac{-\Theta^2}{2\bar{a}\Lambda}\right], \tag{78}$$

where

$$\bar{a} = 2 \lim_{X\to\infty}\left[\frac{1}{X}\int_0^X (X - z)g(z)\, dz\right], \tag{79}$$

and g is given by (58). If, as we assume,

$$\lim_{X\to\infty}\left[\frac{1}{X}\int_0^X zg(z)\, dz\right] = 0, \tag{80}$$

then

$$\bar{a} = 2\int_0^\infty g(z)\, dz. \tag{81}$$

It follows from (78) that

$$\langle \exp [jr\omega\gamma\epsilon\theta(\Lambda/\epsilon^2)]\rangle \approx \exp(-\tfrac{1}{2}\bar{a}r^2\omega^2\gamma^2\Lambda). \tag{82}$$

Hence, for $0 < \epsilon \ll 1$ and $l = \Lambda/\epsilon^2$, the asymptotic expressions for $\langle T(\omega, l)\rangle$, $\langle |T(\omega, l)|^2\rangle$ and $\langle |T(\omega, l)|^4\rangle$ can be obtained from (53) through (55) on replacing $e(r, \omega, l)$ by

$$h(r, \omega, l) = \exp\{-jr\omega\gamma l - \tfrac{1}{2}r^2\omega^2\gamma^2\epsilon^2\bar{a}l\}, \tag{83}$$

where $\bar{a}$ is given by (81). For the random telegraph case $g(z) = \exp(-2bz)$, so that $\bar{a} = 1/b$, and the consistency of (75) with the above result is noted.

V. TWO COUPLED TRANSMISSION LINES

We consider here the case of two coupled transmission lines described in eqs. (14) through (16), but with identical propagation constants so that

$$\Gamma_1 = \Gamma = \Gamma_2 . \tag{84}$$

Then Γ is independent of x in (17), and

$$\mathbf{A} = \begin{bmatrix} 0 & (K_1/K_2)^{\frac{1}{2}} \\ (K_2/K_1)^{\frac{1}{2}} & 0 \end{bmatrix}. \tag{85}$$

Thus, the eigenvalues of $\mathbf{A}$ are $\lambda = \pm 1$. Let

$$\xi(x) = \int_0^x c(y)\, dy, \tag{86}$$

it being assumed that $c(x)$ is real, and that $K_1K_2 > 0$. Then, from (14), (18), and (19), the general solution of (7), subject to (11), for this case may be written in the form

$$v_1 = K_1^{\frac{1}{2}}[e^{j\xi(x)}(Be^{-\Gamma x} + Ce^{\Gamma x}) + e^{-j\xi(x)}(Ae^{-\Gamma x} + De^{\Gamma x})], \tag{87}$$

$$v_2 = K_2^{\frac{1}{2}}[e^{j\xi(x)}(Be^{-\Gamma x} - Ce^{\Gamma x}) - e^{-j\xi(x)}(Ae^{-\Gamma x} - De^{\Gamma x})], \tag{88}$$

$$g_1 = K_1^{-\frac{1}{2}}[e^{j\xi(x)}(Be^{-\Gamma x} - Ce^{\Gamma x}) + e^{-j\xi(x)}(Ae^{-\Gamma x} - De^{\Gamma x})], \tag{89}$$

$$g_2 = K_2^{-\frac{1}{2}}[e^{j\xi(x)}(Be^{-\Gamma x} + Ce^{\Gamma x}) - e^{-j\xi(x)}(Ae^{-\Gamma x} + De^{\Gamma x})]. \tag{90}$$

We suppose that the coupled lines extend from $x = 0$ to $x = l$, so that (87) through (90) hold for $0 \leqq x \leqq l$. For $x < 0$ and $x > l$ we suppose that the transmission system consists of uncoupled lines with constant propagation constants Γ_{p0} and Γ_{pl}, and constant characteristic

impedances K_{p0} and K_{pl}, respectively, $(p = 1, 2)$. Then, for an incoming wave on line 1,

$$\mathcal{V}_1 = (e^{-\Gamma_{10}x} + R_1 e^{\Gamma_{10}x}), \qquad \mathcal{V}_2 = R_2 e^{\Gamma_{20}x}, \tag{91}$$

$$\mathcal{I}_1 = \frac{1}{K_{10}}(e^{-\Gamma_{10}x} - R_1 e^{\Gamma_{10}x}), \qquad \mathcal{I}_2 = \frac{-R_2}{K_{20}} e^{\Gamma_{20}x}, \tag{92}$$

for $x \leqq 0$, and

$$\mathcal{V}_1 = T_1 e^{-\Gamma_{1l}(x-l)}, \qquad \mathcal{V}_2 = T_2 e^{-\Gamma_{2l}(x-l)}, \tag{93}$$

$$\mathcal{I}_1 = \frac{T_1}{K_{1l}} e^{-\Gamma_{1l}(x-l)}, \qquad \mathcal{I}_2 = \frac{T_2}{K_{2l}} e^{-\Gamma_{2l}(x-l)}, \tag{94}$$

for $x \geqq l$.

The boundary conditions are that $\mathcal{V}_p$ and $\mathcal{I}_p$, $(p = 1, 2)$, must be continuous at $x = 0$ and $x = l$. The calculation of reflection and transmission coefficients is tedious, but straightforward, so we omit the details and merely state the result. Let

$$\mu_p = \frac{K_{p0}}{K_p}, \qquad \nu_p = \frac{K_{pl}}{K_p}, \qquad (p = 1, 2), \tag{95}$$

and

$$\kappa = (K_2/K_1)^{\frac{1}{2}}, \qquad \chi = \xi(l). \tag{96}$$

Also, define

$$\begin{aligned}\Delta = [&(1 + \mu_1)(1 + \mu_2)(1 + \nu_1)(1 + \nu_2)e^{2\Gamma l} - 2(\mu_1 - \mu_2)(\nu_1 - \nu_2) \\ &+ (1 - \mu_1)(1 - \mu_2)(1 - \nu_1)(1 - \nu_2)e^{-2\Gamma l} \\ &- 2(1 - \mu_1\mu_2)(1 - \nu_1\nu_2)\cos 2\chi].\end{aligned} \tag{97}$$

Then it is found that

$$\Delta T_1 = 4\nu_1[(1 + \mu_2)(1 + \nu_2)e^{\Gamma l} - (1 - \mu_2)(1 - \nu_2)e^{-\Gamma l}]\cos\chi, \tag{98}$$

$$\Delta T_2 = 4j\kappa\nu_2[(1 + \mu_2)(1 + \nu_1)e^{\Gamma l} + (1 - \mu_2)(1 - \nu_1)e^{-\Gamma l}]\sin\chi, \tag{99}$$

$$\begin{aligned}\Delta R_1 = [&(1 - \mu_1)(1 + \mu_2)(1 + \nu_1)(1 + \nu_2)e^{2\Gamma l} + 2(\mu_1 + \mu_2)(\nu_1 - \nu_2) \\ &+ (1 + \mu_1)(1 - \mu_2)(1 - \nu_1)(1 - \nu_2)e^{-2\Gamma l} \\ &- 2(1 + \mu_1\mu_2)(1 - \nu_1\nu_2)\cos 2\chi],\end{aligned} \tag{100}$$

and

$$\Delta R_2 = -4j\kappa\mu_2(1 - \nu_1\nu_2)\sin 2\chi. \tag{101}$$

We remark that the reflection and transmission coefficients corresponding to an incoming wave on line 2 may be obtained by appropriate interchange of subscripts.

Since the coupling function $c(x)$ is real, χ is also real, from (86) and (96). If k is an integer, then $T_2 = 0$ and $R_2 = 0$ for $\chi = k\pi$, and $T_1 = 0$ and $R_2 = 0$ for $\chi = (k + \frac{1}{2})\pi$. The oscillatory behavior of T_1 and T_2 has been found earlier by Foschini,[10] in the case of matched lines with equal characteristic impedances, i.e., $\mu_p = 1 = \nu_p$, $(p = 1, 2)$, and $\kappa = 1$. In this case there are no reflections.

In the general case, the expressions for the transmission and reflection coefficients T_1, T_2 and R_1, R_2 may be expanded in Fourier series in χ. Thus, in the case of random coupling between the lines, the problem of calculating the expectations of the transmission and reflection coefficients reduces to that of calculating $\langle \exp jr\chi \rangle$ where, from (86) and (96),

$$\chi = \int_0^l c(x)\, dx. \tag{102}$$

We have seen in Section IV how to carry out this calculation if $c(x)$ is a Gaussian or a random telegraph process. Similar remarks apply also to the calculation of the expected transmitted and reflected powers, and their variances. We do not give the results for the general case, although the calculations are straightforward, since the final expressions are somewhat lengthy.

However, we will consider the case of matched lines, for which no reflections occur. Thus, with $\mu_p = 1$, $\nu_p = 1$, $(p = 1, 2)$, we have, from (97) through (99),

$$T_1 = e^{-\Gamma l} \cos \chi, \qquad T_2 = j\kappa e^{-\Gamma l} \sin \chi. \tag{103}$$

This is for unit input voltage on line 1. Interchanging subscripts, for unit input voltage on line 2 we have

$$T_1 = \frac{j}{\kappa} e^{-\Gamma l} \sin \chi, \qquad T_2 = e^{-\Gamma l} \cos \chi, \tag{104}$$

using (96). Thus, if $\mathcal{V}_1(0) = v_1$ and $\mathcal{V}_2(0) = v_2$, then

$$\begin{aligned} \mathcal{V}_1(l) &= T_1 = e^{-\Gamma l}\left(v_1 \cos \chi + \frac{j}{\kappa} v_2 \sin \chi\right), \\ \mathcal{V}_2(l) &= T_2 = e^{-\Gamma l}(j\kappa v_1 \sin \chi + v_2 \cos \chi). \end{aligned} \tag{105}$$

Note that

$$|\kappa T_1|^2 + |T_2|^2 = e^{-(\Gamma + \Gamma^*)l}(|\kappa v_1|^2 + |v_2|^2). \tag{106}$$

We let

$$c(x) = cN(x), \tag{107}$$

where $N(x)$ is a dimensionless stochastic process with zero mean. Then, from (37) and (102),

$$\chi = c\theta(l). \tag{108}$$

Let us consider the case when $N(x)$ is a Gaussian process. Then, from (52),

$$\langle e^{irx} \rangle = e^{-\frac{1}{2}r^2c^2\sigma^2(l)}, \tag{109}$$

where $\sigma^2(l)$ is given by (51). Hence, from (105),

$$\langle T_1 \rangle = v_1 e^{-\Gamma l} e^{-\frac{1}{2}c^2\sigma^2(l)}, \qquad \langle T_2 \rangle = v_2 e^{-\Gamma l} e^{-\frac{1}{2}c^2\sigma^2(l)}. \tag{110}$$

Also, after some algebra, it is found that

$$\langle | T_1 |^2 \rangle = \frac{e^{-(\Gamma+\Gamma^*)l}}{2 | \kappa |^2} [(| \kappa v_1 |^2 + | v_2 |^2) + (| \kappa v_1 |^2 - | v_2 |^2) e^{-2c^2\sigma^2(l)}], \tag{111}$$

$$\langle | T_2 |^2 \rangle = \tfrac{1}{2} e^{-(\Gamma+\Gamma^*)l} [(| \kappa v_1 |^2 + | v_2 |^2) - (| \kappa v_1 |^2 - | v_2 |^2) e^{-2c^2\sigma^2(l)}], \tag{112}$$

and

$$\begin{aligned} \langle | \kappa T_1 |^4 \rangle - \langle | \kappa T_1 |^2 \rangle^2 &= \langle | T_2 |^4 \rangle - \langle | T_2 |^2 \rangle^2 \\ &= \tfrac{1}{8} e^{-2(\Gamma+\Gamma^*)l} [1 - e^{-4c^2\sigma^2(l)}] \{ | v_2^2 - \kappa^2 v_1^2 |^2 [1 + e^{-4c^2\sigma^2(l)}] \\ &\quad - 2(| v_2 |^2 - | \kappa v_1 |^2)^2 e^{-4c^2\sigma^2(l)} \}. \end{aligned} \tag{113}$$

The first equality in (113) is a consequence of (106). From (111) through (113) it follows that the ratio of the standard deviation to the mean of $|T_p|^2$ approaches

$$\sum_\infty = \frac{| v_2^2 - \kappa^2 v_1^2 |}{2(| \kappa v_1 |^2 + | v_2 |^2)} \tag{114}$$

as $c^2\sigma^2(l) \to \infty$, both for $p = 1$ and $p = 2$.

Analogous results may be obtained when $N(x)$ is the random telegraph process, by using (73), and also in the case of weak general coupling and long sections of the coupled lines, by using (82). We remark that we have previously[2] calculated the average modal powers in these cases for two coupled lines with unequal propagation constants, corresponding to equation (16), using entirely different methods.

VI. TIME DOMAIN STATISTICS

In this section we conclude the study of our models by calculating some of their time domain statistics. Consider first the single line and an incident wave moving to the right in $x < 0$ of the form

$$V(t - \gamma_0 x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{V}(\omega) e^{j\omega(t-\gamma_0 x)}\, d\omega,$$

$$I(t - \gamma_0 x) = \frac{1}{K_0} V(t - \gamma_0 x). \tag{115}$$

Then the transmitted wave in $x > l$ is

$$V_T(t - \gamma_l x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} T(\omega, l)\mathcal{V}(\omega) e^{j\omega(t-\gamma_l(x-l))}\, d\omega,$$

$$I_T(t - \gamma_l x) = \frac{1}{K_l} V_T(t - \gamma_l x), \tag{116}$$

where $T(\omega, l)$ is the transmission coefficient. If we substitute expression (41) for $T(\omega, l)$ into (116) and formally interchange summation and integration, we obtain

$$\begin{aligned} V_T(t - \gamma_l x) &= \lambda \sum_{r=0}^{\infty} \rho^r \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{V}(\omega) \exp\{j\omega[t - \gamma_l(x - l) \\ &\qquad - (2r + 1)\gamma(l + \epsilon\theta(l))]\}\, d\omega \\ &= \lambda \sum_{r=0}^{\infty} \rho^r V(t - \gamma_l(x - l) - (2r + 1)\gamma(l + \epsilon\theta(l))). \end{aligned} \tag{117}$$

Therefore,

$$\langle V_T(t - \gamma_l x)\rangle = \lambda \sum_{r=0}^{\infty} \rho^r \langle V(t - \gamma_l(x - l) - (2r + 1)\gamma(l + \epsilon\theta(l)))\rangle. \tag{118}$$

In this formulation, the randomness appears just as we should expect in a random change in the electrical length of the central transmission line. From (117) it follows that

$$\begin{aligned} &\langle |\, V_T(t - \gamma_l x)\,|^n\rangle \\ &\quad = \lambda^n \left\langle \left| \sum_{r=0}^{\infty} \rho^r V(t - \gamma_l(x - l) - (2r + 1)\gamma(l + \epsilon\theta(l))) \right|^n \right\rangle. \end{aligned} \tag{119}$$

To better understand some of the implications of these formulas,

let $V(u)$ be a rectangular pulse,

$$\begin{aligned} V(u) &= 1, \qquad 0 \leqq u \leqq \tau, \\ &= 0, \qquad u < 0, \tau < u. \end{aligned} \tag{120}$$

Then if $N(x)$ is one of the Gaussian processes discussed in Section IV,

$$\begin{aligned} &\langle V(t - \gamma_l(x - l) - (2r + 1)\gamma(l + \epsilon\theta(l)))\rangle \\ &\quad = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\theta^2/2\sigma^2)} V(t - \gamma_l(x - l) - (2r + 1)\gamma(l + \epsilon\theta))\, d\theta \\ &\quad = \frac{1}{\sqrt{\pi}} \int_{[t-\gamma_l(x-l)-(2r+1)\gamma l-\tau/\sqrt{2}(2r+1)\epsilon\gamma\sigma(l)]}^{[t-\gamma_l(x-l)-(2r+1)\gamma l/\sqrt{2}(2r+1)\epsilon\gamma\sigma(l)]} e^{-\omega^2}\, d\omega \\ &\quad = \frac{1}{2}\left[\operatorname{erfc}\left\{\frac{t - \gamma_l(x - l) - (2r + 1)\gamma l - \tau}{\sqrt{2}\,(2r + 1)\epsilon\gamma\sigma(l)}\right\}\right. \\ &\qquad \left. - \operatorname{erfc}\left\{\frac{t - \gamma_l(x - l) - (2r + 1)\gamma l}{\sqrt{2}\,(2r + 1)\epsilon\gamma\sigma(l)}\right\}\right]. \end{aligned} \tag{121}$$

In (121) erfc (x) is the complementary error function.[11]

Equation (121) shows that the average field violates causality, since at $x = l$, for example, $\langle V(t - (2r + 1)\gamma(l + \epsilon\theta(l)))\rangle$ is positive for all $-\infty < t < \infty$. This is really a consequence of the fact that for a Gaussian process, at any point x, there is a positive probability that a sample function is less than $-\epsilon^{-1}$. Hence, in our model the inductance and capacitance can both become negative, leading to the violation of causality.

If $N(x)$ is the random telegraph process, then $\theta(l)$ has the probability density function given in (71). It follows that in this case

$$\begin{aligned} &\langle V(t - \gamma_l(x - l) - (2r + 1)\gamma(l + \epsilon\theta(l)))\rangle \\ &\quad = \tfrac{1}{2}e^{-bl}\left[V(t - \gamma_l(x - l) - (2r + 1)\gamma l(1 + \epsilon))\right. \\ &\qquad + V(t - \gamma_l(x - l) - (2r + 1)\gamma l(1 - \epsilon)) \\ &\qquad + b\int_{-l}^{l}\left\{I_0[b(l^2 - \theta^2)^{\frac{1}{2}}]\right. \\ &\qquad \left.\left. + \frac{lI_1[b(l^2 - \theta^2)^{\frac{1}{2}}]}{(l^2 - \theta^2)^{\frac{1}{2}}}\right\} V(t - \gamma_l(x - l) - (2r + 1)\gamma(l + \epsilon\theta))\, d\theta\right]. \end{aligned} \tag{122}$$

If $V(u)$ is the rectangular pulse (120) which arrives at $x = 0$ at $t = 0$, then from (118) and (122), the average transmitted signal is a train of nonrectangular pulses. It is seen easily from (122) that

$$\langle V(t - \gamma_1(x - l) - (2r + 1)\gamma(l + \epsilon\theta(l))) \rangle$$

is a pulse which begins at $t = \gamma_1(x - l) + (2r + 1)\gamma l(1 - \epsilon)$ and ends at $t = \gamma_1(x - l) + (2r + 1)\gamma l(1 + \epsilon) + \tau$. This shows that as long as $\epsilon < 1$, causality is preserved for the average transmitted signal since then $\gamma_1(x - l) + \gamma l(1 - \epsilon) > 0$ for all $l \leqq x$.

The duration of the rth average transmitted pulse is $\Delta t = \tau + 2(2r + 1)\gamma l\epsilon$, while the time between the end of the rth pulse and beginning of the $(r + 1)$th pulse is $2\gamma l - 4(r + 1)\gamma l\epsilon - \tau$. Thus, each pulse in the train is longer than the pulse which preceeded it and all of these are longer than the incident pulse. Furthermore, no matter how short the incident pulse is, the average transmitted pulses eventually begin to overlap.

The transmitted power can be treated in the same way. However, due to the overlapping of the pulses, the analysis is tedious and we do not discuss it here.

Analogous results may be obtained in the case of two coupled transmission lines with identical propagation constants, if we consider lossless lines and suppose that the phase constants and the coupling coefficient are proportional to the frequency, so that

$$\Gamma = j\omega\gamma, \qquad c = \omega\gamma\epsilon, \tag{123}$$

and

$$\Gamma_{p0} = j\omega\gamma_{p0}, \qquad \Gamma_{pl} = j\omega\gamma_{pl}, \qquad (p = 1, 2). \tag{124}$$

It is also assumed that the characteristic impedances K_p, K_{p0}, and K_{pl}, $(p = 1, 2)$ are independent of ω. We will confine our attention to the case of matched lines, so that no reflected waves occur.

Let us consider an incident wave on line 1 moving to the right in $x < 0$, and of the form

$$V_1(t - \gamma_{10}x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{V}_1(\omega) e^{j\omega(t - \gamma_{10}x)}\, d\omega,$$

$$I_1(t - \gamma_{10}x) = \frac{1}{K_{10}} V_1(t - \gamma_{10}x). \tag{125}$$

Then, from (93), (94), (96), (103), (108), and (123) through (125), the transmitted waves in $x > l$ are

$$V_{1T}(t - \gamma_{11}x) = K_{11}I_{1T}(t - \gamma_{11}x)$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{V}_1(\omega) e^{-j\omega\gamma l} \cos(\omega\gamma\epsilon\theta(l)) e^{j\omega(t-\gamma_{11}(x-l))}\, d\omega$$
$$= \tfrac{1}{2}\{V_1(t - \gamma_{11}(x - l) - \omega\gamma[l - \epsilon\theta(l)])$$
$$+ V_1(t - \gamma_{11}(x - l) - \omega\gamma[l + \epsilon\theta(l)]\}, \tag{126}$$

and

$$V_{2T}(t - \gamma_{21}x) = K_{21}I_{2T}(t - \gamma_{21}x)$$
$$= \frac{j}{\sqrt{2\pi}} \left(\frac{K_2}{K_1}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} \mathcal{V}_1(\omega) e^{-j\omega\gamma l} \sin(\omega\gamma\epsilon\theta(l)) e^{j\omega(t-\gamma_{21}(x-l))}\, d\omega$$
$$= \frac{1}{2}\left(\frac{K_2}{K_1}\right)^{\frac{1}{2}} \{V_1(t - \gamma_{21}(x - l) - \omega\gamma[l - \epsilon\theta(l)])$$
$$- V_1(t - \gamma_{21}(x - l) - \omega\gamma[l + \epsilon\theta(l)])\}. \tag{127}$$

The transmitted waves in $x > l$ corresponding to an incident wave on line 2 moving to the right in $x < 0$, and of the form

$$V_2(t - \gamma_{20}x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{V}_2(\omega) e^{j\omega(t-\gamma_{20}x)}\, d\omega,$$
$$I_2(t - \gamma_{20}x) = \frac{1}{K_{20}} V_2(t - \gamma_{20}x), \tag{128}$$

are obtained by interchanging the subscripts 1 and 2 in (126) and (127). The transmitted waves corresponding to incident waves on both lines are obtained by linear superposition. The ensemble averages of the transmitted waves may be calculated as before.

APPENDIX

We here apply a limit theorem due to R. Z. Khas'minskii,[3] in order to determine the limiting probability density function for the process $\epsilon\theta(\xi/\epsilon^2)$, for bounded ξ, and $\epsilon \to 0$. Now, from (37),

$$\frac{d}{dx}[\epsilon\theta(x)] = \epsilon N(x), \qquad \epsilon\theta(0) = 0. \tag{129}$$

We assume that $N(x)$ is a bounded, zero-mean, wide-sense stationary stochastic process, with correlation function given by (58). Since θ is a scalar rather than a vector, and moreover the derivative of θ depends only on x, and not on θ, we have just about the simplest nontrivial application of the limit theorem.

There is only one quantity to be considered, namely,

$$a(x, y) = \langle N(x)N(y)\rangle = g(|\, x - y \,|), \tag{130}$$

from (58). Then, in accordance with Khas'minskii's definition,

$$\begin{aligned} \bar{a} &= \lim_{X\to\infty}\left[\frac{1}{X}\int_{x_0}^{X+x_0}\int_{x_0}^{X+x_0} a(x, y)\, dx\, dy\right] \\ &= 2\lim_{X\to\infty}\left[\frac{1}{X}\int_0^X (X - z)g(z)\, dz\right], \end{aligned} \tag{131}$$

after some integrations by parts. As required by the hypotheses of the limit theorem, $\bar{a}$ is independent of x_0. It is also required that the stochastic process $N(x)$ satisfy a certain strong mixing condition, and the reader is referred to Khas'minskii's paper for a precise statement of this condition.

For the case under consideration, the limit theorem states that, on the interval $0 \leqq \xi \leqq \xi_0$, where ξ_0 is an arbitrary positive number, the process $\epsilon\theta(\xi/\epsilon^2)$ converges weakly as $\epsilon \to 0$ to a Markov diffusion process $\Theta(\xi)$ with zero drift and diffusion coefficient $\bar{a}$. The drift coefficient $\bar{K}$ is zero since the right-hand side of eq. (129) is independent of θ. The probability density function of the limit process satisfies the equation

$$\frac{\partial p}{\partial \xi} = \frac{1}{2}\bar{a}\frac{\partial^2 p}{\partial \Theta^2}, \qquad p(\Theta, 0) = \delta(\Theta), \tag{132}$$

in view of the initial condition $\Theta(0) = 0$. Thus,

$$p(\Theta, \xi) = \frac{1}{(2\pi\bar{a}\xi)^{\frac{1}{2}}}\exp\left[\frac{-\Theta^2}{2\bar{a}\xi}\right]. \tag{133}$$

REFERENCES

1. Rowe, H. E., and Young, D. T., "Transmission Distortion in Multimode Random Waveguides," IEEE Trans. Microwave Theory and Techniques, *MTT-20*, No. 6 (June 1972), pp. 349–365.
2. Morrison, J. A., and McKenna, J., "Coupled Line Equations with Random Coupling," B.S.T.J., *51*, No. 1 (January 1972), pp. 209–228.
3. Khas'minskii, R. Z., "A Limit Theorem for the Solutions of Differential Equations with Random Right-Hand Sides," Theory Prob. and Appl., *11*, (1966), pp. 390–406.
4. Kuznetsov, P. I., and Stratonovich, R. L., *The Propagation of Electromagnetic Waves in Multiconductor Transmission Lines*, New York: Macmillan, 1964, p. 91.
5. Wohlers, M. R., *Lumped and Distributed Passive Networks*, New York: Academic Press, 1969, p. 24.
6. Davenport, W. B., and Root, W. L., *Random Signals and Noise*, New York: McGraw-Hill, 1958, p. 155.
7. Davenport, W. B., and Root, W. L., ibid, p. 61.
8. McFadden, J. A., IRE National Convention Record, *7*, part 4, (1959), pp. 164–169.

9. Magnus, W., Oberhettinger, F., and Soni, R. P., *Formulas and Theorems for the Special Functions of Mathematical Physics*, New York: Springer, 1966, p. 411.
10. Foschini, G. J., "Noise and Crosstalk-Interaction of Voltages in a Multipair Environment," unpublished work.
11. Abramowitz, M., and Stegun, I. A., *Handbook of Mathematical Functions*, Washington: National Bureau of Standards, 1964, p. 297.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

The Application of Dither to the Quantization of Speech Signals

By N. S. JAYANT and L. R. RABINER

(Manuscript received February 15, 1972)

By adding a pseudo-random "dither" noise to a signal X that is to be quantized, and by subtracting an identical noise sequence from the quantizer output, it is possible to break up undesirable signal-dependent patterns in the quantization error sequence, without increasing the variance of the error E. The idea has been widely discussed in the context of picture coding, and it is the purpose of this paper to demonstrate application of the technique to the quantization of speech signals. Computer simulations have shown how the use of dither whitens the quantization error sequence in PCM encoding, and renders it more acceptable than signal-correlated errors of equal variance. We demonstrate, for conditions of dither and no dither, typical speech recordings, illustrative error waveforms, and data on signal-to-error correlation C, and indicate how the advantage of dithering increases monotonically with crudeness of signal quantization and becomes significant when the number of bits per sample is less than about six. While the parameter C is a simple criterion for demonstrating the effect of dither, it must be emphasized that the truly relevant criterion is the statistical independence of E and X, and not merely the decorrelation of these functions. Thus, for example, we show that for the case of a reciprocal PDF (probability density function) for X, a zero value of C can be achieved without dither. For purposes of implementation, it is desirable to employ dither noise values characterized by a discrete PDF, with a support that is equal to an integral multiple of the step-size Δ_X in the quantizer. We show that for effective dithering, the step-size Δ_N in the noise PDF need be no smaller, typically, than $\Delta_X/4$. Finally, we indicate an application of dither to the quantization of speech signals by delta modulation.

I. INTRODUCTION

Signal quantizers, in general, produce quantization error sequences that have signal-dependent patterns. The perceptibility of such patterns

tends to be very small for quantizations that are fine enough to provide practically useful signal-to-error ratios; while with relatively cruder quantizations, the perceptibility of signal-dependent errors increases to a point where techniques that can make the errors independent of signal samples become very attractive, even if they do not decrease the error variance itself. Dithering[1] is precisely such a scheme. It is based on the concept of forcing the quantization error E, conditional to a given input X, to be a zero-mean random variable, rather than a deterministic function of X. The randomization of conditional error $E(X)$ is accomplished by the addition of a random dither noise sample N to the input, and quantizing $(X + N)$ instead of X. The use of a pseudo-random dither sample N permits the subtraction of N from the quantizer output $(X + N)_Q$, and this insures an error variance that is essentially no greater than that in the undithered system. Roberts[1] provided an excellent demonstration of the above concept in his pioneering paper on the use of dither for picture coding. Subsequent work on dither[2,3] has also referred to picture signals. Specifically, Limb[2] has studied application to differential quantizers, and Lippel, et al.,[3] have demonstrated the use of two-dimensional, non-random dither patterns the inherent low visibility of which makes dither subtraction from $(X + N)_Q$ irrelevant, perceptually.

The purpose of this paper is to demonstrate the utility of dithering for the quantization of speech signals. We have confined our attention to the use of a pseudo-random dither of the type Roberts[1] employed, but we have considered application to differential quantization also; specifically, to the simplest type thereof, viz, delta modulation.

Section II will describe results from a computer simulation which studied dither for uniform quantizers of the PCM type, and showed that the use of dither whitens the quantization error sequence without increasing its variance, and renders the errors more acceptable than the signal-dependent errors in the undithered system. Results are in the form of speech recordings, error waveforms, and data on the signal-to-error correlation C. These data show how the utility of dithering increases monotonically with crudeness of quantization, and becomes significant for quantizers operating with less than about six bits per sample. The parameter C is a simple criterion for our demonstration, but it is emphasized that the truly relevant criterion is the statistical independence of E and X, and not merely the decorrelation of these functions. In fact, we show that in the example of a reciprocal PDF (probability density function) for X, a zero value of C can be achieved without dither. Section III discusses how, for implementation, it is

desirable to employ dither noise values characterized by a discrete PDF with a support that is equal to an integral multiple of the step-size Δ_X in the quantizer, and shows that for effective dithering, the step-size Δ_N in the noise PDF need be no smaller, typically, than $\Delta_X/4$.

Finally, in Section IV, we indicate a possible application of dither to the quantization of speech signals by delta modulation.

II. DITHERING FOR PCM QUANTIZATION OF SPEECH

Referring to Fig. 1, the input X was speech sampled at the Nyquist rate (6 kHz) and included about 6,000 samples from a 1-second male utterance, "Have you seen Bill?". The dither noise N had a uniform PDF with a zero mean and a range equal to the step-size of the B-bit quantizer:

$$p(N) = \frac{1}{\Delta}\,;\ -\frac{\Delta}{2} < N < \frac{\Delta}{2}\,, \tag{1}$$

$$\Delta = \frac{\text{Peak-to-peak value of } X}{2^B}\,. \tag{2}$$

The uniform quantizer was described by the output-input relation

$$X_Q = \left[\frac{X}{\Delta}\right]\cdot\Delta + \frac{\Delta}{2}\,, \tag{3}$$

where the square brackets denote the "greatest integer in." The input X also has an integral part X_I and a fractional part X_F :

$$X = \left[\frac{X}{\Delta}\right]\cdot\Delta + X_F = X_I + X_F\,;\qquad 0 \leqq X_F < \Delta \tag{4}$$

and the quantization error E_0 , *without dither*, is simply the difference between $\Delta/2$ and X_F :

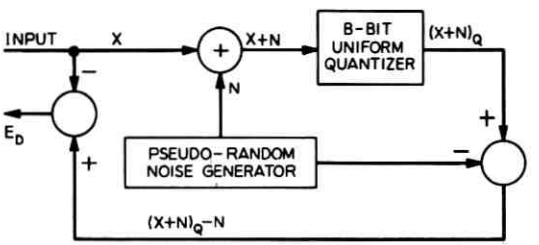


Fig. 1—PCM quantizer with dither.

$$E_0 = X_Q - X = \frac{\Delta}{2} - X_F \,. \tag{5}$$

The signal-independent, random dither noise N has the effect of making the quantizing error *with dither*,

$$E_D = (X + N)_Q - N - X, \tag{6}$$

statistically independent of X; and the subtraction of N from the quantizer output insures that the variance of E is no more (forgetting a correction term for the end steps of the quantizer) than that in the undithered case, which is given by the well-known expression (assuming a uniform distribution)

$$\langle E^2 \rangle = \int_{-\Delta/2}^{\Delta/2} e^2 \cdot p(e)\, de = \int_{-\Delta/2}^{\Delta/2} e^2 \cdot \frac{1}{\Delta}\, de = \frac{\Delta^2}{12}\,, \tag{7}$$

where $\langle \cdot \rangle$ denotes "average value of." The reader is referred to Roberts' paper[1] for a demonstration of the effect of dither on the properties of E, but we will briefly indicate here how E is statistically independent of X.

Let us rewrite eq. (6) in the form

$$E_D = (X + N)_Q - (X + N). \tag{8}$$

Referring to the example in Fig. 2b, one sees that for any given X, the dithered quantizer input $(X + N)$ has a uniform PDF of width Δ, centered around X. In general, a portion of this range (the hatched area) falls outside of the quantizer slot that included X. In view of eqs. (8) and (3), this portion is equivalent, for error calculations, to a corresponding portion (the horizontally striped area) in the quantizer slot including X. In other words, the fractional part (4) of $(X + N)$ can have any value between 0 and Δ, irrespective of the value of X. Hence, the error E_D (8) has the following X-independent distribution:

$$p(E_D/X) = \frac{1}{\Delta}\,; \; -\frac{\Delta}{2} < E_D < \frac{\Delta}{2}\,; \quad \text{any} \quad X \tag{9}$$

while the error E_0 [Fig. 2a and eq. (5)] is a deterministic function of X.

Figure 3 illustrates typical waveforms of E_D and E_0, the quantization errors with and without dither for three illustrative values of B. The following observations emerge:

(*i*) The perceived signal dependence of E_0 is a monotonically decreasing function of B.

(*ii*) The introduction of dither serves to decorrelate the error E_D from the input X even for the worst case of $B = 1$.

(*iii*) A broad similarity between the E_0 and E_D waveforms, attained at $B = 5$, suggests that for larger values of B, the advantage gained by dither tends to become insignificant.

These observations are confirmed by speech recordings of several sentences, which compare, for values of B in the range 1 to 10, the speech output $\{X_0\}$ without dither, and the speech output $\{X_D\}$ with dither. The quantizing error in the latter has an obvious white-noisy nature, while that in the undithered case is perceived to be signal dependent, especially for crude quantizations ($B < 5$ or 6). Recordings of the respective error waveforms E_0 and E_D confirm the point; the E_0 waveforms begin to sound more and more like speech as the quantization gets coarser; and for all values of B, the signal-dependent distortion in X_0 is more degrading than the white-noise in X_D. Incidentally, we have also verified that, for all B, the use of dither has no effect on the error-variance itself, as shown by Roberts.[1]

We have compiled, as a further quantitative description of the effect

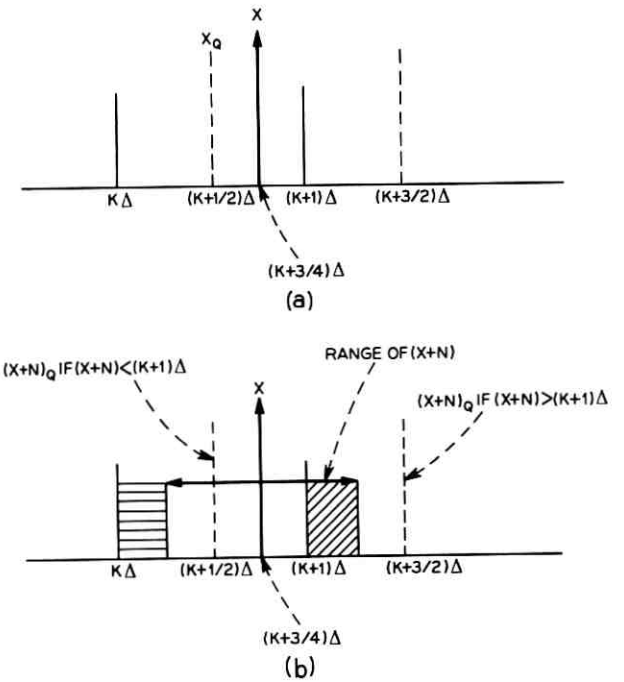


Fig. 2—Illustration of quantization error characteristics (a) without dither, (b) with dither.

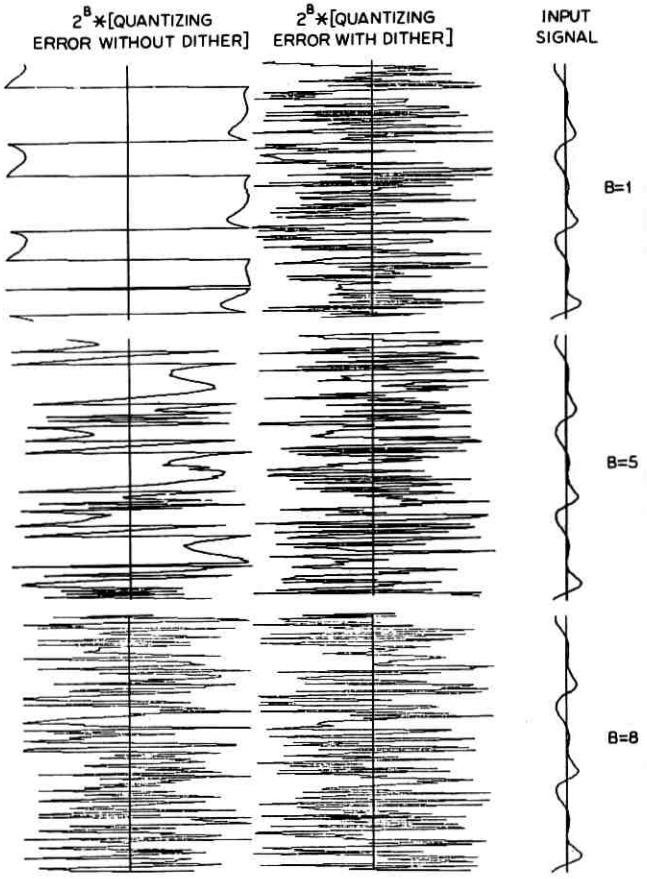


Fig. 3—Waveforms of quantization error.

of dither, values of the signal-to-error correlation (both X and E are assumed to be zero-mean functions)

$$C = \frac{\langle XE \rangle}{\sqrt{\langle X^2 \rangle \langle E^2 \rangle}}. \tag{10}$$

Figure 4 plots C_0 (without dither) and C_D (with dither) as functions of B. It is clear once again, that there is a value of B, say 6, below which the perceptibility of signal dependence in E_0 (as reflected by C_0) is large enough for the decorrelating effect of dither to be significant.

Notice that C_D oscillates, without any obvious structure, in the range

$(-0.01, 0.01)$. Speech recordings mentioned earlier indicate that the ear cannot really resolve colorations corresponding to different values of C in the above range; and, in fact, we believe that a useful criterion for perceptually sufficient signal-to-error decorrelation would be an empirical requirement of the form

$$-0.01 < C < 0.01. \tag{11}$$

We should emphasize, however, that while the correlation measure C is very demonstrative, the truly relevant criterion in question is the statistical independence of E and X, and not merely the decorrelation of these quantities. As a matter of fact, C can be forced to be zero even without dither, as seen in the following example:

Assume that X has a reciprocal PDF. Let us compute, for all values of X in the Kth quantizer slot, the expected value of $X \cdot E_0$:

$$\langle XE_0 \rangle = \int_{K\Delta}^{(K+1)\Delta} \{(K + \tfrac{1}{2})\Delta - X\} \cdot X \cdot p(X)\, dX. \tag{12}$$

Obviously, if $p(X) = 1/X$, $\langle XE_0 \rangle_K$ vanishes, for all K, and, as per eq. (10), C_0 will be zero, without the use of dither! To reiterate, therefore, the idea of using dither is not merely to decorrelate E and X, but to make E statistically independent of X; which, of course, also ensures that C is zero, by definition.

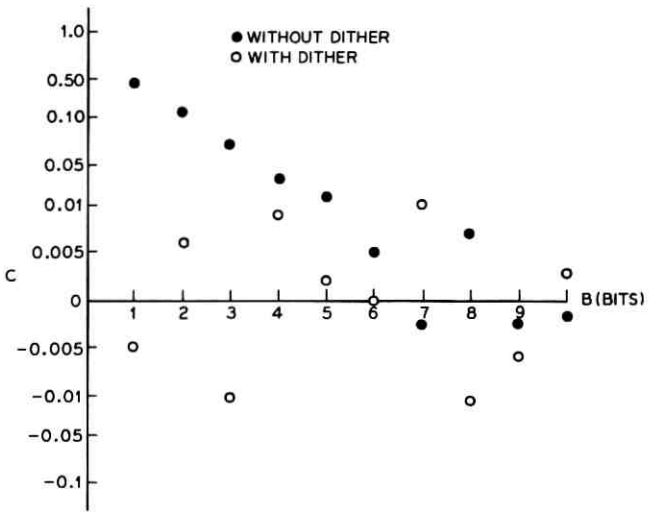


Fig. 4—Signal-to-error correlations.

III. IMPLEMENTATION

The discussion of the previous section has assumed a dither noise with a continuum of sample values uniformly distributed with a zero mean and a support equal to Δ. It is clear, however, that the error randomization or signal-smearing mechanism of Fig. 2 can work even if the support of the noise distribution is equal to an integral multiple of Δ; further, that the noise distribution can be discrete, allowing only a finite number of equiprobable, equally spaced, values; and that for successful dithering, the step-size Δ_N in the noise generator (spacing between consecutive allowed values of N) must be much smaller than the step-size of the signal quantizer itself.

$$\Delta_N \ll \Delta. \tag{13}$$

The above description of the dither noise turns out to be an important one for purposes of practical implementation. Computer simulations were carried out to demonstrate the condition (13) above. Results appear in Fig. 5 which plots the signal-error correlation C_D as a function of (Δ_N/Δ) for different values of B. (The speech input used here was different from that of Fig. 4, but this is immaterial.) Recall now, from (11), that a criterion for perceptually sufficient signal-to-error decorre-

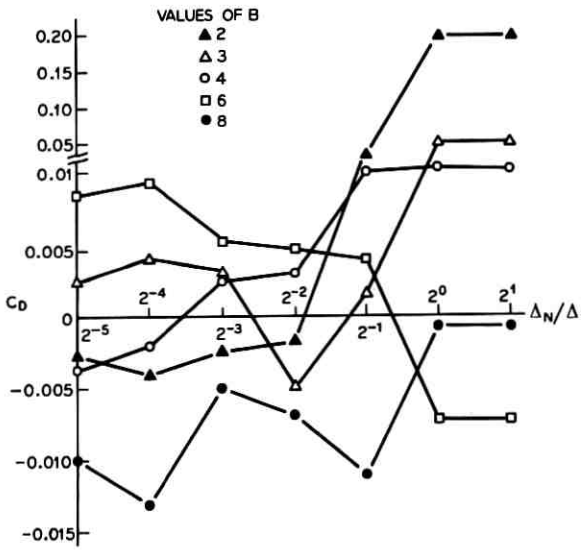


Fig. 5—The effect of discrete dither noise.

lation is the requirement that C lie within the range $(-0.01, 0.01)$. It is hence apparent from Fig. 5 that dither is quite ineffective for relatively fine quantization ($B > 6$) while, for coarse quantization, a safe requirement for achieving the error-smearing effect of dither can be expressed typically in the form

$$\frac{\Delta_N}{\Delta} \gtrsim \frac{1}{4}. \tag{14}$$

It is interesting that (14) is also borne out by previously mentioned literature on picture-coding.[2,3]

Finally, the observation that C_D does not decrease monotonically with Δ/Δ_N is very interesting, especially since it can be shown that there are other important properties of discrete dither which are indeed monotonically related to Δ/Δ_N. For example, it can be shown that the PDF of the quantization error E_D, with discrete dither, is given simply by eq. (9) with a correction term that is inversely proportional to Δ/Δ_N.

IV. THE USE OF DITHER IN DELTA MODULATION

Limb[2] has mentioned the applicability of dither to differential quantizers for picture coding. The differential quantizer that we will discuss here for speech, is a simple one-bit differential quantizer, or a delta modulator (DM); and the dither noise that we will consider is the simple pseudo-random noise considered by Roberts[1] and discussed in Sections II and III.

Figure 6 is a block diagram of a simple delta modulator,[4] which builds a staircase approximation Y to a band-limited input X on the basis of the equation

$$Y_r = Y_{r-1} + \delta_r \operatorname{sgn} (X_r - Y_{r-1}). \tag{15}$$

In other words, each increment in Y follows the direction of the difference between the current value of X_r and the latest staircase approximation to it. With a linear delta modulator (LDM), the step-size δ_r is time-invariant and is tailored to the slope statistics of the input for optimal encoding:[5]

$$\delta_r = \delta_{\mathrm{OPT}}, \tag{16}$$

while in adaptive delta modulation (ADM), δ_r is allowed to follow the slope variations in the input.[5,6] As a result, encoding errors in ADM not only exhibit a smaller variance than in LDM (for a given sampling rate), but are also less dependent on the input signal. The dependence of

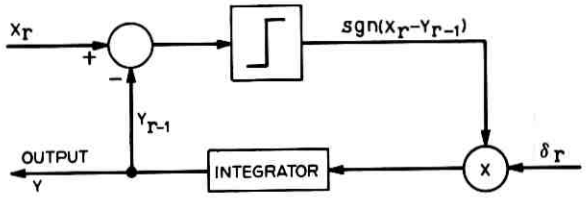


Fig. 6—Block diagram of a linear delta modulator.

encoding error on the input in LDM takes the very specific form of "slope-overload" distortion encountered in the encoding of relatively steep segments of the input. The dither experiment to be mentioned was thus motivated by LDM; and indeed, it proved to be less applicable for ADM.

We ought to emphasize, before proceeding, that for useful encoding, the sampling rate in LDM is at least an order of magnitude times greater than the Nyquist frequency of the input, and the perceptually relevant part of the encoding error in DM is the "in-band" noise as obtained by low-pass-filtering the high-frequency DM noise to the frequency band of the band-limited input. This error is shown as E_D in Fig. 7; here, the input was the utterance, "This is a recording of delta modulated speech," about 3 seconds long, and band-limited to 3.3 kHz. It was sampled for LDM at 60 kHz. The optimum step-size δ_{OPT} was determined in an earlier simulation using signal-to-in-band-error-ratio as a criterion of

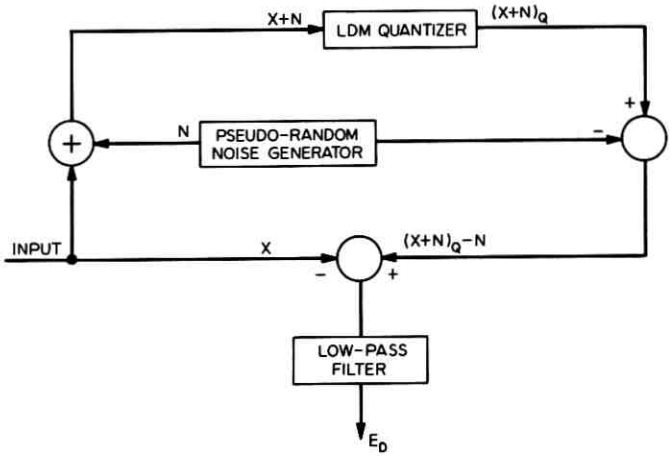


Fig. 7—LDM quantizer with dither.

good encoding. The pseudo-random dither was a zero-mean, uniformly distributed quantity (also sampled at 60 kHz), and had a variable range δ_N:

$$p(N) = \frac{1}{\delta_N}\,;\, \frac{-\delta_N}{2} < N < \frac{\delta_N}{2}. \tag{17}$$

The low-pass filter was a recursive 3-pole filter with a nominal 3.3-kHz cutoff.

Figure 8 plots the signal-to-error ratio SNR as a function of dither range δ_N. One notices a peak SNR advantage at the value

$$\delta_N \sim \delta_{\text{OPT}}\,. \tag{18}$$

In the figure, the case of $\delta_N = 0$ represents the case of no dither.

The SNR advantage due to dither in LDM is significant in that there is no parallel result in PCM. In the quantizers reported by earlier workers, and in Section II, the role of a pseudo-random dither was merely to smear the error sequence, without changing its variance; and this left the SNR unaltered in spite of dither. The same preservation of error variance is expected to hold in LDM, but only with reference to the unfiltered (high-frequency) encoding error. Once again, the role of dither is to smear or whiten this error sequence. But since the unfiltered error in LDM is expected to have considerable low-frequency components (recall the signal-dependent slope-overload distortion), the whitening of this error has the effect of decreasing the error variance within the signal band, hence, the increase in signal-to-error ratio.

A comparison of LDM speech, without dither and with an optimal dither (18), reflects the 2-dB SNR advantage in Fig. 8, and, more obviously, a desirable whitening of the encoding error.

Without going into details, it should be mentioned that the advantage

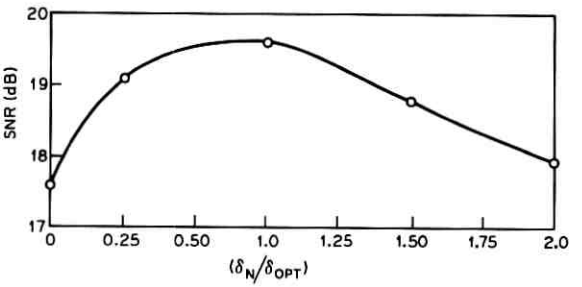


Fig. 8—LDM performance with dither.

of dither was considerably less in evidence when sampling rates of less than 60 kHz were employed, or when the delta modulation was adaptive. This is probably due, in the first case, to a lesser out-of-band noise rejection factor, and, in the second, to the fact that ADM starts out with much lesser signal-to-error dependencies than LDM, and hence, has less to gain from the dither technique.

V. CONCLUSION

The concept of an error-whitening dither noise, utilized so far generally for picture quantization, has been shown to be applicable to the coding of speech signals via PCM and LDM. The demonstrated advantages of dither have considerable practical significance at values of B (bits-per-sample) in the range 4 to 6, for PCM; and typically, for 60-kHz sampling in LDM. The qualities of speech encoding in the two cases are comparable, but they both fall short of toll-quality. However, they still represent a quality range that is obviously quite usable; and the error-whitening property of dither appears to be a very efficient way of enhancing the acceptability of speech in this quality range.

VI. ACKNOWLEDGMENTS

The programming assistance rendered by Mrs. K. Shipley is greatly appreciated. We would also like to thank J. L. Flanagan for suggesting the idea that dithering might be successfully applied to speech.

REFERENCES

1. Roberts, L. G., "Picture Coding Using Pseudo-Random Noise," IRE Transactions on Information Theory, *IT-8*, No. 2 (February 1962), pp. 145–154.
2. Limb, J. O., "Design of Dithered Waveforms For Quantized Visual Signals," B.S.T.J., *48*, No. 7 (September 1969), pp. 2555–2582.
3. Lippel, Z., Kurland, M., and Marsh, A. H., "Ordered Dither Patterns For Coarse Quantization of Pictures," Proc. IEEE (Letters), (March 1971), pp. 429–431.
4. De Yager, F., "Delta Modulation, A Method of PCM Transmission Using a 1-Unit Code," Phillips Res. Rep., No. 6 (December 1952), pp. 442–466.
5. Abate, J. E., "Linear and Adaptive Delta Modulation," Proc. IEEE, *55*, No. 3 (March 1967), pp. 298–307.
6. Jayant, N. S., "Adaptive Delta Modulation With a One-Bit Memory," B.S.T.J., *48*, No. 3 (March 1970), pp. 321–342.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

Queues With Mixed Renewal and Poisson Inputs

By A. KUCZURA

(Manuscript received February 2, 1972)

In a queueing system with two independent input streams, as exists, for example, when first-routed and overflow traffic streams are offered to a common sender-group, the state of the system encountered by the two different types of customers upon their arrival will generally be different. Consequently, in a system where delayed customers wait for service, the service rendered to the individual streams may also be different.

The delay distribution in a single-server queue for each type of customer is derived under the assumption that one stream is Poissonian and the other is described by a renewal process. The difference in service received by the two streams is examined with the aid of numerical examples for two interarrival time distributions of the renewal stream. We show for two cases that a practical indicator of service received by the renewal customers is the coefficient of variation of their interarrival time distribution. If the coefficient is less than unity, then the renewal customers receive better service than the Poissonian customers. The converse is true when the coefficient exceeds unity.

The stationary distribution of the number of busy servers in an infinite-server system as seen by the two types of customers is also derived.

I. INTRODUCTION

The concept of a piecewise Markov process[1] is used to analyze two queueing systems the inputs of which are composed of two independent streams. One of the streams is Poissonian with intensity λ and the other (called a GI stream because of its General Independent Distribution of intervals between arrivals) is assumed to be a renewal process with intensity ν. We assume the service times of all the customers are independent and identically distributed according to an exponential distribution with mean μ^{-1}. Such models are denoted by GI + M/M/c in Kendall's notation, where the "M" refers to the Markovian character

of Poissonian arrivals and of exponential service and c refers to the number of servers.

The state of the system (number of customers waiting and in service) seen by an arriving Poissonian customer will generally differ from that of the GI customer. Consequently, in a system where delayed customers wait for service such as in the GI + M/M/1 queue, the service received by the two types of customers will differ. Whether the GI customers receive better or worse service than the Poissonian customers depends on the variability of the interarrival times of the GI stream.

In Section II, we analyze the GI + M/M/1 queue. The delay distribution with order-of-arrival service for the two types of customers is derived for GI streams the interarrival time distributions of which have rational Laplace–Stieltjes transforms. The coefficient of variation of the interarrival time distribution is introduced as a practical measure of the variability properties of a stream of customers. Its usefulness in predicting service is evaluated through some numerical examples and its relation to the common measure called peakedness is examined. In Section III, in order to describe the intrinsic character of the two streams of customers as they would be observed in a system without delay, we derive the stationary distribution of the number of busy servers seen by the two types of arriving customers in a GI + M/M/∞ system.

II. THE GI + M/M/1 QUEUE

Let $Y(t)$ be the number of customers in the system (those waiting and in service) at time t. Since the GI stream is a renewal process and $Y(t)$ is Markovian between any two consecutive arrival epochs of the GI customers, $\{Y(t), t \geqq 0\}$ is a piecewise Markov process[1] with state space $\{0, 1, 2, \cdots\}$. The regeneration points are the arrival epochs of the GI customers. Thus the distribution of the length of the Markovian segments is given by $A(\xi)$, the interarrival time distribution of the GI customers. The regeneration matrix is given by

$$(p_{ij}) = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ & & \cdots & & \end{bmatrix} \tag{1}$$

The elements of this matrix are the transition probabilities across a regeneration point; that is, p_{ij} is the probability that immediately

after regeneration the process is in state j, given that, immediately prior to regeneration, the process was in state i.

The Markov process operating within the segments is a birth-death process with birth rate λ and death rate μ, the same for all segments. The transition probability functions of this process are given in Ref. 2, page 13.

$$P_{ij}(t) = \rho^{\frac{1}{2}(j-i)}e^{-(\lambda+\mu)t}\Big\{I_{i-j}(2t\sqrt{\lambda\mu}) + \rho^{-\frac{1}{2}}I_{i+j+1}(2t\sqrt{\lambda\mu}) + (1-\rho)\sum_{k=1}^{\infty}\rho^{\frac{1}{2}(k+1)}I_{i+j+k+1}(2t\sqrt{\lambda\mu})\Big\},$$

where $\rho = \lambda/\mu$ $(\lambda/\mu < 1)$ and $I_j(\xi)$ is the modified Bessel function

$$I_j(\xi) = \left(\frac{\xi}{2}\right)^j \sum_{k=0}^{\infty} \frac{1}{k!\,(j+k)!}\left(\frac{\xi}{2}\right)^{2k}.$$

The regeneration matrix (1), the distribution $A(\xi)$, and the transition functions $P_{ij}(t)$ determine the piecewise Markov process completely. We will first derive the distribution of the state of the system seen by an arbitrary GI arrival and then use the rate conservation principle to find the stationary distribution

$$q_i = \lim_{t\to\infty} P\{Y(t) = j \mid Y(0) = i\}, \qquad j = 0, 1, 2, \cdots, \quad \text{for all} \quad i, \tag{2}$$

which is the same as the distribution of the state of the system seen by a Poissonian arrival. Having found the two distributions, we can readily determine the individual delay distributions for an order-of-arrival service discipline.

2.1 *Delay Sustained by GI Customers*

Let $\{p_j\}$ be the stationary distribution of the Markov chain imbedded at points immediately preceding a GI arrival. If r_{ij} is the one-step transition probability from state i to state j, then

$$r_{ij} = \int_0^{\infty} P_{i+1,j}(\xi)\, dA(\xi), \qquad i, j = 0, 1, 2, \cdots.$$

The distribution $\{p_j\}$ satisfies the Chapman–Kolmogorov equations

$$p_j = \sum_{i=0}^{\infty} p_i r_{ij}, \qquad j = 0, 1, 2, \cdots,$$

and hence the solution to the linear system

$$p_j = \int_0^\infty \sum_{i=0}^{\infty} p_i P_{i+1,j}(\xi)\, dA(\xi), \qquad j = 0, 1, 2, \cdots, \tag{3}$$

subject to the normalization condition

$$\sum_{i=0}^{\infty} p_i = 1$$

is the stationary distribution $\{p_i\}$.

Let

$$p(u) = \sum_{i=0}^{\infty} p_i u^i$$

be the probability generating function of $\{p_i\}$. Multiplying (3) by u^j and summing over all j, we obtain

$$p(u) = \int_0^\infty \sum_{i=0}^{\infty} p_i \Gamma_{i+1}(u, \xi)\, dA(\xi), \tag{4}$$

where

$$\Gamma_i(u, \xi) = \sum_{j=0}^{\infty} P_{ij}(\xi) u^j$$

is the probability generating function of $P_{ij}(\xi)$, $j = 0, 1, 2, \cdots$. This function is not readily available. Its Laplace transform

$$\gamma_i(u, s) = \int_0^\infty e^{-s\xi} \Gamma_i(u, \xi)\, d\xi,$$

though, has the following form[2]:

$$\gamma_i(u, s) = \frac{u^{i+1} - (1 - u)\eta^{i+1}(s)/[1 - \eta(s)]}{u[s - h(u)]}$$

where

$$h(u) = \frac{1}{u}(1 - u)(\mu - \lambda u),$$

and

$$\eta(s) = \frac{\lambda + \mu + s - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda}.$$

We will transform the real integral

$$\int_0^\infty \Gamma_{i+1}(u, \xi)\, dA(\xi) \tag{5}$$

appearing in (4) into a complex integral involving $\gamma_{i+1}(u, s)$ and $\alpha(s)$, the Laplace–Stieltjes transform of $A(\xi)$.

If we set

$$dA(\xi) = e^{-a\xi}\, dB(\xi), \qquad a > 0,$$

then (5) becomes the Laplace–Stieltjes transform integral

$$\int_0^\infty e^{-a\xi}\Gamma_{i+1}(u, \xi)\, dB(\xi). \tag{6}$$

But the transform of a product is the complex convolution of the transforms of its factors and so (6) becomes

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \gamma_{i+1}(u, z)\beta(a - z)\, dz,$$

where c is any positive number, $i = \sqrt{-1}$, and

$$\beta(s) = \int_0^\infty e^{-s\xi}\, dB(\xi), \tag{7}$$

the Laplace–Stieltjes transform of $B(\xi)$. But

$$\beta(s) = \int_0^\infty e^{-\xi(s-a)}\, dA(\xi) = \alpha(s - a),$$

and we finally have

$$\int_0^\infty \Gamma_{i+1}(u, \xi)\, dA(\xi) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \gamma_{i+1}(u, z)\alpha(-z)\, dz. \tag{8}$$

Substituting for $\gamma_{i+1}(u, z)$ on the right-hand side of (8), and using the resulting identity in (4), we obtain the following integral equation for $p(u)$:

$$p(u) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \left[\frac{u^2p(u) - (1 - u)H(z)}{u[z - h(u)]}\right]\alpha(-z)\, dz, \tag{9}$$

where

$$H(z) = \frac{\eta^2(z)p[\eta(z)]}{1 - \eta(z)}.$$

We will evaluate the complex integral in (9) for the class of $\alpha(s)$, the members of which are rational functions; but first we note some properties of the integrand which suggest a contour to be used in applying the calculus of residues.

Since $P_{ij}(\xi)$ is a probability function, $\Gamma_i(u, \xi)$ is uniformly convergent

for $|u| \leqq 1$ and $\gamma_i(u, z)$ is holomorphic for $|u| \leqq 1$ and $\Re(z) > 0$, where $\Re(z)$ denotes the real part of z. Hence the bracketed part of the integrand in (9) is holomorphic for $|u| \leqq 1$ and $\Re(z) > 0$ since it is the convergent series

$$\sum_{i=0}^{\infty} p_i \gamma_{i+1}(u, z).$$

Since $A(\xi)$ is a probability distribution function, $\alpha(z)$ is holomorphic for $\Re(z) > 0$. For $\Re(z) < 0$, $\alpha(z)$ may or may not be holomorphic. The predominant case is where $\alpha(z)$ is meromorphic in the half-plane $\Re(z) < 0$ and we shall address ourselves to this case.

Let $-z_1, -z_2, \cdots, -z_n$ be the poles of $\alpha(z)$. Since the poles of $\alpha(z)$ are in the left-half plane, the poles of $\alpha(-z)$ are in the right-half plane. Hence, the integrand in (9) is meromorphic in the right-half plane and we can use the residue theorem to evaluate the integrand over the contour consisting of the line $(c + iR, c - iR)$ and a semicircle of radius R in the right-half plane which connects $c - iR$ with $c + iR$. We choose c and R such that all the poles of $\alpha(-z)$ are interior to this contour.

Since $\alpha(z)$ is meromorphic, we can write

$$\alpha(z) = \frac{Q_m(z)}{(z + z_1)^{k_1}(z + z_2)^{k_2} \cdots (z + z_n)^{k_n}},$$

where k_i is the order of the pole at z_i, $Q_m(z)$ is a polynomial of degree m, and we assume that $m + 1 \leqq k_1 + k_2 + \cdots + k_n$. Two examples are:

(*i*) $A(\xi)$ is the gamma distribution with density

$$a(\xi) = \frac{k\nu}{(k-1)!}(k\nu\xi)^{k-1}e^{-k\nu\xi}, \qquad \xi > 0,$$

and transform

$$\alpha(z) = \left(\frac{k\nu}{k\nu + z}\right)^k; \tag{10}$$

(*ii*) $A(\xi)$ is the mixture of exponentials (hyperexponential)

$$A(\xi) = \sum_{i=1}^{n} a_i(1 - e^{-\nu_i \xi}), \qquad \xi > 0,$$

where $a_i > 0$, $\nu_i > 0$ and $a_1 + a_2 + \cdots + a_n = 1$, and transform

$$\alpha(z) = \sum_{i=1}^{n} \frac{a_i \nu_i}{\nu_i + z}. \tag{11}$$

If we now let $R \to \infty$, then the contour integral along the line $z = c$ tends to the desired limits and it can be shown that the integral along the semicircle tends to zero. Hence, using Cauchy's integral formula with

$$f(z) = \frac{u^2 p(u) - (1 - u)H(z)}{u[z - h(u)]} Q_m(-z),$$

we have

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{(-1)^k f(z)}{(z - z_1)^{k_1}(z - z_2)^{k_2} \cdots (z - z_n)^{k_n}} dz$$

$$= (-1)^{k+1} \sum_{i=1}^{n} \frac{1}{(k_i - 1)!} g_i^{(k_i - 1)}(z_i), \qquad (12)$$

where $k = k_1 + k_2 + \cdots + k_n$ and

$$g_i(z) = \frac{f(z)}{(z - z_1)^{k_1} \cdots (z - z_{i-1})^{k_{i-1}}(z - z_{i+1})^{k_{i+1}} \cdots (z - z_n)^{k_n}}.$$

If all the poles are simple, such as in example (*ii*), then the integral in (12) is equal to

$$(-1)^{n+1} \sum_{i=1}^{n} g_i(z_i). \qquad (13)$$

If z_1 is the only pole and it is of order k_1, such as in example (*i*), then the integral in (12) is equal to

$$\frac{(-1)^{k_1+1}}{(k_1 - 1)!} f^{(k_1 - 1)}(z_1). \qquad (14)$$

Detailed analysis for the case $k = 2$ will be carried out later when we give numerical examples. The case of two simple poles will also be analyzed.

In general, after the integral has been evaluated, and the result substituted into (9), an equation in $p(u)$ results. This equation can be solved for $p(u)$. The values of the unknown function $H(z)$ at the poles z_i and the values of its derivatives may be determined by first applying the normalization condition $p(1) = 1$ and then, since $p(u)$ is holomorphic in $|u| \leqq 1$, forcing the zeros of the numerator in the unit circle to coincide with the zeros of the denominator in the expression for $p(u)$. This procedure will be illustrated later by an example.

We return to the delay distributions. Let W be the delay, or waiting time, from the arrival instant until the beginning of service for an

arbitrary GI customer. The GI delay distribution is defined by

$$D(t) = P[W \leqq t], \quad t \geqq 0.$$

If $D_j(t)$ is the conditional delay distribution, given that the arriving customer finds j other customers in the system, then

$$D(t) = \sum_{j=0}^{\infty} D_j(t) p_j .$$

Obviously, $D_0(t) = 1$. Since service is in order of arrival,

$$D_1(t) = 1 - e^{-\mu t},$$

and $D_j(t)$ for $j > 1$ is the convolution of j identical exponential distributions, each with mean μ^{-1}. If $\delta(s)$ is the Laplace–Stieltjes transform of $D(t)$, then

$$\begin{aligned} \delta(s) &= p_0 + \sum_{j=1}^{\infty} p_j \int_0^{\infty} e^{-st}\, dD_j(t) \\ &= p_0 + \sum_{j=1}^{\infty} p_j \left(\frac{\mu}{s+\mu}\right)^j, \end{aligned}$$

since the transform of a convolution is the product of the transforms. Consequently we have

$$\delta(s) = p\left(\frac{\mu}{s+\mu}\right) .$$

This equation can now be inverted to obtain the delay distribution for the GI customers.

2.2 *Delay Sustained by Poissonian Customers*

Let $E(t)$ be the delay distribution of Poissonian customers; that is, $E(t)$ is the probability that an arbitrary, arriving Poissonian customer will be delayed no more than t units of time. If $\epsilon(s)$ is its Laplace–Stieltjes transform, then, by the same argument as above,

$$\epsilon(s) = q\left(\frac{\mu}{s+\mu}\right) ,$$

where

$$q(u) = \sum_{j=0}^{\infty} q_j u^j$$

is the probability generating function of $\{q_j\}$, the distribution of state seen by an arbitrary, arriving Poissonian customer.

Using the rate conservation principle[1] we can write

$$\nu p_j + \lambda q_j = \mu q_{j+1}, \quad j = 0, 1, 2, \cdots, \tag{15}$$

where the left-hand side is the asymptotic rate of transition from j to $j + 1$ and the right-hand side is the asymptotic rate of transition from $j + 1$ to j. Multiplying both sides of (15) by u^j and summing over all j, we obtain the following relation between $p(u)$ and $q(u)$:

$$\nu p(u) + \lambda q(u) = \frac{\mu}{u}[q(u) - q_0].$$

Thus, if $p(u)$ has been found, then $q(u)$ is given by

$$q(u) = \frac{q_0 + u\sigma p(u)}{1 - \rho u},$$

where $\sigma = \nu/\mu$ and $\rho = \lambda/\mu$. Applying the normalization condition $q(1) = 1$, we get q_0, i.e.,

$$q_0 = 1 - \rho - \sigma,$$

and hence

$$q(u) = \frac{1 - \rho - \sigma + u\sigma p(u)}{1 - \rho u}. \tag{16}$$

In terms of the probabilities themselves, we solve (15) for $\{q_j\}$ and obtain

$$q_j = (1 - \rho - \sigma)\rho^j + \sigma \sum_{i=0}^{j-1} p_i \rho^{j-i-1}, \qquad j = 1, 2, \cdots.$$

2.3 *Some Relations Between the Two Results*

If l_1 is the mean of the distribution $\{p_j\}$ and l_2 is the mean of the distribution $\{q_j\}$, then using (16) and differentiating, we obtain

$$l_2 = \frac{\rho}{1 - \rho} + \frac{\sigma}{1 - \rho}(1 + l_1). \tag{17}$$

Note that the first term on the right-hand side of (17) is the mean number of customers in the system in the M/M/1 queue with traffic intensity ρ. Hence, the second term may be considered to be an increase in the mean number due to the presence of the GI customers.

Since the probability distribution functions $p(u)$ and $q(u)$ are related through (16), the transforms of the delay distributions are related through the following equation:

$$\epsilon(s) = \frac{s + \mu}{s + \mu - \lambda}(1 - \rho - \sigma) + \frac{\nu}{s + \mu - \lambda}\delta(s). \tag{18}$$

Taking the inverse transform of both sides of this equation and rearranging the terms, we obtain

$$E(t) = 1 - \rho e^{-(\mu-\lambda)t} - \sigma\left\{1 + \mu \int_0^t [\rho - D(t - \xi)]e^{-(\mu-\lambda)\xi}\, d\xi\right\}. \tag{19}$$

If τ_1 is the mean delay of the GI customers and τ_2 is the mean delay of the Poissonian customers, then from (18) we have

$$\tau_2 = \frac{\rho}{\mu(1 - \rho)} + \frac{\sigma}{\mu(1 - \rho)}(1 + \mu\tau_1).$$

Again, the second term on the right-hand side of this equation can be thought of as the added mean delay due to the presence of the GI customers.

2.4 *Measures of Variability of a Traffic Stream*

The delay suffered by an arbitrary customer in an input traffic-stream depends on how the customers' arrival epochs are distributed. Roughly, it can be said that the less "variation" in the arrival epochs, the better is the service received by the customers. We discuss two measures of this variation.

Let $A(\xi)$ be the interarrival time distribution of a traffic stream and μ_i the jth moment of $A(\xi)$:

$$\mu_i = \int_0^\infty \xi^i\, dA(\xi), \qquad j = 1, 2, \cdots .$$

We define V, the coefficient of variation of $A(\xi)$, by

$$V = \frac{\sqrt{\mu_2 - \mu_1^2}}{\mu_1}.$$

This measure is dimensionless and depends only on the properties of the stream itself.

Another measure which is used extensively in telephone traffic theory is defined with the aid of an infinite server system. The traffic stream is offered to an infinite number of servers with exponentially distributed service times. The ratio of the variance to the mean of the number of busy servers in statistical equilibrium is taken as a measure of the variation of the traffic stream. This number is called peakedness and is customarily denoted by Z.

We shall classify all traffic streams into two categories: smooth and peaked. A stream is smooth if $V < 1$ and peaked if $1 < V$.* The same dichotomy is effected by the inequalities $Z < 1$ and $1 < Z$; however, Z is a function of the stream's intensity and thus is not as convenient in measuring variability properties. For example, if $\alpha(s)$ is the Laplace–Stieltjes transform of $A(\xi)$, ρ is the stream's traffic intensity, and μ^{-1} is the mean holding time, then it follows (Ref. 3, Chapter 3) that

$$Z = \frac{1}{1 - \alpha(\mu)} - \rho. \tag{20}$$

Using Jensen's inequality,[4] one can show that for a fixed λ, the stream's peakedness (20) attains a minimum whenever $A(\xi)$ is the one-point distribution

$$A(\xi) = \begin{cases} 0, & \xi < \frac{1}{\lambda} \\ 1, & \frac{1}{\lambda} \leqq \xi. \end{cases}$$

In this case,

$$\alpha(s) = e^{-(1/\lambda)s},$$

at $s = \mu$,

$$\alpha(\mu) = e^{-1/\rho}$$

and with this substitution into (20) we see that Z can vary from $\frac{1}{2}$ (when $\rho = \infty$) all the way up to unity (when $\rho = 0$). In contrast, $V = 0$, independent of ρ. We conclude that Z is not a desirable measure of variation for smooth streams. We shall see later that for peaked traffic, Z turns out to be a good measure. Incidentally, the above argument also shows that with exponential holding times the minimum possible value which peakedness can attain is $\frac{1}{2}$.

2.5 *Examples*

We now give two examples: a case with GI being smooth and another with GI being peaked.

Example 1: GI Smooth. Let $\{X_i\}$ be a sequence of independent, exponentially distributed random variables with corresponding means $\{\theta_i\}$. Let the interarrival time distribution of the GI stream be given by the

* A Poissonian stream ($V = 1$) is considered the norm to which the relative properties of smoothness and peakedness are compared.

distribution of the sum

$$Y = \sum_{i=1}^{n} X_i ,$$

where $n > 1$. Then the GI stream is smooth since

$$V = \frac{\sqrt{\theta_1^2 + \theta_2^2 + \cdots + \theta_n^2}}{\theta_1 + \theta_2 + \cdots + \theta_n} < 1. \tag{21}$$

In particular, if all the means are the same, then Y has a gamma distribution and GI is E_n , an Erlangian stream. For this case, we have

$$V = \frac{1}{\sqrt{n}} ,$$

which, moreover, is also the minimum value of (21). For our numerical example, we take $n = 2$.

From (10) we have

$$\alpha(z) = \left(\frac{2\nu}{z + 2\nu}\right)^2$$

and hence (9) becomes

$$p(u) = -\frac{4\nu^2}{u} \frac{d}{dz} \left\{ \frac{u^2 p(u) - (1 - u)H(z)}{z - h(u)} \right\} ,$$

with the derivative being evaluated at $z = 2\nu$. This follows from (14). Carrying out the differentiation, collecting terms, and solving for $p(u)$, we obtain

$$p(u) = \frac{R(u)}{T(u)} , \tag{22}$$

where

$$R(u) = 4\sigma^2\{\lambda\mu H_1 u^2 - [(\lambda - \mu - 2\nu)H_1 - H_2]u + \mu H_1\},$$

$$T(u) = \rho^2 u^3 - (4\sigma^2 + \rho^2 + 2\rho + 4\rho\sigma)u^2 + (2\rho + 4\sigma + 1)u - 1,$$

$$\rho = \frac{\lambda}{\mu} ,$$

$$\sigma = \frac{\nu}{\mu} ,$$

and

$$H_1 = H'(2\nu), \quad H_2 = H(2\nu).$$

It is easy to verify the following inequalities:

$$T(0) < 0, \quad T(1) > 0, \quad T(1/\rho) < 0.$$

Since $\rho < 1$ and $T(u)$ is a polynomial of third degree, these inequalities imply that $T(u)$ has three positive roots with one root being less than unity. Hence, if u_1 , u_2 , u_3 are the roots of $T(u)$, then $0 < u_1 < 1 < u_2 < u_3$. Since $p(u)$ is a holomorphic function of u for $|u| \leqq 1$, the root u_1 must also be a root of $R(u)$. Using this requirement and the normalization condition $p(1) = 1$ we can find H_1 and H_2 .

Omitting the intervening algebra we see that

$$p(u) = \frac{(1 - \rho u_1 u)K}{(1 - \omega_2 u)(1 - \omega_3 u)}, \tag{23}$$

where

$$K = \frac{(1 - \omega_2)(1 - \omega_3)}{(1 - \rho u_1)},$$

and

$$\omega_2 = \frac{1}{u_2}, \qquad \omega_3 = \frac{1}{u_3}.$$

Note that p_0 , the probability of not being delayed at all, is given by K. Recall that for the Poissonian customers this quantity was given by $q_0 = 1 - \rho - \sigma$.

Expanding (23) in powers of u, we obtain

$$p(u) = \frac{K}{(\omega_2 - \omega_3)} \sum_{k=0}^{\infty} [(\omega_2 - \rho u_1)\omega_2^k - (\omega_3 - \rho u_1)\omega_3^k]u^k,$$

and hence, the distribution $\{p_i\}$ is given by

$$p_i = \frac{K}{(\omega_2 - \omega_3)} (\omega_2 - \rho u_1)\omega_2^i - (\omega_3 - \rho u_1)\omega_3^i, \qquad j = 0, 1, 2, \cdots .$$

The mean of this distribution can be computed by differentiation. We have

$$p'(1) = \frac{\omega_2}{1 - \omega_2} + \frac{\omega_3}{1 - \omega_3} - \frac{\rho u_1}{1 - \rho u_1}.$$

The corresponding quantities for the Poissonian stream of customers can easily be obtained using differentiation and relation (16).

If $D(t)$ is the delay distribution for the GI customers, then, as we

already have seen, its Laplace–Stieltjes transform is given by

$$\delta(s) = p\left(\frac{\mu}{s + \mu}\right).$$

From (23) we have

$$\delta(s) = \frac{K(s + \mu)[s + (1 - \rho u_1)\mu]}{[s + (1 - \omega_2)\mu][s + (1 - \omega_3)\mu]}.$$

Inverting this transform, we obtain the delay distribution for the GI customers:

$$D(t) = 1 - A_2 e^{-(1-\omega_2)\mu t} + A_3 e^{-(1-\omega_3)\mu t}, \tag{24}$$

where

$$A_2 = \frac{\omega_2(1 - \omega_3)(\omega_2 - \rho u_1)}{(1 - \rho u_1)(\omega_2 - \omega_3)}$$

$$A_3 = \frac{\omega_3(1 - \omega_2)(\omega_3 - \rho u_1)}{(1 - \rho u_1)(\omega_2 - \omega_3)}.$$

Using (19) we can obtain $E(t)$, the delay distribution for the Poissonian customers. Performing the indicated integration, we have

$$E(t) = 1 - \rho e^{-(1-\rho)\mu t} - \sigma\left[1 + \frac{A_2}{\rho - \omega_2} - \frac{A_3}{\rho - \omega_3}\right] e^{-(1-\rho)\mu t} + \frac{\sigma A_2}{\rho - \omega_2} e^{-(1-\omega_2)\mu t} - \frac{\sigma A_3}{\rho - \omega_3} e^{-(1-\omega_3)\mu t}. \tag{25}$$

Figure 1 shows complementary delay distributions for the GI customers and the Poissonian customers in the E_2 + M/M/1 model. Note that E_2 customers receive better service since they arrive in a smoother stream ($V = 1/\sqrt{2}$). For the Poissonian stream we, of course, have $V = 1$. While significant at low traffic intensities, this advantage diminishes as the traffic intensity increases.

Example 2: GI Peaked. We can generate a peaked traffic stream using the interrupted Poisson process.[5] For a given σ and Z, the interarrival time distribution of such a stream is given by Ref. 5.

$$A(\xi) = k_1(1 - e^{-r_1\xi}) + k_2(1 - e^{-r_2\xi}),$$

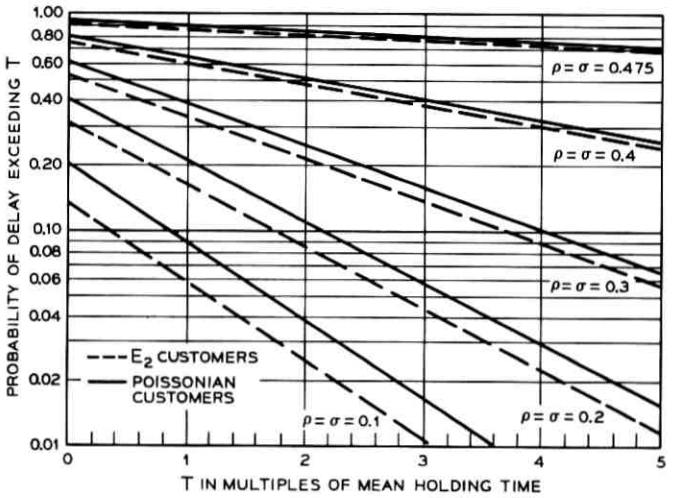


Fig. 1—Delay in the $E_2 + M/M/1$ queue.

where

$$r_1 = \tfrac{1}{2}\{\tau + \omega + \gamma + \sqrt{(\tau + \omega + \gamma)^2 - 4\tau\omega}\},$$

$$r_2 = \tfrac{1}{2}\{\tau + \omega + \gamma - \sqrt{(\tau + \omega + \gamma)^2 - 4\tau\omega}\},$$

$$k_1 = \frac{\tau - r_2}{r_1 - r_2},$$

$$k_2 = 1 - k_1,$$

and

$$\tau = \sigma Z + 3Z(Z - 1),$$

$$\omega = \frac{\sigma}{\tau}\left[\frac{\tau - \sigma}{Z - 1} - 1\right],$$

$$\gamma = \left(\frac{\tau}{\sigma} - 1\right)\omega.$$

The Laplace–Stieltjes transform of $A(\xi)$ is given by

$$\alpha(s) = \frac{k_1 r_1}{s + r_1} + \frac{k_2 r_2}{s + r_2}.$$

Carrying out the same steps as in the previous example, we see that the probability generating function of $\{p_i\}$ has the same form, that is,

$$p(u) = \frac{(1 - \omega_2)(1 - \omega_3)(1 - \rho u_1 u)}{(1 - \omega_2 u)(1 - \omega_3 u)(1 - \rho u_1)},$$

where

$$\omega_2 = \frac{1}{u_2}, \qquad \omega_3 = \frac{1}{u_3},$$

and u_1, u_2, u_3 are now the roots of the following polynomial:

$$\rho(\rho + k_1 r_1 + k_2 r_2)u^3 - [r_1 r_2 + k_1 r_1 + k_2 r_2 + (r_1 + r_2)\rho + 2\rho + \rho^2]u_2 + (r_1 + r_2 + 2\rho + 1)u - 1.$$

Hence, the delay distributions for the GI customers and the Poissonian customers have the same forms as those in Example 1, but with different values of u_1, ω_2, and ω_3.

Figure 2 shows the same information as Fig. 1 of the previous example. Note that the peaked stream with $Z = 3.0$ ($V = 2.3$) receives poorer service than the Poissonian stream. Again, this effect diminishes at higher traffic intensities.

Figure 3 shows the effect of peakedness and smoothness on the quality of service received by all customers. Holding the total traffic intensity constant ($\rho + \sigma = 0.8$) and comparing the results with the case when all the customers arrive in a Poissonian stream (GI = M),

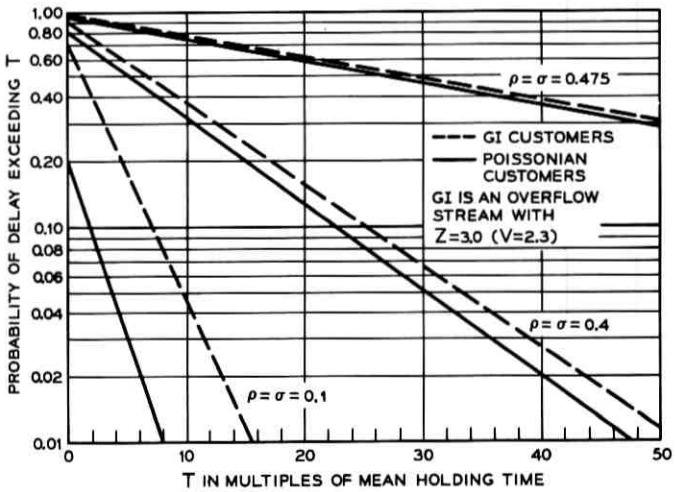


Fig. 2—Delay in the GI + M/M/1 queue when GI is an overflow stream.

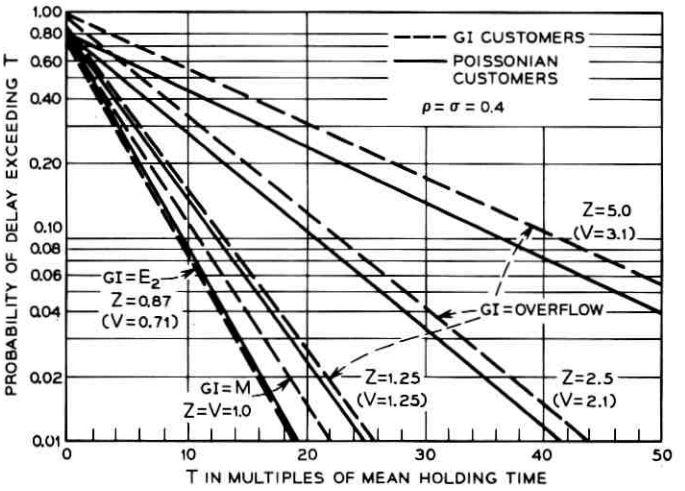


Fig. 3—The effect of peakedness on delay in the GI + M/M/1 queue.

we see that the service provided to *all* the customers either improves or deteriorates according to whether the GI stream is smooth or peaked.

Figure 4 shows the effect of mixing the traffic. Note that as the proportion of peaked traffic in the total offered traffic stream increases, the service deteriorates for everyone.

We now clarify a remark made earlier about the suitability of the peakedness factor Z for measuring variability properties of a stream. It was found numerically that the relation between Z and V in a peaked stream is monotone and nearly independent of the stream's intensity. In fact, by direct computation from $A(\xi)$, we have the two relations

$$Z = \frac{1 + V^2}{2} - \frac{V^2 - 1}{3(1 + V^2) + 2\sigma}$$

and

$$V = \left\{1 + \frac{2(Z - 1)}{1 - 1/(3Z + \sigma)}\right\}^{\frac{1}{2}}.$$

Figure 5 shows this relation for two streams, one with traffic intensity of 1 erlang and the other with 100 erlangs. Since the inverse relation is also monotone and nearly independent of the traffic intensity, the peakedness seems to be a suitable measure of the variational properties of a peaked stream.

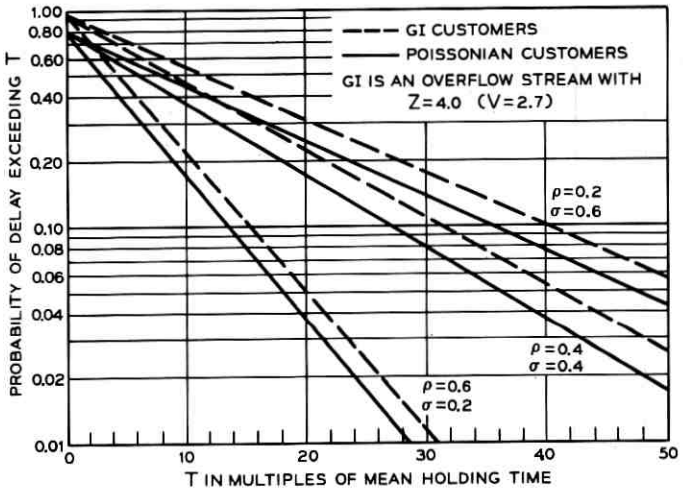


Fig. 4—Effect of mixture of input on delay in the GI + M/M/1 queue.

III. THE GI + M/M/∞ SYSTEM

Let $Y(t)$ be the number of customers in the system at time t. Since there are an infinite number of servers, $Y(t)$ is equal to the number of busy servers at time t. Again, we see that $\{Y(t), t \geqq 0\}$ is a piecewise Markov process. The regeneration points are the arrival epochs of the GI customers; the distribution of the length of Markovian segments is given by $A(\xi)$, the interarrival time distribution of the GI customers; and the regeneration matrix is given by (1). This is the same identification as that made in the GI + M/M/1 queue. The difference here is the Markov process operating within the segment.

The Markovian development of the process within the segments is governed by a birth-death process with birth rate λ, the arrival intensity of the Poissonian stream, and death rate $j\mu$ when the process is in state j, where μ^{-1} is the mean service time. The transition functions of this process are given by the transient solution to the M/M/∞ system (Ref. 2, page 24):

$$P_{ij}(t) = \sum_{k=0}^{j} \binom{i}{j-k} \frac{\rho^k}{k!} e^{-\rho g(t)} [1 - g(t)]^{i-k} g^{i-j+2k}(t), \qquad (26)$$

where

$$g(t) = 1 - e^{-\mu t}, \qquad \rho = \frac{\lambda}{\mu}.$$

Let $\{p_i\}$ be the stationary distribution of the Markov chain imbedded at points immediately preceding a GI arrival. This distribution satisfies the Chapman–Kolmogorov eqs. (3). If $p(u)$ is the probability generating function of $\{p_i\}$, then from (4) we have

$$p(u) = \int_0^\infty \sum_{i=0}^{\infty} p_i G_{i+1}(u, \xi)\, dA(\xi),$$

where

$$G_i(u, \xi) = \sum_{j=0}^{\infty} P_{ij}(\xi) u^j.$$

But from Ref. 2 we have

$$G_i(u, \xi) = (1 - e^{-\mu\xi} + ue^{-\mu\xi})^i e^{-\rho(1-u)a(\xi)},$$

and we see that $p(u)$ satisfies the following integral equation:

$$p(u) = e^{-\rho(1-u)} \int_0^\infty r(u, \xi) p[r(u, \xi)] e^{-\rho(1-u)e^{-\mu\xi}}\, dA(\xi), \tag{27}$$

where

$$r(u, \xi) = 1 - e^{-\mu\xi} + ue^{-\mu\xi}.$$

Now let X_1 be the number of Poissonian customers and X_2 the number of GI customers in the system seen by an arbitrary GI arrival. Since the number of servers is infinite, X_1 and X_2 are independent. If $b(u)$ is the probability generating function of the distribution of X_2, then

$$p(u) = e^{-\rho(1-u)} b(u). \tag{28}$$

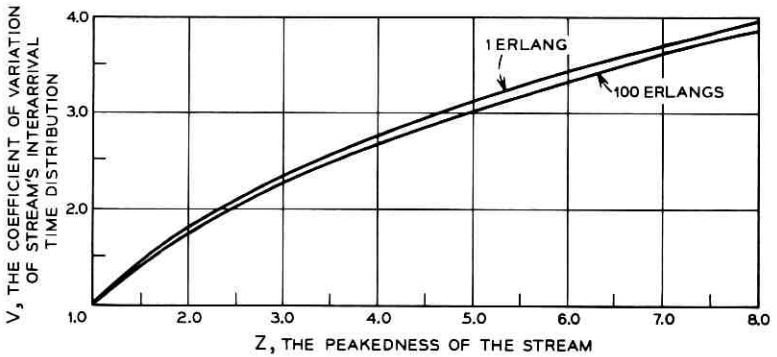


Fig. 5—Relation between V and Z in an overflow stream for two different means.

Since the arriving GI customers are sampling a Markov process according to a renewal process, they will see the stationary distribution of the process[1] and, hence, X_1 has the Poisson distribution

$$P\{X_1 = k\} = e^{-\rho}\frac{\rho^k}{k!}, \tag{29}$$

the stationary distribution of the number of customers in an M/M/∞ system.

Substituting (28) into (27) we obtain an equation for $b(u)$:

$$b(u) = \int_0^\infty r(u, \xi)b[r(u, \xi)]\, dA(\xi). \tag{30}$$

In solving for $b(u)$, it is convenient to consider the expansion of $b(u)$ about $u = 1$, rather than about $u = 0$. Hence, we set

$$b(u) = \sum_{i=0}^{\infty} d_i(u - 1)^i,$$

where

$$d_i = \frac{b^{(i)}(1)}{j!}.$$

If

$$b(u) = \sum_{i=0}^{\infty} b_i u^i$$

is the expansion of $b(u)$ about the origin, then the two sets of coefficients are mutually related through the equations

$$\begin{aligned} d_i &= \sum_{k=i}^{\infty}\binom{k}{j}b_k \\ b_i &= \sum_{k=i}^{\infty}\binom{k}{j}(-1)^{k-i}\, d_k\,. \end{aligned} \tag{31}$$

Differentiating (30) j times and setting $u = 1$, we get

$$\begin{aligned} d_i &= d_i\alpha_i + d_{i-1}\alpha_i\,, \qquad j = 1, 2, \cdots, \\ d_0 &= 1, \end{aligned}$$

where

$$\alpha_i = \alpha(j\mu) = \int_0^\infty e^{-i\mu\xi}\, dA(\xi).$$

Solving for d_i, we obtain

$$d_i = \prod_{i=1}^{i} \frac{\alpha_i}{1 - \alpha_i}, \qquad j = 1, 2, \cdots .$$

Thus, we have

$$p(u) = e^{-\rho(1-u)} \sum_{j=0}^{\infty} d_j(u - 1)^j$$

$$= e^{-\rho(1-u)} \sum_{j=0}^{\infty} b_j u^j,$$

and the distribution $\{p_i\}$ is given by the convolution of $\{b_i\}$ with (29):

$$p_i = e^{-\rho} \sum_{k=0}^{i} b_{i-k} \frac{\rho^k}{k!}.$$

Now

$$q_i = \lim_{t \to \infty} P\{Y(t) = j\}, \qquad j = 0, 1, 2, \cdots ,$$

the distribution of the number of busy servers seen by an arbitrary Poissonian arrival, can be found using the rate conservation principle. If ν^{-1} is the mean of $A(\xi)$ and $\sigma = \nu/\mu$, then equating the asymptotic rate of transition out of the set of states $\{0, 1, 2, \cdots , j\}$ to the rate into that set we get

$$\sigma p_i + \rho q_i = (j + 1)q_{i+1}, \quad j = 0, 1, 2, \cdots . \tag{32}$$

If $q(u)$ is the probability generating function of $\{q_i\}$, then it follows from (32) that $q(u)$ satisfies the differential equation

$$q'(u) = \rho q(u) + \sigma p(u).$$

This equation has the general solution

$$q(u) = e^{\rho u}\sigma \int e^{-\rho\xi} p(\xi)\, d\xi.$$

Substituting for $p(\xi)$, carrying out the integration and using the normalization condition $q(1) = 1$ to determine the integration constant, we obtain

$$q(u) = e^{-\rho(1-u)}\left\{1 - \sigma \sum_{i=0}^{\infty} \frac{b_i}{j + 1} + \sigma \sum_{j=0}^{\infty} b_i \frac{u^{i+1}}{j + 1}\right\}.$$

Since $q(u)$ is the product of two probability generating functions, the distribution $\{q_i\}$ can now easily be determined by convolution. Another

way of obtaining $\{q_i\}$ is by solving (32), i.e.,

$$q_j = \frac{\rho^j}{j!}\left\{q_0 + \sigma \sum_{i=0}^{j-1} \frac{i!\, p_i}{\rho^{i+1}}\right\}, \qquad j = 1, 2, \cdots,$$

$$q_0 = e^{-\rho}\left\{1 - \sigma \sum_{i=0}^{\infty} \frac{\rho^i}{i!} \sum_{i=0}^{i-1} \frac{i!\, p_i}{\rho^{i+1}}\right\}.$$

The means of the two distributions are given by

$$\sum_{j=1}^{\infty} jq_j = q'(1) = \rho + \sigma$$

$$\sum_{j=1}^{\infty} jp_j = p'(1) = \rho + \frac{\alpha(\mu)}{1 - \alpha(\mu)}.$$

We make two observations. First, the mean of $\{q_j\}$ is independent of the form of the interarrival time distribution of the GI customers; it depends only on ν^{-1}, the mean interarrival time. Second, using Jensen's inequality one can show that the mean of $\{p_j\}$ is minimized whenever the interarrival times of the GI customers are constant. For this special case we have

$$p'(1) = \rho + \frac{1}{e^{1/\sigma} - 1},$$

$$q'(1) = \rho + \sigma,$$

and hence

$$p'(1) < q'(1)$$

since

$$1 + \frac{1}{\sigma} < e^{1/\sigma}$$

for $\sigma > 0$.

REFERENCES

1. Kuczura, A., "Piecewise Markov Processes," to appear in SIAM J. Appl. Math.
2. Prabhu, N. V., *Queues and Inventories*, New York: John Wiley & Sons, 1965.
3. Takaćs, L., *Introduction to the Theory of Queues*, New York: Oxford University Press, 1962.
4. Feller, W., *An Introduction to Probability Theory and its Applications*, Vol. 2, New York: John Wiley & Sons, 1966.
5. Kuczura, A., "Loss Systems with Mixed Renewal and Poisson Inputs," unpublished work.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

Optimal Equalization of Wideband Coaxial Cable Channels Using "Bump" Equalizers

By YO-SUNG CHO

(Manuscript received November 8, 1971)

Two methods are described for the optimal equalization of a channel with "Bump" Equalizers composed of several adjustable-gain Bode Networks. The first method is a general one and applies a steepest descent algorithm which minimizes the total mean-squared error (MSE) of the equalized channel. It requires continuous gradient information on the error-gain relationship in order to determine exactly the optimum equalizer adjustments and involves a relatively complicated procedure to calculate the gradient. However, the second method, which also applies a steepest descent algorithm, develops the necessary gradient information with knowledge of the error signal only at selected frequencies across the bandwidth occupied by the channel. Under idealized assumptions, it is shown that the gradients obtained by the second method are exact. When the assumptions do not apply exactly, it is shown by computer simulation that the difference between the gradients obtained by the two methods is very small. A significant potential advantage of the second method lies in the hardware realization which only requires the measurement of the channel error at $2M - 1$ frequencies at the equalizing station (where M is the number of Bode Networks in the equalizer). From these frequency domain errors, the gradients can be generated as real-time signals and applied to the appropriate adjustable elements to obtain the optimum gain settings for minimum MSE.

I. INTRODUCTION

The ideal communication channel exhibits a constant input-output gain characteristic over the entire transmission band. In the case of a 3/8-inch coaxial cable system, the cable loss varies from 4 dB/mile at 1 MHz, to 30 dB/mile at 60 MHz; and to compensate the cable loss, repeaters are required at periodic intervals. As is well known, the cascaded repeaters cannot exactly compensate the cable loss and this

mismatch results in the so-called "misalignment" of the cable system. In addition to the mismatch which is present initially, the channel misalignment is affected by the seasonal temperature variation and the aging of the components in the system. The objective of the main cable equalizers is that, after the equalization, the total input-output gain characteristic of the system should be, at all times, as nearly zero dB as possible over the entire message band.

Since the transmission of amplitude information is of major importance in analog coaxial cable systems, various schemes of input-output "amplitude-only" equalization have been studied in the past.[1,2] It should be noted that the analog signal in the coaxial cable channel may contain either voice or digital information, and the transmission of voice-type information can be accomplished without phase equalization. For the transmission of high-speed digital information, the necessary phase equalization is usually furnished in the digital terminals and not in the main coaxial cable path.

In this paper, the Bump Equalizer,[2] which is an "amplitude-only" equalizer, is studied and a new adjustment method is presented. The Bump Equalizer is composed of a number of Bode Networks,[3] each of which can be controlled independently without affecting the other networks in the set (Fig. 1). A typical Bode Network is shown in Fig. 2a and its transfer function has the form of a bump shape in the frequency domain. In the Bump Equalizer of Fig. 3, several Bode Networks are

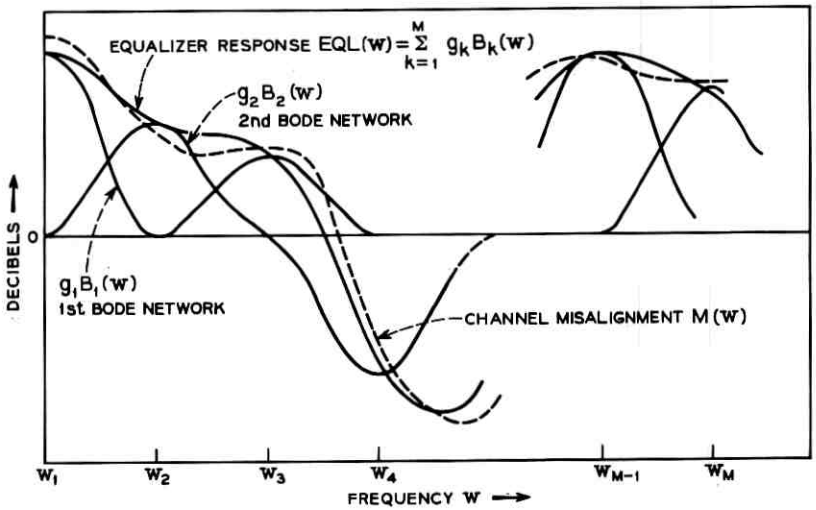


Fig. 1—Channel equalization with Bump Equalizer.

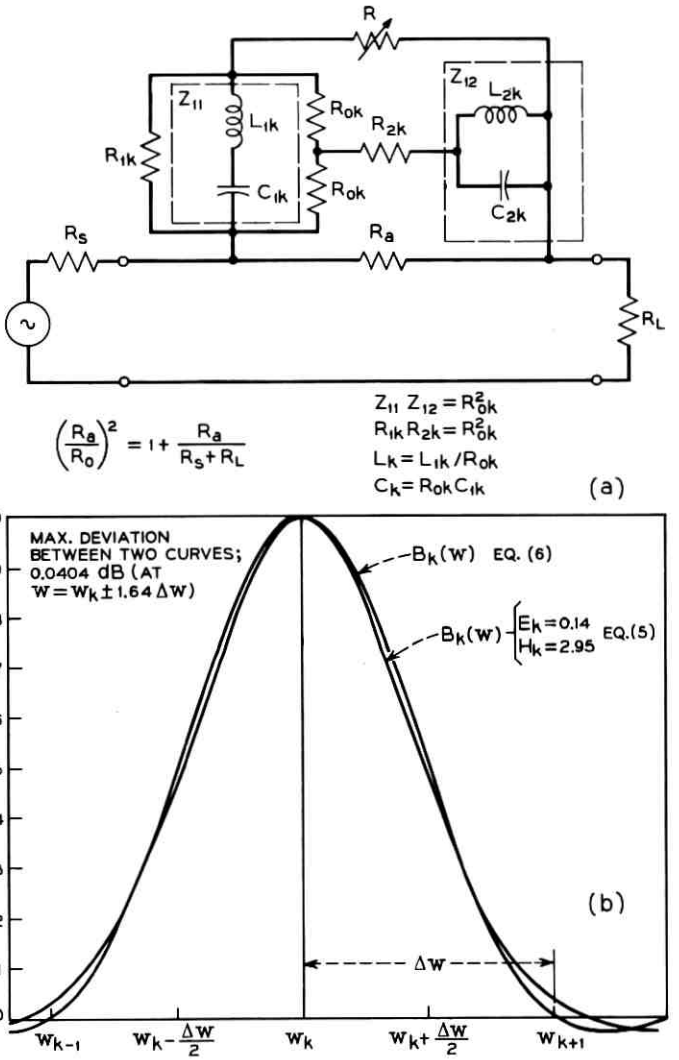


Fig. 2—Bode Network and its input-output transfer function.

connected in the feedback and feedforward paths of linear wideband amplifiers and the transfer function of the equalizer can be expressed by

$$\mathrm{EQL}\,(w) = \sum_{k=1}^{M} g_k B_k(w) \quad (\mathrm{dB}), \tag{1}$$

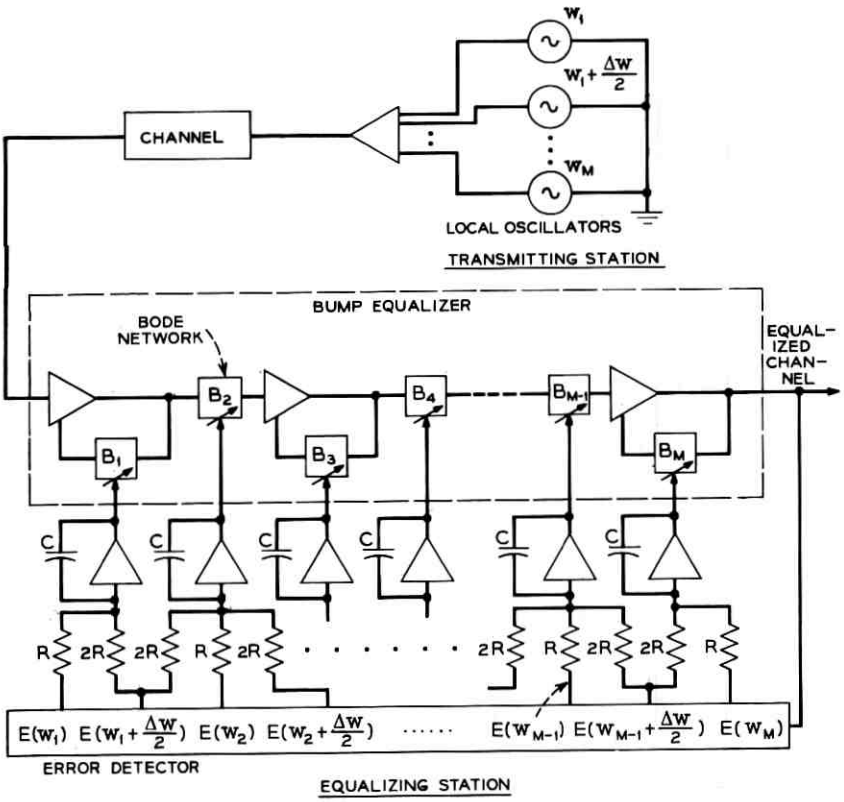


Fig. 3—Equalizer block diagram.

where w indicates the frequency, g_k and B_k represent the gain and response of the kth Bode Network respectively, and M is the number of Bode Networks in the Bump Equalizer.

In the past, the so-called "Zero-Forcing" (ZF) method has been used for the adjustment of Bump Equalizers.[2] Although the ZF method results in zero error at the center frequency of each Bode Network, relatively large errors may exist at other frequencies, and as a result the ZF method may not be optimal in any overall sense. A better error criterion is to minimize the mean-squared error (MSE) over the entire bandwidth, and it is this error criterion which is used in this paper for equalizer gain adjustment.

In Section II, the channel and the transfer function of the equalizer

are characterized, and several assumptions are made before deriving the algorithms which will minimize the MSE of the equalized channel. Before the steepest descent method[4] is used to obtain the minimum MSE, it is discussed in Section III that the MSE of the equalizer has a unique minimum point (no local minima) in the gain parameter space of the equalizer, so that this method is assured of finding the true minimum MSE at all times. It is shown that under certain conditions, the gradient of the MSE, with respect to each gain, can be obtained as a real-time signal by measuring the error at only three points in the frequency domain. The gradient signal derived can be applied to an integrator to adjust the appropriate equalizer gain setting until the gradient becomes zero, and the desired optimization is achieved. This equalization method is called the "simplified MSE" algorithm. Hardware implementation of the algorithm is also discussed in Section III.

Various computer simulations have been carried out and some of the results are discussed in Section IV. Both the conventional expression for the transfer function of a Bode Network and the measured transfer function of a physically realized network have been used in the simulation to verify the effectiveness of the derived algorithm when used in practical applications. One of the channel misalignments used in the simulation resulted from measurements on a working coaxial cable system in the field.

The general steepest descent MSE algorithm is applied to all cases to obtain the absolute minimum MSE for each case; and the resulting values are compared with the MSE obtained by the simplified MSE algorithm. The computer results verify that the simplified algorithm derived under the idealized conditions is, in fact, sufficiently close to the general algorithm in each case so that the former, which permits simplified hardware implementation, can be used as an effective means to achieve optimal control of the Bump Equalizer.

II. CHARACTERIZATION OF CHANNEL AND BUMP EQUALIZER

The coaxial channel is discussed in this paper principally with respect to analog signal transmission. Due to the characteristic of the coaxial cable, the bandwidth of the transmitted signal can be quite wide. The objective of the equalization discussed is to achieve a constant input-output gain characteristic over the entire message band, and the transfer function derived is concerned only with the amplitude characteristic and not the phase characteristic. Since the transfer

function of the Bode Network is symmetric on the "log f" plane, the transformed frequency w, to be used in this section is given by *log f* where f is the natural frequency in Hertz.

2.1 *Characterization of Coaxial Communication Channel*

Let $M(w, t)$ represent the time-varying channel misalignment which is a real-valued function of frequency in units of dB. From the practical point of view, however, the channel can be represented as simply $M(w)$, since the time-variance can be assumed negligible during the interval of any equalization process. Assume also that the Fourier transform of the channel misalignment $M(w)$ is limited by a positive constant, since the Bump Equalizer to be used is strictly a frequency domain equalizer.* Hence, the channel can be characterized in the frequency domain by the following series:

$$M(w) = \sum_{n=0}^{\infty} C_n \frac{\sin (2\pi p_N(w - w_n))}{2\pi p_N(w - w_n)} \quad \text{(dB)}, \tag{2}$$

where C_n , p_N , and w_n are certain real numbers, $w = \log f$, and $w_{n+1} - w_n = 1/2p_N$ for all $n = 0, 1, \cdots$.

Eq. (2) also may be expressed as

$$\begin{aligned} M(w) &= \int_0^1 \sum_{n=0}^{\infty} C_n \cos (2\pi p_N(w - w_n)x)\, dx \\ &= \int_0^1 \Big\{ \sum_{n=0}^{\infty} C_n \cos (2\pi p_N w_n x) \cos (2\pi p_N w x) \\ &\quad + \sum_{n=0}^{\infty} C_n \sin (2\pi p_N w_n x) \sin (2\pi p_N w x) \Big\}\, dx \\ &= \int_0^1 \{F(x) \cos (2\pi p_N w x) + H(x) \sin (2\pi p_N w x)\}\, dx, \end{aligned} \tag{3}$$

where

$$F(x) = \sum_{n=0}^{\infty} C_n \cos (2\pi p_N w_n x) \quad \text{and} \quad H(x) = \sum_{n=0}^{\infty} C_n \sin (2\pi p_N w_n x).$$

Since $0 \leqq x \leqq 1$, eq. (3) implies that the shortest frequency domain ripple period found in the channel $M(w)$ is $1/p_N$.

* It should be noted that the Fourier transform of $M(w)$ does not result in an impulse response of the channel because of "dB" dimension of $M(w)$. For this assumption, however, there is an implicit dual relationship with time domain equalizers, e.g., the tapped delay line equalizers, in which a frequency band limitation of the channel is assumed.

2.2 *Characterization of the Bump Equalizer*

As briefly discussed in the introduction, the frequency domain response of the equalizer can be written as

$$\text{EQL}\ (w) = \sum_{k=1}^{M} g_k B_k(w) \quad \text{(dB)}, \tag{4}$$

where M is the number of Bode Networks in the Bump Equalizer.

A typical Bode Network is shown in Fig. 2 and its loss is controlled by the resistor R. The transfer function, $B_k(w)$, can be analytically derived and, with a suitable flat-gain amplifier, can be expressed by the following equation:

$$B_k(w) = \frac{[E_k(1 + E_k) + D_k(w)]^2 - D_k(w)}{[(1 + E_k)^2 + D_k(w)]^2} \quad \text{(dB)}, \tag{5}$$

where

$$E_k = \frac{R_{0k}}{R_{1k}},$$

$$D_k(w) = \frac{(w/w_k)H_k}{(w/w_k)^2 - 1},$$

$$H_k = \frac{C_k}{R_{0k}^2 L_k},$$

and

$$w_k = \log\ (1/2\pi\sqrt{L_k C_k}).$$

Since $B_k(w)$ of eq. (5) is a quite complicated function of w, the following assumption is made before analyzing the equalizer in detail.

Assumption 1: Let $B_k(w)$ be approximated by

$$\text{cosinc}\left(\frac{\pi}{\Delta w}(w - w_k)\right) = \frac{\sin\frac{\pi}{\Delta w}(w - w_k)}{\frac{\pi}{\Delta w}(w - w_k)} \cdot \frac{\cos\frac{\pi}{\Delta w}(w - w_k)}{1 - 4\left(\frac{w - w_k}{\Delta w}\right)^2} \quad \text{(dB)}. \tag{6}$$

Moreover, if there are M Bode Networks in the equalizer, and if they are spaced equally on the w scale at intervals Δw,* such that $\Delta w = w_{k+1} - w_k$ for all $k = 1, \cdots M - 1, M$, then the transfer function of

* Usually the number of Bode Networks, M, is determined from the practical consideration of equalization objectives. p_N defined in eq. (2) determines the Δw which is the interval between two adjacent Bode Networks.

the equalizer can be expressed by

$$\text{EQL}\,(w) = \sum_{k=1}^{M} g_k \operatorname{cosinc}\left(\frac{\pi}{\Delta w}(w - w_k)\right) \quad \text{(dB)}. \tag{7}$$

Equations (5) and (6) are plotted in Fig. 2b; and it can be seen that cosinc $(\pi/\Delta w(w - w_k))$ approximates the actual transfer function of Bode Network given in eq. (5) reasonably well. The maximum difference between the two curves of Fig. 2b is 0.0327 dB when $| w - w_k | \leqq \Delta w$ and 0.0404 dB when $| w - w_k | > \Delta w$.

III. GAIN OPTIMIZATION USING MEAN-SQUARED ERROR CRITERION

After a Bump Equalizer has been physically realized and connected to the channel, the Δw, which is characteristic of the particular set of Bode Networks, cannot be easily altered in the equalizer even though the channel misalignment $M(w)$ (and hence p_N) may vary. The optimization here consists of determining the gain parameters g_k which will minimize the value of MSE defined in this section. One approach to the optimization is the employment of the steepest descent method.[4] In seeking the minimum MSE by this method, the present values of g_k's are changed by small amounts in the opposite direction of gradients which are the partial derivatives of MSE with respect to each gain parameter g_k . The process is continued until all the gradients of the MSE with respect to the gains g_k reach zero or a stationary point. Hence, it is implicit in the use of the steepest descent method that the surface of MSE in the gain parameter space is a bowl shape, and that there exists a unique stationary point which is the global minimum. The unique existence of such a stationary point is established before a general steepest descent algorithm is derived for the Bump Equalizer; then a simplified algorithm is obtained which is shown to be equivalent to the general algorithm. Finally, hardware implementation of the simplified algorithm is discussed.

3.1 *General Mean-Squared Error Algorithm*

On the dB scale, the residual error after equalization will be

$$E(w) = \sum_{k=1}^{M} g_k B_k(w) - M(w) \quad \text{(dB)}. \tag{8}$$

If $C(w)$ is the channel characteristic, the channel misalignment is defined by $M(w) = -C(w)$ and the MSE can be represented in the frequency domain by

$$\text{MSE} = \int_{-\infty}^{\infty} E(w)^2 \, dw. \tag{9}$$

It should be noted that the definition in (9) is a bounded functional, since $E(w)$ can be made small as $w > w_M$, where w_M is the upper limit of the band to be equalized and, moreover, there is complete freedom for the assumption of $M(w)$ for $w > w_M$.

Let G_k be a gradient of MSE with respect to gain g_k. Then, G_k is obtained by differentiating MSE with respect to g_k:

$$G_k = \frac{\partial(\text{MSE})}{\partial g_k} = 2 \int_{-\infty}^{\infty} B_k(w) E(w) \, dw. \tag{10}$$

This may be stated as: the gradient G_k with respect to gain g_k is found by cross-correlating the Bode Network function, $B_k(w)$, and the error function, $E(w)$. The cross-correlation method to obtain the gradient in practice has been used elsewhere and can be found in Refs. 5 and 6.

Theorem 1 (General MSE Algorithm): Let $G_{k,j}$ be the gradient of MSE with respect to the kth gain, g_k, measured at time $t = j$ (also, let $g_{k,j}$ indicate the value of the kth gain at time $t = j$). For the Bump Equalizer, the next gain setting of g_k (denoted by $g_{k,j+1}$) which will reduce the MSE is a function of the gain setting $g_{k,j}$ and the gradient $G_{k,j}$, as given by

$$g_{k,j+1} = g_{k,j} - \Delta c G_{k,j} \tag{11}$$

for all $k = 1, 2 \cdots M$ where Δc is a small positive constant. As the iterative process described by eq. (11) is continued, the gradients $G_k \to 0$ and the equalizer reaches the optimum state.

Proof: Since an equalizer described by (4) is composed of linearly independent networks, there exists a unique set of g_k's which satisfies $G_k = 0$ for all $k = 1, 2, \cdots M$, and this set of g_k's results in the minimum MSE defined in eq. (9) (see Chapter 2 of Ref. 7). Hence, a steepest descent algorithm described by (11) must bring the gains to this optimum stationary point. A general theory on the steepest descent algorithm is given in Ref. 4.

Since no specific assumptions were made on the channel and the Bump Equalizer, the gradient required for the optimum gain adjustment may be obtained by eq. (10).

3.2 *Simplified MSE Algorithm and Hardware Implementation*

In the general steepest descent method described by Theorem 1, the gradient G_k is obtained by the cross-correlation of the error $E(w)$

and the Bode Network function $B_k(w)$. For the hardware implementation of the gradient calculation, the error function is multiplied by the Bode Network response $B_k(w)$ and the product is integrated in the frequency domain. Computation of frequency domain cross-correlation should be done on a real-time basis, and this fact may prohibit the practical application of Theorem 1. However, the following theorem shows that the gradient G_k can be obtained by measuring $E(w)$ only at three different frequencies for each of the given Bode Networks $B_k(w)$.

Theorem 2 (Simplified MSE Algorithm): Let the Bump Equalizer satisfy Assumption 1 and let the interval Δw between adjacent Bode Networks be no greater than the shortest frequency domain ripple period in the channel shown in (3), i.e.,

$$\Delta w \leqq \frac{1}{p_N}. \tag{12}$$

*Then, the optimum gain setting of the k*th *Bode Network is obtained by repeating the following process:*

$$g_{k,j+1} = g_{k,j} - \Delta c\left\{\frac{1}{2}E_j\left(w_k - \frac{\Delta w}{2}\right) + E_j(w_k) + \frac{1}{2}E_j\left(w_k + \frac{\Delta w}{2}\right)\right\}, \tag{13}$$

where

$k = 1, 2, 3, \cdots M$,
Δc *is a small positive constant, and*
$E_j(w_k - \Delta w/2)$, $E_j(w_k)$, *and* $E_j(w_k + \Delta w/2)$ *are the frequency domain errors measured at time* $t = j$ *at* $w = w_k - \Delta w/2$, $w = w_k$, *and* $w_k + \Delta w/2$ *respectively.*

Proof: The proof is given in the Appendix.

In the derivation of the results stated in Theorem 2, it is specified that the channel has a shortest frequency domain ripple period $1/p_N$ which is greater than, or equal to, the interval Δw between two adjacent Bode Networks. If the channel has ripples having components of shorter periods than Δw, then the gradient shown in Theorem 2 only approximates the true gradient which would be obtained by the cross-correlation technique in eq. (10). However, the accuracy of the approximation depends on the amplitude of ripples whose periods are shorter than Δw, and if there exist such ripples of large amplitude, one cannot expect a satisfactory equalization even if a general algorithm is employed.

The feedback equation for the optimum gain adjustment using the simplified MSE technique is as follows:

$$g_k(T) = g_k(0) - \Delta c \cdot \int_0^T \left\{ \frac{1}{2} E\left(w_k - \frac{\Delta w}{2}\right) + E(w_k) + \frac{1}{2} E\left(w_k + \frac{\Delta w}{2}\right) \right\} dt, \qquad (14)$$

where $k = 1, 2, 3, \cdots M$, and $g_k(0)$ is an initial value of gain g_k. The hardware implementation of (14) can be achieved in various ways. In the block diagram shown in Fig. 3, the input signal source at the transmitting station is composed of $2M - 1$ unit-amplitude sinusoidal oscillators whose frequencies are w_1, $w_1 + \Delta w/2$, w_2, $w_2 + \Delta w/2$, $\cdots$, $w_M - \Delta w/2$, and w_M. In the block diagram, the band to be equalized extends from $w = w_1$ to $w = w_M$. [It is assumed that $E(w_1 - \Delta w/2)$ and $E(w_M + \Delta w/2)$ are zero.] It is possible that in some cases the frequencies of the oscillators could be located in the guard bands of the channel such that interference between the oscillators and the message is avoided and the process could be carried out "in service," i.e., in the presence of a "live" message load. At the equalizing station, the gradient can be generated by adding the errors with the weighting indicated in Theorem 2. Now this gradient, which is a real-time signal, is fed to the integrator the output of which can be used to adjust the corresponding gain until the optimum condition is reached with respect to MSE. In Fig. 3, the function of addition and integration is combined by using operational amplifiers.

IV. RESULTS OF COMPUTER SIMULATION

In the previous section, two algorithms were derived to obtain the MSE optimization. The general algorithm (Theorem 1) can be applied for the adjustment of a Bump Equalizer to obtain the minimum MSE, but a complex hardware implementation of this scheme may prohibit its practical application. Consequently, a simplified algorithm which is relatively simple to implement and equivalent to the general one was derived. To demonstrate the equivalence of the two algorithms, the following two conditions were assumed:

(*i*) The channel is represented by a sin x/x series in the frequency domain.

(*ii*) The Bode Network transfer function $B_k(w)$ can be approximated as in Assumption 1.

In practice, these restrictions on the channel and the equalizer are

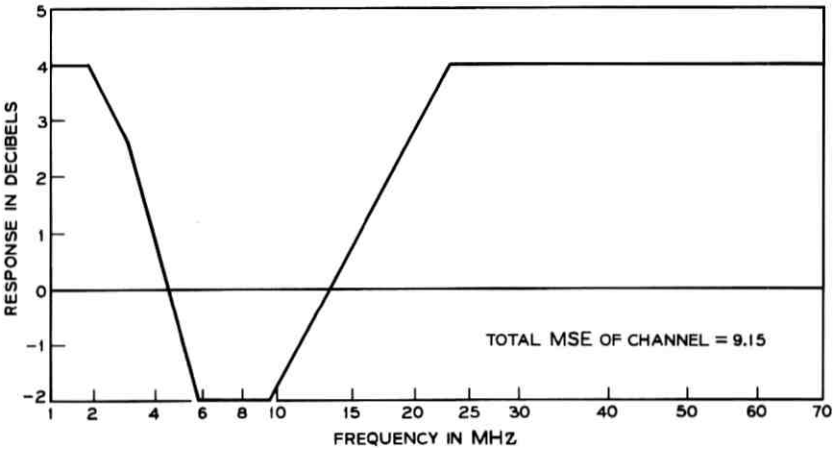


Fig. 4—Channel misalignment, case 1.

not met exactly. Bode Networks spaced equally in w, for example, do not necessarily give the best results. In fact, it has been shown experimentally that, when the number of Bode Networks is limited, unequally spaced Bode Networks are likely to achieve better equalization.[2] Hence, to verify the effectiveness of the new algorithm, the practical limitations were simulated on a digital computer, and the resulting performance compared for the two algorithms.

Case 1: The assumed channel misalignment $M(w)$ is shown in Fig. 4, over the natural frequency range from $f = 1$ MHz to $f = 65$ MHz. Transforming the lower and upper ends of the natural frequency band to the logarithmic scale, such that the message band extends from $w = 0$ $(= \log_e 1)$ to $w = 4.1744$ $(= \log_e 65)$, ten Bode Networks are specified and spaced equally on the w-axis. The transfer function of a physically realized Bode Network was measured and stored in the computer for this simulation.

The results of the simulation are shown in Fig. 5. The total MSE of $M(w)$ within the message band is 9.115. The values of MSE obtained by the simplified algorithm and the general algorithm are 0.45 and 0.42, respectively, with the difference resulting mainly from the following;

(*i*) Since there are sharp corners in $M(w)$, there exist some ripples whose frequency domain periods are shorter than $1/p_N$ [contradiction to the inequality (12)].

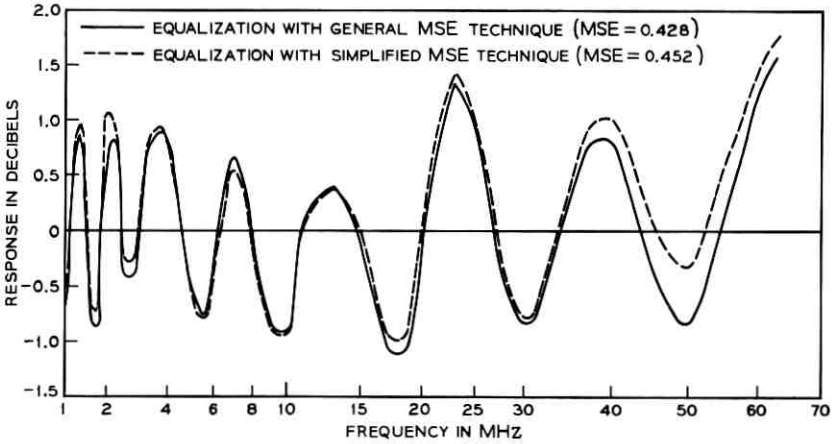


Fig. 5—Equalized channel, case 1.

(*ii*) The measured transfer function of the actual Bode Networks used in the simulation differs slightly from the cosinc function of Assumption 1.

Case 2: The channel misalignment $M(w)$ used is one actually measured on an existing 20-MHz coaxial cable system, with the bandwidth

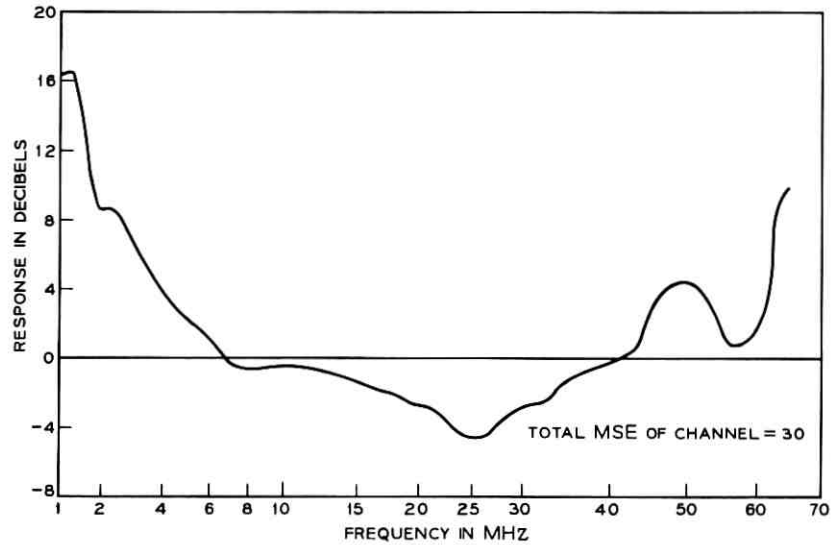


Fig. 6—Channel misalignment $M(w)$, case 2.

arbitrarily extended to 65 MHz. The misalignment $M(w)$ used in this case is shown in Fig. 6. Ten Bode Networks represented by eq. (5) are used in the equalizer, and, in this case, the center frequency of each Bode Network is initially optimized for the specified $M(w)$. Consequently, the resultant array is not spaced equally on the w-axis. To apply the simplified MSE algorithm, however, the errors are measured at 19 frequencies, 10 of which are the center frequencies of Bode Networks and 9 fall between the center frequencies. Total MSE of $M(w)$ is 30. Applying the simplified MSE algorithm, MSE = 0.201 is obtained for the equalized channel and this result is shown in Fig. 7. When the general MSE algorithm is used, the absolute minimum MSE = 0.186 is obtained. The equalized channel with the general MSE algorithm applied is also shown in Fig. 7.

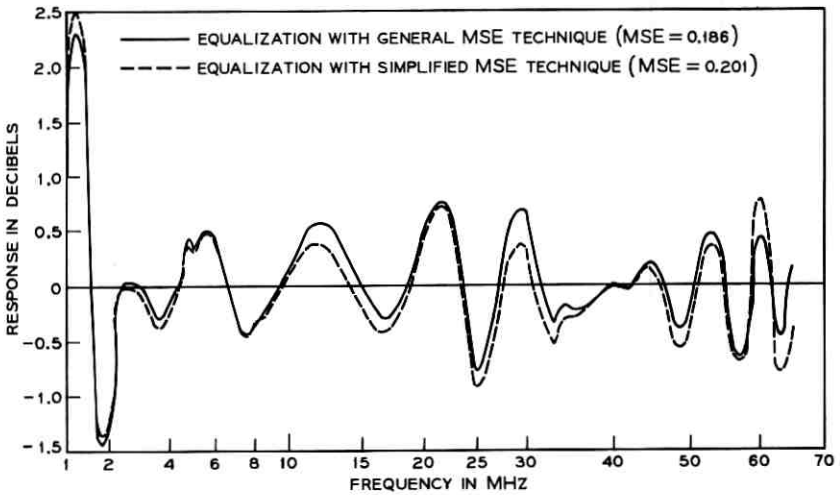


Fig. 7—Equalized channel, case 2.

V. CONCLUSIONS

Two algorithms based on the steepest descent method are presented in this paper for the optimal gain control of a Bump Equalizer. In both cases, the performance index used to evaluate the equalized channel is the MSE. The first algorithm is a general MSE algorithm and requires MSE gradient information with respect to each gain. The required gradients are obtained by a frequency domain cross-correlation between the error and the Bode Network to be adjusted. For this algorithm to be used, the error signal must be known at all

frequencies and this requirement can be a difficult one to implement physically. The simplified MSE algorithm derived in Section III, however, needs error information at only $2M - 1$ frequencies to form all of the gradients. Hence, the hardware implementation of the algorithm is more easily achieved. To derive the algorithm for the simplified case, two basic assumptions were made regarding the loss shape of the Bode Network and the characteristic of the channel. Under these assumptions, the true gradient of the kth Bode Network is given by the weighted sum of the error signal measured at frequencies w_k, $w_k + \Delta w/2$, and $w_k - \Delta w/2$, where $\Delta w = w_{k+1} - w_k$ and w_{k+1} is the center frequency of the next higher frequency Bode Network. For the hardware realization, the gradient information is applied to integrators, the outputs of which in turn control the gain settings, the process being continued until all the inputs to the integrator, i.e., gradients, become zero (see Fig. 3).

The computer results given in Section IV show that the reduction of total MSE of the equalized channel is negligible by changing from the simplified MSE algorithm to the general MSE algorithm, verifying the reasonableness of the assumptions made in Section II.

VI. ACKNOWLEDGMENT

The author wishes to thank F. C. Kelcourse with whom the author had many stimulating and encouraging discussions.

APPENDIX

Proof of Theorem 2

If we can show that the gradient

$$G_k = \Delta w\left\{\frac{1}{2}E\left(w_k - \frac{\Delta w}{2}\right) + E(w_k) + \frac{1}{2}E\left(w_k + \frac{\Delta w}{2}\right)\right\}, \qquad (15)$$

the theorem is proved by the result of Theorem 1. From eq. (10),

$$\begin{aligned} G_k &= 2\int_{-\infty}^{\infty} B_k(w)E(w)\,dw \\ &= 2\int_{-\infty}^{\infty} B_k(w)[\mathrm{EQL}\,(w) - M(w)]\,dw \\ &= 2(G_{k1} - G_{k2}), \end{aligned} \qquad (16)$$

where

$$G_{k1} = \int_{-\infty}^{\infty} B_k(w)\,\mathrm{EQL}\,(w)\,dw$$

and

$$G_{k2} = \int_{-\infty}^{\infty} B_k(w)M(w)\, dw.$$

By Parserval's relationship,[8]

$$G_{k1} = \int_{-\infty}^{\infty} 2\pi b_k(t) \text{ eql } (-\text{t})\, dt, \tag{17}$$

where

$$\begin{aligned} b_k(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} B_k(w) \exp(jwt)\, dw \\ &= \frac{\Delta w}{4\pi}\left(1 + \cos\left(\frac{\Delta w}{2}\, t\right)\right) \exp(jw_k t) \quad \text{for} \quad |t| \leqq \frac{2\pi}{\Delta w} \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{for} \quad |t| > \frac{2\pi}{\Delta w} \end{aligned}$$

and

$$\begin{aligned} \text{eql } (t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{EQL } (w) \exp(jwt)\, dw \\ &= \frac{\Delta w}{4\pi}\left(1 + \cos\left(\frac{\Delta w}{2}\, t\right)\right) \sum_{i=1}^{M} g_i \exp(jw_i t) \quad \text{for} \quad |t| \leqq \frac{2\pi}{\Delta w} \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{for} \quad |t| > \frac{2\pi}{\Delta w}. \end{aligned}$$

Hence,

$$G_{k1} = \frac{\Delta w^2}{8\pi}\int_{-2\pi/\Delta w}^{2\pi/\Delta w} \sum_{i=1}^{M} g_i \exp(j(w_k - w_i)t)\left(1 + \cos\left(\frac{\Delta w}{2}\, t\right)\right)^2 dt. \tag{18}$$

Since $w_k - w_i = (k - i)\Delta w$ and the integration with the imaginary term in (18) is zero,

$$\begin{aligned} G_{k1} &= \frac{\Delta w^2}{8\pi}\int_{-2\pi/\Delta w}^{2\pi/\Delta w} \sum_{i=1}^{M} g_i \cos((k - i)\, \Delta w t) \\ &\quad \cdot\left\{\frac{3}{2} + 2\cos\left(\frac{\Delta w}{2}\, t\right) + \frac{1}{2}\cos(\Delta w t)\right\} dt \\ &= \frac{\Delta w^2}{8\pi}\int_{-2\pi/\Delta w}^{2\pi/\Delta w} \left\{\frac{3}{2}\, g_k + \frac{1}{4}\, g_{k-1} + \frac{1}{4}\, g_{k+1}\right\} dt \end{aligned}$$

$$= \frac{\Delta w}{2}\left\{\frac{3}{2}\, g_k + \frac{1}{4}\, g_{k-1} + \frac{1}{4}\, g_{k+1}\right\}. \tag{19}$$

However,

$$\mathrm{EQL}\left(w_k - \frac{\Delta w}{2}\right) = \frac{1}{2}\,(g_{k-1} + g_k), \qquad \mathrm{EQL}\,(w_k) = g_k\,,$$

and

$$\mathrm{EQL}\left(w_k + \frac{\Delta w}{2}\right) = \frac{1}{2}\,(g_{k+1} + g_k).$$

Hence, eq. (19) becomes

$$G_{k1} = \frac{\Delta w}{2}\left[\frac{1}{2}\mathrm{EQL}\left(w_k - \frac{\Delta w}{2}\right) + \mathrm{EQL}\,(w_k) + \frac{1}{2}\mathrm{EQL}\left(w_k + \frac{\Delta w}{2}\right)\right]. \tag{20}$$

Now

$$G_{k2} = \int_{-\infty}^{\infty} B_k(w)M(w)\, dw. \tag{21}$$

From eqs. (3) and (6),

$$\begin{aligned} G_{k2} &= \int_{-\infty}^{\infty} \mathrm{cosinc}\left(\frac{\pi}{\Delta w}\,(w - w_k)\right) \int_0^1 \{F(x)\cos(2\pi p_N w x) \\ &\quad + H(x)\sin(2\pi p_N w x)\}\, dx\, dw \\ &= \int_{-\infty}^{\infty} \mathrm{cosinc}\left(\frac{\pi}{\Delta w}\,(w - w_k)\right) \int_0^1 \{f(x)\cos(2\pi p_N(w - w_k)x) \\ &\quad + h(x)\sin(2\pi p_N(w - w_k)x)\}\, dx\, dw, \end{aligned} \tag{22}$$

where

$$F(x) = f(x)\cos(2\pi p_N w_k) - h(x)\sin(2\pi p_N w_k)$$

and

$$H(x) = f(x)\cos(2\pi p_N w_k) + h(x)\sin(2\pi p_N w_k).$$

Since cosinc $(\pi/\Delta w(w - w_k))$ and sin $(2\pi p_N(w - w_k)x)$ are even and odd functions, respectively, with respect to $w = w_k$ axis,

$$\int_0^1 \int_{-\infty}^{\infty} h(x)\,\mathrm{cosinc}\left(\frac{\pi}{\Delta w}\,(w - w_k)\right)\sin(2\pi p_N(w - w_k)x)\, dw\, dx = 0.$$

Replacing $w = u + w_k$ and changing "cos" to exponential form, eq. (22) becomes

$$\frac{1}{2}\int_{-\infty}^{\infty} \operatorname{cosinc}\left(\frac{\pi}{\Delta w}u\right)\int_0^1 f(x)(\exp(j2\pi p_N ux) + \exp(-j2\pi p_N ux))\,dx\,du$$
$$= \frac{1}{2}\int_0^1 f(x)\int_{-\infty}^{\infty} \operatorname{cosinc}\left(\frac{\pi}{\Delta w}u\right)$$
$$\cdot(\exp(j2\pi p_N ux) + \exp(-j2\pi p_N ux))\,du\,dx. \quad (23)$$

Since

$$\int_{-\infty}^{\infty} \operatorname{cosinc}\left(\frac{\pi}{\Delta w}u\right)\exp(jut)\,du$$
$$= \frac{\Delta w}{2}\left(1 + \cos\left(\frac{\Delta w}{2}t\right)\right) \quad \text{for} \quad |t| \leqq \frac{2\pi}{\Delta w}$$

and

$$= 0 \quad \text{for} \quad |t| > \frac{2\pi}{\Delta w},$$

eq. (23) becomes

$$G_{k2} = \frac{\Delta w}{2}\int_0^1 f(x)\{1 + \cos(\pi p_N\,\Delta w x)\}\,dx \quad \text{for} \quad p_N x \leqq \frac{1}{\Delta w}$$

and

$$= 0 \quad \text{for} \quad p_N x > \frac{1}{\Delta w}.$$

However, $0 \leqq x \leqq 1$ and $\Delta w \leqq 1/p_N$ by definition (12). Hence,

$$G_{k2} = \frac{\Delta w}{2}\int_0^1 f(x)\{1 + \cos(\pi p_N\,\Delta w x)\}\,dx$$
$$= \frac{\Delta w}{2}\left\{\frac{1}{2}M\left(w_k - \frac{\Delta w}{2}\right) + M(w_k) + \frac{1}{2}M\left(w_k + \frac{\Delta w}{2}\right)\right\}. \quad (24)$$

Combining (21) and (24), the gradient in (16) becomes

$$G_k = 2(G_{k1} - G_{k2})$$
$$= \Delta w\left\{\frac{1}{2}\left(\mathrm{EQL}\left(w_k - \frac{\Delta w}{2}\right) - M\left(w_k - \frac{\Delta w}{2}\right)\right) + (\mathrm{EQL}(w_k) - M(w_k))\right.$$
$$\left. + \frac{1}{2}\left(\mathrm{EQL}\left(w_k + \frac{\Delta w}{2}\right) - M\left(w_k + \frac{\Delta w}{2}\right)\right)\right\}$$
$$= \Delta w\left\{\frac{1}{2}E\left(w_k - \frac{\Delta w}{2}\right) + E(w_k) + \frac{1}{2}E\left(w_k + \frac{\Delta w}{2}\right)\right\}, \quad (25)$$

which is equal to (15). This proves the theorem.

REFERENCES

1. Ketchledge, R. W., et al., "Equalization and Regulation," B.S.T.J., *32*, No. 4 (July 1953), pp. 833–878.
2. Kelcourse, F. C., et al., "Equalizing and Main Station Repeaters," B.S.T.J., *48*, No. 4 (April 1969), pp. 889–952.
3. Bode, H. W., "Variable Equalizers," B.S.T.J., *17*, No. 2 (April 1938), pp. 229–244.
4. Kelley, H., "Method of Gradients," Chap. 6 of *Optimization Techniques*, G. Leitman, ed., New York: Academic Press, 1962.
5. Narendra, K. S., et al., "Multiparameter Self-Optimizing Systems Using Correlation Techniques," IEEE Trans. Auto. Cont., (January 1964), pp. 31–38.
6. Lucky, R. W., "Automatic Equalization for Digital Communication," B.S.T.J., *45*, No. 2 (February 1966), pp. 255–286.
7. Rice, J. R., *The Approximation of Functions*, Vol. 1, New York: Addison-Wesley Pub. Co., 1964.
8. Papoulis, A., *The Fourier Integral and its Application*, New York: McGraw-Hill, 1962.

The Bell System Technical Journal
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

Dimensionality of Crosstalk Functions

By T. A. LENAHAN

(Manuscript received February 7, 1972)

We study the dimensionality for the class of near-end crosstalk functions in a cable. The dimensionality is closely related to the distribution of eigenvalues for a particular integral operator that we call the energy operator. We find bounds for these eigenvalues in terms of the eigenvalues associated with the prolate spheroidal waveforms studied by Landau, Pollak, and Slepian. The important technical observation, permitting us to use their results, is that though the crosstalk functions are not bandlimited, the degree to which they are band-concentrated can be uniformly specified.

I. INTRODUCTION

The class of functions bandlimited to the interval $(-W, W)$ and considered over the interval $(-T, T)$ has long been held to have essentially $[2WT]$ degrees of freedom. This goes back at least as far as the discovery of the sampling (or cardinal) series, since exactly this number of terms in the series is available with knowledge of the function over the interval $(-T, T)$.[1] The notion was made precise and validated by Landau and Pollak.[2] The fundamental quantity in their approach was the energy (or L^2-norm) of the bandlimited function over $(-T, T)$. The energy is computed as a quadratic form of the function and to this there corresponds a positive definite, compact operator. We shall call this the energy operator. The distribution of the eigenvalues for the energy operator, i.e., the energy eigenvalues, determine the approximate number of degrees of freedom or dimensionality of this class of functions. The idea is that energy eigenfunctions with small enough eigenvalues (or energy) can contribute only minimally to the energy in the interval; hence, they can be disregarded. They find that $[2WT]$ energy eigenfunctions span this space of functions within an error bound which they compute.

To be more definite, let D_T denote the operator which acts on square integrable functions as follows:

$$D_T f(t) = \begin{cases} f(t) & t \in (-T, T) \\ 0 & t \notin (-T, T) \end{cases}$$

and let B_W denote the operator which similarly chops off the Fourier transform of the function outside $(-W, W)$; thus, if $F(\omega)$ is the Fourier transform of $f(t)$, then

$$B_W f(t) = \frac{1}{2\pi} \int_{-2\pi W}^{2\pi W} e^{i\omega t} F(\omega)\, d\omega.$$

Whenever f is bandlimited to $(-W, W)$, then $B_W f = f$. The energy of this function in the interval $(-T, T)$ is

$$|| D_T f ||^2 = || D_T B_W f ||^2 = (D_T B_W f, D_T B_W f) = (f, D_T B_W f),$$

where we have made use of the fact that D_T is a projection operator and $B_W f = f$. Note:

$$\left(|| g(t) ||^2 = \int_{-\infty}^{\infty} | g(t) |^2 \, dt \quad \text{and} \quad (g, h) = \int_{-\infty}^{\infty} g(t)\bar{h}(t)\, dt \right).$$

The combined operator $D_T B_W$ is the energy operator, and its eigenvalues are the energy eigenvalues studied by Landau and Pollak.

This paper concerns the generalization of these results on dimensionality to the class of functions representing near-end crosstalk transfer functions within a multipair cable. As an approximation of the coupling within a cable, it follows from the telegrapher's equation[3,4] that these transfer functions have the form

$$N(\omega) = i\omega \int_0^l e^{-2\Gamma(\omega)x} u(x)\, dx,$$

where $\Gamma(\omega) = i\beta(\omega) + \alpha(\omega)$ $(\alpha(\omega) \geqq 0)$ denotes the propagation function for a pair in the cable, l is the cable length, and the coupling function between two pair, $u(x)$, satisfies

$$\int_0^l | u(x) |^2 \, dx < \infty.$$

The physical meaning of $N(\omega)$ is specified in more detail in Section III. We wish to find the approximate dimensionality for the class of such functions either as viewed over some finite interval or, more generally, as weighted by some fixed square integrable function $F(\omega)$. Our approach, again, is to set up an energy operator, then study the energy eigenvalues to reach conclusions about the dimensionality.

The paper goes from the general to the particular. We first introduce a class of compact, integral operators and derive upper bounds for their eigenvalues. Next, we show that this class includes the energy operator corresponding to the crosstalk equation; this gives us bounds on the energy eigenvalues. From these bounds, one can draw quantitative conclusions about the dimensionality of the class of crosstalk functions.

II. A CLASS OF INTEGRAL OPERATORS

The first problem is to determine the distribution of eigenvalues for a special class of integral operators on the space of square integrable functions, $L^2(0, l)$. We characterize these operators by kernels of the form

$$K(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} | F(\omega) |^2 P(x, \omega) P^*(y, \omega) \, d\omega \qquad 0 \leqq x, y \leqq l,$$

where $F(\omega)$ is bounded and square-integrable, and $P(x, \omega)$ has the following two properties:

(*i*) The function $P(x, \omega)$ is bounded and the integral $\int_{\infty}^{\infty} \int_0^l |P(x, \omega)|^2 \, dx \, d\omega$ is bounded. As a consequence:
  (a) The ω-function, $U(\omega) \equiv \int_0^l P(x, \omega) u(x) \, dx$ is square-integrable when $u(x)$ is square-integrable over $(0, l)$. We assume its norm is nonzero.
  (b) $U(\omega)$ has a Fourier transform.
  (c) The operator, B_Y , limiting the Fourier transform to the interval $(-Y, Y)$ can be applied to the functions $U(\omega)$.

(*ii*) For all $u(x)$ in $L^2(0, l)$,

$$|| F(\omega)(I - B_Y)[U(\omega)] ||_W \leqq \epsilon(Y) \, || u(x) ||_X$$

and $\epsilon(Y) \to 0$ as $Y \to \infty$. [Note: $|| \cdot ||_W$ denotes the standard norm on $L^2(-\infty, \infty)$ and $|| \cdot ||_X$ on $L^2(0, l)$. Also, "I" denotes the identity operator on $L^2(-\infty, \infty)$, i.e., $I[G(\omega)] = G(\omega)$.]

Suppose K is the operator having the kernel $K(x, y)$ above; then, since the kernel is square-integrable jointly in x and y, K is compact; i.e., it has a sequence of eigenvalues, say λ_n , $n = 0, 1, \cdots$, which approach zero as n gets large (see Ref. 5, p. 264). We shall bound the λ_n in terms of the eigenvalues, say $\lambda_n^0(F, Y)$, of the compact operator denoted by M_Y which acts on functions in $L^2(-Y, Y)$ and has the kernel

$$M_Y(\eta, \xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} | F(\omega) |^2 e^{i\omega(\eta-\xi)} \, d\omega.$$

These eigenvalues are more directly accessible and better studied than the λ_n , and when

$$F(\omega) = \begin{cases} 1 & | \omega | \leqq 2\pi W \\ 0 & | \omega | > 2\pi W \end{cases}$$

they are exactly the eigenvalues of the operator $D_Y B_W$ studied by Landau and Pollak. Also included in our bounding expression will be the eigenvalues, μ_n , $n = 0, 1, \cdots$, of the compact integral operator L acting on functions in $L^2(0, l)$ and having kernel

$$L(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} P(x, \omega) P^*(y, \omega) \, d\omega.$$

[*Note:* Using Schwarz's inequality on the ω-integral above and then property "(a)" for $P(x, \omega)$, we conclude that $L(x, y)$ is jointly square-integrable in x and y, and this assures the compactness of the operator L.]

Theorem: For the eigenvalues λ_n , $\lambda_n^0(Y, F)$, *and* μ_n *just defined, we have for each* $n = 0, 1, \cdots$

$$\lambda_n \leqq \min \left(\inf_Y [(a_n \lambda_n^0(Y, F))^{\frac{1}{2}} + \epsilon(Y)]^2, \, b_n \mu_n \right),$$

where $0 \leqq a_n \leqq \mu_0$, $0 \leqq b_n \leqq \max_\omega | F(\omega) |^2$, a_n *approaches zero as* n *becomes large.*

Proof: The Weyl–Courant Lemma (See Ref. 6, p. 251) implies that

$$\lambda_n = \inf_{S_n} \left(\sup_{u \perp S_n} \frac{(Ku, u)}{(u, u)} \right)$$

for any $n = 0, 1, 2, \cdots$, where S_n denotes an n-dimensional subspace of $L^2(0, l)$ and the infimum is taken over all such subspaces. Since (Lu, u) is nonnegative for all $u(x)$ in $L^2(0, l)$, we have

$$\lambda_n = \inf_{S_n} \sup_{u \perp S_n} \frac{(Ku, u)}{(Lu, u)} \frac{(Lu, u)}{(u, u)}.$$

So for all choices of S_n ,

$$\lambda_n \leqq \sup_{u \perp S_n} \frac{(Ku, u)}{(Lu, u)} \sup_{u \perp S_n} \frac{(Lu, u)}{(u, u)}.$$

Choose S_n so as to minimize the latter factor. But the minimum value,

by the Weyl–Courant Lemma, is exactly the nth eigenvalue of L. Thus, when S_n' is the appropriate subspace, then

$$\lambda_n \leqq \left(\sup_{u \perp S_{n'}} \frac{(Ku, u)}{(Lu, u)} \mu_{n'} \right) \equiv b_n \mu_n .$$

We claim $(Ku, u) = || F(\omega)U(\omega) ||_w^2$ and $(Lu, u) = || U(\omega) ||_w^2$ with

$$U(\omega) = \int_0^l P(x, \omega)u(x)\, dx,$$

from which it follows that

$$0 \leqq b_n \leqq \max_{\omega} | F(\omega) |^2 \quad \text{for all} \quad n.$$

To prove the claim, we have

$$\begin{aligned}(Ku, u) &= \frac{1}{2\pi} \int_0^l u(y) \int_0^l u(x) \int_{-\infty}^{\infty} | F(\omega) |^2 P(x, \omega)P^*(y, \omega)\, d\omega\, dx\, dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} | F(\omega) |^2 \left(\int_0^l P(x, \omega)u(x)\, dx \right) \left(\int^l P^*(y, \omega)u(y)\, dy \right) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} | F(\omega)U(\omega) |^2\, d\omega = || F(\omega)U(\omega) ||_w^2\end{aligned}$$

and similarly for (Lu, u). (The integrand above is clearly absolutely integrable, so the conditions for the Fubini Theorem are met, and the order of integration can be switched freely.)

For the second part of the bound, the Minkowski inequality implies

$$(Ku, u) \leqq (|| F(\omega)B_Y[U(\omega)] ||_w + || F(\omega)(I - B_Y)[U(\omega)] ||_w)^2.$$

So by assumption (ii) on the function $P(x, \omega)$ we get

$$\frac{(Ku, u)}{(u, u)} \leqq \left(\frac{|| F(\omega)B_Y[U(\omega)] ||_w}{|| u ||} + \epsilon(Y) \right)^2.$$

Now choose the subspace S_n so as to minimize the quantity

$$\sup_{u \perp S_n} \frac{|| F(\omega)B_Y[U(\omega)] ||_w}{|| B_Y[U(\omega)] ||_w}.$$

But this minimum value is upper bounded by the nth eigenvalue of the operator M_Y, defined previously, because the Weyl–Courant Lemma implies

$$\lambda_n^0(F, Y) = \inf_{R_n} \sup_{V \perp R_n} \frac{(M_Y V, V)_Y}{(V, V)_Y},$$

where R_n represents a subspace of $L^2(-Y, Y)$; and by the calculation above

$$(M_Y V, V) = || F(\omega)\tilde{V}(\omega) ||_w^2 ,$$

where

$$\tilde{V}(\omega) = \int_{-Y}^{Y} e^{i\omega y} V(y)\, dy = B_Y[\tilde{V}(\omega)]$$

and the argument is finished by applying the Plancherel Theorem,

$$(V, V)_Y = [\tilde{V}(\omega), \tilde{V}(\omega)]_w .$$

Therefore, when S_n'' is the appropriate subspace of $L^2(0, l)$,

$$\lambda_n \leqq \sup_{u \perp S_n''} \frac{(Ku, u)}{(u, u)} \leqq [(a_n \lambda_n^0(F, Y))^{\frac{1}{2}} + \epsilon(Y)]^2,$$

where

$$a_n = \sup_{u \perp S_n''} \frac{|| B_Y[U(\omega)] ||_w^2}{|| u(x) ||_x^2}.$$

Note that

$$0 \leqq a_n \leqq \sup_u \frac{|| U(\omega) ||_w^2}{|| u(x) ||_x^2} = \left(\sup_u \frac{(Lu, u)}{(u, u)}\right) = \mu_0 .$$

Also, since L is compact, $a_n \to 0$ as $n \to \infty$. The inequality is good for all values of Y, so it is good for the infimum over Y. We have two upper bounds for the λ_n ; thus the minimum of the two is also an upper bound. This proves the theorem. Q.E.D.

III. APPLICATION TO CROSSTALK

The near-end crosstalk equation for multipair cable leads to an integral operator of the type in the theorem. The crosstalk transfer function $N(\omega)$ (in the frequency domain) is related to a coupling function along the cable, $u(x)$, by

$$N(\omega) = i\omega \int_0^l e^{-2\Gamma(\omega)x} u(x)\, dx,$$

where l denotes the cable length and $\Gamma(\omega)$ denotes the propagation function of a pair in the cable. A good approximation to $\Gamma(\omega)$ over the frequency range 0.1 to 10 mHz is

$$\Gamma(\omega) = k_1 \sqrt{|\omega|} + i k_2 \omega + i \operatorname{sgn}(\omega) k_1 \sqrt{|\omega|},$$

where the exact values of the constants k_1 and k_2 depend on the gage of the wire. More precisely, when a pair in a cable is excited by a signal with spectrum $G(\omega)$, the coupling within the cable will produce a signal with spectrum $G(\omega)N(\omega)$ at the near end of an unexcited pair. The energy of the crosstalk signal is

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} |\, G(\omega)N(\omega)\,|^2\, d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\, \omega G(\omega)\,|^2 \left|\int_0^l e^{-2\Gamma(\omega)x}u(x)\,dx\right|^2 d\omega$$

By Fubini's Theorem [applicable when $\omega G(\omega) \in L^2(-\infty, \infty)$]we get,

$$\int_0^l u(y)\int_0^l u(x)\frac{1}{2\pi}\int_{-\infty}^{\infty} |\, \omega G(\omega)\,|^2\, e^{-2\Gamma(\omega)x-2\Gamma^*(\omega)y}\,d\omega\,dx\,dy = (Ku, u).$$

Here K is a compact integral operator with the kernel

$$K(x, y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\, \omega G(\omega)\,|^2\, e^{-2\Gamma(\omega)x-2\Gamma^*(\omega)y}\,d\omega.$$

We call K the energy operator and its eigenvalues, say, λ_n, $n = 0, 1, \cdots$, the energy eigenvalues. With a slightly more restrictive assumption on the function $G(\omega)$, we obtain for the λ_n the same bounds as before.

Proposition: When $(1 + |\,\omega\,|^q)\omega G(\omega)$ is in $L^2(-\infty, \infty)$ for $q > \frac{1}{4}$ and asymptotically for large ω, $Re(\Gamma(\omega)) \sim \omega^r$ with $r \geqq \frac{1}{2}$, then the previous bounds apply to the energy eigenvalues.

Proof: Let

$$P(x, \omega) = \frac{e^{-2\Gamma(\omega)x}}{1 + |\,\omega\,|^q};$$

then the previous properties assumed for $P(x, \omega)$ are satisfied. To demonstrate this: first, it is clear that $P(x, \omega)$ is bounded and it is tailored (i.e., the q values are just large enough) to be square-integrable jointly in x and w. Since $\int_0^l e^{-2zx}u(x)\,dx$ is analytic in z, it follows from the Plancherel Theorem that $U(\omega)$ has nonzero norm for all $u(x)$ in $L^2(0, l)$ with nonzero norm. Finally, using the Fubini Theorem,

$$\begin{aligned} B_Y\left[\int_0^l P(x, \omega)u(x)\,dx\right] &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\sin Y(\eta - \omega)}{(\eta - \omega)}\int_0^l P(x, \omega)u(x)\,dx\,d\omega \\ &= \frac{1}{2\pi}\int_0^l u(x)\int_{-\infty}^{\infty}\frac{\sin Y(\eta - \omega)}{(\eta - \omega)}P(x, \omega)\,d\omega\,dx \\ &= \int_0^l B_Y[P(x, \omega)]u(x)\,dx; \end{aligned}$$

but $B_Y[P(x, \omega)]$ is uniformly bounded in ω for all values of x in the interval $[0, l]$, so

$$\rho_Y(x, \omega) \equiv (I - B_Y)[P(x, \omega)]$$

is square-integrable in x and for all ω,

$$| (I - B_Y)[U(\omega)] |^2 \leqq \int_0^l | \rho_Y(x, \omega) |^2 \, dx \int_0^l | u(x) |^2 \, dx.$$

Therefore,

$$|| F(\omega)(I - B_Y)[U(\omega)] ||_w \leqq \epsilon(Y) \, || u(x) ||_x ,$$

where

$$\epsilon(Y) = \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^l | F(\omega)\rho_Y(x, \omega) |^2 \, dx \, d\omega\right]^{\frac{1}{2}}.$$

Put $F(\omega) = (1 + | \omega |^q)\omega G(\omega)$; then

$$K(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} | F(\omega) |^2 \frac{e^{-2\Gamma(\omega)x}}{1 + | \omega |^q} \frac{e^{-2\Gamma^*(\omega)y}}{1 + | \omega |^q} \, d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} | F(\omega) |^2 P(x, \omega)P^*(y, \omega) \, d\omega.$$

This is exactly the same form as before and so the energy eigenvalues have exactly the same bound. Q.E.D.

Corollary: When

$$F(\omega) = \begin{cases} 1 & | \omega | \leqq W \\ 0 & | \omega | > W \end{cases},$$

$$\lambda_n \leqq \min \left(\inf_Y [(\mu_0\alpha_n(2YW))^{\frac{1}{2}} + \epsilon(Y, W)]^2, \mu_n\right),$$

where α_n denotes the eigenvalues associated with the prolate spheroidal waveforms (Cf. Ref. 2), and $\epsilon(Y, \omega)$ is $\epsilon(Y)$ with the W-dependence indicated.

Proof: Since $a_n \leqq \mu_0$, we can substitute μ_0 for a_n . The eigenvalues $\lambda_n^0(Y, F)$ in this special case are those studied by Landau and Pollak; they depend on the product of Y and W. Finally, $b_n \leqq 1$ in this case. Q.E.D.

IV. DIMENSIONALITY OF CROSSTALK FUNCTIONS

When approximating crosstalk functions over some interval (or as weighted by some square-integrable function) by linear combinations of functions, two practical questions arise. What is the most efficient set of functions, i.e., the one requiring the fewest number of functions to approximate any function in the class to a specified error tolerance? Then, for a given error tolerance, how many of these functions are required, i.e., what is the dimensionality?

There are at least two practical ways that information on crosstalk dimensionality can be used. For a specified error tolerance, the dimensionality gives the minimum number of independent measurements required to determine the crosstalk as a function of frequency. Thus it provides useful information to crosstalk measurement programs. Also, in efforts to reduce crosstalk over a specific range of frequencies, by subtracting linear combinations of fixed functions, the dimensionality indicates the minimum number of independent controls needed to meet a given criterion. Thus, dimensionality is a general concept, not tied to any particular method, either for measuring or for controlling crosstalk.

Before discussing dimensionality further, we answer the first question stated above. We show that the most efficient set is the set of eigenfunctions for the energy operator, i.e., the energy eigenfunctions. This result is a variant of Theorem I in Landau and Pollak's paper.[2] Our proof differs from theirs; and also we strive for the greatest generality by considering crosstalk functions multiplied by an arbitrary square-integrable function, $G(\omega)$. Later, in dealing with dimensionality, we shall take $G(\omega)$ as zero outside the finite frequency range of interest.

Theorem: The quantity

$$\sup_{N(\omega)} \min_{\{a_j\}_0^{J-1}} \left\| N(\omega)G(\omega) - \sum_{j=0}^{J-1} a_j N_j(\omega)G(\omega) \right\|_w^2 ,$$

where the supremum is taken over all crosstalk functions with normalized coupling function, is minimized by choosing $N_j(\omega)$, $j = 0, \cdots, J-1$, *as the crosstalk functions with coupling functions* $u_j(x)$ *equal to the* $(j+1)$th *normalized energy eigenfunction (ordered according to decreasing eigenvalue).*

Proof: Suppose $N_j(\omega)$, $j = 0, \cdots, J-1$, are linearly independent crosstalk functions and let P_J denote the projection operator onto the J-dimensional subspace in $L^2(-\infty, \infty)$ spanned by $\{G(\omega)N_j(\omega)\}_0^{J-1}$; put $P_J' = I - P_J$. Then,

$$\min_{\{a_j\}_0^{J-1}} \left|\left| N(\omega)G(\omega) - \sum_{j=0}^{J-1} a_j N_j(\omega)G(\omega) \right|\right|_w^2 = || P_J' N(\omega)G(\omega) ||_w^2 .$$

Let A denote the operator taking coupling functions in $L^2(0, l)$ to $L^2(-\infty, \infty)$ such that

$$Au(x) = i\omega G(\omega) \int_0^l e^{-2\Gamma(\omega)x} u(x)\, dx = G(\omega)N(\omega).$$

(Note: If A^* denotes the adjoint of A, then $A^*A = K$, the energy operator.) The problem is to choose $u(x)$ (with $|| u ||_x = 1$) to maximize the quantity $|| P_J' Au(x) ||_w^2$ and then to choose the minimizing projection P_J' .

Since P_J' is a projection operator,

$$|| P_J' Au ||_w^2 = (P_J' Au, P_J' Au)_w = (A^* P_J' Au, u)_w$$

and the maximization over $u(x)$ gives the operator norm of the operator $A^*P_J'A$, i.e., the greatest eigenvalue. We denote this by $| A^*P_J'A |$. But

$$| A^*P_J'A | = |(A^*P_J')(P_J'A)| = |(P_J'A)(A^*P_J')| = | P_J'AA^*P_J' | ,$$

which follows from the more general result that BB^* and B^*B have the same nonzero eigenvalues when B is compact (Cf. Ref. 5, p. 262). The Weyl–Courant Lemma implies that $| P_J'AA^*P_J' |$ is minimized when P_J corresponds to the subspace spanned by the first J eigenfunctions of AA^* with the $(J + 1)$th eigenvalue as the minimal value. Since $K = A^*A$ and AA^* have the same nonzero eigenvalues, this minimal value is λ_J , i.e.,

$$\inf_{P_J} | A^*P_J'A | = \lambda_J .$$

When P_J is associated with the subspace spanned by $\{G(\omega)N_i(\omega)\}_0^{J-1}$, where the $N_i(\omega)$ are crosstalk functions with the first J energy eigenfunctions for coupling functions, then clearly

$$| A^*P_J'A | = \lambda_J .$$

Thus, this set of functions is most efficient. Q.E.D.

We note that the supremum in the theorem has been taken over all $N(\omega)$ (with associated coupling function having unit norm, i.e., $\int_0^l | u(x) |^2\, dx = 1$). One can show that the suitable approximating functions are the same even if the supremum is taken over $N(\omega)$ with the additional constraint, $|| N(\omega)G(\omega) ||_\omega = b$ for a fixed value of b.

The main issue is dimensionality. Again, let $N(\omega)$ be an arbitrary crosstalk function. Suppose one wishes to approximate $N(\omega)G(\omega)$ by a

linear combination of fixed linearly independent functions to within a mean square error of δ^2. The preceding theorem indicates that, for greatest efficiency, one should use the energy eigenfunctions $N_i(\omega)$. The dimensionality relative to $G(\omega)$ is the smallest integer D such that

$$\min_{\{a_i\}} \left|\left| G(\omega)N(\omega) - \sum_{i=0}^{D-1} a_i G(\omega)N_i(\omega) \right|\right|_w^2 \leqq \delta^2$$

for all $N(\omega)$ with normalized coupling function. In terms of the energy eigenvalues this means

$$\lambda_D \leqq \delta \quad \text{but} \quad \lambda_{D-1} > \delta.$$

This is so because for any such function $N(\omega)$ there are coefficients $\{b_i\}$ such that

$$N(\omega) = \sum_{i=0}^{\infty} b_i N_i(\omega) \quad \text{and} \quad \sum_{i=0}^{\infty} | b_i^2 | = 1.$$

Thus,

$$\min_{\{a_i\}} \left|\left| G(\omega)N(\omega) - \sum_{i=0}^{D-1} a_i G(\omega)N_i(\omega) \right|\right|_\omega^2 = \left|\left| \sum_{i=D}^{\infty} G(\omega) b_i N_i(\omega) \right|\right|_\omega^2$$

$$= \sum_{i=D}^{\infty} b_i^2 \lambda_i^2 .$$

Since the eigenvalues λ_i are in descending order, this quantity is maximized by putting $b_i = \begin{cases} 0 & i \neq D \\ 1 & i = D \end{cases}$ and, therefore, it must be less than δ^2 for all choices of $\{b_i\}$.

Now we wish to work out the dimensionality for the practical case where one is concerned with a fixed frequency interval, $(-W/2\pi, W/2\pi)$. We modify the assumption of the corollary in Section III to say that

$$(1 + | k_2\omega |^a)\, \omega G(\omega) = \begin{cases} 1 & | \omega | \leqq W \\ 0 & | \omega | \geqq W \end{cases}.$$

The corollary indicates that the dimensionality (relative to δ) is upper bounded by the smallest integer D such that

$$(((\mu_0 \alpha_D(2YW))^{\frac{1}{2}} + \epsilon(Y, W))^2, \mu_D) < \delta \quad \text{(for some } Y)$$

and the modification means only that ϵ and the μ_D undergo a corresponding change. The eigenvalues $\alpha_D(2WY)$ are tabulated and plotted in Ref. 7. They decrease rapidly for increasing D greater than the

threshold value, $2C/\pi \equiv 2/\pi[2YW]$, as indicated there. This is generalized in Ref. 8. What must be calculated is the behavior of the eigenvalue sequence, $\{\mu_D\}$, and the function $\epsilon(Y, W)$.

Assume $q = 2$ and

$$\Gamma(\omega) = k_1(\sqrt{|\omega|} + \text{sgn}(\omega) i\sqrt{|\omega|}) + ik_2\omega.$$

Though we leave k_2 unassigned, we assume relative to mile units, $k_1^2 = k_2$. (A typical value of k_2 is 8.0×10^6 second/mile for 19-gage wire.) An upper bound for $\epsilon(Y, W)$ is $\epsilon(Y, \infty)$ and naturally it is tighter for larger values of W. Explicitly,

$$\epsilon^2(Y, W) \leqq \frac{1}{2\pi}\int_0^l \int_{-\infty}^{\infty} \left| (I - B_Y)\left(\frac{e^{-2\Gamma(\omega)x}}{1 + (k_2\omega)^2}\right)\right|^2 d\omega\, dx.$$

To estimate this, note that the Fourier transform of $e^{-2\Gamma(\omega)x}/1 + (k_2\omega)^2$ is the convolution of the transforms for each factor taken separately, each of which is standard. Denote this by $f(y, x)$; then

$$f(y, x) = \int_{-\infty}^{y_0} \frac{1}{4k_2} e^{-|z|/k_2} \frac{xk_1}{\sqrt{\pi}} (y_0 - z)^{-\frac{3}{2}} \exp\left(-\frac{2x^2k_1^2}{y_0 - z}\right) dz,$$

where $y_o = y - 2k_2x$. By the Plancherel theorem, we have

$$\epsilon^2(Y, W) \leqq \int_0^l \int_{|y|>Y} |f(y, x)|^2\, dy\, dx.$$

To arrive at specific bounds on $\epsilon(Y, W)$, we bound the convolution and then perform the y and x integrations on this bound. The derivation appears in the Appendix; the result is

$$\epsilon^2(Y, W) \leqq \frac{l}{k_2}\left[\frac{e^{-2P}}{64} + \frac{3e^{-2P/3}}{64} + \frac{9l^2}{64\pi P^2} + \min\left\{\frac{9le^{-P/3}}{32P^{\frac{1}{2}}}, \frac{3le^{-P/3}}{8P^{\frac{1}{2}}}\right\}\right],$$

where $P = (Y - 2k_2l)/k_2$.

The remaining unspecified quantities in the bounding expression for the energy eigenvalues, λ_n, are the eigenvalues, μ_n, for the operator L. Note first that these depend on the two parameters, k_2 and l. The former is easily handled:

$$L(x, y) = \int_{-\infty}^{\infty} \frac{e^{-2\Gamma(\omega)x}}{1 + (k_2\omega)^2} \frac{e^{-2\Gamma^*(\omega)y}}{1 + (k_2\omega)^2} \frac{d\omega}{2\pi} \qquad 0 \leqq x, \quad y \leqq l.$$

Put $k_2\omega = \xi$; then we obtain

$$L(x, y) = \frac{2}{k_2}\int_0^{\infty} e^{-2(x+y)\sqrt{\xi}} \cos\left[2(\xi + \sqrt{\xi})(x - y)\right] \frac{d\xi}{(1 + \xi^2)^2}$$

and so the k_2-dependence is completely specified. The l-dependence, on the other hand, is not so easily isolated. The most important element of the sequence, $\{\mu_n\}$, is μ_0 because it appears in conjunction with $\epsilon(Y, W)$. In fact,

$$\lambda_n \leqq \mu_0[(\alpha_n(2YW))^{\frac{1}{2}} + \epsilon(Y, W)/\mu_0^{\frac{1}{2}}]^2 \quad \text{for all} \quad Y.$$

Thus knowledge of μ_0 gives us one completely specified bounding expression for each λ_n. At this point, the question is whether this bound is better or worse than the μ_n for a given n. We do not calculate the μ_n here and leave this question open. Rather we shall study the former expression in an example.

Let $l = 0.1$ mile. In this case, we have done a computer calculation for μ_0; the result is $1.57(k_2)^{-1}$. The problem is to find the value of Y which minimizes the bound for a given W and k_2. For a fixed W, the first term, $\sqrt{\alpha_n(2YW)}$, is reduced by decreasing Y, and $\epsilon(Y, W)$ is reduced by increasing Y; therefore, the best Y is some compromise value. Let $W = (m\pi/k_2 l)$ and $Y = sk_2 l$ for $m > 0$ and $s > 2$; then

$$\lambda_n \leqq \mu_0\left[(\alpha_n(2\pi sm))^{\frac{1}{2}} + \left(\frac{1}{15.7}\left(\frac{e^{-2P}}{64} + \frac{3e^{-2P/3}}{64} + \frac{3}{64(s-2)^2} + \frac{3e^{-P/3}}{8\sqrt{10(s-2)}}\right)\right)^{\frac{1}{2}}\right]^2.$$

For the first eigenvalue, $(n = 0)$ when $\pi m \geqq 1$, then $\alpha_0(2\pi sm) \approx 1$ since $s > 2$ (Cf. Ref. 7, Fig. 2); consequently, μ_0 is the best bound available here. But when, for example, $\pi m = 0.25$ and s is chosen as 2.5, then

$$(\alpha_0(2\pi sm))^{\frac{1}{2}} \approx 0.8 \qquad \text{(Cf. Ref. 7, Fig. 2)}$$

and

$$\lambda_0 \leqq \mu_0[0.80 + 0.17]^2 \approx \mu_0(0.94) < \mu_0 .$$

In this case, it behooves us to use the more complicated bound. For 19-gage wire, this choice of m corresponds to a highest frequency of about 5×10^4 Hz.

For the tenth eigenvalue $(n = 9)$ when $\pi m = 1$, choose $s = 5.5$; then $(\alpha_0(2\pi sm))^{\frac{1}{2}} \approx 0.99$ and

$$\lambda_9 \leqq \mu_0[0.09 + 0.09]^2 \approx 0.032\ \mu_0 .$$

This is in the vicinity of the best choice for s. This means that for $W = (1/k_2 l)$ and $l = 0.1$ mile, the dimensionality of the crosstalk

functions relative to an error criterion, $0.032\ \mu_0$, is at most 10. It may be less than 10, but one needs lower bounds on the eigenvalues to determine that. Our technique does not carry over in any obvious way to a determination of lower bounds.

V. CONCLUSIONS

We have calculated bounds on the energy eigenvalues to determine the dimensionality of the crosstalk functions. Our analysis uses ideas developed by Landau and Pollak,[2] but our problem has a different character. Since any crosstalk function approaches zero for increasing frequency at a certain minimal rate, independent of the coupling function $u(x)$, the eigenvalues are insensitive to increases in the bandwidth W after a certain point, i.e., they saturate. This is indicated in the bounding expression by the presence of the μ_n which are independent of W. Hence, the question of which part of the bound is better depends on W: if W is large enough, the λ_n will have nearly saturated to μ_n and these are better, whereas smaller W-values are better handled by the more complicated expression which is sensitive to changes in W. This phenomena came up in our example for $n = 0$.

The dimensionality, as we have seen, presumes an error criteria. Given this, one can calculate from the bound, in any specific case, an upper bound on the dimensionality. The significance of this in a measurement program or in a crosstalk control scheme is to provide a realistic goal for reducing the number of independent measurements or controls, respectively. Since we have not derived lower bounds for the λ_n, the tightness of the upper bounds remains in question. Thus, the possibility of achieving greater reductions than our results would indicate is open.

VI. ACKNOWLEDGMENTS

It is a pleasure to acknowledge N. A. Strakhov for the original idea of generalizing the famous work of Landau and Pollak to dissipative systems. Also I am grateful to H. J. Landau for the benefit of much valuable criticism in the course of this work.

APPENDIX

We wish to bound the function

$$f(y, x) = \frac{x\mathrm{k}_1}{4\mathrm{k}_2\sqrt{\pi}} \int_{-\infty}^{y_0} e^{-|z|/\mathrm{k}_2}(y_0 - z)^{-\frac{3}{2}} \exp\left(-\frac{2x^2\mathrm{k}_1^2}{y_0 - z}\right) dz$$

by a more convenient function of y and x, both when $y > Y$ and when $y < -Y$. First suppose $y_0 \geqq 0$; then

$$\int_{-\infty}^{y_0/3} e^{-|z|k_2}(y_0 - z)^{-\frac{3}{2}} \exp\left(-\frac{2x^2k_1^2}{y_0 - z}\right) dz \leqq y_0^{-\frac{3}{2}}\left(\frac{2}{3}\right)^{-\frac{3}{2}} \int_{-\infty}^{\infty} e^{-|z|/k_2}\, dz = 3k_2 y_0^{-\frac{3}{2}}\left(\frac{3}{2}\right)^{\frac{1}{2}}.$$

Also,

$$\int_{y_0/3}^{y_0} e^{-z/k_2}(y_0 - z)^{-\frac{3}{2}} \exp\left(-\frac{2x^2k_1^2}{y_0 - z}\right) dz \leqq e^{-y_0/3k_2} \int_0^{\infty} \gamma^{-\frac{1}{2}} \exp(-2x^2k_1^2\gamma)\, d\gamma = e^{-y_0/3k_2} \frac{(\pi/2)^{\frac{1}{2}}}{xk_1},$$

where we have put $\gamma = 1/(y_0 - z)$ in the change of variables. Thus

$$f(y, x) \leqq \frac{3xk_1}{4}\left(\frac{3}{2\pi}\right)^{\frac{1}{2}} y_0^{-\frac{3}{2}} + \frac{1}{4k_2\sqrt{2}}\, e^{-y_0/3k_2}.$$

Next, suppose $y_0 = y_0'$ with $y_0' > 0$. Then, after changing z to $-z$ in the integral above, we have

$$f(y, x) \leqq \frac{xk_1}{4k_2\sqrt{\pi}}\, e^{-y_0'/k_2} \int_{y_0'}^{\infty} (z - y_0')^{-\frac{3}{2}} \exp\left(-\frac{2x^2k_1^2}{z - y_0'}\right) dz$$

$$\leqq \frac{1}{4k_2\sqrt{2}}\, e^{-|y_0|/k_2}.$$

The details of this calculation parallel the former case.

Since

$$\epsilon^2(Y, W) \leqq \int_0^l \int_{|y|>Y} |f(y, x)|^2\, dy\, dx,$$

we can bound $\epsilon^2(Y, W)$ by performing the integrations on the above bounds for $f(y, x)$. First we have

$$\int_{|y|>Y} |f(y, x)|^2\, dy \leqq \frac{1}{k_2}\left[\frac{e^{-2P(x)}}{64} + \frac{3e^{-2P(x)/3}}{64} + \frac{27x^2}{64\pi P(x)^2} + \min\left\{\frac{9xe^{-P(x)/3}}{16(P(x))^{\frac{1}{2}}}, \frac{3xe^{-P(x)/3}}{4(P(x))^{\frac{1}{2}}}\right\}\right],$$

where $P(x) = (Y - 2k_2x)/k_2$. In the latter term we have upper bounded $3/\pi$ by 1 and have put $(Y - 2k_2x)$ for y_0 (in the former case before

the y-integration and in the latter after one integration by parts). Before doing the x-integration, replace $P(x)$ by $P \equiv P(l)$, thus obtaining a greater bound; then the x-integration gives us

$$\epsilon^2(Y, W) \leqq \frac{l}{k_2}\left[\frac{e^{-2P}}{64} + \frac{3e^{-2P/3}}{64} + \frac{9l^2}{64\pi P^2} + \min\left\{\frac{9le^{-P/3}}{32P^{\frac{1}{2}}}, \frac{3le^{-P/3}}{8P^{\frac{1}{2}}}\right\}\right].$$

We shall always assume that $Y > 2k_2 l$, i.e., that $P > 0$.

REFERENCES

1. Whittaker, E. T., "On the Functions which are Represented by the Expansions of the Interpolation Theory," Proc. Roy. Soc. Edinburgh, Scotland, *35*, 1915, pp. 181–194.
2. Landau, H. J., and Pollak, H. O., "Prolate Spheriodal Wave Functions, Fourier Analysis, and Uncertainity III," B.S.T.J., *41*, No. 4 (July 1962), pp. 1295–1336.
3. Jewett, F. B., "Dr. George A. Campbell," B.S.T.J., *14*, No. 4 (October 1935), pp. 553–572.
4. Lechleider, J. W., "The Frequency Variability of Direct Crosstalk," unpublished work.
5. Kato, T., *Perturbation Theory of Linear Operators*, New York: Springer-Verlag, 1966.
6. Riesz, F., and Sz-Nagy, *Functional Analysis*, New York: Ungar, 1955.
7. Slepian, D., and Sonnenblick, E., "Eigenvalues Associated with Prolate Spheroidal Wave Functions of Zero Order," B.S.T.J., *44*, No. 8 (October 1965), pp. 1745–1759.
8. Landau, H. J., "The Eigenvalue Behavior of Certain Convolution Equations," Trans. Amer. Math. Soc., *115*, (March 1965), pp. 242–256.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 6, July-August, 1972
Printed in U.S.A.

The Camp-On Problem for Multiple-Address Traffic

By ERIC WOLMAN

(Manuscript received March 3, 1972)

A communication system for multiple-address messages is described, in which a message waits in parallel queues until it can be transmitted simultaneously to all the addressed receivers. An idealized mathematical model of this system leads to a nonlinear integral equation for the stationary distribution of delays in receiver queues. A phase-plane analysis shows this equation to have a one-parameter family of solutions, one member of which is found to be the unique limiting distribution of receiver delays. Even though service times (message lengths) are not bounded, the receiver queues in this model can operate in the steady state at critical load. Under these conditions, the probability that a server is idle is positive; and all moments of the delay distribution are finite. Computation of the delay distribution is discussed, numerical examples are given, and the behavior of the transmitter queues is analyzed.

Predictions of this model are compared with performance parameters of simulated systems. The model is shown to be very accurate up to its critical load. For higher loads, performance depends strongly upon the number of receivers in the system. The model's discontinuity in receiver occupancy is not physically realizable, but is approached asymptotically as the number of receivers tends to infinity.

Contents

I. INTRODUCTION

The frequency of conference calls in voice telephony is very low; but many of the messages carried by some data communication systems are directed from a single transmitter to two or more destinations. In the mathematical analysis of some types of data traffic, it is the presence of these multiple-address (MA) messages that raises problems essentially different from those encountered in the classical congestion theory of telephone systems.

Methods of coping with MA traffic fall into three classes. One involves *message switching* (also called *storage*, or the store-and-forward method), in which a transmitter sends a message to a switching center or other central location at which it is stored. Copies of the message are then retransmitted more or less independently to the desired receivers. A second class makes use of *selective calling*. In the simplest example, all stations are connected to a single channel, which may be thought of as a loop without a central switch; a message on this channel may be directed to any or all of the receivers. The number of simultaneous transmissions that is possible in a system of this class is limited by the number of space-, time-, or frequency-division channels in the loop. The third class of methods is based on *line switching*, in which a switching center merely establishes connections between the terminals of a system instead of providing storage for messages in transit. Line switching for MA traffic has itself two extreme forms, *sequential* and *simultaneous*, between which lie many other types of line switching. The sequential technique requires the transmitter to send each copy of a message to the proper receiver as a separate transmission, so that the transmitter sees an MA message as a sequence of single-address messages. With simultaneous line-switching, all copies of a message are sent at once after the transmitter has been connected to all the addressed receivers.

The term "camp-on" refers to one natural implementation of simultaneous line-switching. Suppose that a transmitter serves messages

* A nonmathematical statement of the problem and of the results obtained may be found in the starred sections.

offered to the system in order of arrival, and that a queue of delayed messages can form at each transmitter. When a message reaches the head of a transmitter queue, it gives its set of addresses to some control device, thereby requesting connection to several receivers. The simplest plan is for such requests to be entered into first-come, first-served queues corresponding to individual receivers. When a message (or copy thereof) reaches the head of a receiver queue, it (i.e., other copies) may still be awaiting other receivers that are not yet idle. Then the message *camps on* the available receiving line, so that the receiver, although idle, appears busy to other traffic. Thus it is possible for a delayed message to be waiting for two receivers neither of which is actually engaged in receiving a message; but reshuffling of the order in which messages are handled, so as to avoid this situation, would require involved computations, not to mention precautions against indefinite postponement of some transmissions.

For systems in which the lines to terminals radiate from a switching center, we may think of message switching, camp-on, and sequential line-switching as the three basic or "pure" techniques for handling MA traffic, in the sense that other schemes are really combinations of these three. There is much to be said for and against each of them. Message switching requires expensive storage-facilities, and saddles the communication system with the responsibility for messages in its possession. Sequential line-switching leads to excessively heavy loading of transmitters; and receivers are used with corresponding inefficiency in camp-on systems. I cite these *dis*advantages in order to point out that, in many situations, a practical system must combine some features of at least two of the pure techniques. In fact, the advantages of many compromise schemes are well-known. One possibility is to use storage for all MA messages, while sending single-address traffic over line-switched connections. In another plan, described as simultaneous transmission to available destinations (STAD), the transmitter is connected to all the addressed receivers that are not occupied. When this transmission is complete, a new attempt is made to reach the remaining addresses. STAD avoids factitious loading of receivers while holding down the number of transmissions needed per message.

The invention, design, and analysis of good arrangements for handling MA traffic are very difficult.* Our understanding of this subject is limited, although approximate analysis, simulation, and field measure-

* These difficulties are greatest in connection with line switching, practical interest in which has decreased as technological changes have made other schemes more attractive.

ment have yielded considerable data. It seems clear that a first step toward an adequate theory must be a thorough comprehension of each of the basic techniques mentioned above.

The presence of storage in message-switching systems decouples the two stages of blocking (at the transmitters and at the receivers) enough to allow application of standard delay-theories to each stage. No satisfactory analysis of sequential line-switching has been published; but the structure of such systems is not of deep theoretical interest, partly because the key question is that of the order in which the addressees of a message are served. The camp-on situation is of crucial importance because of its combinatorial structure. This structure is described in the next section. Here it is enough to say that the service process in the transmitter queues depends on the way in which messages interact in the receiver queues.

This paper describes the analysis of a very simple and idealized model of a line-switching system with camp-on. The answer obtained is a description, valid in the limit of very large systems, of message delays during steady-state operation. This problem is one of a large class of problems, characterized by complex interactions between queues in parallel, which forms an important domain on the frontier of congestion theory. The analysis presented here is important for two reasons: It yields insight into one of the basic procedures for handling multiple-address traffic, and should therefore lead toward an understanding of more realistic models; and it embodies a modification, which may prove useful in similar problems, of a standard method in queuing theory. In addition, the remarkable behavior exhibited by this idealized model renders it of independent interest.

The problem treated here, and others not dissimilar, have been discussed by other authors. The most relevant paper is by Haenschke.[1] But discussion of other work is deferred until Section VII so that different approaches may be compared in detail.

II. A CAMP-ON MODEL

2.1 *The System*

We first consider a system with one switching-center, X transmitters, and R receivers. Messages are offered to the system in X independent arrival-processes, one for each transmitter; and each process is Poisson with rate α. Every message is addressed to exactly m of the receivers. This m is an integer greater than 1. The m addresses of a message are chosen at random uniformly from the R possibilities, and the addresses

of any two messages are chosen independently. The unit of time is the mean length of the messages, and these lengths are independent random variables with the negative exponential distribution. By symmetry, all the receivers experience identical arrival processes. Let their mean rate be λ. If the system can operate in equilibrium, the number of copies of messages that must leave the transmitters is $m\alpha X$ per unit time. Since λR copies per unit time arrive in the receiver queues, the conservation equation in equilibrium must be $m\alpha X = \lambda R$. A useful form of this equation is

$$\frac{R}{X} = \frac{m\alpha}{\lambda} = k, \tag{1}$$

where we have introduced the parameter k to denote the ratio of the numbers of receivers to transmitters.

The previous paragraph, together with the description of camp-on given in Section I, completely specifies the traffic characteristics of a communication system. Except for its mean rate λ, we are ignorant of the arrival process at the receiver queues; but it is determined by the stated requirements. We also know nothing of the conditions for the existence of a steady state, although clearly λ and α cannot be too large. The "service time" of a message in a transmitter queue consists of all the time spent at the head of that queue before departure–that is, of the transmission time plus the time spent waiting for the last of the addressed receivers to become available. The situation is shown in Fig. 1 for the case in which $m = 2$. Each circle at the upper level represents a message in a transmitter queue, and at the lower level a copy in a receiver queue. The lines in the figure connect each message-symbol at the head of a transmitter queue to the symbols for the two corresponding copies. The head messages at transmitters I and V are being transmitted, to receivers 1 and 2 in the former case and 5 and 6 in the latter. The message at II is awaiting receiver 2 and camping

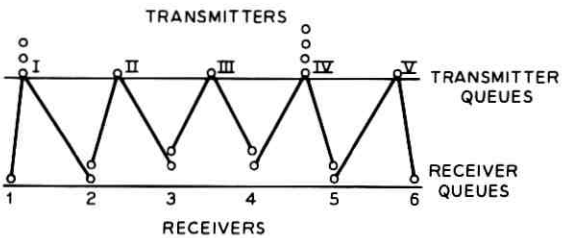


Fig. 1—The queuing discipline in a camp-on system.

on 3; and similarly for the head message at IV. Although neither of receivers 3 and 4 is receiving a message, both are being camped on and act occupied. Therefore the message at III, which announced its addresses after the head messages at II and IV did so, is, as shown in Fig. 1, effectively in third place in queuing for receivers.

The system just described is unrealistically simple in several ways. First, it consists of a single switching-center with lines to receivers and transmitters radiating from it, whereas a physical system could have many switching-centers interconnected by trunks. This difference is not critical, because we are studying principally the effect of having to wait simultaneously in m receiver-queues. Blocking due to inadequacy of trunk facilities could be made negligible in comparison. Second, signaling and switching times are taken to be zero, a mathematical convenience of long standing in queuing theory and often justified in practice. Third, each message has the same number of addresses. This assumption is far indeed from being realistic, but we shall see in Section 5.2 that it does not seriously restrict the usefulness of the model. Fourth, the transmitters and receivers are equally and independently loaded. This is a genuine restriction, especially for the receivers. A wide variation among receiver loads would represent reality better, particularly if a destination with heavy traffic could have several receivers sharing the load. Traffic flow in some physical systems is further complicated by more restrictive geometries. For example, a transmitter and receiver can be connected to a switching-center by a single line not capable of full-duplex operation (simultaneous transmission in both directions). The simple model adopted here is required to avoid prohibitive mathematical complexity, but it has the corresponding virtue of introducing no complications other than that inherent in the camp-on discipline itself.

In order to proceed, we need symbols for various portions of the time that a message spends in the system. Figure 2 shows this time-interval for a particular message, and is drawn from the viewpoint of *one* of the m receivers to which the message is addressed. At the point A in Fig. 2, the message arrives and joins a transmitter queue.

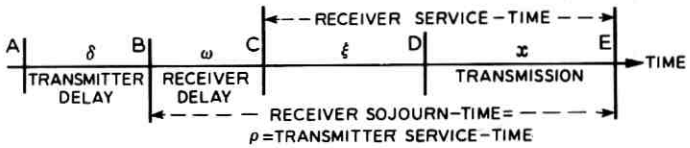


Fig. 2—Composition of the time spent by a message in the system.

At B the message reaches the head of this queue and notifies the control of its addresses; thus the message arrives in the distinguished receiver's queue (and in those of the other $m - 1$ receivers) at time B also. The interval from A to B is the *transmitter delay*, and is represented by the random variable δ.

At C, after the *receiver delay* (or waiting time) ω, the message reaches the head of the receiver queue and camps on the receiver. The longest of the m receiver-delays ends at time D, and transmission then begins. The random variable ξ represents the excess of the longest of the receiver delays over the delay at this receiver (the one for which the diagram is drawn). Of course $\xi = 0$ if the longest delay occurs at this receiver. If $\omega^{(i)}$ is the delay at the ith receiver, then ξ is defined by the equation

$$\omega^{(j)} + \xi^{(j)} = \max (\omega^{(1)}, \cdots, \omega^{(m)}) \quad \text{for each} \quad j = 1, \cdots, m.$$

The dummy indices run, of course, over the addresses of one message. Omitting superscripts for "our own" receiver, we write the simpler formula

$$\omega + \xi = \max_{i \leqq m} \omega^{(i)}. \tag{2}$$

The message is actually transmitted during the interval of length x, with density function $\exp(-x)$, that runs from D to E. At time E the message ends and departs from the system.

The interval from B to E, of random length $\rho = \omega + \xi + x$, is the *sojourn time* of the message in the receiver queue. During this entire time the message occupies its transmitter, so that ρ is also the *service time* in the transmitter queue. Indeed, the transmitter queue is a single-server queue, with Poisson arrivals, whose service times ρ have an unknown distribution and are not independent.

The reader who is interested in the results of this investigation, rather than in their derivation, may now skip to Section VII, which contains a summary of the argument and a discussion of its conclusions.

2.2 *Passage to a Limit*

The trick that makes this model amenable to analysis is this: We let R and X both go to infinity while holding their ratio constant. In fact–see eq. (1)–we keep k, α, λ, and m all fixed as the system gets larger. In general terms, the purpose of this trick is to decouple the queues that operate in parallel, while preserving the essential interaction caused by each message's having to wait in m queues at once. The detailed consequences of this procedure are three in number, all

of them necessary to further progress. The first of these is that the receiver queues become independent. In other words, in the infinitely large system obtained by letting R and X approach infinity, the m copies of the random variable ω that are encountered by each message are independent. I have no rigorous proof of this fact, which is, however, buttressed by the following heuristic argument.

Consider the m receivers addressed in a particular message. Before this message reaches the head of a transmitter queue, the m receiver-queues may have various lengths. Then, at time B in the sketch of Fig. 2, the message arrives simultaneously in all m queues. When, at time E, the message departs from the system, it leaves behind it the same *expected* number of messages in each of the m queues: For this message had the same sojourn-time in each of these queues, which are all subjected to arrival processes of the same intensity λ. Let us then call the departure of this message a *synchronizing event* (*S-event*) for these m queues, using "synchronizing" in a rather weak sense.

Each receiver queue participates in λ S-events per unit time, on the average, if the system is in equilibrium. Consider a fixed pair of receivers which have just participated in the same S-event. The probability that any single subsequent message addressed to receiver 1 will generate an S-event involving receiver 2 is $(m - 1)/(R - 1)$, since the $m - 1$ addresses of such a message (other than receiver 1) are chosen with equal probability from among the $R - 1$ other receivers. The S-events involving receiver 1 form a sequence of Bernoulli trials with respect to the probability of involving receiver 2, because the addresses of different messages are chosen independently. Therefore the mean interval between successive S-events involving two particular receivers is $(R - 1)/(m - 1)$ events (corresponding to one of the receivers), and this is equal to $(R - 1)/[\lambda(m - 1)]$ units of time. This quantity increases without bound when we let $R \to \infty$ while holding λ and m constant. In other words, in the infinitely large system obtained by means of our trick, the mean time between successive "synchronizations" of two given receiver queues is infinite.

Indirect methods of "synchronization" also affect the receiver queues. If receivers 1 and 3 participate in an S-event, and if shortly thereafter receivers 2 and 3 do so, we might say that the queues at receivers 1 and 2 are connected by "an S-chain of length 2." The effect of such chains is, of course, to increase the degree of statistical dependence between the states of the queues involved. But the longer an S-chain is, the less effective it can be in "synchronizing" two queues, because of the time-lag between the S-events that form the ends of

the chain. Here we do not investigate the relative efficacy, in increasing the dependence of the states of any pair of receiver queues, of S-chains of different lengths. But the time required, on the average, to complete an S-chain of given length between two specified queues clearly approaches infinity as $R \to \infty$, so that the efficacy of any such chain vanishes in the infinite model.

This argument for the independence of receiver queues can be summarized by saying that as $R \to \infty$ the frequency of occurrence, of events through which the state of one receiver queue can influence the state of any other particular receiver queue, goes to zero. Also, each receiver in this statement can be replaced by any fixed, finite set of receivers.

We come now to the second consequence of passing to an infinitely large system, which is that the service times of successive messages in any transmitter queue are independent. In the finite system, the probability that the sets of m addresses of two messages have one or more receivers in common is easily found to be $m^2/R + O(R^{-2})$. This quantity tends to zero as $R \to \infty$. A similar calculation shows that, in the infinite system, the address-sets of any finite group of messages overlap with probability zero. The service times of any set of messages in a transmitter queue, which are just their respective sojourn-times in the receiver queues, therefore depend on the states of nonoverlapping (with probability one) sets of receiver queues; and these we have already found to be independent. Thus each transmitter queue is of type M/G/1–that is, has Poisson arrivals, one server, and general independent service-times.

(Notice that both these arguments are not uniform in λ: Fixing R, we can choose λ large enough to ensure a significant degree of dependence between the states of various receiver queues and also between successive transmitter service-times. Thus we must allow R to go to infinity before varying λ.)

The third consequence of our trick is that the arrival process at each receiver is Poisson. In the finite system each such process, or, in Khinchin's[2] words, stream of events, is the union of X substreams, each of intensity λ/X or $m\alpha/R$. Each event in one substream (except for an arrival at an idle transmitter) coincides with the departure of a message from a transmitter queue, for that is the instant at which the next message, if any, announces its addresses and joins m receiver-queues. Such a substream is obtained by selecting at random a fraction m/R of the events in the departure process from one transmitter, this process being itself a stream of intensity α. The resulting substreams

which make up the arrival process at one receiver have, in the steady state, these properties: They are orderly and stationary (see Khinchin[2]), and, by the arguments given above, they approach mutual independence as R and $X \to \infty$. A limit theorem (see Ref. 2, Chap. 5, and the similar but slightly earlier results of Cox and Smith[3]) suggests that as $X \to \infty$ the arrival stream becomes Poisson. This result requires that one technical condition [Ref. 2, p. 50, condition (2)] be fulfilled; this is easily verified. It also requires that Khinchin's version of the theorem hold when the substreams of decreasing intensity λ/X are not independent, but merely *approach* independence as $X \to \infty$; and this is assumed here.

Let us review the properties of our final model, focusing, as in Fig. 2, on what happens to a single message. The queuing system consists of two stages. The first stage is an ordinary M/G/1 queue with an unknown distribution of service times. Service in this queue consists of a sojourn in the second-stage queue. The latter has Poisson arrivals and an exponential departure-mechanism (known as transmission) with unit mean. But transmission in the second queue begins only after the expiration of the longest of m independent, identically distributed intervals, each of which is what would ordinarily be called the waiting time in the second queue.

We constructed this model by starting with a very simple but physically realizable finite system and modifying it in an appropriate way. The properties of the final model were deduced informally. But this model is of interest in its own right, and could have been proposed for study at the start. Its properties are derived here in order to demonstrate its connection with the camp-on problem. It is also fair to say that the interactions which it is the purpose of our limiting-procedure to remove–namely, dependence of receiver delays, dependence of transmitter service-times, and deviation from Poisson character of receiver arrivals–are demonstrably small in a large finite system at moderate loads. (A system is large in this sense, if a receiver gets hardly any of its messages from any single transmitter and every message is addressed to a very small fraction of all receivers.) The effects of such interactions are discussed further in Section 5.4.

2.3 *Equation for Delay Distribution in Equilibrium*

If we knew the distribution of sojourn times in the receiver queues, which are identical with the service times in the transmitter queues, we could in principle determine the distribution of transmitter delays from the well-known formula of Pollaczek and Khinchin [Khinchin,[2]

p. 116; Cox and Smith,[4] Section 2.6, eq. (28)]. Lacking closed forms for the various quantities involved, we could still find the mean transmitter-delay, given the mean and variance of the receiver sojourn times [Khinchin,[2] p. 117; Cox and Smith,[4] Section 2.6, eq. (22)]. We therefore turn our attention exclusively to the receiver queues, which can be studied without reference to the larger system of which they are a part. The results of this investigation are then used to describe the transmitter queues, which are not mentioned again until Section VI.

We now reproduce in Fig. 3 the relevant part of Fig. 2, drawn as before for one receiver, but extended so as to show the relation between two successive messages addressed to that receiver. Subscripts refer to messages, in order of arrival; in particular, ω_n is the delay suffered by the nth message and ω_{n+1} that of the next. We also need symbols for the as-yet-unknown distribution functions for the receiver delays ω and sojourn times ρ. These are respectively F and G. Thus G_n is the sojourn-time distribution for the nth message, and F_{n+1} the delay distribution for its successor.

Proceeding in the spirit of Lindley[5] (or see Cox and Smith,[4] Section 5.3), we now relate the delay distributions of successive customers. We can do this by using two simple integral-relations, each of which amounts merely to a definition. Note first that ρ_n is the sum of two random variables: x_n , whose density function is the unit exponential, and $\omega_n + \xi_n$, which we know from (2) to be the maximum of m independent copies of ω_n . The probability that this maximum does not exceed t is the product of the probabilities that each of the m variables ω_n does not exceed t. Thus the distribution function for $\omega_n + \xi_n$ is

$$\Pr\{\omega_n + \xi_n \leqq t\} = [F_n(t)]^m.$$

The usual convolution-formula gives for $G_n(t)$, the probability that ρ_n does not exceed t, the value

$$G_n(t) = \int_0^t e^{-(t-u)} F_n^m(u)\, du, \tag{3}$$

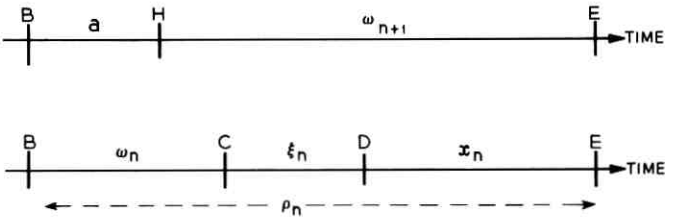


Fig. 3—Receiver delays of successive messages.

where $F_n^m(u) \equiv [F_n(u)]^m$.

An equally simple calculation yields F_{n+1} . Let a in Fig. 3 be the interval between the arrival time of the nth message (at B) and that of the next (at H). The delay ω_{n+1} runs from time H to time E, stopping then because the nth message departs. (Message $n + 1$ can be transmitted only after a further interval of length ξ_{n+1} , which may of course be zero.) Certainly ω_{n+1} is less than t if ρ_n is less than t, since time H cannot precede time B. Suppose, on the other hand, that ρ_n has the value $u > t$. Then ω_{n+1} is less than t if and only if a is greater than $u - t$. But a is an interarrival time in a Poisson process of intensity λ. It exceeds $u - t$ if and only if no arrival occurs during an interval of length $u - t$, and this event has probability $\exp(-\lambda(u - t))$. Rewriting this argument in symbols, we get the formula

$$F_{n+1}(t) = G_n(t) + \int_t^\infty e^{-\lambda(u-t)}\, dG_n(u). \tag{4a}$$

Integration by parts gives us the useful equivalent formula

$$F_{n+1}(t) = \lambda \int_t^\infty e^{-\lambda(u-t)} G_n(u)\, du. \tag{4b}$$

We can represent F_{n+1} directly in terms of F_n by eliminating G_n between eqs. (3) and (4). We note first from (3) that

$$G_n(0) = 0. \tag{5}$$

This says that no message has a sojourn time of length zero; this is to be expected because ρ_n is at least as long as x_n , whose distribution has no mass at the origin. Equation (5) also removes any doubts about the integration by parts that yields (4b), in case $t = 0$.

Equation (3) also tells us that G_n has a derivative. Differentiating the right-hand member explicitly, and denoting time-differentiation by a dot, we find that

$$\dot{G}_n(t) = -G_n(t) + F_n^m(t). \tag{6}$$

Writing $dG_n(u) \equiv \dot{G}_n(u)\, du$, we can substitute (3) and (6) into (4a). Using (4b) to eliminate the remaining appearance of G_n , we obtain the important equation

$$F_{n+1}(t) = \frac{\lambda}{1 + \lambda}\left[\int_0^t e^{-(t-u)} F_n^m(u)\, du + \int_t^\infty e^{-\lambda(u-t)} F_n^m(u)\, du\right]. \tag{7}$$

A useful special case of (7) comes from setting $t = 0$. This is

$$F_{n+1}(0) = \frac{\lambda}{1+\lambda}\int_0^\infty e^{-\lambda u}F_n^m(u)\,du. \tag{8}$$

If we multiply both members of this equation by exp (λt) and divide the interval of integration into two parts at t, we find that

$$F_{n+1}(0)e^{\lambda t} = \frac{\lambda}{1+\lambda}\int_0^t e^{-\lambda(u-t)}F_n^m(u)\,du + \frac{\lambda}{1+\lambda}\int_t^\infty e^{-\lambda(u-t)}F_n^m(u)\,du.$$

We can use this to eliminate the integral from t to ∞ in (7), obtaining the "retrospective" formula

$$F_{n+1}(t) = F_{n+1}(0)e^{\lambda t} - \frac{\lambda}{1+\lambda}\int_0^t [e^{\lambda(t-u)} - e^{-(t-u)}]F_n^m(u)\,du. \tag{9}$$

It is possible to base the analysis of this problem entirely on the equations obtained so far, but much of the argument is simpler if we use an equivalent formulation which involves differential instead of integral equations. Differentiation of eq. (4b) shows that $\dot{F}_{n+1}$ exists for $t > 0$ and is given by

$$\dot{F}_{n+1}(t) = \lambda^2\int_t^\infty e^{-\lambda(u-t)}G_n(u)\,du - \lambda G_n(t).$$

For $t = 0$ this must be interpreted as a derivative on the right; the jump in F_{n+1} at $0-$ can be ignored if we remember that $F_{n+1}(0) > 0$. Substitution of (4b) and suppression of the argument t yield

$$\dot{F}_{n+1} = \lambda F_{n+1} - \lambda G_n\,. \tag{10}$$

Since both terms of the right member have derivatives, it is also true that

$$\ddot{F}_{n+1} = \lambda\dot{F}_{n+1} - \lambda\dot{G}_n\,,$$

where $\ddot{F}_{n+1}$ is also a right-derivative at $t = 0$. The sum of the last two equations is

$$\ddot{F}_{n+1} + (1-\lambda)\dot{F}_{n+1} - \lambda F_{n+1} = -\lambda(\dot{G}_n + G_n).$$

Using (6) to eliminate G_n and its derivative, we obtain

$$\ddot{F}_{n+1} + (1-\lambda)\dot{F}_{n+1} - \lambda F_{n+1} = -\lambda F_n^m, \tag{11a}$$

the differential version of (7). This must be accompanied by the boundary condition

$$\dot{F}_{n+1}(0) = \lambda F_{n+1}(0), \tag{11b}$$

which comes from (10) with the aid of (5). Also, since F_n is a distribution function, $0 \leqq F_n(t) \leqq 1$ for all $t \geqq 0$. Substitution of 1 as an upper bound for F_n in (7), where both kernels are positive, easily yields the fact that

$$0 \leqq F_{n+1}(t) < 1 \quad \text{for all } t \geqq 0. \tag{11c}$$

This inductive result, based essentially on the assumption that F_1 is a distribution function, serves as a second boundary-condition for (11a), one of whose homogeneous solutions is exp (λt).

The derivations of (7) and (11) given above are chosen for brevity, and both make use of equations involving G_n . The discussion that follows rests on the fact that, as self-contained descriptions of the sequence $\{F_n\}$, (7) and (11) are equivalent. We should therefore verify that each can be obtained from the other. We got (11c) from (7) anyway, and explicit differentiation of (7) quickly yields (11b) and (11a). On the other hand, treatment of (11a) as an inhomogeneous linear equation with driving function $-\lambda F_n^m$ allows us to use the standard formula for its general solution. Application of (11b) then yields the representation (9), which is the natural form when integrating in the positive t-direction. As it stands, equation (9) seems to contain the wrong homogeneous solution, but actually the integral term is a negative correction which offsets the growth of exp (λt). We can rearrange (9) so as to represent $F_{n+1}(t)$ as the sum of two quantities: One is the right-hand member of (7), which we know is bounded when $0 \leqq F_n \leqq 1$, and the other is the correction

$$e^{\lambda t}\left[F_{n+1}(0) - \frac{\lambda}{1+\lambda}\int_0^\infty e^{-\lambda u} F_n^m(u)\, du\right].$$

Since the bracketed quantity is independent of t, either it vanishes or $| F_{n+1}(t) |$ grows without limit as $t \to \infty$. The condition (11c) selects the former possibility, proving that both (8) and (7) hold. Thus the systems (7) and (11) are indeed equivalent.

Again following Lindley,[5] we impose the condition of equilibrium by specifying that two successive messages must have the same delay-distribution. The distinction between the nth message and its successor can be removed by omitting the subscripts in (7). This leaves the formula

$$F(t) = \frac{\lambda}{1+\lambda}\left[\int_0^t e^{-(t-u)} F^m(u)\, du + \int_t^\infty e^{-\lambda(u-t)} F^m(u)\, du\right]. \tag{12a}$$

This is our fundamental nonlinear integral equation for F, the equilibrium distribution of receiver delays. (Thus the distribution function for $\omega + \xi$ is F^m.) Since (12a), unlike (7), makes no implicit reference to the initial distribution F_1, we must adjoin to it the explicit condition

$$0 \leqq F(t) \leqq 1. \tag{12b}$$

III. SOLUTIONS

3.1 *Reduction to a First-Order System*

We are looking for distribution functions F that satisfy our basic system of equations (12) on the interval $0 \leqq t < \infty$. We expect on physical grounds that (12) has solutions for sufficiently small positive λ. We also expect that, for large enough values of λ, (12) has no solutions; for these large arrival-rates the queue lengths and delays should increase without limit, and no steady state should exist.

In this section we show that indeed there exists a critical arrival-rate Λ above which (12) has no solutions, and that for $\lambda \leqq \Lambda$ it has infinitely many solutions. In the latter case the solutions fall into two groups, one of which contains a single distribution and the other a one-parameter family of distributions. One's natural inclination to choose the distinguished solution as the answer to the physical problem encounters a major difficulty: The set of distinguished solutions obtained by varying λ has the property that their initial values, which represent the respective probabilities of finding a receiver idle, are bounded away from zero as λ approaches Λ from below. The paradoxical implication that the receivers have spare capacity at their critical load makes it unclear whether the idealized problem (with $R = \infty$) has a unique solution at all. This issue is discussed further below and resolved in Section IV.

We begin by observing two consequences of the system (12). First, its solutions must be what we may call *unlimited*. According to (12b), the right member of (12a), evaluated for $F = 1 = F^m$, is an upper bound for F; explicitly,

$$\frac{\lambda}{1+\lambda}\left[e^{-t}\int_0^t e^{u}\,du + e^{\lambda t}\int_t^\infty e^{-\lambda u}\,du\right] = 1 - \frac{\lambda}{1+\lambda}e^{-t}. \tag{13}$$

[This is the formula alluded to in the derivation of eq. (11c).] This result rules out functions F (called *limited*) which reach 1 at finite values of t. They would correspond, if they existed, to operation with receiver delays bounded above by some finite limit; and this situation

seems physically unreasonable because transmission times are not bounded. In particular, F cannot reach the value 1 with positive slope.

Improper equilibrium-distributions, corresponding to a positive probability of infinite delay, also cannot occur. This situation is represented by distributions which never increase beyond some positive value less than 1. That this cannot happen is shown next; but here we note in passing the trivial exception that $F \equiv 0$ is a formal solution of (12).

By defining the kernel

$$k_\lambda(t, u) \equiv \begin{cases} [\lambda/(1 + \lambda)] \exp(-(t - u)), & u \leqq t, \\ [\lambda/(1 + \lambda)] \exp(-\lambda(u - t)), & u \geqq t, \end{cases} \tag{14}$$

we can write (12a) in the form

$$F(t) = \int_0^\infty k_\lambda(t, u) F^m(u)\, du. \tag{15}$$

Suppose there were a bound $b < 1$ such that $0 \leqq F \leqq b$ for some solution F of (12a) and (15). Then, k_λ being positive,

$$F(t) \leqq b^m \int_0^\infty k_\lambda(t, u)\, du < b^m,$$

with the final inequality coming from (13). Substitution of this result into (15) shows that $F(t) < b^{m^2}$; thus $F(t) < b^{m^j}$ for arbitrarily large j by iteration, and so $F(t) \equiv 0$. Therefore all nontrivial solutions have the property that

$$F(t) \to 1 \quad \text{as} \quad t \to \infty \tag{16}$$

and are called *proper*.

One other useful formula comes directly from (12a). Differentiation of that equation tells us that

$$\dot{F}(t) = \lambda F(t) - \lambda \int_0^t e^{-(t-u)} F^m(u)\, du.$$

Since the last term is nonpositive,

$$\dot{F}(t) \leqq \lambda F(t), \quad t \geqq 0. \tag{17}$$

It is now convenient to switch to the differential formulation of the problem. Omitting all subscripts from (11), we get for equilibrium the system

$$\ddot{F} + (1 - \lambda)\dot{F} = \lambda F - \lambda F^m, \tag{18a}$$

$$\dot{F}(0) = \lambda F(0), \tag{18b}$$

$$0 \leqq F(t) < 1 \quad \text{for} \quad 0 \leqq t < \infty, \tag{18c}$$

which can also be derived from (12) and the fact that solutions of (12) are unlimited. Since t does not appear explicitly in (18a), we can transform it into an equation of first order by introducing a function $p(F)$ with these properties:

$$p(F) = \dot{F}; \qquad \ddot{F} = \frac{d\dot{F}}{dt} = \frac{dp}{dt} = \frac{dp}{dF}\cdot\frac{dF}{dt} = p'p, \tag{19}$$

where $p' \equiv dp/dF$. We also use the notation $F_0 \equiv F(0)$ so that the integration for $p(F)$ can begin with $p(F_0)$; the previous subscripts should cause no confusion, since there is no zeroth message. Equations (18a) and (19) yield, after some rearrangement,

$$p' = \frac{\lambda(F - F^m)}{p} - (1 - \lambda). \tag{20a}$$

The boundary condition (18b) becomes

$$p(F_0) = \lambda F_0 . \tag{20b}$$

Admissible solutions of (20a) must have three other properties. Since we are seeking only solutions corresponding to distribution functions, which must be monotone, p must be nonnegative. Second, (17) holds. We also have the condition (18c), except that, since every solution of (18) satisfies (16) but is unlimited, every solution of (20a) must include the point $(F, p) = (1, 0)$ (though with parameter $t = \infty$). In symbols,

$$0 \leqq p \leqq \lambda F; 0 \leqq F < 1 \quad \text{except at} \quad (F, p) = (1, 0). \tag{20c}$$

Our original problem is now reduced to that of solving the first-order system (20). One more integration then yields $F(t)$, as we see from the first part of (19), which can be written $dt = dF/p(F)$.

3.2 *Trajectories in the Phase Plane*

Equation (20a) is best studied in the Fp-plane, which is properly called the phase plane for (18) because $p = \dot{F}$. Figure 4 shows the relevant region of this plane for the case in which $\lambda < \Lambda$. According to (20c), meaningful solutions must lie above the F-axis, within the strip $0 \leqq F \leqq 1$, and below the oblique line $p = \lambda F$. The formula (20a) defines a vector field, and we must study those of its integral curves that lie within the triangle just described. These integral curves, or trajectories, are parametrized by t according to the results of the

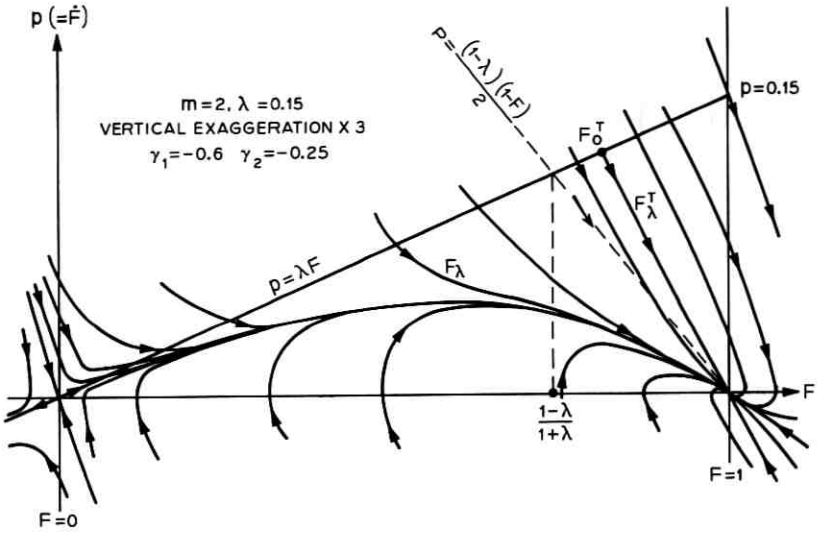


Fig. 4—Phase-plane trajectories for subcritical load.

second integration. The boundary condition (20b) tells us that the desired trajectories begin (with $t = 0$ and $F = F_0$) on the line $p = \lambda F$.

What does the vector field look like? Since F is nondecreasing, the field points to the right (for increasing t), or vertically if $p = 0$. Along the right edge of the triangle, where $F = 1$ and $p > 0$, (20a) shows that $p' = -(1 - \lambda)$. Some of these field vectors are shown in Fig. 4. They are drawn for $\lambda < 1$; this inequality is shown below to be appropriate. Along the top edge of the triangle, (20a) reduces to $p' = \lambda - F^{m-1}$. This agrees with the right-edge field at the corner $(1, \lambda)$, and points increasingly upward as we progress toward the origin. The field points into the triangle everywhere on the open top edge; in the limit it is parallel to the top edge at (0, 0). Along the bottom edge, $p' = +\infty$. The plus sign is chosen because, for fixed F and small positive p, p' assumes large positive values. The field is singular at the two corners (0, 0) and (1, 0), since p' assumes different limiting values depending on the direction from which each of these points is approached.

In studying these two singularities we mainly follow Chapter 15, "Perturbation Theory of Two-Dimensional Real Autonomous Systems," of Coddington and Levinson;[6] however, the present discussion is self-contained, though succinct. We begin by writing (20a) as a pair of

simultaneous parametric equations. The standard form, obtained from (19) and (20a), is

$$\begin{aligned} \dot{F} &= p, \\ \dot{p} &= \lambda F - (1 - \lambda)p - \lambda F^m. \end{aligned} \tag{21}$$

The system (21) is suitable for studying the singularity at (0, 0), but it is convenient to examine the one at (1, 0) first. For this purpose we introduce the variable $\phi = 1 - F$, which transforms (21) into

$$\begin{aligned} \dot{\phi} &= -p, \\ \dot{p} &= \lambda(m - 1)\phi - (1 - \lambda)p - \lambda \sum_{i=2}^{\infty} \binom{m}{j}(-\phi)^i \end{aligned} \tag{22}$$

in the ϕp-plane. (Of course the binomial expansion terminates when m is an integer. The significance of nonintegral m is discussed in Section 5.2.) Omission of the higher-order terms leaves the unperturbed system

$$\begin{aligned} \dot{\phi} &= -p, \\ \dot{p} &= \lambda(m - 1)\phi - (1 - \lambda)p, \end{aligned} \tag{23}$$

which is linear and has its only singularity at $(\phi, p) = (0, 0)$. As on p. 371 of Ref. 6, this can be written in the form

$$\begin{bmatrix} \dot{\phi} \\ \dot{p} \end{bmatrix} = A_1 \begin{bmatrix} \phi \\ p \end{bmatrix}, \qquad A_1 = \begin{bmatrix} 0 & -1 \\ \lambda(m - 1) & -(1 - \lambda) \end{bmatrix}. \tag{23a}$$

The determinant of A_1 is $\lambda(m - 1)$, which does not vanish because $\lambda > 0$ and $m > 1$.

The eigenvalues of A_1 satisfy the quadratic equation

$$\gamma^2 + (1 - \lambda)\gamma + \lambda(m - 1) = 0 \tag{24}$$

with discriminant $(1 - \lambda)^2 - 4\lambda(m - 1)$, and can therefore be written

$$\begin{aligned} \gamma_1 &= -(1/2)\{1 - \lambda + [(1 + \lambda)^2 - 4\lambda m]^{\frac{1}{2}}\}, \\ \gamma_2 &= -(1/2)\{1 - \lambda - [(1 + \lambda)^2 - 4\lambda m]^{\frac{1}{2}}\}. \end{aligned} \tag{25}$$

The discriminant of these expressions vanishes when $\lambda^2 - 2(2m - 1)\lambda + 1 = 0$; that is, at these two values of λ:

$$\Lambda(m) = 2m - 1 - 2(m^2 - m)^{\frac{1}{2}}, \tag{26a}$$

$$\Upsilon(m) = 2m - 1 + 2(m^2 - m)^{\frac{1}{2}}. \tag{26b}$$

(These symbols are mnemonic for "lower" and "upper"; also $\Lambda(m)$ turns out to be the critical load Λ mentioned above.) The discriminant in (25) represents a parabola in λ opening upwards, and it takes the values 1 at $\lambda = 0$ and $4(1 - m)$ at $\lambda = 1$. Since $m > 1$, we see that $0 < \Lambda(m) < 1 < \Upsilon(m)$ and that the discriminant is negative only when λ lies between Λ and Υ. When the discriminant is positive it is smaller than $(1 - \lambda)^2$, so that the eigenvalues γ_i are of one sign. The range of arrival rates λ can now be divided as follows, according to the nature of the eigenvalues of A_1 :

Case 1. $0 < \lambda < \Lambda$: γ_i distinct, real, negative;
Case 2. $\lambda = \Lambda$: both $\gamma_i = (\lambda - 1)/2 < 0$;
Case 3. $\Lambda < \lambda < \Upsilon$: γ_i complex conjugate;
Case 4. $\lambda = \Upsilon$: both $\gamma_i = (\lambda - 1)/2 > 0$;
Case 5. $\lambda > \Upsilon$: γ_i distinct, real, positive.

This classification allows us to put A_1 into the canonical forms listed in Ref. 6. We treat the cases in order, beginning with the first and most important, which is represented in Fig. 4.

Case 1: We introduce new coordinates (x_1 , y_1) by means of the linear transformation

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = T_1 \begin{bmatrix} \phi \\ p \end{bmatrix} . \tag{27}$$

If we choose for this operator the real nonsingular matrix

$$T_1 = \begin{bmatrix} \lambda(m - 1) & \gamma_2 \\ \lambda(m - 1) & \gamma_1 \end{bmatrix} \tag{28a}$$

with inverse

$$T_1^{-1} = \delta_1^{-1} \begin{bmatrix} \gamma_1 & -\gamma_2 \\ -\lambda(m - 1) & \lambda(m - 1) \end{bmatrix} , \tag{29a}$$

where the determinant $\delta_1 < 0$, then the linear system in (23a) becomes

$$\begin{bmatrix} \dot{x}_1 \\ \dot{y}_1 \end{bmatrix} = J_1 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} , \qquad J_1 = T_1 A_1 T_1^{-1} = \begin{bmatrix} \gamma_2 & 0 \\ 0 & \gamma_1 \end{bmatrix} . \tag{30a}$$

The transformed matrix J_1 has the canonical form listed as (II) on p. 371 of Ref. 6; the trajectories of (30a) near the x_1y_1-origin (an improper node) are sketched in Fig. 5 on p. 373 of Ref. 6. All these trajectories

approach $(x_1, y_1) = (0, 0)$ as $t \to \infty$. The positive and negative y_1-axes are trajectories; all the other integral curves have zero as their limiting slope at the origin and reach the origin tangent to the x_1-axis.

Since the nonlinear system (22) can be written

$$\begin{bmatrix} \dot{\phi} \\ \dot{p} \end{bmatrix} = A_1 \begin{bmatrix} \phi \\ p \end{bmatrix} + \begin{bmatrix} 0 \\ -\lambda \sum_{j=2}^{\infty} \binom{m}{j} (-\phi)^j \end{bmatrix},$$

the corresponding transformed system is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{y}_1 \end{bmatrix} = J_1 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + T_1 \begin{bmatrix} 0 \\ -\lambda \sum_{j=2}^{\infty} \binom{m}{j} (-\delta_1)^{-j} (\gamma_1 x_1 - \gamma_2 y_1)^j \end{bmatrix},$$

in which, by virtue of (27) and (29a), the expression $\delta_1^{-1}(\gamma_1 x_1 - \gamma_2 y_1)$ has been substituted for ϕ. Thus the canonical form of the nonlinear system is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{y}_1 \end{bmatrix} = J_1 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} -\lambda\gamma_2 \Sigma_2^{\infty} \binom{m}{j} (-\delta_1)^{-j} (\gamma_1 x_1 - \gamma_2 y_1)^j \\ -\lambda\gamma_1 \Sigma_2^{\infty} \binom{m}{j} (-\delta_1)^{-j} (\gamma_1 x_1 - \gamma_2 y_1)^j \end{bmatrix}. \tag{31a}$$

If we call the perturbing vector in this equation (f_1, f_2), we can apply Theorem 5.1 on p. 384 of Ref. 6 by establishing two properties of the f_i. Let $r^2 = x_1^2 + y_1^2$. Then f_1 and f_2 must be $o(r)$ as $r \to 0$; and also $\partial f_1/\partial x_1$ and $\partial f_2/\partial x_1$ must exist and be continuous in a neighborhood of the x_1y_1-origin. It is easy to see from the expressions in (31a) that the perturbing functions f_i satisfy these hypotheses. The theorem cited then says that the trajectories of (31a) have the same topological behavior near the x_1y_1-origin as those of (30a); namely, that all trajectories near the origin approach the origin as $t \to \infty$: one each becoming tangent to the positive and negative y_1-axes at the origin, and all the rest with limiting slope zero.

This improper node in the x_1y_1-plane can be transplanted back to the Fp-plane through application of T_1^{-1} and then of the substitution $F = 1 - \phi$. From (27) and (29a), and because (24) shows that $\gamma_1\gamma_2 = \lambda(m - 1)$, the negative y_1-axis has direction $(1, -\gamma_1)$ in the first quadrant of the ϕp-plane; likewise the positive x_1-axis has the ϕp-direction $(1, -\gamma_2)$. These directions then become lines of slope γ_1 and γ_2 respectively, passing through the point $(F, p) = (1, 0)$. Thus exactly

one trajectory of (21) reaches $(F, p) = (1, 0)$ with slope γ_1 from inside the triangle of Fig. 4; and infinitely many trajectories reach this singular point, from inside the triangle, with slope γ_2. (Note that $\gamma_1 < \gamma_2 < 0$.) Typical trajectories near $(F, p) = (1, 0)$ are shown in Fig. 4.

A similar analysis covers the singularity at the Fp-origin. The unperturbed form of (21) involves the matrix

$$A_0 = \begin{bmatrix} 0 & 1 \\ \lambda & -(1 - \lambda) \end{bmatrix} \tag{23b}$$

with determinant $-\lambda < 0$. The eigenvalues are -1 and λ–that is, real and of opposite sign for all relevant values of λ. The real nonsingular matrix

$$T_0 = \begin{bmatrix} \lambda & -1 \\ 1 & 1 \end{bmatrix} \tag{28b}$$

with inverse

$$T_0^{-1} = (1 + \lambda)^{-1} \begin{bmatrix} 1 & 1 \\ -1 & \lambda \end{bmatrix} \tag{29b}$$

brings in new coordinates (x_0, y_0) as in (27), and puts A_0 in the canonical form

$$J_0 = T_0 A_0 T_0^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & \lambda \end{bmatrix} \tag{30b}$$

listed as (IV) on p. 371 of Ref. 6. The corresponding singularity is a saddle point, as sketched in Fig. 9 on p. 374 of Ref. 6. The nonlinear system (21) transforms into

$$\begin{bmatrix} \dot{x}_0 \\ \dot{y}_0 \end{bmatrix} = J_0 \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \lambda \begin{bmatrix} [(x_0 + y_0)/(1 + \lambda)]^m \\ -[(x_0 + y_0)/(1 + \lambda)]^m \end{bmatrix}. \tag{31b}$$

Calling the perturbing functions f_1 and f_2 again, we invoke part (a) of Theorem 6.1 on p. 387 of Ref. 6. The required hypothesis is that both f_i be $o(r)$ (with $r^2 = x_0^2 + y_0^2$), which is true because $m > 1$. Then the trajectories of (31b) behave as follows: One pair of integral curves leaves the origin tangent to the y_0-axis; *at least* one pair approaches the origin along the x_0-axis; and the other nearby trajectories resemble modified hyperbolae. (Whether the pair reaching the origin tangent to the x_0-axis is unique or not, depends on m; but this question need

not concern us.) From (29b), the direction of the positive y_0-axis is $(1, \lambda)$ in the Fp-plane–that is, along the upper edge of the triangle of Fig. 4. The x_0-axis has Fp-direction $(1, -1)$. Thus the field near the Fp-origin has integral curves as shown in Fig. 4.

To support the remaining features of Fig. 4, we need only show that the trajectories passing into the triangle through some (leftmost) portion of its oblique upper edge do in fact reach the point (1, 0). [The contrary possibility would be that all such trajectories pass out through the right edge of the triangle, while the family of trajectories that have been shown to reach (1, 0) are among those that enter the triangle vertically through its bottom edge.] In fact all trajectories beginning to the left of $F = (1 - \lambda)/(1 + \lambda)$ have the desired property; for this point is the intersection of the upper edge $p = \lambda F$ with a straight line of slope $-(1 - \lambda)/2$ passing through (1, 0), and we now show that no trajectory can cross this line from left to right.

To prove this we must demonstrate that the vector field points more steeply downward, along this straight line, than the slope of the line itself; the field is so shown in Fig. 4. If we substitute the equation of the line in question, $p = (1 - \lambda)(1 - F)/2$, into the expression (20a) for p', we obtain $[2\lambda(F - F^m)/(1 - \lambda)(1 - F)] - (1 - \lambda)$ as the slope of the vector field along this straight line. This quantity must be more negative than the slope $-(1 - \lambda)/2$ of the straight line. The corresponding inequality can be written, after rearrangement,

$$(F - F^m)/(1 - F) < (1 - \lambda)^2/4\lambda, \quad (1 - \lambda)/(1 + \lambda) \leqq F < 1. \tag{32}$$

The left member of this inequality is an increasing function of F on $0 < F < 1$ for all real $m > 1$. (The proof is trivial for integral m, when the division can be carried out explicitly. For nonintegral m, it is perhaps easiest to show that the numerator of the derivative of this fraction is a strictly decreasing function which reaches zero at $F = 1$.) This function of F approaches $m - 1$ as $F \to 1$. Therefore (32) is satisfied if and only if

$$m - 1 \leqq (1 - \lambda)^2/4\lambda. \tag{33}$$

But this statement is true (with strong inequality) for Case 1, as the discussion on pp. 1381–2 shows.

We have shown that a trajectory can cross the line $p = (1 - \lambda)(1 - F)/2$ only *downward*. Since the slope of this line is $(\gamma_1 + \gamma_2)/2$, the qualitative aspects of Fig. 4 (which is in fact drawn to scale for $m = 2$, $\lambda = 0.15$) are verified. In particular, the unique trajectory that reaches $(F, p) = (1, 0)$ with the steeper slope γ_1 must

begin, with $t = 0$ on the line $p = \lambda F$, at an abscissa F_0^T that lies to the right of the intercept $(1 - \lambda)/(1 + \lambda)$.

Before discussing the bearing of Fig. 4 on the original problem, we investigate the form of the integral curves of (21) for the remaining ranges of λ.

Case 2: When $\lambda = \Lambda$ and $\gamma_1 = \gamma_2 = \gamma < 0$, the trajectories near the Fp-origin are as shown in Fig. 4, since the analysis on pp. 1384–5 goes through unchanged. For the other singularity we follow the calculation of the previous case, beginning with the introduction of new coordinates (x_2 , y_2) at the ϕp-origin by means of the matrix

$$T_2 = \begin{bmatrix} \gamma & 1 \\ 1 & 4/(\lambda - 1) \end{bmatrix}. \tag{28c}$$

It has determinant unity and the inverse

$$T_2^{-1} = \begin{bmatrix} 4/(\lambda - 1) & -1 \\ -1 & \gamma \end{bmatrix}, \tag{29c}$$

and so A_1 in (23a) assumes the canonical form

$$J_2 = \begin{bmatrix} \gamma & 0 \\ 1 & \gamma \end{bmatrix} \tag{30c}$$

listed as (III) on p. 371 of Ref. 6. The trajectories of the linear system are as shown in Fig. 7 on p. 373 of Ref. 6, and all reach the origin tangent to the y_2-axis in the first and third quadrants of the x_2y_2-plane. From (29c), we can substitute $\phi = (2x_2/\gamma) - y_2$ into (22) and obtain the canonical form

$$\begin{bmatrix} \dot{x}_2 \\ \dot{y}_2 \end{bmatrix} = J_2 \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + \begin{bmatrix} -\lambda \Sigma_2^{\infty} \binom{m}{j} (-\gamma)^{-j} (2x_2 - \gamma y_2)^j \\ -(2\lambda/\gamma) \Sigma_2^{\infty} \binom{m}{j} (-\gamma)^{-j} (2x_2 - \gamma y_2)^j \end{bmatrix} \tag{31c}$$

for the nonlinear system.

This system is treated in problem 10 (p. 346), Chapter 13 of Ref. 6. (No general theorem, analogous to the one used for Case 1, is given in Chapter 15. Rather than establishing a new framework suitable for importing such a theorem from another source, it seems simpler to use this weaker but adequate result.) Again calling the vector that perturbs the linear system (f_1 , f_2), we must show that both the f_i have continuous first partial derivatives in a neighborhood of the x_2y_2-origin; that the f_i and their first partials vanish at the origin; and that

each f_i is $O([\,|\,x_2\,|\,+\,|y_2\,|\,]^b)$ for some $b > 1$. These hypotheses are easily verified. Then the answer given to the problem just cited is that all trajectories in a neighborhood of the origin can be represented parametrically in the form

$$\begin{aligned} x_2(t) &= c_1 \exp(\gamma t) + O(\exp((\gamma - \epsilon)t)), \\ y_2(t) &= c_1 t \exp(\gamma t) + c_2 \exp(\gamma t) + O(\exp((\gamma - \epsilon)t)), \end{aligned} \tag{34}$$

for some $\epsilon > 0$. (The coefficients c_1 and c_2 are appropriate constants, and every choice of c_1 and c_2 corresponds to a trajectory.) Thus, in particular, all these trajectories approach the x_2y_2-origin, as $t \to \infty$, tangent to the y_2-axis.

As in Case 1, T_2^{-1} and the relation $F = 1 - \phi$ enable us to move this (degenerate) improper node back to the Fp-plane. The y_2-axis becomes a straight line of slope γ passing through $(F, p) = (1, 0)$. The trajectories differ from those shown in Fig. 4 only in that all those reaching (1, 0) from inside the triangle have limiting slope γ. But again there is an uppermost trajectory, which corresponds to the y_2-axis for the unperturbed system $(\dot{x}_2\ ,\ \dot{y}_2) = J_2(x_2\ ,\ y_2)$ and to the choice $c_1 = 0$ in (34). This distinguished trajectory crosses the edge $p = \lambda F$ (with $t = 0$) to the right of the abscissa $(1 - \lambda)/(1 + \lambda)$ by the argument used in Case 1, since (33) holds (with equality) in Case 2. Higher trajectories pass out of the triangle through its right-hand edge.

Case 3: When λ lies between Λ and Υ, the eigenvalues of A_1 are complex conjugates. The transformed matrix J_3 assumes the canonical form (V) of Ref. 6, p. 371; except that when $\lambda = 1$, so that from (25) the real part of γ_i vanishes, form (VI) occurs. In these situations, the x_3y_3-origin is a spiral point and a center, respectively, for the unperturbed system. According to Theorems 2.2 (p. 376) and 4.1 (p. 382) of Ref. 6, the nonlinear system also has either a center or a spiral point at $(x_3\ ,\ y_3) = (0, 0)$, and therefore behaves likewise at $(F, p) = (1, 0)$. In this case, the system (21) can have no trajectories which reach this point from the edge $p = \lambda F$ and stay within the triangle of Fig. 4.

Cases 4 and 5: Similar arguments cover the situation when $\lambda \geqq \Upsilon > 1$; but it is even easier to observe from (20a) that $p' > 0$ everywhere in the relevant triangle, so that each trajectory which starts with $t = 0$ on the edge $p = \lambda F$ must pass out through the right-hand edge of the triangle with an ordinate exceeding $p(F_0)$ at $F = 1$.

* * *

We saw in Section 3.1 that an admissible solution to the systems (12) and (18) corresponds to a parametrized trajectory of (20a) that

starts (with $t = 0$) on the line $p = \lambda F$, proceeds to the right within the triangle shown in Fig. 4, and ends (with $t = \infty$) at the point $(F, p) = (1, 0)$. The preceding discussion of Case 1 shows that, when $\lambda < \Lambda(m)$, there are two classes of admissible solutions. The first contains a single member corresponding to the unique trajectory, shown in Fig. 4, which reaches (1, 0) with slope γ_1 and starts with abscissa F_0^T. We call this solution F_λ^T. This function is represented in the phase plane by the graph of a function $p(F)$, but the latter corresponds through another integration to a distribution $F_\lambda^T(t)$: hence the symbol. The superscript is explained below; and the function F_λ^T also depends on the suppressed parameter m.

The second class of solutions is a one-parameter family represented in the phase plane by the trajectories which reach (1, 0) with slope γ_2 and begin on $p = \lambda F$ with abscissae lying in the open interval $(0, F_0^T)$. Each member of this family is called an F_λ . The superscript T was introduced as a mnemonic for "top": Not only does F_λ^T lie on top of all the F_λ in the phase plane, but also, because $dt = dF/p(F)$, it has a smaller value of the parameter t for given F than any F_λ and therefore lies above all the F_λ in the tF-plane.

The integral curves of (20a) in the triangle of Fig. 4, including those that enter through the F-axis and those above F_λ^T that correspond to distributions that are not unlimited, constitute a partition of the interior of the triangle: That is, each point of the interior lies on exactly one integral-curve. Nevertheless, those that end at (1, 0) do so only with the slopes given by (25). An F_λ starting very close to F_0^T, for instance, lies just below F_λ^T and turns, very near (1, 0), to make the angle $\arctan \gamma_2$ with the F-axis.

As shown in the discussion of Case 2, the situation is similar when $\lambda = \Lambda(m)$. The only difference is that F_λ^T and all the trajectories F_λ reach (1, 0) with the same slope $\gamma = \gamma_1 = \gamma_2$. On the other hand there are no admissible solutions when $\lambda > \Lambda$, for in Cases 3, 4, and 5 every trajectory corresponds to a distribution which is not unlimited. That is, when λ exceeds Λ, there is no equilibrium distribution of receiver delays; and this is why we call Λ the critical load.

Is it possible that F_λ^T and all the members of $\{F_\lambda\}$ are meaningful "answers" to the original problem when $\lambda \leqq \Lambda$? Certainly it is not clear how one would select the "correct" solution: On the one hand, the family $\{F_\lambda\}$ forms an open set corresponding to the open interval $(0, F_0^T)$ containing its end-points, and it is hard to see why any particular member of such a family should be "better" than its fellows. On the other hand, there is a substantial objection to the idea that

F_λ^T, the only distinguished solution of (20), should be the unique answer. For if F_λ^T is the meaningful solution, then F_0^T represents the probability of finding a receiver idle, and $(F_0^T)^m$ the probability of finding all m addressed receivers idle. Is it possible that this quantity should be positive even when $\lambda = \Lambda(m)$–that is, when even the slightest increase in load would preclude the possibility of an equilibrium distribution of receiver delays? On the contrary, innumerable other problems in queuing theory suggest that the value of F_0 corresponding to the true solution of (12) should approach zero as $\lambda \uparrow \Lambda(m)$. Yet it seems that even the idealized problem should dictate a unique solution.

IV. UNIQUENESS

We are now faced with a one-parameter family of stationary delay-distributions satisfying eqs. (12). But not every *stationary* distribution need be a *limiting* distribution; and so we return to (7) and other equations of Section 2.3 to study the sequence of delay distributions corresponding to successive messages arriving at a particular receiver queue. (Since the argument of this section is more suited to formal exposition than the preceding material, the results are given as theorems rather than discursively.)

Some additional notation is required, and we begin with a restatement of the key equation

$$F_{n+1}(t) = \frac{\lambda}{1+\lambda}\left[\int_0^t e^{-(t-u)}F_n^m(u)\,du + \int_t^\infty e^{-\lambda(u-t)}F_n^m(u)\,du\right]. \tag{7}$$

It is convenient to define the power operator $\mathcal{P}_m$ such that

$$F^m = \mathcal{P}_m F \tag{35a}$$

for real m and the integral operator

$$\mathfrak{M}_\lambda \equiv \int_0^\infty du\, k_\lambda(t, u) \tag{35b}$$

where the kernel k_λ is given by (14), so that, with

$$\mathfrak{N}_\lambda \equiv \mathfrak{M}_\lambda \mathcal{P}_m\,, \tag{35c}$$

we can write (7) in the form

$$F_{n+1} = \mathfrak{N}_\lambda F_n\,. \tag{36}$$

(The dependence of $\mathfrak{N}_\lambda$ on m is suppressed.) All functions of t mentioned in this section have domain $[0, \infty)$.

Lemma 1: If F_n is a (nontrivial) distribution function (d.f.), then F_{n+1} is strictly increasing.

Proof: By hypothesis F_n is nondecreasing, so that $F_n^m(t) \geqq F_n^m(u)$ for $0 \leqq u \leqq t$. Thus eq. (3) shows that for all t ($\geqq 0$, of course)

$$G_n(t) \leqq F_n^m(t) \int_0^t e^{-(t-u)}\, du = F_n^m(t)(1 - e^{-t})$$
$$< F_n^m(t).$$

[We write this $G_n < F_n^m$; such inequalities between functions, written without restriction, are to be interpreted pointwise on $[0, \infty)$.] Therefore by (6) $\dot{G}_n > 0$, and so also, using (4a),

$$\int_t^\infty e^{-\lambda(u-t)}\, dG_n(u) = F_{n+1}(t) - G_n(t) > 0.$$

[Notice that this inequality holds even when, because $F_n(v) = 0$, the preceding strong inequalities fail for $u \leqq v$.] From (10) we see that $\dot{F}_{n+1} > 0$. □

The next lemma shows that F_{n+1} is unlimited–that is, assigns positive probability to arbitrarily long delays.

Lemma 2: If $F_n \leqq 1$, then $F_{n+1} < 1$.

Proof: The statement of the lemma is contained in (11c). □

Lemma 3: If $F_n \to 1$ as $t \to \infty$, then $F_{n+1}(t) \to 1$ as $t \to \infty$. (In words, F_{n+1} is proper if F_n is proper.)

Proof: We showed in proving Lemma 1 that $F_{n+1} > G_n$. Thus $G_n < 1$ by Lemma 2. If G_n were bounded away from 1, then by (6), $\dot{G}_n$ would be positive and bounded away from 0 for sufficiently large t. This is impossible; therefore $G_n \to 1$ as $t \to \infty$. Since $F_{n+1} > G_n$, $F_{n+1} \to 1$ also. □

[The fact that solutions of (12) must be proper and unlimited was discussed on pp. 1377–8.]

Theorem 1: If F_1 is a proper distribution function, then the remainder of the sequence generated by (7) and (36) consists of d.f.s which are proper, unlimited, and strictly increasing.

Proof: Apply Lemmata 1, 2, and 3 inductively on n. □

The hypothesis of Theorem 1 covers the case in which $F_1 = 1$; that is, $F_1(t) = 1\ \forall t \geqq 0$. In this case, with probability one the first message

suffers no delay at the receiver. In physical terms, the first message finds the system idle. We now study the system's simplest kind of transient behavior, beginning with no messages present and letting the receivers run indefinitely at a fixed subcritical load. In the discussion that follows, we represent by $F_n[F_1, \lambda]$ the nth element of the sequence generated by (36); this notation specifies both the starting function F_1 and the parameter λ of the operator $\mathfrak{N}_\lambda$. The value of this function at t is $F_n[F_1, \lambda](t)$.

We need this preliminary result:

Lemma 4: Let F_1 be a d.f. If $F_n[F_1, \lambda] \to L$ as $n \to \infty$, then L is a solution of (12).

Proof: Because L is a pointwise limit of functions which satisfy (11c), L itself must satisfy (12b). Using (35) and (36), we can write (7) in the form*

$$F_{n+1}(t) = \int_0^\infty k_\lambda(t, u)F_n^m(u)\, du, \tag{37}$$

which corresponds to the equilibrium equation (15). For each t, the sequence $\{k_\lambda(t, u)F_n^m(u)\}$ approaches $k_\lambda(t, u)L^m(u)$ as $n \to \infty$ and is bounded by $k_\lambda(t, u)$, which is an integrable function of u according to (13). These are the hypotheses of Lebesgue's Dominated Convergence Theorem, which tells us that

$$\lim_{n\to\infty} \int_0^\infty k_\lambda(t, u)F_n^m(u)\, du = \int_0^\infty k_\lambda(t, u)L^m(u)\, du.$$

By (37) and the hypothesis of the present lemma, the left member of the last equation is $L(t)$. Thus L indeed satisfies (12a). □

We are now ready to prove

Theorem 2: Let $F_1 = 1$ and choose a fixed $\lambda \leqq \Lambda(m)$. Then the sequence $\{F_n[F_1, \lambda]\}$ approaches F_λ^T.

Proof: Setting $F_1 = 1$ in (7) shows as in (13) that $F_2(t) = 1 - [\lambda/(1 + \lambda)] \exp(-t)$, so that $F_2 < F_1$. From (35) and (36), $F_{n+1} = \mathfrak{N}_\lambda F_n^m$. Writing this also for F_n, we get by subtraction

* Since the difference kernel k_λ is the density function of the difference $x - a$ between a transmission time and an interarrival time, we recognize (37) as a consequence of the equation $\omega_{n+1} = \max[0, \max(\omega_n^{(1)}, \cdots, \omega_n^{(m)}) + x_n - a_n]$. This relation generalizes the familiar recurrence for a single-server queue, in which $\max \omega_n^{(i)}$ reduces to ω_n.

$$F_n - F_{n+1} = \mathfrak{M}_\lambda(F_{n-1}^m - F_n^m). \tag{38}$$

Since $\mathfrak{M}_\lambda$ is monotone–that is, has a strictly positive kernel–the left member of (38) is strictly positive if the operand is so. Therefore, by induction on n, $F_{n+1} < F_n \ \forall n$.

The sequence $\{F_n(t)\}$ is strictly decreasing and bounded below by zero. Thus it has a limit $L(t)$, and this defines a function L on $[0, \infty)$ such that $F_n \to L$ as $n \to \infty$. By Lemma 4, L is stationary under $\mathfrak{N}_\lambda$. Thus L must be either 0, one of the F_λ, or F_λ^T.

Let H be any solution of (12), so that $H = \mathfrak{M}_\lambda H^m$. If $F_n > H$ then $F_n^m > H^m$, so that $F_{n+1} > H$ because, as in (38), $F_{n+1} - H = \mathfrak{M}_\lambda(F_n^m - H^m)$. Every solution of (12), being unlimited, is less than 1, so that $H < F_1$. By induction on n, $H < F_n \ \forall n$. That is, H is a pointwise lower bound for $\{F_n\}$. As we found on p. 1388, the largest solution of (12) is F_λ^T, so that $F_\lambda^T(t)$ is the greatest of the lower bounds represented by $H(t)$. Therefore $L = F_\lambda^T$. □

We can write Theorem 2 as the statement that $F_n[1, \lambda] \downarrow F_\lambda^T$ when $\lambda \leqq \Lambda$. (Actually the convergence is uniform in this case, but there is no need to prove this.) This result is in essence the answer to our problem, although we must now generalize it considerably in order to remove the strong restriction that $F_1 = 1$.

Before going on to strengthen Theorem 2, we observe that receiver delays increase indefinitely if the system continues to operate with more-than-critical load. This fact is stated accurately in

Theorem 3: Choose a fixed $\lambda > \Lambda(m)$. Then the sequence $\{F_n[1, \lambda]\}$ approaches zero.

Proof: The proof of Theorem 2 applies here with one exception. Again $\{F_n\}$ has a limit L which is stationary under $\mathfrak{N}_\lambda$, but with $\lambda > \Lambda$ the only solution of (12) is identically zero. □

In this case the convergence is not uniform; the probability masses associated with successive members of $\{F_n\}$ are located farther and farther to the right. We may conclude our discussion of the supercritical case with

Corollary 1: Let F_1 be any d.f. and choose a fixed $\lambda > \Lambda(m)$. Then the sequence $\{F_n[F_1, \lambda]\}$ approaches zero.

Proof: Since F_1 is a d.f., $F_1 \leqq 1$. Thus $F_1^m \leqq 1^m$, and so $F_2[F_1, \lambda] - F_2[1, \lambda] = \mathfrak{M}_\lambda(F_1^m - 1^m) \leqq 0$. Likewise $F_n[F_1, \lambda] \leqq F_n[1, \lambda] \ \forall n$ by induction. But $F_n[F_1, \lambda] \geqq 0$, and $F_n[1, \lambda] \to 0$ as $n \to \infty$ by Theorem 3. Therefore $F_n[F_1, \lambda] \to 0$ as $n \to \infty$. □

Thus, no matter what the delay distribution encountered by the first message, delays increase without bound when λ exceeds Λ.

Our first step in generalizing Theorem 2 is to consider the case in which the system, instead of starting empty, is first allowed to reach statistical equilibrium with a receiver arrival-rate $\mu \leqq \Lambda$. In effect, the system starts operation at time $-\infty$ and comes to equilibrium before the arrival (in receiver queues) of the message we label number one. At that instant we change the receiver arrival-rate to a new value $\lambda \leqq \Lambda$ and thenceforth keep it fixed. This mathematical model is a very idealized one, since it is not clear how the receiver interarrival-times can suddenly acquire the density $\lambda \exp(-\lambda t)$ at a specified instant. With k and m fixed, the arrival rate can change from μ to λ only through a change in the rate α of transmitter arrivals; and the effect of a change in α will be felt gradually as the number of occupied transmitters increases to its new equilibrium-value. However, since eqs. (7) and (36) form our only tool for studying the evolution of receiver delays, the best we can do in the present case is to take $F_1 = F_\mu^T$ (by Theorem 2) and to generate $\{F_n[F_\mu^T, \lambda]\}$ by applying $\mathfrak{N}_\lambda$. We begin with

Lemma 5: If $0 < \lambda < \mu < \Lambda(m)$, *then* $0 > \gamma_2(\lambda) > \gamma_2(\mu) > \gamma_1(\mu) > \gamma_1(\lambda) > -1$. *If* $\mu = \Lambda$, *then* $\gamma_2(\mu) = \gamma_1(\mu) \equiv \gamma(\mu) = (\Lambda - 1)/2$. *If* $\lambda = 0$, *then* $\gamma_2(\lambda) = 0$ *and* $\gamma_1(\lambda) = -1$.

Proof: The result for $\mu = \Lambda$ was stated on p. 1382. The result for $\lambda = 0$ is obvious from (25). From (25),

$$2\frac{d\gamma_1}{d\lambda} = 1 + \frac{2m - (1 + \lambda)}{[(1 + \lambda)^2 - 4\lambda m]^{\frac{1}{2}}}$$

which is clearly positive because $m > 1$ and $\lambda \leqq \Lambda < 1$. Thus $\mu > \lambda$ implies that $\gamma_1(\mu) > \gamma_1(\lambda) > -1$. By definition $\gamma_2(\mu) > \gamma_1(\mu)$ when $\mu < \Lambda$. To show that $0 > \gamma_2(\lambda) > \gamma_2(\mu)$, we need only prove that

$$2\frac{d\gamma_2}{d\lambda} = 1 - \frac{2m - (1 + \lambda)}{[(1 + \lambda)^2 - 4\lambda m]^{\frac{1}{2}}} < 0,$$

and this follows after trivial manipulation from the fact that $m > 1$. □

The next lemma expresses another kind of "monotonicity" property of $\mathfrak{N}_\lambda$, this time with respect to the parameter λ.

Lemma 6: Let H_μ *be any nonzero d.f. stationary under* $\mathfrak{N}_\mu$. *If* $\lambda < \mu$, *then* $\mathfrak{N}_\lambda H_\mu > H_\mu$; *and if* $\lambda > \mu$, *then* $\mathfrak{N}_\lambda H_\mu < H_\mu$.

Proof: We express the action of $\mathfrak{N}_\lambda$ as in eqs. (35), with k_λ given by (14). First assume that $\lambda < \mu$, and define τ as the value of u at which

k_λ and k_μ intersect; see Fig. 5. That is, $k_\lambda(t, \tau) = k_\mu(t, \tau)$, so that τ is a function of t. [Explicitly, $\tau = t + (\mu - \lambda)^{-1} \cdot \ln([1 + \lambda^{-1}]/[1 + \mu^{-1}])$.] It is easy to see that, as indicated in Fig. 5, $k_\lambda(t, u) < k_\mu(t, u)$ when $u < \tau$ and $k_\lambda(t, u) > k_\mu(t, u)$ when $u > \tau$.

For any fixed $t \geqq 0$,

$$\begin{aligned}
\mathfrak{N}_\lambda H_\mu(t) - H_\mu(t) &= \mathfrak{N}_\lambda H_\mu(t) - \mathfrak{N}_\mu H_\mu(t) \\
&= [(\mathfrak{N}_\lambda - \mathfrak{N}_\mu) H_\mu^m](t) \\
&= \int_0^\infty [k_\lambda(t, u) - k_\mu(t, u)] H_\mu^m(u)\, du \\
&= \int_\tau^\infty [k_\lambda(t, u) - k_\mu(t, u)] H_\mu^m(u)\, du \\
&\quad - \int_0^\tau [k_\mu(t, u) - k_\lambda(t, u)] H_\mu^m(u)\, du.
\end{aligned}$$

We know that in each of the integrals the bracketed quantity is positive. Also H_μ is nondecreasing; thus $H_\mu(\tau) \leqq H_\mu(u)$ for $u > \tau$ and $H_\mu(\tau) \geqq H_\mu(u)$ for $u < \tau$. Therefore

$$\begin{aligned}
\mathfrak{N}_\lambda H_\mu(t) - H_\mu(t) &\geqq H_\mu^m(\tau) \left\{ \int_\tau^\infty [k_\lambda(t, u) - k_\mu(t, u)]\, du \right. \\
&\quad \left. - \int_0^\tau [k_\mu(t, u) - k_\lambda(t, u)]\, du \right\} \\
&= H_\mu^m(\tau) \int_0^\infty [k_\lambda(t, u) - k_\mu(t, u)]\, du \\
&= H_\mu^m(\tau) \{\mathfrak{N}_\lambda(1) - \mathfrak{N}_\mu(1)\} \\
&= H_\mu^m(\tau) e^{-t} \left\{ \frac{\mu}{1 + \mu} - \frac{\lambda}{1 + \lambda} \right\},
\end{aligned}$$

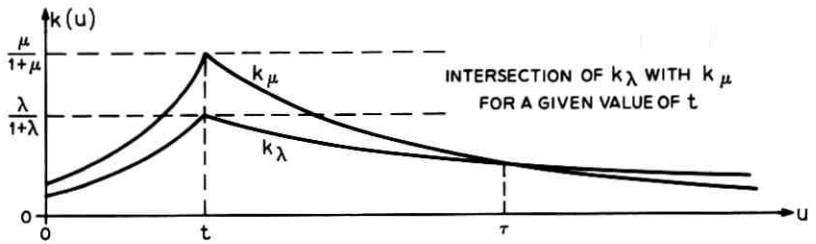


Fig. 5—The kernels k_λ and k_μ cross at τ.

where the last equality comes from (13) as noted in the proof of Theorem 2. The explicit formula for τ shows that $\tau > 0 \ \forall t$, so that $H_\mu(\tau) > 0$ because H_μ is nonzero. [By the equilibrium form of (8) on p. 1375, $H_\mu(0) > 0$.] Since $\lambda < \mu$, each factor of the last expression displayed above is positive, and thus $\mathfrak{N}_\lambda H_\mu(t) > H_\mu(t) \ \forall t$. This proves the first conclusion of the lemma.

When $\lambda > \mu$, the proof goes through as before except that the appropriate inequalities are reversed. The modified proof shows that $\mathfrak{N}_\lambda H_\mu(t) - H_\mu(t)$ is less than or equal to the final expression in the calculation above, and that expression is negative in this case. □

Lemma 7: If $\lambda < \mu \leqq \Lambda(m)$, *then* $F_\lambda^T > F_\mu^T$.

Proof: Because F_μ^T is unlimited, $1 > F_\mu^T$. Applying $\mathfrak{N}_\lambda$, we find that $F_2[1, \lambda] \equiv \mathfrak{N}_\lambda 1 > \mathfrak{N}_\lambda F_\mu^T > F_\mu^T$, where the first inequality comes from the argument used with (38) and the second from Lemma 6. Repeated application of $\mathfrak{N}_\lambda$ shows that $F_n[1, \lambda] > \mathfrak{N}_\lambda F_\mu^T \ \forall n$, so that, using Theorem 2,

$$F_\lambda^T \equiv \lim_{n \to \infty} F_n[1, \lambda] \geqq \mathfrak{N}_\lambda F_\mu^T > F_\mu^T \ . \ \square$$

Lemma 8: Let F be any solution of (12), and let the corresponding phase-plane trajectory $p(F)$ *reach* $(F, p) = (1, 0)$ *with slope* $\gamma < 0$. *Then for any* $\epsilon > 0$, *there exist constants* c_3 , c_4 , *and* T *such that, whenever* $t > T$,

$$c_3 e^{(\gamma - \epsilon)t} < 1 - F(t) < c_4 e^{(\gamma + \epsilon)t}. \tag{39}$$

Proof: From (19), $dt = dF/p(F)$, so that we can write

$$t(F) = \int_{F_0}^{F} \frac{du}{p(u)}. \tag{40}$$

Given $\epsilon > 0$, there is an abscissa F_ϵ such that the integral curve $p(F)$ lies in the wedge $-(\gamma + \epsilon)(1 - F) < p < -(\gamma - \epsilon)(1 - F)$ when $F_\epsilon < F \leqq 1$. Choose T so large that $F(T) > F_\epsilon$. For $t > T$, (40) becomes

$$t(F) = \int_{F_0}^{F(T)} \frac{du}{p(u)} + \int_{F(T)}^{F} \frac{du}{p(u)} \ ,$$

and the first of these integrals equals T by definition. When $F > F(T)$, $p(F)$ satisfies the inequalities defining the wedge just mentioned, and so

$$T - \int_{F(T)}^{F} \frac{du}{(\gamma - \epsilon)(1 - u)} < t(F) < T - \int_{F(T)}^{F} \frac{du}{(\gamma + \epsilon)(1 - u)}.$$

Thus

$$T + \frac{1}{\gamma - \epsilon} \ln \frac{1 - F(t)}{1 - F(T)} < t < T + \frac{1}{\gamma + \epsilon} \ln \frac{1 - F(t)}{1 - F(T)},$$

and the conclusion (39) follows after rearrangement. The constants are

$$c_3 = [1 - F(T)]e^{-(\gamma - \epsilon)T}, \quad c_4 = [1 - F(T)]e^{-(\gamma + \epsilon)T}. \quad \square$$

[The facts that $1 - F(t)$ behaves nearly exponentially for large t, that the decay constant γ must take one of the forms (25), and that no admissible solution exists when $\lambda > \Lambda(m)$, can all be inferred non-rigorously by considering linearized versions of (12a), (18a), or (20a) for very small values of $1 - F$.]

Lemma 9: When $\lambda < \mu$, no member of the open family $\{F_\lambda\}$ lies above F_μ^T.

Proof: For any given F_λ and $\epsilon > 0$ and all sufficiently large t, we know from Lemma 8 that there exist constants c_λ and c_μ such that

$$F_\lambda(t) < 1 - c_\lambda e^{[\gamma_2(\lambda) - \epsilon]t} \quad \text{and} \quad F_\mu^T(t) > 1 - c_\mu e^{[\gamma_1(\mu) + \epsilon]t}.$$

By Lemma 5, $\gamma_2(\lambda) > \gamma_1(\mu)$. So long as the ϵ that appears above is small enough to make $\gamma_2(\lambda) - \epsilon > \gamma_1(\mu) + \epsilon$, we can find τ such that $F_\mu^T(t) > F_\lambda(t)$ when $t > \tau$. $\square$

Theorem 4: Let $\lambda \leqq \Lambda(m)$ and $\mu \leqq \Lambda(m)$. Then $F_n[F_\mu^T, \lambda] \to F_\lambda^T$ as $n \to \infty$.

Proof: If $\lambda = \mu$, the result is obvious because F_μ^T is stationary under $\mathfrak{N}_\mu$.

If $\lambda > \mu$, then $F_2[F_\mu^T, \lambda] \equiv \mathfrak{N}_\lambda F_\mu^T < F_\mu^T$ by Lemma 6. Therefore, by the argument based on monotonicity and induction that was used in proving Theorem 2, the sequence $\{F_n[F_\mu^T, \lambda]\}$ is strictly decreasing. Thus it has a limit $L \geqq 0$ which by Lemma 4 is stationary under $\mathfrak{N}_\lambda$. Exchanging λ with μ in Lemma 7, we see that $F_\lambda^T < F_\mu^T$. Continuing the argument of Theorem 2, we see that F_λ^T is a lower bound for the sequence $\{F_n[F_\mu^T, \lambda]\}$, and greater than any of the other lower bounds F_λ. Thus $F_n[F_\mu^T, \lambda] \downarrow F_\lambda^T$.

If $\lambda < \mu$, a similar argument shows that F_λ^T is an upper bound for the strictly increasing sequence $\{F_n[F_\mu^T, \lambda]\}$, which again has a limit stationary under $\mathfrak{N}_\lambda$. By Lemma 9, no other function F_λ that is stationary under $\mathfrak{N}_\lambda$ can be an upper bound for this sequence. Therefore $F_n[F_\mu^T, \lambda] \uparrow F_\lambda^T$. $\square$

It is natural to consider the communication system as initially idle and thenceforth, after the arrival of the first message, subjected to a load of constant intensity. Theorem 4 generalizes this situation by

considering changes from a previously attained equilibrium state. But the most realistic and therefore useful statement about limiting distributions would involve initialization of some nearly arbitrary kind, so that any queue-length distribution might be encountered by message number one. We now study a model of this kind in two lemmata and a theorem. In this model the system begins empty, receives messages at a sequence of different arrival-rates which may be supercritical, and is then subjected to a single subcritical load indefinitely. A real system operating with time-varying load would experience a sequence of receiver interarrival-times with different and presumably non-exponential distributions. We consider an idealized counterpart in which these distributions are indeed exponential.

Lemma 10: Let the sequence $\{F_n[1, \{\lambda_i\}]\}$ *be generated by applying* $\mathfrak{N}_{\lambda_i}$ k_i *times,* $i = 1, \cdots, j$, *so that* $n = 1 + \sum_{i=1}^{j} k_i$, *with each* $\lambda_i > 0$. *Then* $F_n[1, \{\lambda_i\}](t) \cong 1 - c_n t^{n-2} \exp(-t)$.

Proof: We suppose that the exact expression for F_n is a series beginning with the two terms shown and continuing with exp $(-t)$ times lower powers of t, exp $(-2t)$ times higher-degree polynomials in t, and so on. We know that $F_2[1, \lambda_1] = 1 - [\lambda_1/(1 + \lambda_1)] \exp(-t)$, which has the assumed form. Also

$$F_n[1, \{\lambda_i\}] = \mathfrak{N}_{\lambda_j}^{k_j} \mathfrak{N}_{\lambda_{j-1}}^{k_{j-1}} \cdots \mathfrak{N}_{\lambda_1}^{k_1}(1) = \mathfrak{N}_{\lambda_j} F_{n-1}[1, \{\lambda_i\}].$$

Assume $F_{n-1}[1, \{\lambda_i\}]$ has the stated form, substitute into (7), expand F_{n-1}^m using the binomial theorem, and integrate explicitly. This very tedious procedure, which, being straightforward, is not recorded here, yields the conclusion of the lemma by induction. The constant c_n is a function of m and of the λ_i. □

Corollary 2: The conclusion of Lemma 10 holds if, in generating $\{F_n\}$, *the receiver arrival-rate is set equal to zero during a finite number of finite intervals.*

Proof: If the load is zero for an interval of length t_0 between the arrivals of messages $n - 1$ and n, then F_n is simply shifted t_0 units to the left; that is, the probability that would have been $F_n[1, \{\lambda_i\}](t)$ becomes $F_n[1, \{\lambda_i\}](t - t_0)$. This expression, with F_n calculated as in Lemma 10, defines a new function of t having the same form as $F_n[1, \{\lambda_i\}]$ but with different constants. This process does not affect the proof of Lemma 10 and can be repeated finitely many times. □

Note that the loads $\{\lambda_i\}$ in Lemma 10 need not be less than or equal to $\Lambda(m)$. In describing transient loads, we measure the length of time

during which λ_i is the receiver arrival-rate by specifying k_i, the number of messages which follow their predecessors by intervals with the density $\lambda_i \exp(-\lambda_i t)$. Since this procedure cannot account for a period with zero arrival-rate, we must assume $\lambda_i > 0$ in Lemma 10 and then show separately in Corollary 2 that the form of the sequence $\{F_n[1, \{\lambda_i\}]\}$ is not affected by the existence of periods during which no messages can arrive. The effect of shutting off transmitter arrivals in an actual system would be approximated by a sequence of λ_i gradually decreasing to zero as the transmitter queues empty out.

Lemma 11: Let F be any element $F_n[1, \{\lambda_i\}]$ of the sequence defined in Lemma 10. Then there exists a member F_Λ, of the open family of d.f.s stationary under $\mathfrak{N}_\Lambda$, such that $F_\Lambda < F$.

Proof: We know from Lemma 5 that $\gamma(\Lambda) > -1$; choose $\epsilon < [1 + \gamma(\Lambda)]/2$, so that $\gamma(\Lambda) - \epsilon > -1 + \epsilon$. According to (39) we can find τ_1 and c such that

$$F_\Lambda^T(t) < 1 - ce^{[\gamma(\Lambda)-\epsilon]t} \quad \text{for} \quad t > \tau_1 .$$

Also, by Lemma 10, we can choose τ_2 and c_n such that

$$F(t) > 1 - c_n t^{n-2} e^{-(1-\epsilon)t} \quad \text{for} \quad t > \tau_2 ,$$

since for large enough t, the error term in Lemma 10 can be made small enough to be bounded by the effect of the factor $\exp(\epsilon t)$. (Here n is the index of F in $\{F_n\}$.) There exists τ_3 such that $c_n t^{n-2}/c < \exp([1 + \gamma(\Lambda) - 2\epsilon]t)$ when $t > \tau_3$. Let $\tau = \max(\tau_1, \tau_2, \tau_3)$. Then $F_\Lambda^T(t) < F(t)$ for $t > \tau$. Since $F_\Lambda < F_\Lambda^T$ for every member of the family $\{F_\Lambda\}$, we have $F_\Lambda(t) < F(t)$ when $t > \tau$ for each F_Λ–that is, uniformly in the parameter $F_\Lambda(0)$ that we use to specify an element of $\{F_\Lambda\}$.

By Theorem 1, every element of the sequence $\{F_n\}$ is proper, and so by (8) $F(0) > 0$. Choose $F_\Lambda(0) < F(0) \exp(-\Lambda\tau)$. From the equilibrium form of (9) we have $F_\Lambda(\tau) < F_\Lambda(0) \exp(\Lambda\tau) < F(0)$. Since F_Λ and F are nondecreasing, $F_\Lambda(t) < F(t)$ when $t \leqq \tau$. We have now proved this inequality for all t, so that $F_\Lambda < F$. □

Theorem 5: Let F be any element $F_n[1, \{\lambda_i\}]$ of the sequence defined in Lemma 10. Choose $\lambda < \Lambda(m)$. Then $F_n[F, \lambda] \to F_\lambda^T$ as $n \to \infty$.

Proof: Given F, choose F_Λ by Lemma 11 so that $F_\Lambda < F$. By Lemma 5, $\gamma_2(\lambda) > \gamma(\Lambda)$. Thus Lemma 9 remains valid with F_μ^T replaced by any member of $\{F_\Lambda\}$. Therefore Theorem 4 is valid with F_μ^T replaced by F_Λ, so that $F_n[F_\Lambda, \lambda] \uparrow F_\lambda^T$. By Theorem 2, $F_n[1, \lambda] \downarrow F_\lambda^T$. By Theorem 1, F is unlimited, and we can write $F_\Lambda < F < 1$. By the monotonicity

argument of Theorem 2, $F_n[F_\Lambda\ ,\ \lambda] < F_n[F,\ \lambda] < F_n[1,\ \lambda]$ for all n. Therefore $\{F_n[F,\ \lambda]\}$ is bounded term by term on both sides by sequences which approach F_λ^T, and so $F_n[F,\ \lambda] \to F_\lambda^T$ as $n \to \infty$. □

This theorem does not cover the case of a system operated *at* its critical load after being subjected to an arbitrary transient of finite duration. Part of the gap is filled by

Corollary 3: Let F be chosen as in Theorem 5 but with $F \geqq F_\Lambda^T$, for example because $\max\{\lambda_i\} \leqq \Lambda$. Then $F_n[F,\ \Lambda] \to F_\Lambda^T$ as $n \to \infty$.

Proof: $F_\Lambda^T \leqq F < 1$; $F_n[1,\ \Lambda] \downarrow F_\Lambda^T$ by Theorem 2; and $F_n[F_\Lambda^T\ ,\ \Lambda] = F_\Lambda^T\ \forall n$. Thus $\{F_n[F,\ \Lambda]\}$ is squeezed between two sequences with the same limit. □

This kind of argument does not work when F is not greater than or equal to F_Λ^T. We can bound F below by a member of the family $\{F_\Lambda\}$, but F_Λ is stationary under $\mathfrak{N}_\Lambda$; thus, if $\{F_n[F,\ \Lambda]\}$ or a subsequence of it has a limit L, we know only that $\forall t\ L(t)$ lies in the interval $[F_\Lambda(t), F_\Lambda^T(t)]$. We could choose $\lambda < \Lambda$ and bound F below by a particular F_λ, but this would not help because $\mathfrak{N}_\Lambda F_\lambda < F_\lambda$ by Lemma 6: $\{F_n[F,\ \Lambda]\}$ would be bounded below by a decreasing sequence. I have not been able to determine the behavior of the sequences in question more precisely than is stated in

Corollary 4: Let F be chosen as in Theorem 5 but with F not bounded below by F_Λ^T. Then there exists an $F_\Lambda \in \{F_\Lambda\}$: $F_n[F,\ \Lambda] > F_\Lambda\ \forall n$; and if $\{F_n[F,\ \Lambda]\}$ has a limit L, then L is either F_Λ^T or a member of $\{F_\Lambda\}$.

Proof: By Lemma 11 we can find $F_\Lambda < F$, and by the monotonicity of $\mathfrak{N}_\Lambda$, this inequality is preserved throughout the sequence $\{F_n[F,\ \Lambda]\}$. By Lemma 4, L is stationary under $\mathfrak{N}_\Lambda$; and our first conclusion shows that L is bounded away from zero. □

The limitations of this result are not surprising. If the system is operated *at* its critical load after being temporarily overloaded, we see that delays still do not increase indefinitely; but on the other hand we have found no assurance that there is a limiting distribution or that, if there is, it agrees with the one (F_Λ^T) that would have resulted if the period of critical load had begun with the system idle.

We defer discussion of all these results to later sections, merely observing that the argument given on pp. 1388–9 makes Theorems 2, 4, and 5 quite remarkable. Another interesting point is raised by the strong dependence of these proofs on the *asymptotic* behavior of the

functions involved–in particular, on the rate at which $F_1(t) \to 1$ for large t, where F_1 is the initial distribution of a sequence $\{F_n[F_1, \lambda]\}$. It may seem that our results are too sensitive to such decay rates–that this feature somehow represents an "instability" of the mathematical model. But in fact the essentials are the general properties recorded in Theorem 1, and the decay behavior of F_n which Lemma 10 shows is dominated by $t^{n-2} \exp(-t)$. All these properties are in turn controlled by the density $\exp(-t)$ of the transmission times, and the expression $t^{n-2} \exp(-t)$ in particular is related to the fact that the sum of the $n - 1$ preceding transmission times *or their tails* has an Erlang distribution of order $n - 1$. The most interesting implication of the dependence of our arguments on asymptotic behavior is that our results are accessible only to analytic techniques: A computer could not be successfully used to study experimentally the properties of sequences generated by eq. (7).

The situation for $m = 1$ puts the argument of this section in perspective. All solutions of (12) and (18) have the form $F(t) = C[1 - \lambda \exp(-(1 - \lambda)t)]$ in this case, with $0 < C \leqq 1$. Only the one of these with $C = 1$ is proper, and it is Erlang's well-known delay-distribution for the single-server queue.

V. DISTRIBUTION OF RECEIVER DELAYS

We have found that the meaningful solutions of the stationarity equations, (12) and (18), are the limiting distributions F_λ^T. This section is devoted to four questions: How can these distributions be calculated explicitly; how do we proceed when the number of addresses per message is a random variable of which m is the mean; what are the properties of the delay distributions; and how closely do these results represent the behavior of systems with finite R?

5.1 *Computation of Delay Distributions and Their Moments*

In answering the first of these questions we take λ and m to be fixed and interpret the symbol F to mean F_λ^T; likewise the symbol γ_1 refers to the quantity defined in (25). The first step is to carry out the first integration, based on

$$p' = [\{\lambda(F - F^m)\}/p] - (1 - \lambda), \tag{20a}$$

which yields the phase-plane trajectory $p(F)$. (See Fig. 4.) This can be done numerically using any standard integration-formula of the

predictor-corrector type, and is best done as follows: Beginning at the corner (1, 0), integrate the vector field to the left, starting at an angle with the F-axis whose tangent is γ_1 . Follow the resulting path until it intersects the edge $p = \lambda F$ at the abscissa $\hat{F}_0$, which is an estimate of the true initial value $F(0)$ of the distribution sought. Since the vector field diverges in the leftward direction, this integration is numerically quite unstable, so that $\hat{F}_0$ need not be an accurate estimate of $F(0)$. Explore the neighborhood of $\hat{F}_0$ for the true value of $F(0)$ by selecting a set of abscissae (spaced about 10^{-5} units apart) and calculating, for each one, the integral curve that passes through it. These calculations are performed by integrating to the right; and since the vector field converges to the right, the resulting curves are quite accurate. For each such curve, find the ordinate at which it intersects the line $F = 1$. These ordinates can be plotted against the abscissae at which the integral curves start on the edge $p = \lambda F$. Beginning at some point well to the right of $\hat{F}_0$, the graph so constructed has an ordinate of considerable positive magnitude. As the abscissa decreases, approaching the true $F(0)$ from above, the ordinate must decrease to zero, remaining there for all lower values of the abscissa; for trajectories beginning to the left of F_0^T must all end at (1, 0). This ideal pattern is perturbed by noise arising from roundoff and from the numerical integration itself, but it is not difficult to find $F(0)$ from this graph of p-intercepts to an accuracy ranging from about 10^{-4} to about 10^{-6} units, so long as λ is not very close to Λ. The function $p(F)$ is then found by integrating the vector field to the right from $F(0)$. [Trajectories lying below F_λ^T cannot be found precisely; they have very large curvatures near (1, 0) because of the singularity there.]

This technique must be modified when $\lambda = \Lambda(m)$ (and in fact when $\lambda/\Lambda \gtrsim 0.9$), for in this case the graph just described is erratic and appears to have quite a high-order contact with its abscissa, the F_0-axis. Less than two decimal-places of precision can be obtained in this way. Barbara R. LaCava suggested looking for the trajectory which has the smallest number of corrector cycles per predictor step in integrating leftward; and this yields an order-of-magnitude improvement. In order to find $F(0)$ to five-place accuracy, we had to obtain an accurate value of $p(F)$ analytically for $F \cong 0.9998$ and to use numerical techniques only for smaller values of F. The analytical values came from a pair of parametric power-series about (1, 0) for the coordinates of the desired trajectory. Such series, describing trajectories in the neighborhood of an improper node, can be found from the method of Picard,[7] of which the existence and relevance were pointed out to me by

A. Kuczura. (This method is also mentioned by Kaplan,[8] who does not cover the case of $\lambda = \Lambda$ in which $\gamma_1 = \gamma_2$.)

The curve $p(F)$ can be parametrized by numerical evaluation of the integral

$$t(F) = \int_{F_0}^{F} \frac{du}{p(u)}. \tag{40}$$

This process yields the inverse $t(F)$ of the desired distribution function. As $(F, p) \to (1, 0)$, the numerical integration becomes inaccurate and must be replaced by analytical approximation. The values $F(t)$ for large t can be estimated by the exponential obtained from (40) by assuming that $p(F)$ represents a straight line to the right of some point (F, p). This can be taken, for example, as the straight line of slope γ_1 that passes through (1, 0) or as the chord from (F, p) to (1, 0).

These procedures* yield a table of triples (F, p, t) from which the distribution F and the density p can be plotted, and the sojourn-time distribution G and its transform computed. The moments M_i of the receiver delay about zero can also be found from $p(F)$, as follows: By definition,

$$M_i(F) \equiv E(\omega^i) \equiv \int_0^\infty t^i \, dF(t) = i \int_0^\infty t^{i-1}[1 - F(t)] \, dt, \tag{41}$$

where E is the expectation operator. The integration by parts that leads to the final expression above is valid whenever $1 - F(t)$ goes to zero faster than t^{-i} as $t \to \infty$. By Lemma 8 this is true $\forall i$ if $\gamma_1 < 0$, as is always the case for $0 < \lambda \leqq \Lambda(m)$ according to Lemma 5. (Indeed, we see that F has finite moments of all orders, as do the members of the other family $\{F_\lambda\}$ of stationary distributions.) We can rewrite (41) as an integral with respect to F, using $dt = dF/p(F)$ from (19):

$$M_i(F) = i \int_{F_0}^{1} [t(F)]^{i-1}\{(1 - F)/p(F)\} \, dF. \tag{42}$$

Because these integrals are hard to evaluate precisely, we divide them into two parts as shown next for the case of M_2 . From (42),

$$M_2(F) = 2 \int_{F_0}^{F(T)} t(F) \frac{1 - F}{p(F)} \, dF + 2 \int_{F(T)}^{1} t(F) \frac{1 - F}{p(F)} \, dF.$$

Following the proof of Lemma 8, for sufficiently large T we can write the approximation

* A simpler and more accurate method can be used when m is an integer.

$$t(F) \cong T + \frac{1}{\gamma_1} \ln \frac{1 - F}{1 - F(T)}.$$

Then the last integral above becomes

$$-2 \int_{F(T)}^{1} \left[T + \frac{1}{\gamma_1} \ln \frac{1 - F}{1 - F(T)} \right] \frac{dF}{\gamma_1},$$

where we have replaced $(1 - F)/p(F)$ by $-1/\gamma_1$ because the approximation $p(F) \cong -\gamma_1(1 - F)$ improves as $p(F) \to 0$. This integral can be written as

$$-\frac{2}{\gamma_1} [1 - F(T)] \left[T + \frac{1}{\gamma_1} \int_0^1 \ln u \, du \right],$$

and so M_2 becomes

$$M_2(F) \cong 2 \int_{F_0}^{F(T)} t(F) \frac{1 - F}{p(F)} \, dF - \frac{2}{\gamma_1} [1 - F(T)][T - (1/\gamma_1)]. \tag{43}$$

The integral term can be calculated from the previous evaluation of $F(t)$. As with F, we get a slightly different estimate of M_2 by using the slope $p(F(T))/[1 - F(T)]$ of the chord, in place of γ_1.

Also, we have not yet specified the value T of the time-parameter at which the integral defining $M_2(F)$ is "broken in two." Each such choice leads to a particular estimate of M_2. If T is too small, $p(F)$ is not accurately approximated by a straight line to the right of $(F(T), p(F(T)))$; but as T increases, accurate evaluation of the integral in (43) becomes more difficult. Thus it is useful to evaluate (43) for many values of T. As T increases, at first the calculated values of M_2 smoothly approach a limit which can be estimated graphically; subsequently these values start to behave erratically as the numerical integration loses precision. (A similar method applies in the simpler case of M_1.)

5.2 *Variable Number of Addresses per Message*

The number m, which we first took to be an integer, enters the problem only through the function F_n^m in (3), by way of the definition

$$\omega + \xi = \max_{i \leq m} \omega^{(i)} \tag{2}$$

and the property that the random variables $\omega^{(i)}$ are independent. But suppose instead that the number of addresses of a message takes the value j with probability ν_j, independently of the numbers of addresses of all other messages; and that

$$m = E(j) \equiv \sum_{j=1}^{\infty} j\nu_j \tag{44}$$

is the average multiplicity. Then with probability ν_j we have $\omega + \xi = \max_{i \leq j} \omega^{(i)}$, so that the distribution of $\omega + \xi$ is F^j with probability ν_j. Thus the true distribution of $\omega_n + \xi_n$ that appears in (3) should be

$$\sum_{j=1}^{\infty} \nu_j F_n^j ,$$

which also appears in the recurrences (7) and (37) and leads to the equilibrium equation

$$F = \mathfrak{M}_\lambda(\Sigma_1^\infty \nu_j F^j). \tag{45}$$

[The operator $\mathfrak{M}_\lambda$ is defined in (35b).]

This equation fits nicely into the framework of our earlier results. We now think of the distribution of $\omega + \xi$ as being represented by a function $\Delta(F)$, defined by

$$\Delta(F) = \Sigma_1^\infty \nu_j F^j \tag{46}$$

rather than by the distribution F^m as in Section 2.3. Then the phase-plane differential equation (20a) becomes

$$p'(F) = \frac{\lambda}{p}[F - \Delta(F)] - (1 - \lambda), \tag{47}$$

which leads to an analogue of (22) with m replaced by $\Delta'(1)$ in the linear term and each binomial coefficient $\binom{m}{j}$ replaced by a multiple of the jth derivative $\Delta^{(j)}(1)$ (which is assumed to exist). The analysis of Cases 1 through 5 proceeds as before. In particular, eqs. (25) and (26a) [for γ_1 and $\Lambda(m)$] are valid with m replaced by $\Delta'(1)$. But here, by (46) and (44), $\Delta'(1) = m$; and so the critical load Λ and the limiting slope γ_1 are meaningful and correct even when m is merely the *average* number of addresses per message! It is this result that justifies treating m as a real number exceeding 1 in the analysis of Sections III and IV, in contrast to the original appearance of m as in integer in eqs. (2) and (3).

We may obtain explicit numerical results for a distribution $\{\nu_j\}$ by means of the procedure of Section 5.1, using (47) to find the phase-plane trajectory. No new difficulty is encountered, because we know γ_1 exactly. It is convenient to have a simple expression for $\Delta(F)$, as is possible when $\{\nu_j\}$ has a form allowing explicit summation of the series in (46). For example, when $\{\nu_j\}$ is geometric with parameter q, so that $m = (1 - q)^{-1}$, we find that $\Delta(F) = F/[m(1 - qF)]$.

Unless $\nu_i = \delta_{im}$, we know that $\Delta(F) > F^m$ when $E(j) = m$ in (46). It follows from (20a) and (47) that $F^T_{\lambda,\Delta} > F^T_{\lambda,m}$, where $F^T_{\lambda,\Delta}$ is the (uppermost) solution of (45) and $F^T_{\lambda,m}$ that of (12) and (20). For our purposes, in view of the invariance of Λ and γ_1 under m-preserving changes in $\{\nu_i\}$, it is sufficient to take $F^T_{\lambda,m}$ as an approximation to $F^T_{\lambda,\Delta}$; the latter could be found numerically if needed for engineering purposes. We could also construct a quantitative theory for approximating $F^T_{\lambda,\Delta}$ by some other $F^T_{\lambda,M}$ with properly chosen M, proceeding by way of (47), (40), and this analogue of (38) (with superscript T suppressed):

$$F_{\lambda,m} - F_{\lambda,\Delta} = \mathfrak{M}_\lambda(F^m_{\lambda,m} - \Delta(F_{\lambda,\Delta})).$$

Such an investigation does not seem worthwhile; we merely note that certainly $M < m$.

5.3 *Numerical Results*

We begin by examining the function

$$\Lambda(m) = 2m - 1 - 2(m^2 - m)^{\frac{1}{2}} \tag{26a}$$

that specifies the maximum rate of receiver arrivals that allows of an equilibrium delay-distribution. This function [eq. (26a)] is plotted in Fig. 6. Its most striking property is its rapid decrease as m increases from 1: Indeed, the slope $d\Lambda/dm$ is $-\infty$ at $m = 1+$. The critical load Λ is down to 0.5 at $m = 1.125$ (corresponding, for example, to one-eighth of the messages having two addresses and all the rest, one); other values are $\Lambda(2) \cong 0.172$, $\Lambda(3) \cong 0.101$, and $\Lambda(10) \cong 0.0263$. Since m ranges from 2 to 3 in a number of practical situations, we see how severely the camp-on discipline limits the possible efficiency of a very large multiple-address system. As discussed in Section VII,

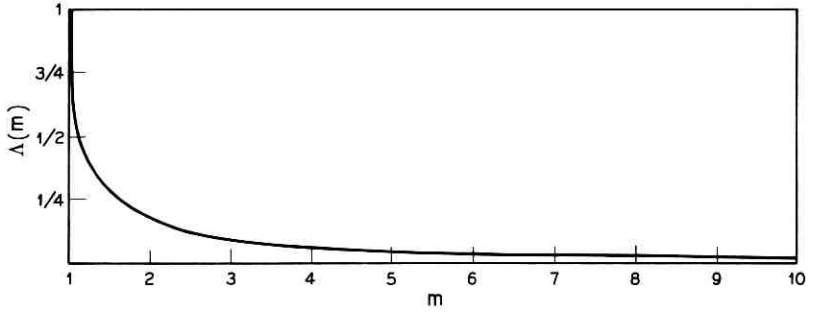


Fig. 6—The critical load for m addresses per message.

this fact was known qualitatively; but explicit knowledge of the function $\Lambda(m)$ is new. It suggests, as mentioned in Section I, that in practice a good system is likely to employ a mixed discipline.

Eight typical functions $F^T_{\lambda,m}$ are shown in Figs. 7–9, as computed–with considerable perseverance, necessitated by the singularity at $(F, p) = (1, 0)$–by LaCava. Figure 7 shows the receiver-delay densities for three values of λ when $m = 2$ addresses per message. The logarithm of the probability density $\dot{F}^T_\lambda$ is plotted against time in units of the mean length of a message, as in Figs. 8 and 9. [The probability F_0 of no delay, and the relative traffic intensity $\lambda/\Lambda(m)$, are shown in Table I for each pair (m, λ) appearing in these figures. This table also shows the mean and the variance for each distribution.] For a given delay, the probability density is least for the lowest load. All these densities appear as not-quite-straight lines, slightly convex, in semi-logarithmic plots. The departure from straightness shows the nonexponential character of these functions, which stems from the nonlinearity of the problem. The uppermost curve in Fig. 7, which corresponds to $\Lambda(2)$, appears to have the greatest curvature near $t = 0$, as the phase-plane geometry suggests. The long straight tails in Figs. 7–9 show how good the exponential approximation is for large t.

The delay densities for $\lambda = 0.10102$ and $m = 1.25$, 2, and 3 appear in Fig. 8. For the uppermost curve, $\lambda = \Lambda(3)$, the critical load.

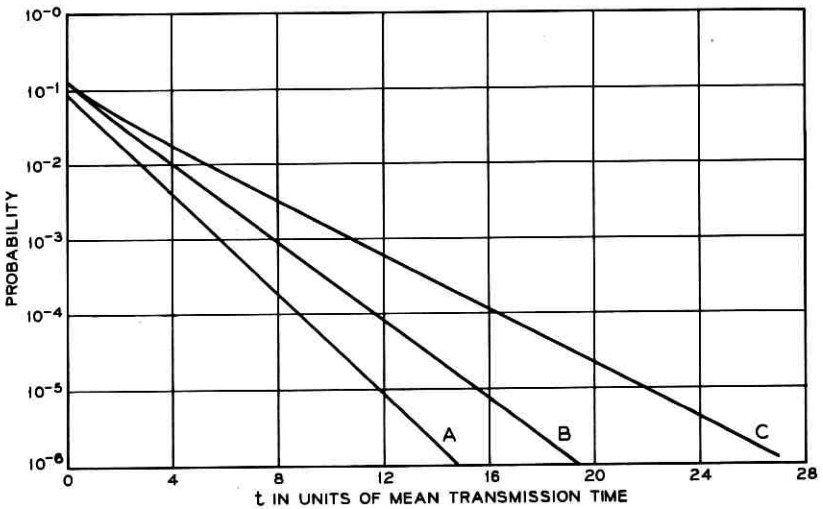


Fig. 7—Delay densities for $m = 2$ and λ equal to (A) 0.10102, (B) 0.15, (C) 0.17157.

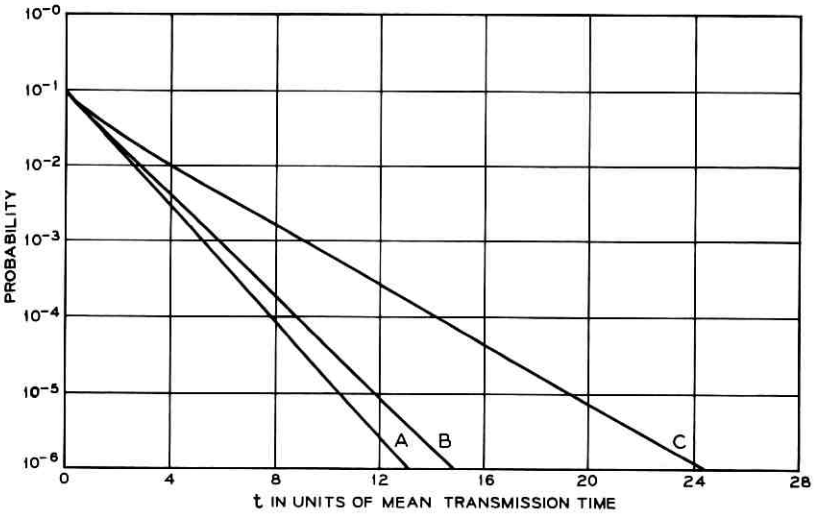


Fig. 8—Delay densities for $\lambda = 0.10102$ and m equal to (A) 1.25, (B) 2, (C) 3.

These functions are shown in Fig. 9 for $\lambda = \Lambda(m)$ and $m = 1.05$, 1.25, 2, 3, and 10. We see that the probability of a given delay *decreases* as m rises; this effect is due of course to the very rapid decrease in critical load with increasing m. In other words, for a large number of addresses per message, equilibrium requires such a small value of offered load that long delays are not likely to occur!

Table I shows that as λ increases for fixed m, or as m increases for fixed λ, F_0 goes down, the mean and variance of the receiver delays increase, and the coefficient of variation goes down. As m increases and λ is kept at its critical value $\Lambda(m)$, all these effects are reversed! However, it is in some ways more instructive to examine the *conditional* coefficient of variation, C_c in Table I: the coefficient of variation of the delay density function, conditional upon the delay being positive. It is C_c that really measures the departure from straightness of the density plots in Figs. 7–9, and we see that these densities become less exponential as λ increases for fixed m, or m increases for fixed λ, or as m *decreases* when $\lambda = \Lambda(m)$.

If the phase-plane trajectory $p(F)$ were straight and the distribution F_λ^T exponential, all the quantities of interest could be calculated exactly without numerical integration. Since C_c is never very far from 1 in Table I, we replace $p(F)$ by a straight line of slope γ_1 and obtain the following approximations for the receiver occupancy and the moments

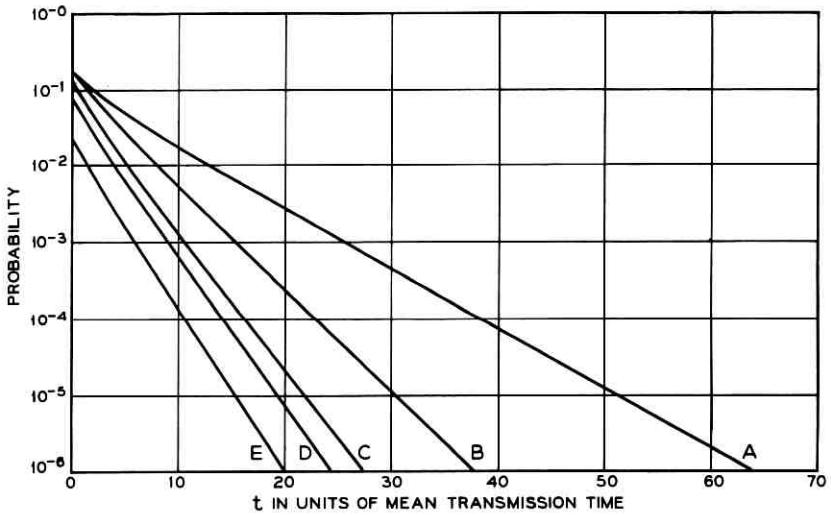


Fig. 9—Delay densities for critical loads and m equal to (A) 1.05, (B) 1.25, (C) 2, (D) 3, (E) 10.

of delay: The line $p = -\gamma_1(1 - F)$ intersects the edge $p = \lambda F$ at the abscissa

$$\tilde{F}_0 = -\gamma_1/(\lambda - \gamma_1). \tag{48a}$$

Likewise we can substitute $-1/\gamma_1$ for $(1 - F)/p(F)$ in (42). When $i = 1$ this yields the estimate $M_1(F) \cong (F_0 - 1)/\gamma_1$, or using (48a),

$$\tilde{E}(\omega) = -\lambda/[\gamma_1(\lambda - \gamma_1)]. \tag{48b}$$

When $i = 2$ it is easier to use (43) directly with $T = 0$, so that $M_2(F) \cong (2/\gamma_1^2)(1 - F_0)$. Substitution of (48a) then gives us

$$\tilde{E}(\omega^2) = 2\lambda/[\gamma_1^2(\lambda - \gamma_1)]. \tag{48c}$$

The estimate (48a) is quite accurate when λ lies well below $\Lambda(m)$. $\tilde{F}_0$ is always too small, but the error does not exceed 1 percent for the three cases in which $\lambda < \Lambda$ in Table I; and for the case $m = 1.25$, $\lambda = 0.10102$ it is only 0.01 percent. For critical load, the error is 19 percent when $m = 1.05$ and $\lambda = 0.64174$, and it decreases to less than 1 percent for $m = 10$.

The error of eq. (48b) is virtually constant when $\lambda = \Lambda(m)$, ranging from 24.5 percent at $m = 1.05$ to 27.3 percent at $m = 10$. $\tilde{E}(\omega)$ is always too large. Like $\tilde{F}_0$, it is a better approximation when $\lambda < \Lambda$, being

TABLE I—PROPERTIES OF RECEIVER DELAY DISTRIBUTIONS AS FUNCTIONS OF LOAD AND MULTIPLICITY

Receiver Load, λ	Mean Number of Addresses per Message, m				
	1.05	1.25	2	3	10
0.64174	$F_0=0.27069$ $C_c=1.075$ $E(\omega)=3.50495$ $C=1.40$ $\mathrm{Var}(\omega)=24.0415$ $\lambda/\Lambda=1$				
0.38197		$F_0=0.50492$ $C_c=1.061$ $E(\omega)=1.41582$ $C=1.82$ $\mathrm{Var}(\omega)=6.6050$ $\lambda/\Lambda=1$			
0.17157			$F_0=0.74481$ $C_c=1.050$ $E(\omega)=0.55612$ $C=2.69$ $\mathrm{Var}(\omega)=2.2383$ $\lambda/\Lambda=1$		
0.15			$F_0=0.80668$ $C_c=1.012$ $E(\omega)=0.31320$ $C=3.08$ $\mathrm{Var}(\omega)=0.9332$ $\lambda/\Lambda=0.8743$		
0.10102		$F_0=0.89606$ $C_c=1.000$ $E(\omega)=0.11941$ $C=4.27$ $\mathrm{Var}(\omega)=0.2602$ $\lambda/\Lambda=0.2645$	$F_0=0.88476$ $C_c=1.003$ $E(\omega)=0.14938$ $C=4.05$ $\mathrm{Var}(\omega)=0.3660$ $\lambda/\Lambda=0.5888$	$F_0=0.84141$ $C_c=1.046$ $E(\omega)=0.32067$ $C=3.49$ $\mathrm{Var}(\omega)=1.2554$ $\lambda/\Lambda=1$	
0.02633					$F_0=0.95599$ $C_c=1.043$ $E(\omega)=0.08281$ $C=6.81$ $\mathrm{Var}(\omega)=0.3184$ $\lambda/\Lambda=1$

off by 6 percent when $m = 2$, $\lambda = 0.15$ and by only 0.15 percent when $m = 1.25$, $\lambda = 0.10102$, and $\lambda/\Lambda = 0.2645$.

The variance of ω, as estimated from (48b) and (48c), is always on the high side and generally less accurate than the approximate mean. The variance is overestimated by 23.5 percent when $m = 1.05$ and by 32.5 percent when $m = 10$, for $\lambda = \Lambda$. The error is 7 percent for $m = 2$, $\lambda = 0.15$, and 0.17 percent for the best case in Table I–that is, when $m = 1.25$ and $\lambda = 0.10102$.

5.4 *Behavior of Systems with Finitely Many Receivers*

In a physical system with fixed finite R, the extent of interdependence among the delays $\omega^{(i)}$ suffered by the copies of one message must increase with λ, as noted on p. 1371. Thus the true distribution of $\omega + \xi$ is less well approximated by F^m as λ grows, and m must be replaced in (26a) by some function of λ which departs increasingly from m. [*Cf.* p. 1404, where $\Delta'(1)$ replaces m.] The critical load for such a system may therefore differ substantially from $\Lambda(m)$.

The predictions of the present theory have been tested against the behavior of physically realizable systems in a modest series of simulation experiments. A few runs were first made to compare the behavior of receiver queues with Poisson arrivals to that found with transmitters in the system. Since there was no perceptible difference with $R = 50$, the remaining experiments simulated only the receiver queues. Arrivals were Poisson; all messages had exactly 2 addresses; and most runs were made with 50 or 200 receivers, although there were several with $R = 100$ and one each with $R = 400$, 500, and 1000.

Representative results are shown in Fig. 10, in which a quantity called "lim F_0" is plotted against load for $m = 2$. The ordinate "lim F_0" represents the asymptotic probability, approached as $t \to \infty$, of finding a receiver idle. In the steady state this quantity is just F_0 ; and for loads above the critical value, for which the queues grow without bound, it is zero. The theory described in this paper (for infinite R) predicts the lower curve in Fig. 10, which is discontinuous at $\Lambda(2) \cong 0.172$. In a system with exactly m receivers (here $R = m = 2$), all messages go to all m receivers, and the m queues behave as identical copies of a single-server queue. In this case, the critical load is unity; $F_0 = 1 - \lambda$; and "lim F_0" is the linear function, reaching zero at $\lambda = 1$, which is shown as the upper curve in Fig. 10. In this case, the dependence among the receiver queues is complete, and it raises the critical load from $\Lambda(2)$ to 1. The intermediate curves are for $R = 50$ and 200 as shown, and the isolated cross marks an approximate (because of the

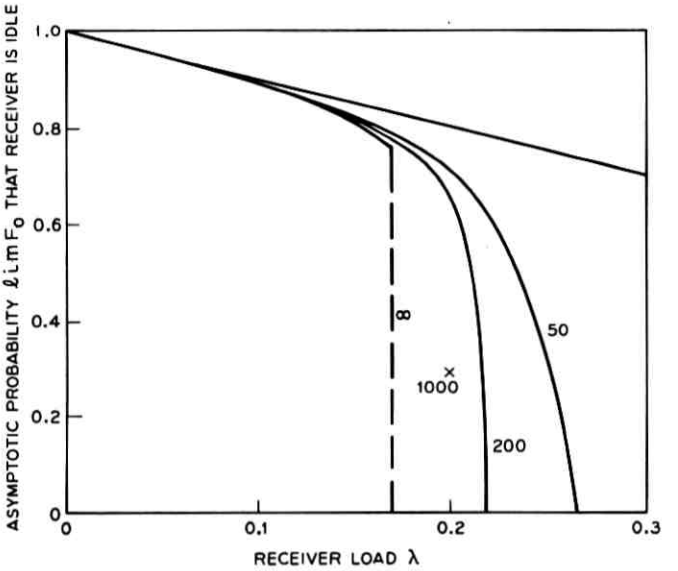


Fig. 10—Idle capacity of receivers as a function of load.

expense of simulation in this case) value for $R = 1000$. We see that the critical load, identified for finite R as the point where "lim F_0" reaches zero, approaches $\Lambda(m)$ from above as R increases. The curves for finite R are shown as slightly above that for infinite R in the region just to the left of $\Lambda(m)$; the difference is small, and the simulated values are not so precise as to guarantee its existence.

Figure 10 indicates that the asymptotic theory is extremely accurate for loads well below Λ and for practical values of R–say from 50 to a few hundred. For loads just below Λ, the behavior of finite systems rapidly approaches that of the infinite model: Even for $R = 50$, the relative error in "lim F_0" at Λ is at most a few percent. The discontinuity in "lim F_0" is not physically realizable: For each finite R, "lim F_0" decreases quickly but smoothly to zero as λ approaches its critical value from below. The family of response curves appears to lie between the straight line for $R = m$ and the discontinuous curve for $R = \infty$ and to approach the latter from the right as $R \to \infty$. This approach is clearly quite slow for $\lambda > \Lambda$, so that the true critical load lies significantly above Λ until R reaches a value at least several thousand. The evidence is similar for the other parameters measured in the simulations: $E(\omega)$, $E(\omega + \xi)$, mean queue-length, and occupancy.

More than this we cannot learn by experiment, at least at acceptable cost.

In summary, the evidence from simulation indicates that our model represents reality well in the range $\lambda \leqq \Lambda$, and that for $\lambda > \Lambda$ it is valid asymptotically but with very slow convergence in R. A precise analysis of the effects of dependence for finite R remains an interesting and difficult question for further research.

VI. TRANSMITTER DELAYS

We now return to the question of transmitter delays which was left behind on p. 1373. The distribution of delay δ could be found numerically from the results of the previous section by evaluating the equilibrium form of (3),

$$G(t) = e^{-t} \int_0^t e^u F^m(u)\, du,$$

to find the distribution G of transmitter service-times ρ; calculating numerically the Laplace transform of G; and numerically inverting the transform, which is obtained from Pollaczek's formula, of the δ-distribution. Such extensive computations do not seem justified in a study of the present kind, and we consider here only the mean delay.

We recall from Fig. 2 and p. 1369 that the receiver sojourn-time is also $\rho = \omega + \xi + x$. The receiver queue can be viewed as a single-server system with delay ω and service time $\xi + x$. Its occupancy is therefore $\lambda E(\xi + x)$ and can also be written $1 - F_0$. Since $\xi + x = \rho - \omega$, we have $\lambda E(\rho - \omega) = 1 - F_0$, so that

$$E(\rho) = E(\omega) + \frac{1 - F_0}{\lambda}. \tag{49}$$

The total receiver-delay suffered by the average *message*, through having to wait in m queues before transmission, is $E(\omega + \xi) = E(\rho - x) = E(\rho) - 1$, or

$$E(\omega + \xi) = E(\omega) - 1 + \lambda^{-1}(1 - F_0). \tag{50}$$

We call this the *receiver queuing-time.*

In order to evaluate $E(\delta)$ by means of the Pollaczek-Khinchin formula, we need two more quantities, of which the first is the transmitter occupancy. Recalling from (1) that the transmitter arrival-rate is $\alpha = k\lambda/m$, where k is the ratio R/X of the numbers of receivers and transmitters, we can write the transmitter occupancy as

$$\alpha E(\rho) = (k/m)[1 - F_0 + \lambda E(\omega)] \tag{51}$$

from (49). The second quantity required is $E(\rho^2)$, the second moment of transmitter service-times. Differentiation of the equilibrium form of (10) yields $\lambda\dot{G} = \lambda\dot{F} - \ddot{F}$, which, after multiplication by t^2 and integration from zero to infinity, becomes

$$\lambda E(\rho^2) = \lambda E(\omega^2) - \int_0^\infty t^2\ddot{F}(t)\,dt.$$

We integrate by parts and divide through by λ to obtain

$$E(\rho^2) = E(\omega^2) + (2/\lambda)E(\omega). \tag{52}$$

The Pollaczek-Khinchin formula (Ref. 2, p. 117) for the mean transmitter-delay involves the transmitter occupancy and the first and second moments of its service time ρ:

$$E(\delta) = \frac{\alpha E(\rho)}{2[1 - \alpha E(\rho)]} \cdot \frac{E(\rho^2)}{E(\rho)}.$$

If we cancel the factors $E(\rho)$ and substitute into the resulting equation the relations (1), (51), and (52), and simplify, we find that

$$E(\delta) = \frac{E(\omega) + (\lambda/2)E(\omega^2)}{(m/k) - [1 - F_0 + \lambda E(\omega)]}. \tag{53}$$

Adding (50) and (53), we obtain the total delay suffered by the average message, prior to transmission, in both transmitter and receiver queues.

We are now in a position to consider the choice of k. The simplest approach is to let $k = m$, which results in equal *utilization* of receivers and transmitters; that is, all terminals spend the same fraction of time with transmission of messages actually taking place. In this case, (53) reduces to

$$E_u(\delta) = \frac{E(\omega) + (\lambda/2)E(\omega^2)}{F_0 - \lambda E(\omega)}. \tag{53a}$$

The denominator of this expression need not be positive. From (49), $F_0 - \lambda E(\omega) = 1 - \lambda E(\rho)$, and the last term is the average number of messages present (waiting and being served) at a single receiver, which need not be less than 1. For fixed m and k, the transmitter delay becomes infinite as the load (α and λ) increases, and this may occur at a value $\lambda < \Lambda(m)$. In other words, for fixed m and $\lambda < \Lambda(m)$, it may be impossible to utilize transmitters as efficiently as receivers: The ratio k may have to be smaller than m. As shown in Table II below, this

phenomenon occurs only for m quite close to 1, so that $\Lambda(m)$ is relatively large.

We could also choose the ratio of numbers of receivers to transmitters by making their *occupancies* equal; that is, all terminals spend the same fraction of time with at least one message present (waiting, camping on, or being transmitted). The condition for this is that $1 - F_0 = \alpha E(\rho)$, which from (51) becomes

$$k = m(1 - F_0)/[1 - F_0 + \lambda E(\omega)]. \tag{54}$$

The corresponding expected transmitter-delay is

$$E_o(\delta) = \frac{(1 - F_0)[E(\omega) + (\lambda/2)E(\omega^2)]}{F_0[1 - F_0 + \lambda E(\omega)]}. \tag{53b}$$

As a third alternative, we could require the mean delays suffered by a message at the transmitter and receiver stages to be equal. The condition for this is that $E(\omega + \xi) = E(\delta)$, which from (50) and (53) yields

$$\frac{m}{k} = 1 - F_0 + \lambda E(\omega) + \frac{\lambda E(\omega) + (\lambda^2/2)E(\omega^2)}{1 - F_0 - \lambda + \lambda E(\omega)}. \tag{55}$$

In this case $E(\delta)$ is given simply by (50).

In a practical context, m and α would be given, and k (and hence λ) would be determined by an appropriate balance of hardware and delay costs. An optimum design can be found only in terms of receiver costs, transmitter costs, and a cost per unit of delay time. We do not pursue this (still oversimplified) system-design problem, but merely illustrate relative performance of transmitters and receivers in terms of the criteria mentioned above. This is done in Table II, which was constructed by calculating the receiver-transmitter ratio k, the transmitter arrival-rate α, the mean transmitter-delay $E(\delta)$, and the mean total delay for a message "Del" $= E(\delta + \omega + \xi)$. These quantities are shown for each design-criterion mentioned above and for each combination of mean number of addresses m and receiver arrival-rate λ listed in Table I. For each choice of m and λ, the mean receiver queuing-time per message $E(\omega + \xi)$, which is not affected by the choice of k, is also shown. This quantity comes from (50); the others come from (54), (55), (1), (53a), and (53b). "Del" is the sum of $E(\delta)$ and $E(\omega + \xi)$.

In the two configurations for which $F_0 - \lambda E(\omega) < 0$, $E(\delta)$ is not defined; it becomes infinite at some smaller value of α. We indicate in Table II how far beyond this singularity α lies by giving the ratio $F_0/(\lambda E(\omega))$. The locus of this singularity could be determined, but is

TABLE II—EXPECTED TRANSMITTER DELAY AND TOTAL DELAY FOR VARIOUS SYSTEM CONFIGURATIONS

Case	Criterion for Choice of $k = R/X$		
	Equal Utilization	Equal Occupancy	Equal Delay
$m = 1.05, \lambda = 0.64174$ $E(\omega + \xi) = 3.641$	$k = 1.05$ $E(\delta) = \infty$ $\alpha = 0.642$ $[F_0/\lambda E(\omega) = 0.120]$	$k = 0.257$ $E(\delta) = 13.71$ $\alpha = 0.157$ Del $= 17.35$	$k = 0.147$ $E(\delta) = 3.641$ $\alpha = 0.090$ Del $= 7.282$
$m = 1.25, \lambda = 0.38197$ $E(\omega + \xi) = 1.712$	$k = 1.25$ $E(\delta) = \infty$ $\alpha = 0.382$ $[F_0/\lambda E(\omega) = 0.934]$	$k = 0.597$ $E(\delta) = 2.897$ $\alpha = 0.183$ Del $= 4.609$	$k = 0.443$ $E(\delta) = 1.712$ $\alpha = 0.135$ Del $= 3.424$
$m = 2, \lambda = 0.17157$ $E(\omega + \xi) = 1.044$	$k = 2$ $E(\delta) = 1.193$ $\alpha = 0.172$ Del $= 2.237$	$k = 1.46$ $E(\delta) = 0.757$ $\alpha = 0.125$ Del $= 1.801$	$k = 1.83$ $E(\delta) = 1.044$ $\alpha = 0.157$ Del $= 2.088$
$m = 2, \lambda = 0.15$ $E(\omega + \xi) = 0.596$	$k = 2$ $E(\delta) = 0.514$ $\alpha = 0.15$ Del $= 1.110$	$k = 1.61$ $E(\delta) = 0.389$ $\alpha = 0.121$ Del $= 0.985$	$k = 2.24$ $E(\delta) = 0.597$ $\alpha = 0.168$ Del $= 1.193$
$m = 1.25, \lambda = 0.10102$ $E(\omega + \xi) = 0.148$	$k = 1.25$ $E(\delta) = 0.151$ $\alpha = 0.101$ Del $= 0.299$	$k = 1.12$ $E(\delta) = 0.133$ $\alpha = 0.091$ Del $= 0.281$	$k = 1.23$ $E(\delta) = 0.148$ $\alpha = 0.100$ Del $= 0.296$
$m = 2, \lambda = 0.10102$ $E(\omega + \xi) = 0.290$	$k = 2$ $E(\delta) = 0.194$ $\alpha = 0.101$ Del $= 0.484$	$k = 1.77$ $E(\delta) = 0.169$ $\alpha = 0.089$ Del $= 0.459$	$k = 2.81$ $E(\delta) = 0.290$ $\alpha = 0.142$ Del $= 0.580$
$m = 3, \lambda = 0.10102$ $E(\omega + \xi) = 0.891$	$k = 3$ $E(\delta) = 0.481$ $\alpha = 0.101$ Del $= 1.372$	$k = 2.49$ $E(\delta) = 0.384$ $\alpha = 0.084$ Del $= 1.275$	$k = 4.78$ $E(\delta) = 0.891$ $\alpha = 0.161$ Del $= 1.782$
$m = 10, \lambda = 0.02633$ $E(\omega + \xi) = 0.754$	$k = 10$ $E(\delta) = 0.091$ $\alpha = 0.0263$ Del $= 0.845$	$k = 9.53$ $E(\delta) = 0.087$ $\alpha = 0.025$ Del $= 0.841$	$k = 61.8$ $E(\delta) = 0.754$ $\alpha = 0.163$ Del $= 1.508$

not of sufficient theoretical interest to calculate here. Suffice it to observe that the delays are in some sense dominated by the transmitter queues for small m and by the receiver queues for large m: For $m = 1.05$ and $\lambda = \Lambda$, transmitter arrival-rates must be much smaller than λ in order for the transmitter queues to be in equilibrium at all; for m greater than some value not much above 1.25, k can equal m for any permissible λ; and for $m = 10$, a very small number of transmitters can handle, without excessive delay, all the traffic that can pass through the receiver queues. This phenomenon is another consequence of the extent to which the camp-on discipline limits receiver utilization at high address-multiplicities. The latter effect is partly illustrated by comparing the mean time spent camping on, $E(\xi)$, with the ordinary receiver-delay $E(\omega)$: The ratio $E(\omega + \xi)/E(\omega)$ generally increases with m, ranging from 1.04 for the first case listed in Table II to 9.1 for the last, and being always a little smaller than m.

The approximation based on assuming F_λ^T to be exponential, with decay-constant $-\gamma_1$, can be used to estimate the quantities of interest in this section as it was in the last. Direct substitution of eqs. (48) yields these "linear" estimates in terms of λ and γ_1 alone: First, from (50),

$$\tilde{E}(\omega + \xi) = -(1 + \gamma_1)/\gamma_1 . \tag{56a}$$

This estimate is always on the high side. Its error is largest in the eighth case in Table II: about 40 percent for $m = 10$, $\lambda = \Lambda$. Next, for the criterion of equal utilization, (53a) leads to

$$\tilde{E}_u(\delta) = \lambda/[\gamma_1(\lambda + \gamma_1)]. \tag{56b}$$

This approximation too is always high, and is worst, with an error of 43 percent, in the third case of Table II, with $m = 2$ and load Λ. [Of course it must get still worse near the singularity of $E_u(\delta)$, which (56b) estimates wrongly as occurring where $\lambda = -\gamma_1$.] For the criterion of equal occupancy we get

$$\tilde{k}_o = m\tilde{F}_0 \tag{56c}$$

from (54) and

$$\tilde{E}_o(\delta) = \lambda/\gamma_1^2 \tag{56d}$$

from (53b). The value $\tilde{k}_o$ is too low in the eight examples considered here, the largest error being about 11 percent at $m = 1.05$; and $\tilde{E}_o(\delta)$ is too large in these examples, with a maximum error of 46 percent at $m = 1.05$. Last, (55) becomes

$$\tilde{k}_d = -m\gamma_1(1 + \gamma_1)/[\lambda(2 + \gamma_1)], \tag{56e}$$

which corresponds to the criterion of equal delay. The error of this estimate takes both signs in these examples, and is worst, at about -10 percent, when $m = 1.05$.

No rigorous analysis of the errors of these approximations has been performed. However, all the estimates in eqs. (48) and (56) are better for $\lambda < \Lambda(m)$ than for critical load; and the cases reported in Tables I and II include a representative sample of values of m from 1.05 to 10, for $\lambda = \Lambda$. Furthermore, all the errors behave monotonically in m for $\lambda = \Lambda(m)$. Since the aforementioned range of m covers all the values that seem likely to arise in communications engineering, it is reasonable to conclude that the estimates (48) and (56), with γ_1 known exactly from (25), can be used in place of exact results of the asymptotic theory whenever a maximum relative error of 50 percent–larger than any of those encountered in these examples–is tolerable in the tabulated quantities.

The present theory predicts no startling qualitative behavior of the transmitter queues, as it does for the receivers. As noted in Section 5.4, the receiver queues behaved no differently in the presence of transmitters than they did with Poisson input, in the few cases simulated (with both sub- and supercritical loads). Thus no detailed records were kept of the transmitter-queue parameters in these simulation runs; and the remarks of Section 5.4 may be considered to apply to the whole system as well as to the receivers alone.

VII. DISCUSSION

This section includes a brief summary of the argument and results of this paper, a discussion of its relation to other literature, and a statement of problems that remain open.

7.1 *Summary*

After an Introduction relating the camp-on problem to the question of engineering for multiple-address traffic in data communication systems, a specific model of a camp-on system is described in Section II. This model is reduced by an informal argument to an idealized mathematical model of receiver delays, of interest in its own right and characterized by Poisson arrivals, exponential transmission-times, and the fact that each message can be transmitted only after the longest of m independent delays in receiver queues has ended. [Key symbols appear in eq. (1) and Fig. 2.] The mathematical model yields the integral

recurrence (7) for the distributions of successive delays at a given receiver. This recurrence leads to an integral equation [eq. (12)] in statistical equilibrium and also to an equivalent differential system [eqs. (18), given in Section III]. The method used to establish the fundamental recurrence, though based on Lindley's equation, is new in its treatment of the delays suffered by a message in parallel queues, and very simple; it should prove useful in other traffic problems involving messages which must wait in several queues at once. Another important element of our approach is a method for separation of the transmitter and receiver delay-analyses, which should be useful in other two-stage queuing problems in which servers in both stages are released (or seized) simultaneously.

In Section III the equation for the equilibrium distribution of receiver delays is reduced to the first-order differential system (20). Analysis of the corresponding vector-field and its topology in the phase plane shows that, for receiver arrival-rates λ not exceeding a critical load Λ, a one-parameter family of distributions exists, each of which satisfies all requirements for a solution to the problem of delays encountered in equilibrium. The uppermost member of this family is qualitatively distinct from all the others. The critical load Λ is found explicitly [eq. (26a)] as a function of the number m of addresses per message; above this load, statistical equilibrium cannot exist.

In Section IV we return to the recurrence (7) to show that, although the integral equation (12) has infinitely many solutions, any reasonable assumption about the previous history of the system leads to a unique limiting-distribution when the load λ is held constant indefinitely at a value smaller than Λ; and the distribution in question is the uppermost of the equilibrium solutions. A slightly weaker result holds when $\lambda = \Lambda$: The existence of the limit can be guaranteed only if the system has not previously been subjected to too great an overload. Even if it has, delays do not increase indefinitely when the load is held at its critical value; instead, the distributions of delay encountered by all subsequent messages are bounded below by some member of the family of stationary distributions corresponding to $\lambda = \Lambda$. (See Theorems 4 and 5 and Corollaries 3 and 4.) So far as I am aware, no other example has been reported of a queuing system which can operate in equilibrium–and with delays having finite moments of all orders–*at* (not just *below*) its critical load when the basic service-process (here exponential) admits of arbitrarily long holding-times. This startling result applies, of course, only to the asymptotically large system which is not physically realizable. The structure of the proofs in Section IV also shows that

analytical techniques are necessary, and that numerical study of the sequences generated by (7) could not solve the uniqueness problem.

A procedure for calculating the limiting distributions of receiver delay is described in Section V. The computation of the phase-plane trajectories near a singular point required considerable effort, and when λ was near Λ rested on the use of a generally neglected series expansion developed by Briot and Bouquet in 1856 and by Picard in 1908.[7] Graphs of eight typical distributions are shown (Figs. 7–9) along with their means, variances, and probabilities of no delay (Table I). The receivers spend so much time being camped on, and this factitious loading so limits their useful capacity, that mean delays and probabilities of delay generally *decrease* as the number m of addresses per message rises and λ stays at the same fraction of critical load. This remarkable behavior shows how inefficient is the camp-on discipline in its pure form. In Section V, it is also shown that the previous analysis can be validly interpreted with minor quantitative changes when m is not an integer. A consequence of this model is the explicit representation of $\Lambda(m)$, which is shown (*cf.* Fig. 6) to decrease so rapidly above $m = 1$ as to account for the curious reduction of delays mentioned above.

The numerical predictions of the asymptotic theory are compared with simulation results in Section 5.4. For a physical system with a finite number R of receivers, the true critical load, along with other indicators of performance such as the probability F_0 of finding a receiver idle, depends on R. Convergence to the predicted behavior as R increases is rapid for $\lambda \leqq \Lambda(m)$, so that in this range the idealized model is very accurate for such values of R as are likely to arise in engineering. When $\lambda > \Lambda(m)$, the effects of interdependence among the receiver queues dissipate very slowly as R approaches infinity. For fixed finite R, the critical load exceeds $\Lambda(m)$ and the idle capacity of the receivers tends to zero as the load increases toward its critical value. The critical load falls with increasing R, approaching $\Lambda(m)$ from above, and the changes in performance parameters that occur as λ approaches its critical value from below become more abrupt. The discontinuity in system behavior at $\Lambda(m)$, that is characteristic of the idealized system, cannot be realized and is approached only asymptotically by response curves for increasing R. Fig. 10 illustrates these effects.

In Section VI the earlier results on receiver performance in the asymptotic model are related to the behavior of the transmitter queues. The choice of the design parameter k is considered (k being the ratio of numbers of receivers to transmitters), and the mean transmitter-delay is calculated for some representative configurations (Table II).

Each transmitter queue behaves as an ordinary M/G/1 system whose service time is the receiver sojourn-time. Critical receiver-loads are so strictly limited for large m that queuing at transmitters is not a severe problem; but when m is near 1, receiver delays contribute so greatly to transmitter occupancies that transmitter arrival rates must be held substantially below those for receivers.

7.2 *Related Literature*

Many of the most important unsolved problems of congestion theory relate to queues in parallel or in series which interact in complex ways. The work reported here is significant partly because it constitutes a partially successful attack on a problem of this type and may, as discussed above, lead to the solution of others. An early example of studies of the same class is the elegant paper by Kingman[9] on two queues in parallel, where each new arrival joins the shorter queue. Studies of other important mechanisms of interaction are gradually becoming more common in the literature. I cite as a recent example involving communications traffic the paper by Cooper and Murray.[10] Hunter[11] surveys the literature on exactly two queues in parallel; and a heavy-traffic approximation for many such queues, with customers randomly choosing which queue to join, is given by Whitt.[12]

It was mentioned in Section I that the camp-on problem itself is treated in a paper by Haenschke.[1] He does not consider the delays encountered at transmitters, or the relation between the transmitting and receiving stages of a camp-on system, but analyzes the receiver queues–the essential component of the problem–by means of a clever approximation. He assumes, as we do, that arrivals at each receiver are Poisson, that transmission times are exponential, and that delays in different receiver queues are independent. He also assumes that all receiving locations have the same number of lines; in our model this number is always 1, but in Haenschke's paper it can be any positive integer. His technique is based on the assumption that the receiver service-time $\xi + x$ is exponentially distributed, so that by Erlang's delay theory (see Ref. 2 or Ref. 4, for example) the receiver delay distribution F is also exponential. The resulting model is in essence a linear approximation to the nonlinear one analyzed here. Haenschke's results (which are not directly comparable with those presented in Section V above) are adequately convincing with regard to the drawbacks of the camp-on discipline in practice. They do not, on the other hand, yield any inkling of the qualitative implications of assuming the receiver queues to behave independently.

This system is also discussed by Weber[13] in an unpublished report which includes extensive simulation results. At loads below our values of $\Lambda(m)$, where the magnitude of R is not very important, his figures agree closely with those presented here. For example, when $m = 2$ and $\lambda = 0.102$, Weber's simulation gives for the receiver occupancy and the mean queuing-time per message [$1 - F_0$ and $E(\omega + \xi)$ in our notation] the values 0.116 and 0.287. Our calculation for $\lambda = 0.101$ (from Tables I and II) yields 0.115 and 0.290 respectively (and see Fig. 10). There being no extant record of the values chosen for R in Weber's runs, his results cannot be compared in detail with ours when $\lambda > \Lambda(m)$. However, examination of his printout shows no inconsistency with the discussion in Section 5.4 above.

It is amusing to observe in closing that our basic differential equation (18a) agrees, except for the sign of the coefficient of F, with the homogeneous (zero driving-term) equation for the anharmonic oscillator. This equation has recently been studied (see Bloembergen,[14] for example) in connection with nonlinear optics. However, the necessary methods do not overlap: We are interested in decaying solutions, while in optics the oscillatory solutions (*cf.* the change of sign just mentioned) are relevant and are obtained by perturbation techniques good only for very small values of the parameter we call λ.

7.3 *Open Questions*

A number of issues raised in this paper are clearly in need of further investigation. Most important, of course, would be an exact analysis of the receiver queues for finite R, quantifying the effects of dependence among them and the rate of convergence to the asymptotic model as $R \to \infty$. It would be particularly useful and interesting to have an analytical expression for the critical load as a function of R as well as m.

Our results should be extended to cover other arrival and transmission-time distributions, especially the case of constant message-length. It would be important to solve the present problem (and many other queuing problems!) without the assumption of complete symmetry–that is, of equal loads on all receiving stations. And the present work should of course be extended to the case of more than one receiving line per location, which was treated by both Weber[13] and Haenschke.[1]

This paper reports on a technique which is new in detail, though not in principle, and describes a curious qualitative result on the behavior of a queuing system (albeit not a physically realizable one) at critical load. It will be interesting to test the technique on other problems

involving complex interactions between queues, and to find the domain of validity of the qualitative result.

VIII. ACKNOWLEDGMENTS

I am grateful to a number of colleagues for their help: to R. W. Hamming, J. A. Morrison, and especially David R. Cox for stimulating discussions and useful advice; to A. Kuczura for acquainting me with Picard's method, which made the computations possible in the case of critical load; to Barbara R. LaCava for obtaining the simulation results for finite R and for her diligent pursuit of the delay distributions; to P. J. Burke, A. Descloux, E. O. Elliott, A. J. Fabens, M. Segal, and the referee for useful comments on various drafts of this paper; and most of all to Henry J. Landau for much help, for the idea which led to the results of Section IV, for penetrating comments on an early draft of this paper, and for constant encouragement, without which this work would not have been completed.

REFERENCES

1. Haenschke, D. G., "Analysis of Delay in Mathematical Switching Models for Data Systems," Bell Sys. Tech. J., *42* (1963), pp. 709–736.
2. Khintchine, A. Y., *Mathematical Methods in the Theory of Queueing*, 2nd ed., London: Griffin, and New York: Hafner, 1969. (First published in Russian, 1955.)
3. Cox, D. R., and Smith, Walter L., "On the Superposition of Renewal Processes," Biometrika, *41* (1954), pp. 91–99.
4. Cox, D. R., and Smith, Walter L., *Queues*, London: Methuen, and New York: John Wiley & Sons, 1961.
5. Lindley, D. V., "The Theory of Queues with a Single Server," Proc. Camb. Phil. Soc., *48* (1952), pp. 277–289.
6. Coddington, E. A., and Levinson, N., *Theory of Ordinary Differential Equations*, New York: McGraw-Hill, 1955.
7. Picard, Émile, *Traité d'Analyse*, 3rd ed., Paris: Gauthier-Villars, 1928, Vol. 3, Chap. II, Sec. II, pp. 30–34.
8. Kaplan, W., *Advanced Calculus*, Reading: Addison-Wesley, 1952.
9. Kingman, J. F. C., "Two Similar Queues in Parallel," Ann. Math. Statist., *32* (1961), pp. 1314–1323.
10. Cooper, R. B., and Murray, G., "Queues Served in Cyclic Order," Bell Sys. Tech. J., *48* (1969), pp. 675–689.
11. Hunter, Jeffrey J., "Two Queues in Parallel," J. Roy. Statist. Soc. (B), *31* (1969), pp. 432–445.
12. Whitt, Ward, "Multiple Channel Queues in Heavy Traffic. III: Random Server Selection," Adv. Appl. Prob., *2* (1970), pp. 370–375.
13. Weber, J. H., "Delays in Line Switching Systems with Multiple Address Traffic," unpublished work (1959).
14. Bloembergen, N., *Nonlinear Optics*, New York: Benjamin, 1965.

Contributors to This Issue

Note: The anniversary article, "Fifty Years of BSTJ," was prepared under the auspices of the BSTJ Editorial Committee. Principal contributors were former Editorial Committee Chairman (1958–66) A. C. Dickieson, and current Committeemen A. E. Joel, Jr., and W. E. Danielson.

Yo-Sung Cho, B.S.E.E., 1962, Seoul (Korea) National University; M.S., 1966, and Ph.D., 1968, Yale University; Honeywell E.D.P. Division, 1964–1965 and 1967–1969; Bell Laboratories, 1969—. Mr. Cho has made equalization studies of the L5 Coaxial Transmission System employing manual and automatic equalizers. He was also engaged in the development of the equalizer adjustment unit which will be used for the optimal equalization of the L5 system. He is currently working on exploratory repeaters for the new long-haul coaxial transmission system. Member, IEEE.

Cecil H. Coker, B.S. (Electrical Engineering), 1954, and M.S. (Electrical Engineering), 1956, Mississippi State University; Ph.D. (Electrical Engineering, minor in Physics), 1960, University of Wisconsin; Electrical Engineering Department, University of Wisconsin, 1960–1961; Bell Laboratories, 1961—. Mr. Coker worked for several years on analysis and synthesis of speech. Subsequent work includes supervision of the development of four laboratory computer facilities, and continued research in the synthesis of speech from English text. Mr. Coker, a supervisor in the Acoustics Research Department, holds several patents and has written a number of papers on speech analysis and synthesis.

James L. Flanagan, B.S., 1948, Mississippi State University; S.M., 1950, and Sc.D., 1955, Massachusetts Institute of Technology. Faculty of Electrical Engineering, Mississippi State University, 1950–1952; Air Force Cambridge Research Center, 1954–1957. Bell Laboratories, 1957—. Mr. Flanagan has worked in speech and hearing research, computer simulation and digital encoding, and acoustics research. He is Head, Acoustics Research Department. Fellow, IEEE; Fellow, Acoustical Society of America; Tau Beta Pi; Sigma Xi; member of several government and professional society boards, including com-

mittees of the National Academy of Sciences and the National Academy of Engineering.

KENZO ISHIZAKA, B.S., 1953, and M.S., 1955, Tohoku University, Sendai, Japan; Toyo Communication Equipment Co., 1955–1962; University of Electro-Communications, Tokyo, Japan, 1962—. Bell Laboratories, 1970–1971. Mr. Ishizaka has been engaged in research on computer simulation of speech. Member, Acoustical Society of America, Acoustical Society of Japan, Institute of Electronics and Communication Engineers of Japan.

NUGGEHALLY S. JAYANT, B.Sc., 1962, University of Mysore (India); B.E. (Distinction), 1965, and Ph.D., 1970, Indian Institute of Science, Bangalore; Research Associate, Stanford Electronics Laboratories, 1967–68; Visiting Scientist, Indian Institute of Science, January–March, 1972; Bell Laboratories, 1968—. Mr. Jayant has worked on digital communication in the presence of burst-noise, and on the detection of fading signals. His current interests include source encoding and pattern discrimination. Member, IEEE.

WALTER J. KROPFL, B.S.E.E., 1956, Pennsylvania State University; M.S.E.E., 1970, Stevens Institute of Technology; Walter Reed Army Institute of Research, 1957–1966; Bell Laboratories, 1966—. Mr. Kropfl is a member of the Sensory and Perceptual Processes Department where he is currently engaged in the design of mini-computer hardware and software systems to control psychological and psychophysical experiments. Member ACM, Eta Kappa Nu.

ANATOL KUCZURA, B.S. (Engineering Physics), 1961, University of Illinois; M.S. (Mathematics), 1963, University of Michigan; M.S.E.E., 1966, New York University; Ph.D. (Mathematics), 1971, Polytechnic Institute of Brooklyn; Bell Laboratories, 1963—. From 1963 to 1966, Mr. Kuczura worked in military systems engineering. Since 1966, he has been engaged in research on the application of probability theory and stochastic processes to the analysis of telephone traffic and queueing. Member, ORSA, SIAM, American Mathematical Society, Mathematical Association of America, AAAS, Chi Gamma Iota, Pi Mu Epsilon.

TERRENCE A. LENAHAN, B.S. and M.S. (Electrical Engineering), 1964, Massachusetts Institute of Technology; Ph.D (Applied Mathe-

matics), 1970, University of Pennsylvania; Bell Laboratories, 1970—. Mr. Lenahan's major interests at Bell Laboratories have been integral and differential equations related to crosstalk in cables. Member, Eta Kappa Nu, Sigma Xi.

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 Laboratories, 1957—. At Siemens and Halske, Mr. Marcuse was engaged in transmission research, studying coaxial cable and circular waveguide transmission. At Bell Laboratories, he has been engaged in studies of circular electric waveguides and work on gaseous masers. He spent one year (1966–1967) on leave of absence from Bell Laboratories at the University of Utah. He is presently working on the transmission aspect of a light communications system. Mr. Marcuse is the author of two books. Member, IEEE, Optical Society of America.

JAMES MCKENNA, B.Sc. (Mathematics), 1951, Massachusetts Institute of Technology; Ph.D. (Mathematics), 1961, Princeton University; Bell Laboratories, 1960—. Mr. McKenna has done research in quantum mechanics, electromagnetic theory, and statistical mechanics. He has recently been engaged in the study of nonlinear partial differential equations that arise in solid state device work, and in the theory of stochastic differential equations.

J. A. MORRISON, B.Sc., 1952, King's College, University of London; Sc.M., 1954, and Ph.D., 1956, Brown University; Bell Laboratories, 1956—. Mr. Morrison has been doing research in a variety of problems in mathematical physics and applied mathematics. His recent interests have included the theory of stochastic differential equations and propagation in random media. He was a Visiting Professor of Mechanics at Lehigh University during the Fall semester, 1968. Member, American Mathematical Society, SIAM, Sigma Xi.

J. R. PIERCE, B.S., 1933, M.S., 1934, and Ph.D., 1936, California Institute of Technology; Bell Laboratories, 1936–1971. Mr. Pierce has worked in the fields of microwave tubes and communication, communication satellites, and acoustics. At Bell Laboratories he became Executive Director–Research, Communication Sciences Division. He is now Professor of Engineering at the California Institute of Technology. Mr. Pierce has received various awards including the Morris

Liebman Memorial Award, the Edison Medal, the Poulsen Medal, and the Cedegren Medal. He has received seven honorary degrees.

LAWRENCE R. RABINER, S.B. and S.M., 1964, and Ph.D. (E.E.), 1967, Massachusetts Institute of Technology; Bell Laboratories, 1962–1964, 1967—. Mr. Rabiner has worked on digital circuitry, military communications problems, and problems in binaural hearing. Since 1967, he has been engaged in research on speech communication, signal analysis, digital filtering, and techniques for waveform processing. Member, Eta Kappa Nu, Sigma Xi, Tau Beta Pi, IEEE; Fellow, Acoustical Society of America. Mr. Rabiner is chairman of the IEEE G-AE Technical Committee on Digital Signal Processing, associate editor of the IEEE Transactions on Audio and Electroacoustics, and member of the technical committees on speech communication of both IEEE and the Acoustical Society.

ERIC WOLMAN, A.B., 1953, A.M., 1954, and Ph.D., 1957, Harvard University; Bell Laboratories, 1957—. Mr. Wolman has been concerned with various aspects of traffic flow in communication systems, and now heads the Traffic Research Department. He is a member of the editorial board of the SIAM Journal on Applied Mathematics, and served as visiting lecturer in applied mathematics at Harvard in the Spring term of 1964. Member, AAAS, AMS, IEEE, ORSA, Phi Beta Kappa, Sigma Xi, SIAM.