

THE NOVEMBER 1982
VOL. 61, NO. 9, PART 1



BELL SYSTEM
TECHNICAL JOURNAL

Feeder Planning Methods for Digital Loop Carrier	2129
B. Bulcha, L. E. Kodrich, D. B. Luber, W. J. Mitchell, M. A. Schwartz, and F. N. Woomer	
Evaluation of Private Networks	2143
R. L. Kaufman and A. J. M. Kester	
A Description of the Bell Laboratories Scanned Acoustic Microscope	2167
P. Sulewski, D. J. Bishop, and R. C. Dynes	
A Statistical Model of Multipath Fading on a Space-Diversity Radio Channel	2185
W. D. Rummeler	
Matrix Analysis of Mildly Nonlinear, Multiple-Input, Multiple- Output Systems With Memory	2221
A. A. M. Saleh	
Computing the Distribution of a Random Variable via Gaussian Quadrature Rules	2245
M. H. Meyers	
A 9.6-kb/s DSP Speech Coder	2263
R. E. Crochiere, R. V. Cox, J. D. Johnston, and L. A. Seltzer	
An Improved Model for Isolated Word Recognition	2289
J. M. Tribolet, L. R. Rabiner, and J. G. Wilpon	
A Theoretical Model of Transient Heat Transfer in a Firestopped Cable Bundle	2313
P. B. Grimado	
Planning and Conducting Field-Tracking Studies	2333
S. J. Amster, G. G. Brush, and B. Saperstein	
LETTER TO THE EDITOR	2365
CONTRIBUTORS TO THIS ISSUE	2367
PAPERS BY BELL LABORATORIES AUTHORS	2373
CONTENTS, DECEMBER 1982	2381

THE BELL SYSTEM TECHNICAL JOURNAL

ADVISORY BOARD

D. E. PROCKNOW, *President*, *Western Electric Company*
I. M. ROSS, *President*, *Bell Telephone Laboratories, Incorporated*
W. M. ELLINGHAUS, *President*, *American Telephone and Telegraph Company*

EDITORIAL COMMITTEE

A. A. PENZIAS, *Chairman*
M. M. BUCHNER, JR. R. A. KELLEY
A. G. CHYNOWETH R. W. LUCKY
R. P. CLAGETT R. L. MARTIN
T. H. CROWLEY J. S. NOWAK
B. P. DONOHUE, III L. SCHENKER
I. DORROS G. SPIRO
J. W. TIMKO

EDITORIAL STAFF

B. G. KING, *Editor*
PIERCE WHEELER, *Managing Editor*
LOUISE S. GOLLER, *Assistant Editor*
H. M. PURVIANCE, *Art Editor*
B. G. GRUBER, *Circulation*

THE BELL SYSTEM TECHNICAL JOURNAL is published by the American Telephone and Telegraph Company, C. L. Brown, Chairman and Chief Executive Officer; W. M. Ellinghaus, President; V. A. Dwyer, Vice President and Treasurer; T. O. Davis, Secretary.

The Journal is published in three parts. Part 1, general subjects, is published ten times each year. Part 2, Computing Science and Systems, and Part 3, single-subject issues, are published with Part 1 as the papers become available.

The subscription price includes all three parts. Subscriptions: United States—1 year \$35; 2 years \$63; 3 years \$84; foreign—1 year \$45; 2 years \$73; 3 years \$94. Subscriptions to Part 2 only are \$10 (\$11 Foreign). Single copies of the Journal are available at \$5 (\$6 foreign). Payment for foreign subscriptions or single copies must be made in United States funds, or by check drawn on a United States bank and made payable to The Bell System Technical Journal and sent to Bell Laboratories, Circulation Dept., Room 1E-335, 101 J. F. Kennedy Parkway, Short Hills, N. J. 07078.

Single copies of material from this issue of The Bell System Technical Journal may be reproduced for personal, noncommercial use. Permission to make multiple copies must be obtained from the editor.

Comments on the technical content of any article or brief are welcome. These and other editorial inquiries should be addressed to the Editor, The Bell System Technical Journal, Bell Laboratories, Room 1J-319, 101 J. F. Kennedy Parkway, Short Hills, N. J. 07078. Comments and inquiries, whether or not published, shall not be regarded as confidential or otherwise restricted in use and will become the property of the American Telephone and Telegraph Company. Comments selected for publication may be edited for brevity, subject to author approval.

Printed in U.S.A. Second-class postage paid at Short Hills, N.J. 07078 and additional mailing offices.

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 61

November 1982

Number 9, Part 1

Copyright © 1982 American Telephone and Telegraph Company. Printed in U.S.A.

Feeder Planning Methods for Digital Loop Carrier

By B. BULCHA, L. E. KODRICH, D. B. LUBER, W. J. MITCHELL,
M. A. SCHWARTZ, and F. N. WOOMER

(Manuscript received May 4, 1982)

This paper describes three new approaches to planning for digital loop carrier (DLC) in the exchange feeder loop network. Two of these are manual planning methods that are used to determine the most economical technology for satisfying facility shortages along a feeder route. Both are based on the concept of a distance-oriented crossover point beyond which DLC is the economic relief technology. One of the methods uses a global crossover point for all routes, and the other uses a route-specific crossover point. The third approach is a mechanized tool, the Pair Gain Planning (PGP) program. PGP, which is implemented as a module within the existing Bell System loop planning system, first identifies the appropriate technology for relief and then synthesizes a specific DLC implementation plan. A discussion of system performance and projected Bell System applications of the various methods is included.

I. INTRODUCTION

For most of the twentieth century the subscriber loop network has been dominated by cable technology. Although carrier systems have been applied successfully in the rural environment¹ for a number of years, it has only been in the very recent past that carrier has gained a foothold in the high-growth suburban environment. However, this situation is changing rapidly, and it is predicted that by the mid-to-late 1980s 50 percent of all loop growth will be served by digital loop carrier (DLC).²

To plan for this growth, a detailed study procedure, the Suburban Pair Gain Planning (SPGP) method, was developed.³ SPGP was designed for general DLC applications where existing Bell System tools, which were designed for rural applications,⁴ were not appropriate. The SPGP procedure uses six tabular forms and the existing Bell System planning tool for conventional cable and structure relief—the Exchange Feeder Route Analysis Program (EFRAP)⁵—to determine a DLC relief plan for a feeder route.

SPGP performs the two planning steps necessary for developing a relief plan. First, it determines the appropriate technology for satisfying facility shortages along a feeder route (i.e., cable or DLC). Second, it selects from the possible alternatives an economic implementation plan for the selected technology. Although not optimal, the method has been shown to produce good relief plans on a wide variety of Bell System feeder routes.

Since SPGP can be very time consuming to perform, three alternative planning systems have recently been developed for Bell System planning applications. Two of these are manual methods that address the technology decision question and provide at least limited evaluation of alternative implementation plans. The third is a mechanized system that performs both the technological decision-making and the implementation plan evaluation steps.

In this paper we describe both the manual and mechanized systems and discuss their applicability and performance. First, we describe the manual methods. Both are based on the concept of a distance-oriented crossover point, a point on a route beyond which DLC is the economic technology for relief. Second, we describe the mechanized system, the Pair Gain Planning (PGP) program. PGP has been developed as a module that is incorporated within the existing EFRAP system. We conclude the paper with a discussion of experience using the various methods and their projected Bell System applications.

II. A TECHNOLOGY DECISION AID—THE CROSSOVER POINT

2.1 Introduction

A simple decision aid is needed to provide telephone company managers with a rapid means of evaluating proposed feeder projects and to provide planning engineers with a simple scale by which they can measure the economic feasibility of any proposal. In the interoffice trunk-planning environment, this need has traditionally been filled by using a length-oriented crossover point. A similar approach for the feeder plant has been developed.

A carrier crossover point is defined as the point on a route beyond which it is generally more economical to relieve shortages with DLC

than with cable. In contrast to the trunk network, which is a point-to-point network with uniform shortages and growths, the feeder plant is characterized as being branchy, with multiple gauges and demand points, tapering facilities, and shortages staggered over time. For this reason, it has been historically difficult to develop a crossover point for the feeder with any reliability. A new approach was needed.

The following sections describe two new approaches to the development of a crossover point. The first addresses the more traditional problem of defining a global point applicable for any feeder route in the Bell System. The second develops a crossover point unique to each feeder route, which is generally less than or equal to the global crossover point. The global crossover has the advantage of being simple to apply in comparison with a route-specific point. However, a global crossover cannot be applied to individual routes since it only applies in the aggregate. Hence, both points are needed, one to perform top-down DLC studies, the other for bottom-up studies.

2.2 Global crossover model

A global crossover point can be developed from two approaches: (i) from sample statistics of feeder route economics, and (ii) from a model representation of a feeder route. The approach described here uses a model representation because we felt a model is more amenable to parametric “what-if” studies. The accuracy of the model can be verified by comparing results with a sample of actual routes. Development of such a model, its results, and verification are reported below.

The Global Crossover Point model presumes an existing two-gauge cable loop from a central office (CO) to a variable end point. Forecasted growth is collected at the end of the loop and at the gauge change point. Relief requirements under both a cable-only and a DLC plan are then determined to meet this growth rate and associated present worth of expenditures (PWE) costs are computed for each plan.

The model provides the means of evaluating the changes in crossover points by determining the PWE differences between an all-cable and a DLC relief solution. The model predicts the impact on the crossover point of relief growth, loop length, and the shortage sequence.

A significant feature of this model is the presence of fine-gauge cable requirements in the DLC plan. This recognizes that any growth between the CO and the crossover point must be served by cable, even in the DLC plan. However, the timing of this “residual cable” is determined by the sequence with which shortages occur along the length of the route. Two different relief-shortage sequences were used in the model, namely: (i) CO to field and (ii) field to CO. For simplicity, no conduit or other structure costs are assumed in the model. Economic sizing of

cables for each growth rate⁶ and the most current estimates of costs, inflation rates, and the impact of special services are used.

To determine the crossover point, a PWE is computed for each loop length, growth rate, and relief sequence studied for both a cable and a DLC relief plan. These PWEs can then be plotted against growth for each loop length. The crossover point for any growth rate is the point where the PWE for the cable and DLC relief plans are equal. The locus of such points across different length routes provides the desired function of crossover length versus growth rate.

One advantage of this model is the ease of evaluating the impact of different cost assumptions. By graphically shifting the PWE curves, the impact on the crossover point is quickly determined. For example, a net circuit cost advantage was assumed for provisioning special services on DLC rather than cable, and the crossover points were recalculated. The study also looked at the effect on the crossover point of the intangible advantages of digital technology. Such advantages might include lower overall maintenance costs and possibly additional revenues from new services. Results of this study for a typical model and for the two relief sequences mentioned above are shown in Table I.

These results show that adding intangible DLC effects and special service advantages can move a crossover point distance by about 10 percent. The relief sequence assumption has a much larger impact.

A number of routes previously selected as candidates for DLC application were used to verify the model. The PGP program (see Section III) was then run on each route. No PWE penalty or DLC advantage for special services was assumed. All conduit requirements were assumed to be satisfied. The distance from the CO to the closest PGP target section was determined. (A target section is a feeder section that is more economical to relieve with DLC than with cable.) Since we are interested in near-term crossovers, only targets that required DLC relief in the first five years were selected.

The verification confirmed the Sequence 1 (34 kft) worst-case crossover. Of all routes studied (and several major branches on some routes), none had the closest PGP target section beyond 34 kft. Target sections ranged, however, from 8 kft to 34 kft. This suggests that a global crossover can only be used as an upper bound, as in Sequence 1 above. Therefore, Sequence 2 results are not useful as a global crossover.

Table I—Typical model results

	Nominal Cross-over Point	With 10% DLC Advantage	With Special Service Advantage
Relief Sequence 1	34 kft	32 kft	31 kft
Relief Sequence 2	28 kft	26 kft	25 kft

2.3 Prescription design approach for route planning

Prescription design (PD) is a manual planning tool that uses a PWE per line criterion to select one or more route-dependent crossover points. It requires a cable and structure relief plan for the route and a schematic layout showing the PWE for cable and structure (P_i) and the average cumulative growth (g_i) in each feeder section over the study period, as shown in Fig. 1.

PD is intended for a non-EFRAP environment. Therefore, the cable and structure plan for the route is obtained using manual cable and structure sizing curves and economic study tools. These data can be used to develop a route economic profile (the piecewise linear curve in Fig. 2) that indicates the approximate cost per growth line to provide cable and structure relief from any point on the route to the CO.

To obtain a similar economic measure for DLC, a separate calculation is performed. The PWE for providing DLC relief during the same study period for various growth rates is represented by the continuous cost curve in Fig. 2.

The area above the DLC curve depicts where DLC is more economical, and the area below the DLC curve depicts where cable is more economical. The portions of the route profile that fall in this region identify the portions of the feeder route that should be relieved by DLC.

In the example, DLC should be used for relief in route sections 3 and 4. Once the sections for DLC relief are identified, potential remote terminal (RT) sites are located at the field ends and are then timed and sized.

In prescription design, timing and sizing algorithms are provided for two DLC deployment strategies, namely: (i) "growth only," where only growth lines are committed to the RT; and (ii) "growth plus existing," where growth lines plus some existing lines are cut over to the RT to make pairs available to relieve shortages closer to the CO. The timing and sizing procedure is straightforward with the growth-only strategy, but the algorithm is more complex if a growth plus existing strategy is employed.

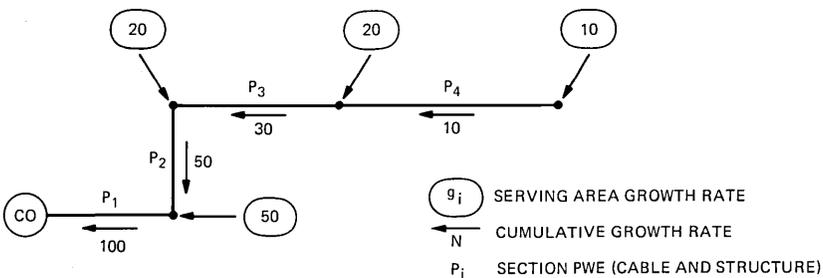


Fig. 1—Route schematic model.

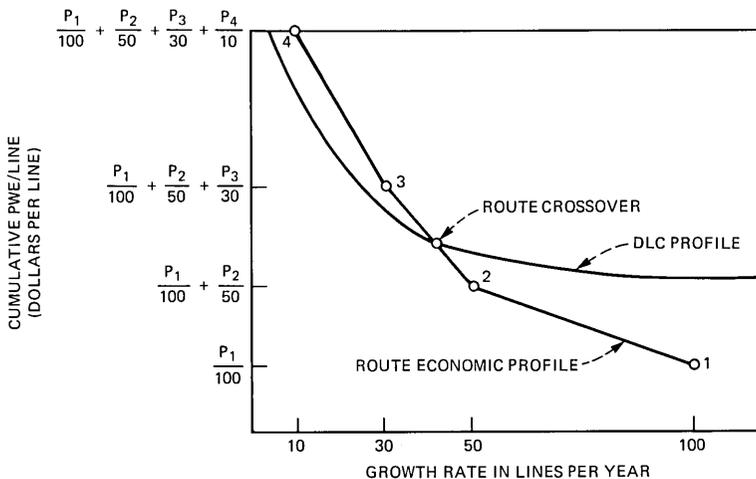


Fig. 2—Prescription design.

III. THE PAIR GAIN PLANNING PROGRAM

3.1 Introduction

The Pair Gain Planning (PGP) program is a new, mechanized planning tool that integrates the technology decision and the selection of the best implementation plan into one mechanized process. After a base EFRAP run, which determines a near optimal all-cable relief plan for a study route, the PGP program uses a network flow model of all existing feeder capacity, proposed cable additions, and potential DLC implementations to find the best combination of cable and carrier to relieve the route, i.e., to make the technology decision. The application of network flow methodology to DLC planning provides near-optimal solutions quickly and efficiently.

After the network flow algorithm has identified sections for DLC relief and a list of potential carrier serving areas (CSAs)⁷ for activation, extensions of SFGP algorithms are used to find the theoretical RT sites that will be activated. The program then times the placement of DLC systems at those sites.

As shown in Fig. 3, the PGP program is designed to be an interactive tool. After the program has developed a carrier relief solution, the user is given the option of modifying the solution to establish a relief plan that is most appropriate to the immediate local environment. When the user is satisfied with the solution, the PGP program creates a modified EFRAP input file that reflects the addition of DLC relief on the route so that EFRAP can be used to determine a new cable relief plan that satisfies any residual shortages that are not economic to relieve with DLC.

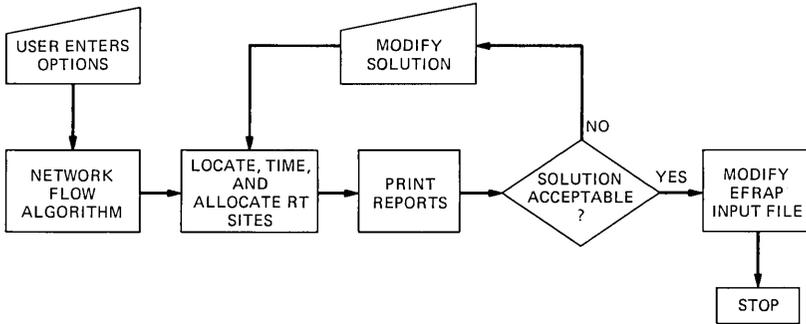


Fig. 3—PGP terminal session.

3.2 Network flow algorithm for target-section selection

The program's first task is to decide which cable and structure placements, as determined by a base EFRAP study, can be relieved more economically with DLC. Generally, it is not economical to use DLC to replace all cables on the route, and the program must locate the economical DLC placements. Since it is not possible to look at each EFRAP section individually to compare cable and DLC alternatives—as EFRAP does in its cable analysis—a global algorithm that examines the entire route at once must be used.

This problem of finding cable relief projects that are more economically relieved with DLC can be solved by a number of algorithms, but the one chosen for the PGP program considers the problem as a minimum-cost network flow problem. The advantages to this approach are that many efficient, easily programmable algorithms have been developed for solving this class of problem⁸ and that standard software already existed for this purpose,⁹ thus shortening the PGP program development time.

The minimum-cost network flow problem has been applied to a wide range of situations, including transportation of goods, design of pipeline systems, and production scheduling.^{8,10,11} The general form of this problem is concerned with the flow of a commodity through a network, which is a directed graph defined by a set of nodes and a set of arcs connecting the nodes. The term "directed" implies that the commodity being studied can flow in only one direction, from a tail node to a head node. For each arc, there is a piecewise linear convex function that defines the cost per unit of flow over this arc as a function of its present flow. Upper and lower bounds of flow are also defined for each arc. Each node is identified as one of three types: (i) a supply node, where flow enters the network; (ii) a demand node, where flow leaves; or (iii) a transshipment node. No storage is permitted at nodes. There is also an objective function that usually minimizes the total cost with

flows that satisfy the upper and lower bounds on each arc and preserve the conservation of flow at each node.

In mathematical form, the network is stated by a node-arc incidence matrix⁸ A (an $I \times J$ matrix if the network has I nodes and J arcs), with elements

$$A_{ij} = \begin{cases} +1 & \text{if arc } j \text{ directed out of node } i, \\ -1 & \text{if arc } j \text{ directed into node } i, \\ 0 & \text{otherwise.} \end{cases}$$

The problem is then stated as:

$$\begin{aligned} & \min cx \\ & \text{such that } Ax = r \\ & 1 \leq x \leq u, \end{aligned}$$

where

- x_j is flow on arc j
- u_j is upper bound on arc j
- l_j is lower bound on arc j
- c_j is cost for arc j (need not be linear in the general case)
- r_i is supply (>0) or demand (<0) at node i .

Figure 4 shows a typical EFRAP feeder route layout, with appropriate dummy sections added for the DLC RT site at the field end of section 1106. The network model for the same route is shown in Fig. 5, including additional arcs used to represent existing and future DLC capacity at potential RT sites.

In the PGP program, the objective function is the PWE cost to provide relief to the route. The commodity flowing in the network is the

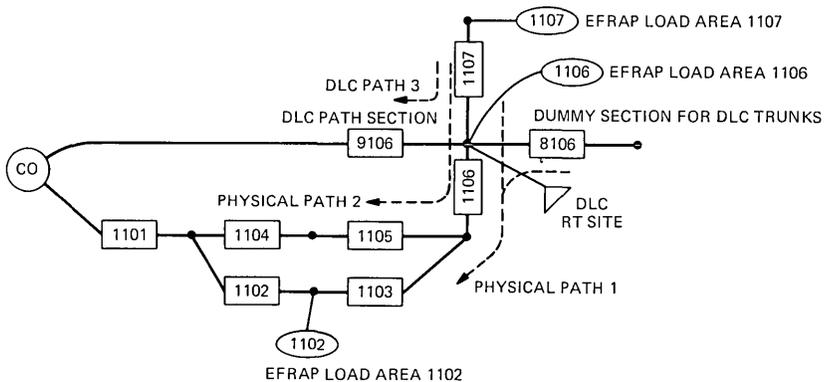


Fig. 4—Typical EFRAP feeder route layout.

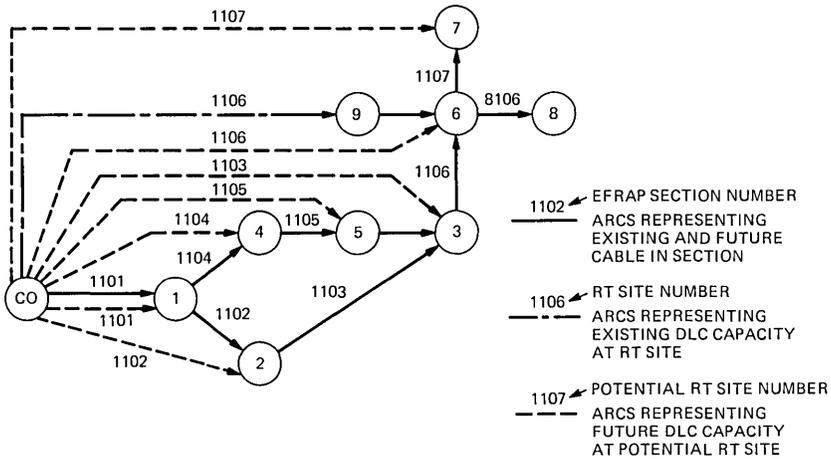


Fig. 5—Network model for typical EFRAP feeder route layout.

demand for loop feeder capacity, both existing and future, which must find its way from the supply node at the CO to the nodes representing the various EFRAP load areas. The demands are the projected requirements for facilities at the end of the study period (typically 20 years). To simplify implementation, this model assumes that all capacity additions and demands for the entire study period occur at the start of the study. The arcs of the model correspond to either (i) EFRAP sections with upper bounds representing existing facilities plus future cable additions from the base EFRAP study, or (ii) future DLC placements with RT sites at their head nodes. All lower bounds are zero.

The cost function is zero for flows below the existing capacity on arcs representing cables or existing DLC in EFRAP sections. When this value is exceeded, a cost is incurred for the facilities that must be placed to support this flow. On cable arcs, this cost increases linearly using the per-pair cost of the cables placed by EFRAP, starting with the least expensive cable placed during the study period. This procedure yields a convex cost function. The cost function for future DLC arcs uses the DLC cost model described above to obtain a per-unit cost with the common equipment and site costs averaged over a full system.

When the solution to the network problem is obtained, those physical arcs whose flow does not exceed their existing capacity are the target sections where DLC relief should be used since the network flow algorithm has found it more economical to route flow over DLC arcs rather than increase flow through these cable arcs (which is equivalent to placing new cable). Also, those future DLC arcs with positive flows represent potential RT sites, since they indicate places where it is more economical to route flow via DLC. The timing and sizing routines

discussed below will select the actual RT sites used in the DLC relief plan from this group.

3.3 Timing and sizing algorithms for RT locations

The procedure of timing and sizing of CSA theoretical RT locations is divided into four parts: (i) the selection of RT sites to activate; (ii) the determination of the amount of DLC, by time, to be placed at the location; (iii) the calculation of the PWE for the electronics placed; and (iv) the production of EFRAP data for the residual cable-relief recommendations.

The following algorithm determines the order for activating RT sites. The goal is to eliminate the earliest shortages on the route first. The route is examined for target sections using the EFRAP path model for the route as a tree. The algorithm searches from the CO to the ends of the route for the branch of the tree that has the target sections with the earliest shortage dates. The RT site that will be activated to relieve this branch is the potential RT site that is on the field side of the most recently selected target section and closest to that target section.

The next step is to determine the amount of DLC needed to be placed at this location. The upper bound on the amount of DLC that is useful to place at that location is determined by the number of assigned pairs and growth pairs in the CSA associated with the RT site and with the cutover strategy that will be used for this RT site. The PGP program allows either a cutover strategy, which places all of the pairs in the CSA on DLC or places just enough pairs to assure that no cable will be needed in the target sections affected by this RT site. To be realistic, this upper bound is decremented to allow for DLC trunk pairs. For each year, the PGP program attempts to cover the shortages in the target sections on this branch with the available DLC.

When enough theoretical RT sites are activated to relieve all target sections, DLC placements at the RT sites are used to revise the original EFRAP input data to reflect the DLC systems placed. These data are processed, and the EFRAP program determines the PWE for the residual cable needed for the relief solution. A comparison of the all-cable and the cable-and-DLC plan PWES then indicates the DLC savings.

3.4 Interactive features

The PGP program is an interactive system. The interactive alternatives are designed to assist the outside plant engineer in forming a DLC relief plan. The user has the choice of calling for an "automatic" solution or of hand-tailoring a solution. Several options are available to modify a solution or to demand a solution. These options create a broad spectrum of possible user control over the program solution. The PGP program interactive options are invoked within a framework

that allows the user to save the status of the work and return to that point at a later date.

3.4.1 Automatic run

The PGP program has the capability of producing a DLC relief plan with very little input from the user at the terminal. The user need only specify the run number of the EFRAP data that will be studied for DLC applications and the name of the DLC system that will be studied. From this input, the PGP program will execute an entire DLC study, producing a summary report that includes:

- RT sites activated
- Number of DLC systems placed
- Schedule for DLC placements
- PWE for DLC placements.

The automatic run of the PGP program produces excellent DLC relief plans and is the recommended starting point for developing a final project.

3.4.2 Solution modification

In addition to the brief output reports given at the terminal with the automatic run, the user can request detailed reports for every RT site that is activated. After studying the output reports for a solution, the outside plant engineer might want to more accurately reflect in the study some conditions of the route. This is accomplished by choosing any of the modifying options:

- Change the DLC cutover strategy
- Change the list of target sections being considered for DLC relief
- Change the list of potential RT sites
- Change the list of forbidden RT sites
- Change the DLC data associated with a particular RT site.

3.4.3 Demand solutions

To compare the costs of making working-pair transfers to the costs of activating additional CSAs, the PGP program demand option is available. During the terminal session, the user associates a first year of activation and a DLC cutover strategy with each demanded RT site. All demanded RT sites are activated by the PGP program in the years specified by the user before any other RT sites that may be necessary. For demanded sites, the timing and sizing algorithm will select either a growth-only cutover strategy or one that cuts all assigned-plus-growth pairs in the CSA onto DLC.

3.5 Results

A side-by-side comparison of the PGP program network flow solution with the PD method was conducted using (EFRAP) route data for ten

Bell System routes. The PGP program DLC plus residual cable and structure solutions were generally lower in cost than was the equivalent PD solution; on an average the network flow solution showed a 7-percent lower PWE. (PWE figures do not include rearrangement costs, which were not computed for either study.)

The PGP program is valuable in that the network flow algorithm provides very good starting solutions, and that the interactive options can be used to easily modify the starting solution to satisfy local requirements. During the PGP program field trial, which was conducted in three Bell System operating companies, users evaluated as many as ten alternative relief plans for a single route. Because the PGP program is a mechanized tool, they were able to do this within two or three days, whereas equivalent SPGP studies would have taken over a month of engineering time.

IV. STUDY PROCEDURE SUMMARY

Since the PGP program performs both phases of feeder planning in an integrated fashion with little manual effort, the Bell System has recommended its use as the primary feeder-planning vehicle for DLC. However, global crossover points and PD both have a place in the planner's arsenal.

Global crossover points are still used in project review to rapidly assess the planner's technology decision. As such, they effectively shift the burden of proof to cable on longer routes. The global crossover can also be used to guide the engineer in the interactive portion of PGP studies. Since many implementations of DLC with similar cost are possible on a route, the crossover can guide the planner to the solution that best meets company objectives, while PGP can measure the economic implications.

For those engineering districts that have not yet implemented the prerequisite EFRAP, PD provides an effective interim procedure for planning digital carrier. (Although circumstances are changing quickly, a private 1981 survey showed that about one third of all Bell System engineering districts fell into this category.) The global crossover can be used in combination with PD to split the route into two parts. Only sections between the crossover and the CO need to be evaluated with PD.

V. CONCLUSION

This paper described three tools for studying feeder relief in the digital age. Each tool has advantages and disadvantages, each has its place. There can no longer be a valid excuse for "business as usual" and the placement of metallic cable exclusively. We expect these tools to provide significant stimulus to the advancement of the digital age.

REFERENCES

1. J. R. Bres, "The Appropriate Application of Analog and Digital Carrier in the Loop Plant," Int. Symp. on Subscriber Loops and Services, Atlanta, Georgia, March 20-24, 1978, Conf. Rec., IEEE Cat. No. 78CH1279-9 COM, pp. 142-145.
2. A. J. Ciesielka and N. G. Long, "New Technology for Loops—A Plan for the '80's," IEEE Trans. Commun., *COM-28*, No. 7 (July 1980), pp. 923-30.
3. B. Bulcha and W. J. Mitchell, "An Economic Planning Method for Pair Gain Systems in Suburban Areas," Nat. Telecommun. Conf., Houston, Texas, November 30-December 4, 1980, Conf. Rec., IEEE Cat. No. 80CH1539-6, pp. 12.2.1-5.
4. B. S. Abrams, A. D. Robinson, and M. R. Plummer, "Long Route Planning Methods," Int. Symp. on Subscriber Loops and Services, Ottawa, Canada, May 20-23, 1974, Conf. Rec., IEEE Cat. No. 74CH0844-1 COMM, pp. 1.4.1-7.
5. C. D. McLaughlin and J. Albers, "Exchange Feeder Route Analysis Program: An Application of Branch and Bound Techniques to Economic Cable Sizing," Int. Symp. on Subscriber Loops and Services, Ottawa, Canada, May 20-23, 1974, Conf. Rec., IEEE Cat. No. 74CH0844-1 COMM, pp. 1.2.1-6.
6. J. Freidenfelds, *Capacity Expansion—Analysis of Simple Models with Applications*, New York: Elsevier North Holland, 1981.
7. T. P. Byrne, R. F. Coburn, H. C. Mazzoni, F. W. Aughenbaugh, and J. L. Duffany, "Positioning the Subscriber Loop Network for Digital Services," IEEE Trans. Commun., *COM-30*, No. 9 (September, 1982), pp. 2006-11.
8. J. L. Kennington and R. V. Helgason, *Algorithms for Network Programming*, New York: John Wiley & Sons, 1980.
9. C. L. Monma and M. Segal, "A Primal Algorithm for Finding Minimum-Cost Flows in Capacitated Networks With Applications," *B.S.T.J.*, 61, No. 6, Part 1 (July-August 1982), pp. 949-68.
10. G. B. Dantzig, *Linear Programming and Extensions*, Princeton, NJ: Princeton University Press, 1963.
11. L. R. Ford and D. R. Fulkerson, *Flows in Networks*, Princeton, NJ: Princeton University Press, 1962.

Evaluation of Private Networks

By R. L. KAUFMAN and A. J. M. KESTER

(Manuscript received March 17, 1982)

This paper describes algorithms for predicting end-to-end performance measurements in private networks. Service characteristics such as end-to-end blocking and delay are calculated based on point-to-point traffic data and a network routing guide. These techniques have been incorporated in the Enhanced Network Administration System. The system is routinely used by AT&T Long Lines and operating company network administrators for private network design and evaluation.

I. INTRODUCTION

An important procedure in network design is performance prediction. In a private network environment, a traffic engineer must recommend a network design that satisfies the customer's performance requirements. A set of computer programs called the Enhanced Network Administration System (ENADS) has been designed for administering Enhanced Private Switched Communications Service (EPSCS) networks and Electronic Tandem Switching (ETS) networks. The ENADS Network Service Evaluator (NETEVAL) uses point-to-point traffic data and a network routing guide to provide detailed service characteristics of a network design. This paper summarizes the NETEVAL algorithms.

In addition to characterizing network service, NETEVAL is used, in conjunction with other ENADS modules,¹ to ensure that the service characteristics chosen by the customer are achieved. The Network Synthesis (NETSYN) module designs a network that is close to the service specified by the customer. The service evaluator is used to ensure that the final network recommendation meets the customer's service requirements.

Traditionally, performance estimates of a large network in an alternate routing environment have been confined to a segment of a network, such as the intermachine trunk portion.² NETEVAL contains

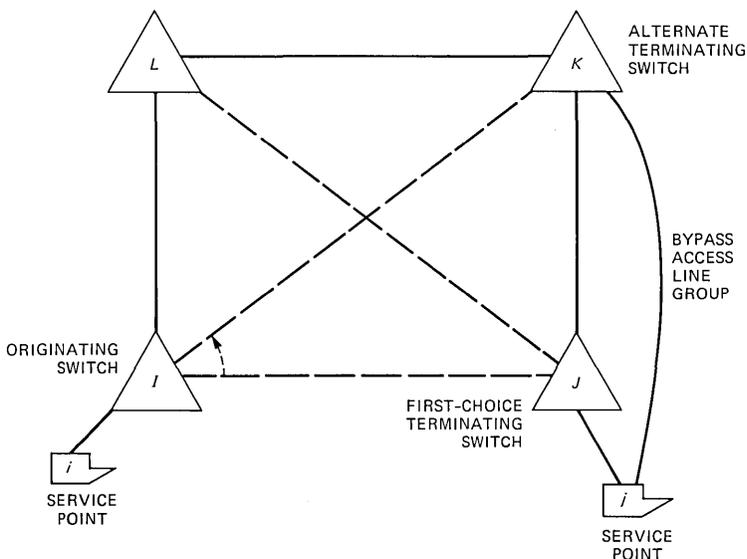


Fig. 1—Private network configuration.

decomposition and aggregation procedures that permit evaluation of complete end-to-end performance of a network as opposed to just a network segment. In addition, the ENADS evaluator generalizes the Katz algorithm to include private network features such as queues, and controlled access to network facilities via Facility Restriction Levels (FRLs). This paper describes the decomposition and aggregation procedures used in the evaluator, as well as the additional tools required to analyze some of the network features germane to private networks.

II. PRIVATE NETWORK CHARACTERISTICS

A private network (see Fig. 1) interconnects customer locations (on-net service points) and other locations (off-net service points). An on-net service point is generally a PBX/Centrex or a key set connected directly to an access line group. Off-net points are served by off-net facilities (Foreign Exchange, WATS, or local off-net access lines). In this paper both access lines and off-net facilities are called end-links. One end of each end-link is associated with a service point and the other end is connected (or homed) to a switch. A network of intermachine trunk groups connects the switches together, forming the intermachine trunk-group portion of the network. Switches permit concentration of point-to-point traffic.

Traffic generally originates at a service point, seizes an access line, and arrives at a switch. (It is possible in an ETS network for a service

point to reside at the same location as the switch, so that an access line is not required.) NETEVAL assumes that the switch examines only the final destination of the traffic and its FRL and then associates an ordered route list of links at that switch. The route list is scanned until an idle circuit is found. If no such circuit exists, the call is blocked (in the absence of queuing). If a circuit is seized, the traffic arrives at the next switch, and the routing procedure continues until the call is either completed or blocked.

Other private network characteristics of interest for evaluation are queuing and FRLs. Queuing permits a call to wait at a specified trunk group for an idle circuit. The call is queued on a trunk group after a search through an ordered route list fails to find an idle circuit. The queued call may still be sent to reorder if the caller's wait exceeds a time-out threshold. Also, the caller may abandon the queue. The queue discipline is first come, first served.

FRLs provide the capability to restrict or expand access to network-facility route lists. The FRL can be based on the calling station and/or an authorization code. This provides the customer the opportunity to offer different grades of service by restrictive routing to different groups of users in the network. The FRL and final destination of a call are used in deciding to which links in the network the call will have access.

III. EARLIER SERVICE EVALUATION ALGORITHMS

Statistical techniques are available for estimating service characteristics of networks. It is well known that if first-offered traffic is generated from a Poisson process and holding times are exponentially distributed, then network characteristics can be obtained by formulating an appropriate Markov chain model and solving the resultant system of birth-death equations. This technique, however, requires an exorbitant amount of storage and computer time and is not considered feasible for large-scale networks.

Another reasonably accurate statistical technique that requires knowledge of only the traffic mean and variance (or the variance-to-mean ratio, called peakedness) is the Katz algorithm,² which was originally designed for evaluating switch-to-switch blocking probabilities. This algorithm requires estimates of switch-to-switch traffic means and variances that are "effectively" offered to each trunk group in the network.

Traffic that is carried on a link may be blocked on subsequent links. The holding time for such traffic on seized circuits is substantially less than that for completed traffic on the network. Effective-offered traffic reflects the shorter holding times of subsequently blocked traffic. The effective-offered load is used to compute the link-blocking probabilities

and traffic-overflow variance for each link. The algorithm is an iterative process that updates effective-offered loads and link parameters at the end of each iteration.

After these initial link parameters have been computed, the switch-to-switch loads are distributed through the network based on the network-alternate routing plan and the link parameters. As this traffic is routed through the network, the effective means and variances of the traffic offered to each link are accumulated to update the link parameters for the subsequent iteration. The blocked traffic is also accumulated to provide probabilities for that iteration. After all the switch-to-switch loads have been distributed throughout the network, the link parameters are updated and the entire procedure is repeated until convergence in the switch-to-switch loss probabilities is obtained.

The Katz algorithm provides switch-to-switch blocking probabilities. Unfortunately, the number of calculations per iteration during the load assignment process is a function of the number of switches in the network. Traffic associated with each switch pair must be offered separately to each link in its routing path. Thus, an N -switch network must distribute $N(N - 1)$ switch-to-switch loads through the inter-machine trunk network. Since private networks generally consist of several hundred service points, it is not computationally feasible to define each service point as a switch and use the Katz algorithm. In addition, the Katz algorithm does not analyze queuing.

IV. NETEVAL ALGORITHM

NETEVAL is an iterative algorithm that successively updates traffic loads and associated traffic characteristics until a convergence criterion is satisfied. Because of the size of the network, NETEVAL decomposes the network into an end-link portion and an extended-trunk portion (see Section 4.4) to compute the traffic characteristics. After convergence the network performance components are aggregated to obtain end-to-end network performance estimates.

Sections 4.1 and 4.2 contain the NETEVAL assumptions and a brief outline of the algorithm. Subsequent sections explain in more detail the decomposition procedure, calculation of network parameters, and the aggregation techniques.

4.1 Model assumptions

The model assumptions for NETEVAL are as follows:

(i) Traffic means and variances provide a sufficient description of the loads.

(ii) The holding time on a link has four components: actual message time, ringing time, link set-up times, and subsequent queue delays.

Under this assumption blocked traffic can contribute positive loads to the network.

(iii) A blocked customer will redial with a specified retrial probability and a variance that is a fraction of the overflow variance.

(iv) No queue time-outs or abandonments occur.

(v) The system is in statistical equilibrium.

(vi) Only the final destination and FRL of a call are used at a switch to determine routing through the network.

4.2 NETEVAL algorithm for service evaluation

The basic algorithm procedure is given below. Some of the terms used are explained more clearly in later subsections.

(i) Decompose the network based on routing into an end-link network portion and an extended trunk portion.

(ii) Associate traffic parcels with each network portion.

(iii) Initialize all parcel blockings. (Zero can be used if no other estimates are available.)

(iv) Compute effective-offered parcels to the end-link network portion.

(v) Calculate blocking probabilities for the end-link parcels.

(vi) Compute total switch-to-switch and switch-to-final-destination parcels offered to the extended trunk portion of the network.

(vii) Calculate switch-to-switch and switch-to-final-destination blocking probabilities.

(viii) If end-link, switch-to-switch, and switch-to-final-destination blocking probabilities change significantly, return to step *iv*. Otherwise, go to step *ix*.

(ix) Compute point-to-point characteristics.

4.3 Decomposition

As we mentioned in Section III, a typical private network is too large to model each service point as a switch and use the Katz algorithm. The network must be decomposed into segments and analyzed separately. To define such a decomposition we must first define a traffic parcel.

The aggregate of point-to-point loads with identical routing, when offered to a particular portion of the network, will be called a traffic parcel for that network segment. For example, if all point-to-point traffic originating on service points homed on switch *I* and destined for service points homed on switch *J* have identical route lists when offered to the inter-machine trunk-group portion of the network, then such traffic forms a switch *I* to switch *J* parcel in the inter-machine trunk-group subnetwork. Any technique using such a traffic aggrega-

tion assumes that the network characteristics of the traffic parcel sufficiently approximate those of the individual point-to-point loads.

One method of decomposition is to separate the network into an inter-machine trunk-group portion and an end-link portion. Switch-to-switch parcels, as defined above, are associated with the inter-machine trunk-group network segment. All point-to-point traffic in the end-link segment, originating from a point to a switch and offered to the same route list of end-links, forms an originating end-link parcel. Similarly, all point-to-point traffic offered in the end-link network from a switch to a service point, using the same route list, is aggregated into a terminating end-link parcel.

Since each parcel's point-to-point load components have identical routing in its associated network segment, blocking can be computed for each parcel using the Katz algorithm. For example, switch-to-switch blocking can be computed for switch-to-switch parcels in the inter-machine trunk-group network segment, and originating and terminating parcel blockings can be computed in the end-link segment.

The above decomposition facilitates aggregation of parcel characteristics to obtain end-to-end characteristics. For example, if points i and j are homed on switches I and J , respectively, then i -to- j blocking (b_{ij}) is

$$b_{ij} = 1 - (1 - B_{i,\text{org}}^I)(1 - B^{IJ})(1 - B_{j,\text{term}}^J). \quad (1)$$

(See Appendix A for a summary of the notations used.)

The three factors in the product represent the probability of call completion for the originating parcel i , the switch I to switch J parcel, and the terminating parcel j , respectively. An implicit assumption in such a decomposition is that all i -to- j traffic must use end-links homed on switch J for call completion. Or equivalently, the last switch, called the terminating switch, that completed i -to- j traffic encounters must be switch J . Figure 1, however, displays a typical private network in which terminating switches are not unique.

It illustrates a two-level hierarchical trunk-structure, with on-net service points homed on lower level switches I and J , respectively. A bypass access line group is used to route traffic from switch K to point j . The two-level trunk hierarchy permits i -to- j traffic to reach point j using either switch J or K as a terminating switch.

Since i -to- j traffic is not required to use the home access line group serving j , the above equation for b_{ij} is inaccurate. The last two terms in the product of the equation do not provide the switch I to point j blocking. To handle routing patterns with nonunique terminating switches, it is necessary to introduce a switch-to-final-destination parcel. Such a parcel represents all point-to-point loads from service points homed on a switch that has identical routing from the switch to

the destination service point. In Fig. 1, a switch I to point j parcel is required to compute i -to- j blocking. If B^{ij} is the blocking probability for such a parcel on the network, then b_{ij} can be accurately estimated by

$$b_{ij} = 1 - (1 - B_{i,\text{org}}^I)(1 - B^{ij}). \quad (2)$$

Thus, the NETEVAL decomposition must classify traffic into four parcels for adequate performance estimation: originating end-link parcels, terminating end-link parcels, switch-to-switch parcels, and switch-to-final-destination parcels. Switch-to-final-destination parcels are used only for service points that do not have unique terminating switches. To compute characteristics for switch-to-switch and switch-to-final-destination parcels, a segment of the network must be defined that contains all links used by these parcels. This subnetwork consists of all inter-machine trunk groups and those end-links contained in the route lists of the switch-to-final-destination parcels. Such a collection of links is called an extended trunk network in this paper. With the extended trunk network and its associated switch-to-switch and switch-to-final-destination parcel means and variances, the appropriate parcel characteristics can be computed. Similarly, the end-link subnetwork and end-link parcels permit the calculation of end-link parcel characteristics. Section 4.6 describes how parcel characteristics can be aggregated to form the desired end-to-end characteristics.

4.4 Link analysis

Both the end-link and switch-to-switch and switch-to-final-destination analyses are based on offering effective-offered loads to a single trunk group and computing associated link characteristics that are a function of overflow parcel means and variances. Whenever alternate routing is involved in a network segment, the Katz algorithm is used on the appropriate section. (Even in a queuing environment this procedure is valid except that a new algorithm to compute link parameters is used.) FRL analysis is included in NETEVAL by stratifying parcel loads by FRL and associating different route lists with each FRL grouping.

Effective-offered loads are computed for each link, from which link parameters such as blocking and overflow variance are estimated. The effective-offered loads associated with a link must take into account interactions with the rest of the network. Traffic blocking both prior and subsequent to a parcel being offered to a link reduces the effective-offered traffic load to the link. However, network setup times, ring times, queue delays, and retrial attempts increase effective-offered link loads. Even though the end-link and extended trunk network segments are analyzed separately in NETEVAL, the effective-offered load equa-

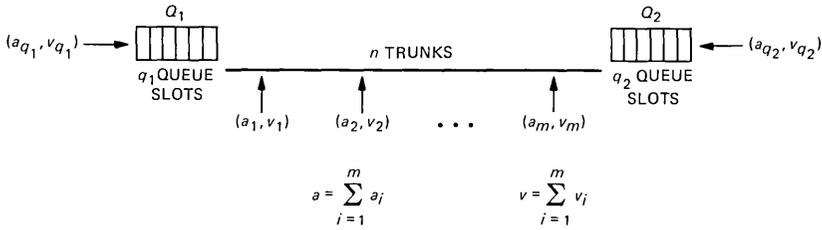


Fig. 2—Link configuration.

tions are designed to reflect the above network characteristics and thus account for the interactions between the network segments. These interactions are updated successively in the form of updated parcels in steps iv and vi as summarized in Section 4.2.

The following subsections describe the computation of link parameters both with and without queues.

4.4.1 Unqueued case

Figure 2 displays a general link configuration with link-offered parcels (a_i, v_i) , $i = 1$ to m . a_i is the mean of the i th parcel and v_i is its variance. [If queues are present, the queue parcels are (a_{q_i}, v_{q_i}) .] The equivalent random method can be used to compute overflow means and variances for aggregate traffic (a, v) . To apportion the aggregate overflow moments to different parcels we use the Katz parcel-splitting approach. If b_i is the i th parcel blocking, then

$$b_i = b_o[1 + k(z_i - z_o)], \quad (3)$$

where,

- b_o = aggregate link blocking,
- z_o = peakedness of aggregate traffic,
- $z_i = v_i/a_i$, and
- k = modification factor.

Several heuristic formulas can be used for k . Katz² expresses k as a function of z_o and b_o . We have selected the heuristic formula from the Defense Communications Engineering Center

$$k = ce^{tn+sb_o}, \quad (4)$$

where $c = 2.249z_o^{-2.82}$, $t = 0.0528z_o^{-4.163}$, $s = 5.456z_o^{-2.025}$, and n is the number of trunks.

Once b_i is known, the parcel overflow and carried mean estimates are $a_i b_i$ and $a_i(1 - b_i)$, respectively. The overflow and carried variance

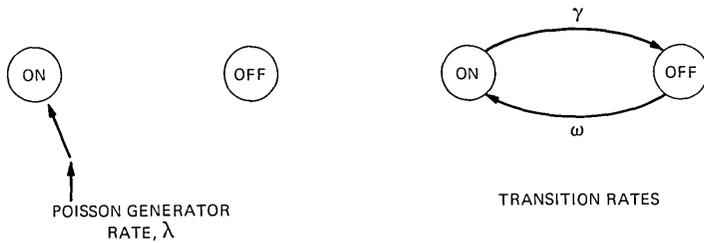


Fig. 3—The IPP process.

are approximated by $z_o^* a_i b_i$ and $v_i - z_o^* a_i b_i$, where z_o^* is the overall overflow peakedness. The overflow variance estimate results in equal peakedness to all parcels.

4.4.2 Queued case

To derive the overflow mean and variance for a parcel offered to a link in the presence of queues, another approach is needed. We approximate the behavior of a link containing queues with a two-dimensional Markov chain for each peaked parcel and a one-dimensional Markov chain for each smooth parcel. The parcel is modeled based on its mean and variance. If the parcel is peaked (peakedness greater than one), an Interrupted Poisson Process (IPP) is used. Smooth parcels (peakedness less than one), which can occur when traffic carried on previous groups is offered to a trunk group, are modeled as a Poisson process with parameters adjusted to better reflect the small peakedness.

4.4.2.1 Interrupted Poisson processes. An IPP, also called switched Poisson, is a Poisson process with rate λ , which alternately for some period of time shuts off all arrivals from the Poisson process, and lets the arrivals go through for another period of time. Both time periods are exponentially distributed, independent of each other and any previous time periods. IPP is used to model traffic with peakedness of one or greater. We say that the switch is ON when the arrivals go through, and otherwise the switch is OFF. The IPP process is shown in Fig. 3. A Poisson process is an IPP with the switch constantly in the ON position. Given (a_i, v_i) , an IPP with an expected on time γ^{-1} , an expected off time ω^{-1} , and a Poisson rate λ (when the IPP is in the ON state) can be obtained from the relationships

$$a_i = \lambda \frac{\omega}{\omega + \lambda} \quad \text{and} \quad (5)$$

$$v_i - a_i + a_i^2 = a_i \lambda \frac{\omega + 1}{\omega + \lambda + 1}, \quad (6)$$

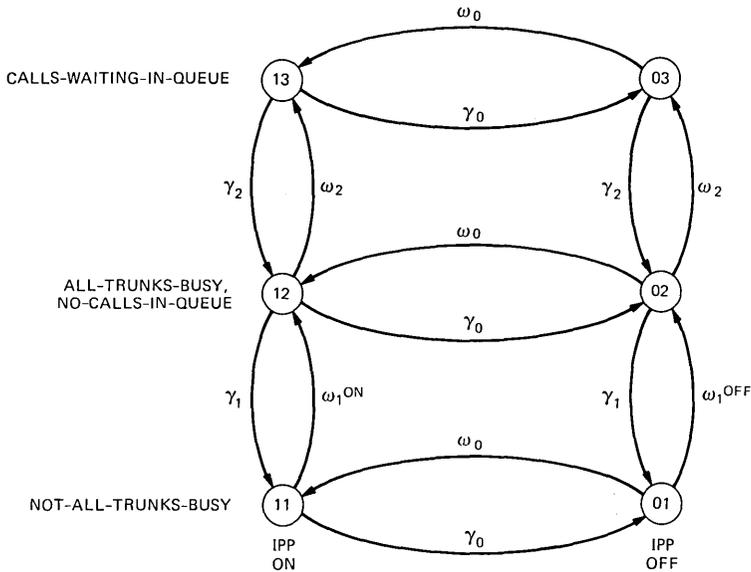


Fig. 4—Markov chain for peaked parcel input.

where λ is arbitrarily set to the larger of a_i and v_i . $v_i - a_i + a_i^2$ is the second factorial moment of the IPP.

This completely specifies the IPP point process. For further properties of the IPP, see Kuczura.³

4.4.2.2 Peaked parcel analysis. The IPP ON and OFF states form one dimension of the Markov chain used in the queued case. The other dimension consists of three superstates:

- (i) Not-all-trunks-busy
- (ii) All-trunks-busy, but no-calls-in-queue
- (iii) Some-calls-waiting-in-queue.

Note that in case (iii), necessarily all trunks are busy, and that the system moves from state 1 only to state 2, from state 2 either to state 1 or to state 3, and from state 3 only to state 2. No transition between states 1 and 3 is possible.

Thus, the Markov chain space is (X, Y) , $X = 0, 1$, $Y = 1, 2, 3$, where X is the IPP status of a parcel offered to a link and Y is one of the three superstates discussed above. Figure 4 displays the flows among the states.

Modeling the complex behavior of the trunk group with queues by compressing all states into superstates and assuming constant transition rates between them is a crucial assumption. The simplified state space is not Markovian any more, and the rates between the states are generally nonconstant. However, for first-moment calculations, such

as the mean of the overflow stream, the replacement of the nonconstant transition rate by the average rate gives exact results. The calculation of the rates is given below. Some of the rates are themselves approximations for the average rate.

Since traffic is offered to the queues only when all trunks are busy in the group, we will consider the queue-offered traffic "conditional on all trunks busy," and use the conditional mean and variance of these streams when computing link characteristics. If (a_q, v_q) is the unconditional queue mean and variance, and p is the probability that all trunks are busy, then the mean and variance conditional on all-trunks-busy is given by

$$a_{cq} = a_q/p, \quad (7)$$

$$v_{cq} = \frac{v_q}{p} - \frac{a_q^2}{p} \left(\frac{1}{p} - 1 \right). \quad (8)$$

These equations are based on the assumption that if the random variable X is the number of busy servers in an infinite trunk group receiving traffic from the queue stream, then $EX^i = pEX_c^i$, $i = 1, 2$.

4.4.2.3 Transition rates. To compute transition rates ω_1^{ON} and ω_1^{OFF} , we need to define ω_1 as the average frequency of transition from not-all-trunks-busy states to all-trunks-busy states. Since this rate does not depend on the presence of queues, we may calculate ω_1 on a trunk group without queues. Let b_T be the time congestion of a trunk group of size n , offered traffic with mean a and variance v . Further, let T_1 and T_2 be the expected sojourn times of the trunk-group system (without queues) in states not-all-trunks-busy and all-trunks-busy, respectively. In that case

$$b_T = \frac{T_2}{T_1 + T_2}. \quad (9)$$

Since $T_2 = n^{-1}$ (assuming unit average holding time) for known b_T , one may find T_1 , and hence $\omega_1 = 1/T_1$, immediately. The time congestion b_T is equal to the blocking, b_i , that would be experienced by a Poisson parcel offered to the trunk group. This is given by the Katz parcel splitting formula

$$b_i = b_o[1 + k(z_i - z_o)], \quad (10)$$

where z_i is equal to 1.

Now that ω_1 is known we can solve for ω_1^{ON} and ω_1^{OFF} . We require that in the long run the number of transitions from states $[(1, 1), (0, 1)]$ to $[(1, 2), (0, 2)]$ in Fig. 4 is ω_1 per unit of time. Since movement among these states does not involve transition rates in or out of states $[(1, 3), (0, 3)]$, it is more convenient to consider a smaller Markov chain

(X^*, Y^*) , $X^* = 0, 1$, $Y^* = 1, 2$, which is a subset of the chain shown in Fig. 4. We can then express the relationship between ω_1 , ω_1^{OFF} , and ω_1^{ON} as

$$\omega_1 = \frac{\rho_{01}}{\rho_{01} + \rho_{11}} \omega_1^{\text{OFF}} + \frac{\rho_{11}}{\rho_{01} + \rho_{11}} \omega_1^{\text{ON}}, \quad (11)$$

where ρ_{ij} is the steady-state probability distribution for the Markov chain (X^*, Y^*) .

We still require an additional constraint to make ω_1^{OFF} and ω_1^{ON} unique. A reasonable additional assumption is that in the absence of queuing, the parcel blocking is the same as that obtained from the parcel splitting formula. Thus, we require

$$b_i = \frac{\rho_{12}}{\rho_{11} + \rho_{12}}, \quad (12)$$

where b_i is obtained from the Katz parcel splitting formula. The ρ_{ij} are themselves expressed in terms of ω_1^{OFF} and ω_1^{ON} . However, it can be shown that the equations above lead to an equation no worse than quadratic.

To compute the transition rate γ_1 , assume the system is in an all-trunks-busy state, with no-calls-in-queue. γ_1 is the rate with which trunks become available. Since we express time in multiples of an average holding time, $\gamma_1 = n$ (the number of trunks). Further, under the assumption that the holding times are all independent, exponentially distributed variables, the transition rate γ_1 is a constant, independent of the time spent in the all-trunks-busy state, with no-calls-in-queue.

To compute ω_2 assume the system is in the state with all-trunks-busy and no-calls-in-queue. Then the frequency with which the state some-calls-waiting-in-queue is entered is the input rate into the queues. Since the queue-input rates a_{cq_1} and a_{cq_2} are given conditionally, it is clear that

$$\omega_2 = a_{cq_1} + a_{cq_2}, \quad (13)$$

where ω_2 is a constant, rather than the average rate when the queue input streams are both Poisson. If the input streams are peaked, they may be represented as interrupted Poisson processes, and a more precise, but more elaborate calculation for ω_2 may be carried out, taking account of the variances of the queue-input streams. This precision was not considered necessary.

Note that in all of the above, it is possible that one or both queues are non-existent. The mean a_{cq} should be set equal to 0 for non-existent queues.

γ_2 , the rate from the state calls-in-queue to the state with all-trunks-

busy and no-calls-in-queue, is the inverse of the mean length of time the system spends in the state with some-calls-in-queue. What we therefore need to calculate is the average duration from the time the first call enters one of the queues until the first subsequent time both queues are empty. We denote this mean duration by Q and call it "the queues busy period." One then has

$$\gamma_2 = Q^{-1}. \quad (14)$$

The calculation of Q is appreciably simpler for links with one queue than for links with two queues. In either case, one models the input stream(s) into the queue(s) as IPP(s), and then finds the simultaneous ergodic distribution of the number of waiting calls and the IPP input state. This ergodic distribution is a conditional distribution, given that all trunks are busy. The input streams are taken to be conditional on this event.

Lastly, ω_o and γ_o are the on and off rates for an IPP parcel offered to a trunk group. Thus, eqs. (5) and (6) can be used to compute ω_o and γ_o .

4.4.2.4 Computation of link characteristics. The overflow mean, overflow variance, and carried variance of a parcel (a_i, v_i) are computed by adding another dimension to the state space, which is the number of calls in an infinite trunk group that receives parcel overflow traffic when the system is in an all-trunks-busy state. Moment-generating functions can be used to compute the desired quantities. (Details are provided in Appendix B.) The overall link blocking is then the ratio of the sum of the parcel overflow means to the sum of the parcel offered loads.

Each of the parcels offered to the queues gives rise to two resultant parcels, namely the overflow parcel and the carried parcel. The queue overflow parcel consists of those calls finding all queue slots occupied, and the carried parcel contains all other calls, since we assume that there are no abandonments or time-outs.

As a result of assumed independence of the parcels offered to the group, the parcels offered to the two queues are independent during the time intervals that all trunks are busy. This means that waiting times and other variables associated with a two-queue trunk group may be estimated through simple Markov chain calculations.

Queue blocking and delay are computed conditional on all-trunks-busy. A Markov chain (j_1, j_2, i_1, i_2), where the j 's are the number in each queue and the i 's represent the on/off conditions of IPP queue streams, is used to compute these quantities. (Appendix C explains the algorithms used to compute the steady-state probabilities of the chain.) For example, if e_{j_1, j_2, i_1, i_2} is the ergodic distribution of the chain and (γ_1, ω_1) represents the on and off rates of the Q_1 arrival stream, then

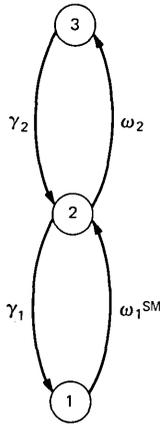


Fig. 5—Markov chain for smooth parcel input.

the Q_1 blocking probability is

$$b_{q_1} = \left(1 + \frac{\gamma_1}{\omega_1} \right) \sum_j (e_{q_1,j10} + e_{q_1,j11}). \quad (15)$$

Since a customer can arrive at Q_1 only when the IPP is ON, eq. (15) is expressed only if the queue arrival stream turned on.

4.4.2.5 Smooth traffic parcel analysis. A smooth traffic parcel is offered to a link when v_i is less than a_i . A somewhat simplistic approach is used to compute the overflow stream characteristic for a smooth parcel. A one-dimensional Markov chain is formed, assuming the parcel is Poisson (see Fig. 5). All rates, except for ω_1^{SM} (used in place of ω_1^{ON} and ω_1^{OFF}), are the same as in the case of the peaked traffic analysis. ω_1^{SM} is computed in the same manner as ω_1 , but with z_i set to the parcel peakedness.

Overflow quantities can be calculated using the same techniques applied to peaked parcels. The corresponding variance calculation, however, leads to too large a variance of the overflow, because the input into the trunk-group characterization switch is taken to be a Poisson stream with mean and variance (a_i, a_i) , instead of the given smooth stream with mean and variance (a_i, v_i) . To correct for this effect, we reduce the calculated variance by multiplying it by the peakedness (less than one), of the offered parcel. The estimated overflow variance is $z_i v_i^{o*}$, where v_i^{o*} is the overflow variance based on Poisson input. The mean overflow is given by

$$\alpha_i \frac{\omega_1^{SM}(\omega_2 + \gamma_2)}{(\omega_1^{SM} + \gamma_1)(\omega_2 + \gamma_2) - \gamma_1 \omega_2}. \quad (16)$$

4.5 Convergence criteria

The convergence criteria for step *viii* of the NETEVAL algorithm (Section 4.2) assumes convergence has occurred if the absolute difference of each end-link, switch-to-switch, and switch-to-final-destination blocking probability in successive iterations is within a specified limit.

4.6 Point-to-point characteristics

Once the network has been decomposed into the extended trunk network and the end-link network, and the components have been analyzed, aggregation techniques are required to compute point-to-point characteristics. This is done by associating two or three traffic parcels, depending on the routing patterns, with each point-to-point pair. If *i-to-j* traffic is part of a switch-to-final-destination parcel when offered to the extended trunk network, then only two parcels are associated with the point pair: the originating parcel at point *i* offered to the end-link network segment, and the switch-to-final-destination parcel in which it is contained. If the *i-to-j* traffic is part of a switch-to-switch parcel in the extended network, then three traffic parcels are associated with the point pair: the originating parcel, the switch-to-switch parcel, and the terminating parcel in which *i-to-j* traffic is contained. Point-to-point characteristics are functions of the corresponding parcel characteristics.

If *i-to-j* traffic is part of a switch-to-switch parcel when traversing the extended trunk network, the *i-to-j* blocking b_{ij} is

$$b_{ij} = 1 - (1 - B_{i,org}^I)(1 - B^{IJ})(1 - B_{j,term}^J). \quad (17)$$

If *i-to-j* traffic, however, is contained in a switch-to-final-destination parcel,

$$b_{ij} = 1 - (1 - B_{i,org}^I)(1 - B^{IJ}). \quad (18)$$

The *i-to-j* expected queue delay, W_{ij} , is a weighted average of the expected queue delay on trunks and end-links. When *i-to-j* traffic is contained in a switch-to-switch parcel,

$$W_{ij} = (1 - B_{i,org}^I)W^{IJ} + (1 - B_{i,org}^I)(1 - B^{IJ})W_j^J. \quad (19)$$

The trunk delay, W^{IJ} , and terminating link delay, W_j^J , are weighted by the probability that *i-to-j* traffic will be offered to a link on which it is queue-eligible. When *i-to-j* traffic is contained in a switch-to-final-destination parcel,

$$W_{ij} = (1 - B_{i,org}^I)W^{IJ}, \quad (20)$$

where W^{IJ} is the expected queue delay for traffic originating at switch *I* destined for point *j*. $1 - B_{i,org}^I$ is the probability that the traffic seizes an end-link and arrives at switch *I*.

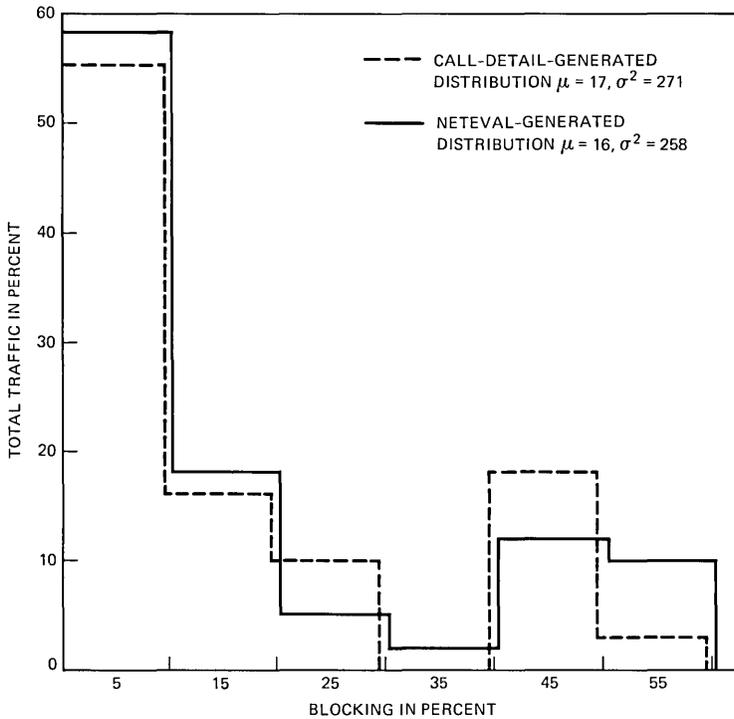


Fig. 6—End-to-end blocking distribution—Network A.

The i -to- j queue delay probabilities, PD_{ij} , are computed analogous to W_{ij} . When i -to- j traffic is contained in a switch-to-switch parcel when traversing the extended trunk network,

$$PD_{ij} = (1 - B_{i,org}^I)PD^{IJ} + (1 - B_{i,org}^I)(1 - B^{IJ} - PD^{IJ})PD_j^J, \quad (21)$$

where $(1 - B_{i,org}^I)(1 - B^{IJ} - PD^{IJ})PD_j^J$, is the probability of traversing the trunk portion of the network without queue delay but incurring queue delay on the terminating end-link.

If i -to- j traffic is contained in a switch-to-final-destination parcel,

$$PD_{ij}^J = (1 - B_{i,org}^I)PD_j^J. \quad (22)$$

Conditional expected queue delays are given by W_{ij}/PD_{ij} , for $PD_{ij} > 0$. Note that by creating an artificial access-line group with no blocking or delay, the above equations remain valid for ETS networks with traffic originating at a tandem switch.

V. APPLICATIONS

EPSCS and ETS networks generate call-detail records. These records contain information on individual calls in the network from which it is

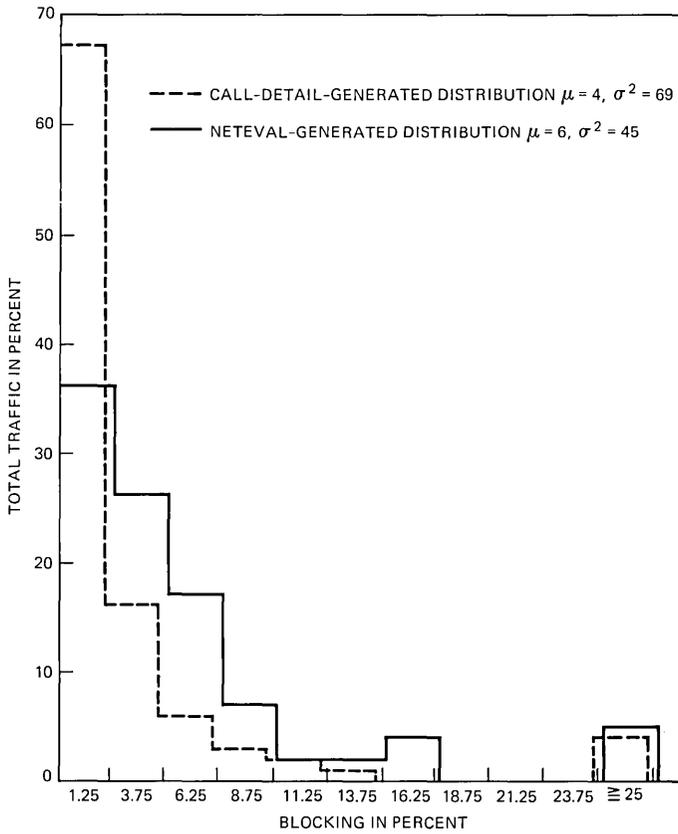


Fig. 7—End-to-end blocking distribution—Network B.

possible to estimate the end-to-end blocking distribution in existing networks. This information can be compared with an ENADS NETEVAL analysis of the same network. Figures 6 and 7 display both call detail and NETEVAL-derived blocking distributions for two existing networks.

The means and variances in Fig. 6 are very close. The variances in Fig. 7 are not as similar as their corresponding means. This is primarily because the midpoint of the last cell in the call-detail-derived distribution is 43 percent, compared with 30 percent in the NETEVAL distribution. Note, however, that the overall shape of the distributions within Figs. 6 and 7 is similar. This suggests that the modeling techniques in our evaluation algorithm provide a reasonable prediction of end-to-end performance.

VI. SUMMARY

Algorithms for computing end-to-end and link characteristics in

private networks have been described in this paper. The algorithms generalize the Katz procedure to provide end-to-end characteristics for large-scale networks and also model private network features such as queuing and FRLS. These techniques have been incorporated into the ENADS Service Evaluator. NETEVAL performance predictions have compared well with actual measurements for several networks. The ENADS NETEVAL module is now routinely used with the companion NETSYN module by AT&T Long Lines and Operating Company private network administrators for network design and evaluation.

VII. ACKNOWLEDGMENTS

The authors acknowledge the preliminary evaluation work of J. K. Andreozzi and the end-to-end blocking distributions provided by W. E. Relyea.

REFERENCES

1. A. L. Kalro, "Traffic Network Design of New Private Switched Services," Telecommun. Networks Planning Symp., Paris, France, 1980.
2. S. S. Katz, "Statistical Performance Analysis of a Switched Communications Network," paper presented at the 5th ITC Congress in New York, 1967.
3. A. Kuczura, "The Interrupted Poisson Process as an Overflow Process," B.S.T.J., 52, No. 3 (March 1973), pp. 437-48.

APPENDIX A

Notations

Much of the notation used in Section IV is summarized below for the reader's convenience.

Parcel Characteristics

Blocking

b_{ij} = Point i to point j blocking.

$B_{i,org}^I$ = Originating end-link blocking for service point i homed on switch I .

$B_{j,term}^J$ = Terminating end-link blocking for service point j homed on switch J .

B^{IJ} = Switch I to switch J blocking.

B^{Ij} = Switch I to final destination point j blocking.

Delays

W_{ij} = Point i to point j queue delay.

W_j^J = Terminating end-link queue delay for service point j homed on switch J .

W^{IJ} = Switch I to switch J queue delay.

W^{Ij} = Switch I to final destination point j queue delay.

PD_{ij} = Point i to point j delay probability.

PD_j^J = Terminating end-link delay probability for service point j homed on switch J .

PD^{IJ} = Switch I to switch J delay probability.

PD^J = Switch I to final destination point j delay probability.

Link Parameters

(a_i, v_i) = Mean and variance of i th parcel offered to trunk group.

z_i = Peakedness of i th parcel offered to trunk group.

(a_{q_i}, v_{q_i}) = Mean and variance of load offered to queue.

(a_{cq_i}, v_{cq_i}) = Mean and variance of load offered to queue conditioned on all trunks busy.

b_i = Blocking of i th parcel offered to trunk group.

(a, v) = Aggregate trunk-group load mean and variance.

z_o = Trunk-group peakedness of offered load.

z_o^* = Trunk-group overflow peakedness.

b_o = Trunk-group blocking.

Markov Transition Rates in Trunk-Group Queuing Model

γ_o = Interrupted Poisson Process off rate.

ω_o = Interrupted Poisson Process on rate.

ω_1^{ON} = Not-all-trunks-busy to all-trunks-busy, no-calls-in-queue flow rate, given Interrupted Poisson Process (IPP) on.

ω_2 = All-trunks-busy, no-calls-in-queue to calls-waiting-in-queue flow rate.

ω_1^{OFF} = Not-all-trunks-busy to all-trunks-busy, no-calls-in-queue flow rate, given IPP off.

γ_2 = Calls-waiting-in-queue to all-trunks-busy, no-calls-in-queue flow rate.

γ_1 = All-trunks-busy, no-calls-in-queue to not-all-trunks-busy flow rate.

APPENDIX B

Overflow Mean and Variance of Peaked Parcel

This appendix gives the details of the calculation of the overflow mean and variance of a parcel, modeled as an IPP, offered to a trunk group, which in turn is modeled by three states and the four transition rates among them. Let the variables λ , γ_o , γ_1 , γ_2 , ω_o , ω_1^{ON} , ω_1^{OFF} and ω_2 be as in Section 4.4.2.

The system to be treated, including the infinite trunk group, is a continuous-time Markov chain, with state space

$$N = \{(k, i, j): k = 0, 1, \dots; i = 0, 1; j = 1, 2, 3\},$$

where k is the number of calls in the infinite trunk group, i is the status (ON = 1, OFF = 0) of the IPP, and j is the state of the trunk group, (not-all-trunks-busy, all-trunks-busy, but no-calls-in-queue, some-

calls-waiting-in queue). Let S_1 denote the status of the IPP and S_2 the status of the trunk group. Designating by p_{kij} the steady-state probability that the system is in state (k, i, j) , we define the generating functions

$$\pi_{ij}(s) = \sum_{k=0}^{\infty} p_{kij} s^k, \quad \text{for } 0 \leq s \leq 1, \quad i = 0, 1, \quad j = 1, 2, 3.$$

The mean number of calls in the infinite trunk group is

$$\alpha^* = \sum_{i,j} \sum_k k p_{kij} = \sum_{i,j} \frac{d}{ds} \pi_{ij}(1) = 1' \frac{d}{ds} \pi(1),$$

where $1'$ stands for the six element row vector of ones and

$$\pi(s) = \begin{pmatrix} \pi_{01}(s) \\ \pi_{02}(s) \\ \pi_{03}(s) \\ \pi_{11}(s) \\ \pi_{12}(s) \\ \pi_{13}(s) \end{pmatrix}.$$

Similarly, the second factorial moment is

$$EX(X-1) = \sum_{i,j} \sum_k k(k-1) p_{kij} = \sum_{i,j} \frac{d^2}{ds^2} \pi_{ij}(1) = 1' \frac{d^2}{ds^2} \pi(1),$$

from which one easily finds the variance. Writing out the steady-state equations for p_{kij} and summing appropriately, one finds that the $\pi_{ij}(s)$ satisfy a system of equations, which we will define in vector form. We define

$$A = \begin{pmatrix} -(\omega_0 + \omega_1^{\text{OFF}}) & \gamma_1 & 0 & \gamma_0 & 0 & 0 \\ \omega_1^{\text{OFF}} & (\omega_0 + \gamma_1 + \omega_2) & \gamma_2 & 0 & \gamma_0 & 0 \\ 0 & \omega_2 & -(\omega_0 + \gamma_2) & 0 & 0 & \gamma_0 \\ \omega_0 & 0 & 0 & -(\gamma_0 + \omega_1^{\text{ON}}) & \gamma_1 & 0 \\ 0 & \omega_0 & 0 & \omega_1^{\text{ON}} & -(\gamma_0 + \gamma_1 + \omega_2) & \gamma_2 \\ 0 & 0 & \omega_0 & 0 & \omega_2 & -(\gamma_0 + \gamma_2) \end{pmatrix},$$

and the diagonal matrix B by

$$B = \text{diag}(0, 0, 0, 0, \lambda, \lambda),$$

where the elements of the diagonal are given inside the parentheses.

The vector $\pi(s)$ satisfies the equation

$$(s-1) \frac{d}{ds} \pi(s) = A\pi(s) + (s-1)B\pi(s). \quad (23)$$

Setting $s = 1$ in (23), one obtains $A\pi(1) = 0$, from which the marginal probabilities $\pi_{ij}(1) = P\{S_1 = i, S_2 = j\}$ are found. Taking derivatives

w.r.t. s in (23) and setting $s = 1$ gives

$$(I - A) \frac{d}{ds} \pi(1) = B\pi(1).$$

Since $1'A = 0$, one gets

$$1' \frac{d}{ds} \pi(1) = 1'(I - A) \frac{d}{ds} \pi(1) = 1'B\pi(1), \quad (24)$$

and hence the mean a^* of the overflow may be found by solving for $A\pi(1) = 0$, with the side condition $1'\pi(1) = 1$. Differentiating eq. (23) twice, one gets

$$2 \frac{d^2}{ds^2} \pi(s) + (s - 1) \frac{d^3}{ds^3} \pi(s) = A \frac{d^2}{ds^2} \pi(s) + 2B \frac{d}{ds} \pi(s) + (s - 1)B \frac{d^2}{ds^2} \pi(s).$$

Setting $s = 1$ and solving for $\frac{d^2}{ds^2} \pi(1)$ gives

$$\frac{d^2}{ds^2} \pi(1) = 2(2I - A)^{-1}B \frac{d}{ds} \pi(1).$$

Again, since $1'(2I - A) = 2 \cdot 1'$, one finds

$$1' \frac{d^2}{ds^2} \pi(1) = 1'B \frac{d}{ds} \pi(1),$$

or, substituting (24),

$$1' \frac{d^2}{ds^2} \pi(1) = 1'B(I - A)^{-1}B\pi(1).$$

Thus, by inverting the 6×6 matrix $I - A$, the variance of the overflow may be found.

The carried traffic mean may be obtained directly from the overflow mean via

$$a_c = a - a^*,$$

where a_c is the carried mean, a the offered traffic mean, and a^* the overflow mean. It can be shown that a_c may alternatively be found via the procedure sketched above, if we change the B matrix to

$$B = \text{diag}(0, 0, 0, \lambda, 0, 0).$$

This change represents the fact that carried traffic leaves when the IPP switch is ON, and the trunk group characterization is in the not-all-trunks-busy state.

With this change in B , the variance of the carried traffic may be calculated in exactly the same way as the overflow variance.

APPENDIX C

Queue Analysis

This appendix gives some details of the queue analysis. Assume a trunk group with n trunks, q_1 queue slots on one side, q_2 on the other, and input streams into the queues given, conditionally on all-trunks-busy, by their IPP parameters $(\lambda_i, \omega_i, \gamma_i)$, $i = 1, 2$. The system is a continuous time Markov chain. This appendix gives a description and a fixed ordering of the states of the system in Section C.1 and algorithms for the steady-state probabilities in Section C.2.

C.1 Ordering of the states

The elements of the different vectors and matrices to be used in this appendix are characterized by a four-dimensional index, $(i, j, o1, o2)$, where each possible index is a state of the system. Index i is the number of calls waiting in the first queue, j the number of calls waiting in the second queue, and $o1$ and $o2$ are status bits, indicating whether the IPP representing the input stream is currently in the OFF state ($o = 0$) or in the ON state ($o = 1$). Status bit $o1$ refers to the first queue, $o2$ to the second. Index i runs between 0 and q_1 , j between 0 and q_2 . All bounds are inclusive.

A fixed order of enumeration is adhered to. There is a macro ordering, for the (i, j) part of the state designations, and within that ordering, the status bits have a fixed order.

The gross order is given by

$$\begin{array}{cccccc} (0, 0), & (1, 0), & (2, 0), & \cdots & (q_1, 0), \\ (0, 1), & (1, 1), & \cdot & \cdots & (q_1, 1), \\ (0, 2), & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ (0, q_2), & (1, q_2), & \cdot & \cdots & (q_1, q_2). \end{array}$$

Within these index elements there are four states indicating the status bits, which are ordered like 00, 01, 10, 11. Thus, the overall order is indicated by 0000, 0001, 0010, 0011, 1000, 1001, 1010, 1011, 2000, etc.

C.2 Steady-State Equations

The steady-state equations form a set of $4(q_1 + 1)(q_2 + 1)$ equations in as many unknowns, with the normalizing side conditions that the probabilities add to 1. In the ordering given above, the structure of these equations is that of a tri-diagonal blocked matrix equation, as follows. Let e_j be the vector of ergodic probabilities associated with the states having j calls in Q_2 , $j = 0, 1, \dots, q_2$, i.e.,

$$e_j = (e_{0j00}, e_{0j01}, e_{0j10}, e_{0j11}, e_{1j00}, \dots, e_{q_1j11}).$$

Then the e_j satisfies

$$e_0 A_e + e_1 C = 0, \quad (25a)$$

$$e_{j-1} B + e_j A_g + e_{j+1} C = 0, \quad j = 1, \dots, q_2 - 1 \quad (25b)$$

$$e_{q_2-1} B + e_{q_2} A_f = 0. \quad (25c)$$

Writing out these equations in matrix form will show the tri-diagonal block structure.

The matrices A_e , A_f , A_g , B , and C are all square of order $4(q_1 + 1)$; in addition, B and C are both diagonal matrices. Below we discuss each of these matrices, and exhibit their structures.

(i) C is associated with a call arriving into Q_2 : the system moves from state (i, j, o_1, o_2) to state $(i, j + 1, o_1, o_2)$, with transition rate λ_2 if and only if $o_2 = 1$. Therefore, $C = \text{diag}(0, \lambda_2, 0, \lambda_2, \dots, 0, \lambda_2)$.

(ii) B signifies the transitions caused by a call waiting in Q_2 being served. The transition is from (i, j, o_1, o_2) to $(i, j - 1, o_1, o_2)$, with rate n if $i = 0$, and with rate n_2 (where n_2 is the effective number of circuits available to calls from Q_2 when Q_1 and Q_2 are both non-empty) if $i \neq 0$. Specifically, $B = \text{diag}(n, n, n, n, n_2, n_2, \dots, n_2)$.

(iii) A_e , A_f , and A_g are all associated with transitions for which the number of calls in Q_2 is constant, this constant being equal to 0 for A_e (empty), q_2 for A_f (full), and general for A_g . These matrices have a tri-diagonal block structure themselves, with blocks of size 4×4 , with the super- and sub-diagonal blocks associated with calls entering Q_1 and leaving Q_1 , respectively, and with the diagonal blocks modeling the transitions between the input status states $(o_1, o_2) = (0, 0), (0, 1), (1, 0), (1, 1)$. A_e , A_f and A_g have slight differences owing to the fact that

(a) The diagonal element for any state contains the negative sum of the transition rates out of that state.

(b) The rate at which calls leave Q_1 depends on whether Q_2 is empty.

Next we show that we may take advantage of the block structure of the $4(q_1 + 1)(q_2 + 1)$ set of steady-state equations, and reduce it to a set of size $4(q_1 + 1)$ to be solved via brute-force matrix inversion methods. Consider equations (25). Since B is a nonsingular diagonal matrix, we may write

$$e_{q_2-1} = -e_{q_2} A_f B^{-1}$$

$$e_{q_2-2} = -e_{q_2} (C B^{-1} + A_f B^{-1} A_g B^{-1}).$$

In general, if $e_j = e_{q_2} D_j$ for $j = k + 1, k + 2, \dots, q_2$, one obtains D_k from equation (25b)

$$D_k = -D_{k+1} A_g b^{-1} - D_{k+2} C B^{-1} \quad (k = 0, 1, \dots, q_2 - 1),$$

with $D_{q_2} = I$, the unit matrix, and $D_{q_2+1} = 0$.

Assume then, that we have expressed e_0 in terms of e_{q_2} . Now we use the top equation to solve for e_{q_2} and get

$$e_{q_2}(D_0A_e + D_1C) = 0. \quad (26)$$

The requirement that all probabilities sum to 1 can likewise be expressed in terms of e_{q_2} only:

$$\sum e_j 1 = \sum e_{q_2} D_j 1 = 1.$$

Thus, we sum the rowsums of the D_j , obtaining

$$x = \sum D_j 1.$$

We now replace one column of $D_0A_e + D_1C$ by the vector x , and the corresponding entry in the 0 vector in (26) by 1, and solve this modified set of equations. We obtain e_{q_2} , and from it may find all ergodic probabilities by postmultiplying e_{q_2} by the D_j 's.

If the number of queue slots for $Q1$ is larger than for $Q2$ ($q_1 > q_2$), we may choose to relabel $Q1$ and $Q2$, and reduce the problem size by performing the recursive procedure described above over the larger of q_1 and q_2 . Thus, we may reduce the size of the matrix to be inverted from $4(q_1 + 1)(q_2 + 1)$ to $4[1 + \min(q_1, q_2)]$.

A Description of the Bell Laboratories Scanned Acoustic Microscope

By P. SULEWSKI,* D. J. BISHOP, and R. C. DYNES

(Manuscript received February 4, 1982)

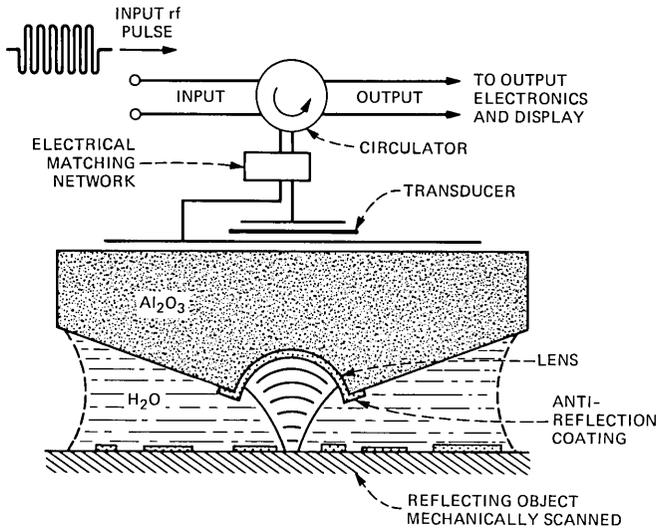
We have developed a working scanning reflection acoustic microscope. In this paper we describe its construction and operation and also present preliminary acoustic micrographs and compare them with equivalent optical and electron micrographs. Our instrument operates at room temperature using 2-GHz acoustic radiation with water as a coupling medium, and has a resolution of approximately 1 μm . We also discuss improvements to be made in future instruments with liquid helium as a coupling medium.

I. INTRODUCTION

Scanning reflection acoustic microscopy uses the amplitude of reflected high-frequency sound waves as a contrast mechanism to generate micrographs with submicrometer resolution. This method was first reported in 1974 by R. A. Lemons and C. F. Quate.¹ Since that time, much progress has been made in the field, including a steady improvement in the resolution of the instrument.² To date, acoustic microscopy has been used to study integrated circuits, biological specimens, and various materials with much success.²

Although acoustic micrographs appear quite similar to their optical counterparts, the source of the acoustic contrast lies in the mechanical properties of the sample. Hence, acoustic micrographs provide information that is fundamentally different from that of optical micrographs. Since the acoustic reflectivity at a surface is a strong function of the layering structure beneath,³ much effort has been directed toward nondestructive analysis and characterization of integrated circuit defects using acoustic microscopy. In our research, it is hoped that variations in stress and crystal orientation near dislocations in crystals

* Princeton University.



r_0 = RADIUS OF LENS = $35 \mu\text{m}$
 f = FOCAL LENGTH OF LENS = $1.13 r_0$
 R = RADIUS OF LENS APERTURE = $0.7 r_0$

Fig. 1—The lens configuration of the scanning acoustic microscope.

will provide sufficient acoustic contrast to image dislocations, providing a powerful tool for nondestructive analysis.

Here we present micrographs from some preliminary surface studies, indicating the performance of our acoustic microscope, along with an explanation of its basic operation and components.

II. PRINCIPLES OF OPERATION

Our micrographs are generated using the acoustic lens pictured in Fig. 1. While using such a simple lens optically would result in severe distortion owing to spherical aberration, such effects are negligible in the acoustic case, since the distortion is proportional to the square of the ratio of the velocity of sound in the two media, sapphire and water. Since $V_{\text{Al}_2\text{O}_3} = 11.1 \text{ km/s}$ and $V_{\text{H}_2\text{O}} = 1.5 \text{ km/s}$, such aberration effects can be ignored.¹ A coupling medium, in our case water, is used since air cannot transmit 2-GHz acoustic waves with acceptable losses.

The lens focuses the acoustic waves to a point, whose actual width is diffraction limited to of the order of λ . With the sample at this focal point, the acoustic waves are reflected back through the lens. The amplitude of the reflected waves provides the contrast for each resolution element in the final picture. Light and dark areas arise from variations in the acoustic reflectivity of the sample. Since the focal

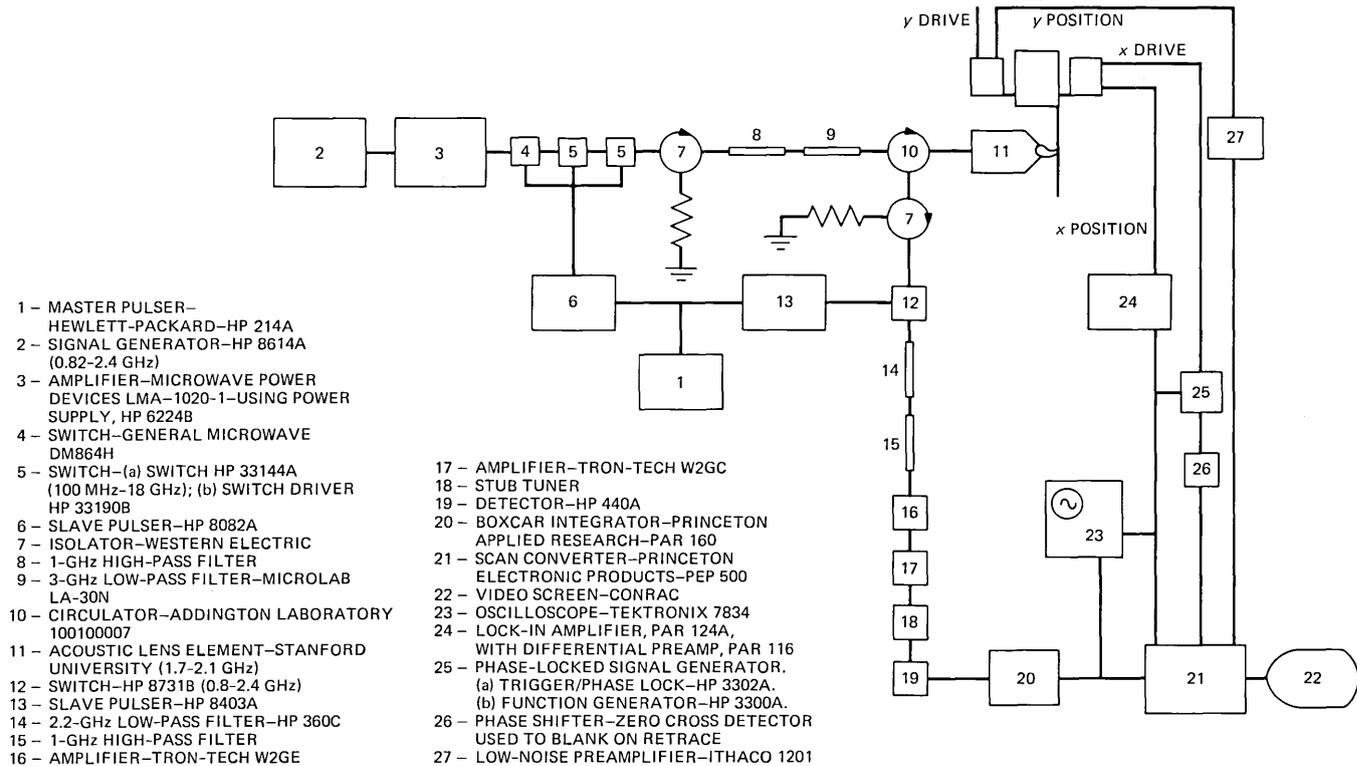


Fig. 2—Block diagram of the electronics for the microscope.

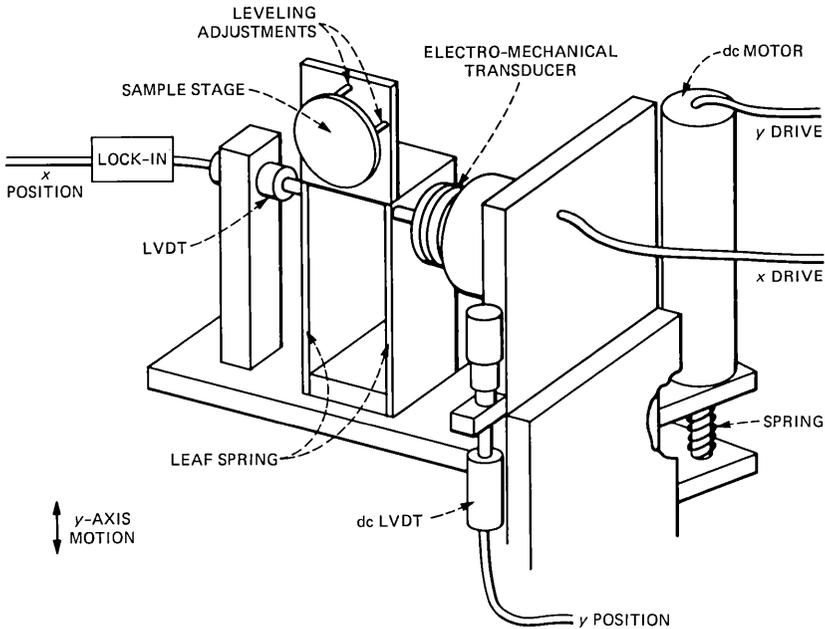


Fig. 3—The mechanical stage for the scanning acoustic microscope.

point has finite width, the resolution of the resulting micrograph is limited to $\approx 0.75\lambda$, as calculated by Wickramasinghe.⁴

To create a picture, the sample is scanned in a raster pattern, line by line. This is accomplished by vibrating the sample at a frequency of 37 Hz and an amplitude of $\pm 250 \mu\text{m}$ in the x direction, while driving the sample through 500 μm in the y direction in 10 to 15 seconds, completely scanning the sample. The amplitude of the reflected acoustic wave is then used together with x and y positioning information given by linear variable differential transformers (LVDT) to create a video image by a scan converter.

The acoustic waves are generated by a piezoelectric transducer on the back of the lens element, which converts microwaves at a frequency near 2 GHz to acoustic waves of the same frequency.

A master pulser, (1) in Fig. 2, governs the timing of the electronics that control and process the microwave power. The master pulser generates trigger pulses at a rate of 500 kHz. The microwaves are generated (2) and then amplified (3) to a power level of 1 watt. Three PIN-diode switches (4, 5) connected in series chop the continuous waves into 18-ns pulses at a rate set by the master pulser. An intermediate slave pulser (6) drives the switches. The use of pulsed waves allows us to isolate the desired information pulse from spurious reflec-

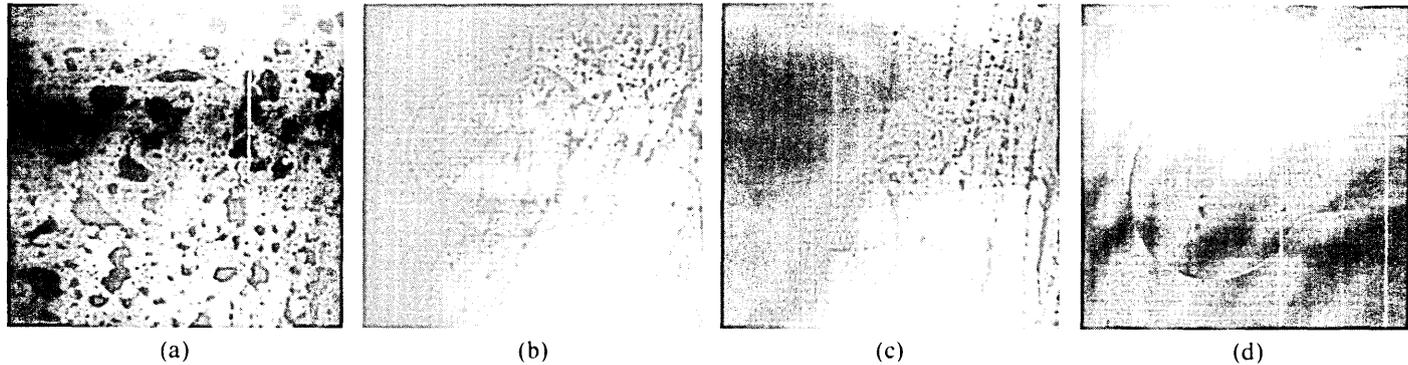


Fig. 4—Acoustic micrograph of gold mask on epitaxially grown LED (layers of InP, InGaAsP, InP). (a) Before cleaning. (b) After swabbing with cotton swab soaked in ethanol. (c) After swabbing with alcohol. (d) With lens off focus slightly (or at different focus).

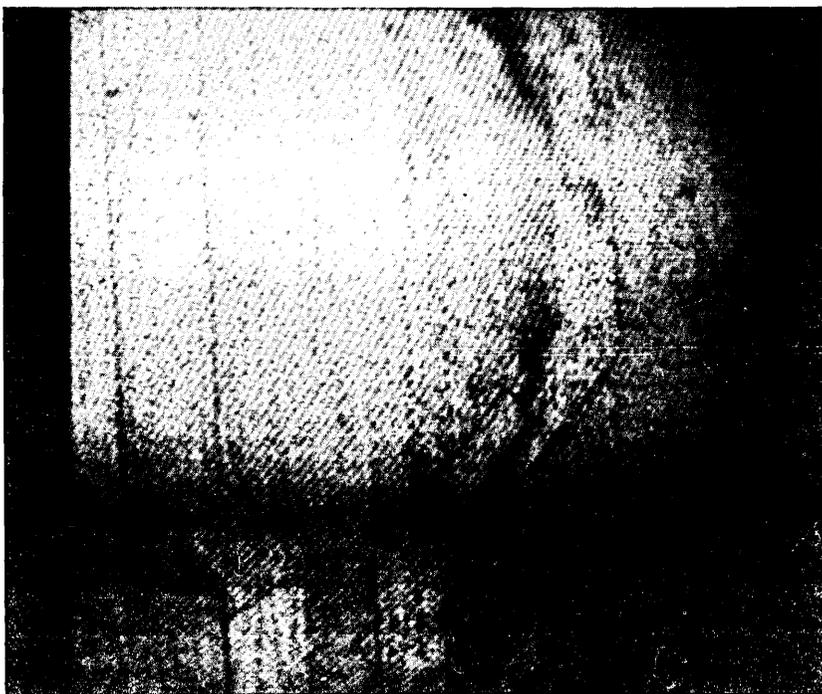


Fig. 5—Gold stripes on silicon.

tions. Three switches are used in series to increase the on-off ratio. Since the power of the desired reflected signal pulse is generally 80 dB below that of the initial pulse, any leakage of microwaves during the off state would produce undesirable interference patterns.

The microwave pulse then passes through an isolator (7) which protects the system from spurious multiple reflections. A 1-GHz high-pass filter (8) and a 3-GHz low-pass filter (9) eliminate low-frequency and high-frequency noise, respectively, from the switches. A circulator (10) directs the pulse into the lens element (11), and then directs the reflected signal pulse into the rest of the circuit.

At the back of the lens, the pulse is electrically matched to the piezoelectric transducer as in Fig. 1. Even with this matching network, there is a large reflection at this point. The transducer, composed of a gold-ZnO₂-gold sandwich, evaporated onto the back of the sapphire, converts the microwave pulse into an acoustic pulse of the same frequency, which then travels through the sapphire to the lens. At a simple lens-water interface, most of the incident radiation would be reflected since there is a large impedance mismatch, as $Z_{\text{H}_2\text{O}} = 1.5 \times 10^5 \text{ gm/cm}^2\text{-s}$ and $Z_{\text{Al}_2\text{O}_3} = 44.0 \times 10^5 \text{ gm/cm}^2\text{-s}$. A one-quarter-wavelength impedance-matching layer of borosilicate glass at this interface

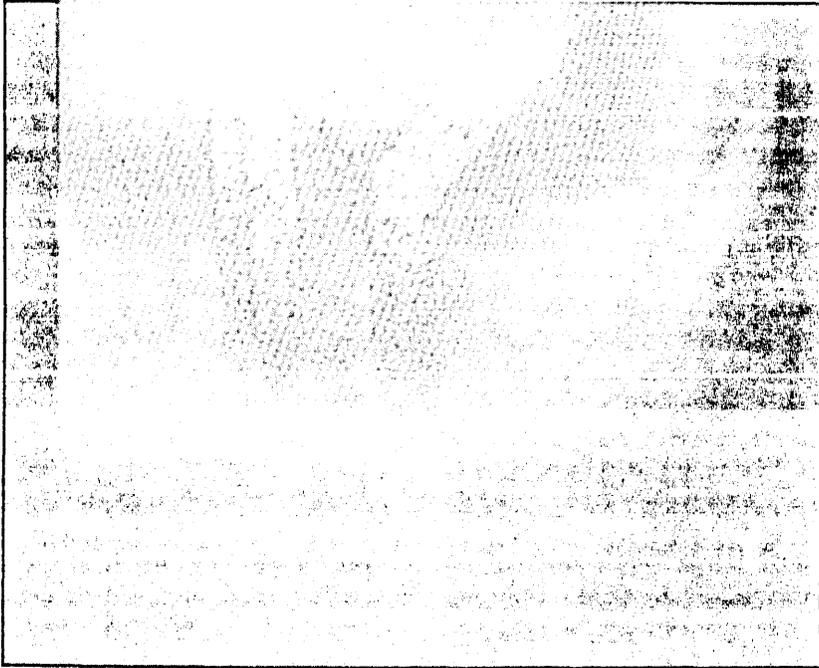


Fig. 6—A different region of gold stripes on silicon pattern shown in Fig. 5.

significantly increases transmission although it does not completely eliminate this reflection.⁵ The transmitted and focused acoustic pulse then reflects off of the sample and back through the lens to the transducer, where it is reconverted into microwaves. Losses in the water amount to ≈ 70 dB. The amplitude of the reflected pulse depends upon the acoustic properties of the sample; hence, the contrast in the resulting micrograph reflects variations in acoustic properties of the sample. The pulse passes through an isolator (7) and a final PIN-diode switch (12), which acts as a window to protect the amplifiers from the large spurious reflections already mentioned. Two filters (14, 15) reduce the noise, especially low frequencies from the switch. The amplifiers (16, 17) give a total of 62-dB amplification. A stub tuner (18) matches the incoming transmission line to the crystal-diode detector (19). The detector rectifies the pulse of microwaves, and a boxcar integrator (20) is used to improve the signal-to-noise ratio. The scan converter (21) uses the x and y positioning information together with the signal pulse amplitude to construct a video picture, which may then be photographed from the screen (22).

The scan electronics provide the means to scan the sample in a



Fig. 7—Gold stripes with scratches on silicon.

raster pattern. The resolution of the LVDTs used must be submicrometer if the resolution in the microscope itself is to be submicrometer.

The sample stage (shown in Fig. 3) is mounted on leaf springs so that vibration of the sample is possible only in the x direction. The frequency of this vibration is held at resonance, approximately 37 Hz, using a phase-locked signal generator (25). Translation in the y direction is accomplished by an optical translation stage driven by a dc motor. Most of the weight of the stage is supported by a spring to relieve the motor, so that it can be operated without stalling.

The x position of the sample is given by an ac LVDT connected to a lock-in amplifier (24). The y position is given by a dc LVDT whose output voltage is proportional to the displacement. It was found that the 10-kHz modulation frequency used to operate the LVDTs leaked into both x and y position signals, causing image degradation and poor resolution. Filtering these signals resulted in much improved images.

Since the theoretical resolution is proportional to λ , increasing the frequency should improve the resolution. However, for water and many other liquids, the acoustic attenuation is proportional to f^2 , so

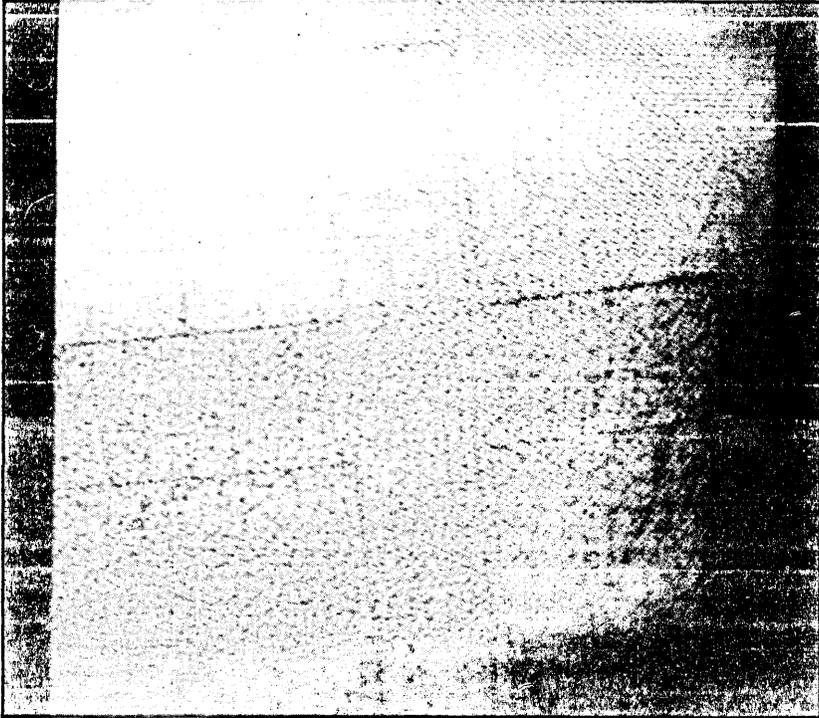


Fig. 8—Gold cross on silicon.

increasing the frequency dramatically increases the power losses within the coupling medium.⁵ To preserve the signal-to-noise ratio, 1.8 to 2.0 GHz is the present optimal frequency range. Other coupling media continue to be investigated.

We use water heated to 60°C, where the acoustic attenuation is lower than at room temperature, resulting in an improved signal-to-noise ratio.

We have found that both samples and lens must be kept very clean if the microscope is to function at all. A very thin film of oil, like that deposited by a fingerprint, can reduce the amplitude of the reflected pulse dramatically. Also any small dust particle, with a diameter of several micrometers, which becomes lodged in the lens will distort and even eliminate the reflected signal pulse.

III. IMAGES

For our initial testing of the microscope, various surface structures were examined, and comparisons performed with optical and electron micrographs.

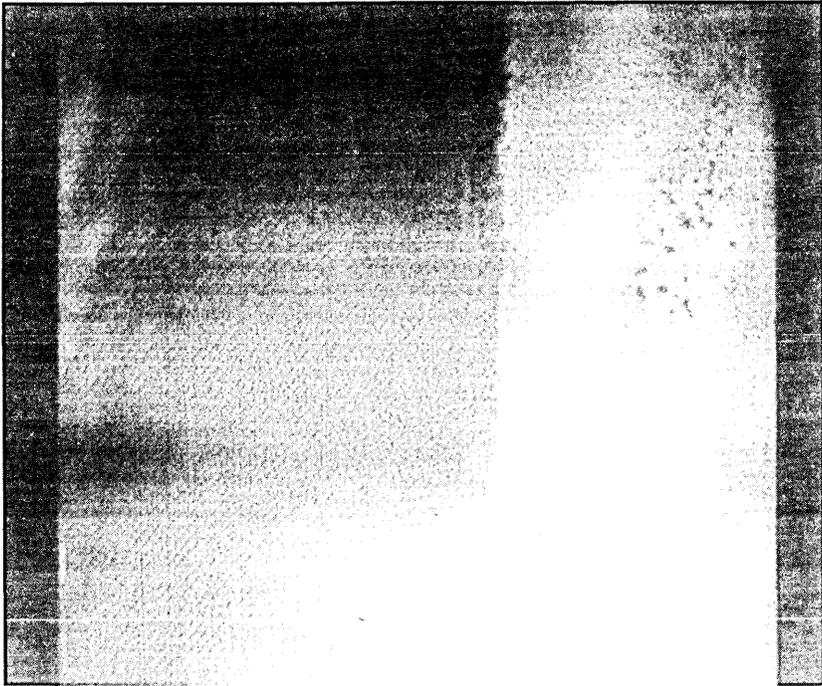


Fig. 9—Bulk-epitaxial transition InGaAsP; left, bulk; right, epitaxial.

The first series of micrographs in Fig. 4 of a gold mask on InGaAsP indicates the importance of surface cleaning. Any oil or grease, for example a fingerprint or any of the vacuum grease used to fix the sample to the stage, shows up quite clearly and obscures the true detail (Fig. 4a). Cleaning the surface with a cotton swab soaked in ethanol significantly reduces the surface structure of the first micrograph (Fig. 4b,c). Fig. 4d presents another method of discerning true structure from mere surface dirt effects. By maintaining the lens slightly out of focus, the effects of surface dirt disappear. Also note the phase shifts at the edges of the gold to InGaAsP transition owing to different acoustic path lengths in the materials. The diagonal lines visible in Fig. 5, 6, and 7 are caused by noise over the signals given by the LVDTs. This effect was eliminated from subsequent micrographs. The noise produced an uncertainty of several micrometers in the sample position, seriously degrading image resolution.

Fig. 8 displays the dramatic improvement in image quality achieved by eliminating the noise over the LVDT signals. The horizontal dark and light bands result when the dc motor alternately slows down or speeds up within the single frame. Near the bottom of the picture,

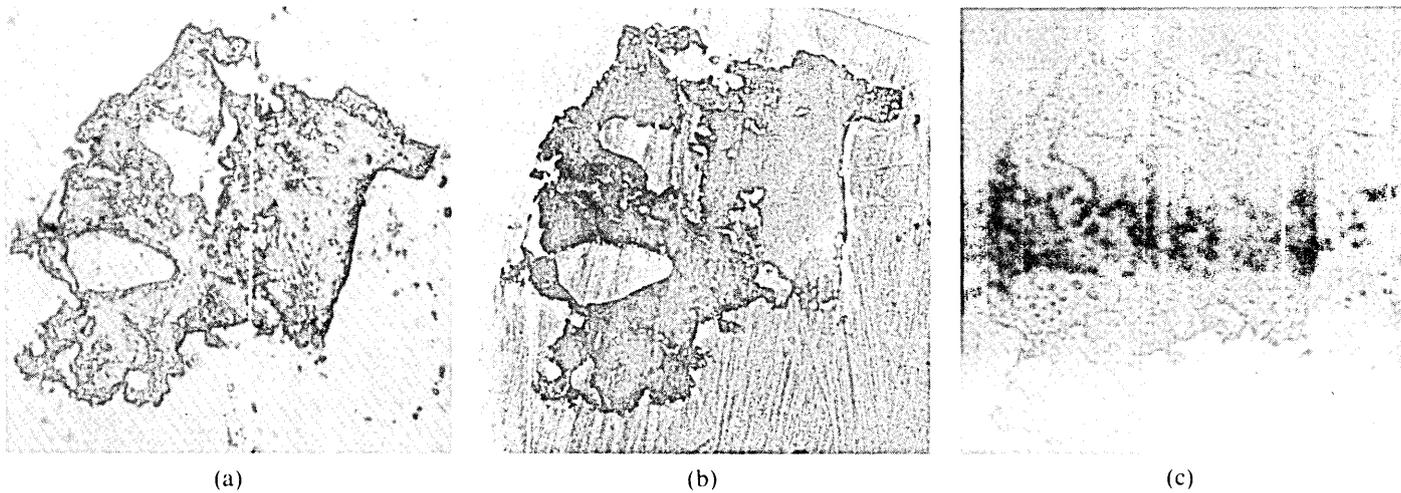
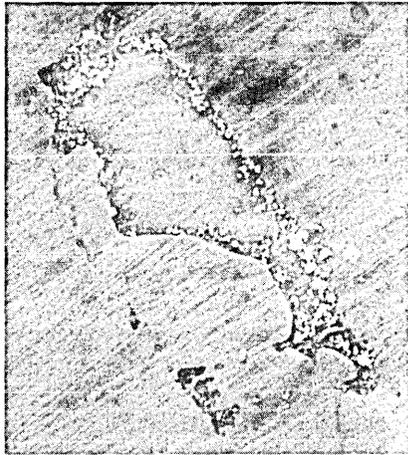


Fig. 10—Indium inclusion in epitaxial InGaAsP. (a) Optical micrograph. (b) Electron micrograph. (c) Acoustic micrograph.



(a)

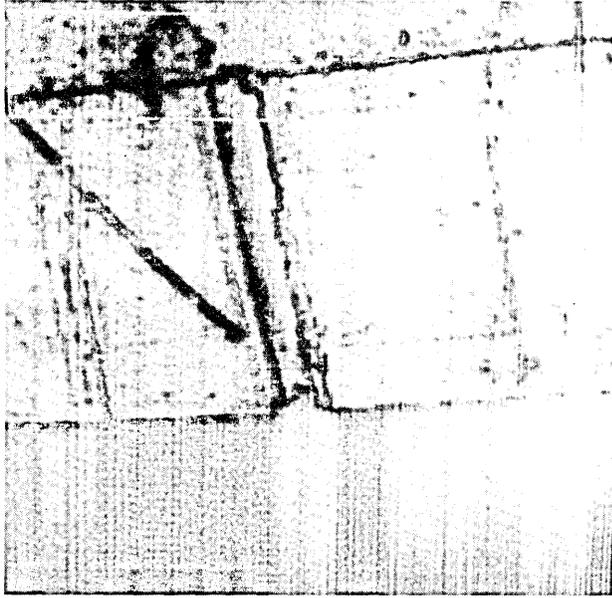


(b)

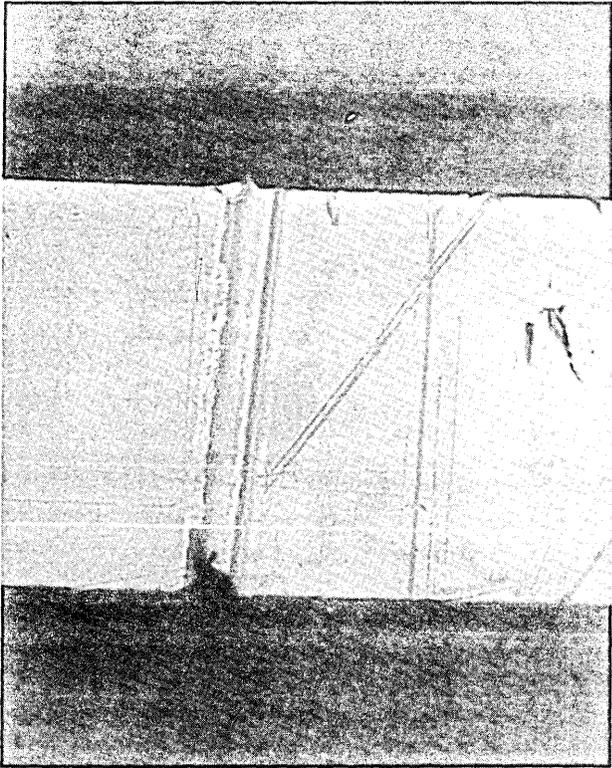


(c)

Fig. 11—Indium inclusion on epitaxial layer. (a) Optical. (b) Electron. (c) Acoustic.



(b)



(a)

Fig. 12—Gold stripe on silicon. (a) Electron. (b) Acoustic.

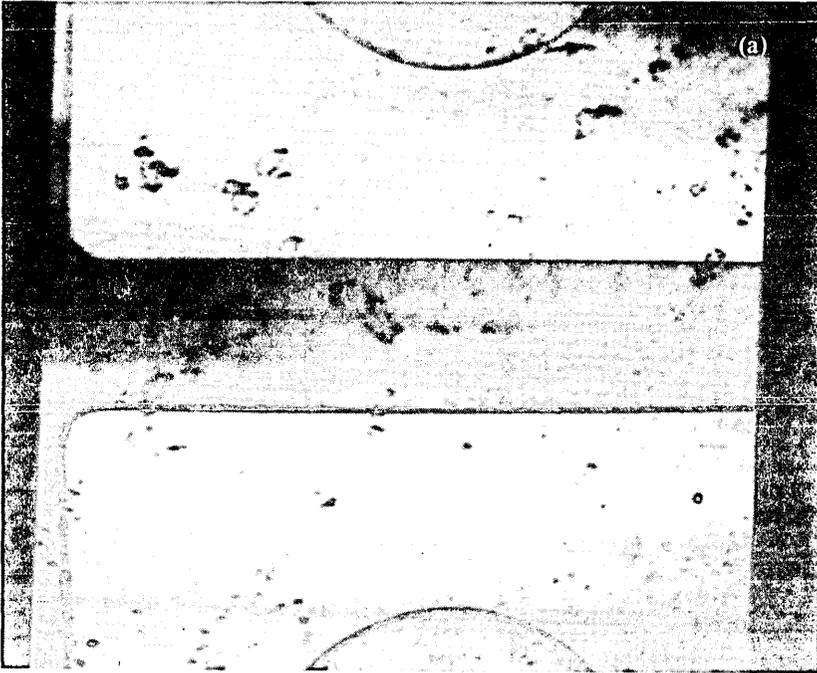


Fig. 13—(a) A sample of InGaAsP is shown with a gold evaporation on the surface. The dark regions are indium inclusions which have formed on the surface. The field of view is $\approx 500 \mu\text{m} \times 500 \mu\text{m}$. (b) A close-up of an inclusion from Fig. 13a at ten times the magnification.

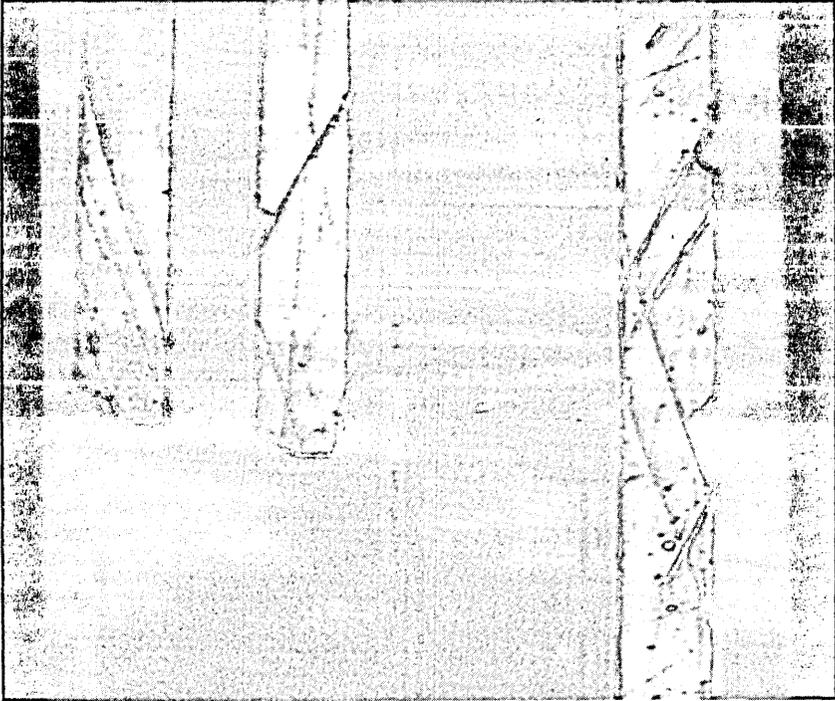


Fig. 14—A more recent micrograph of gold on silicon showing the quality of the acoustic micrographs we have been able to obtain.

single scan lines are visible. This problem has been overcome with a better mechanical drive.

One area of application of the acoustic microscope is the evaluation of subsurface defects within epitaxial layers. Preliminary to such studies, we have examined various samples of epitaxially grown wafers with inclusions visible on the surface. The first picture in the series, Fig. 9, shows a transition from bulk to epitaxial InGaAsP, from left to right.

In Fig. 10, we compare optical, electron, and acoustic micrographs of an indium inclusion in an epitaxial layer of InGaAsP. In the acoustic picture, in the lower right opening in the inclusion, several round spots a few micrometers across appear, which do not appear in either the electron or optical micrographs. These could be subsurface features that only the acoustic microscope can image.

The next series of pictures in Fig. 11 show another indium inclusion on the same sample. Again while more surface features are visible with optical and electron microscopy, there is a round spot 5 μm in diameter near the lower edge of the inclusion in the acoustic micrograph that does not appear in either of the other pictures.

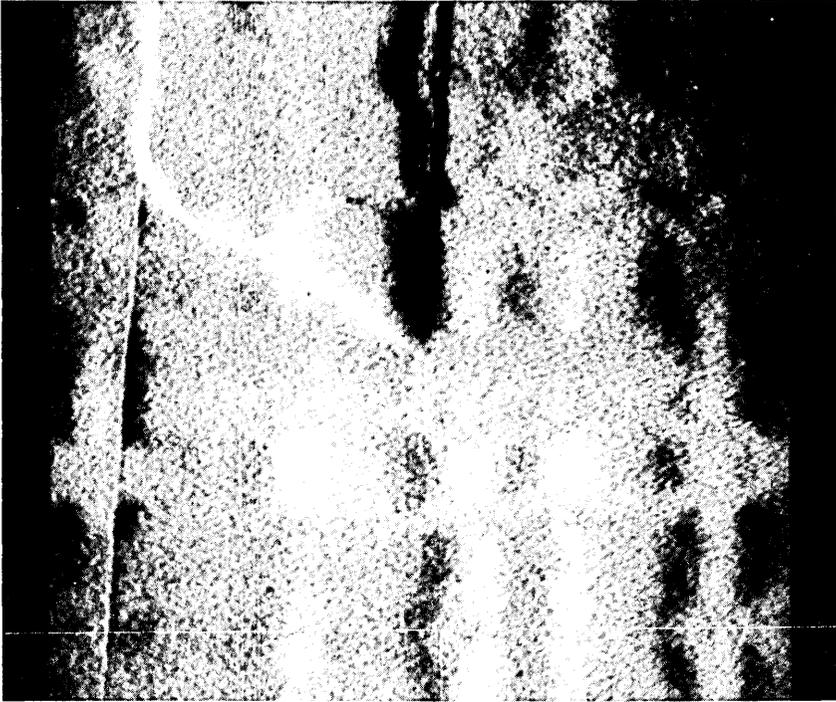


Fig. 15—This micrograph shows an image where the derivative of the acoustic intensity is plotted. This type of photo provides a sense of depth perception to the photos and generates a microscopic contour map.

In Fig. 12, we present a comparison between electron and acoustic micrographs of a 50- μm gold strip evaporated onto silicon. The scratch marks provide a good measure of the resolution of the instrument. The top two horizontal scratches are 5 μm apart, and easily discerned in the acoustic micrograph. The bottom three scratches are each 2.5 μm apart from the other and also easily discernible. However, the two diagonal scratches, separated by 1 μm , are not distinguishable in this micrograph.

In Fig. 13, we show a recent micrograph of an InGaAsP sample with a gold overlayer. The instrument has been optimized for this series of pictures and gives some indications of the quality that it is capable of. The small dark regions are indium inclusions which have formed on top of the sample. Figure 13b, is a micrograph of a single inclusion seen in the center of the upper photo.

In Fig. 14, we show a picture of gold stripes evaporated on a silicon substrate. Finally in Fig. 15, we show the results of a form of derivative microscopy. In that picture, we have displayed the derivative of the

acoustic amplitude. This produces an image which is a micro-contour map of the substrate.

IV. CONCLUSIONS

We have successfully operated a scanning acoustic microscope at 60°C with resolution approaching the diffraction limit using water as a transmitting medium. Our instrument operates at 1.9 GHz with a resolution of $\approx 1 \mu\text{m}$. We have shown some preliminary acoustic micrographs indicating how our present instrument might be used in studies of structural defects in various materials of technological importance.

Work is currently under way to develop an instrument using liquid He⁴ as an acoustic transmitting medium, which will allow us to use much higher acoustic frequencies. Such a device will be the first acoustic microscope that will not have diffraction-limited resolution.

ACKNOWLEDGMENTS

We wish to thank R. A. Lemons, J. E. Heiserman, C. F. Quate, A. K. Chin, and A. Atlar for useful discussion. In addition we would like to thank C. F. Quate for providing the acoustic lens used in these studies.

REFERENCES

1. R. A. Lemons and C. F. Quate, "Acoustic Microscope—Scanning Version," *Appl. Phys. Lett.*, **24**, No. 4 (February 1974), pp. 163-5.
2. See for example R. A. Lemons and C. F. Quate, "Acoustic Microscopy," *Physical Acoustics*, **14**, R. N. Thurston (Editor), Academic Press (to be published).
3. A. Atlar, "An Angular-Spectrum Approach to Contrast in Reflection Acoustic Microscopy," *J. Appl. Phys.*, **49**, No. 10 (October 1978), pp. 5130-9.
4. H. K. Wickramasinghe, "Contrast and Imaging Performance in the Scanning Acoustic Microscope," *J. Appl. Phys.*, **50**, No. 2 (February 1979), pp. 664-72.
5. R. A. Lemons, *Acoustic Microscopy by Mechanical Scanning*, Ph.D. dissertation, Stanford University, 1975.

A Statistical Model of Multipath Fading on a Space-Diversity Radio Channel

By W. D. RUMMLER

(Manuscript received February 22, 1982)

The joint probability of occurrence of frequency-selective fades on a pair of spatially separated receiving antennas is modeled for a typical line-of-sight microwave radio path in the 6-GHz band. The model was developed from observations of the transmission in a 24.2-MHz band during all multipath-fading occurrences in a 30-day period on a 26.4-mile path. By fitting the observations of every scan on both antennas with a simplified three-ray channel modeling function, the joint transmission at each observation is characterized by six parameters, three for each antenna. The joint occurrence of these six parameters is described by simple statistical distribution functions, allowing one to associate with any pair of channel transmission shapes the fraction of a year, or number of seconds in a year, that such a channel state will be encountered. The model represents the frequency selectivity or shape of the fades on the two antennas as statistically independent. Only the average fade levels on the two antennas are statistically related. Either antenna is more likely to experience a fade deeper than the median when selectivity is observed on it or when the other antenna is experiencing deeper fading than the median. The (marginal) statistics of fading on each of the antennas separately, as derived from the diversity model, are essentially the same as those described by a nondiversity statistical fading model, which has been used successfully to predict the multipath outage of digital radio systems. The model developed here will allow performance to be estimated in a diversity configuration.

I. INTRODUCTION

Occurrences of multipath fading limit the performance quality of high-speed digital radio systems operating on line-of-sight microwave radio paths. Extensive field measurement programs have been imple-

mented to evaluate the performance of a number of digital radio systems operating in different configurations under various conditions.¹⁻⁵ These studies have indicated the universal need for some form of dynamic channel equalization, and for space diversity reception, on many paths, to meet performance requirements. While field measurements provide a good means of evaluating the operation of radio systems, they require considerable time, effort, and expense. Furthermore, they suffer from the vagaries of nature in that multipath fading is a randomly occurring phenomenon with variable characteristics from month to month.^{5,6}

To reduce the need for field measurements, a statistical model of multipath fading was developed.^{6,7} This model, used in conjunction with characterization measurements performed on a radio in the laboratory, allows predictions to be made of the system performance under multipath conditions when operating in a nondiversity configuration.⁸⁻¹⁰ The results presented here extend the previous work by providing a statistical model for multipath fading in a space-diversity configuration.

The data base used for modeling was obtained by transmitting a wideband (8-PSK digital radio) signal at 6 GHz over a 26.4-mile path from Atlanta to Palmetto, Georgia. The received power at a number of frequencies in a 24.2-MHz band was measured simultaneously on both a horn reflector and a parabolic dish antenna separated by 30 feet. Spectra were observed at rates up to five times per second during the occurrences of multipath propagation in a 30-day period in August to September, 1977. The received voltages on both the horn and dish at each observation, relative to unfaded or free-space propagation conditions, are represented as a function of frequency by the simplified three-path modeling function that has the form

$$H(j2\pi f) = a[1 - be^{-j2\pi(f-f_0)\tau}]. \quad (1)$$

The diversity channel model provides a joint statistical representation of the occurrence of the parameters of the function (1) as fitted for both the horn and dish.

The choice of a modeling function for representing selective multipath fades over a restricted frequency band is not unique. Such a function needs only to be capable of representing the characteristics of a multipath fade. The parameter statistics will depend strongly on the choice of function. Greenstein and Czekaj¹¹ have used a complex polynomial in frequency to represent multipath fades and have developed statistics for the coefficients of the polynomial for a nondiversity data base. Although other modeling functions have been proposed,¹²⁻¹⁷ none has been successfully represented on a statistical basis. The modeling function (1) used here has the virtues of providing an

excellent representation of the observed multipath fades and of being convenient for synthesis in the laboratory for the stressing of radio systems for performance appraisal. In the present work we show that a further advantage of this function is that the joint statistics in a diversity configuration are well behaved and easily represented.

The statistical channel model is summarized in Section II. The data base used and the fitting of observations with the modeling function are described in Section III. In Section IV we provide the methodology for developing and verifying the statistical model. Concluding remarks are provided in Section V.

II. MODEL SUMMARY

2.1 Modeling function

During multipath fading, the voltage transfer functions of the paths to the horn reflector and to the parabolic dish antenna are modeled by

$$H_H(j2\pi f) = a_H[1 - b_H e^{-j2\pi(f-f_{0H})\tau}] \quad (2)$$

and

$$H_D(j2\pi f) = a_D[1 - b_D e^{-j2\pi(f-f_{0D})\tau}], \quad (3)$$

respectively. These transfer functions are measured relative to the unfaded, or free-space, transfer functions, which are both taken as unity at all frequencies, f . For convenience and for consistency with previous work we fix the delay τ at 6.3 ns. These functions may be interpreted as the responses of channels with direct transmission paths with amplitudes a_H and a_D , and second paths with relative amplitudes b_H and b_D , both respectively. The second path in each case has a relative delay of 6.3 ns, and a phase of $2\pi f_{0H}\tau + \pi$ and $2\pi f_{0D} + \pi$ (independently controllable) at the center frequency of the channel.

A typical plot of the attenuation produced by such modeling functions is shown in Fig. 1. The a and b parameters control the depth and shape of the simulated fades, respectively. The parameters f_{0H} and f_{0D} determine the frequencies of the transmission minima, or notches, of the simulated fades. With a simulated minimum within the channel, the modeling functions can simulate a wide range of levels and notch widths. With a simulated minimum out of the channel band, the modeling functions can generate a wide range of combinations of levels, slopes, and curvatures of the in-band responses.

For convenience we work with the following related parameters: the fade-level parameters (in decibels)

$$A_H = -20 \log a_H \quad (4)$$

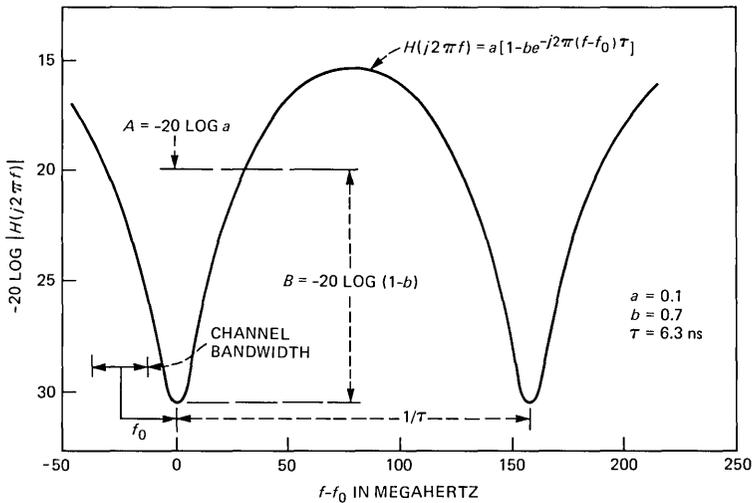


Fig. 1—Attenuation of channel modeling function.

and

$$A_D = -20 \log \alpha_D; \tag{5}$$

the relative notch depth parameters (in decibels),

$$B_H = -20 \log(1 - b_H) \tag{6}$$

and

$$B_D = -20 \log(1 - b_D); \tag{7}$$

and the notch frequency parameters, which we measure in degrees,

$$\phi_H = 360f_0H\tau \tag{8}$$

and

$$\phi_D = 360f_0D\tau. \tag{9}$$

With τ equal to 6.3 ns, one degree in ϕ corresponds to 0.44 MHz. We measure notch frequencies, ϕ , from the center frequency of the channel, so that ϕ covers the range from -180 to $+180$ degrees, corresponding to the 158.4-MHz period of the functions (2) and (3).

2.2 Parameter statistics

The number of seconds in a year that the six parameters ($A_H, A_D, B_H, B_D, \phi_H, \phi_D$) are in a differential element of the six-dimensional parameter space is shown to be given by

$$T(A_H, A_D, B_H, B_D, \phi_H, \phi_D) = T_0 p_{A/B}(A_H, A_D/B_H, B_D) p_{B_H}(B_H) p_{B_D}(B_D) p_H(\phi_H) p_D(\phi_D), \quad (10)$$

where the functions $p(\cdot)$ are all probability density functions. The time scale factor T_0 , under the assumption that events scale with the classical scaling of the incidence of multipath fading,¹⁸ is given by the expression:

$$T_0 = 52800c(f/6)(D/25)^3 \quad (11)$$

where

f is the frequency in GHz,
 D is the path length in miles, and
 c is the terrain factor, varying between 0.25 and 4.

The probability density functions of horn and dish relative notch depths are given, respectively, by

$$p_{B_H}(B_H) = 0.76711(2B_H)(0.10258)e^{-0.10258B_H^2} + 0.23289(0.23281)e^{-0.23281B_H} \quad (12)$$

$$p_{B_D}(B_D) = 0.82295(2B_D)(0.07668)e^{-0.07668B_D^2} + 0.17705(0.21786)e^{-0.21786B_D}. \quad (13)$$

The joint probability density function of A_H and A_D is conditioned on the values of B_H and B_D and is given by

$$p_{A/B}(A_H, A_D | B_H, B_D) = \frac{1}{2\pi\sigma_H\sigma_D\sqrt{1-\rho^2}} \exp \frac{-1}{2(1-\rho^2)} \left[\frac{(A_H - g_H)^2}{\sigma_H^2} - \frac{2\rho(A_H - g_H)(A_D - g_D)}{\sigma_H\sigma_D} + \frac{(A_D - g_D)^2}{\sigma_D^2} \right], \quad (14)$$

where

$$g_H = g_H(B_H) = 23.956(701.11 + B_H^4)/(1320.6 + B_H^4) \quad (15)$$

$$g_D = g_D(B_D) = 27.139(1223.8 + B_D^4)/(2650.9 + B_D^4) \quad (16)$$

and

$$\begin{aligned} \sigma_H &= 6.8268 \\ \sigma_D &= 7.0272 \\ \rho &= 0.64995. \end{aligned} \quad (17)$$

The probability density functions for the horn and dish notch frequencies are given, respectively, by

$$p_H(\phi_H) = \begin{cases} \frac{5}{1080} & |\phi_H| < 90 \\ \frac{1}{1080} & 90 \leq |\phi_H| < 180 \end{cases} \quad (18)$$

and

$$p_D(\phi_D) = \begin{cases} \frac{8}{1620} & |\phi_D| < 90 \\ \frac{1}{1620} & 90 \leq |\phi_D| < 180. \end{cases} \quad (19)$$

2.3 Interpretive discussion

The salient features of the model, which are verified in Section 4, are easily stated. The selective components of multipath fading as seen by the horn and the dish are modeled as independent processes. This means that detailed knowledge of the transmission "shape" present at a given instant on one antenna provides no information concerning the shape that will be present on the other. The only coupling between the fading on the two antennas is provided through the joint conditional A -distribution of (14). The form of this conditional probability density function implies that the fade-level parameters, or fading levels, on the two antennas are related. Deeper fades on one tend to be accompanied by deeper fades on the other. The conditioning on the relative notch depth parameters implies that the fade-level parameters depend on the fade shapes. This is similar to the coupling provided in the nondiversity model^{6,7} in that the fade-level parameter is correlated with relative notch depth, or that deeper fades are more likely to occur when shape is present. For the diversity model we find that the existence of a shapely fade on one antenna is more likely to be accompanied by deeper than average fading on the other.

There are two important limitations of the proposed model. The most obvious limitation is that we have not explicitly represented the variability in fading statistics that would accompany changes in relative spacing in the two receiving antennas. At first blush, we ascribe this to the existence of only one data base for the particular configuration tested; hence, the model is only valid for antenna spacings of thirty feet. Upon closer inspection, one notes that the only coupling parameter that is unique to the diversity configuration is the parameter ρ in the joint conditional A -parameter distribution. (With $\rho = 0$, the model would factor into two independently fading probability models.) One could, in principle, relate ρ to antenna separation by calculating the single-frequency fade statistics of a simulated diversity switch as a function of ρ and compare these results with the known results for various antenna separations.¹⁹ While an extrapolation of the model to

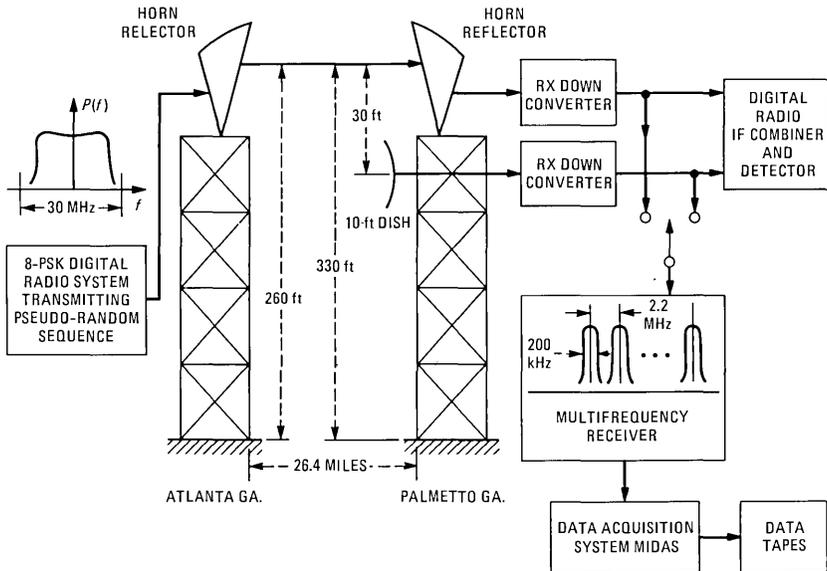


Fig. 2—Experimental configuration for diversity propagation measurements, August 8 to September 6, 1977.

larger separations is straightforward, one would expect that for sufficiently small antenna separations the fade shapes observed on the two antennas would be correlated, leading to a more complicated model.

The other model limitation results from the lack of phase information in the channel probing measurements. We model the channel as a minimum phase channel at all times; that is,²⁰ we choose a solution with $0 \leq b \leq 1$ and assume a minus sign for the $f\tau$ term in the exponential in (2) and (3). This limitation makes it difficult to assess the characteristics of the combined signal for a continuously adaptive space-diversity combining algorithm.

III. PROPAGATION MEASUREMENTS AND THEIR REPRESENTATION

3.1 Description of the propagation experiment

The propagation measurements used in this study were obtained from an experiment conducted on a 26.4-mile path from Atlanta to Palmetto, Georgia, during the period from August 8 to September 6, 1977. Many of the parameters of the experiment are summarized in Fig. 2. The radiated signal source was a general trade 78-Mb/s, 8-PSK digital radio operating at a nominal center frequency of 6034.2 MHz. The signal was received at Palmetto on both a standard horn reflector and a 10-foot diameter parabolic dish located 30 feet below the horn. The spectral energy received by each antenna was measured at 12

frequencies separated by 2.2 MHz and spanning 24.2 MHz. The receiving filter at each of these frequencies had a 200-kHz bandwidth.

During fading activity the received power at each frequency on both antennas was measured either five times a second or once every two seconds, depending on how rapidly the channel was changing. Sampled power, quantized in 1-dB steps, was recorded by the Multiple Input Data Acquisition System (MIDAS), constructed by G. A. Zimmerman. During nonfading periods the power was recorded at a rate of once every thirty seconds. Based on a two-hour measurement period spanning noon on each day, free-space, or nonfaded, received power levels were determined for each frequency.

3.2 Diversity data base

Over the duration of this experiment, multipath fading occurred in fourteen separate time periods. The measurements made in each of these time periods were calibrated and collected into a computer data base for further analysis. The overall data base includes 85,410 scans of both the horn and the dish, and encompasses 44,386 seconds of fading activity.

Although propagation was monitored for approximately one month in the heavy fading season for this path, the observed fading activity was about twice the amount that would be expected in a typical heavy fading month. Figure 3 shows the time-faded statistics for the horn; it shows the number of seconds that power was faded to or below the level specified by the abscissa. (For purposes of analysis, it is assumed that measured or calculated parameters hold a constant value from one observation time instant until the next observation time instant at which the value may make a stepwise change.) The four curves shown represent the power measured at a frequency near the upper end of the frequency band, one near mid-band, and a third near the lower edge of the band. The fourth curve represents the fading of the average power, based upon a wideband measurement of the received signal. We observe that the four curves are virtually coincident down to the 40-dB level, where the rms power characteristically rolls off more rapidly. The coincidence of the curves indicates a good mix of fading events with no dominant events causing an excess of fading activity at any particular frequency in the band.

Also shown in Fig. 3 is the fading activity predicted for this hop in a heavy fading month.¹⁸ The observed fading statistics match the predicted L^2 slope of the predicted curve; however, for this period we have obtained twice as much time at a given level as one would expect in a heavy fading month.

Figure 4 shows the time-faded statistics as observed on the dish antenna. The general observations made for the horn apply here also,

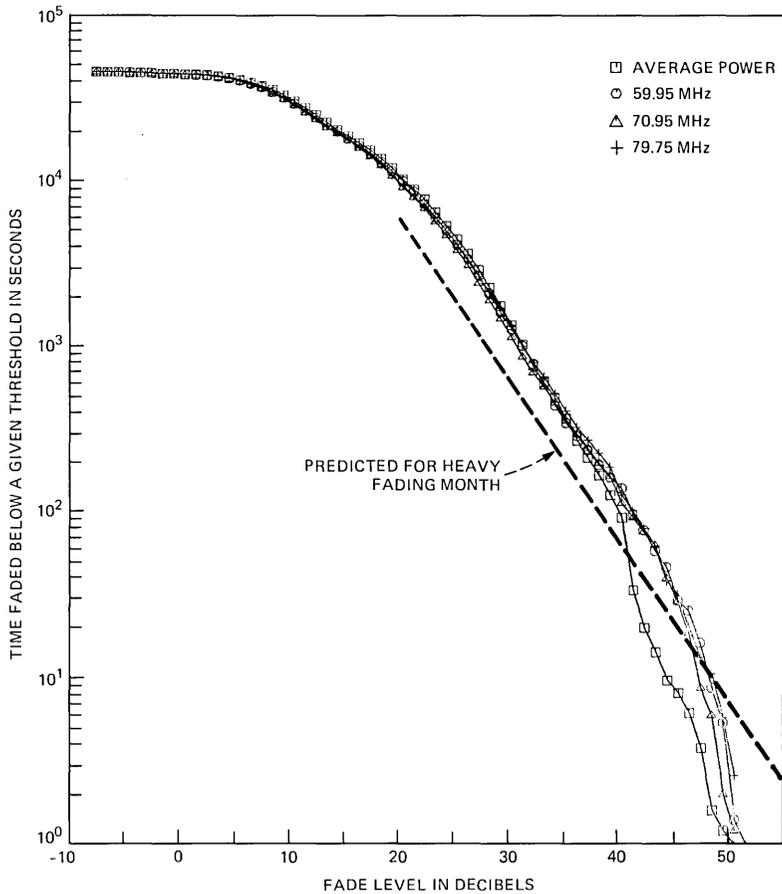


Fig. 3—Time-faded statistics of received power on the horn reflector.

except that we note that the lower frequencies in the band show more fading activity than the upper frequencies at fade levels near 40 dB. The effects of this will become more apparent in Section 4.4. Comparing the dish fading statistics with the predicted fading curve, we would expect the observed fading activity in 2.5 heavy fading months.

As a further consistency check, we consider the in-band power difference (IBPD) statistics, which have also been used for sizing data bases of fading observations.¹⁰ When fading is monitored at a number of frequencies in a band, one can characterize the transmission shape of an observed channel by IBPD, which is the difference, in decibels, between the largest and smallest attenuation of the observed frequencies at a given time. Figure 5 shows the IBPD statistics for the horn and the dish; that is, it shows the number of seconds that the IBPD equaled

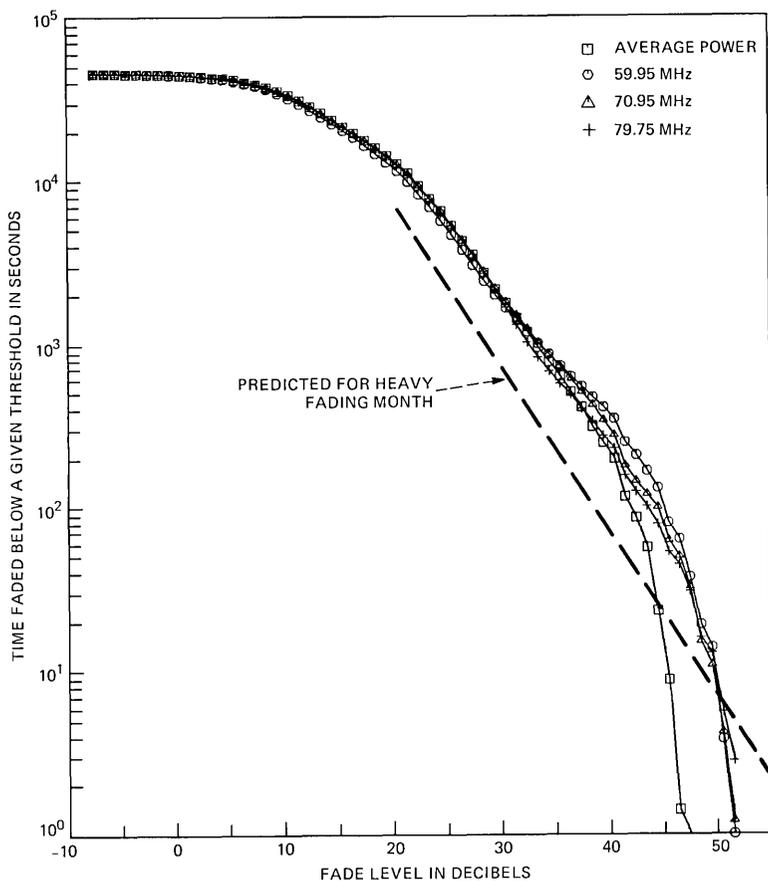


Fig. 4—Time-faded statistics of received power on the dish antenna.

or exceeded the value specified by the abscissa. As a reference curve we also show, in Fig. 5, the IBPD curve (see Ref. 8, Fig. 20) derived from the data base used for the nondiversity model⁷ as scaled to a heavy fading month. The reference IBPD curve was derived from 23 observed frequencies spanning a 25.3-MHz band. Although the IBPD ascribed to a given channel condition depends upon the frequency band spanned by the observations and, to a lesser extent, on the number of frequencies observed, the bandwidth difference is small, and the effects of frequency spacing may be minimized by concentrating on the more modest values of IBPD, 5 to 10 dB. Over this range of IBPD there are 1.8 to 2.3 times as many seconds at a given level for the horn, and 2.3 to 2.7 times as many seconds for the dish. The midpoints of these ranges are very nearly equal to the scaling factors determined previously from the time-faded statistics.

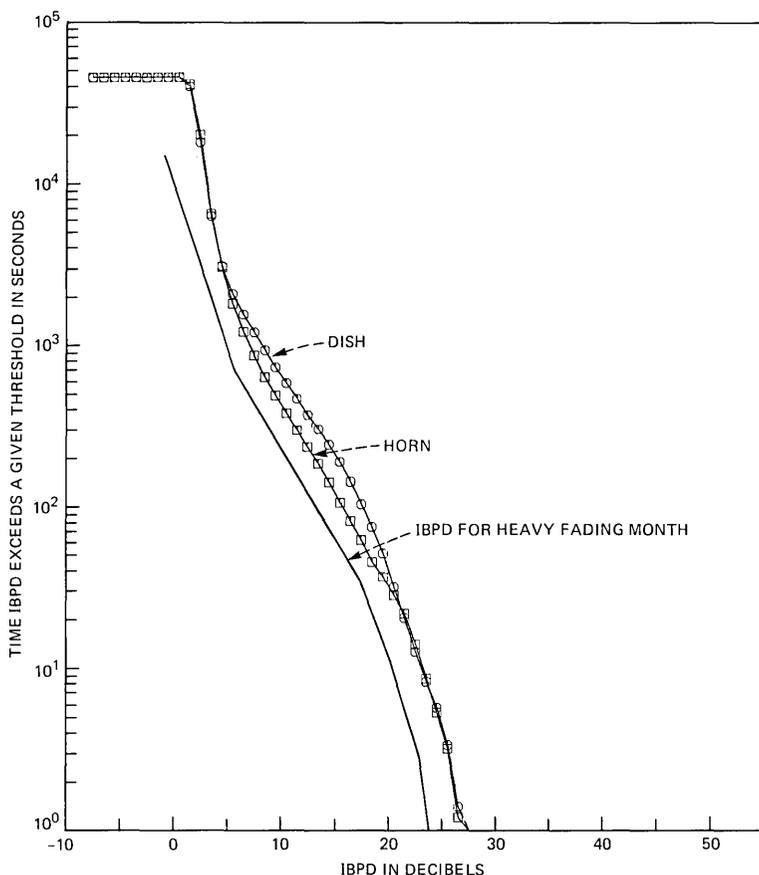


Fig. 5—Time-faded statistics of in-band power difference (IBPD) of the horn and dish.

Splitting the difference between the 2.0 and 2.5 months-of-fading estimates for the horn and dish, respectively, the data base is taken as representing 2.25 months of multipath fading. On the basis that three heavy fading months are equivalent to the fading activity in a year, we take the data base as representing 0.75 of the expected annual multipath-fading activity for this path.

3.3 Representation of spectral measurements

Each scan of the received power levels of the horn and dish is represented by the channel transfer functions (2) and (3), respectively. To obtain the parameters of the functions for each scan, we fit the squared magnitudes of the functions to the received powers. The fitting procedure minimizes the weighted mean squared error between the observed power and the estimated power. The weighting function

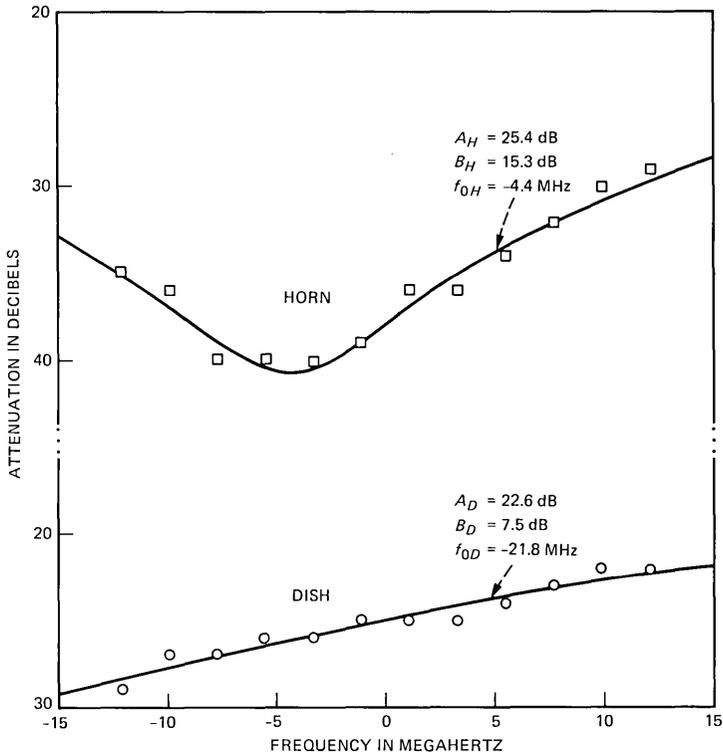


Fig. 6—Representation of fades observed on the horn and dish on August 21, 1977, at 1:28:3.2.

is proportional to the reciprocal of the square of the received power at each frequency. This provides a weighting that is approximately logarithmic, to match the instrumentation errors which are independent and approximately Gaussian on a logarithmic scale. The reader is referred to Ref. 7 for additional details of the fitting procedure.

Figure 6 shows the fits to horn and dish scans observed concurrently. At this particular instant there was a notch present on the horn at a frequency of 4.4 MHz below the band center, and a 6-dB slope present on the dish. The parameter values producing the fitting functions are shown on the plots. The rms error between the observed levels and the values of the fitted function at these frequencies may be taken as a measure of the quality of the representation of the fade. For the horn scan shown, the rms error is 0.68 dB, for the dish 0.50 dB. These values are typical for the measurement system. The power measurement at each frequency has associated with it a fluctuation, or noise, that is additive and approximately Gaussian on a decibel scale. This noise is

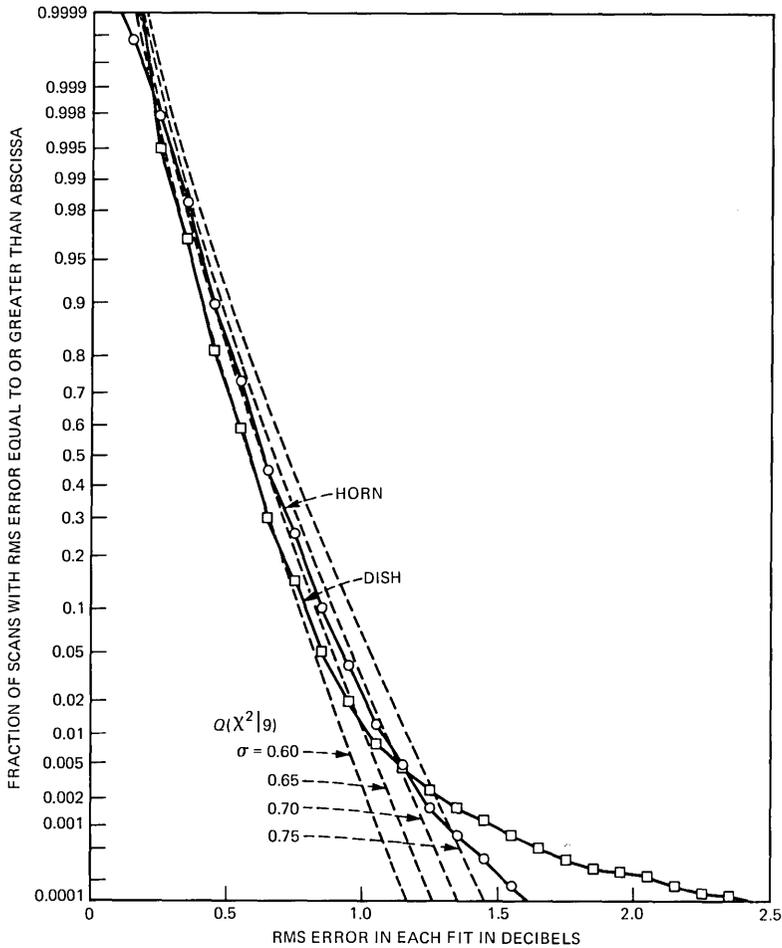


Fig. 7—Distribution of rms fit errors.

independent both frequency-to-frequency and scan-to-scan, and is large enough (0.6 to 0.7 dB) to mask quantization errors.

If the differences between measured and estimated powers were due solely to Gaussian fluctuations with a standard deviation, σ , the quantity $12E_{\text{rms}}^2/\sigma^2$ would be a χ^2 random variable with nine degrees of freedom,^{6,7} where E_{rms} is the rms error in fitting a scan. Thus, one can determine the quality of the modeling by comparing the distribution of the values of E_{rms} for all the horn or dish scans with that of a χ^2 variable. The distribution of the horn and dish rms errors are plotted in Fig. 7 along with the χ^2 distributions for several σ values. The excess error, the difference between a sample distribution and a χ^2 distribution

that matches it at the median, is more than 0.1 dB for less than 1 percent of all scans. Because the time between scans is not identical, there is no precise method of scaling the percentages in Fig. 7 to seconds per year. However, an approximate scaling is achieved by applying the percentage to the time covered by the data base and interpreting the resultant time as representing 0.75 of a year. On this basis the excess error in fitting the dish data exceeds 0.5 dB for about 40 seconds a year. For the horn data the excess error is less than 0.4 dB over the entire data base. We conclude that the fitting is exceptionally good for the horn and better than average⁶ for the dish.

IV. VERIFICATION OF MODEL STATISTICS

4.1 Overview

We shall begin our discussion with a description of the general modeling problem. By drawing on the properties of probability density functions and on past experience in modeling-selective fading, we will simplify the problem, somewhat. We conclude this subsection with a statement of the objectives of the remainder of the section.

By representing each scan of the horn and the dish with (2) and (3), we obtain a reduced data base consisting of 85,410 sextuples of values of $(A_H, A_D, B_H, B_D, \phi_H, \phi_D)$. Each of these sextuples has associated with it a time weighting corresponding to the time interval until the next scan in the same fading event. We wish to describe this data by a function, $T_D(A_H, A_D, B_H, B_D, \phi_H, \phi_D)$, whose values are equal to the number of seconds the six parameters were in a differential element of the parameter space, centered on the point $(A_H, A_D, B_H, B_D, \phi_H, \phi_D)$. Normalizing T_D to the data base time span, we obtain a probability density function,

$$p(A_H, A_D, B_H, B_D, \phi_H, \phi_D) = T_D(A_H, A_D, B_H, B_D, \phi_H, \phi_D)/44386. \quad (20)$$

It is this probability density function that we wish to determine. We will ultimately show that it may be approximated by the product of the probability functions in (10).

To simplify (20), let us first rewrite it as

$$\begin{aligned} p(A_H, A_D, B_H, B_D, \phi_H, \phi_D) \\ = p_{AB/\phi}(A_H, A_D, B_H, B_D | \phi_H, \phi_D) p_\phi(\phi_H, \phi_D). \end{aligned} \quad (21)$$

In previous work where multipath fading on a single antenna was statistically modeled,^{6,7} it was found that the notch frequency statistics were not related to the relative notch depth statistics or to the fade-level statistics. An examination of the current data base revealed the same properties, that is, the statistics of ϕ_H do not depend on those of A_H or B_H , those of ϕ_D do not depend on those of A_D or B_D . Under the

assumption that a cross-coupling (between ϕ_H and A_D , for instance) is even less likely, it was assumed at the outset that (21) can be written as

$$p(A_H, A_D, B_H, B_D, \phi_H, \phi_D) = p_{AB}(A_H, A_D, B_H, B_D)p_\phi(\phi_H, \phi_D). \quad (22)$$

We rewrite (22) as

$$p(A_H, A_D, B_H, B_D, \phi_H, \phi_D) = p_{A/B}(A_H, A_D | B_H, B_D)p_B(B_H, B_D)p_\phi(\phi_H, \phi_D). \quad (23)$$

In the remainder of this section we shall derive the functional form of each of the probability density functions in (23). The ultimate objective is to show that (23) can be represented by the factors multiplying T_0 , on the right-hand side of (10), with the various probability density functions as defined in (12) to (19). To this end, we consider the joint distribution of B_H and B_D in Section 4.3; we show that it can be represented as the distribution of two independent variables with distributions given by (12) and (13). The form of the conditional distribution $p_{A/B}(A_H, A_D | B_H, B_D)$, as given in (14) to (17), is derived in Section 4.3. In Section 4.4, we consider the notch frequency distribution and show that it can be modeled by independent random variables as given by (18) and (19).

As part of the process of developing a multidimensional statistical model one must make many choices that may seem arbitrary. However, we have proceeded with the philosophy that we should represent the data well whenever there is a significant degree of fading present in either antenna. To accurately represent the most severe events, we must develop our cumulative distribution functions from the more severe to the less severe fades, i.e., the complement of the usual cumulative distribution function. Our goal is to find the simplest probability functions that match these distribution functions, where we define simplest functions as those having the fewest possible number of parameters. In assessing how well these objectives are achieved, we view the composite data base as a member of an ensemble of all possible data bases. Thus, the parameters obtained from the data base must be considered as random variables of this ensemble of fading events.

4.2 Notch depth statistics

The objective of the diversity model is to accurately represent the transmission shapes present on both the horn and dish at any time that deep or shapely fading was present on either. In previous work with the nondiversity model it was only necessary to represent the notch depth parameter at values large enough to produce several

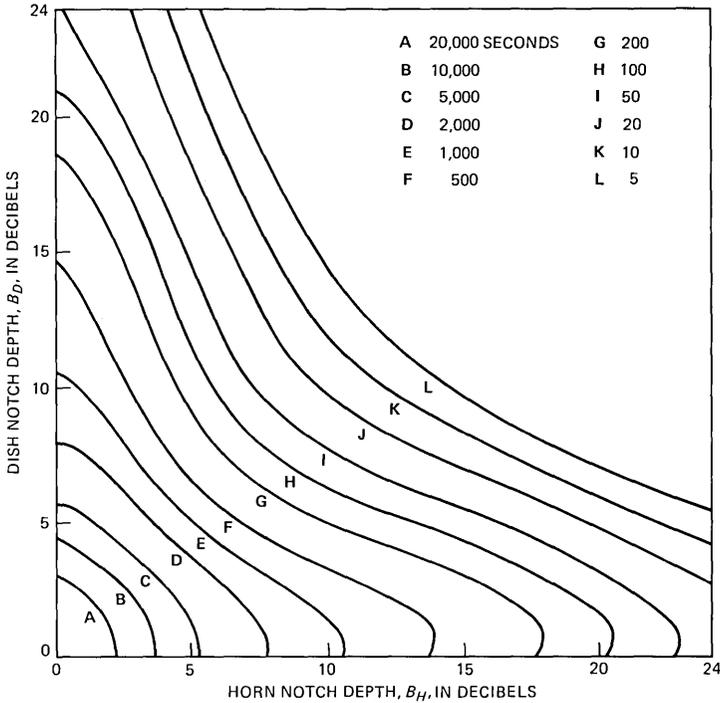


Fig. 8—Contour plot of $\hat{F}(B_H, B_D)$, smoothed fit to cumulative joint distribution function of horn and dish relative notch depths.

decibels of shape in the band. For the diversity model considered here, we must represent the joint distribution of horn and dish notch depth parameters at all values. (While one could give less importance to the distribution in regions where both notch depth parameters are small, this is found to be unnecessary.) We first develop the complement of the cumulative distribution of the sample values of the horn and dish relative notch depths, B_H and B_D , respectively. We define this two-dimensional function, $F(x, y)$, as the number of seconds in the data base that B_H equaled or exceeded a value x , and B_D equaled or exceeded y :

$$F(x, y) = \text{Number of seconds: } B_H \geq x, B_D \geq y. \quad (24)$$

To provide a focus for the ensuing discussions, we develop the function $\hat{F}(x, y)$, which is a smooth function fitted to the multiply-discontinuous function $F(x, y)$. Figure 8 shows a contour plot of the surface $\hat{F}(x, y)$. It shows, for instance, that there are fewer than 20 seconds in the data base during which both the horn and dish notch depths simultaneously exceeded 10 dB. There are fewer than 5 seconds

for which the notch depth on one antenna equals or exceeds 15 dB, while that on the other is 10 dB or greater.

Figure 8 was derived by first determining the values of the function $F(x, y)$ on a square grid of points, x_i, y_j , where

$$\begin{aligned} x_i &= \begin{cases} 0 & i = 1 \\ i - 0.5 & i > 1 \end{cases} \\ y_j &= \begin{cases} 0 & j = 1 \\ j - 0.5 & j > 1. \end{cases} \end{aligned} \quad (25)$$

Since $F(x, y)$ falls off approximately exponentially with increasing values of x and y , we approximate it with $\hat{F}(x, y)$, where

$$\hat{F}(x, y) = \exp\left(-\sum_{m+n \leq N} a_{mn} x^m y^n\right). \quad (26)$$

The coefficients a_{mn} in (26) were determined by minimizing the mean square error between $\ln \hat{F}(x, y)$ and $\ln F(x, y)$ over all x_i, y_j , less than 24 dB, for which there were five or more seconds in the data base, $F(x_i, y_j) \geq 5$. [The 5-second limit was chosen to avoid the region of the x, y plane where data is becoming sparse, causing the sample function, $F(x, y)$, to have increasingly extensive flat areas.] Figure 8 shows the equal value contours of $\hat{F}(x, y)$ as defined by (26) with $N = 6$, for which 28 parameters, a_{mn} , were determined.

The function $\hat{F}(x, y)$ with $N = 6$ provides an excellent representation of the sample function, $F(x, y)$; the rms error between $\ln \hat{F}(x, y)$ and $\ln F(x, y)$ is approximately 0.092, which corresponds to an rms error of 9.2 percent over the fitting region. While one can reduce the fitting error by increasing the dimension, N , of $\hat{F}(x, y)$, the reduction is not great. The minimum error, 6.7 percent, is obtained with $N = 9$ (55 coefficients, a_{mn}). Furthermore, the fitted functions, the $\hat{F}(x, y)$'s, lose the appearance of distribution functions for N greater than 6. Note that the distribution function, $F(x, y)$, must satisfy the inequality

$$F(x, y) \geq F(x', y') \quad \text{for } x' \geq x, y' \geq y. \quad (27)$$

It may be seen that $\hat{F}(x, y)$ for $N = 6$, as shown in Fig. 8, violates this inequality for B_D near 1 dB and B_H greater than 12 dB.

An extensive study was undertaken to find a distribution function that would fit $F(x, y)$. While a function with a polynomial exponent such as $\hat{F}(x, y)$ can be fitted by solving a system of $(N + 1)(N + 2)/2$ linear equations, the more general distribution functions were fitted using a modified gradient search routine.²¹ For practical reasons the class of functions was limited to those having no more than eight parameters. Of those tried, the best was of the form

$$F_M(x, y) = z_0(e^{-z_1 x^2} + z_3 e^{-z_2 x})(e^{-z_4 y^2} + z_6 e^{-z_5 y}). \quad (28)$$

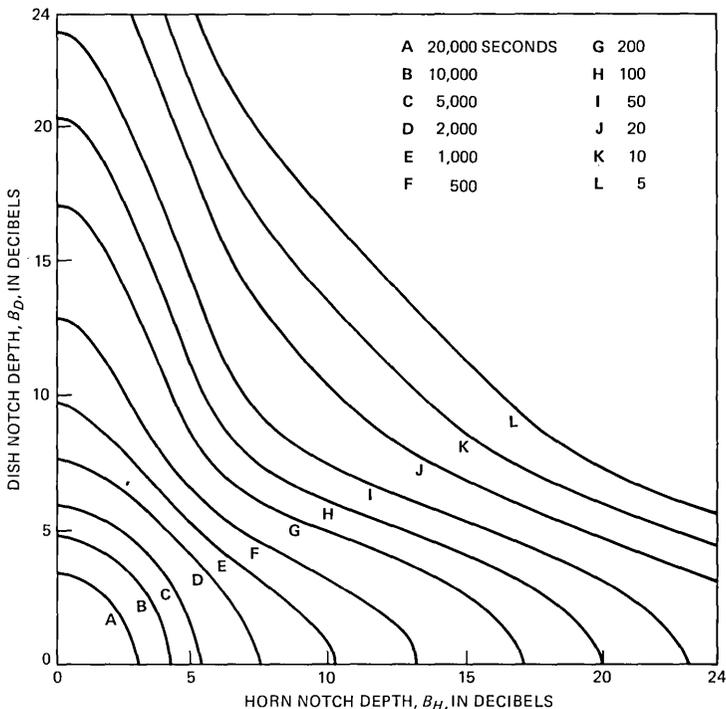


Fig. 9—Contour plot of $F_M(B_H, B_D)$, model cumulative joint distribution function of horn and dish relative notch depths.

For the parameter values corresponding to those in (12) and (13), the rms fitting error was 18.7 percent. The contour plot of this function, shown in Fig. 9, is seen to match the smoothed function of Fig. 8 quite closely. To show how well $F_M(x, y)$ matches the original sample distribution, we plot the distribution of horn notch depth conditioned on dish notch depth in Fig. 10, along with the values of the sample distribution, $F(x_i, y_j)$, being matched. A similar plot of the distribution of dish notch depth conditioned on horn notch depth is shown in Fig. 11.

Comparing Figs. 8 and 9, we see that the model function, $F_M(x, y)$, has most of the properties of the function $\hat{F}(x, y)$. From Figs. 10 and 11, one obtains an appreciation of the irregularities in the trends in the distribution of the data points; these irregularities contribute substantially to the fitting error. While one could argue that the modeling is acceptable on the basis of Figs. 8 to 11, there are more compelling reasons for accepting this distribution function as representing the data distribution, as we shall outline in the following paragraphs.

If the function used to represent the data distribution, $F(x, y)$, can be factored as a product of a function of x and a function of y , the

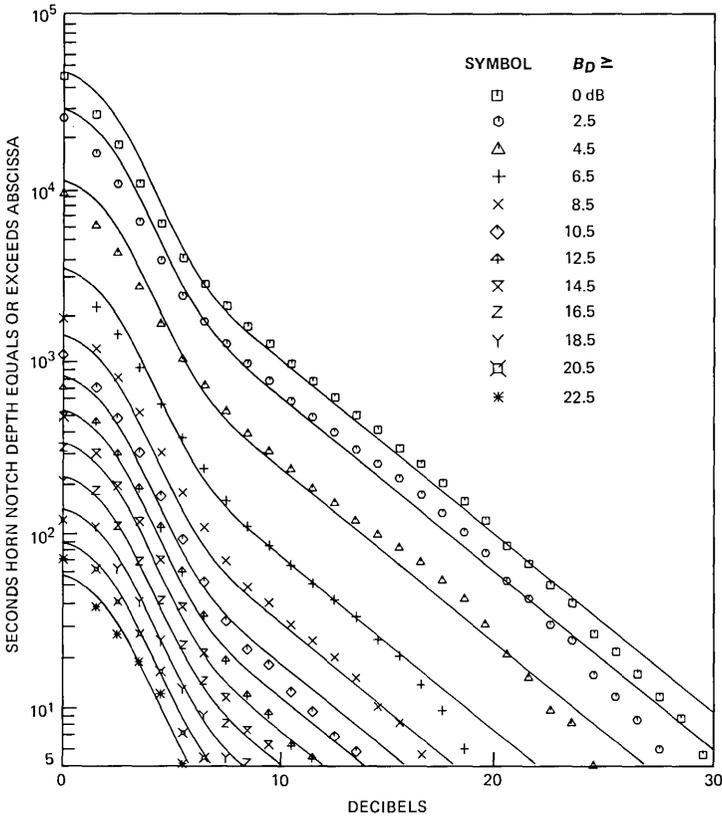


Fig. 10—Distribution of relative notch depth for the horn, conditioned on that for the dish.

resulting probability model will be one in which B_H and B_D are statistically independent. We will first show that the modeling function, $F_M(x, y)$, provides close to the minimum achievable error under the factorability assumption. Subsequently, we shall show that there is no reasonable alternative.

If the sample distribution were factorable, we could represent it exactly by

$$F_F(x, y) = c_0 \left[\exp\left(-\sum_{m=1}^N a_m x^m\right) \right] \left[\exp\left(-\sum_{n=1}^N b_n y^n\right) \right], \quad (29)$$

since we are only representing the sample distribution at a finite number of points. Fitting a function of the form (29) to $F(x, y)$, as described previously, we find that the minimum rms error is 17.4 percent for $N = 7$, which corresponds to 15 terms. For larger dimensionality the error increases, presumably, because of loss of precision

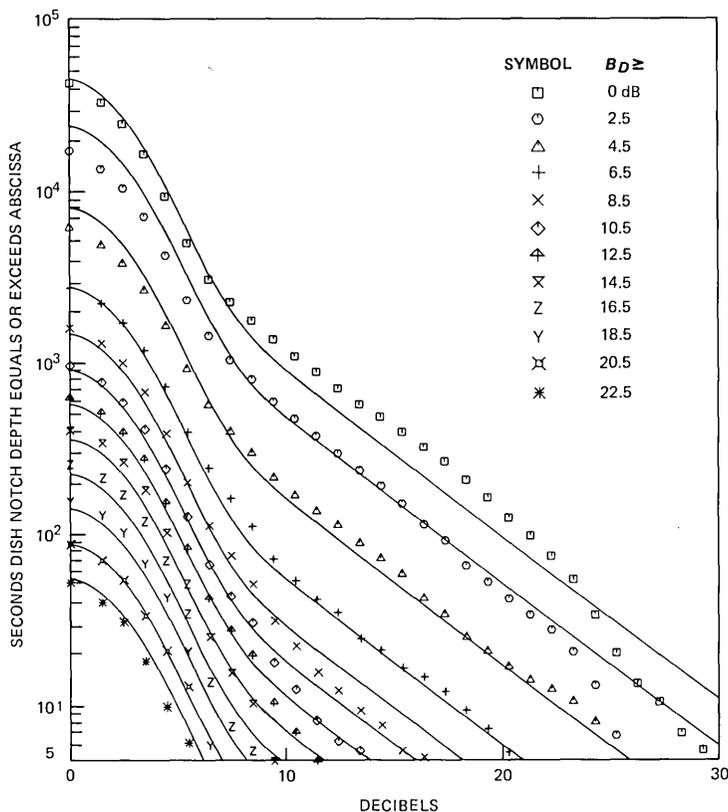


Fig. 11—Distribution of relative notch depth for the dish, conditioned on that for the horn.

in the double precision solution to the set of linear equations for the coefficients. The model solution achieves an rms error of 18.7 percent with fewer than half the number of coefficients.

As a check of the ruggedness of the model solution, we varied the region over which the model function was fitted. If the upper limit of x and y is reduced from 24 to 22 dB, the parameters change by less than one percent. If the region where both x and y are less than 5 dB is removed, the change is even smaller. We conclude that $F_M(x, y)$ provides an accurate and stable approximation to $F(x, y)$.

There is no way of testing whether B_H and B_D are statistically independent. One check, which we can apply, is to estimate their correlation coefficient. By examining the data base, we find that the coefficient of correlation between B_H and B_D is 0.0306. While there are standard tests for the significance of correlation coefficients,²² they pertain to the case of independent sets of samples, whereas the time

series samples of the fading parameters we are working with are correlated.²³ In the appendix, we show that the effective number of independent time samples of B_H and B_D in the data base is approximately 1516, and that one would expect a correlation of this magnitude (0.0306) or greater about 27 percent of the time for this sample size.

Various subpopulations of the B -parameters were also examined for correlations. For instance, consider the set of B_H and B_D observations for which both were equal to or greater than 6 dB. The correlation coefficient is -0.118 for this subpopulation. It is shown in the appendix that a correlation coefficient of this magnitude or greater would be expected to occur 53 percent of the time.

Thus, we have based our choice of the model function $F_M(x, y)$ on the following grounds: (i) the model function (Fig. 9) captures the essential morphology of the sample distribution (Fig. 8); and (ii) no candidate distribution function providing correlation between B_H and B_D and employing a similar number of parameters represents the distribution function as well. While the rms error between the model distribution and the sample distribution is considerably larger than is that between the best functional representation, $\hat{F}(x, y)$, and the sample distribution, this is to be expected because $\hat{F}(x, y)$ has many more degrees of freedom and is not constrained to be a distribution function. In other words, $\hat{F}(x, y)$ represents the data within the region of interest by following all minor irregularities; immediately outside this region, this function exhibits large amplitude oscillations. By examining the correlation of B_H and B_D for various subpopulations we have established that there are correlations within these subpopulations. These correlations correspond to variations in the sample distribution surface that a factorable function, such as the model function, is incapable of matching, but which have been shown to be without significance. We conclude that there is no basis for choosing a more complicated function than (28) for representing the sample joint distribution.

4.3 Fade-level statistics

For the nondiversity fading model, the fade level or A -distribution was Gaussian with a mean dependent on the relative notch depth. A generalization of this, a two-dimensional Gaussian probability density function, describing the joint probability of A_H and A_D conditioned on B_H and B_D would be given by (14), with g_H , g_D , σ_H , σ_D , and ρ being functions of B_H and B_D . To obtain, for the diversity model, a probability density function that will easily reduce to that of the nondiversity model, we assume that g_H depends only on B_H , g_D depends only on B_D , and that σ_H , σ_D , and ρ are independent of both B_H and B_D .

As the first step in verifying this hypothesis, we must determine the

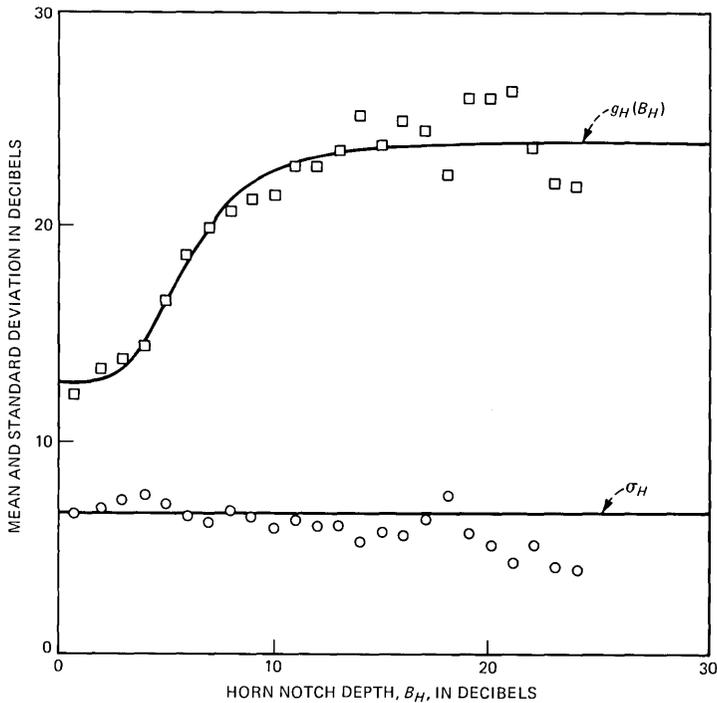


Fig. 12—Mean and standard deviation of A_H , fade-level parameter for the horn, as modeled and estimated for horn relative notch depth in 1-decibel intervals.

functional dependence of the means, g_H and g_D , on their respective variables, B_H and B_D . We do this by estimating the value of, say, g_H at a set of values of B_H , and fit a function to these sample means. The value of $g_H(x)$ is the average value of A_H for B_H equal to x . We estimate $g_H(x)$ by taking the expectation, or average value of A_H in the data base for all times that B_H is between $x - \delta$ and $x + \delta$. Specifically, we work with 1-dB intervals and estimate g_H by

$$\hat{g}_H \left(\frac{x_i + x_{i+1}}{2} \right) = E \{ A_H : x_i \leq B_H \leq x_{i+1} \}, \quad (30)$$

where $E \{ \cdot \}$ denotes expectation, and the x_i 's are defined by (25).

The values of $\hat{g}_H(x)$ are indicated by squares in Fig. 12, which also shows the function $g_H(B_H)$ of (15), which we use as the conditional mean of A_H in the model. We use a meromorphic function, $g_H(B_H)$, to represent this conditional mean to ensure that it approaches a constant at large values of horn notch depth. Note that the accuracy of the estimates of \hat{g}_H decrease at large values of the horn notch depth because the number of samples decreases. (The approximating function g_H was obtained from a weighted least-squares fit to the estimates,

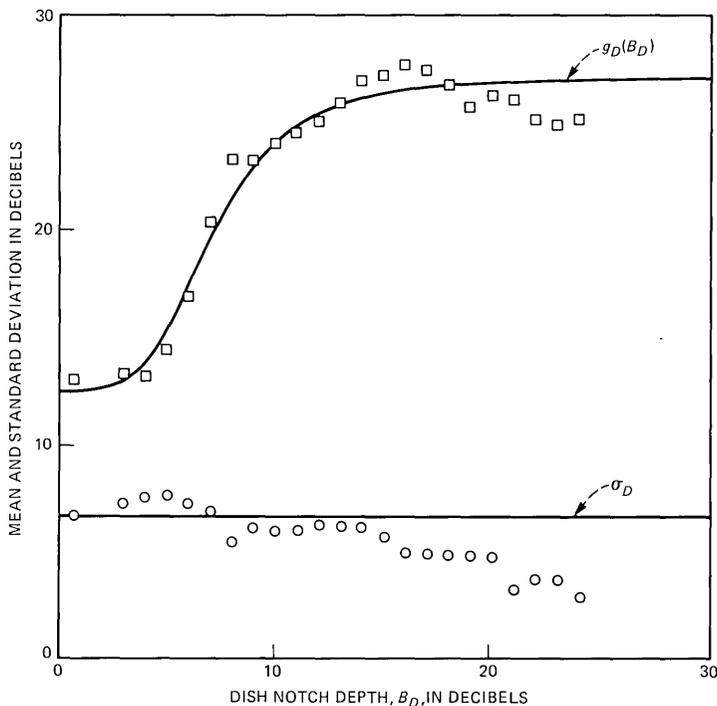


Fig. 13—Mean and standard deviation of A_D , fade-level parameter for the dish, as modeled and estimated for dish relative notch depth in 1-decibel intervals.

\hat{g}_H . The weighting was in proportion with the square root of the number of seconds of data in the notch depth interval.)

Figure 12 also shows the results of estimating the standard deviation of A_H conditioned on B_H in the same set of intervals. The straight line represents the (unconditional) standard deviation (6.8268) of $A_H - g_H(B_H)$ for the whole data base. Figure 13 shows the results of duplicating for the dish parameters the calculations leading to Fig. 12.

It is a simple matter to test the validity of the hypothesized model. If $A_H - g_H(B_H)$ and $A_D - g_D(B_D)$ are jointly Gaussian with zero means, correlation ρ , and respective standard deviations σ_H and σ_D , they may be linearly transformed into a pair of zero mean, unit-variance, independent, Gaussian random variables. We shall develop the transformation in two steps: a rotation of axes, followed by a scale change. Taking advantage of hindsight, we plot in Fig. 14 contours of the joint probability density for A_H and A_D of (14). The rotated axes x and y are defined by the transformation:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} A_H - g_H(B_H) \\ A_D - g_D(B_D) \end{bmatrix}. \quad (31)$$

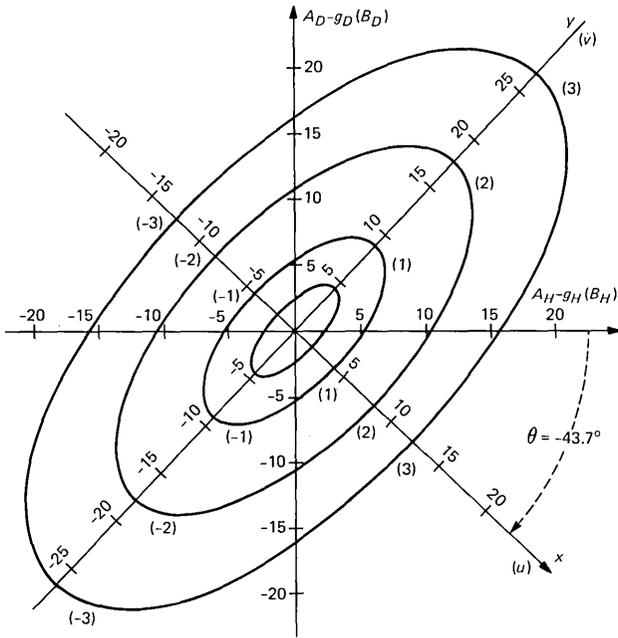


Fig. 14—Equal value contours of the joint conditional probability density function for fade-level parameters, A_H and A_D , showing variable transformations.

The chosen angle, θ ($= -43.7$ degrees), satisfies the relation

$$\tan 2\theta = \frac{2\rho\sigma_H\sigma_D}{\sigma_H^2 - \sigma_D^2}, \quad (32)$$

which ensures that x and y are uncorrelated. Their variances are given by

$$\sigma_x^2 = \sigma_H^2 \cos^2 \theta + \sigma_D^2 \sin^2 \theta + 2\rho\sigma_H\sigma_D \sin \theta \cos \theta \quad (33)$$

and

$$\sigma_y^2 = \sigma_H^2 \sin^2 \theta + \sigma_D^2 \cos^2 \theta - 2\rho\sigma_H\sigma_D \sin \theta \cos \theta, \quad (34)$$

respectively. From the variances and correlation in (17) and the value of θ given above, we calculate the values of σ_x and σ_y , as 4.097 and 8.900, respectively. Note that the major axes of the ellipses of concentration shown in Fig. 14 lie along the y -axis, and the minor axes along the x -axis. By rescaling the x - and y -axes we obtain the desired zero mean, unit variance, independent variables, u and v , as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1/\sigma_x & 0 \\ 0 & 1/\sigma_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (35)$$

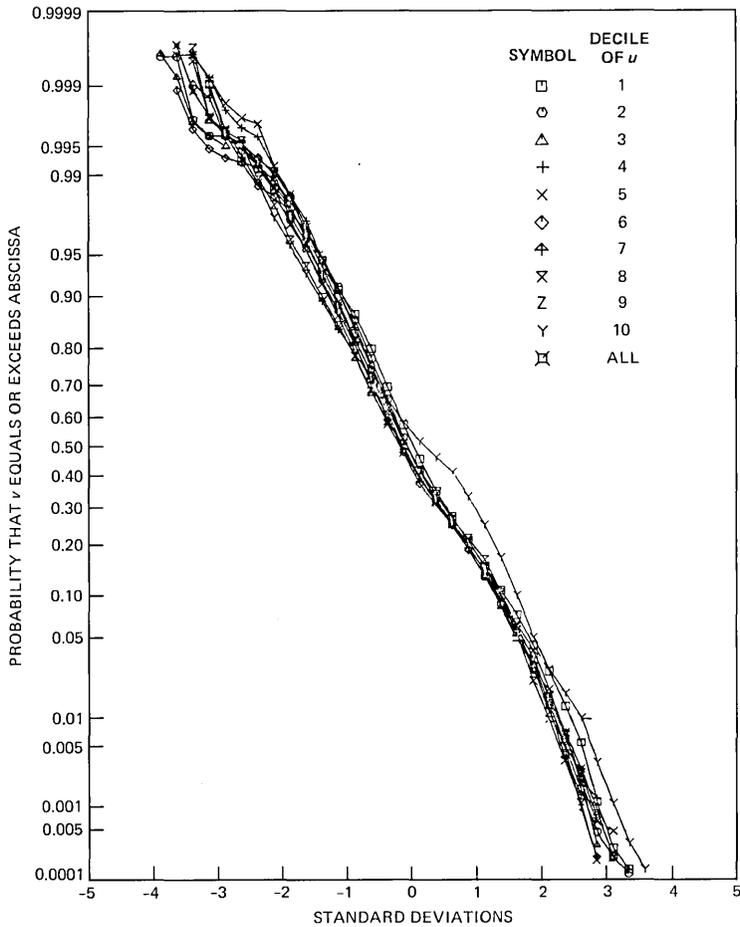


Fig. 15—Distribution of canonic sum parameter, v , conditioned on deciles of canonic difference parameter, u .

One may generate the sample distribution of u and v from the data base by calculating their values at each scan using eqs. (15) to (17) and (31) to (35). The conditional cumulative distribution functions of these two variables are shown in Figs. 15 and 16. Each plotted curve represents the cumulative distribution function of all values of one of the variables conditioned on the other variable being in a given decile of a Gaussian distribution; e.g., the first decile of u contains all u values that are less than -1.28155 . Figure 15 shows the distributions of u conditioned on v ; Fig. 16 shows v conditioned on u . Both families of distributions are closely grouped and approximately Gaussian within the range of -3 to $+3$ standard deviations. This is remarkably good agreement for a sample of this size.

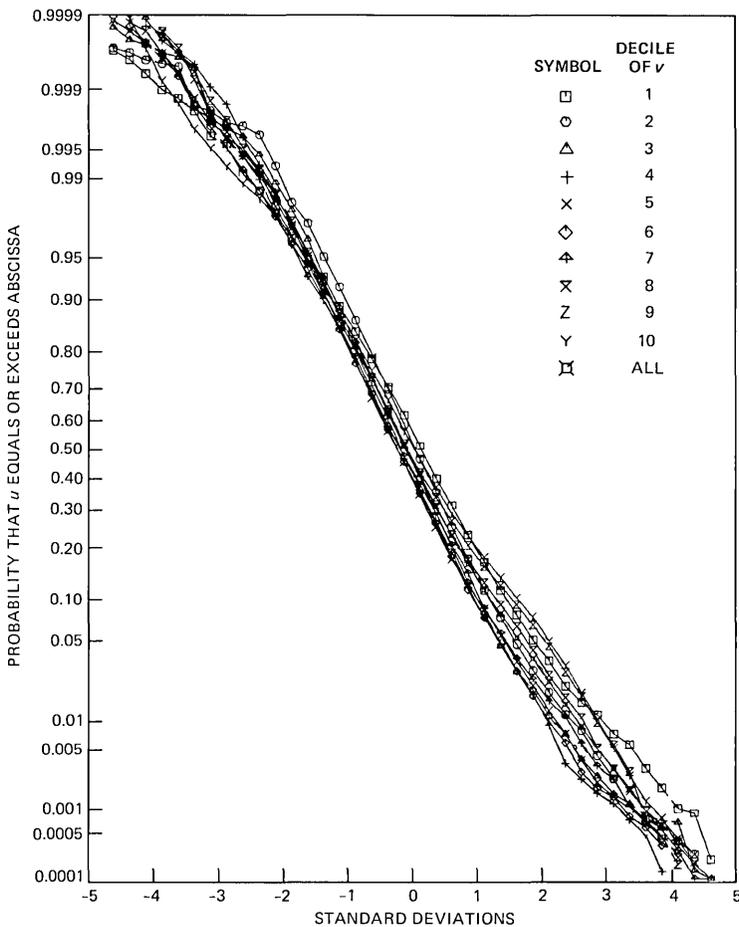


Fig. 16—Distribution of canonic difference parameter, u , conditioned on deciles of canonic sum parameter, v .

Figures 15 and 16 provide good confirmation of the assumption that A_H and A_D are jointly Gaussian variables. The only assumptions that were studied further were those relating to the functional form of g_H and g_D . The procedures outlined above were carried through under the assumption that g_H and g_D were both functions of both B_H and B_D . The resultant u and v conditional distribution functions were not noticeably different from Figs. 15 and 16. The only notable difference was that the value of σ_H was reduced from 6.8268 to 6.6815, that of σ_D from 7.0272 to 6.954, and the correlation coefficient increased from 0.650 to 0.681. This trivial difference in the coefficients would be achieved at considerable cost in complexity because the probability density function, (14), would become extremely difficult to use. The

effect of this generalized assumption is not pronounced on the functions g_H and g_D . For instance, the fractional variation of $g_H(B_H, B_D)$, as B_D varied its values along any line of constant B_H , was on the order of 10 percent. Hence, the development of a more complicated distribution cannot be justified.

4.4 Notch frequency statistics

As a first look for dependency between the horn and dish notch frequency parameters, the correlation coefficient between these two variables, ϕ_H and ϕ_D , was determined from the data base to be -0.0281 . Using the techniques described in the appendix, one would expect a correlation magnitude larger than this value to occur 7.3 percent of the time in a data base of this size if ϕ_H and ϕ_D were statistically independent. While this correlation is small, it is large enough to be considered "almost significant" and to warrant a more detailed study.

The simplest way to look for the existence of any interdependency between the horn and dish notch frequencies is to plot distributions of one conditioned on the other. For instance, one chooses all those time intervals when the value of ϕ_D was between -85 and -65 degrees (or f_{0D} was between 37.4 and 28.6 MHz below the center frequency of the channel). One then plots the fraction of this time interval (2615 seconds) that the horn notch frequency exceeded a given value. The resulting conditional distribution is labeled by octagons on the composite plot of Fig. 17. The other distributions in Fig. 17 were obtained by conditioning the notch frequency on other intervals of dish notch frequency, as indicated. For reference, the span of the channel measurements is from -27.5 to $+27.5$ degrees (± 12.1 MHz).

Since the overall spread of this family of distributions is small (less than 10 percent), over the range of ϕ_H , we represent the entire family by a single distribution, (18), shown dashed. Thus, we describe the distribution as uniform at two levels with values of $|\phi_H|$ less than 90 degrees being five times as likely as values greater than 90 degrees.

Figure 18 shows a set of distributions of the dish notch frequency conditioned on the horn notch frequency. This family of distributions is very tightly clustered, implying that the horn notch position had no influence on the distribution of dish notch position. The family of dish notch distributions does not fit the two-level uniform approximation, shown dashed in Fig. 18, as well as does the horn data. The relatively large deviation (about 15 percent near -30 degrees) between the data and the modeled distributions results from an asymmetry in the data, where, on a physical basis, none should be found. For atmospheric multipath, one expects transmission notches to be equally likely at any frequency in the neighborhood; hence, the notch frequency probability density functions should be symmetric and the cumulative distribution

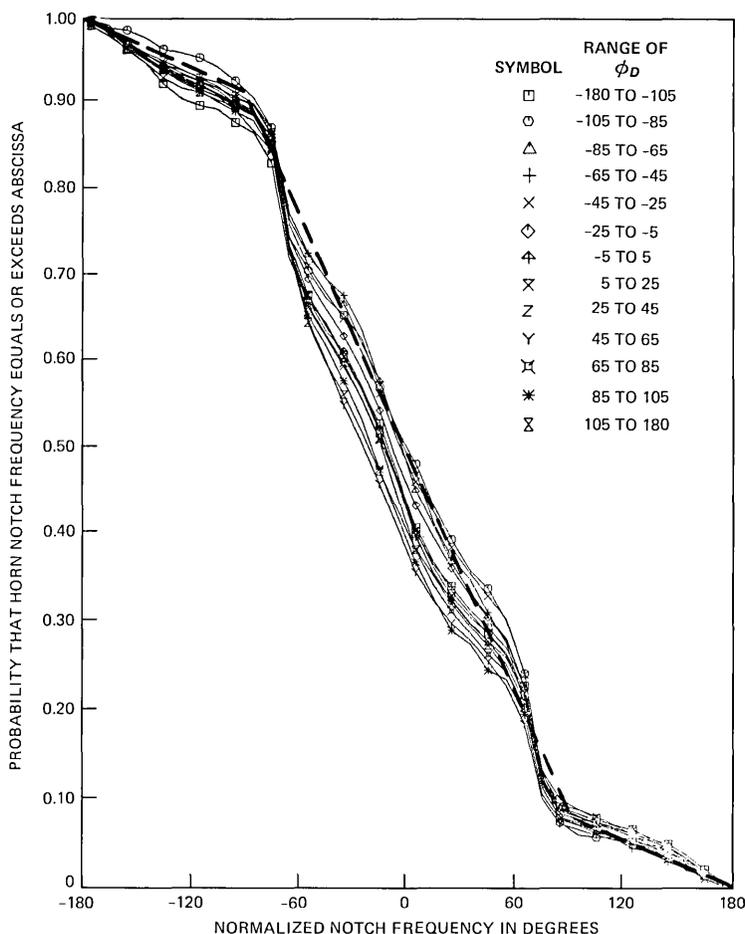


Fig. 17—Distribution of horn notch frequency conditioned on dish notch frequencies in specified intervals.

functions shown should be antisymmetric about the (0 degree, 50 percent) point.

In the central 60-degree region of the horn notch frequency distribution, the distributions conditioned on positive dish notch frequencies cluster separately from those conditioned on negative frequencies. This is an integrated effect, in that it is not apparent in the conditional probability density functions (not shown). There may be some relation between this spread in the horn notch frequency distribution, the spreading of the time-faded statistics for the dish (Fig. 4), and the symmetry properties of the observed dish notch frequency distribution (noted in Fig. 18). However, these differences are small and no attempt

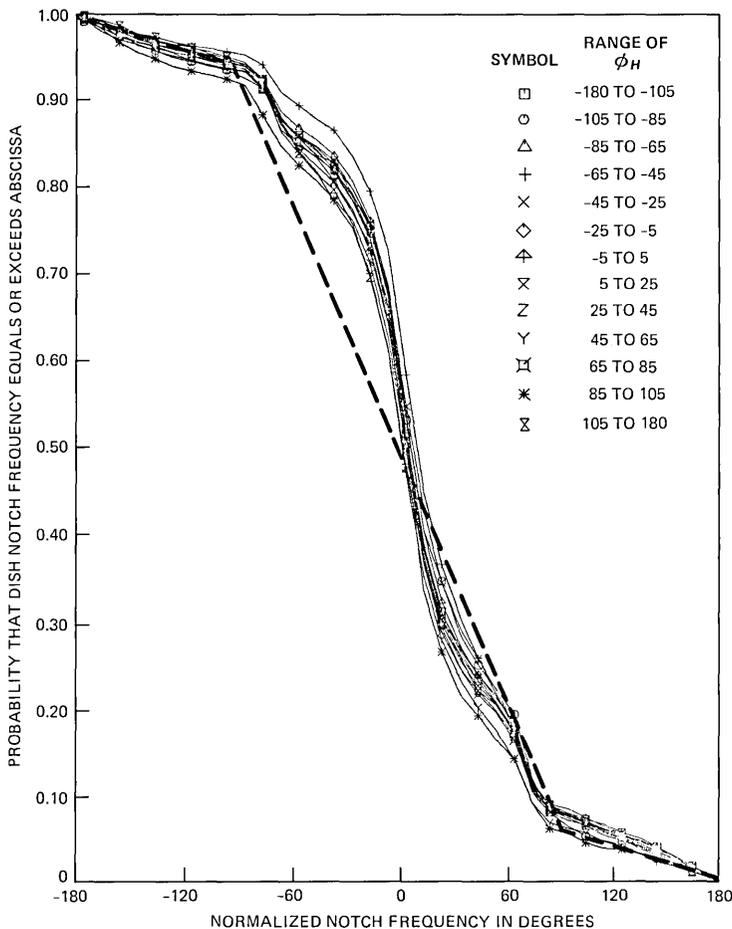


Fig. 18—Distribution of dish notch frequency conditioned on horn notch frequencies in specified intervals.

was made to isolate the events giving rise to these effects or to incorporate them into the model.

V. CONCLUDING REMARKS

We have provided in Section II a statistical model of multipath fading as observed in a diversity configuration over an extended period of time. Supporting evidence of the accuracy of the statistical model was presented in Section IV, along with a description of the methodologies employed. The transmission path to each antenna was represented by a function synthesizing a simplified three-path fade. While one could statistically represent the joint occurrence of the parameters

of these two functions in the observed data base with greater precision using a more complex model, all attempts to establish the significance of higher order extrapolations of the proposed statistical model have been negative. There would seem to be little virtue in representing, with a particular statistical model, features that would not be found in a fading data base corresponding to a different observation period.

For the proposed model, the only correlation in the fading of the two antennas is in the level of fade. This simplicity of dependence is a direct result of the form of the function used to represent fading on the two antennas. More, or different, interdependencies might become apparent or significant with other fade representations. The limited dependence is a virtue of the model since it places all of the impact of diversity antenna separation in a single parameter, for separations in the range of practical interest. However, for sufficiently small separation one would ultimately expect to see correlations in the shapes of the fades observed on the two antennas.

The proposed model was only checked against the data; however, it reduces to the form of the nondiversity model, which has been extensively verified. A model representing fading on the horn is obtained by integrating the model statistics over the dish parameters, and vice versa. Some of the parameters of the horn and dish models derived from the diversity model are different from those that have been derived previously, but the differences are not great, on the order of 10 percent, at most. As a consequence, one would not be surprised by 20 to 30 percent differences between expectations calculated with this model and corresponding expectations calculated with the nondiversity model. This merely reflects the month-to-month and year-to-year variability in the nature and severity of multipath fading.

ACKNOWLEDGMENT

This study would not have been possible without the contributions of many individuals. The radio equipment was installed, aligned, and maintained by R. A. Hohmann, C. W. Lundgren, and L. J. Morris, with instrumentation designed by G. A. Zimmerman. The data base was assembled and calibrated by D. W. Rowland under the guidance of M. V. Pursley, who also generated Figs. 3 to 5. C. H. Menzel provided much useful advice for the efficient utilization of computer resources and maintained the viability of these resources frequently at odd hours in the face of rolling waves of hardware/software catastrophies. The insights into the data would not have been possible without the plotting packages developed by A. K. Ranade. For many stimulating discussions of this work and comments on this manuscript, I would like to thank W. T. Barnett, S. H. Lin, W. W. Toy, T. S. Giuffrida, V.

K. Prabhu, and other members of the Transmission Systems Characterization Department.

REFERENCES

1. W. T. Barnett, "Multipath Fading Effects on Digital Radio," *IEEE Trans. Commun.*, *COM-27*, No. 12 (December 1979), pp. 1842-48.
2. C. W. Anderson, S. G. Barber, and R. N. Patel, "The Effect of Selective Fading on Digital Radio," *IEEE Trans. Commun.*, *COM-27*, No. 12 (December 1979), pp. 1870-76.
3. T. S. Giuffrida, "Measurements of the Effects of Propagation on Digital Radio Systems Equipped with Space Diversity and Adaptive Equalization," in *Conf. Rec.*, 1979 Int. Conf. on Commun., Vol. 3, pp. 48.1.1-6.
4. W. W. Toy, "The Effects of Multipath Fading on 16 QAM Digital Radio," in *Conf. Rec.*, 1980 Int. Conf. Commun., Vol. 3, pp. 52.1.1-5.
5. T. S. Giuffrida, "The Effects of Multipath Fading Upon Adjacent Channel Operation of an 8 PSK 6 GHz Digital Radio," in *Conf. Rec.*, 1980 Int. Conf. on Commun., Vol. 2, pp. 34.1.1-5.
6. W. D. Rummler, "More on the Multipath Fading Channel Model," *IEEE Trans. Commun.*, *COM-29*, No. 3 (March 1981), pp. 346-52.
7. W. D. Rummler, "A New Selective Fading Model: Application to Propagation Data," *B.S.T.J.*, *58*, No. 5 (May-June 1979), pp. 1037-72.
8. C. W. Lundgren and W. D. Rummler, "Digital Radio Outage Due to Selective Fading—Observation vs. Prediction From Laboratory Simulation," *B.S.T.J.*, *58*, No. 5 (May-June 1979), pp. 1073-1100.
9. A. J. Giger and W. T. Barnett, "Effects of Multipath Propagation on Digital Radio," *IEEE Trans. Commun.*, *COM-29*, No. 9, (September 1981), pp. 1345-52.
10. T. S. Giuffrida and W. W. Toy, "16 QAM and Adjacent Channel Interference," in *Conf. Rec.*, 1981 Int. Conf. on Commun., Vol. 1, pp. 13.1.1-5.
11. L. J. Greenstein and B. A. Czekaj, "A Polynomial Model for Multipath Fading Channel Responses," *B.S.T.J.*, *59*, No. 7 (September 1980), pp. 1197-1225.
12. W. C. Jakes, Jr., "An Approximate Method to Estimate an Upper Bound on the Effect of Multipath Delay Distortion on Digital Transmission," *IEEE Trans. Commun.*, *COM-27*, No. 1 (January 1979), pp. 76-81.
13. L. J. Greenstein and V. K. Prabhu, "Analysis of Multipath Outage with Applications to 90 Mbit/s PSK Systems at 6 and 11 GHz," *IEEE Trans. Commun.*, *COM-27*, No. 1 (January 1979), pp. 68-75.
14. J. Sandberg, "Extraction of Multipath Parameters from Swept Measurements on a Line-of-Sight Path," *IEEE Trans. Antennas and Propagation*, *AP-28*, No. 6 (November 1980), pp. 743-50.
15. L. Martin, "Etude de la Selectivite des Evanouissements dus aux Trajets Multiples," *Ann. Telecommunic.*, *35*, No. 11-12 (November-December 1980), pp. 482-7.
16. F. Andreucci, F. Fedi, and P. G. Marchetti, "Analytical Method to Evaluate the Effect of Multipath Fading on Long-Haul High-Capacity Digital Radio Relay Systems," *Ann. Telecommunic.*, *35*, No. 11-12 (November-December 1980), pp. 488-93.
17. E. Damosso and S. DePadova, "A Statistical Model for the Evaluation of the Impairments Due to Multipath Fades on Digital Radio Links," in *Conf. Rec.*, 1981 Int. Conf. on Commun., Vol. 4, pp. 68.7.1-5.
18. W. T. Barnett, "Multipath Propagation at 4, 6, and 11 GHz," *B.S.T.J.*, *51*, (February 1972), pp. 321-61.
19. A. Vigants, "Space-Diversity Engineering," *B.S.T.J.*, *54*, No. 1 (January 1975), pp. 103-41.
20. W. D. Rummler, "Time- and Frequency-Domain Representation of Multipath Fading on Line-of-Sight Microwave Paths," *B.S.T.J.*, *59*, No. 5 (May-June 1980), pp. 763-796.
21. P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*, New York: McGraw-Hill Book Company, 1969.
22. H. Cramer, *Mathematical Methods of Statistics*, Princeton, N.J.: Princeton University Press, 1946.
23. W. D. Rummler, "Advances in Multipath Channel Modeling," in *Conf. Rec.*, 1980 Int. Conf. on Commun., Vol. 3, pp. 52.3.1-5.

APPENDIX

Correlation and Significance

The purpose of this appendix is to develop the methods of testing for the significance of the correlation between two random variables when the sample values of each variable are taken from a time series with known autocorrelation. Consider two stationary, independent, zero mean, unit variance random processes, $x(t)$ and $y(t)$. Assume that we have samples of each at a large number, N , of time instants, where the i th time instant is taken as

$$t_i = i\Delta t \quad i = 1, 2, 3, \dots, N. \quad (36)$$

We are interested in the correlation coefficient of x and y as determined from the set of samples $x_i = x(t_i)$ and $y_i = y(t_i)$, where

$$\rho = \frac{1}{N} \sum_{i=1}^N x_i y_i. \quad (37)$$

Because $x(t)$ and $y(t)$ are independent processes by assumption, the expected value of ρ is zero:

$$\bar{\rho} = \frac{1}{N} \sum_{i=1}^N \overline{x_i y_i} = 0, \quad (38)$$

where the overbar implies an ensemble average.

We wish to determine the variance of the estimate (37). We write the expected value of this sample variance as

$$\sigma_\rho^2 = \overline{\rho^2} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \overline{x_i y_i x_j y_j}. \quad (39)$$

Under the hypothesis that $x(t)$ and $y(t)$ are independent processes, we may rewrite (39) as

$$\sigma_\rho^2 = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \overline{x_i x_j y_i y_j} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \rho_x(i-j) \rho_y(i-j), \quad (40)$$

where

$$\rho_x(i) = \overline{x_k x_{k+i}} \quad (41)$$

and

$$\rho_y(i) = \overline{y_k y_{k+i}} \quad (42)$$

are the autocorrelation functions of $x(t)$ and $y(t)$, respectively, at time difference $i\Delta t$. We may rewrite (40) as

$$\sigma_\rho^2 = \frac{1}{N^2} \left[N + 2 \sum_{k=1}^N (N-k) \rho_x(k) \rho_y(k) \right]. \quad (43)$$

If the time samples of $x(t)$ and $y(t)$ were independent, the autocor-

relation functions would be unity for $k = 0$ and zero elsewhere, that is,

$$\sigma_p^2 = \frac{1}{N} \quad \text{if } \rho_x(i) = \rho_y(i) = 0 \quad \text{for } i \neq 0. \quad (44)$$

Let us define an effective sample size, N_{eff} , such that, for a given sample size, N , and given autocorrelation functions, $\rho_x(i)$ and $\rho_y(i)$,

$$\sigma_p = \frac{1}{\sqrt{N_{eff}}}. \quad (45)$$

Then,

$$\frac{N}{N_{eff}} = 1 + 2 \sum_{k=1}^N \left(\frac{N-k}{N} \right) \rho_x(k) \rho_y(k). \quad (46)$$

Our object is to use (45) and (46) for the diversity data base. As noted in Section III, the data base is not uniformly sampled. Furthermore, multipath fading is not a stationary random process. One may, however, define a lagged autocorrelation function for the samples of such a process²³ at delay τ as

$$\rho_x(\tau) = \frac{\left(\frac{1}{M} \sum x_m x_{m+\tau} \right) - \left(\frac{1}{M} \sum x_m \right) \left(\frac{1}{M} \sum x_{m+\tau} \right)}{\left[\left(\frac{1}{M} \sum x_m^2 \right) - \left(\frac{1}{M} \sum x_m \right)^2 \right]^{1/2} \left[\left(\frac{1}{M} \sum x_{m+\tau}^2 \right) - \left(\frac{1}{M} \sum x_{m+\tau} \right)^2 \right]^{1/2}}, \quad (47)$$

where the sample denoted $x_{m+\tau}$ is τ seconds delayed from the sample x_m , and the sums are taken over all $M = M(\tau)$ pairs of samples in the data base with a delay difference of τ .

Figure 19 shows plots of (47) for: (a) the relative notch depth of the horn, and (b) the relative notch depth of the dish. Figure 20 shows $M(\tau)$, the sample size for the data base, as a function of the delay, τ , for delays of integer numbers of seconds. Using these two figures as examples, we can approximate the quantity N/N_{eff} . We describe the autocorrelation functions by

$$\rho_x(\tau) = \alpha_x e^{-\tau/\tau_x} = \alpha_x \beta_x^\tau \quad (48)$$

and

$$\rho_y(\tau) = \alpha_y e^{-\tau/\tau_y} = \alpha_y \beta_y^\tau. \quad (49)$$

The samples in the data base are taken nonuniformly in a set of disjoint intervals. With 85,410 samples in 44,386 seconds we have an average sample spacing of 0.5 second. Hence, we approximate the uniform-sampling window function of (46), $(N-k)/N$, by taking an approximation to $M(\tau)/N$, as shown in Fig. 20 and given by

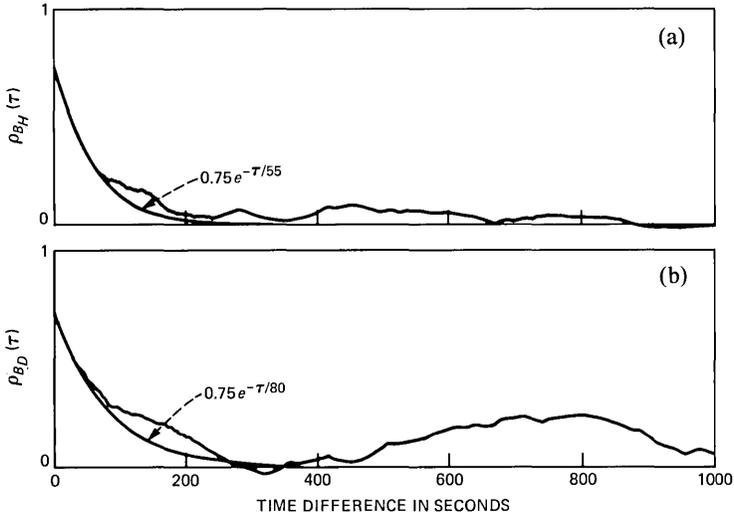


Fig. 19—Autocorrelation of lagged samples of relative notch depth for (a) horn and (b) dish.

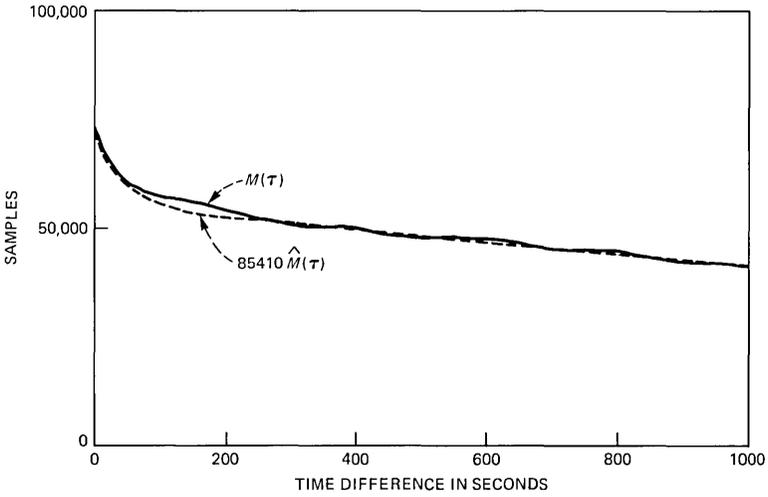


Fig. 20—Sample size for autocorrelation estimates in Fig. 19.

$$\hat{M}(\tau) = \frac{M(\tau)}{N} = \frac{3.5}{17} e^{-\tau/45} + \frac{11}{17} \left(1 - \frac{\tau}{4000} \right) \quad \tau \leq 4000. \quad (50)$$

Using (48) to (50) with $\tau = 0.5 k$ corresponding to the 0.5 second average sampling interval, we rewrite (46) as

$$\frac{N}{N_{eff}} = 1 + 2 \sum_{k=1}^N \hat{M}(k/2) \alpha_x \alpha_y (\beta_x \beta_y)^{k/2} \quad (51)$$

or

$$\frac{N}{N_{eff}} = 1 + \frac{2\alpha_x\alpha_y}{17} \left\{ \frac{3.5r e^{-1/90}}{1 - r e^{-1/90}} + \frac{11r}{1 - r} \left[1 - \frac{1}{8000(1 - r)} \right] \right\}, \quad (52)$$

where

$$r = (\beta_x\beta_y)^{0.5}. \quad (53)$$

Consider the horn and dish relative notch depths, B_H and B_D , respectively. Letting $B_H = x$ and $B_D = y$,

$$\alpha_x = \alpha_y = 0.75$$

$$\beta_x = e^{-1/55}$$

$$\beta_y = e^{-1/80}$$

$$r = 0.98478. \quad (54)$$

Therefore, from (51) to (54),

$$\frac{N}{N_{eff}} \approx 56.3. \quad (55)$$

For the data base, the calculated value of the correlation coefficient of B_H and B_D was 0.0306 for the entire set of 85,410 samples. From (55), $N_{eff} = 1516$. Using this in (45), we find $\sigma_\rho = 0.0257$. Under the unfavorable assumption²² that the distribution of ρ is normal, we would expect to find a value of $|\rho|$ this large ($1.19\sigma_\rho$) or greater to occur for less than 27 percent of the samples of this size.

Similarly, for the case of a subpopulation of 1608 samples of B_H and B_D , whose correlation coefficient was -0.118 , we find from (55) that $N_{eff} = 28.5$, and from (45) that $\sigma_\rho = 0.187$. We would expect to find a value of $|\rho|$ this large ($0.63\sigma_\rho$) or greater to occur for less than 53 percent of the samples of this size. One concludes that in neither of these two cases does the correlation differ significantly from zero.

Matrix Analysis of Mildly Nonlinear, Multiple-Input, Multiple-Output Systems With Memory

By A. A. M. SALEH

(Manuscript received April 29, 1982)

A matrix method of analysis is developed for mildly nonlinear, multiple-input, multiple-output systems with memory (e.g., nonlinear multiport networks and multichannel communication systems). The method is based on a Volterra-series representation whose kernels are two-dimensional matrices rather than multidimensional arrays. This is made possible through the use of the Kronecker product of matrices, which results in a compact formulation. The response of the aforementioned systems to multiple sinusoidal excitations is also studied. Moreover, formulas are given for various system operations (e.g., addition, cascading, inversion, and feedback), which can be used to describe a complex system as an interconnection of simple subsystems.

I. INTRODUCTION

Communication, control, and instrumentation systems employ components, such as amplifiers and mixers, which are inherently nonlinear. Even when the nonlinearities are mild, as is often the case, they can produce bothersome signal distortion that limits the system performance. The nonlinear components themselves, and the other linear components used in the system, are generally frequency-dependent, i.e., they have memory. Numerous studies are available in the literature for the analysis of mildly nonlinear systems with memory through the use of Volterra-series expansions.¹⁻²¹ The classic paper by Bedrosian and Rice,⁷ the recent paper by Chua and Ng,¹⁴ and the book by Weiner and Spina¹⁷ cover that subject very thoroughly. Also, the paper by Gopal, Njakhla, Singhal, and Vlach¹² is interesting in that it evaluates the range of accuracy of the Volterra-series approach by comparing it with a nearly exact, but quite involved, method of analysis. The book¹⁸ and paper¹⁹ by Schetzen deal mainly with random inputs. The condi-

tions for the existence of a Volterra-series representation have recently been studied rigorously by Sandberg.²¹

For the most part, the studies mentioned above are limited to systems with one input and one output, i.e., "scalar" systems. This scalar representation is usually not easily applicable to Multiple-Input, Multiple-Output (MIMO) systems. Such systems include, for example, nonlinear multiport networks, multichannel communication systems, and transmitting or receiving systems employing multibeam antennas. In principle, one can represent these systems by a set of dependent scalar Volterra equations. This was done, for example, in the papers by Narayanan³ and by Bussgang, Ehrman, and Graham,⁹ where node equations were used to analyze nonlinear, two-port network models of bipolar transistor amplifiers. This method of analysis is tractable only when the numbers of nodes and of nonlinear elements in the network are small. For example, when the above authors considered the analysis of two-stage transistor amplifiers, they were forced by the complexity of the cascade equations involved to assume that the interaction between the stages, i.e., the loading effect of one stage on the other, is linear. While this might have been a reasonable approximation in their particular case, it is not valid in general. A symbolic matrix inversion algorithm that simplifies the computational aspects of the nodal method of analysis was recently discussed by Thapar and Leon.^{15,16}

To conveniently handle the problem of two-port networks, or to analyze nonlinear multiport networks in general, one needs to use a black-box representation of the network, as is usually done in linear networks. For example, consider a nonlinear, two-port network, which has two independent port variables (e.g., the port currents) and two dependent port variables (e.g., the port voltages). One should be able to express the latter variables in terms of the former (e.g., by a nonlinear impedance representation). Furthermore, one should be able to perform transformations among various network representations (e.g., from impedance to cascade parameters), and to carry out the computations involved in interconnecting several networks together to form a complex network (e.g., through cascading). The same operations are also needed in the analysis of other nonlinear MIMO systems.

The purpose of this paper is to develop a method for analyzing mildly nonlinear MIMO systems with memory. This method, which employs Volterra-series whose kernels are two-dimensional matrices, facilitates the systematic performance of various useful system operations, such as addition, cascading, inversion, and feedback. The application of the results of this study to the analysis of mildly nonlinear multiport networks will be the subject of a future paper.

Actually, Weiner and Naditch,¹⁰ and Gopal, Nakhla, Singhal, and Vlach¹² used multidimensional arrays of Volterra kernels to represent

nonlinear, two-port networks. The same was suggested by Chua and Ng¹⁴ for extending their results to multiple-input systems. All of these analyses can also be generalized to multiport networks and other MIMO systems. The resulting notation is similar to the index notation discussed in the beginning of the next section and in Appendix A. This notation, though more natural in its initial formulation, turns out to be cumbersome when attempting to perform the aforementioned system operations.

II. REPRESENTATION OF NONLINEAR MEMORYLESS MIMO SYSTEMS

A nonlinear, memoryless scalar system is characterized by its instantaneous input-output transfer function. When this function is analytic, as is the usual case encountered in practice, it can be represented by the power-series expansion

$$w = P^{(1)}u + P^{(2)}u^2 + P^{(3)}u^3 + \dots, \quad (1)$$

where $u = u(t)$ is the input, $w = w(t)$ is the output, and $P^{(k)}$, $k = 1, 2, 3, \dots$, are system constants. The corresponding representation of a nonlinear, memoryless, MIMO system with n inputs, $u_j = u_j(t)$, $j = 1, 2, \dots, n$, and m outputs, $w_i = w_i(t)$, $i = 1, 2, \dots, m$, is

$$\begin{aligned} w_i = & \sum_{j_1=1}^n P_{ij_1}^{(1)}u_{j_1} + \sum_{j_1=1}^n \sum_{j_2=1}^n P_{ij_1j_2}^{(2)}u_{j_1}u_{j_2} \\ & + \sum_{j_1=1}^n \sum_{j_2=1}^n \sum_{j_3=1}^n P_{ij_1j_2j_3}^{(3)}u_{j_1}u_{j_2}u_{j_3} + \dots, \quad i = 1, 2, \dots, m, \end{aligned} \quad (2)$$

where $P_{ij_1 \dots j_k}^{(k)}$, $k = 1, 2, 3, \dots$, are $(k + 1)$ -dimensional $m \times n \times \dots \times n$ arrays of system constants. The notation used in (2) will be referred to as the "index notation." It is similar to that used in Refs. 10 and 12, but the superscripts and subscripts are interchanged. We now proceed to represent (2) in the "matrix notation."

Let

$$\mathbf{u} = \mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_n(t) \end{bmatrix}, \quad \mathbf{w} = \mathbf{w}(t) = \begin{bmatrix} w_1(t) \\ w_2(t) \\ \vdots \\ w_m(t) \end{bmatrix} \quad (3)$$

be the $n \times 1$ and $m \times 1$ input and output vectors, respectively. The first (i.e., linear) term in (2) can be written as an ordinary product of matrices in the form $\mathbf{w} = \mathbf{P}^{(1)} \cdot \mathbf{u}$, where $\mathbf{P}^{(1)}$ is the $m \times n$ matrix $[P_{ij}^{(1)}]$. We will now show that the remaining terms in (2) can also be written in a matrix form through the use of the Kronecker product of

matrices.²²⁻²⁴ Appendix B defines this product and gives some of its useful properties. Actually, Harper and Rugh¹¹ employed the Kronecker product in conjunction with state variables to study factorable, scalar, nonlinear systems. Also, Brockett^{25,26} used a reduced form of the Kronecker product (to be explained shortly) in the state-variable representation of scalar, time-varying, nonlinear systems that are linear in the control variable.

As is explained below, the elements of the $(k + 1)$ -dimensional, $m \times n \times \dots \times n$ arrays, $\{P_{i_1 \dots j_k}^{(k)}\}$, can be reorganized to form two-dimensional, $m \times n^k$ matrices, $\{\mathbf{P}^{(k)}\}$, such that (2) can be written in the matrix form

$$\mathbf{w} = \mathbf{P}^{(1)} \cdot \mathbf{u} + \mathbf{P}^{(2)} \cdot (\mathbf{u} \times \mathbf{u}) + \mathbf{P}^{(3)} \cdot (\mathbf{u} \times \mathbf{u} \times \mathbf{u}) + \dots, \quad (4)$$

where “ \times ” is the Kronecker-product sign. As mentioned in Appendix B, we will employ *left* Kronecker products.

To understand (4), we note from (61) that the k -fold Kronecker product $\mathbf{u} \times \mathbf{u} \times \dots \times \mathbf{u}$ results in an $n^k \times 1$ vector whose j th element is given by

$$[\mathbf{u} \times \mathbf{u} \times \dots \times \mathbf{u}]_j = u_{j_1} u_{j_2} \dots u_{j_k}, \quad (5)$$

where j_1, j_2, \dots, j_k are uniquely determined from

$$j = j_1 + n(j_2 - 1) + \dots + n^{k-1}(j_k - 1). \quad (6)$$

Thus, to make (4) equivalent to (2), the i - j element of the $m \times n^k$ matrix $\mathbf{P}^{(k)}$ should be given by

$$[\mathbf{P}^{(k)}]_{ij} = P_{i_1 \dots j_k}^{(k)}, \quad (7)$$

where j is given by (6).

For example, if $m = n = 2$, (5)-(7) give

$$\mathbf{u} \times \mathbf{u} = \begin{bmatrix} u_1 u_1 \\ u_2 u_1 \\ u_1 u_2 \\ u_2 u_2 \end{bmatrix}, \quad \mathbf{u} \times \mathbf{u} \times \mathbf{u} = \begin{bmatrix} u_1 u_1 u_1 \\ u_2 u_1 u_1 \\ u_1 u_2 u_1 \\ u_2 u_2 u_1 \\ u_1 u_1 u_2 \\ u_2 u_1 u_2 \\ u_1 u_2 u_2 \\ u_2 u_2 u_2 \end{bmatrix}, \quad (8)$$

$$\mathbf{P}^{(2)} = \begin{bmatrix} P_{111}^{(2)} & P_{121}^{(2)} & P_{112}^{(2)} & P_{122}^{(2)} \\ P_{211}^{(2)} & P_{221}^{(2)} & P_{212}^{(2)} & P_{222}^{(2)} \end{bmatrix}, \quad (9)$$

$$\mathbf{P}^{(3)} = \begin{bmatrix} P_{1111}^{(3)} & P_{1211}^{(3)} & P_{1121}^{(3)} & P_{1221}^{(3)} & P_{1112}^{(3)} & P_{1212}^{(3)} & P_{1122}^{(3)} & P_{1222}^{(3)} \\ P_{2111}^{(3)} & P_{2211}^{(3)} & P_{2121}^{(3)} & P_{2221}^{(3)} & P_{2112}^{(3)} & P_{2212}^{(3)} & P_{2122}^{(3)} & P_{2222}^{(3)} \end{bmatrix}. \quad (10)$$

Note that each of the Kronecker-product vectors given in (8) has redundant entries. In Brockett's notation cited earlier, these entries would be removed. For example, the 4×1 vector, $\mathbf{u} \times \mathbf{u}$, would be replaced by the 3×1 vector $[u_1^2, u_1 u_2, u_2^2]$, and the corresponding 2×4 matrix, $\mathbf{P}^{(2)}$, given in (9) would be reduced to a 2×3 matrix, etc. However, when the system has memory, no redundant entries occur, since, as can be seen for example from (11), one needs to evaluate Kronecker products of the form $\mathbf{u}(t_1) \times \mathbf{u}(t_2)$, etc., where $t_1 \neq t_2$.

In the remainder of the paper, we will employ the compact matrix notation used in (4) rather than the index notation used in (2). However, on some occasions, it is helpful to keep track of the interrelation between the two notations. Thus, some key equations in the paper are rewritten in Appendix A in the index notation.

III. REPRESENTATION OF NONLINEAR MIMO SYSTEMS WITH MEMORY

The usual Volterra-series expansion used to represent nonlinear, time-invariant, scalar systems with memory¹⁻²¹ can be generalized through the use of the notation of (4) to represent MIMO systems by the matrix equation

$$\begin{aligned} \mathbf{w}(t) = & \int_{-\infty}^{\infty} \mathbf{p}^{(1)}(\tau_1) \cdot \mathbf{u}(t - \tau_1) d\tau_1 \\ & + \iint_{-\infty}^{\infty} \mathbf{p}^{(2)}(\tau_1, \tau_2) \cdot [\mathbf{u}(t - \tau_1) \times \mathbf{u}(t - \tau_2)] d\tau_1 d\tau_2 \\ & + \iiint_{-\infty}^{\infty} \mathbf{p}^{(3)}(\tau_1, \tau_2, \tau_3) \\ & \quad \cdot [\mathbf{u}(t - \tau_1) \times \mathbf{u}(t - \tau_2) \times \mathbf{u}(t - \tau_3)] d\tau_1 d\tau_2 d\tau_3 \\ & + \dots, \end{aligned} \tag{11}$$

where $\mathbf{u}(t)$ and $\mathbf{w}(t)$ are, respectively, the $n \times 1$ and $m \times 1$ input and output vectors given by (3), and where $\mathbf{p}^{(k)}(\tau_1, \dots, \tau_k)$, $k = 1, 2, 3, \dots$, are two-dimensional, $m \times n^k$ matrices of system kernels. Note that if $\mathbf{p}^{(k)}(\tau_1, \dots, \tau_k) = \mathbf{P}^{(k)} \delta(\tau_1) \dots \delta(\tau_k)$, where $\delta(\tau)$ is the unit impulse function, then the system becomes memoryless, and (11) reduces to (4).

As is the case for linear systems, it is more convenient to represent (11) in the frequency domain. To do this, we introduce the dummy time variables, t_1, t_2, \dots, t_k , and rewrite the k th order output component in (11) as

$$\mathbf{w}^{(k)}(t_1, \dots, t_k) = \int_{-\infty}^{\infty} \dots \int \mathbf{p}^{(k)}(\tau_1, \dots, \tau_k) \cdot [\mathbf{u}(t_1 - \tau_1) \times \dots \times \mathbf{u}(t_k - \tau_k)] d\tau_1 \dots d\tau_k. \quad (12)^\dagger$$

Thus, (11) becomes

$$\mathbf{w}(t) = \mathbf{w}^{(1)}(t) + \mathbf{w}^{(2)}(t, t) + \mathbf{w}^{(3)}(t, t, t) + \dots \quad (13)$$

Now, we introduce the single-dimensional Fourier-transform pair

$$\mathbf{X}(f) = \int_{-\infty}^{\infty} \mathbf{x}(t) \exp(-j2\pi ft) dt, \quad (14a)$$

$$\mathbf{x}(t) = \int_{-\infty}^{\infty} \mathbf{X}(f) \exp(j2\pi ft) df, \quad (14b)$$

to represent the transformations $\mathbf{u}(t) \leftrightarrow \mathbf{U}(f)$ and $\mathbf{w}(t) \leftrightarrow \mathbf{W}(f)$. Similarly, we introduce the multi-dimensional Fourier-transform pair

$$\mathbf{Y}(f_1, \dots, f_k) = \int_{-\infty}^{\infty} \dots \int \mathbf{y}(t_1, \dots, t_k) \cdot \exp[-j2\pi(f_1 t_1 + \dots + f_k t_k)] dt_1 \dots dt_k, \quad (15a)$$

$$\mathbf{y}(t_1, \dots, t_k) = \int_{-\infty}^{\infty} \dots \int \mathbf{Y}(f_1, \dots, f_k) \cdot \exp[j2\pi(f_1 t_1 + \dots + f_k t_k)] df_1 \dots df_k, \quad (15b)$$

to represent the transformations $\mathbf{p}^{(k)}(\tau_1, \dots, \tau_k) \leftrightarrow \mathbf{P}^{(k)}(f_1, \dots, f_k)$ and $\mathbf{w}^{(k)}(t_1, \dots, t_k) \leftrightarrow \mathbf{W}^{(k)}(f_1, \dots, f_k)$. It can be shown from (14) and (15) that (12) can be written in the frequency domain as (see Refs. 1 through 4, 7 through 9, 14 and 17)

$$\mathbf{W}^{(k)}(f_1, \dots, f_k) = \mathbf{P}^{(k)}(f_1, \dots, f_k) \cdot [\mathbf{U}(f_1) \times \dots \times \mathbf{U}(f_k)]. \quad (16)^\dagger$$

The Fourier transform of the output becomes

$$\begin{aligned} \mathbf{W}(f) &= \mathbf{W}^{(1)}(f) + \int_{-\infty}^{\infty} \mathbf{W}^{(2)}(f_1, f - f_1) df_1 \\ &+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{W}^{(3)}(f_1, f_2, f - f_1 - f_2) df_1 df_2 + \dots \quad (17) \end{aligned}$$

[†] All equations in the paper marked by a dagger are rewritten in the index notation in Appendix A, where the same equation numbers are used.

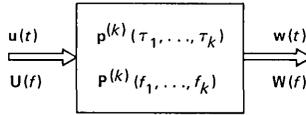


Fig. 1—A nonlinear, time-invariant, MIMO system with memory having n inputs and m outputs.

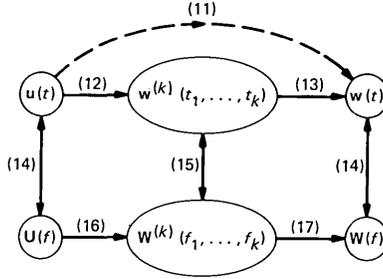


Fig. 2—Interrelations among the input, output components, and total output in the time and frequency domains for the nonlinear MIMO system of Fig. 1. The numbers in parentheses represent equation numbers in the text.

Note from (13) and (17) that the single-dimensional Fourier transform of $\mathbf{w}^{(k)}(t, \dots, t)$ is given by the k th term in (17), which is *not* equal to $\mathbf{W}^{(k)}(f, \dots, f)$ unless $k = 1$.

A schematic diagram of the system represented by (11) through (17) is given in Fig. 1. The interrelations among the input, the output components, and the total output in the time and frequency domains, and the corresponding equation numbers, are indicated in the flow-chart of Fig. 2.

IV. KERNEL SYMMETRIZATION

The representation of the response of a nonlinear scalar system to sinusoidal and Gaussian excitations is greatly simplified if each of the kernels, $P^{(k)}(f_1, \dots, f_k)$, or equivalently, $p^{(k)}(\tau_1, \dots, \tau_k)$, is a symmetric function of its arguments.^{7-9,14} The generalization of this symmetry requirement to nonlinear MIMO systems is somewhat more involved. Following the reasoning given in the aforementioned references, one can show that it is the output components, $\mathbf{W}^{(k)}(f_1, \dots, f_k)$ given by (16), or equivalently, $\mathbf{w}^{(k)}(t_1, \dots, t_k)$ given by (12), that are required to be symmetric functions of their arguments. For example, for $k = 2$, it is required that $\mathbf{W}^{(2)}(f_1, f_2) = \mathbf{W}^{(2)}(f_2, f_1)$; and thus, from (16),

$$\mathbf{P}^{(2)}(f_1, f_2) \cdot (\mathbf{u}_1 \times \mathbf{u}_2) = \mathbf{P}^{(2)}(f_2, f_1) \cdot (\mathbf{u}_2 \times \mathbf{u}_1), \quad (18a)$$

where $\mathbf{U}(f_i)$ is replaced by \mathbf{u}_i for generality. Similarly, for $k = 3$, it is required that $\mathbf{W}^{(3)}(f_1, f_2, f_3) = \mathbf{W}^{(3)}(f_\alpha, f_\beta, f_\gamma)$; and thus, from (16),

$$\mathbf{P}^{(3)}(f_1, f_2, f_3) \cdot (\mathbf{u}_1 \times \mathbf{u}_2 \times \mathbf{u}_3) = \mathbf{P}^{(3)}(f_\alpha, f_\beta, f_\gamma) \cdot (\mathbf{u}_\alpha \times \mathbf{u}_\beta \times \mathbf{u}_\gamma), \quad (18b)$$

where α, β, γ assume all permutations of 1, 2, 3. For a scalar system, (18a) and (18b) are indeed equivalent to requiring the corresponding system kernels to be symmetric functions of their arguments, as mentioned above.

To find the symmetry requirement implied by (18) on the kernels of a MIMO system, we need to introduce the $n^2 \times n^2$ "reversing" matrix, \mathbf{R} , and the six $n^3 \times n^3$ "permutation" matrices, $\Phi_{\alpha\beta\gamma}$, where α, β, γ assume all permutations of 1, 2, 3. These matrices have properties such that if $\mathbf{u}_1, \mathbf{u}_2$ and \mathbf{u}_3 are $n \times 1$ vectors, then

$$\mathbf{R} \cdot (\mathbf{u}_1 \times \mathbf{u}_2) = \mathbf{u}_2 \times \mathbf{u}_1, \quad (19)$$

$$\Phi_{\alpha\beta\gamma} \cdot (\mathbf{u}_1 \times \mathbf{u}_2 \times \mathbf{u}_3) = \mathbf{u}_\alpha \times \mathbf{u}_\beta \times \mathbf{u}_\gamma. \quad (20)$$

Appendix C defines these matrices and gives some of their useful properties.

Finally, (18) through (20) give the required symmetry conditions of the kernels as

$$\mathbf{P}^{(2)}(f_1, f_2) = \mathbf{P}^{(2)}(f_2, f_1) \cdot \mathbf{R}, \quad (21a)^\dagger$$

$$\mathbf{P}^{(3)}(f_1, f_2, f_3) = \mathbf{P}^{(3)}(f_\alpha, f_\beta, f_\gamma) \cdot \Phi_{\alpha\beta\gamma}. \quad (21b)^\dagger$$

The generalization of (21) to higher-order kernels requires the introduction of permutation matrices of more than three indices.

If the given system kernels, say, $\hat{\mathbf{P}}^{(2)}(f_1, f_2)$ and $\hat{\mathbf{P}}^{(3)}(f_1, f_2, f_3)$, are unsymmetric, they can be symmetrized, i.e., made to satisfy (21), through the use of the relations

$$\mathbf{P}^{(2)}(f_1, f_2) = \frac{1}{2} [\hat{\mathbf{P}}^{(2)}(f_1, f_2) + \hat{\mathbf{P}}^{(2)}(f_2, f_1) \cdot \mathbf{R}], \quad (22a)^\dagger$$

$$\mathbf{P}^{(3)}(f_1, f_2, f_3) = \frac{1}{6} \sum_{\alpha, \beta, \gamma} \hat{\mathbf{P}}^{(3)}(f_\alpha, f_\beta, f_\gamma) \cdot \Phi_{\alpha\beta\gamma}, \quad (22b)^\dagger$$

where the summation is performed over α, β, γ assuming all 6 permutations of 1, 2, 3. These symmetrization relations are generalizations of those discussed in Refs. 7, 9, and 14 for scalar kernels.

V. RESPONSE TO SINUSOIDAL EXCITATION

The response of a nonlinear scalar system to multiple-sinusoidal excitation has been studied by several authors including Bedrosian and Rice,⁷ Goldman,⁸ and Chua and Ng.¹⁴ Here we generalize some of their results to nonlinear MIMO systems.

5.1 Multiple-exponential excitation

Let the input vector be

$$\mathbf{u}(t) = \sum_{i=1}^l \mathbf{u}_i \exp(j2\pi f_i t), \quad (23)$$

where the \mathbf{u}_i 's are time-independent, complex, $n \times 1$ vectors. The Fourier transform of $\mathbf{u}(t)$ is

$$\mathbf{U}(f) = \sum_{i=1}^l \mathbf{u}_i \delta(f - f_i). \quad (24)$$

Substituting (24) into (16), and using (15b), one obtains the k th order output component

$$w^{(k)}(t) \equiv \mathbf{w}^{(k)}(t, \dots, t) = \sum_{i_1=1}^l \dots \sum_{i_k=1}^l \{[\mathbf{P}^{(k)}(f_{i_1}, \dots, f_{i_k}) \cdot (\mathbf{u}_{i_1} \times \dots \times \mathbf{u}_{i_k})] \exp[j2\pi(f_{i_1} + \dots + f_{i_k})t]\}. \quad (25)$$

Finally, the output, $\mathbf{w}(t)$, is obtained from (13), i.e., by summing $\mathbf{w}^{(k)}(t)$ from $k = 1$ up to any desired order. Note that (25) is valid whether or not the system kernels are symmetric.

5.2 Single-frequency excitation

Let the (real) $n \times 1$ input vector be

$$\begin{aligned} \mathbf{u}(t) &= \text{Real}[\mathbf{a} \exp(j2\pi ft)] \\ &= \frac{1}{2} \mathbf{a} \exp(j2\pi ft) + \frac{1}{2} \mathbf{a}^* \exp(-j2\pi ft), \end{aligned} \quad (26)$$

where the asterisk refers to complex conjugation. Comparing (26) to (23), one obtains $l = 2$, $\mathbf{u}_1 = \frac{1}{2} \mathbf{a}$, $\mathbf{u}_2 = \frac{1}{2} \mathbf{a}^*$, $f_1 = f$, and $f_2 = -f$. Thus, using (25), and assuming that the kernels are symmetric, i.e., that (18) is satisfied, one obtains the following expressions for the various k th order output components:

$$\begin{aligned} \mathbf{w}^{(1)}(t) &= \frac{1}{2} [\mathbf{P}^{(1)}(f) \cdot \mathbf{a}] \exp(j2\pi ft) \\ &\quad + \frac{1}{2} [\mathbf{P}^{(1)}(-f) \cdot \mathbf{a}^*] \exp(-j2\pi ft). \end{aligned} \quad (27a)$$

$$\begin{aligned} \mathbf{w}^{(2)}(t) &= \frac{1}{2} [\mathbf{P}^{(2)}(f, -f) \cdot (\mathbf{a} \times \mathbf{a}^*)] \leftarrow (d - c \text{ term}) \\ &\quad + \frac{1}{4} [\mathbf{P}^{(2)}(f, f) \cdot (\mathbf{a} \times \mathbf{a})] \exp[j2\pi(2f)t] \\ &\quad + \frac{1}{4} [\mathbf{P}^{(2)}(-f, -f) \cdot (\mathbf{a}^* \times \mathbf{a}^*)] \exp[-j2\pi(2f)t]. \end{aligned} \quad (27b)$$

$$\begin{aligned} \mathbf{w}^{(3)}(t) &= \frac{3}{8} [\mathbf{P}^{(3)}(f, f, -f) \cdot (\mathbf{a} \times \mathbf{a} \times \mathbf{a}^*)] \exp(j2\pi ft) \\ &\quad + \frac{3}{8} [\mathbf{P}^{(3)}(-f, -f, f) \cdot (\mathbf{a}^* \times \mathbf{a}^* \times \mathbf{a})] \exp(-j2\pi ft) \\ &\quad + \frac{1}{8} [\mathbf{P}^{(3)}(f, f, f) \cdot (\mathbf{a} \times \mathbf{a} \times \mathbf{a})] \exp[j2\pi(3f)t] \\ &\quad + \frac{1}{8} [\mathbf{P}^{(3)}(-f, -f, -f) \cdot (\mathbf{a}^* \times \mathbf{a}^* \times \mathbf{a}^*)] \exp[-j2\pi(3f)t]. \end{aligned} \quad (27c)$$

Note that the asterisks on the \mathbf{a} 's correspond in number and location

to the negative signs in the frequency arguments of the associated kernels.

If the system is real, i.e., if $\mathbf{p}^{(k)}(\tau_1, \dots, \tau_k)$, $k = 1, 2, 3, \dots$, are real, then it can be shown from (15) that

$$\mathbf{P}^{(k)}(f_1, \dots, f_k) = [\mathbf{P}^{(k)}(-f_1, \dots, -f_k)]^*. \quad (28)$$

As expected, (28) implies that all the output components given in (27) are real. In that case, it can be shown through generalizing (27) that the total m th harmonic output term, $\mathbf{w}_m(t)$, $m = 0, 1, 2, \dots$, is given by

$$\mathbf{w}_m(t) = \epsilon_m \text{Real} \left[\exp(j2\pi mf) \sum_{k=m, m+2, \dots} \left\{ 2^{-k} \binom{k}{2} \cdot \underbrace{\mathbf{P}^{(k)}(f, \dots, f, -f, \dots, -f)}_{\substack{(k+m)/2 \\ (k-m)/2}} \cdot [\mathbf{a}^{[(k+m)/2]} \times (\mathbf{a}^*)^{[(k-m)/2]}] \right\} \right], \quad (29)$$

where $\mathbf{a}^{[l]}$ is the l -fold Kronecker product $\mathbf{a} \times \dots \times \mathbf{a}$, ϵ_m is the Neuman factor (which is equal to 1 when $m = 0$, and is equal to 2 when $m \neq 0$), and $\mathbf{P}^{(0)}$ is defined to be zero.

Because of the symmetry conditions of (21), the kernels used in (27) and (29) satisfy the relations

$$\mathbf{P}^{(2)}(f, f) = \mathbf{P}^{(2)}(f, f) \cdot \mathbf{R}, \quad (30a)^\dagger$$

$$\mathbf{P}^{(2)}(f, -f) = [\mathbf{P}^{(2)}(f, -f)]^* \cdot \mathbf{R}, \quad (30b)^\dagger$$

$$\mathbf{P}^{(3)}(f, f, -f) = \mathbf{P}^{(3)}(f, f, -f) \cdot \Phi_{213}, \quad (30c)^\dagger$$

$$\mathbf{P}^{(3)}(f, f, f) = \mathbf{P}^{(3)}(f, f, f) \cdot \Phi_{\alpha\beta\gamma}. \quad (30d)^\dagger$$

In addition to the kernel symmetry requirement, (30b) is based on the assumption that the system is real, i.e., that (28) is satisfied. The implication of (30) is that the elements of each of the system kernels are not all independent. For example, if $n = 2$, (30a) through (30d) imply, respectively, that (i) columns 2 and 3 of $\mathbf{P}^{(2)}(f, f)$ are equal; (ii) column 2 of $\mathbf{P}^{(2)}(f, -f)$ is the complex conjugate of column 3, and columns 1 and 4 are real; (iii) columns 2 and 3 of $\mathbf{P}^{(3)}(f, f, -f)$ are equal, and so are columns 6 and 7; and (iv) columns 2, 3, and 5 of $\mathbf{P}^{(3)}(f, f, f)$ are equal, and so are columns 4, 6, and 7. It is worth mentioning that (30a) and (30d), respectively, would also be satisfied by $\mathbf{P}^{(2)}$ and $\mathbf{P}^{(3)}$ of the memoryless system represented by (4).

5.3 Two-frequency excitation

Let the (real) $n \times 1$ input vector be

$$\mathbf{u}(t) = \text{Real}[\mathbf{a} \exp(j2\pi f_a t) + \mathbf{b} \exp(j2\pi f_b t)]. \quad (31)$$

We assume that the system is real, and that the kernels are symmetric, i.e., that (28) and (18) are satisfied. One can use (25) to obtain the output corresponding to (32) by following the same steps used to derive (29). The leading terms at some of the various output frequencies are:

$$\mathbf{w}(t)|_{d-c} \approx \frac{1}{2}[\mathbf{P}^{(2)}(f_a, -f_a) \cdot (\mathbf{a} \times \mathbf{a}^*) + \mathbf{P}^{(2)}(f_b, -f_b) \cdot (\mathbf{b} \times \mathbf{b}^*)]. \quad (32a)$$

$$\begin{aligned} \mathbf{w}(t)|_{f_a} \approx \text{Real}\{ & \exp(j2\pi f_a t)[\mathbf{P}^{(1)}(f_a) \cdot \mathbf{a} \\ & + \frac{3}{4}\mathbf{P}^{(3)}(f_a, f_a, -f_a) \cdot (\mathbf{a} \times \mathbf{a} \times \mathbf{a}^*) \\ & + \frac{3}{2}\mathbf{P}^{(3)}(f_a, f_b, -f_b) \cdot (\mathbf{a} \times \mathbf{b} \times \mathbf{b}^*)]\}. \end{aligned} \quad (32b)$$

$$\begin{aligned} \mathbf{w}(t)|_{2f_a-f_b} \approx \text{Real}\{ & \frac{3}{4} \exp[j2\pi(2f_a - f_b)t] \\ & \cdot \mathbf{P}^{(3)}(f_a, f_a, -f_b) \cdot (\mathbf{a} \times \mathbf{a} \times \mathbf{b}^*)\}. \end{aligned} \quad (32c)$$

$$\begin{aligned} \mathbf{w}(t)|_{lf_a \pm mf_b} \approx \text{Real} \left\{ & 2^{-(l+m-1)} \binom{l+m}{l} \exp[j2\pi(lf_a \pm mf_b)t] \right. \\ & \left. \cdot \mathbf{P}^{(l+m)}(\underbrace{f_a, \dots, f_a}_l, \pm \underbrace{f_b, \dots, f_b}_m) \cdot [\mathbf{a}^{[l]} \times (\mathbf{b}^\pm)^{[m]}] \right\}, \end{aligned} \quad (32d)$$

where $l > 0$ and $m \geq 0$, and where we defined $\mathbf{b}^+ \equiv \mathbf{b}$ and $\mathbf{b}^- \equiv \mathbf{b}^*$.

5.4 Three-frequency excitation

Let the (real) $n \times 1$ input vector be

$$\mathbf{u}(t) = \text{Real}[\mathbf{a} \exp(j2\pi f_a t) + \mathbf{b} \exp(j2\pi f_b t) + \mathbf{c} \exp(j2\pi f_c t)]. \quad (33)$$

Again, we assume that the system is real, and that the kernels are symmetric. Following the same steps used to derive (29) and (32), one can obtain the following leading terms at some of the various output frequencies:

$$\begin{aligned} \mathbf{w}(t)|_{d-c} \approx \frac{1}{2}[\mathbf{P}^{(2)}(f_a, -f_a) \cdot (\mathbf{a} \times \mathbf{a}^*) + \mathbf{P}^{(2)}(f_b, -f_b) \\ \cdot (\mathbf{b} \times \mathbf{b}^*) + \mathbf{P}^{(2)}(f_c, -f_c) \cdot (\mathbf{c} \times \mathbf{c}^*)]. \end{aligned} \quad (34a)$$

$$\begin{aligned} \mathbf{w}(t)|_{f_a} \approx \text{Real}\{ & \exp(j2\pi f_a t)[\mathbf{P}^{(1)}(f_a) \cdot \mathbf{a} + \frac{3}{4}\mathbf{P}^{(3)}(f_a, f_a, -f_a) \\ & \cdot (\mathbf{a} \times \mathbf{a} \times \mathbf{a}^*) + \frac{3}{2}\mathbf{P}^{(3)}(f_a, f_b, -f_b) \cdot (\mathbf{a} \times \mathbf{b} \times \mathbf{b}^*) \\ & + \frac{3}{2}\mathbf{P}^{(3)}(f_a, f_c, -f_c) \cdot (\mathbf{a} \times \mathbf{c} \times \mathbf{c}^*)]\}. \end{aligned} \quad (34b)$$

$$\begin{aligned} \mathbf{w}(t)|_{f_a+f_b-f_c} \approx \text{Real}\{ & \frac{3}{2} \exp[j2\pi(f_a + f_b - f_c)t] \\ & \cdot \mathbf{P}^{(3)}(f_a, f_b, -f_c) \cdot (\mathbf{a} \times \mathbf{b} \times \mathbf{c}^*)\}. \end{aligned} \quad (34c)$$

$$\begin{aligned}
\mathbf{w}(t) &|_{kf_a \pm lf_b \oplus mf_c} \\
&\approx \text{Real} \left\{ 2^{-(k+l+m-1)} \frac{(k+l+m)!}{k!l!m!} \exp[j2\pi(kf_a \pm lf_b \oplus mf_c)t] \right. \\
&\quad \cdot \mathbf{P}^{(k+l+m)} \left(\underbrace{f_a, \dots, f_a}_k, \underbrace{\pm f_b, \dots, \pm f_b}_l, \underbrace{\oplus f_c, \dots, \oplus f_c}_m \right) \\
&\quad \left. \cdot [\mathbf{a}^{[k]} \times (\mathbf{b}^{\pm})^{[l]} \times (\mathbf{c}^{\oplus})^{[m]}] \right\}, \tag{34d}
\end{aligned}$$

where $k, l, m \geq 0$, but at least one of them being nonzero, and where the sign symbols \pm and \oplus are each consistent throughout the equation, but are otherwise independent.

VI. SYSTEM OPERATIONS

6.1 Operational notation

Let the input-output relations given in (11) through (17) be written symbolically as

$$\mathbf{W}_m = \{\mathbf{P}_{m,n}^{(k)}\} o \mathbf{U}_n, \tag{35}$$

where “ o ” means “operating on.” The frequency dependence has been omitted for simplicity. The subscripts n and m are included to emphasize the numbers of inputs and outputs. On some occasions, these subscripts will be eliminated.

If the system is linear, i.e., if $\mathbf{P}^{(k)} = \mathbf{0}$ for $k > 1$, the operation in (35) reduces to an ordinary matrix product. Thus,

$$\mathbf{W} = \{\mathbf{P}^{(1)}\} o \mathbf{U} = \mathbf{P}^{(1)} \cdot \mathbf{U}. \tag{36}$$

The operational notation of (35), and the three system operations of addition, cascading, and inversion, which are discussed in the next three subsections, form an algebraic structure that permits a shorthand description of complex interconnections of nonlinear MIMO systems. The laws of this algebra¹ are identical to the algebra of linear systems (i.e., the algebra of matrices) with two important exceptions—the left distributive law does not hold, and the laws of multiplication by a scalar constant are more complex.

6.2 Addition

Two systems, $\{\mathbf{P}_{m,n}^{(k)}\}$ and $\{\mathbf{Q}_{m,n}^{(k)}\}$, having the same number of inputs, n , and the same number of outputs, m , are said to be “added” if they share the same input vector, \mathbf{U}_n , and if their respective outputs are added to form the final output vector, \mathbf{W}_m . This operation, which is shown schematically in Fig. 3, is represented by

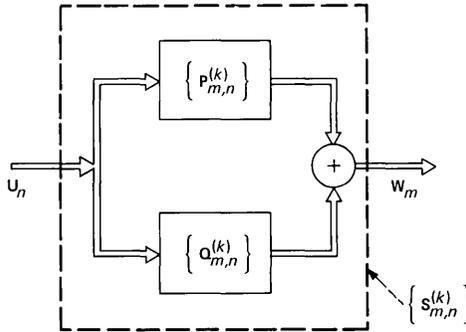


Fig. 3—A schematic representation of the addition operation $\{S_{m,n}^{(k)}\} = \{P_{m,n}^{(k)}\} + \{Q_{m,n}^{(k)}\}$.

$$\begin{aligned}
 W_m &= \{P_{m,n}^{(k)}\} \circ U_n + \{Q_{m,n}^{(k)}\} \circ U_n \\
 &= [\{P_{m,n}^{(k)}\} + \{Q_{m,n}^{(k)}\}] \circ U_n \\
 &= \{S_{m,n}^{(k)}\} \circ U_n.
 \end{aligned} \tag{37}$$

The kernels of the sum system,

$$\{S_{m,n}^{(k)}\} = \{P_{m,n}^{(k)}\} + \{Q_{m,n}^{(k)}\}, \tag{38}$$

are given by

$$S^{(k)}(f_1, \dots, f_k) = P^{(k)}(f_1, \dots, f_k) + Q^{(k)}(f_1, \dots, f_k), \tag{39}$$

where the plus sign refers to matrix addition.

One can define a subtraction operation in an obvious manner. A multiplication operation,^{1,14} which is more involved, can also be defined.

6.3 Cascading

When the output vector, W_m , of a system, $\{P_{m,n}^{(k)}\}$, is used as an input vector to a second system, $\{Q_{l,m}^{(k)}\}$, whose output vector is X_l , the two systems are said to be in "cascade." This operation, which is shown schematically in Fig. 4, is represented by

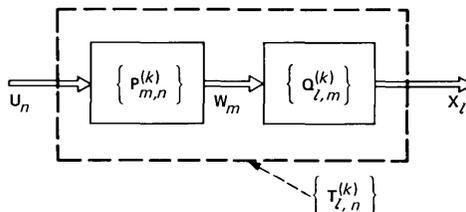


Fig. 4—A schematic representation of the cascade operation $\{T_{l,n}^{(k)}\} = \{Q_{l,m}^{(k)}\} * \{P_{m,n}^{(k)}\}$.

$$\begin{aligned}
\mathbf{X}_l &= \{\mathbf{Q}_{l,m}^{(k)}\} \circ \mathbf{W}_m \\
&= \{\mathbf{Q}_{l,m}^{(k)}\} \circ [\{\mathbf{P}_{m,n}^{(k)}\} \circ \mathbf{U}_n] \\
&= [\{\mathbf{Q}_{l,m}^{(k)}\} * \{\mathbf{P}_{m,n}^{(k)}\}] \circ \mathbf{U}_n \\
&= \{\mathbf{T}_{l,n}^{(k)}\} \circ \mathbf{U}_n, \tag{40}
\end{aligned}$$

where the asterisk refers to the cascade operation. The kernels of the cascade system,

$$\{\mathbf{T}_{l,n}^{(k)}\} = \{\mathbf{Q}_{l,m}^{(k)}\} * \{\mathbf{P}_{m,n}^{(k)}\}, \tag{41}$$

can be obtained by substituting the output expression of the first system into the system equations of the second system, as was done in Refs. 3 and 9 to derive the cascade relations of scalar systems. This procedure is straightforward, but somewhat tedious. A simpler approach is to employ the harmonic probing method discussed in Refs. 7 and 9, and the expression for the response of nonlinear vector systems to multiple-exponential excitation given in (25). The resulting relations for the cascade kernels are

$$\mathbf{T}^{(1)}(f_1) = \mathbf{Q}^{(1)}(f_1) \cdot \mathbf{P}^{(1)}(f_1), \tag{42a}^\dagger$$

$$\begin{aligned}
\mathbf{T}^{(2)}(f_1, f_2) &= \mathbf{Q}^{(1)}(f_1 + f_2) \cdot \mathbf{P}^{(2)}(f_1, f_2) \\
&\quad + \mathbf{Q}^{(2)}(f_1, f_2) \cdot [\mathbf{P}^{(1)}(f_1) \times \mathbf{P}^{(1)}(f_2)], \tag{42b}^\dagger
\end{aligned}$$

$$\begin{aligned}
\hat{\mathbf{T}}^{(3)}(f_1, f_2, f_3) &= \mathbf{Q}^{(1)}(f_1 + f_2 + f_3) \cdot \mathbf{P}^{(3)}(f_1, f_2, f_3) \\
&\quad + \mathbf{Q}^{(2)}(f_1, f_2 + f_3) \cdot [\mathbf{P}^{(1)}(f_1) \times \mathbf{P}^{(2)}(f_2, f_3)] \\
&\quad + \mathbf{Q}^{(2)}(f_1 + f_2, f_3) \cdot [\mathbf{P}^{(2)}(f_1, f_2) \times \mathbf{P}^{(1)}(f_3)] \\
&\quad + \mathbf{Q}^{(3)}(f_1, f_2, f_3) \cdot [\mathbf{P}^{(1)}(f_1) \times \mathbf{P}^{(1)}(f_2) \times \mathbf{P}^{(1)}(f_3)]. \tag{42c}^\dagger
\end{aligned}$$

A generalization of (42) for arbitrary k is given in Appendix D.

If the kernels of the cascaded systems are symmetric, i.e., satisfy (21), then it can be shown that the resulting second-order kernel given by (42b) is also symmetric. However, the resulting third-order kernel given by (42c) is not symmetric. This fact is indicated by the presence of the circumflexes.

As mentioned in Section IV, it is desirable to deal with symmetric kernels. Thus, using the symmetrization relation given in (22b), assuming that the kernels of the cascaded systems are symmetric, and employing the properties of the reversing and permutation matrices given in Appendix C, one obtains the symmetric form of (42c) as

$$\begin{aligned}
\mathbf{T}^{(3)}(f_1, f_2, f_3) &= \mathbf{Q}^{(1)}(f_1 + f_2 + f_3) \cdot \mathbf{P}^{(3)}(f_1, f_2, f_3) \\
&\quad + \frac{2}{3} \{ \mathbf{Q}^{(2)}(f_1, f_2 + f_3) \cdot [\mathbf{P}^{(1)}(f_1) \times \mathbf{P}^{(2)}(f_2, f_3)] \\
&\quad + \mathbf{Q}^{(2)}(f_2, f_3 + f_1) \cdot [\mathbf{P}^{(1)}(f_2) \times \mathbf{P}^{(2)}(f_3, f_1)] \cdot \Phi_{231} \\
&\quad + \mathbf{Q}^{(2)}(f_1 + f_2, f_3) \cdot [\mathbf{P}^{(2)}(f_1, f_2) \times \mathbf{P}^{(1)}(f_3)] \} \\
&\quad + \mathbf{Q}^{(3)}(f_1, f_2, f_3) \cdot [\mathbf{P}^{(1)}(f_1) \times \mathbf{P}^{(1)}(f_2) \times \mathbf{P}^{(1)}(f_3)], \tag{42c}^\dagger
\end{aligned}$$

where Φ_{231} is defined in (68) and (69).

If the first system, $\{\mathbf{P}_{m,n}^{(k)}\}$, is linear, (42) reduces to

$$\mathbf{T}^{(k)}(f_1, \dots, f_k) = \mathbf{Q}^{(k)}(f_1, \dots, f_k) \cdot [\mathbf{P}^{(1)}(f_1) \times \dots \times \mathbf{P}^{(1)}(f_k)]. \quad (43)$$

On the other hand, if the second system, $\{\mathbf{Q}_{l,m}^{(k)}\}$, is linear, (42) reduces to

$$\mathbf{T}^{(k)}(f_1, \dots, f_k) = \mathbf{Q}^{(k)}(f_1 + \dots + f_k) \cdot \mathbf{P}^{(k)}(f_1, \dots, f_k). \quad (44)$$

6.4 Inversion

Let the numbers of inputs and outputs in the system represented by (35) be equal, i.e., $m = n$. Suppose that it is required to find the input vector, \mathbf{U}_n , in terms of the output vector, \mathbf{W}_n . This inversion operation is represented by

$$\mathbf{U}_n = \{\mathbf{P}_{n,n}^{(k)}\}^{-1} \circ \mathbf{W}_n = \{\mathbf{Q}_{n,n}^{(k)}\} \circ \mathbf{W}_n. \quad (45)$$

To find the kernels of the inverse system,

$$\{\mathbf{Q}_{n,n}^{(k)}\} = \{\mathbf{P}_{n,n}^{(k)}\}^{-1}, \quad (46)$$

it is helpful to use the interpretation given in Fig. 5, which defines the inversion operation in terms of the cascade operation and the identity system, $\{\mathbf{1}_n\}$, where $\mathbf{1}_n$ is the $n \times n$ identity matrix. Thus, applying the symmetric cascade relations of (42) to Fig. 5b by interchanging the roles of \mathbf{P} and \mathbf{Q} , setting $\mathbf{T}^{(1)} = \mathbf{1}_n$, and $\mathbf{T}^{(k)} = \mathbf{0}$ for $k > 1$, and solving for $\mathbf{Q}^{(k)}$, one obtains the symmetric inversion relations

$$\mathbf{Q}^{(1)}(f_1) = [\mathbf{P}^{(1)}(f_1)]^{-1}, \quad (47a)$$

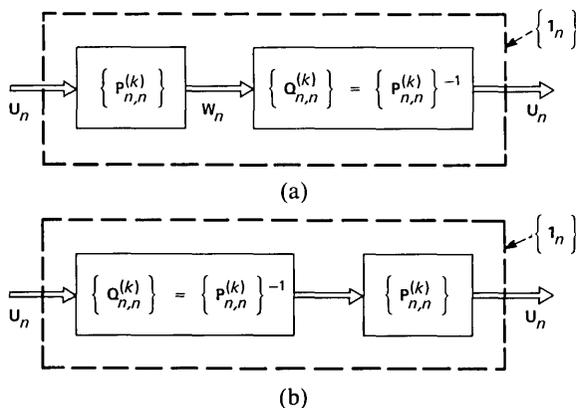


Fig. 5—Two equivalent interpretations of the inversion operation $\{\mathbf{Q}_{n,n}^{(k)}\} = \{\mathbf{P}_{n,n}^{(k)}\}^{-1}$: (a) $\{\mathbf{Q}_{n,n}^{(k)}\} * \{\mathbf{P}_{n,n}^{(k)}\} = \{\mathbf{1}_n\}$, and (b) $\{\mathbf{P}_{n,n}^{(k)}\} * \{\mathbf{Q}_{n,n}^{(k)}\} = \{\mathbf{1}_n\}$, where $\{\mathbf{1}_n\}$ is the identity system.

$$\mathbf{Q}^{(2)}(f_1, f_2) = -\mathbf{Q}^{(1)}(f_1 + f_2) \cdot \mathbf{P}^{(2)}(f_1, f_2) \cdot [\mathbf{Q}^{(1)}(f_1) \times \mathbf{Q}^{(1)}(f_2)], \quad (47b)$$

$$\begin{aligned} \mathbf{Q}^{(3)}(f_1, f_2, f_3) = & -\mathbf{Q}^{(1)}(f_1 + f_2 + f_3) \cdot \{ \mathbf{P}^{(2)}(f_1, f_2 + f_3) \\ & \cdot [\mathbf{Q}^{(1)}(f_1) \times \mathbf{Q}^{(2)}(f_2, f_3)] \\ & + \mathbf{P}^{(2)}(f_2, f_3 + f_1) \cdot [\mathbf{Q}^{(1)}(f_2) \times \mathbf{Q}^{(2)}(f_3, f_1)] \cdot \Phi_{231} \\ & + \mathbf{P}^{(2)}(f_1 + f_2, f_3) \cdot [\mathbf{Q}^{(2)}(f_1, f_2) \times \mathbf{Q}^{(1)}(f_3)] \} \\ & + \mathbf{P}^{(3)}(f_1, f_2, f_3) \cdot [\mathbf{Q}^{(1)}(f_1) \times \mathbf{Q}^{(1)}(f_2) \times \mathbf{Q}^{(1)}(f_3)]. \quad (47c) \end{aligned}$$

Note that the inverse system exists if and only if $\mathbf{P}^{(1)}(f)$ is nonsingular.

6.5 Feedback

As an application of the three system operations discussed in the previous subsections, consider the nonlinear, feedback, MIMO system shown schematically in Fig. 6, where both the forward, $\{\mathbf{P}_{m,n}^{(k)}\}$, and reverse, $\{\mathbf{Q}_{n,m}^{(k)}\}$, branches are nonlinear. Using the operational notation of (35), one obtains

$$\mathbf{W}_m = \{\mathbf{P}_{m,n}^{(k)}\} \circ \mathbf{X}_n, \quad (48a)$$

$$\mathbf{X}_n = \mathbf{U}_n + \{\mathbf{Q}_{n,m}^{(k)}\} \circ \mathbf{W}_m, \quad (48b)$$

where \mathbf{U}_n , \mathbf{X}_n and \mathbf{W}_m are the $n \times 1$ input vector, the $n \times 1$ intermediate vector, and the $m \times 1$ output vector, respectively. Substituting \mathbf{W}_m from (48a) into (48b), solving for \mathbf{X}_n in terms of \mathbf{U}_n , and substituting the result in (48a), one obtains the feedback system equation

$$\mathbf{W}_m = \{\mathbf{F}_{m,n}^{(k)}\} \circ \mathbf{U}_n, \quad (49)$$

where

$$\{\mathbf{F}_{m,n}^{(k)}\} = \{\mathbf{P}_{m,n}^{(k)}\} * [\{\mathbf{I}_n\} - \{\mathbf{Q}_{n,m}^{(k)}\} * \{\mathbf{P}_{m,n}^{(k)}\}]^{-1}. \quad (50)$$

Thus, the kernels of the feedback system can be obtained by applying the subtraction, cascade, and inversion operations discussed above. However, the explicit formulas for these kernels will not be given here.

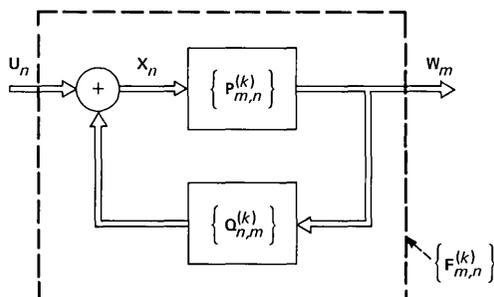


Fig. 6—A schematic representation of a nonlinear MIMO feedback system.

Actually, special cases of these formulas have been obtained for scalar systems in Refs. 2, 6, and 7.

Note from (50) and Section 6.4, that the feedback kernels exist if and only if the $n \times n$ matrix $[\mathbf{1}_n - \mathbf{Q}_{m,n}^{(1)}(f) \cdot \mathbf{P}_{m,n}^{(1)}(f)]$ is nonsingular. Note also that if $m = n$, and if $\mathbf{P}_{n,n}^{(1)}(f)$ is nonsingular, then (50) reduces to

$$\{\mathbf{F}_{n,n}^{(k)}\} = [\{\mathbf{P}_{n,n}^{(k)}\}^{-1} - \{\mathbf{Q}_{n,n}^{(k)}\}]^{-1}. \quad (51)$$

If the system in Fig. 6 is changed to a negative feedback system, then the minus signs in (50) and (51) should be changed to plus signs.

VII. CONCLUSIONS

A method of analysis has been presented for mildly nonlinear MIMO systems with memory. The method utilizes Volterra series whose kernels are two-dimensional matrices. The analysis was made possible through the use of the Kronecker product of matrices, which is a simple but powerful tool in matrix theory. This results in a compact representation of the system equations, and facilitates the systematic performance of various useful system operations, such as addition, cascading, inversion, and feedback. These operations can be used to describe a complex, nonlinear MIMO system as an interconnection of simple subsystems.

ACKNOWLEDGMENT

I thank Harrison E. Rowe for introducing me to the wonderful world of the Kronecker products of matrices.

REFERENCES

1. D. A. George, "Continuous Nonlinear System," Tech. Report 355, Research Lab. of Elec., M.I.T., Cambridge, MA., July 24, 1959.
2. H. L. Van Trees, "Functional Techniques for the Analysis of the Nonlinear Behavior of Phase-Locked Loops," Proc. IEEE, Vol. 52, No. 8 (August 1964), pp. 894-911.
3. S. Narayanan, "Transistor Distortion Analysis Using Volterra Series Representation," B.S.T.J., 46, No. 5 (May-June 1967), pp. 991-1024.
4. R. E. Maurer and S. Narayanan, "Noise Loading Analysis of a Third-Order Nonlinear System with Memory," IEEE Trans. Comm. Tech., COM-16, No. 5 (October 1968), pp. 701-12.
5. S. Narayanan, "Intermodulation Distortion of Cascaded Transistors," IEEE J. of Solid-State Circuits, SC-4, No. 3 (June 1969), pp. 97-106.
6. S. Narayanan, "Application of Volterra Series to Intermodulation Distortion Analysis of Transistor Feedback Amplifiers," IEEE Trans. Circuit Theory, CT-17, No. 4 (November 1970), pp. 518-27.
7. E. Bedrosian and S. O. Rice, "The Output Properties of Volterra Systems (Nonlinear Systems with Memory) Driven by Harmonic and Gaussian Inputs," Proc. IEEE, 59, No. 12 (December 1971), pp. 1688-1707.
8. J. Goldman, "A Volterra Series Description of Crosstalk Interference in Communications Systems," B.S.T.J., 52, No. 5 (May-June 1973), pp. 649-668.
9. J. J. Bussgang, L. Ehrman, and J. W. Graham, "Analysis of Nonlinear Systems with Multiple Inputs," Proc. IEEE, 62, No. 8 (August 1974), pp. 1088-1119.
10. D. D. Weiner and G. H. Naditch, "A Scattering Variable Approach to the Volterra

- Analysis of Nonlinear Systems," IEEE Trans. Microwave Theory Tech., *MTT-24*, No. 7 (July 1976), pp. 422-33.
11. T. R. Harper and W. J. Rugh, "Structural Feature of Factorable Volterra Systems," IEEE Trans. Automat. Control, *AC-21*, No. 6 (December 1976), pp. 822-32.
 12. K. Gopal, M. S. Nakhla, K. Singhal, and J. Vlach, "Distortion Analysis of Transistor Networks," IEEE Trans. Circuits and Systems, *CAS-25*, No. 2 (February 1978), pp. 99-106.
 13. S. Benedetto, E. Biglieri, and R. Daffara, "Modeling and Performance Evaluation of Nonlinear Satellite Links—A Volterra Series Approach," IEEE Trans. Aerospace and Electronic Sys., *AES-15*, No. 4 (July 1979), pp. 494-507.
 14. L. O. Chua and C.-Y. Ng, "Frequency Domain Analysis of Nonlinear Systems: General Theory," IEE Journal on Electronic Circuits and Systems, *3*, No. 4 (July 1979), pp. 165-85.
 15. H. K. Thapar and B. J. Leon, "Computer-Aided Distortion and Spectrum Analysis of Nonlinear Circuits and Systems," Proc. 1979 Int. Sym. on Circuits and Sys., (July 1979), pp. 49-51.
 16. ———, "The Computational Aspects of Volterra Series," Proc. 1980 Int. Sym. on Circuits and Sys. (April 1980), pp. 532-34.
 17. D. D. Weiner and J. F. Spina, *Sinusoidal Analysis and Modeling of Weakly Nonlinear Circuits*, New York: Van Nostrand, 1980.
 18. M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, New York: John Wiley, 1980.
 19. ———, "Nonlinear System Modeling Based on the Wiener Theory," Proc. IEEE, *69*, No. 12 (December 1981), pp. 1557-73.
 20. P. J. Lawrence, "Estimation of the Volterra Functional Series of a Nonlinear System Using Frequency-Response Data," IEE Proc., *128*, Pt.D, No. 5 (September 1981), pp. 206-10.
 21. I. W. Sandberg, "Expansions for Nonlinear Systems," B.S.T.J., *61*, No. 2 (February, 1982), pp. 159-99.
 22. F. A. Graybill, *Introduction of Matrices with Applications in Statistics*, Belmont, California: Wadsworth, 1969, pp. 196-209.
 23. R. Bellman, *Introduction to Matrix Analysis*, New York: McGraw Hill, 1970, pp. 235-40.
 24. J. W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," IEEE Trans. Circuits and Sys. *Cas-25*, No. 9 (September 1978), pp. 772-81.
 25. R. W. Brockett, "On the Algebraic Structure of Bilinear Systems," in *Theory and Applications of Variable Structure Systems*, R. R. Mohler and A. Ruberti, eds., New York: Academic Press, 1972, pp. 153-68.
 26. ———, "Volterra Series and Geometric Control Theory," Automatica, *12* (March 1976), pp. 167-76.

APPENDIX A

Index Notation

Here we rewrite, in the index notation, some of the key equations marked by a dagger ([†]) in the body of the paper. The same equation numbers are used here as are used in the text. Before doing so, however, we note from (7) that, for MIMO systems with memory, the matrix kernels used in the matrix notations are related to the array kernels used in the index notation by the relations

$$[\mathbf{P}^{(k)}(\tau_1, \dots, \tau_k)]_{ij} = P_{ij_1 \dots j_k}^{(k)}(\tau_1, \dots, \tau_k), \quad (52)$$

$$[\mathbf{P}^{(k)}(f_1, \dots, f_k)]_{ij} = P_{ij_1 \dots j_k}^{(k)}(f_1, \dots, f_k), \quad (53)$$

where j is given by (6).

A list of the equations in question follows.

$$w_i^{(k)}(t_1, \dots, t_k) = \sum_{j_1=1}^m \cdots \sum_{j_k=1}^m \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P_{ij_1 \dots j_k}^{(k)}(\tau_1, \dots, \tau_k) \cdot u_{j_1}(t_1 - \tau_1) \cdots u_{j_k}(t_k - \tau_k) d\tau_1 \cdots d\tau_k \right\} \quad (12)^\dagger$$

$$W_i^{(k)}(f_1, \dots, f_k) = \sum_{j_1=1}^m \cdots \sum_{j_k=1}^m P_{ij_1 \dots j_k}^{(k)}(f_1, \dots, f_k) \cdot U_{j_1}(f_1) \cdots U_{j_k}(f_k). \quad (16)^\dagger$$

$$P_{ij_1 j_2}^{(2)}(f_1, f_2) = P_{ij_2 j_1}^{(2)}(f_2, f_1). \quad (21a)^\dagger$$

$$P_{ij_1 j_2 j_3}^{(3)}(f_1, f_2, f_3) = P_{ij_\alpha j_\beta j_\gamma}^{(3)}(f_\alpha, f_\beta, f_\gamma). \quad (21b)^\dagger$$

$$P_{ij_1 j_2}^{(2)}(f_1, f_2) = \frac{1}{2} [\hat{P}_{ij_1 j_2}^{(2)}(f_1, f_2) + \hat{P}_{ij_2 j_1}^{(2)}(f_2, f_1)]. \quad (22a)^\dagger$$

$$P_{ij_1 j_2 j_3}^{(3)}(f_1, f_2, f_3) = \frac{1}{6} \sum_{\alpha, \beta, \gamma} \hat{P}_{ij_\alpha j_\beta j_\gamma}^{(3)}(f_\alpha, f_\beta, f_\gamma). \quad (22b)^\dagger$$

$$P_{ij_1 j_2}^{(2)}(f, f) = P_{ij_2 j_1}^{(2)}(f, f). \quad (30a)^\dagger$$

$$P_{ij_1 j_2}^{(2)}(f, -f) = [P_{ij_2 j_1}^{(2)}(f, -f)]^*. \quad (30b)^\dagger$$

$$P_{ij_1 j_2 j_3}^{(3)}(f, f, -f) = P_{ij_2 j_1 j_3}^{(3)}(f, f, -f). \quad (30c)^\dagger$$

$$P_{ij_1 j_2 j_3}^{(3)}(f, f, f) = P_{ij_\alpha j_\beta j_\gamma}^{(3)}(f, f, f). \quad (30d)^\dagger$$

$$T_{ij_1}^{(1)}(f) = \sum_{\alpha=1}^m Q_{i\alpha}^{(1)}(f) P_{\alpha j_1}^{(1)}(f). \quad (42a)^\dagger$$

$$T_{ij_1 j_2}^{(2)}(f_1, f_2) = \sum_{\alpha=1}^m Q_{i\alpha}^{(1)}(f_1 + f_2) P_{\alpha j_1 j_2}^{(2)}(f_1, f_2) + \sum_{\alpha=1}^m \sum_{\beta=1}^m Q_{i\alpha\beta}^{(2)}(f_1, f_2) P_{\alpha j_1}^{(1)}(f_1) P_{\beta j_2}^{(1)}(f_2). \quad (42b)^\dagger$$

$$\begin{aligned} \hat{T}_{ij_1 j_2 j_3}^{(3)}(f_1, f_2, f_3) &= \sum_{\alpha=1}^m Q_{i\alpha}^{(1)}(f_1 + f_2 + f_3) P_{\alpha j_1 j_2 j_3}^{(3)}(f_1, f_2, f_3) \\ &+ \sum_{\alpha=1}^m \sum_{\beta=1}^m [Q_{i\alpha\beta}^{(2)}(f_1, f_2 + f_3) P_{\alpha j_1}^{(1)}(f_1) P_{\beta j_2 j_3}^{(2)}(f_2, f_3) \\ &+ Q_{i\alpha\beta}^{(2)}(f_1 + f_2, f_3) P_{\alpha j_1 j_2}^{(2)}(f_1, f_2) P_{\beta j_3}^{(1)}(f_3)] \\ &+ \sum_{\alpha=1}^m \sum_{\beta=1}^m \sum_{\gamma=1}^m Q_{i\alpha\beta\gamma}^{(3)}(f_1, f_2, f_3) P_{\alpha j_1}^{(1)}(f_1) P_{\beta j_2}^{(1)}(f_2) P_{\gamma j_3}^{(1)}(f_3). \quad (42c)^\dagger \end{aligned}$$

$$\begin{aligned}
T_{i_1 j_1 j_2 j_3}^{(3)}(f_1, f_2, f_3) &= \sum_{\alpha=1}^m Q_{i\alpha}^{(1)}(f_1 + f_2 + f_3) P_{\alpha j_1 j_2 j_3}^{(3)}(f_1, f_2, f_3) \\
&+ \frac{2}{3} \sum_{\alpha=1}^m \sum_{\beta=1}^m [Q_{i\alpha\beta}^{(2)}(f_1, f_2 + f_3) P_{\alpha j_1}^{(1)}(f_1) P_{\beta j_2 j_3}^{(2)}(f_2, f_3) \\
&+ Q_{i\alpha\beta}^{(2)}(f_2, f_3 + f_1) P_{\alpha j_2}^{(1)}(f_2) P_{\beta j_3 j_1}^{(2)}(f_3, f_1) \\
&+ Q_{i\alpha\beta}^{(2)}(f_3, f_1 + f_2) P_{\alpha j_3}^{(1)}(f_3) P_{\beta j_1 j_2}^{(2)}(f_1, f_2)] \\
&+ \sum_{\alpha=1}^m \sum_{\beta=1}^m \sum_{\gamma=1}^m Q_{i\alpha\beta\gamma}^{(3)}(f_1, f_2, f_3) P_{\alpha j_1}^{(1)}(f_1) P_{\beta j_2}^{(1)}(f_2) P_{\gamma j_3}^{(1)}(f_3). \quad (42c)^\dagger
\end{aligned}$$

APPENDIX B

Kronecker Product of Matrices

Here we define the Kronecker product of matrices and summarize some of its properties that are used in this paper. More extensive coverage of this topic is given in Refs. 22 through 24.

Let $\mathbf{A} = [a_{i_a j_a}]$ and $\mathbf{B} = [b_{i_b j_b}]$ be $m_a \times n_a$ and $m_b \times n_b$ matrices, respectively. Their Kronecker product results in the $m_a m_b \times n_a n_b$ matrix, $\mathbf{C} = [c_{i_c j_c}]$, given by

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = \begin{bmatrix} \mathbf{A}b_{11} & \mathbf{A}b_{12} & \cdots & \mathbf{A}b_{1n_b} \\ \mathbf{A}b_{21} & \mathbf{A}b_{22} & & \mathbf{A}b_{2n_b} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \mathbf{A}b_{m_b 1} & \mathbf{A}b_{m_b 2} & \cdots & \mathbf{A}b_{m_b n_b} \end{bmatrix}, \quad (54)$$

where “ \times ” is the Kronecker-product symbol. Thus,

$$c_{i_c j_c} = a_{i_a j_a} b_{i_b j_b}, \quad (55a)$$

where

$$i_c = i_a + m_a(i_b - 1), \quad (55b)$$

$$j_c = j_a + n_a(j_b - 1). \quad (55c)$$

Note that, since $i_a \leq m_a$ and $j_a \leq n_a$, (55b) and (55c) have unique solutions for i_a , i_b , j_a and j_b in terms of i_c and j_c . Actually, (54) and (55) define the left Kronecker product.²² One can also define a right Kronecker product,^{23,24} which, however, is not used in this paper. In general,

$$\mathbf{A} \times \mathbf{B} \neq \mathbf{B} \times \mathbf{A}. \quad (56)$$

It can be shown that the Kronecker product has the following properties:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \times \mathbf{C} \equiv \mathbf{A} \times \mathbf{B} \times \mathbf{C}. \quad (57)$$

$$(\mathbf{A} + \mathbf{B}) \times \mathbf{C} = (\mathbf{A} \times \mathbf{C}) + (\mathbf{B} \times \mathbf{C}). \quad (58)$$

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) + (\mathbf{A} \times \mathbf{C}). \quad (59)$$

$$(\mathbf{A} \cdot \mathbf{B}) \times (\mathbf{C} \cdot \mathbf{D}) = (\mathbf{A} \times \mathbf{C}) \cdot (\mathbf{B} \times \mathbf{D}). \quad (60)$$

$$(\mathbf{A} \times \mathbf{B})^{-1} = \mathbf{A}^{-1} \times \mathbf{B}^{-1}. \quad (61)$$

$$(\mathbf{A} \times \mathbf{B})^T = \mathbf{A}^T \times \mathbf{B}^T. \quad (62)$$

In the above equations, “ T ” refers to matrix transposition, and the dot implies ordinary matrix multiplication. The dimensions of the various matrices are arbitrary, but of course, should be consistent with the requirements of the inversion, addition, and ordinary multiplication operations, where applicable.

APPENDIX C

Reversing and Permutation Matrices

Here we define the $n^2 \times n^2$ reversing matrix, $\mathbf{R}^{(n)}$, and the six $n^3 \times n^3$ permutation matrices, $\Phi_{\alpha\beta\gamma}^{(n)}$, which satisfy (19) and (20). The superscript “ (n) ” is used in this appendix to emphasize the dimensions. It can be shown from (19) and (55) that $\mathbf{R}^{(n)}$ is given by (cf. Ref. 24)

$$R_{i+n(j-1), k+n(l-1)}^{(n)} = \delta_{il}\delta_{jk}, \quad i, j, k, l = 1, 2, \dots, n, \quad (63)$$

where $\delta_{\alpha\beta}$ is the Kronecker delta, which is equal to 1 if $\alpha = \beta$, and 0 if $\alpha \neq \beta$. For example,

$$\mathbf{R}^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (64)$$

It can be verified that

$$\mathbf{R}^{(n)} = [\mathbf{R}^{(n)}]^T = [\mathbf{R}^{(n)}]^{-1}, \quad (65)$$

where “ T ” refers to matrix transposition. Moreover, if \mathbf{M}_1 and \mathbf{M}_2 are $m \times n$ matrices, then

$$\mathbf{R}^{(m)} \cdot (\mathbf{M}_1 \times \mathbf{M}_2) \cdot \mathbf{R}^{(n)} = \mathbf{M}_2 \times \mathbf{M}_1, \quad (66)$$

which is a generalization of (19).

It can be shown from (19), (20) and (60) that

$$\Phi_{123}^{(n)} = [\Phi_{123}^{(n)}]^T = [\Phi_{123}^{(n)}]^{-1} = \mathbf{1}_{n^3}, \quad (67a)$$

$$\Phi_{132}^{(n)} = [\Phi_{132}^{(n)}]^T = [\Phi_{132}^{(n)}]^{-1} = \mathbf{1}_n \times \mathbf{R}^{(n)}, \quad (67b)$$

$$\Phi_{213}^{(n)} = [\Phi_{213}^{(n)}]^T = [\Phi_{213}^{(n)}]^{-1} = \mathbf{R}^{(n)} \times \mathbf{1}_n, \quad (67c)$$

$$\Phi_{231}^{(n)} = [\Phi_{312}^{(n)}]^T = [\Phi_{312}^{(n)}]^{-1} = [\mathbf{1}_n \times \mathbf{R}^{(n)}] \cdot [\mathbf{R}^{(n)} \times \mathbf{1}_n], \quad (67d)$$

$$\Phi_{312}^{(n)} = [\Phi_{231}^{(n)}]^T = [\Phi_{231}^{(n)}]^{-1} = [\mathbf{R}^{(n)} \times \mathbf{1}_n] \cdot [\mathbf{1}_n \times \mathbf{R}^{(n)}], \quad (67e)$$

$$\begin{aligned} \Phi_{321}^{(n)} &= [\Phi_{321}^{(n)}]^T = [\Phi_{321}^{(n)}]^{-1} = [\mathbf{1}_n \times \mathbf{R}^{(n)}] \cdot [\mathbf{R}^{(n)} \times \mathbf{1}_n] \cdot [\mathbf{1}_n \times \mathbf{R}^{(n)}] \\ &= [\mathbf{R}^{(n)} \times \mathbf{1}_n] \cdot [\mathbf{1}_n \times \mathbf{R}^{(n)}] \cdot [\mathbf{R}^{(n)} \times \mathbf{1}_n], \end{aligned} \quad (67f)$$

where $\mathbf{1}_n$ and $\mathbf{1}_{n^3}$ are the $n \times n$ and $n^3 \times n^3$ identity matrices, respectively. Also, it can be shown from (20) and (55) that, if α, β, γ are any permutation of 1, 2, 3, then the i - j element of $\Phi_{\alpha\beta\gamma}^{(n)}$ is given by

$$[\Phi_{\alpha\beta\gamma}^{(n)}]_{ij} = \delta_{i_1 j_1} \delta_{i_2 j_2} \delta_{i_3 j_3}, \quad (68a)$$

where

$$i = i_\alpha + n(i_\beta - 1) + n^2(i_\gamma - 1), \quad (68b)$$

$$j = j_1 + n(j_2 - 1) + n^2(j_3 - 1), \quad (68c)$$

and where $i_1, j_1, i_2, j_2, i_3, j_3 = 1, 2, \dots, n$. For example, (67d) and (68) gives

$$\Phi_{231}^{(2)} = [\Phi_{312}^{(2)}]^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (69)$$

It can be verified that if $\mathbf{M}_1, \mathbf{M}_2$ and \mathbf{M}_3 are $m \times n$ matrices, then

$$\Phi_{\alpha\beta\gamma}^{(m)} \cdot (\mathbf{M}_1 \times \mathbf{M}_2 \times \mathbf{M}_3) \cdot [\Phi_{\alpha\beta\gamma}^{(n)}]^T = \mathbf{M}_\alpha \times \mathbf{M}_\beta \times \mathbf{M}_\gamma, \quad (70)$$

which is a generalization of (55). Also, if \mathbf{M} and \mathbf{N} are $m \times n$ and $m^2 \times n^2$ matrices, respectively, then

$$\Phi_{231}^{(m)} \cdot (\mathbf{M} \times \mathbf{N}) \cdot [\Phi_{231}^{(n)}]^T = \mathbf{N} \times \mathbf{M}, \quad (71a)$$

$$\Phi_{312}^{(m)} \cdot (\mathbf{N} \times \mathbf{M}) \cdot [\Phi_{312}^{(n)}]^T = \mathbf{M} \times \mathbf{N}. \quad (71b)$$

Moreover, if \mathbf{M} and \mathbf{K} are $m \times n$ and $m \times n^2$ matrices, respectively, then

$$\mathbf{R}^{(m)} \cdot (\mathbf{M} \times \mathbf{K}) \cdot [\Phi_{231}^{(n)}]^T = \mathbf{K} \times \mathbf{M}, \quad (72a)$$

$$\mathbf{R}^{(m)} \cdot (\mathbf{K} \times \mathbf{M}) \cdot [\Phi_{312}^{(n)}]^T = \mathbf{M} \times \mathbf{K}. \quad (72b)$$

Finally, if \mathbf{M} and \mathbf{L} are $m \times n$ and $m^2 \times n$ matrices, respectively, then

$$\Phi_{231}^{(m)} \cdot (\mathbf{M} \times \mathbf{L}) \cdot \mathbf{R}^{(n)} = \mathbf{L} \times \mathbf{M}, \quad (73a)$$

$$\Phi_{312}^{(m)} \cdot (\mathbf{L} \times \mathbf{M}) \cdot \mathbf{R}^{(n)} = \mathbf{M} \times \mathbf{L}. \quad (73b)$$

APPENDIX D

General Cascade Relation

Here we give a generalization of the unsymmetric cascade relations given in (42a), (42b), and (42c) for arbitrary k , (cf. Refs. 7 and 9 for scalar systems)

$$\hat{\mathbf{T}}^{(k)}(f_1, \dots, f_k) = \sum_{l=1}^k \left\{ \sum_{\substack{k_1, k_2, \dots, k_l=1 \\ (k_1+k_2+\dots+k_l=k)}}^{k-l+1} \mathbf{Q}^{(l)}(f_1 + \dots + f_{k_1}, f_{k_1+1} + \dots + f_{k_1+k_2}, \dots, f_{k-k_l+1} + \dots + f_k) [\mathbf{P}^{(k_1)}(f_1, \dots, f_{k_1}) \times \mathbf{P}^{(k_2)}(f_{k_1+1}, \dots, f_{k_1+k_2}) \times \dots \times \mathbf{P}^{(k_l)}(f_{k-k_l+1}, \dots, f_k)] \right\}. \quad (74)$$

Note that the second summation contains $\binom{k-1}{l-1}$ terms, and that the frequency arguments always appear in the order f_1, f_2, \dots, f_k . As is the case with (42c), the cascade relation of (74) does not preserve kernel symmetry for $k \geq 3$. The symmetric form of (74), which would generalize (42c), will not be given since it requires the use of permutation matrices of more than three indices, which have not been introduced yet.

Computing the Distribution of a Random Variable via Gaussian Quadrature Rules

By M. H. MEYERS

(Manuscript received March 26, 1982)

Using the technique of Gaussian quadrature rules, a new estimator is proposed for approximating the distribution of a random variable given only a finite number of its moments. The estimator is shown by numerous examples to be accurate on the tails of both continuous and discrete distributions. Efficient algorithms exist for computing the estimator from the first $2N$ moments of the random variable. A robust implementation of the estimator is presented, along with rules that provide additional protection against computer roundoff errors.

I. INTRODUCTION

In this paper we present a method for computing the Cumulative Distribution Function (CDF) of an arbitrary random variable. Using the theory of Gaussian Quadrature Rules (GQRs), we derive an estimator that converges asymptotically to the true CDF. In practice, convergence is obtained without excessive computation. A general estimator is developed here that is applicable to a wide class of problems.

Section 2.1 begins with a review of GQR analysis as it has traditionally been used for numerical integration. Several authors have shown the existence of extremely efficient algorithms for computing the parameters of the GQR. An efficient and robust procedure for obtaining the GQR parameters is presented in the appendix. Two CDF estimators based on GQR are derived in Sections 2.3 and 2.4. The first estimator is most suited to numerical integration schemes and estimation of discrete distributions, while the second is appropriate for continuous distributions such as Gaussian noise or crosstalk. Section III gives numerous examples that show the inherent accuracy of the technique for continuous, discrete, and mixed distributions. Computational methods for deriving the required moments are discussed, along with

modifications that tend to mitigate the roundoff errors that plague GQR analysis of nonsymmetric distributions.

II. THEORY AND PROPERTIES OF GQR

2.1 Classical use of GQR

The GQR has traditionally been used as a numerical integration procedure and is particularly efficient for computing integrals of the form

$$\int_a^b f(x)w(x)dx$$

where the integrand has been factored into a non-negative term $w(x)$ and a strongly continuous term $f(x)$.

The first application of GQR in the communications literature¹ was motivated by the work of Golub and Welsch² and Sack and Donovan³, who showed that the non-negative factor $w(x)$ need not even be completely known to compute the desired integral. Only a finite number of the moments of $w(x)$ are required to find the desired integration rule. Benedetto et al.¹ noticed that the problem of error probability evaluation in the presence of intersymbol interference (ISI) could be posed in this form. Other applications of the GQR technique can be found in Refs. 4 through 9.

In this paper, we apply the GQR technique to a larger class of problems where $f(x)$ need not be continuous. We begin by reviewing a fundamental result in the theory of GQR.

Theorem: Let $w(x)$ be a non-negative weight function defined on (a, b) . Then if $f(x)$ has continuous derivatives up to order $2N$ (see Refs. 10 through 13),

$$\begin{aligned} J &= \int_a^b f(x)w(x)dx \\ &= \sum_{i=1}^N A_i f(t_i) + R_N(\xi) \quad a < \xi < b \\ &\quad \alpha < t_i < b \quad i = 1, 2 \dots N, \end{aligned} \quad (1)$$

where

$$R_N(\xi) = \frac{f^{(2N)}(\xi)}{(2N)!(h_N)^2}, \quad a < \xi < b, \quad (2)$$

$f^{(2N)}(x)$ is the $2N$ th derivative of $f(x)$ and $(2N)!$ is $2N$ factorial. The nodes $\{t_i\}$ are the distinct real roots of the unique N th degree polynomial

$$p_N(x) = k_N \prod_{i=1}^N (x - t_i), \quad k_N > 0. \quad (3)$$

The polynomials $p_n(x)$ are orthonormal with respect to $w(x)$, i.e.,

$$\int_a^b w(x) p_m(x) p_n(x) dx = \begin{cases} 1 & m = n \\ 0 & m \neq n. \end{cases}$$

The strictly positive weights (or Christoffel numbers) are in turn given by

$$A_i = \frac{-k_{N+1}}{k_N} \frac{1}{p_{N+1}(t_i) p'_N(t_i)} \quad i = 1, 2 \dots N, \quad (4)$$

where

$$p'_N(t_i) = \left. \frac{dp_N(t)}{dt} \right|_{t=t_i}.$$

The $2N$ -tuple $\{A_i, t_i\}_{i=1}^N$ is known as the N -point rule corresponding to $w(x)$.

If $f(x)$ is a polynomial of degree $(2N - 1)$ or less, the remainder $R_N(\xi)$ equals zero and the GQR is exact. This affords the maximum degree of precision (i.e., the maximum degree polynomial that is integrable with no error for an N -point rule) possible with a quadrature formula of the form of (1).¹⁰⁻¹² When the remainder is not zero, it can be bounded in magnitude to obtain upper and lower bounds on J . The bounds obtained in Ref. 1 for the ISI and Gaussian noise problem are often loose though, and convergence of the N -term summation in (2) is usually much faster than might be inferred from bounds on $R_N(\xi)$.

2.2 Methods for computing GQR

Several algorithms are known for efficiently computing the rule for an arbitrary weight function $w(x)$. Extremely useful procedures have been discovered by Golub and Welsh,² Sack and Donovan,³ and Gautchi.¹⁴ The outstanding merit of these techniques is that the N -point rule corresponding to a given $w(x)$ can be computed from the moments

$$\mu_i = \int_a^b x^i w(x) dx. \quad (5)$$

Because explicit knowledge of the weight function $w(x)$ is not required, the GQR procedure is a powerful tool for the analysis of communications systems.

Details of an algorithm for computing GQR are given in the appendix. Our algorithm is a modification of Gautchi's procedure,¹⁴ which tends

to reduce computer roundoff errors. The critical stage in the algorithm is the Cholesky decomposition of a positive definite matrix of moments. The standard Cholesky decomposition used in Refs. 1 and 2 fails when, because of limitations of machine accuracy, the matrix is no longer positive definite due to roundoff errors. Improved accuracy is obtained by using an alternate method of performing the Cholesky decomposition* that avoids taking a square root at each step in the algorithm.¹⁵ Combining the alternate Cholesky decomposition with the modified moment algorithm of Gautchi¹⁴ yields an extremely stable method for obtaining GQR. Further discussion of techniques to mitigate computer roundoff errors is found in the appendix.

2.3 Computing the distribution of a random variable via GQR

In Ref. 1 GQRs are used to obtain the exact probability of error for digital transmission in the presence of ISI and Gaussian noise. The problem was reduced, via the GQR approach, to computing the moments of the ISI and letting $f(x)$ in (1) be the probability of error caused by Gaussian noise conditioned on the ISI. The ISI moments can be computed via Prabhu's method¹⁶ when the data symbols are independent. For a large class of correlated data, the moments can be efficiently computed via the modified Cariolaro-Pupolin algorithm.^{17,18} Both of the above procedures are easily implemented and have a complexity that grows only linearly with pulse duration.

While there have been numerous applications of GQR to problems in the literature, all those known to us have had the restriction that the function $f(x)$ has continuous derivatives up to order $2N$. Presumably, this is because of the desire for strict bounds on the error term in (2). If we are willing to forego the analytical error term and consequently accept an empirical convergence of (1), we can apply the GQR technique to a larger class of problems with excellent results.

The following theorem shows that no continuity requirements need be imposed on $f(x)$.

Theorem: (see Ref. 19) *If $W(x)$ is a fixed, nondecreasing function with infinitely many points of increase and the Riemann-Stieltjes integral*

$$\int_a^b f(x) dW(x)$$

exists, then

$$\int_a^b f(x) dW(x) = \lim_{N \rightarrow \infty} \sum_{i=1}^N A_i f(t_i), \quad (6)$$

* Applying the alternate Cholesky decomposition to the GQR problem was suggested by L. Kaufman. Subsequently, the same approach was found to have been independently proposed in Ref. 4.

where $\{A_i, t_i\}_{i=1}^N$ is the GQR corresponding to the moments

$$\mu_i = \int_a^b x^i dW(x) \quad i = 0, 1, 2, \dots, 2N.$$

The function $f(x)$ is arbitrary as long as the integral in (6) exists.

Because the CDF of a random variable is a nondecreasing function, we write the statistical expectation of the function $f(x)$ as

$$E[f(x)] = \int_a^b f(x) dW(x), \quad (7)$$

where $W(x)$ is a probability measure with infinitely many points of rise. Choosing $f(x)$ to be the indicator function

$$\begin{aligned} f(x) &= \phi_\alpha(x) \\ &= \begin{cases} 1 & x \leq \alpha \\ 0 & x > \alpha, \end{cases} \end{aligned} \quad (8)$$

we obtain the distribution function of the random variable as

$$\begin{aligned} \int_a^b \phi_\alpha(x) dW(x) &= \lim_{N \rightarrow \infty} \sum_{S_N^\alpha} A_i \\ &= \lim_{N \rightarrow \infty} \hat{W}_N(\alpha), \end{aligned} \quad (9)$$

where

$$\hat{W}_N(\alpha) = \sum_{S_N^\alpha} A_i \quad (10)$$

and

$$S_N^\alpha = \{i \mid t_i \leq \alpha\}$$

is the set of indices for which $t_i \leq \alpha$.

Since the rule can be obtained from the $\{\mu_i\}_{i=0}^{2N}$, we have a means of constructing an approximation to the CDF of a random variable from its moments. In the limit as N approaches infinity, eq. (9) is exact at each point α .

This leads us to propose the following estimator

$$W(x) \approx \hat{W}_N(x) = \sum_{S_N^x} A_i. \quad (11)$$

This estimator gives a staircase approximation to the true cumulative distribution that becomes increasingly fine as N increases. Equivalently, each (A_i, t_i) can be considered a point mass of a discrete approximation to the true probability density function.

While Szeġo's theorem proves the asymptotic convergence of the estimator to the true CDF when $W(x)$ has an infinite number of points of rise, a different result holds for a discrete distribution with a finite number of points of increase.

Theorem: If $W(x)$ is a fixed, nondecreasing function with $M < \infty$ points of increase, then

$$\int_a^b \phi_\alpha(x) dW(x) = \lim_{N \rightarrow M} \widehat{W}_N(\alpha). \quad (12)$$

Proof: An alternate formulation of Gaussian Quadrature¹⁴ is as the purely algebraic solution to

$$\mu_j = \sum_{i=1}^N A_i (t_i)^j \quad j = 0, 1, \dots, 2M. \quad (13)$$

Now we assume the unknown discrete PDF is of the form

$$w(x) = \sum_{i=1}^M A_i \delta(x - t_i).$$

The moments of this random variable are given by

$$\sum_{i=1}^M A_i (t_i)^j,$$

which is identical to (13) for $N = M$.

Finally, we consider the behavior of the GQR for N larger than the number of points of increase M . The result is that the algorithm breaks down entirely. This is because a discrete distribution that takes on exactly M values is completely characterized by its first $2M$ moments and the addition of redundant moments to the problem causes the procedure to fail when the Hankel matrix of moments [eq. (21)] becomes nonsingular.

2.4 A modified GQR estimator

The following is a modification of the estimator $\widehat{W}_N(\alpha)$ that has been found to be more accurate in many applications. Instead of assuming that the approximation PDF is composed of point masses, we assume that each area of mass A_i is more accurately modeled by a narrow, even symmetric, distribution centered around the point t_i . Thus, we propose the smoothed estimator $W_N^*(\alpha)$ which, when evaluated at a node, equals

$$W_N^*(t_i) = \widehat{W}_N(t_i) - \frac{A_i}{2}. \quad (14)$$

Between nodes, $W_N^*(\alpha)$ is given by any "smooth" interpolation routine. A simple linear interpolation was found to be sufficient in the examples

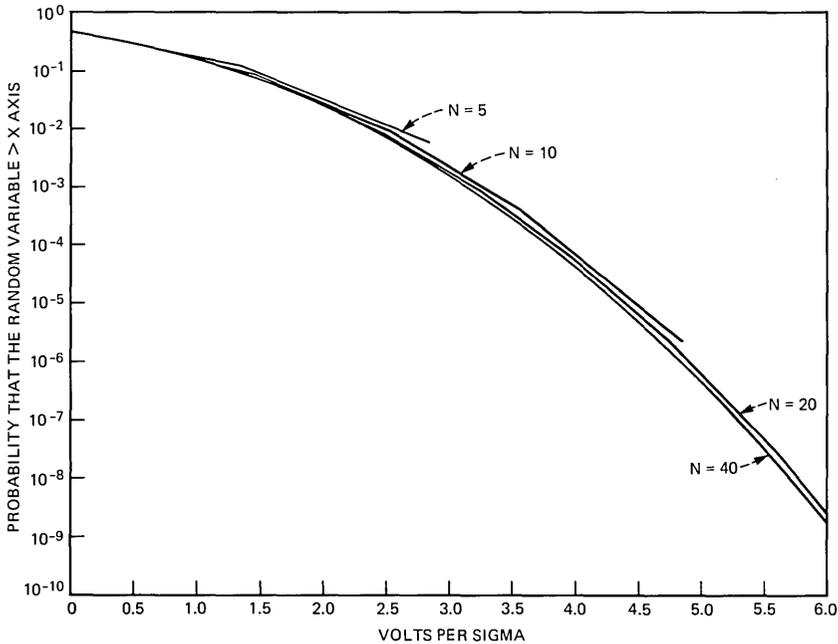


Fig. 1—Convergence of GQR estimator for Gaussian PDF.

that follow. This estimator does not have the jump discontinuities of the estimator $\hat{W}_N(\alpha)$ and is intuitively more satisfying because it fits a smoother distribution to $W(x)$. $W_N^*(\alpha)$ has been found to give more accurate results when applied to known continuous distributions and to discrete distributions when M is much greater than N .

III. APPLICATION TO ARBITRARY DISTRIBUTIONS

3.1 Known continuous random variable case

To show the convergence properties of the GQR technique, we illustrate the behavior of $\hat{W}_N(\alpha)$ and $W_N^*(\alpha)$ with some examples. We begin with the Gaussian distribution. Assuming a zero mean, unit variance random variable X , we compute the GQR estimators for various values of N in Fig. 1. Reasonably accurate results were obtained at the 10^{-6} point for $N > 10$. This empirical rate of convergence is also typical of distributions that have near-Gaussian statistics. The GQR algorithm, using the Cholesky decomposition described in the appendix, returned accurate results for all $N \leq 60$, where $N = 60$ was the dimensionality limit in the computer program.

In general, the GQR algorithm performs well for zero mean, symmetric distributions. To illustrate the problems that can occur with non-symmetric distributions, consider the lognormal distribution related to

the Gaussian distribution by $Y = e^X$. Straightforward computation of the moments yields

$$\mu_k = \exp(k^2/2). \quad (15)$$

Using these moments in the GQR algorithm, the algorithm breaks down at $N = 17$ because of roundoff errors in computation. This problem is solved by a transformation that symmetrizes the distribution. We then compute the GQR corresponding to the symmetrized distribution and take the inverse transform to obtain the original distribution.

For the lognormal distribution, we form a new PDF

$$w_e(y) = 1/2[w(y) + w(-y)], \quad (16)$$

which corresponds to the even part of $w(y)$. The moments of $w_e(y)$ are obtained by setting the odd moments of $w(y)$ equal to zero. The symmetric moments are then used in the GQR algorithm to obtain $W_e(y)$, which is easily transformed back to the desired CDF via the relation

$$W(y) = \begin{cases} 2W_e(y) - 1 & x \geq 0 \\ 0 & x < 0. \end{cases} \quad (17)$$

Experience with these procedures suggests that it is well worth the effort to transform distributions that are not symmetric and even (see Fig. 2). The modified moments produce the same robust accuracy seen with the Gaussian distribution above.

As another example, consider a uniform distribution defined on the interval $(-1, 1)$. The convergence of the GQR estimator $\hat{W}_N(\alpha)$ is shown in Fig. 3. Since the distribution has only finite support, by eq. (1) we know that all the nodes will lie in the interval $(-1, 1)$. In the limit as $N \rightarrow \infty$, the nodes will become more densely packed in this interval and

$$\lim_{i \rightarrow \infty} \{\max_i |t_i|\} = 1. \quad (18)$$

Thus, the GQR algorithm can be used to find the maximum value that a random variable attains, i.e., the largest node t_{\max} . This can be used, for example, to find the maximum eye degradation in a digital regenerator caused by correlated intersymbol interference.

All the examples so far have been trivial applications since we knew the real distributions *a priori*. A more interesting application is determining the distribution of the sum of K lognormal random variables. This problem has a long history and no closed form solution is known. This PDF is related to the distribution of crosstalk power in paired cable transmission systems and also results from transmission over certain types of fading channels. Utilizing the GQR technique, we can find the desired distribution if we can compute the necessary moments.

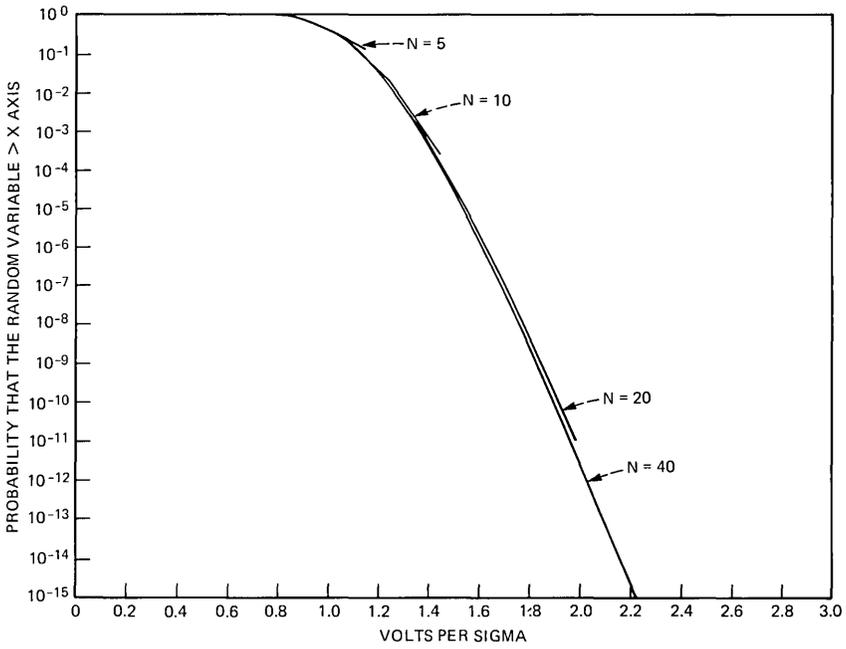


Fig. 2—Convergence of GQR estimator for lognormal PDF.

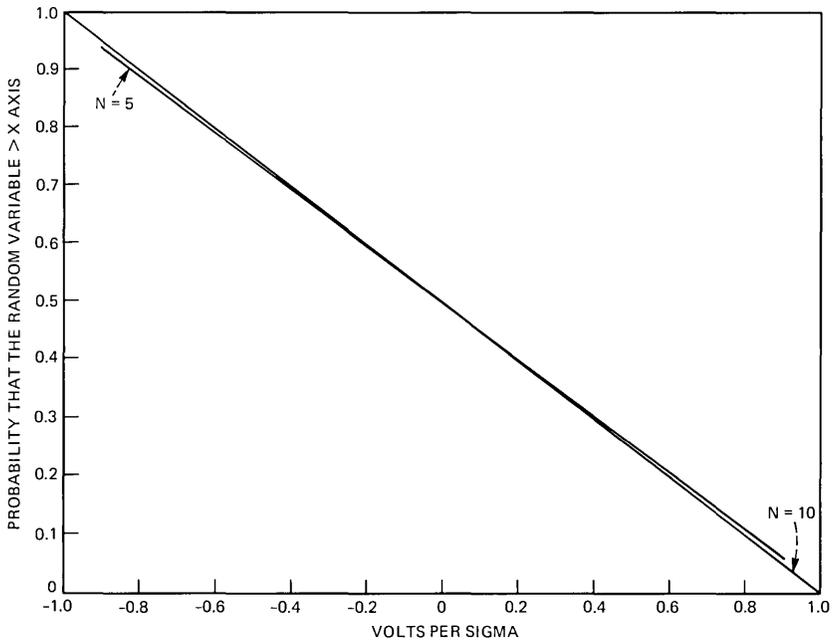


Fig. 3—Convergence of GQR estimator for uniform PDF.

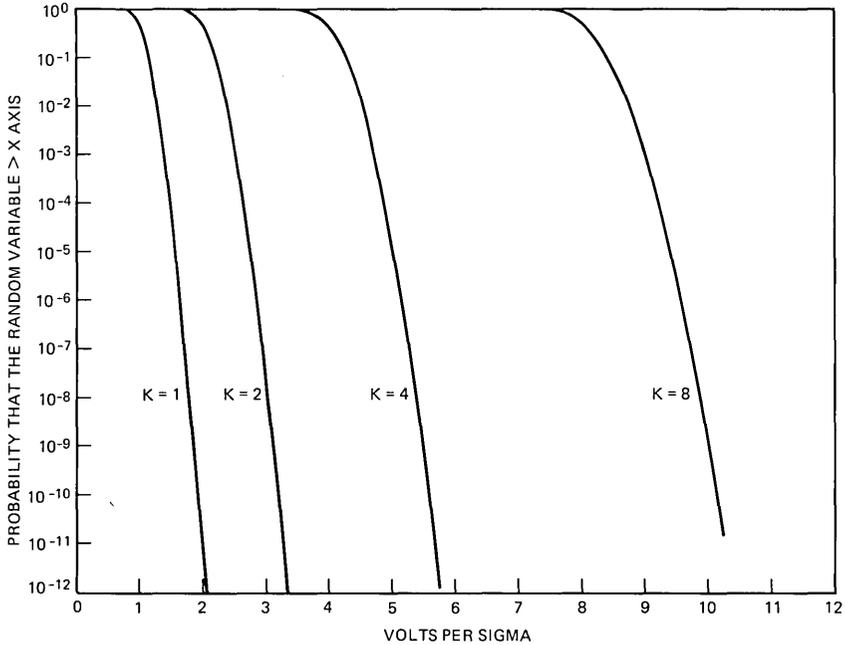


Fig. 4—Sum of K lognormal ($K = 1, 2, 4, 8$).

Assuming that the lognormal random variables are independent, and following Prabhu,¹⁶ we find that the moments of

$$V_K = \sum_{i=1}^K Y_i$$

are given by the recurrence relation

$$E[(V_K)^l] = \sum_{i=0}^l \binom{l}{i} E[(V_{K-1})^i] \mu_{l-i}, \quad (19)$$

where $\{\mu_i\}_{i=1}^{2N}$ are the moments of the independent, identically distributed lognormal random variables. Figure 4 shows the resulting distributions for $K = 2, 4, 8$, and 16. As we mentioned above, the distribution was symmetrized and inverse transformed to reduce the effects of roundoff errors. This technique can be applied to any number of arbitrary distributions for which the required moments can be computed.

3.2 Known discrete random variable case

In this section, we apply the GQR estimator $W_N^*(\alpha)$ to discrete distributions. First we consider the case of a mixed distribution composed of a Gaussian distribution plus discrete components. The weights

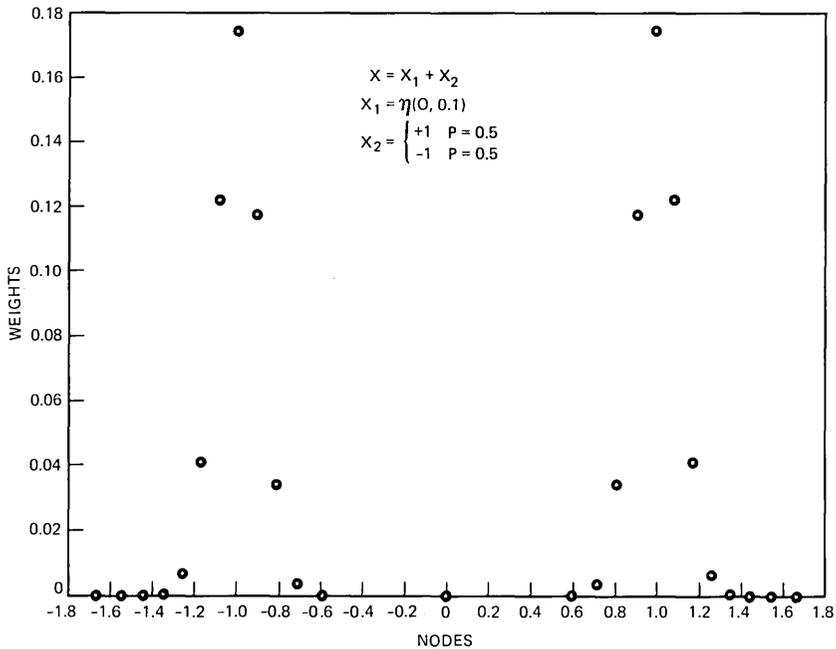


Fig. 5—Mixed Gaussian distribution.

and nodes of the estimator are shown in Fig. 5, where the second moment of the discrete part equals ten times that of the Gaussian component. As we can readily discern, the GQR procedure is useful in identifying the discrete components of a PDF.

As the final example of a known distribution, we consider the sum of nine equally spaced delta functions

$$w(x) = 1/9 \sum_{i=1}^9 \delta(x - x_i) \quad x_i = -5 + i \quad i = 1, 2, \dots, 9. \quad (20)$$

The convergence of the GQR estimator $W_N^*(x)$ is shown in Fig. 6, where the $N = 9$ estimator is exact since the distribution is uniquely defined by the first $2N = 18$ moments. For $N > 9$, the algorithm breaks down.

IV. SUMMARY

An estimator based on GQR has been proposed, which converges rapidly to the CDF of a random variable and requires only knowledge of the moments of the random variable in question. The technique is generally applicable to a large class of communications problems and provides a practical solution to many analytically intractable problems. The technique works equally well for discrete and continuous distributions and assumes no *a priori* knowledge of the distribution.

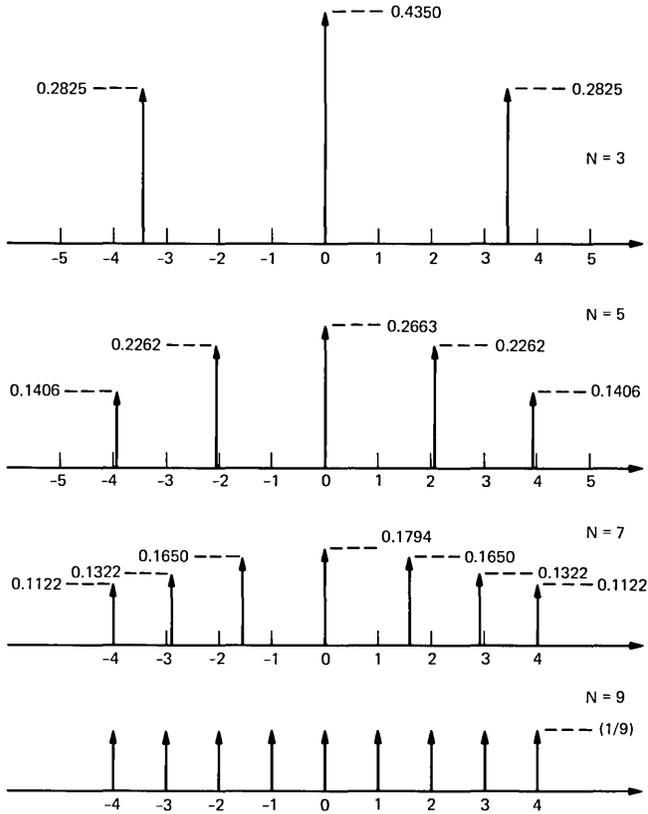


Fig. 6—GQR for discrete distribution.

V. ACKNOWLEDGMENT

It is a pleasure to acknowledge many discussion of GQR analysis with G. Turner.

REFERENCES

1. S. Benedetto, G. DeVincentiis, and A. Luvison, "Error Probability in the Presence of Intersymbol Interference and Additive Noise for Multilevel Digital Signals," *IEEE Trans. on Comm.*, *COM-21*, No. 3 (March 1973), pp. 181-90.
2. G. H. Golub and J. H. Welsch, "Calculation of Gauss Quadrature Rules," *Math. Comp.*, 26, No. 106 (April 1969), pp. 221-30.
3. R. A. Sack and A. F. Donovan, *An Algorithm for Gauss Quadrature Given Generalized Moments*, Salford, England: Dept. of Mathematics, Univ. of Salford, 1969.
4. R. Dogliotti, A. Luvison, and G. Pirani, "Error Probability in Optical Fiber Transmission Systems," *IEEE Trans. Information Theory*, *IT-25*, No. 2 (March 1979), pp. 170-78.
5. S. Benedetto, G. DeVincentiis, and A. Luvison, "Application of Gauss Quadrature Rules to Digital Communication Problems," *IEEE Trans. Comm.*, *COM-21* (October 1973), pp. 1159-65.

6. S. Y. Kwon and N. M. Shedadeh, "Analysis of Incoherent Frequency-Shift Keyed Systems," IEEE Trans. Comm., COM-23 (November 1975), pp. 1331-9.
7. E. Biglieri and R. Dogliotti, "Error Probability for Multi-level PAM Transmission with Intersymbol Interference, Gaussian and Impulsive Noise," IEEE Int. Conf. Comm. (June 1972), pp. 29.12-17.
8. S. Benedetto, E. Biglieri, and V. Castellani, "Combined Effects of Intersymbol, Interchannel and Co-channel Interferences in M-ary CPSK Systems," IEEE Trans. Comm., COM-21, No. 9 (September 1973), pp. 997-1007.
9. H. M. Carr and G. E. Cook, "Explicit Inversion of the Laplace Transform by a Jacobi-Gauss Quadrature Approximation," IEEE Trans. Auto. Control, AC-14, No. 6 (December 1969), pp. 777-9.
10. A. Stroud and D. Secrest, *Gaussian Quadrature Formulas*, Englewood Cliffs, NJ: Prentice-Hall, 1966.
11. V. I. Krylov, *Approximate Calculation of Integrals*, New York: The MacMillan Co., 1962.
12. Z. Kopal, *Numerical Analysis*, New York: John Wiley & Sons, Inc., 1955.
13. P. Davis and P. Rabinowitz, *Numerical Integration*, Waltham, MA: Blaisdell Publishing Company, 1967.
14. W. Gautschi, "On the Construction of Gaussian Quadrature Rules from Modified Moments," Math. Comp., 24 (April 1970), pp. 245-60.
15. R. S. Martin, G. Peters, and J. H. Wilkinson, "Symmetric Decomposition of a Positive Definite Matrix," Numer. Math., 7, No. 5 (December 1965), pp. 362-83.
16. V. K. Prabhu, "Some Considerations of Error Bounds in Digital Systems," B.S.T.J., 50 (December 1971), pp. 3127-52.
17. M. H. Meyers, "A Reduced Complexity Moment Algorithm for Correlated Digital Signals," IEEE Trans. Comm., COM-30, No. 7 (July 1982), pp. 1798-9.
18. G. L. Cariolaro and S. G. Pupolin, "Moments of Correlated Digital Signals for Error Probability Evaluation," IEEE Trans. of Information Theory, IT-21 (September 1975), pp. 558-68.
19. G. Szego, *Orthogonal Polynomials*, Providence, RI: Amer. Math. Soc. Collog. Publ., 23, Amer. Math. Soc., 1959.
20. W. Gautschi, "Construction of Gauss-Christoffel Quadrature Formulas," Math. Comp., 22 (April 1968), pp. 251-70.

APPENDIX

Details of the GQR Algorithm

In this appendix, we outline the algorithm used to compute Gaussian Quadrature Rules. The procedure combines Gautschi's modified moment technique¹⁴ with the Cholesky decomposition suggested by Martin et al.¹⁵ The resulting algorithm has been implemented using double precision arithmetic and has proven stable and robust.

To compute the $2N$ unknowns $\{A_i\}_{i=1}^N$ and $\{t_i\}_{i=1}^N$, we first form the matrix of modified moments

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} \cdots m_{1,N+1} \\ m_{2,1} & \cdot \\ \vdots & \vdots \\ m_{2,N+1} & \cdots m_{N+1,N+1} \end{bmatrix}, \quad (21)$$

where m_{ij} is given by the inner product

$$\begin{aligned} m_{ij} &= (T_{i-1}, T_{j-1}) \\ &= \int_a^b T_{i-1}(x)T_{j-1}(x)dW(x) \quad i, j = 1, 2 \cdots N + 1 \end{aligned} \quad (22)$$

and $\{T_i\}_{i=0}^N$ are the first $N + 1$ members of an arbitrary set of polynomials satisfying the recurrence relation

$$\begin{aligned} xT_j(x) &= a_j T_{j+1}(x) + b_j T_j(x) + c_j T_{j-1}(x) \quad j = 0, 1, 2, \dots, N \\ T_{-1}(x) &= 0, \quad a_j \neq 0. \end{aligned} \quad (23)$$

The orthogonal Tchebycheff polynomials determined by (23) constitute a convenient choice, with

$$\begin{aligned} a_0 &= 1 \\ a_j &= c_j = 1/2 \quad j = 1, 2, \dots \\ b_j &= 0 \quad j = 1, 2, \dots \end{aligned} \quad (24)$$

The modified moments m_{ij} in (23) are simply linear combinations of the moments

$$\mu_j = \int_a^b x^j dW(x)$$

and can be simplified for the case of the Tchebycheff polynomials by using the relation

$$T_i(x)T_j(x) = 1/2\{T_{i+j}(x) + T_{i-j}(x)\} \quad i \geq j. \quad (25)$$

Thus, if we define

$$\nu_k = \int T_k(x) dW(x),$$

then

$$m_{ij} = 1/2\{\nu_{i+j-2} + \nu_{i-j}\} \quad i \geq j. \quad (26)$$

It is not necessary in theory for the $T_k(x)$ to be orthogonal. The formulation by Golub and Welsch² used the unmodified moments corresponding to $T_k(x) = x^k$ and, hence, $a_j = 1$, $b_j = 0$, and $c_j = 0$ for all j . As Gautchi shows,¹⁴ the use of modified moments results in less sensitivity to computer roundoff errors.

We next form the tridiagonal matrix

$$\mathbf{J} = \begin{bmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \cdot & \cdot & \\ & \cdot & \cdot & \alpha_{N-1} & \beta_{N-1} \\ 0 & & & \beta_{N-1} & \alpha_N \end{bmatrix}, \quad (27)$$

where

$$\begin{aligned} \alpha_j &= b_j + \frac{r_{j,j+1}}{r_{j,j}} a_j - \frac{r_{j-i,j}}{r_{j-1,j-1}} a_{j-1} \quad j = 1, 2, \dots, N \\ \beta_j &= \frac{r_{j+1,j+1}}{r_{j,j}} a_j \quad j = 1, 2, \dots, N - 1. \end{aligned}$$

The r_{ij} are found from the relation

$$\mathbf{M} = \mathbf{R}^T \mathbf{R}. \quad (28)$$

The matrix \mathbf{R} is an upper triangular matrix and theoretically is positive definite if \mathbf{M} is positive definite. In practice, however, \mathbf{M} can be ill-conditioned and finite precision arithmetic will cause the matrix to appear singular.

The elements of \mathbf{R} are related to the moment matrix \mathbf{M} by the relations

$$\begin{aligned} r_{ii} &= \left(m_{ii} - \sum_{k=1}^{i-1} r_{ki}^2 \right)^{1/2} \\ r_{ij} &= \left(m_{ij} - \sum_{k=1}^{i-1} r_{ki} r_{kj} \right) / r_{ii} \quad i < j \\ & \quad i, j = 1, 2 \dots N. \end{aligned} \quad (29)$$

In practice, the computation of \mathbf{R} from (29) will fail at relatively small values of N when the square root of a negative number is attempted.

A refined Cholesky decomposition¹⁵ overcomes this problem by only requiring square roots to be computed at the end of the decomposition and not at each step as in (29). If we define \mathbf{R}^* by the relation

$$\mathbf{R} = \mathbf{R}^* \text{diag}(r_{ii}),$$

then \mathbf{R}^* will be a unit upper triangular matrix and

$$\begin{aligned} \mathbf{M} &= \mathbf{R}^T \mathbf{R} \\ &= \mathbf{R}^{*T} \text{diag}(r_{ii}^2) \mathbf{R}^* \\ &= \mathbf{R}^{*T} \mathbf{D} \mathbf{R}^*, \end{aligned} \quad (30)$$

where \mathbf{D} is a positive diagonal matrix. Then, defining the auxiliary quantities

$$m_{ij}^* = r_{ij}^* d_j, \quad (31)$$

the following solution is obtained

$$\begin{aligned} m_{ij}^* &= m_{ij} - \sum_{k=1}^{j-1} m_{ik}^* r_{jk}^* \quad j = 1, 2 \dots i - 1 \\ d_i &= m_{ii} - \sum_{k=1}^{i-1} m_{ik}^* r_{ik}^*. \end{aligned} \quad (32)$$

The advantage of the alternate decomposition is that square roots are not required until the final step, when the positive diagonal matrix

Table I—Comparison of three implementations

	Standard Cholesky	Alternate Cholesky	Alternate Cholesky Modified Moments
Gaussian Random Variable	17	60	60
Uniform Random Variable	13	40	38
Lognormal Random Variable	—	17	14
Symmetrized, Lognormal Random Variable	—	60	60

D in (30) is factored. Along with the modified moment procedure, the alternate Cholesky decomposition yields accurate results even for large values of N .

Several implementations of the GQR algorithm have been examined to elucidate the features that contribute to the reduction of computer roundoff errors. These include:

- (i) Standard Cholesky
- (ii) Alternate Cholesky
- (iii) Alternate Cholesky with modified moments
- (iv) All of the above using symmetrized moments.

Each approach was evaluated in double precision arithmetic.

The value of N at which the Cholesky decomposition fails was chosen as the measure of robustness for a variety of input probability density functions. Some of these results are tabulated in Table I. The standard Cholesky consistently had the poorest performance for all of the distributions considered. For symmetric distributions, the alternate Cholesky scheme provided a significant reduction of computer error. For the Gaussian distribution, the procedure was accurate for all $N \leq 60$, where 60 was the dimensionality limit imposed on the computer routine by storage requirements. The addition of the modified moment approach resulted in virtually no improvement relative to the alternate Cholesky implementation alone. None of the first three approaches proved satisfactory for nonsymmetric distributions (e.g., lognormal). The solution to this obstacle for one-sided distributions is to symmetrize the distribution according to (16), find the GQR estimate for the symmetrized distribution, and then obtain the desired distribution using (17). As we see in Table I, this renders the lognormal estimate as robust as the symmetric Gaussian distribution.

The final step in obtaining the nodes and weights involves finding the eigenvalues and eigenvectors of the matrix J in (27). The eigenvector q_j corresponding to the eigenvalue t_j is found from the equation

$$Jq_j = t_j q_j \quad j = 1, 2, \dots, N. \quad (33)$$

The eigenvalues $\{t_j\}_{j=1}^N$ are the nodes of the GQR and the positive weights are given by

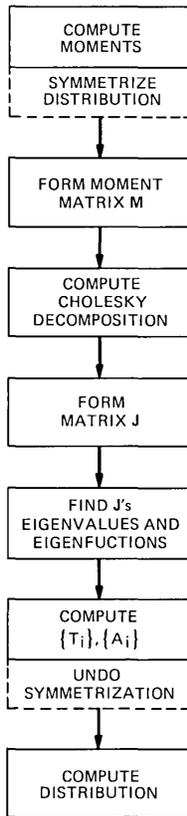


Fig. 7—Flowchart of GQR algorithm.

$$A_j = q_{1j}^2 \mu_0, \quad (34)$$

where

$$\underline{q}_j^T = (q_{1j}, q_{2j} \cdots q_{N,j}).$$

A flowchart of the steps used to compute GQR is shown in Fig. 7.

A 9.6-kb/s DSP Speech Coder

By R. E. CROCHIERE, R. V. COX, J. D. JOHNSTON, and
L. A. SELTZER

(Manuscript received April 30, 1982)

A digital speech coder has been designed for real-time operation at a data rate of 9.6 kb/s. The design is based on a combination of two speech compression techniques: Time-Domain Harmonic Scaling (TDHS) and Sub-Band Coding (SBC). It is a highly modularized hardware implementation using five Bell Laboratories Digital Signal Processor (DSP) integrated circuits as the key processing elements. Three DSPs are used in the encoder for pitch detection, TDHS compression, and sub-band encoding. Another two DSPs are used in the receiver for sub-band decoding and TDHS expansion. In this paper we describe the overall design of the system and discuss some of the techniques used to realize it in the DSP hardware in real time. General issues of the algorithm design, software implementation, and hardware design are considered.

I. INTRODUCTION

The subject of digital speech encoding and bit-rate compression has been one of considerable interest in recent years. Attention has focused strongly on bit rates in the range of 9.6 kb/s for applications where good "communications quality" is required and where robustness across a broad range of background noise conditions and speaker variations is necessary.^{1,2} This bit rate appears, at present, to be about the lowest practical rate at which this standard of quality and robustness can be reliably achieved. Below 9.6 kb/s presently known techniques have a noticeable synthetic quality and are considerably more fragile to differences in speakers and background conditions.¹

Several encoding methods for achieving "communications quality" at 9.6 kb/s have been proposed and studied. Most of these methods involve a considerable amount of signal processing to meet these goals and are thus referred to as "high-complexity" algorithms. Their imple-

mentation in real time typically requires the use of specially designed high-speed digital hardware or array-processing digital computers.^{3,4}

In recent studies,² we found that a combined technique of sub-band coding (SBC) and time-domain harmonic scaling (TDHS) leads to an encoding approach whose quality is comparable to, or better than, some of the previously studied "high-complexity" algorithms. The interesting aspect of this approach is that it is a combination of two relatively "low-complexity" algorithms that are amenable to real-time implementation using currently available technology. Thus, this TDHS/SBC approach appears to be an attractive, economical candidate for real-time implementation of 9.6-kb/s speech encoding using a relatively small amount of hardware.

In this paper we discuss the design of a 9.6-kb/s speech coder based on the above TDHS/SBC approach. The design is highly modularized around the use of the Bell Laboratories Digital Signal Processor (DSP) integrated circuit.^{5,6} Three DSPs are used in the encoder for pitch detection, time-domain harmonic-scaling compression, and sub-band encoding, respectively. Another two DSPs are used in the receiver for synchronization, sub-band decoding, and time-domain harmonic-scaling expansion. Essentially all of the signal processing in the coder is performed by the DSPs with only a minimal amount of support hardware for interfacing, clock generation, and input/output (I/O) buffering.

This paper discusses aspects of the overall design and the design parameters of the coder. Other papers⁷⁻⁹ discuss in more detail the software implementation of the pitch, TDHS, and SBC algorithms on the DSPs and the architecture of the multiple-DSP hardware used to implement the coder.

II. THE TDHS/SBC ALGORITHM

The TDHS/SBC system is basically a cascade of two different speech-compression algorithms, TDHS and SBC. Figure 1 gives a basic block diagram of this approach. The sampled input signal $s(n)$ is first compressed in bandwidth and sampling rate by the TDHS algorithm to form the intermediate signal $s_c(n)$. This processing is performed in a pitch-synchronous manner. Consequently, a pitch detector is required in the system. The SBC encoder digitally encodes the compressed signal, $s_c(n)$, to form the encoded data. This data, multiplexed with the encoded pitch and appropriate frame-synchronization information, forms a 9.6-kb/s bit-stream which is sent over the digital channel. In the receiver the digital signal is first synchronized into frames and demultiplexed into pitch and SBC data. The SBC data is decoded by the SBC decoder to form the intermediate signal, $\hat{s}_c(n)$, which is a quantized version $s_c(n)$. It is then expanded back to its original bandwidth and

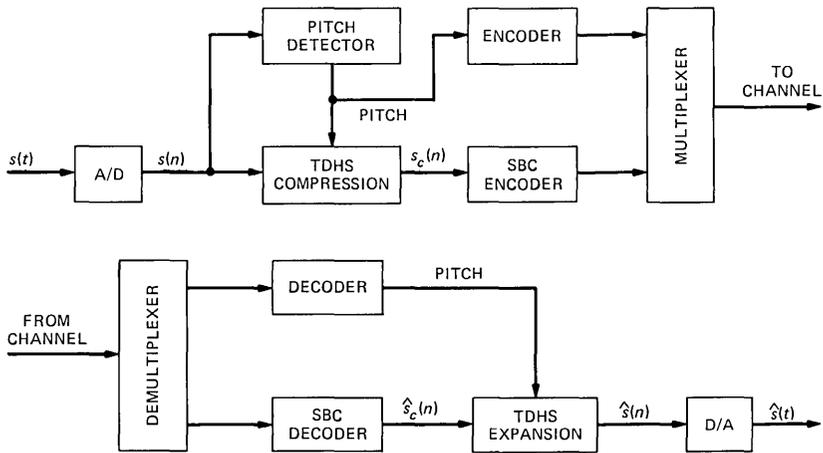


Fig. 1—Basic block diagram of the TDHS/SBC system.

sampling rate by a TDHS expansion algorithm to produce the signal $\hat{s}(n)$, which is a decoded replica of the input $s(n)$.

The two algorithms operate on different properties of redundancy of the speech signal to achieve their compression. The TDHS algorithm takes advantage of the pitch structure or pseudoperiodic nature of speech through a pitch synchronous process.¹⁰ It effectively interpolates two pitch periods of signal into one period to achieve a signal-compression factor of two. Further details of this approach are reviewed in Section V.

The SBC algorithm is a waveform coding technique that achieves a bit-rate compression by adaptively quantizing the speech in frequency bands. It takes advantage of the properties of temporal nonstationarity, spectral-formant structure, and auditory masking in speech production and perception.^{1,9,11} Further details of this method are outlined in Section VI.

Because the TDHS and SBC algorithms operate on different properties of redundancy of speech, they are highly complementary and “noncompetitive” in their operation. Thus, the overall compression of the cascaded system is effectively the product of the two individual compression factors.

A second advantage of this cascaded approach is that the degradations introduced by the two compression techniques are perceptually different. Degradations introduced by TDHS compression appear as a form of reverberance, whereas degradations introduced by SBC coding appear in the form of quantizing noise and intermodulation distortion. Since these degradations combine in different perceptual “dimensions,” the overall perceived degradation tends to be less objec-

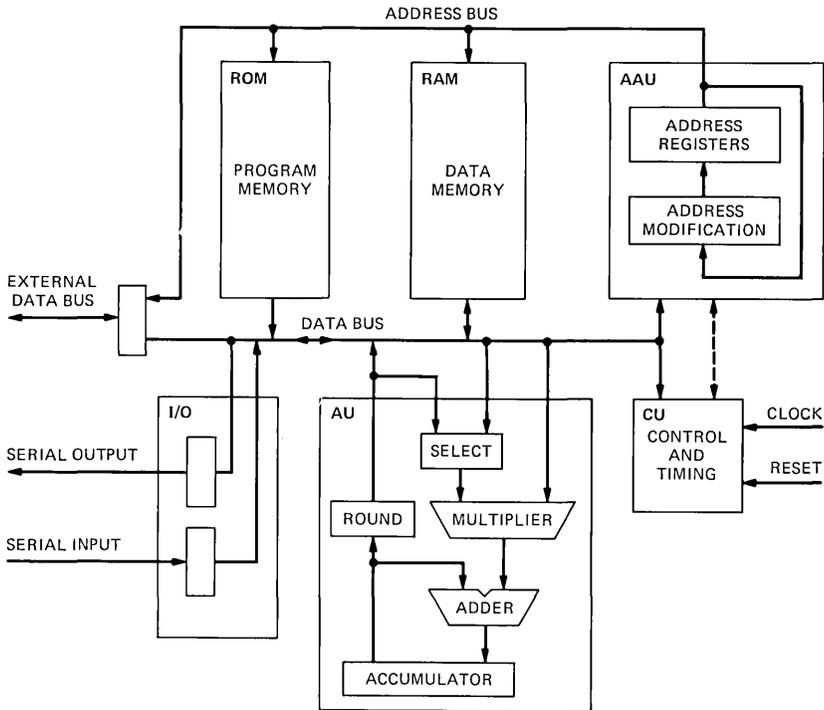


Fig. 2—Block diagram of the DSP.

tionable than if they were combined along the same perceptual “dimension.”²

Finally, a third desirable feature of this algorithm is that the compression introduced by the TDHS algorithm allows the SBC algorithm to be computed at effectively one-half the computation rate of that which would be required with an uncompressed signal. Thus, it leads to a system that is efficient computationally, as well as one that can be modularized into a system of smaller algorithmic units.

III. BASIC HARDWARE CONFIGURATIONS

The hardware design is highly modularized using the DSP as the key processing element in each module. Therefore, it is useful to review the basic characteristics of this device. Figure 2 shows a basic block diagram of the DSP.⁵ The main elements are: (i) a 1024-word \times 16-bit read-only memory (ROM) for instruction and coefficient storage, (ii) a 128-word \times 20-bit random-access memory (RAM) for variable data storage, (iii) an address arithmetic unit (AAU) with address registers for controlling memory access, (iv) a data arithmetic unit (AU) with

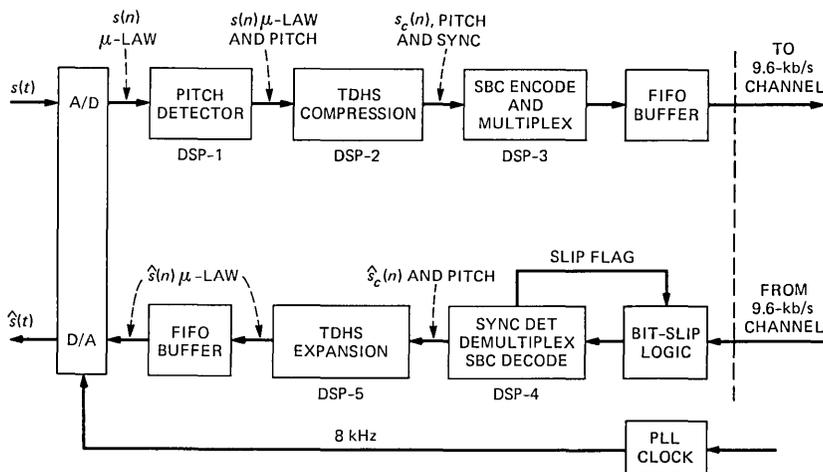


Fig. 3—Block diagram of the hardware configuration for the 9.6-kb/s coder.

provision for multiplication, full product accumulation, rounding, and overflow protection, (*v*) an I/O unit to control serial data transmission in and out of the circuit, and (*vi*) a control unit that provides instruction decoding and processor synchronization.

The processor operates with an 800-ns machine cycle time, which is established by a 5-MHz clock. In one machine cycle it can: (*i*) decode an instruction, (*ii*) fetch data and perform a 16×20 -bit multiplication, (*iii*) accumulate the output products from the multiplier, and (*iv*) store data in memory.

Figure 3 is a simplified block diagram of the hardware architecture for the 9.6-kb/s coder. It consists of cascaded connections of DSP modules with data passing from one DSP to the next in multiplexed form. The encoder contains three DSPs in series with an analog-to-digital (A/D) converter at the input and a first-in-first-out (FIFO) buffer at the output to the channel. The decoder contains two DSPs with a FIFO buffer between the second DSP and the output to the digital-to-analog (D/A) converter. In addition, a logic circuit is necessary at the receiver between the input of the 9.6-kb/s channel and the first DSP for use in bit slipping for synchronization on startup.

The analog input, $s(t)$, is first converted to digital form $s(n)$ by a μ -law A/D converter.¹² An 8-KHz clock signal for the sampling rate of the A/D converter is generated from the 9.6-KHz channel clock using a phase-locked loop (PLL) circuit. The first DSP is used to implement the pitch detector. It passes the μ -law signal $s(n)$ through to the second DSP along with multiplexed pitch information. The second DSP is used for the TDHS compression algorithm. Its output consists of the com-

pressed signal, $s_c(n)$, multiplexed with pitch and synchronization information. The third DSP is used for the SBC encoder and for multiplexing the pitch, synchronization, and SBC-encoded data into a 9.6-kb/s serial bit stream. Its output is passed to the channel in the form of 16-bit serial words. A FIFO memory is used to control the flow of data between the DSP and the channel.

In the decoder the fourth DSP receives the serial bit stream directly from the 9.6-kb/s channel in the form of 16-bit words. A bit-slipping logic circuit is used to interface the channel to the DSP; it is controlled by a flag from the DSP. The circuit is used to align the first bit of a frame of data to the first bit of a 16-bit word in the synchronization locking mode of the decoder. Once frame synchronization is established, the fourth DSP is used to demultiplex the pitch and SBC data and perform the computation for the SBC decoder. The SBC decoded output, $\hat{s}_c(n)$, and the pitch are then passed to the fifth DSP, which computes the TDHS expansion algorithm. The output, $\hat{s}(n)$, of this DSP is the decoded version of $s(n)$ and it is passed to the D/A (in μ -law PCM format) through a second FIFO buffer.

In the next five sections we discuss each of the components of this system in more detail. Section IV discusses the design of the pitch detector, Section V discusses the TDHS algorithms, and Section VI discusses the design of the SBC. The overall framing and multiplexing structure for the coder is discussed in Section VII, and Section VIII discusses the synchronization detection algorithm used for frame alignment in the receiver. Finally, Section IX discusses the performance of the system.

IV. THE PITCH DETECTOR

The pitch detector is based on a modification of the autocorrelation-type pitch-detection algorithm,¹³ and it is designed for implementation in a single DSP.⁷ Figure 4 illustrates this basic approach. The input speech signal $s(n)$ is first converted to a linear PCM signal and lowpass filtered to remove spectral energy above 1 KHz. The resulting low-pass filtered signal, $x(n)$, is then used to compute the autocorrelation-function estimate at time n

$$r_n(m) = \sum_{\ell=-\infty}^{\infty} f(n-\ell)x(\ell)x(\ell-m), \quad (1)$$

where m denotes the autocorrelation lag and $f(n)$ corresponds to the analysis window over which $r_n(m)$ is computed. The pitch period is then defined as the lag m_0 over the range of the allowed set of lag values $\{m\}$ for which $\tilde{r}_n(m)$ is maximum, i.e.,

$$\tilde{r}_n(m_0) = \max_m [r_n(m) \cdot g(m)] \quad (2a)$$

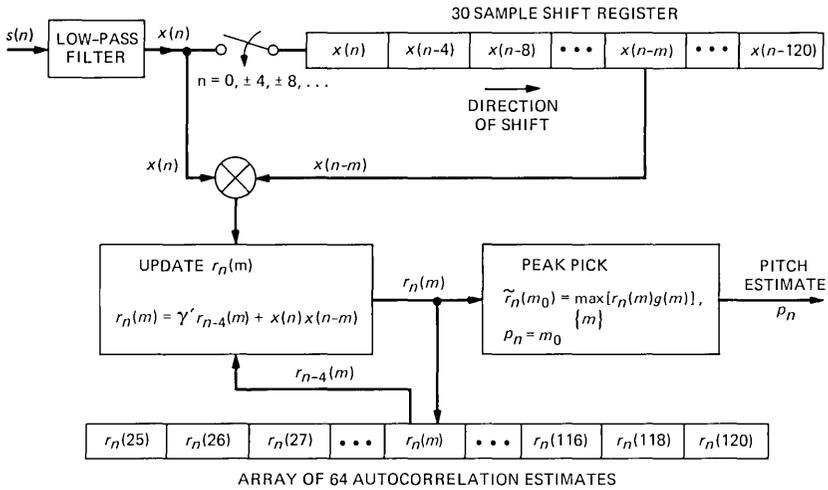


Fig. 4—Block diagram of the pitch-detector implementation.

$$\text{pitch} = p_n = m_0, \quad (2b)$$

where $g(m)$ is a weighting factor used to control the behavior of the algorithm.⁷

A total of 64 autocorrelation coefficients are included in the set $\{r_n(m)\}$ corresponding to values

$$\{m\} = \{25, 26, 27, \dots, 55, 56, 58, 60, 62, \dots, 116, 118, 120\}. \quad (3)$$

This allows the pitch period to be encoded into a 6-bit word for transmission over the channel. With a sampling rate of 8 KHz, the pitch values in the set $\{m\}$ correspond to a pitch frequency range from 66.7 Hz to 320 Hz, which spans the range of most speakers. Note also from eq. (3) that the allowed values of pitch period are more closely spaced (quantized) for low values, m , than for high values. This allows for a more uniform percentage of accuracy of pitch for high- and low-pitch speakers.

The autocorrelation estimate $r_n(m)$ is computed using an exponential window function

$$f(n) = \begin{cases} \gamma^n & n \geq 0 \\ 0 & n < 0. \end{cases} \quad (4)$$

This forms allows $r_n(m)$ to be sequentially updated according to the relation

$$r_n(m) = \gamma r_{n-1}(m) + x(n)x(n-m). \quad (5)$$

In this manner, only two multiplications and one addition are required to update each autocorrelation coefficient.

To further reduce the amount of computation and storage required in the DSP, eq. (5) was modified so that $r_n(m)$ is updated only every fourth sample, i.e.,

$$r_n(m) = \gamma' r_{n-4}(m) + x(n)x(n - m), \quad (6)$$

where γ' typically has a value of $\gamma' = 0.95$. This reduces the total amount of computation by a factor of 4. It also reduces storage requirements for the delayed signal $x(n - m)$ by allowing $x(n)$ to be decimated (reduced in sampling rate) by a factor of 4. Thus, at any one sample time only certain values of $r_n(m)$ can be updated. Computation is therefore distributed over a four-cycle process in which four input samples $s(n)$ are received, all values of $r_n(m)$ are updated, and a new estimate of pitch is determined every four sample times. A more detailed description of this computational structure and the manner in which it is implemented in the DSP is described in Ref. 7. Further details concerning the weighting parameters $g(m)$ and the performance of this design are also available in the same reference.

The speech signal $s(n)$ is passed from the pitch detector to the TDHS algorithm along with the 6-bit encoded pitch information, which is inserted after every fourth speech sample. Since the TDHS algorithm does not require information regarding the voiced or unvoiced nature of the speech, no voiced/unvoiced decision is made by the pitch detector. During voiced regions, the pitch detector measures the speech periodicity, and during the unvoiced and silence regions it gives the best estimate of any long-term correlation that may exist in the signal, even though this correlation may be low.

V. THE TDHS ALGORITHM

The TDHS algorithm compresses the input signal $s(n)$ by a factor of two such that the compressed signal $s_c(n)$ contains one-half of the original number of samples.^{2,8,10} Since the sampling rate of $s(n)$ is 8 KHz, the sampling rate of $s_c(n)$ is, therefore, 4 KHz. This compression process can be interpreted in the frequency domain in terms of a 2:1 compression of the spectral bandwidth such that the original 0- to 4-KHz bandwidth of the signal $s(n)$ is scaled to a 0- to 2-KHz bandwidth allowing the sampling rate to be reduced by the factor of two. The compression is achieved by reducing the frequency spacing between pitch harmonics and appropriately scaling the envelope of the spectrum by a factor of two.

In an alternative time-domain interpretation, the signal $s(n)$ is compressed by a factor of two by computing one pitch period of the compressed signal $s_c(n)$ from a weighted average of every two pitch periods of the input signal $s(n)$ in a pitch-synchronous manner. The

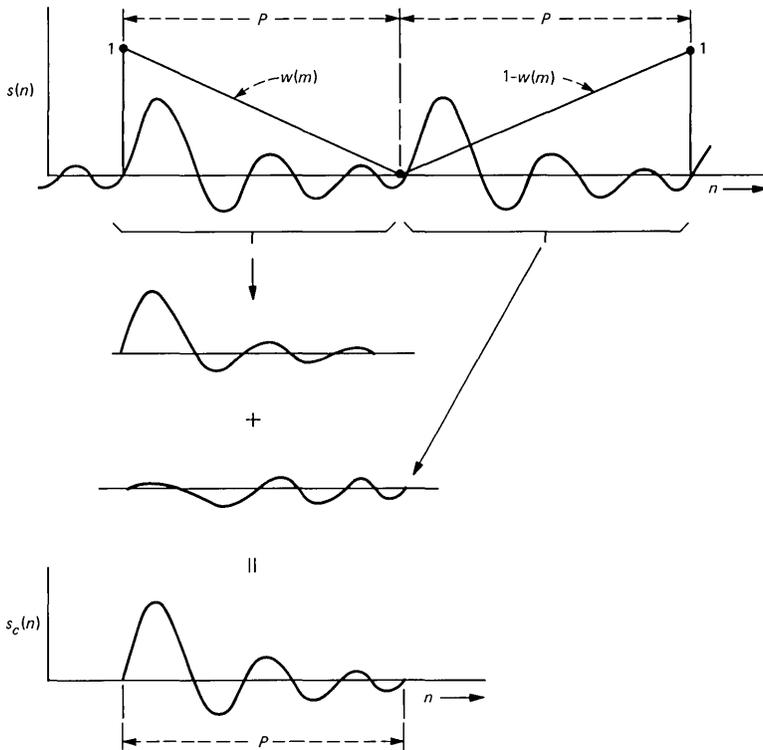


Fig. 5—TDHS compression.

TDHS algorithm is implemented according to this interpretation. Figure 5 illustrates this interpolation process for the compression algorithm. Given the pitch period, $p = p_n$, from the pitch detector, the input speech $s(n)$ is divided into blocks of $2p$ samples. One block of p samples of compressed signal $s_c(n)$ is then computed from these $2p$ samples according to the following process. The first block of p samples of $s(n)$ is weighted by a p -sample window, $w(m)$, $m = 0, 1, \dots, p - 1$, which linearly decreases from a value of 1 to 0 across the block. The second block of p samples of $s(n)$ is similarly weighted with a window $1 - w(m)$ that linearly increases from 0 to 1 across the block. The sum of the two weighted blocks then produces one block of p samples of the compressed signal $s_c(n)$, as illustrated in Fig. 5. The waveform of $s_c(n)$, therefore, looks mostly like the first block of $s(n)$ at its beginning and mostly like the second block of $s(n)$ at the end. In this way the concatenation of the blocks of $s_c(n)$ forms a continuous waveform without end effects from block to block. The next block of $s_c(n)$ is computed in the same manner as above using the next $2p = 2p_n$ samples of $s(n)$.

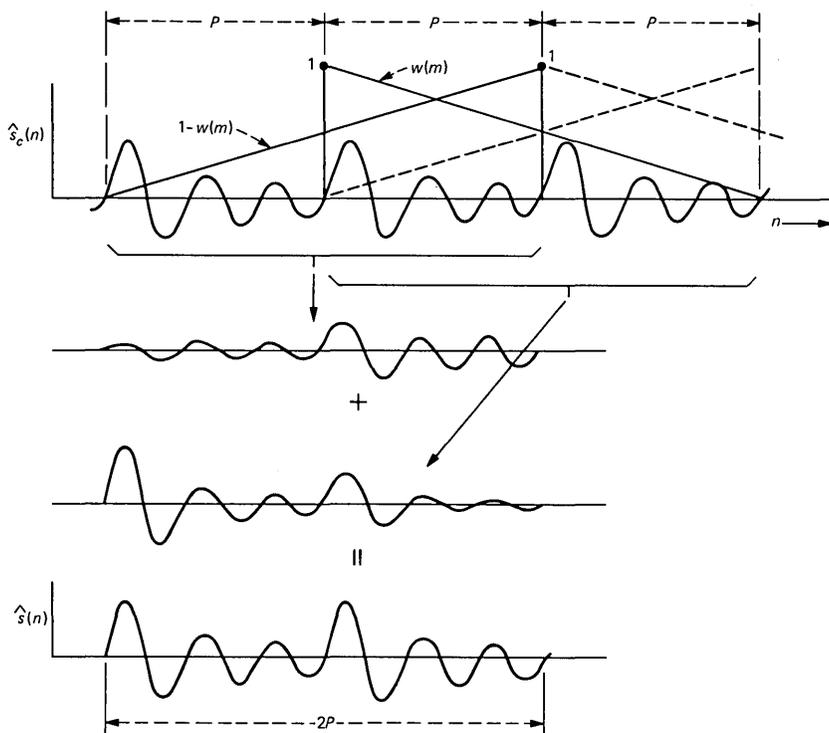


Fig. 6—TDHS expansion.

In the TDHS expansion of $\hat{s}_c(n)$ to $\hat{s}(n)$ a similar pitch-synchronous interpolation is performed. Figure 6 illustrates this process. In this case, $3p$ samples of $\hat{s}_c(n)$ are used to compute $2p$ samples of $\hat{s}(n)$ using the $2p$ -sample overlapped windows shown by the solid lines in Fig. 6. The windows are then moved over by p samples, as shown by the dashed lines in Fig. 6, and the next $2p$ samples of $\hat{s}(n)$ are computed in a similar process. Thus, for every p samples of new input signal, $\hat{s}_c(n)$, $2p$ samples of the expanded signal $\hat{s}(n)$ are computed. A careful analysis reveals that this process results in an output waveform $\hat{s}(n)$ that is continuous across the concatenated output blocks without end effects.

Although the TDHS compression and expansion algorithms have been discussed above in terms of block-processing operations, they are more conveniently implemented in the DSP in a stream-processing manner.⁸ That is, for every two input samples of $s(n)$ in the TDHS compression, one sample of $s_c(n)$ is computed. Similarly, for every input sample of $\hat{s}_c(n)$ in the TDHS expansion, two outputs of $\hat{s}(n)$ are computed. These operations are performed using the structure shown

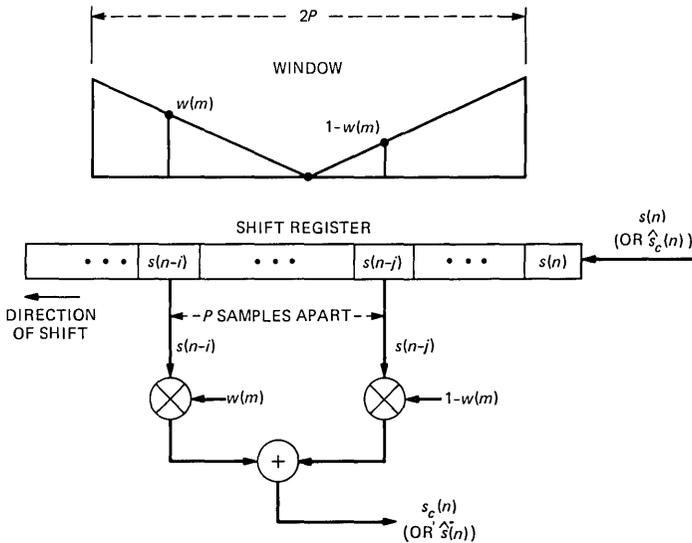


Fig. 7—Block diagram of the implementation of TDHS compression (or expansion).

in Fig. 7. A length $2P$ shift register is used to hold the input data $s(n)$ (for compression) or $\hat{s}_c(n)$ (for expansion), where

$$P = \max p_n = 120. \quad (7)$$

Input samples enter from the right and are shifted one sample to the left for each input. The computation of each output sample involves the multiplication of samples located at indices i and j in the shift register with the respective window coefficients $w(m)$ and $1 - w(m)$, as illustrated in Fig. 7. The sum of these two products is then the interpolated output sample. It also should be clear from the above discussion that the indices i and j are always spaced p samples.

For TDHS compression the data in the shift register moves two samples to the left and the indices i and j are decremented (shifted left) by one sample for each output. At the start of a new block of p samples, index i is initialized to the center of the shift register, p is initialized to p_n , j is initialized to $i + p_n$ and $w(m)$ is initialized to $w(0) = 1$. The process is then repeated for the next p samples.

For TDHS expansion the opposite process occurs. Since the output samples are generated at twice the rate of the input samples, the indices i and j must be incremented (shifted right) by 1 after each output sample. At the start of a new block, index j is initialized to the center of the shift register and i is initialized to $j = p$. Further details on how these operations are implemented in the DSP are covered in another paper.⁸

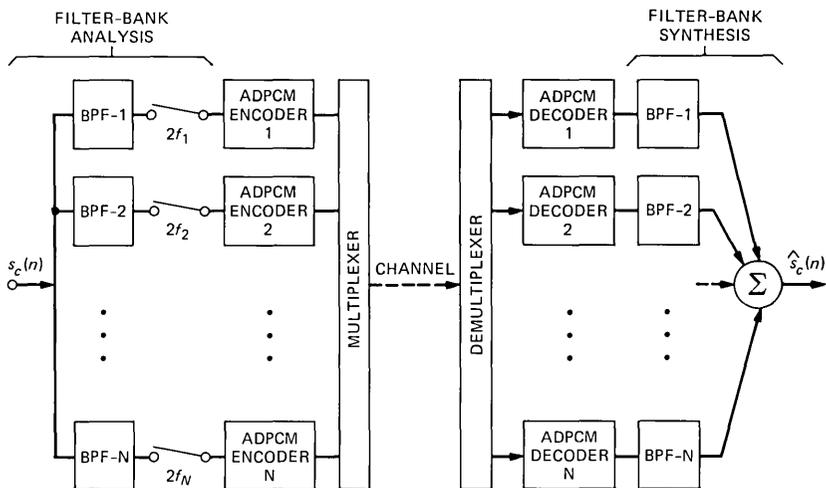


Fig. 8—Block diagram of the SBC encoder and decoder.

Another novel aspect of the above DSP implementation of TDHS compression and expansion algorithms involves the amount of memory required for the shift register. From eq. (7) and Fig. 7 we can see that the required length of the shift register is $2P = 240$ samples, whereas the DSP contains only 128 locations of RAM. This conflict is conveniently solved by using the capability of the DSP to convert between 8-bit μ -law and 20-bit linear PCM formats. By storing the data in the shift register in 8-bit μ -law format and packing two 8-bit words into each 20-bit RAM location, the effective memory of the DSP is doubled, allowing the TDHS algorithms to be completely implemented within individual DSPs without external memory. Also, by carefully organizing the "high" and "low" 8-bit words of memory it is possible in the TDHS algorithms to associate "high" memory with one of the indices i or j and "low" memory with the other. This greatly simplifies the software implementation.⁸

VI. THE SUB-BAND CODER

The sub-band coder (SBC) encodes the compressed signal $s_c(n)$ into a digital bit stream. It is a waveform coding technique that takes advantage of the temporal and spectral properties of speech production and speech perception by partitioning the signal into a set of sub-bands by a filter bank (see Refs. 1, 9, 11, and 14). Each sub-band is effectively bandpass filtered, low-pass translated to dc , sampled at its Nyquist rate (twice the width of the sub-band), and then digitally encoded using adaptive differential PCM (ADPCM). Figure 8 illustrates

Table I—Sub-band coder design

Band	Frequency Range (Hz)		Sampling Rate (Hz)	Bits/Sample	Bit Rate
	Uncompressed	Compressed			
1	0-500	0-250	500	5	2500
2	500-1000	250-500	500	4	2000
3	1000-2000	500-1000	1000	2	2000
4	2000-3000	1000-1500	1000	2	2000
TOTAL	0-3000				8500

a simplified block diagram of this process. In the receiver the digital signal $\hat{s}_c(n)$ is reconstructed by decoding the sub-band signals, interpolating them, translating them back to their original spectral locations, and then summing them to form the decoded signal $\hat{s}_c(n)$.

The sub-band framework offers several advantages. Quantization noise is contained in bands to prevent masking of one frequency band by quantizing noise in another frequency band. Separate adaptive quantizer step sizes are used so that bands with lower signal energy have lower quantizer step sizes and contribute less quantization noise. By appropriately allocating bits in different bands, the shape of the quantization noise can be controlled in frequency. In the lower frequency bands, where pitch and formant structure must be accurately preserved, a larger number of bits/sample are used, whereas in upper frequency bands, where fricative and noise-like sounds occur in speech, fewer bits/sample are used.

A four-band SBC design is used in the 9.6-kb/s coder. Table I summarizes the choice of bands, sampling rates, and bits/sample used in this design. Bandwidths are given with respect to both the uncompressed and compressed frequency scales. Therefore, the total bandwidth of the coder is 3 KHz, relative to the uncompressed frequency scale. The total bit rate of the SBC coder is 8.5 kb/s, leaving the remaining 1.1 kb/s for transmission of pitch and framing information.

Figure 9a illustrates the manner in which the SBC analysis filter bank is implemented. The design is based on the use of quadrature mirror filter (QMF) designs,^{9,15} which are implemented in terms of polyphase structures.¹⁶ The QMF approach allows a signal band to be divided into two equally spaced, high-pass filtered (HPF) and low-pass filtered (LPF) sub-bands, which are each reduced in sampling rate by a factor of two. This process is accomplished with a pair of symmetric finite impulse response (FIR) high-pass and low-pass filters. Because of the symmetry and the quadrature mirror relationship of these two filters, their coefficients are identical except for the signs of the odd-numbered coefficients. This property allows the computation to be shared between the two filters by separately computing the even taps, $h_e(n)$, and the odd taps, $h_o(n)$, of the low-pass filter $h(n)$.^{9,15} The sum of these two partial computations gives the output for the lower band

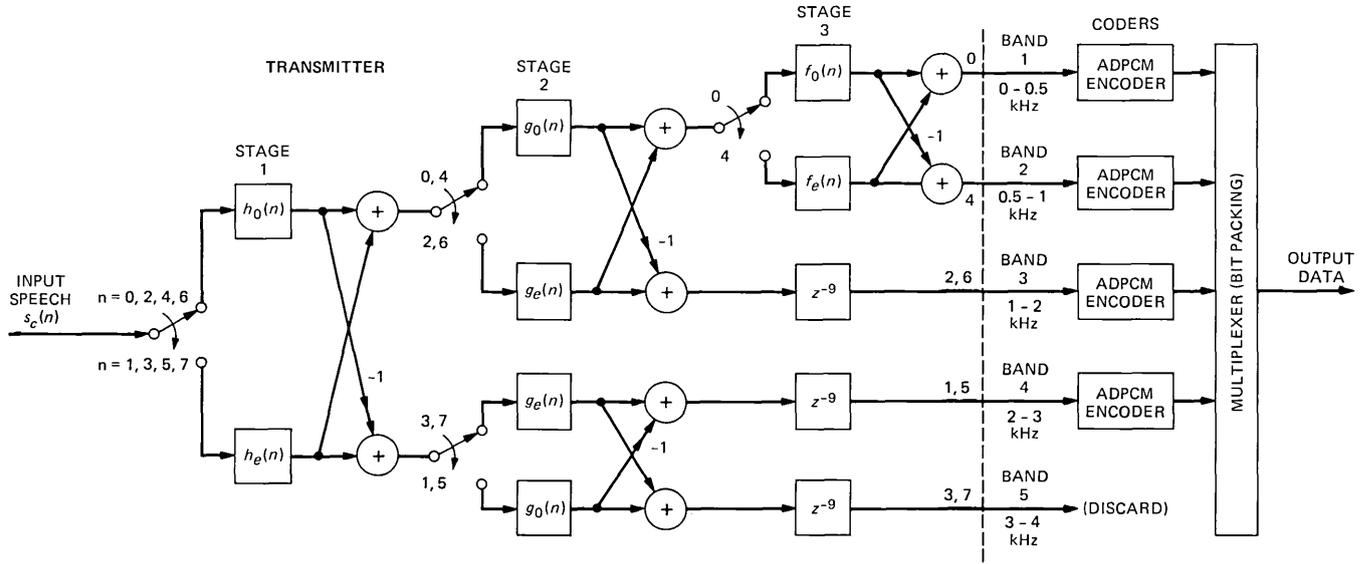


Fig. 9 (a)—QMF filter bank tree structure for analysis.

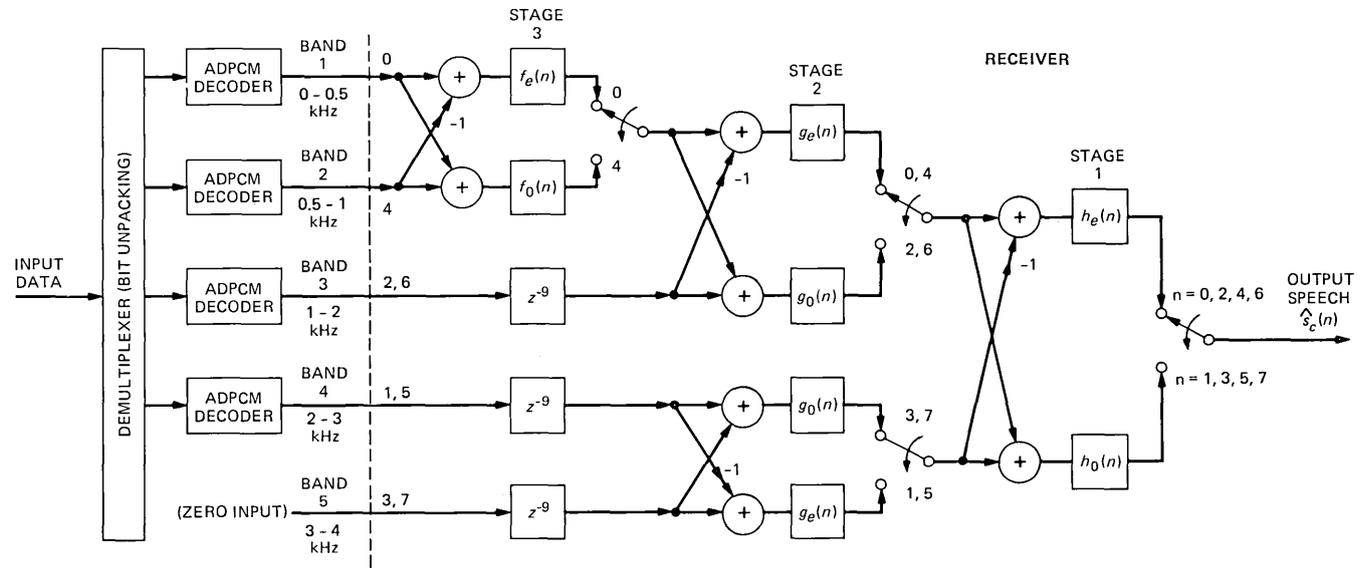


Fig. 9 (b)—QMF filter bank tree structure for synthesis.

Table II—Filter coefficients for $h(n)$, $g(n)$, and $f(n)$

n	$h(n)$	$g(n)$ and $f(n)$
0	0.0006506	0.0065257
1	-0.0013508	-0.0204875
2	-0.0012601	0.0019911
3	0.0041581	0.0464768
4	0.0014272	-0.0262756
5	-0.0093636	-0.0992955
6	-0.0001722	0.1178666
7	0.0178820	0.4721122
8	-0.0041094	
9	-0.0311553	
10	0.0144688	
11	0.0529093	
12	-0.0392449	
13	-0.0998001	
14	0.1284651	
15	0.4664583	

and the difference gives the output for the upper band. By applying this two-band splitting process in the tree structure in Fig. 9a, the band structure in Table I is obtained where the highest band (band 5) is discarded in the process. In the receiver a similar tree structure (Fig. 9b) is used in reverse order to recombine the sub-bands into a full-band signal. The delays in bands 3 and 4 in Fig. 9 are used to compensate for the processing delay caused by the extra split used to generate bands 1 and 2.

The above QMF filter-bank design offers two advantages. First, it leads to an efficient means of computing the filters by sharing computation among bands. Second, because of the quadrature nature of the design, aliasing terms, which are generated in the process of sampling rate reduction in the encoder, are canceled by imaging terms, which are generated in the filter-bank interpolation process in the receiver. This property allows the use of relatively low-order finite impulse response (FIR) filter designs in the coder.

The low-pass filter design, $h(n)$, for the first stage of QMF splitting is accomplished with a 32-tap FIR filter. The coefficients for this design¹⁷ are given in Table II, column 2. Note that only the first symmetric half of the coefficients is given. The remaining coefficients can be obtained from the relation $h(n) = h(31 - n)$ for $n = 16$ to 31. The filter designs $g(n)$ and $f(n)$ for the second and third stages of QMF splitting are accomplished with identical 16-tap FIR filters. Table II, column 3, gives the coefficients for this design. Again, only the first symmetric half of the coefficients is given and coefficients from $n = 8$ to 15 can be obtained from the relation $g(n) = g(15 - n)$. All coefficient values are quantized to the 16-bit accuracy of the DSP in the actual implementation.

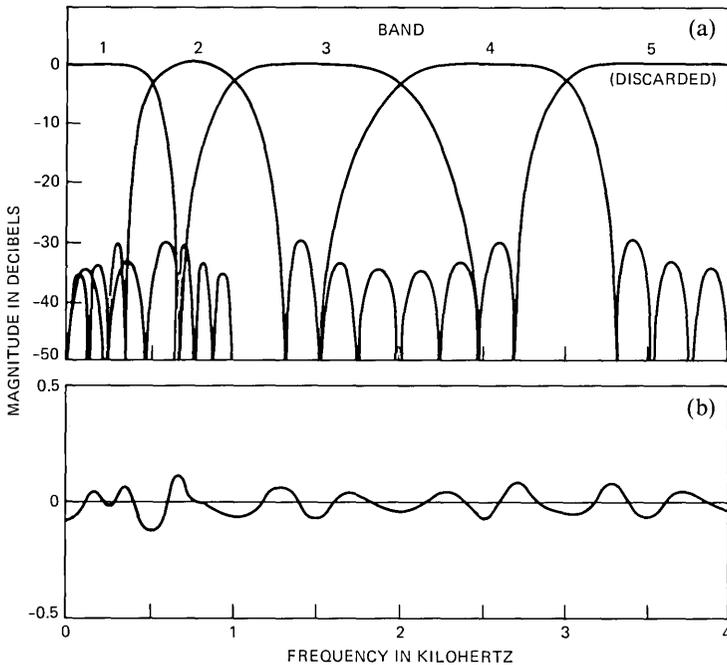


Fig. 10—Frequency response for (a) the composite five-band filter bank, and (b) the analysis/synthesis reconstruction based on the filter bank.

The frequency response for the combined filter bank is shown in Fig. 10a, where the band structure given in Table I is clearly apparent. Note also that because the filters in the second and third stages of the structures in Fig. 9 are implemented at lower sampling rates, their frequency responses are respectively scaled to narrower widths in the combined system. Finally, Fig. 10b shows a plot of the total frequency response of the back-to-back filter bank (without ADPCM coding). A total reconstruction error of less than 0.25 dB is observed.

The ADPCM coders for quantizing the sub-band signals are identical to the DSP design discussed in Ref. 18, except for the choice of design parameters. Therefore, we will only briefly review the design and discuss the choice of parameters. Figure 11 is a block diagram of this design. The main elements of the coder are: (i) a b -bit PCM quantizer (where b can be varied from 2 to 5), (ii) tables to store the step size and inverse step size, (iii) a step-size adaptation circuit to control the choice of step size (i.e., the addresses to the tables), and (iv) a first-order fixed predictor.

A predictor signal $p(n)$ is first computed by scaling the previous decoded sub-band signal $\hat{s}_c(n-1)$ by the predictor coefficient β , where for telephone speech, values of β will be close to zero (e.g., 0, -0.4, 0,

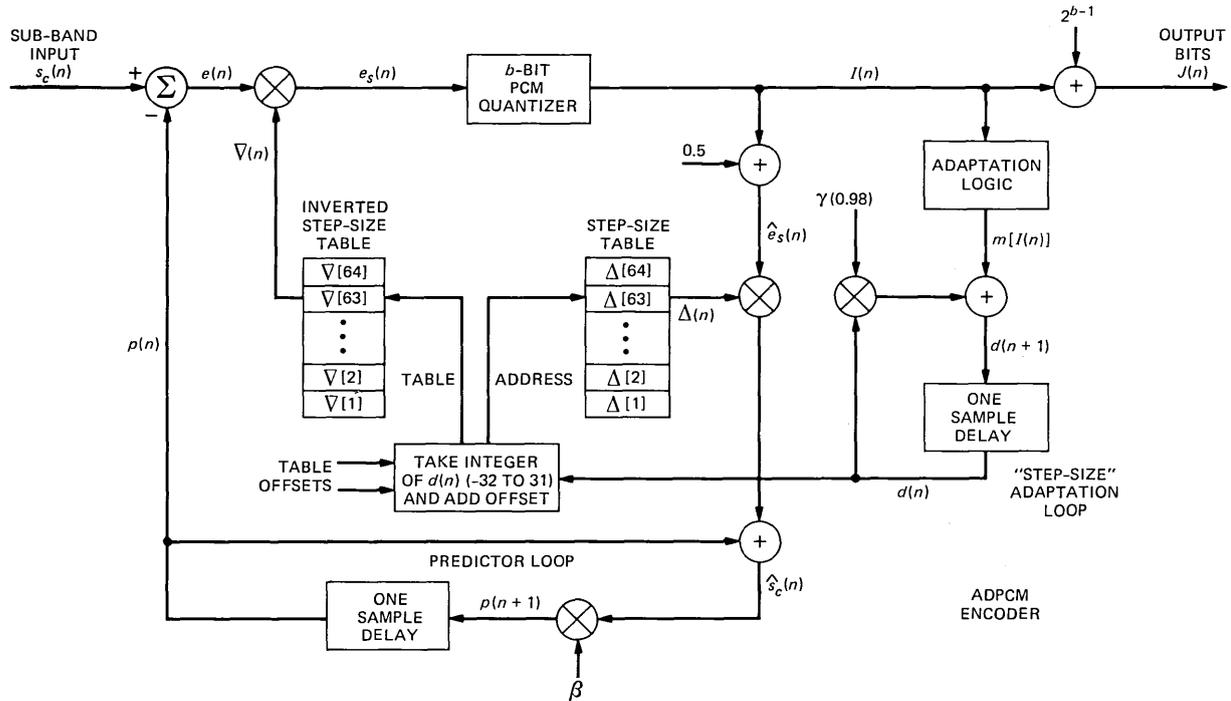


Fig. 11 (a)—Block diagram of the ADPCM encoder.

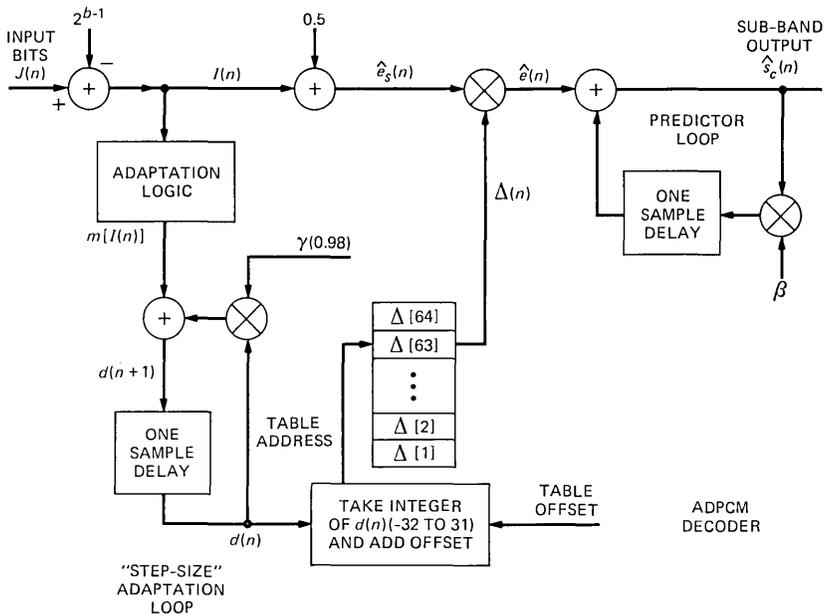


Fig. 11 (b)—Block diagram of the ADPCM decoder.

0 for bands 1 to 4^{11}). The signal $p(n)$ is then subtracted from the sub-band signal $s_c(n)$ to produce the difference $e(n)$, which is adaptively quantized. This is accomplished by first scaling $e(n)$ down by the inverse step size $\nabla(n)$ and limiting and rounding (PCM coding) to the b -bit integer $I(n)$. By adding the value 2^{b-1} to $I(n)$ a positive b -bit code word, $J(n)$, is obtained, which can be conveniently multiplexed with code words from other sub-bands (as well as pitch and synchronization information).

The b -bit integer, $I(n)$, is also used in the encoder and decoder for obtaining the decoded output $\hat{s}_c(n)$ and for step-size control. Decoding is accomplished by adding 0.5 to $I(n)$ and scaling it back up by the step size $\Delta(n)$. By adding the predictor value $p(n)$ to this quantized difference we obtain the decoded signal $\hat{s}_c(n)$, which is the output of the decoder. It is also used in both the encoder and decoder for generating the next predictor value.

The step size $\Delta(n)$ and its inverse $\nabla(n)$ are determined by the table address, which is adaptively incremented or decremented by the step-size control. The size of the tables are 64 words and they span a 60-dB step-size range (i.e., 0.9375 dB/word). Their address is obtained from the integer part of the variable $d(n)$ (see Fig. 11), which is limited to the range -32 to $+31$. The value of $d(n)$ is stored in a "leaky integrator" with a leak factor of 0.98. In this way $d(n)$ "drifts" toward zero (the

Table III—Quantizer adaptation parameters

No. of Bits	m_1	m_2	Attack Rate (dB/sample)	Decay Rate (dB/sample)
2	4.88	-0.98	4.6	-0.92
3	5.47	-1.27	5.1	-1.2
4	5.47	-1.27	5.1	-1.2
5	5.47	-1.27	5.1	-1.2

center of the tables) and mitigates the effects of channel error. If an outermost (positive or negative) quantizer level is used, a positive value, $m[I(n)] = m_1$ is added to $d(n)$, and if an innermost (positive or negative) level of the quantizer is used, a negative value, $m[I(n)] = m_2$ is added to $d(n)$. In this way the step size is dynamically varied to match the range of $e_s(n)$ to the range of the PCM quantizer. Table III shows typical values of m_1 and m_2 for $b = 2$ to $b = 5$ bit quantizers in SBC and their corresponding "attack" and "decay" rates (in dB/sample) at which they can respectively expand or contract their step sizes.

The manner in which the computation in the SBC is performed is strongly determined by the three-stage framework of the filter bank. Since the sampling rate is reduced by a factor of two at each stage, computation in different stages must be performed at different rates. This is accomplished by distributing the total computation of the coder (and decoder) over an 8-cycle process. In each cycle one input sample $s_c(n)$ is received in the coder, one path of the filter band "tree" is computed, and one ADPCM coder is implemented. The cycle numbers $n = 0, 1, 2, \dots, 7$ in Fig. 9 indicate the paths of computation taken in each cycle. For example, in cycle 0 of the SBC encoder, filters $h_0(n)$, $g_0(n)$, and $f_0(n)$ in the uppermost branch of the tree are updated and the ADPCM coder in band 1 is computed. As can be seen from this structure, stages and coders in the tree that have higher sampling rates are computed more often and those with lower sampling rates are computed less often.

The tree structure also determines the manner in which output bits from the ADPCM coders are multiplexed and framed for transmission. Table IV lists the cycle number and the sub-band that is coded in each cycle. After eight cycles of processing, eight input samples are received and 17 output bits are generated. The process is then repeated.

VII. MULTIPLEXING AND FRAMING OF THE DATA

The 9.6-kb/s data from the TDHS/SBC coder is multiplexed into frames of 96 bits each for serial transmission over the channel. Each frame corresponds to 10 ms of encoded speech. The multiplexing is performed in the third DSP along with the multiplexing of the sub-band coder data. The output from this DSP is in the form of 16-bit words

Table IV—Order of computation in the sub-band coder

Cycle	Band	Bits
0	1	5
1	4	2
2	3	2
3	(5)	—
4	2	4
5	4	2
6	3	2
7	(5)	—

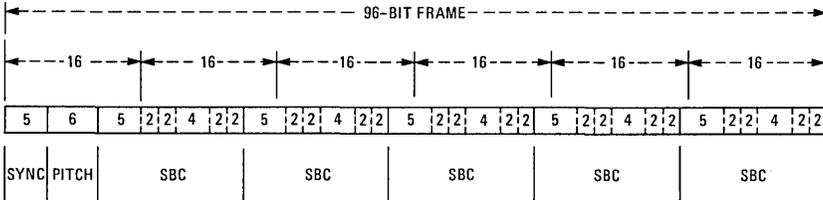


Fig. 12—Frame structure for the 9.6-kb/s bit stream.

such that six words form one 96-bit frame. Figure 12 summarizes this multiplexing structure for one frame of data.

The first five bits in the frame represent a synchronization header, which is used for frame locking in the receiver. The next six bits are used to transmit pitch from the pitch detector. The remaining bits are used for the transmission of five groups of 17 bits each of SBC data, where each group comprises one 8-cycle computational loop in the SBC coder. From Fig. 12 we can see that one 16-bit word can be obtained from the DSP after encoding band 1 in each 8-cycle loop of the SBC coder, and a final 16-bit word can be transmitted at the end of the frame. In the receiver a similar decoding process is performed in the demultiplexer of the SBC decoder.

VIII. SYNCHRONIZATION DETECTION AND DATA ALIGNMENT

Before the data in the receiver can be decoded, the framing structure in the bit stream must first be identified and the data must be aligned with the 16-bit input words according to Fig. 12. This synchronization detection and data alignment is accomplished in the first DSP of the receiver with the aid of the bit-slipping logic at the input from the channel (see Fig. 3).

Figure 13 shows a block diagram of this process. The algorithm is divided into two modes of operation, a synchronization search mode and a run mode. In the search mode the DSP receives input from the channel in the form of 16-bit words. If the first five bits match the bit

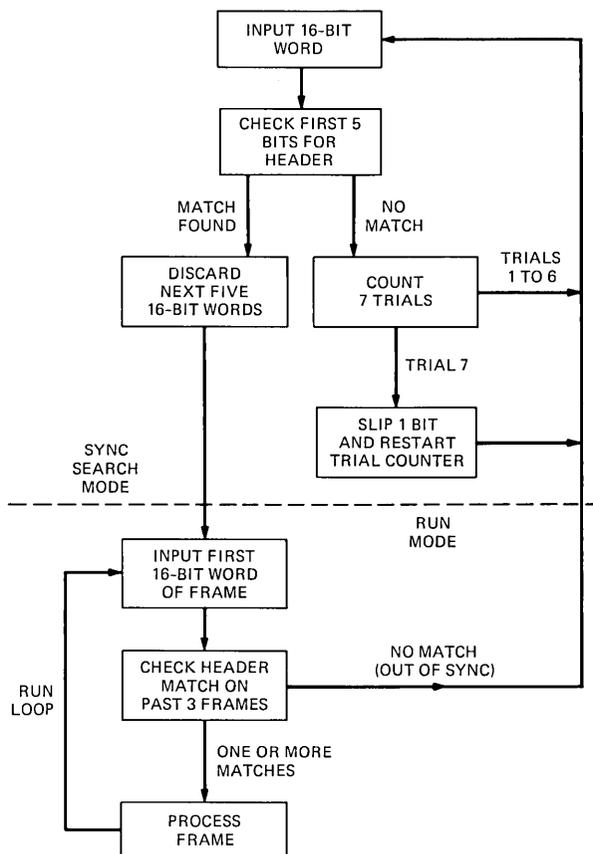


Fig. 13—Block diagram of the synchronization detection and data alignment algorithm.

pattern of the header, it is assumed that the first word of a frame has been found. It then discards the next five words of the frame and proceeds to the run mode.

If no match is found, it waits for the next 16-bit input and tries again. If it is unsuccessful in finding a header after seven trials it assumes that the frame structure is out of alignment with the word structure. It then sends a flag to the bit-slipping logic, which discards one bit in the bit stream. The search process is then repeated. A total of seven trials are used to avoid the possibility of skipping a bit in the middle of the header.

Once the algorithm is in run mode it continues to check the header at the beginning of each frame. If a match is found in one or more of the past three frames it assumes that the algorithm is synchronized

and proceeds to process the frame. In this way the algorithm is more resistant to channel errors that may occur in the header. If no match is found after three consecutive frames, it assumes that the system is out of synchronization (or had received a false start) and it returns to the search mode.

The above algorithm works well and uses a minimum amount of code in the DSP. The bit pattern for the header was chosen to be the 5-bit sequence 10010.

IX. PERFORMANCE

In terms of quality, the real-time DSP version of the TDHS/SBC coder matches the quality of the computer simulations.² The degradations introduced by TDHS compression/expansion are generally perceived as a form of reverberance, and degradations introduced by the SBC coder are generally perceived as a form of quantization noise and harmonic distortion.

In a recent study involving TDHS/SBC coding of telephone network speech it was found that the perceived reverberance caused by the TDHS processing was more noticeable than that for high-quality microphone speech.¹⁹ This implies that some caution must be exercised in applications involving a direct tandem connection of TDHS/SBC with the telephone network environment. Later we will show how this effect can be mitigated to some extent. For applications where the characteristics of the transducer can be controlled, this poses no problem.

The reason for the increased reverberance for telephone network speech was found to be a consequence of the strong pre-emphasis caused by the 500-type telephone set specifications and the tight band-pass filtering (200 to 3200) of the *D*-channel bank. Figure 14 shows the resulting frequency response of the network environment.¹⁹ This can be compared to the flat response of a high-quality microphone.

The network frequency response of Fig. 14 has two effects on the perceived quality of TDHS processing. Both effects stem from the fact that TDHS depends strongly on an accurate pitch measurement for its performance, as well as the assumption that voiced speech can be modeled as a pseudoperiodic signal. For example, if Δf represents the error in the measurement of the fundamental pitch harmonic (caused by measurement error or quantization of the pitch) the *n*th harmonic will have an error of $n\Delta f$. This means that TDHS processing will always degrade the high-frequency regions of speech more than the low-frequency regions. For high-quality microphone speech this high-frequency reverberance is not very apparent because for voiced sounds the high-frequency regions are relatively low in amplitude. In addition, the strong energy in the first formant helps to perceptually mask any degradations in the high-frequency regions. During unvoiced regions

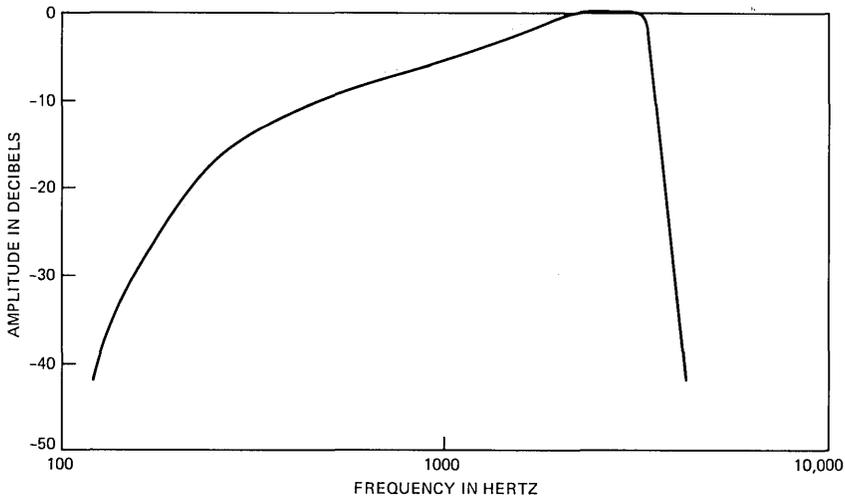


Fig. 14—Frequency response of the telephone network environment.

the noise-like character of speech tends to mask most of the effects of TDHS processing.

The effect of the telephone pre-emphasis in Fig. 14 is to amplify the high-frequency region of the speech by a factor of 15 to 20 dB relative to the low-frequency region. This also amplifies the high-frequency reverberance of TDHS as well. The effect of the high-pass filtering from 0 to 200 Hz effectively removes most of the fundamental harmonic of the speech and consequently reduces the masking process. Thus, both effects contribute to enhancing the perceived reverberance of the TDHS processing. Therefore, the problem is one of auditory perception rather than the TDHS algorithm performing differently on telephone or microphone speech. We found that if high-quality microphone speech is filtered with the same filter response as that of Fig. 14, either before or after TDHS/SBC coding, we get the same effect as with telephone network speech.

The effects of the pre-emphasis in Fig. 14 on TDHS processing can be mitigated to some degree by undoing some of the pre-emphasis. It cannot restore the fundamental harmonic, however, which is removed by 0- to 200-Hz high-pass filtering. Thus, the perceived reverberance can be reduced in amplitude (along with the amplitude of the high-frequency content of the speech) but the masking effects caused by the fundamental harmonic cannot be restored. The deemphasis filter used in the DSP was a simple first-order filter with the difference equation

$$y(n) = 0.65y(n - 1) + 0.667x(n) + 0.25x(n - 1). \quad (8)$$

This filter can be inserted or removed in the DSP realization by controlling one of the input flags to the DSP with a switch.

A final observation that we have made is that the TDHS/SBC coder appears to perform slightly better with electret microphone speech than with carbon-button microphone speech. We speculate that this is due to the increased harmonic distortion of the carbon-button microphone over that of the electret.

X. CONCLUSIONS

In this paper we have discussed the overall aspects of the design of the 9.6-kb/s TDHS/SBC speech coder. An attractive feature of the design is that it can be broken down into a set of five highly modularized parts, each of which is implemented in a single DSP and effectively fully utilizes the capabilities of the DSP. It therefore leads to a relatively efficient and low-complexity approach to a 9.6-kb/s coder. It offers good "communications" quality that is competitive with that of other higher complexity systems.

The overall processing delay of the coder and decoder is about 80 ms or about 160 ms for a full-duplex system. This amount of delay is fairly typical compared with other competitive techniques at this bit rate.

REFERENCES

1. J. L. Flanagan, M. R. Schroeder, B. S. Atal, R. E. Crochiere, N. S. Jayant, and J. M. Tribolet, "Speech Coding," *IEEE Trans. on Commun.*, COM-27, No. 4 (April 1979), pp. 710-37.
2. D. Malah, R. E. Crochiere, and R. V. Cox, "Performance of Transform and Sub-Band Coding Systems Combined with Harmonic Scaling of Speech," *IEEE Trans., Acoustics, Speech, and Signal Processing*, ASSP-20, No. 2 (April 1981), pp. 273-83.
3. R. V. Cox and R. E. Crochiere, "Real-Time Simulation of Adaptive Transform Coding," *IEEE Trans. Acoustics, Speech, and Signal Processing*, ASSP-29, No. 2 (April 1981), pp. 147-54.
4. J. J. Wolf and K. D. Field, "Real-Time Speech Coder Implementation at 9.6 and 16 KB/S," *Proc. 1981 IEEE Int. Conf. Acoust., Speech, and Signal Processing* (April 1981), pp. 607-10.
5. J. R. Boddie et al., "Digital Signal Processor: Architecture and Performance," *B.S.T.J.*, 60, No. 7 (September 1981), pp. 1449-62.
6. Special Issue on the DSP, *B.S.T.J.*, 60, No. 7 (September 1981).
7. R. V. Cox and R. E. Crochiere, "A Single Chip Speech Periodicity Detector," *Proc. IEEE 1982 Int. Conf. on Acoustics, Speech, and Signal Processing* (May 1982), pp. 525-8.
8. R. V. Cox, R. E. Crochiere, and J. D. Johnston, "An Implementation of Time Domain Harmonic Scaling with Application to Speech Coding," *Proc. IEEE 1982 Int. Conf. on Communications* (June 1982), pp. 4G.1.1-4.
9. R. E. Crochiere, "Digital Signal Processor: Sub-Band Coding," *B.S.T.J.*, 60, No. 7 (September 1981), pp. 1633-53. See also: R. E. Crochiere, M. Randolph, J. W. Upton, and J. D. Johnston, "Real-Time Implementation of Sub-Band Coding on a Programmable Integrated Circuit," *Proc. IEEE 1981 Int. Conf. on Acoustics, Speech, and Signal Processing*, (March 1981), pp. 455-8.
10. D. Malah, "Time-Domain Algorithms for Harmonic Bandwidth Reduction and Time Scaling of Speech Signals," *IEEE Trans. Acoust., Speech, and Signal Processing* ASSP-27 (April 1979), pp. 121-33.

11. R. E. Crochiere, S. A. Weber, and J. L. Flanagan, "Digital Coding of Speech in Sub-Bands," *B.S.T.J.*, 55, No. 8 (October 1976), pp. 1069-85.
12. D. G. Marsh, B. K. Ahuja, T. Misawa, M. R. Dwarakanath, P. E. Fleischer, and V. R. Saari, "A Single-Chip CMOS PCM Coder with Filters," *IEEE J. of Solid State Circuits*, SC-16, No. 4 (August 1981), pp. 308-15.
13. L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Englewood Cliffs, NJ: Prentice-Hall, 1978.
14. R. E. Crochiere, R. V. Cox, J. D. Johnston, "Real-Time Speech Coding," *IEEE Trans. on Commun.*, COM-30, No. 4 (April 1982), pp 621-34.
15. D. Esteban and C. Galand, "Application of Quadrature Mirror Filters to Split Band Voice Coding Schemes," *Proc. 1977 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing* (May 1977), pp. 191-5.
16. M. G. Bellanger, G. Bonnerot, and M. Coudreuse, "Digital Filtering by Polyphase Network: Application to Sample Rate Alteration of Filter Banks," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-24, No. 2 (April 1976), pp. 109-114.
17. J. D. Johnston, "A Filter Family Designed for Use in Quadrature Mirror Filter Banks," *Proc. IEEE Int. Conf. ASSP* (April 1980), pp. 291-4.
18. J. R. Boddie, J. D. Johnston, C. A. McGonegal, J. W. Upton, R. E. Crochiere, and J. L. Flanagan, "Digital Signal Processor: "Adaptive Differential Pulse-Code-Modulation Coding," *B.S.T.J.*, 60, No. 7, Part 2 (September 1981), pp. 1547-61.
19. W. R. Daumer, "Subjective Evaluation of Several Efficient Speech Coders," *IEEE Trans. on Commun.*, COM-30, No. 4 (April 1982), pp. 655-62.

An Improved Model for Isolated Word Recognition

By J. M. TRIBOLET,* L. R. RABINER, and J. G. WILPON

(Manuscript received April 19, 1982)

Current models for isolated word recognition perform very well on small vocabularies of distinctly different sounding words. However, when we are confronted with vocabularies of similar sounding words (e.g., the letters of the alphabet), the performance of isolated word recognizers decreases dramatically. By carefully reexamining the model used for isolated word recognition we have identified some of the inherent deficiencies. In this paper we propose an improved word-recognition model that is inherently capable of accurately recognizing words from almost any vocabulary. We have investigated a simple implementation of the model that preserves most of the structure of a linear predictive coding (LPC)-based version of the canonic isolated word model. In an experimental evaluation of the improved model, using an alpha-digit vocabulary, recognition accuracy improvements of from 1 to 5.7 percent were obtained for four talkers. The improvements were due to changes in both the analysis model and the decision procedure. The strengths and weaknesses of the improved model are discussed.

I. INTRODUCTION

Although the goal of continuous speech recognition by machine remains far out of reach, the one area of speech recognition that is practical with today's technology and understanding is that of isolated word recognition.¹⁻⁶ What is interesting about this area is that the general approach used to solve the isolated word-recognition problem (i.e., the statistical-pattern-recognition approach) bears little relationship to the way in which humans understand speech. As a result the vocabularies for which the isolated word recognizers can achieve good

* Work performed while a consultant to Bell Laboratories.

performance are severely constrained in both size and complexity.⁷ If we are interested in using a vocabulary for which the performance of the isolated word recognizer is less than perfect (e.g., the letters of the alphabet), then we have to rely on the syntax and semantics of the recognition task to provide the desired level of performance from the overall system.⁸⁻¹⁰

In an effort to improve word-recognition accuracy for arbitrary vocabularies, we have re-examined the word-recognition model and proposed a somewhat more general structure. The proposed changes in the model include an improved feature analysis in which both long-time and short-time features are measured, and an improved decision box in which the two-pass decision rule of Rabiner and Wilpon¹¹ is adapted to the speaker-trained case.

The implementation of the improved word-recognition model, which we have studied, is based on the standard linear predictive coding (LPC) word recognizer as originally proposed by Itakura.¹² In an effort to retain as much of the original structure as possible, we have used the standard LPC analysis as the long-time features, and a new LPC analysis based on 15-ms frames as the short-time features. Experimentation with the improved model, using a 39-word vocabulary of the alphabet, the digits, and three command words in a speaker-trained mode, showed improvements in accuracy of from 1 to 5.7 percent for four talkers. An analysis of the results showed that the improved feature analysis provided only small improvements in accuracy (from 0 to 1.3 percent), whereas the two-pass decision rule provided somewhat larger improvements in accuracy (from 1 to 4.4 percent).

The outline of this paper is as follows. In Section II we briefly review the canonic isolated word-recognition model and discuss its strengths and weaknesses. We also discuss, in this section, the implementation of the model based on LPC feature analysis and an LPC distance measure. In Section III we present the improved word-recognition model and discuss how it was implemented within the structure of the LPC-based recognizer. In Section IV we describe the experimental evaluation of the improved model based on the alpha-digit vocabulary. Finally, in Section V we discuss the results and their implications for practical systems.

II. THE CANONIC MODEL FOR ISOLATED WORD RECOGNITION

Figure 1 is a block diagram of the canonic (statistical-pattern-recognition) model for isolated word recognition. The three basic components of the model include:

(i) Feature measurement in which the speech signal is analyzed to provide a set of Q features (e.g., filter bank energies, LPC coefficients, etc.) once every M samples. If the isolated word is of duration $L \times M$

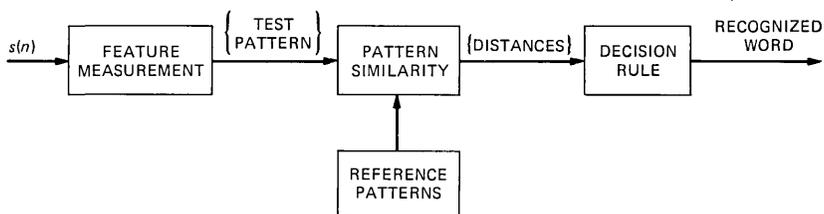


Fig. 1—Block diagram of standard isolated word-recognition model.

samples, then a total of L sets of features characterize the word. The matrix of $Q \times L$ features is called the test pattern.

(ii) Pattern similarity measurement in which a score (similarity or distance) relating the similarity of the test pattern to each of a set of V reference patterns is computed. Pattern similarity involves both time alignment (registration) of the test and reference pattern, and distance computation along the alignment path. The output of the pattern similarity box is a set of V distance scores, i.e., one for each reference pattern.

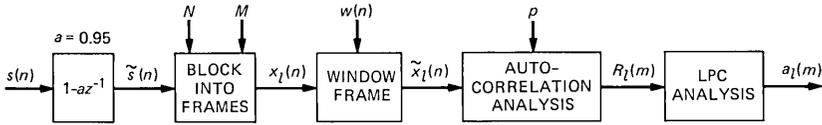
(iii) A decision rule in which the distance scores are used to provide an ordered (by distance) list of recognition candidates. Generally, the candidate with the smallest distance is chosen as the “recognized word.”

Rather than dwelling further on the canonic model we now review the LPC implementation of this model, as we will be relying on this structure throughout this paper. We will return to the canonic model in Section 2.2 when we discuss its limitations and propose the improved model.

2.1 The LPC-based implementation of the word recognizer

Figure 2 is a block diagram of the feature measurement for an LPC-based analyzer. The digitized speech signal (digitized at a 6.67-kHz rate) is first preemphasized using a simple first-order digital network and then blocked into overlapping frames of N (300) samples with consecutive frames overlapping by 200 samples. Thus, a frame spacing of $M = 100$ samples is used (i.e., 67 frames/second). Each speech frame is then windowed by a 300-sample Hamming window, and a p th-order ($p = 8$) autocorrelation analysis is performed. A full LPC analysis (using the autocorrelation method¹³) is then performed giving the set of $(p + 1)$ LPC coefficients as the features for each frame.

The pattern similarity processing is carried out using a dynamic time-warping (DTW) algorithm in which the test pattern is simultaneously time aligned with each reference pattern, and a distance along the time-alignment path is computed. One of the major features of this



$$\begin{aligned} \tilde{s}(n) &= s(n) - as(n-1) \\ x_l(n) &= \tilde{s}(Ml + n), \quad l = 0, 1, \dots, L-1 \\ &\quad n = 0, 1, \dots, N-1 \\ \tilde{x}_l(n) &= x_l(n) \cdot w(n) \\ R_l(m) &= \sum_{n=0}^{N-1-m} \tilde{x}_l(n) \tilde{x}_l(n+m) \end{aligned}$$

Fig. 2—Block diagram of LPC analysis system.

processing is the local distance measure, which relates the distance between a frame of the test pattern and a frame of the reference pattern, of the form¹²

$$d(T, R) = \log \left[\frac{\mathbf{a}_R \mathbf{V}_T \mathbf{a}_R^t}{\mathbf{a}_T \mathbf{V}_T \mathbf{a}_T^t} \right], \quad (1)$$

where \mathbf{a}_R and \mathbf{a}_T are the LPC feature sets of reference and test, respectively, and \mathbf{V}_T is the autocorrelation coefficient set of the test. The distance measure of eq. (1), called the LPC log-likelihood measure, can be computed using only $(p + 1)$ multiplications and additions, and one logarithm.¹² Furthermore, the LPC distance of eq. (1) has been shown to have reliable and well understood statistical properties.^{14,15} In particular, if both \mathbf{a}_R and \mathbf{a}_T are derived from the same underlying stationary random process, then $d(T, R)$ is precisely χ^2 distributed with p degrees of freedom. This statistical behavior of the LPC distance holds for fricative sounds. For voiced speech, although the model is inexact on a frame-by-frame basis, the statistical properties are approximately correct on a time-average basis.

To compute the pattern similarity between the test and each reference pattern using the DTW algorithm with the distance measure of eq. (1), a solution to the minimization of

$$D^* = \min_{w(n)} \left[\sum_{n=1}^{NT} d(T_n, R_{w(n)}) \right] \quad (2)$$

must be found where NT is the number of frames in the test, and $w(n)$ is the warping function relating frame n of the test to frame $w(n)$ of the reference. Efficient recursive procedures for solving eq. (2) have been described in the literature.^{12,16-18}

Finally, the decision box orders the distance scores for each reference pattern and chooses either the reference with the minimum distance

(the nearest neighbor rule) or the reference whose average of the K -best scores (for multiple-template systems) is minimum (the K -nearest neighbor rule) as the recognized word. When the recognized word is unique (i.e., only a single reference gets a small distance score), this simple decision rule is sufficient. However, for complex vocabularies, generally several references achieve small distance scores, and reliable recognition using the smallest distance cannot be achieved. In such cases a two-pass decision rule¹¹ has been shown to increase accuracy by deferring the final recognition decision to a discrimination analysis in a second pass of the decision rule. This discrimination analysis has only been applied to speaker-independent systems because of the problems associated with obtaining appropriate word discrimination weights.¹¹

2.2 Strengths and limitations of the word-recognition model

The strengths of the canonic word-recognition model of Fig. 1 are as follows:

(i) It is invariant to different speech vocabularies, users, feature sets, pattern similarity algorithms, and decision rules.

(ii) It is easy to implement.

(iii) It works well in practice.

The weaknesses of the model include:

(i) The feature analysis only adequately represents long-time stationary events in the speech signal; nonstationary and transient events are only poorly represented.

(ii) The model does not perform well for complex vocabularies with acoustically similar words.

We now consider the first weakness of the model. By way of example Fig. 3 shows waveform plots of the beginning regions of two distinct words. Word 1 shows a silence followed by the onset of voiced speech. Word 2 shows a short (15 ms) transient of low-level, unvoiced speech (e.g., a plosive sound) followed by the onset of voiced speech. Figure 3 also shows the placement of the first two long-time speech segments (frames), which contain identical data except for the first 15 ms of the first segment, in which one frame has silence and one frame has a short plosive. It should be clear that for a long-time analysis such as the LPC model of Section 2.1, the low-level differences in the first 15 ms of frame 1 will be swamped out by the high-level voiced speech in the last 30 ms of the frame. Thus, in a long-time stationary framework accurate recognition of differences between short transients and other nonstationary regions (e.g., as occur during onsets and offsets of voicing) is greatly limited. Thus, to ameliorate this weakness, the feature-detection algorithm must be enhanced to include some representation of short-time nonstationary events.

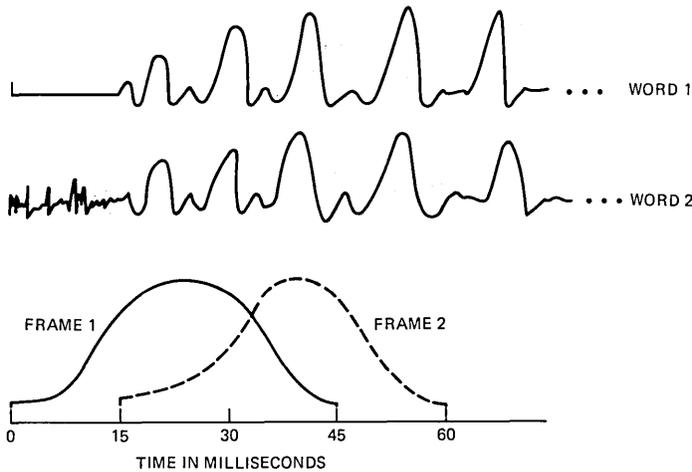


Fig. 3—How a short transient in a word can be swamped out by a voiced region in the long-time analysis model.

Consider now the second weakness of the model. The reason that acoustically similar words are easily confused is that the pattern-similarity measure (the DTW distance) gives equal weight to all frames of the word. For differentiating words of one equivalence class from words of another equivalence class this procedure is reasonable. However, within a class of acoustically similar words a discrimination analysis rather than a straight recognition is required. Such an analysis has been proposed by Rabiner and Wilpon¹¹ for the case of speaker-independent recognition of words.

For speaker-trained recognizers this two-pass decision rule must be modified so that the optimal weighting curves for word discrimination could be obtained directly from the robust training procedure.¹⁹

With the incorporation of the expanded feature analysis, a modified DTW algorithm, and an expanded decision rule, the basic weaknesses of the canonic word recognizer can be overcome to some extent. In the next section we describe an “improved” model for word recognition and show how the improvements can be incorporated directly into the LPC framework of Section 2.1.

III. THE IMPROVED WORD-RECOGNITION MODEL

Based on the discussion of Section 2.2, the improved word-recognition model would have a structure of the type shown in Fig. 4. The major differences in the model, from that of Fig. 1, are:

(i) The feature measurement box is expanded into three sub-blocks, namely long-time feature measurements, short-time feature

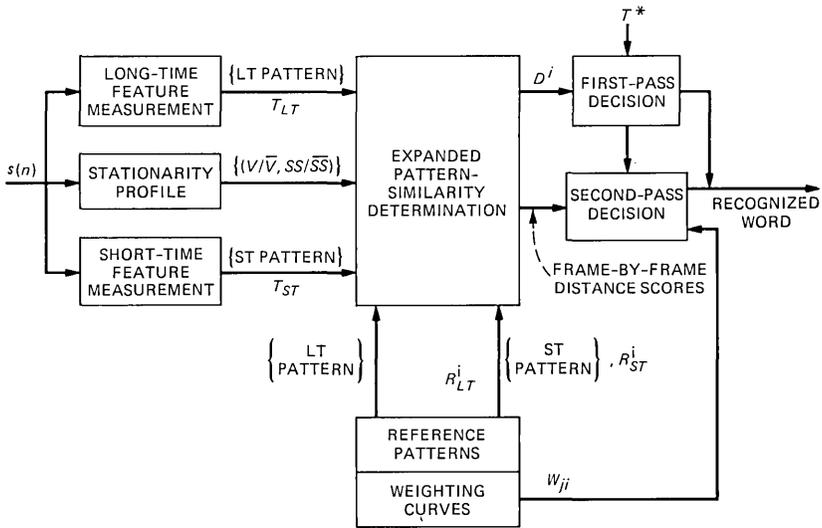


Fig. 4—Block diagram of the improved, isolated word-recognition model using both long-time and short-time features, and a two-pass discrimination model.

measurements, and a stationarity profile. The long-time features are essentially those of the original model, although the rate at which they are measured will generally be higher for this new model than for the original model. The short-time features are intended to characterize transients and other nonstationary events in the speech signal. Some typical short-time features include zero or level crossing counts over short-time intervals, wideband (short-impulse response), filter bank analyses, short-time LPC analyses, etc. The stationarity profile decides which feature set (either long-time or short-time) is used to characterize a given frame of speech, and hence is used for the distance measure of the pattern-similarity algorithm.

(ii) The dtw algorithm is expanded to use both long-time and short-time patterns, for both test and reference patterns, in determining similarity of a given reference pattern to the test pattern. The stationarity profile is used to guide the alignment and to choose which feature set is used in making a given distance computation.

(iii) The decision box is implemented as a two-pass decision. In the first-pass decision the distance scores for each reference pattern are ordered, and if the best distance is smaller than the second best distance by a threshold T^* , the decision phase is terminated. If, however, the top two or more references are within T^* in distance, a second-pass decision rule is used in which the similar words are compared using a discriminant analysis and the recognized word is chosen on the basis of this analysis. To implement the discriminant

analysis, a set of distance-weighting curves discriminating word i from word j (for all i, j) must be saved along with the reference patterns. We now describe how the improved model was implemented in the framework of the LPC analysis system.

3.1 The LPC basic improved word-recognition model

Using the LPC analysis framework, the expanded feature measurement was implemented as follows. The long-time analysis was implemented as described in Section 2.1 except that the shift parameter, M , was changed from $M = 100$ to $M = 33$, and the analysis frame length, N , was changed from $N = 300$ to $N = 297$. Thus, for the long-time analysis, analysis frames were computed every 5 ms rather than every 15 ms, thereby giving a frame rate three times larger. The analysis frame was changed to 297 samples so as to be an integral multiple of M , the shift parameter. We denote the long-time LPC features as T_{LT} .

For convenience the short-time analysis was implemented with the same processing (i.e., that of Fig. 2) as that of the long-time analysis, except that N was changed to 99 (15-ms analysis frames) and M was again set to 33 (5-ms frame shifts). The order of the LPC analysis was kept at 8 for the short-time as well as the long-time analysis. We denote the short-time LPC features as T_{ST} .

To understand how the stationarity profile, μ , is generated within the framework of the LPC analysis, we must first define a characterization of the types of speech segments that are encountered. For this purpose we define two binary features that characterize the source and the dynamics of the vocal tract. The first feature describes the excitation for the frame of speech and we denote voiced speech as V , and unvoiced speech as \bar{V} . The second feature describes the vocal tract dynamics and we denote the stationary, steady-state case as SS , and the nonstationary, time-varying case as \bar{SS} . Thus, a given frame of speech is characterized by the notation $(V/\bar{V}, SS/\bar{SS})$.

The determination of whether a frame is voiced or unvoiced is fairly straightforward and is readily obtained from any number of pitch-detection algorithms. The determination of whether a frame is stationary or nonstationary is somewhat more complicated. This computation is made as follows. The basic idea is to compare both the long-time and short-time features of frames j and i , where j represents the frame occurring 15 ms before frame i . A distance comparing frames i and j is made as

$$\alpha_i = \frac{d[T_{LT}(i), T_{LT}(j)] + d[T_{LT}(j), T_{LT}(i)] + d[T_{ST}(i), T_{ST}(j)] + d[T_{ST}(j), T_{ST}(i)]}{4}, \quad (3)$$

i.e., the average of the long- and short-time LPC distances between

Table I—Feature sets used for similarity determination

Test Frame Status	Feature Set	Frame Spacing	Speech Example
(V, SS)	LT analysis	15 ms	Vowels, steady-state sounds
(\overline{V}, SS)	LT analysis	5 ms	Onset, offset of voicing transitions
($\overline{V}, \overline{SS}$)	LT analysis	15 ms	Steady fricatives
(\overline{V}, SS)	ST analysis	5 ms	Transients

frames i and j and between frames j and i (recall that the LPC distance is not symmetric). The distance score, α_i , is then compared with a threshold (different for voiced and unvoiced frames), and the stationarity value is given as

$$SS = \begin{cases} 1 & \text{if } \overline{V} \text{ and } \alpha_i \leq \text{THV} \\ 1 & \text{if } \overline{V} \text{ and } \alpha_i \leq \text{THU} \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where 1 represents a stationary frame, and 0 represents a nonstationary frame, and THV and THU are voiced and unvoiced thresholds, respectively.

Once a frame has been characterized with the two-feature code, ($V/\overline{V}, SS/\overline{SS}$), the only remaining step is to specify which feature set and frame spacing should be used in the DTW distance computation.

It should be clear that for voiced frames, ($V, -$), the long-time analysis should be used to avoid potential bias caused by the pitch period. Similarly, for all nonstationary frames, ($-, \overline{SS}$), a frame spacing of 5 ms should be used to track the fast dynamics of such frames. Finally, for unvoiced, nonstationary frames, ($\overline{V}, \overline{SS}$), the short-time analysis is most appropriate to follow transients and other brief events.

Table I shows a summary of the feature sets and frame spacings, for each of the four types of frames, as used to determine word and reference template similarity.

To illustrate the above analysis, Fig. 5 shows a series of plots of (a) the waveform, (b) the log energy (in dB), (c) the pitch, and (d) the average of long- and short-time LPC distance [eq. (3)] for the word /B/. It can be seen in Fig. 5a that the LPC distance becomes large at the beginning of voicing (point A in the plot), at the termination of voicing (point B in the plot), and at the end of the word (point C in the plot). Such frames (and their neighborhoods) are the nonstationary regions of the word, and generally correspond well with transients, onset and offset of voicing, and rapidly varying vocal-tract dynamics.

To determine the stationarity thresholds intelligently, THV and THU, histograms of the behavior of α_i for voiced and unvoiced frames, had to be measured. Such histograms are shown in Fig. 6. The data in this figure were obtained by computing α_i every 5 ms for all the frames of a 39-word vocabulary of letters of the alphabet plus the digits. Based

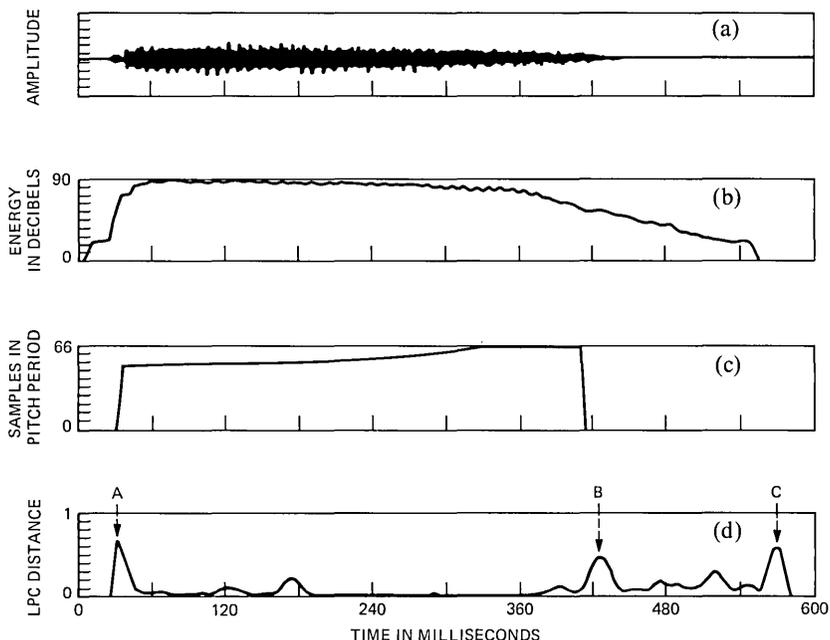


Fig. 5—An example (the word *B*) showing: (a) the waveform, (b) its energy profile, (c) its pitch contour, and (d) the LPC distance comparing adjacent frames.

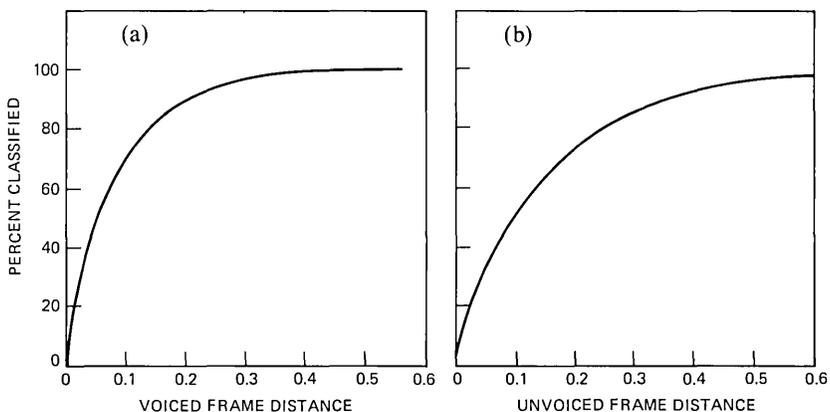


Fig. 6—Histograms of values of (a) LPC distance for voiced speech, and (b) unvoiced speech. Thresholds THV and THU are chosen to give desired percentages of nonstationary classification.

on the data of Fig. 6, values for THV and THU can be chosen, so as to obtain any desired average probabilities of occurrence of voiced or unvoiced classification. For example, if we assume that, on average, only 10 percent of the voiced frames should be classified as *SS*, then

a threshold of $\text{THV} = 0.2$ should be used. Similarly, for non-voiced frames a threshold of $\text{THU} = 0.3$ yields an average of 10 percent of the frames being classified as nonstationary. If the thresholds, THV and THU , are both set to infinity, then all frames are classified as stationary and hence the feature analysis is essentially identical to that of the original model. Similarly, if the thresholds are both set to zero, all frames are classified as nonstationary and a 5-ms frame spacing is used with both short- and long-time feature sets.

3.2 Modifications to the DTW algorithm for the improved word model

As discussed above, the basic changes made in the feature measurement were inclusion of both short- and long-time LPC analyses, and an increase in the frame rate of the analysis from once every 15 ms to once every 5 ms. These analysis changes required some modifications to the DTW algorithm to properly handle the raw data structure. The modifications primarily involve reformulation of the local path constraints to account for the different possible frame spacings (i.e., nonuniform sampling in time), and modifications to the distance computation to handle both long- and short-time LPC distances and their appropriate weights.

We denote the long-time test pattern as $\{T_{LT}(n), n = 1, 2, \dots, NT\}$, the short-time test pattern as $\{T_{ST}(n), n = 1, 2, \dots, NT\}$, and the stationarity distance (on which the stationarity profile is based) as $\{\alpha_n, n = 1, 2, \dots, NT\}$. Similarly, we denote the long-time reference pattern as $\{R_{LT}(m), m = 1, 2, \dots, NR\}$ and the short-time reference pattern as $\{R_{ST}(m), m = 1, 2, \dots, NR\}$.

We wish to solve for the optimum warping path of the form $m = w(n)$, defined for values of n that satisfy either of the following conditions:

$$(n - 1) \oplus 3 = 0 \quad (5a)$$

or

$$\alpha_n > \text{TH} \quad \text{or} \quad \alpha_{n-1} > \text{TH} \quad \text{or} \quad \alpha_{n-2} > \text{TH}. \quad (5b)$$

Equation (5a) says we solve for $m = w(n)$ at each standard 15-ms time slot. This constraint essentially guarantees a grid spacing, between adjacent DTW frames, of no more than three frames. It also guarantees that, in the limit, as the entire word is classified as stationary, the new analysis becomes identical to the previous analysis. Equation (5b) says we solve for $m = w(n)$ at each frame, n , in which the stationarity distance, α_n , of that frame *or* either of its two predecessors falls below the specified threshold, TH. (For voiced frames the threshold TH is set to THV, and for nonvoiced frames the threshold TH is set to THU). Cases in which eq. (5b) is satisfied (i.e., one of the distances is

above threshold) correspond to voiced frames with a rapidly changing spectrum (transitions), or unvoiced frames with nonstationary excitation.

For each frame n that satisfies one of the constraints of eq. (5) we must solve the DTW recursion

$$D_A(n, m) = w(n - n_L)\hat{d}(n, m) + \min_{\tilde{m}_L \leq m_0 \leq \tilde{m}_H} [D_a(n - n_L, m_0)], \quad m_L \leq m \leq m_H, \quad (6)$$

where

n_L = last value of n for which a DTW recursion was done.

\hat{n}_L = next-to-last value of n for which a DTW recursion was done.

$w(n - n_L)$ = weighting function on the local distance to account for the nonuniform frame spacing.

$\hat{d}(n, m)$ = local frame distance for reference frame m and test frame n .

\tilde{m}_L = smallest value of m at $n = n_L$ from which a valid path can go to the grid point (n, m) .

\tilde{m}_H = largest value of m at $n = n_L$ from which a valid path can go to the grid point (n, m) .

m_L = smallest value of m at frame n for which DTW recursion is solved.

m_H = largest value of m at frame n for which DTW recursion is solved.

The values of m_L and m_H are determined from the global path constraints which specify that all valid DTW paths must lie within a parallelogram defined from lines of slope 2 and slope 1/2 beginning at grid point $(0, 0)$ and ending at grid point (NT, NR) . Thus, m_L and m_H satisfy the path constraints

$$m_L = \max[(n - 1)/2 + 1, 2 \times (n - NT) + NR, 1] \quad (7a)$$

$$m_H = \min[2 \times (n - 1) + 1, (n - NT)/2 + NR, NR]. \quad (7b)$$

The values of \tilde{m}_L and \tilde{m}_H are those which guarantee that the path to grid point (n, m) satisfies the local constraint that the average slope be no less than one half nor more than 2. If we define a path increment function, $\Delta(m)$, as

$\Delta(m)$ = increment in m along the best path to grid point (n_L, m) ,

i.e., if the best path to grid point (n_L, m) comes from grid point $[\hat{n}_L, m - \Delta(m)]$, then values of m_0 in the DTW recursion [eq. (6)] must satisfy the local path constraint

$$\frac{(n - \hat{n}_L)}{2} \leq \Delta(m_0) + (m - m_0) \leq 2(n - \hat{n}_L). \quad (8)$$

Since $\Delta(m_0)$ also satisfies the constraint

$$\Delta(m_0) \leq 2(n_L - \hat{n}_L), \quad (9)$$

we can rewrite the inequalities of eq. (8) as

$$\tilde{m}_H - \Delta(\tilde{m}_H) \leq \frac{m - (n - \hat{n}_L)}{2} \quad (10a)$$

$$\tilde{m}_L \geq m - 2(n - n_L). \quad (10b)$$

Equation (10a) must be checked for each possible m value to find its solution, whereas eq. (10b) can be used directly.

The weighing function $w(n - n_L)$ is simply

$$w(n - n_L) = (n - n_L) \quad (11)$$

to give more weight to longer frame separations, and the distance $\hat{d}(n, m)$ of the form

$$\hat{d}(n, m) = \begin{cases} d[T_{LT}(n), R_{LT}(m)] & \text{if } (V, SS), (\bar{V}, SS) \\ & \text{or } (V, \bar{SS}) \\ d[T_{ST}(n), R_{ST}(m)] & \text{if } (\bar{V}, \bar{SS}). \end{cases} \quad (12a)$$

$$(12b)$$

The complicated form of the DTW recursion is due to the nonuniform sampling rate at which the recursion is solved. If we translate eqs. (6) through (12) into words we can say that for each frame n for which the recursion is solved we compute $D_A(n, m)$ for a range of m from $m = m_L$ to $m = m_H$, as determined by the global path constraints. For each m the optimal path is determined as the weighted local distance, $\hat{d}(n, m)w(n - n_L)$, (as determined by the stationarity profile at frame n) plus the best accumulated distance to a predecessor frame that is a valid candidate for a path to frame m (i.e., $\tilde{m}_L \leq m_0 \leq \tilde{m}_H$). The range on m_0 is chosen to guarantee that the local path constraints of a warping curve slope of between 1/2 and 2 are met. Since the number of frames between the current frame n and the predecessor frame n_L , for which the DTW recursion was last solved, is variable (ranging from 1 to 3), the local path constraints must use this range, along with information as to how much the local path rose $[\Delta(m)]$ at frame (n_L, m_0) to set the local path constraints correctly.

The DTW recursion of eq. (6) is solved for all valid points from $n = 1$ to $n = NT$, and the total DTW solution is then given as

$$D^* = D_A(NT, NR) \quad (13)$$

and the average path distance is

$$\bar{D} = \frac{D_A(NT, NR)}{NT}. \quad (14)$$

3.3 The improved decision rule

As we discussed earlier a two-pass decision rule is used to improve recognition accuracy. The task of the first-pass decision rule is to determine the set of vocabulary words that are acoustically similar to the test word (i.e., the set of confusions). The task of the second-pass decision rule is then to resolve these confusions.

The key idea behind the operation of the second-pass decision rule is that the DTW distance scores between the test pattern and those reference patterns that are acoustically close to each other and to the test pattern consist of a χ^2 random component and a Gaussian random component. The χ^2 random component is associated with the averaging of distance scores between frames with the same basic spectrum, and therefore has a χ^2 distribution with p degrees of freedom. The Gaussian random difference is associated with the averaging of large distance scores between frames with dissimilar spectra.

In cases where the size of the dissimilar region is small (such as in comparing a $/B/$ to a $/D/$) compared to the size of the similar region, the χ^2 component distance often outweighs the Gaussian component, thereby causing potential recognition errors.

The purpose of the second-pass decision rule is to enhance the role of the Gaussian component associated with spectrally dissimilar regions in determining the final decision. This is accomplished using a distance-weighting function that enhances the discrimination power of the frame-by-frame distance scores.

By way of example, consider a simple confusion list of two references, R_i and R_j , for test word T . Let the DTW frame-by-frame distance and warping path be specified as

$$d_k(n) = d\{T(n), R_k[w(n)]\} \quad (15)$$

and

$w_k(n)$ = Warping path comparing frame n of the test with reference R_k .

We now define two distance-weighting functions,

$$\begin{aligned} &\{W^{i,j}(n), n = 1, 2, \dots, NR_i\} \\ &\{W^{j,i}(n), n = 1, 2, \dots, NR_j\}, \end{aligned}$$

where $W^{i,j}(n)$ is the weighting to discriminate R_j from R_i , and $W^{j,i}(n)$ is the weighting to discriminate R_i and R_j . (Reference 11 shows that

these weighting functions are generally not symmetric). We defer a discussion of how the weights are generated, in a speaker-trained system, to Section 3.4.

The basic hypothesis is that the test pattern, T , corresponds to either R_i or R_j , and we wish to come up with a discrimination score that aids in this decision. If we define a discrimination score, $\delta(T, R_i | T \in R_j)$, as the weighted distance between the T and R_i , assuming that T actually corresponds to R_j , then we get

$$\delta(T, R_i | T \in R_j) = \frac{\sum_{k=1}^{NT} W^{i,j}[w_i(k)]d\{T(k), R_i[w(k)]\}}{\sum_{k=1}^{NT} W^{i,j}[w_i(k)]}, \quad (16)$$

and similarly we get

$$\delta(T, R_j | T \in R_i) = \frac{\sum_{k=1}^{NT} W^{j,i}[w_j(k)]d\{T(k), R_j[w(k)]\}}{\sum_{k=1}^{NT} W^{j,i}[w_j(k)]}. \quad (17)$$

The weighted distance corresponding to the hypothesis $T \in R_j$ [i.e., eq. (16)] is shown in Fig. 7. The frame-by-frame distance is multiplied by the weighting function reflected through the warping curve to give the discrimination score δ .

The discrimination distances of eqs. (16) and (17) have the following important property. If T and R_i are from the same word (different replications) then the frame-by-frame distances, $d(\cdot, \cdot)$ are all χ^2 distributed (theoretically) and thus $\delta(T, R_i | T \in R_j)$ will be theoretically "independent" of the weighting function. If, however, T and R_j (instead of R_i) are from the same word, then $\delta(T, R_i | T \in R_j)$ will reflect to a greater extent the Gaussian-distributed component of the original distance score, $d(T, R_i)$, since it primarily consists of distance in regions where R_j and R_i differ significantly, even though they may be quite short.

Thus, in the simple case of a confusion between two references, R_i and R_j , the final decision is made on the basis of the discrimination scores of eqs. (16) and (17).

More generally, if the confusion list associated with test pattern T has Q candidates, $\{R_{i_1}, R_{i_2}, \dots, R_{i_Q}\}$, then the following procedure is followed:

(i) Compute all pairs of discriminations

$$\delta(T, R_{i_a} | T \in R_{i_b}), \quad b \neq a, \quad a, b = 1, 2, \dots, Q.$$

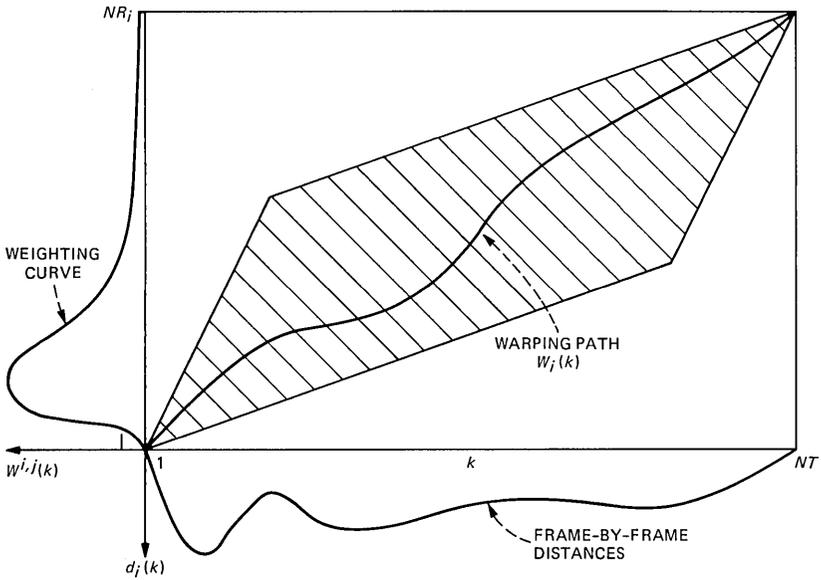


Fig. 7—The time warping plane of a test and reference pattern along with the distance of each frame and the weighting curve on distance.

(ii) Form the average discrimination distance

$$\bar{\delta}(T, R_{i_a}) = \frac{1}{Q-1} \sum_{\substack{b=1 \\ b \neq a}}^Q \delta(T, R_{i_a} | T = R_{i_b}), \quad a = 1, 2, \dots, Q.$$

(iii) Define the most likely candidate, R_{i_a} , as the candidate with the minimum average discrimination distance, i.e.,

$$\delta_{\text{MIN}} = \min_a \{\bar{\delta}(T, R_{i_a})\}.$$

Similarly, a least likely candidate with maximum distance is defined as

$$\delta_{\text{MAX}} = \max_a \{\bar{\delta}(T, R_{i_a})\}$$

(iv) Given the original (i.e., first-pass) distance scores for all Q candidates, $d(T, R_{i_a})$, with smallest distance d_{MIN} and largest distance d_{MAX} , a second-pass set of distances scores is computed by retaining second-pass ordering with first-pass distances. This procedure is illustrated in Fig. 8. A reference with second-pass discrimination score $\bar{\delta}(T, R_i)$ is given distance $\bar{d}(T, R_i)$ by linearly interpolating along the line of Fig. 8.

3.4 Determination of the weighting curves in the speaker-trained case

The determination of the weighting curves, $W^{j,i}$ and $W^{i,j}$, is readily

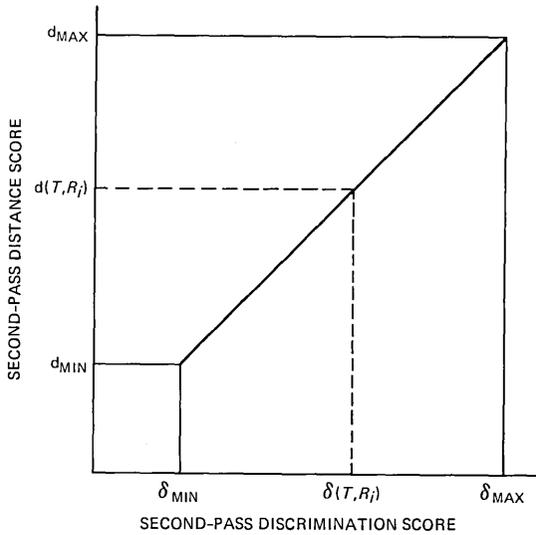


Fig. 8—Linear transformation between second-pass distance score and second-pass discrimination score.

performed in the training phase for speaker-trained systems. Given reference templates R_i and R_j , as obtained using the robust training procedure of Rabiner and Wilpon,¹⁹ a simple way of obtaining $W^{i,j}$ is to warp R_i to R_j , giving

$$W^{i,j}(k) = d\{R_j(n), R_i[w_j(n)]\}, \quad (18)$$

where $w_j(n)$ denotes the warping path. Thus, the frame weights (W) are essentially the frame-by-frame warped DTW distances between the reference templates. Figure 9 shows weighting functions for references corresponding to the words /I/ and /Y/. When compared with the speaker-independent weights of Rabiner and Wilpon,¹¹ we immediately see the statistical effects of small samples. It is evidence that the curves of Fig. 9 need some smoothing to reduce the statistical variance. The resulting of applying a 3-point smoother (a triangular window) to the data of Fig. 9 is given in Fig. 10. A good deal of the statistical variation in the curves is smoothed out.

An alternative, more statistically meaningful, way of obtaining smoother weighting curves is to use all P replications of each word in the training set to determine the weights. Basically, we obtain a weighting function for each pair of training tokens such that each token is close in distance to the appropriate reference. The final weighting curve is then obtained by averaging the individual weighting curves, with appropriate time alignments. We use the term subweights to denote the set of weights obtained by averaging all training tokens,

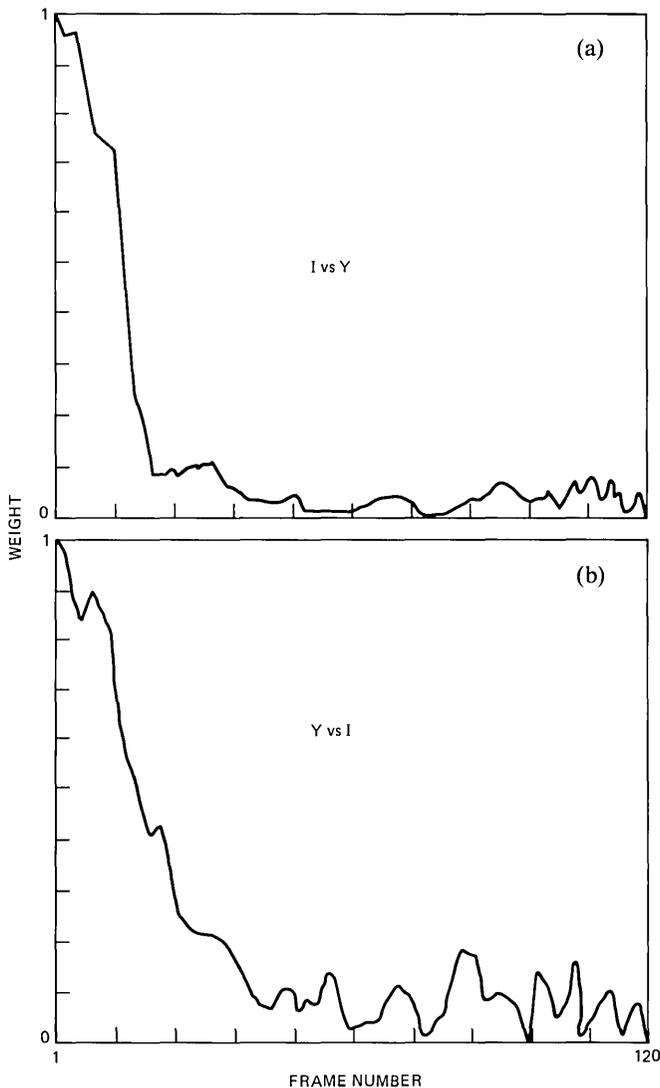


Fig. 9—Typical weighting curves for (a) *I* vs *Y*, and (b) *Y* vs *I*, derived from the robust training tokens.

and we use the notation *S* to refer to this set. Figure 11 illustrates the (sub) weighting curves for *I*, *Y* comparisons based on a set of five training tokens for each word.

IV. EXPERIMENTAL EVALUATION OF THE IMPROVED MODEL

To measure the performance of the improved, LPC-based, isolated word-recognition model, a small evaluation test was performed. Each

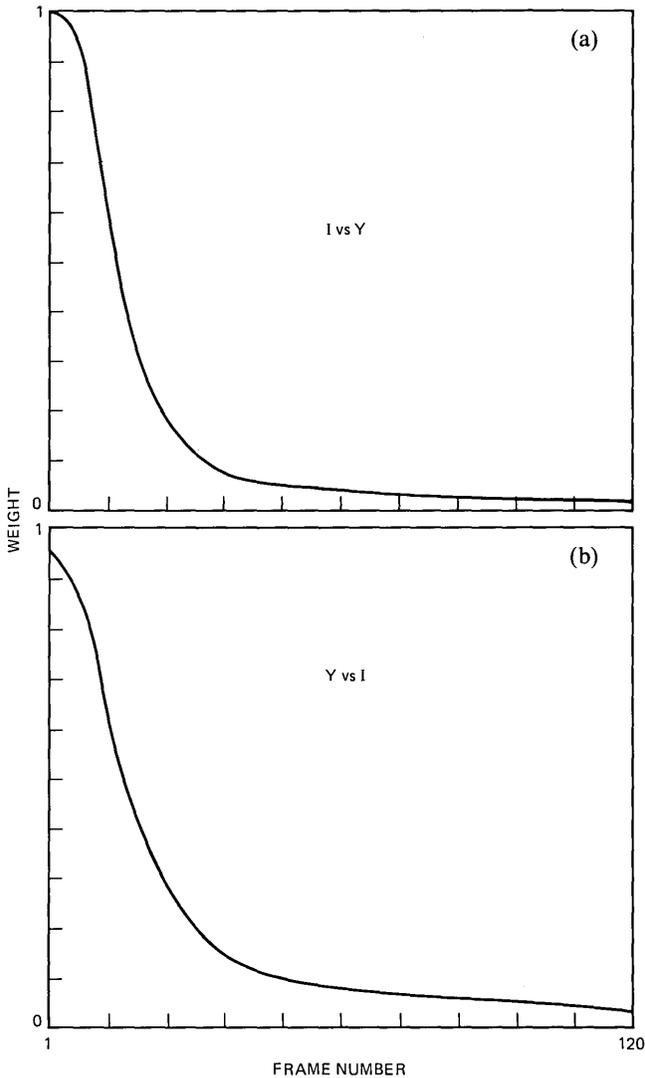


Fig. 10—Smoothed weighting curves for (a) *I* vs *Y*, and (b) *Y* vs *I*, derived from the robust training tokens and a 3-point smoother.

of four talkers (two male, two female—all experienced with speech-recognition systems) trained the recognizer on a 39-word alpha-digit vocabulary by saying each vocabulary word five times during the course of a single training session. The word-reference patterns, the normal discrimination weights, W , and subweights, S , were determined from the training data using the robust training procedure of Rabiner and Wilpon.¹⁹

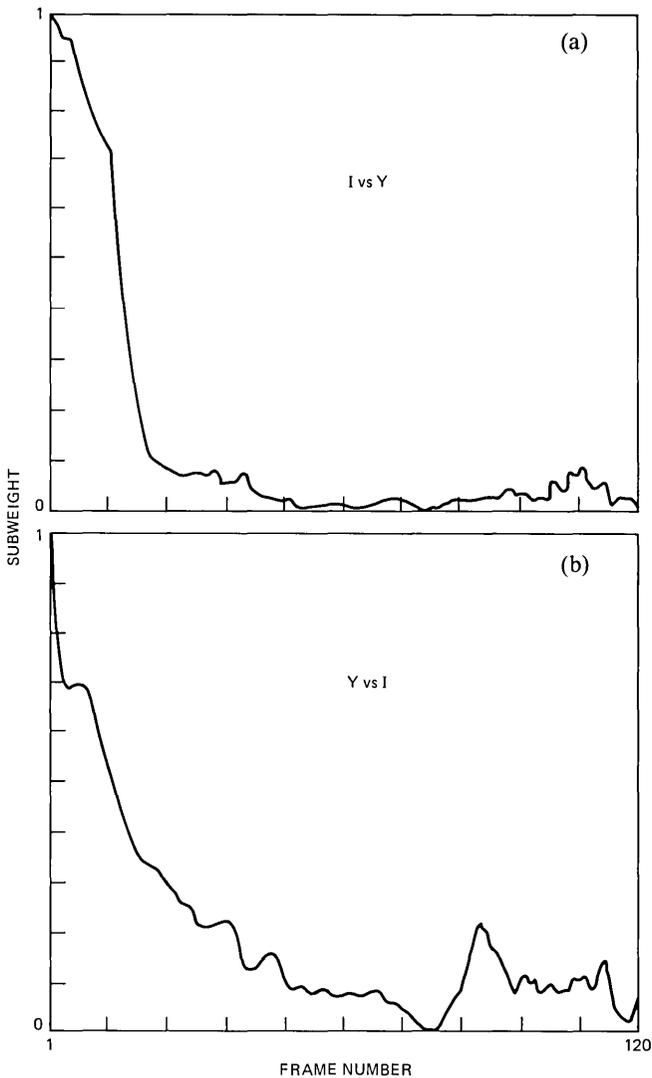


Fig. 11—Subweight curves for (a) *I vs Y*, and (b) *Y vs I*, derived from using all training tokens.

For evaluation purposes the 39-word vocabulary was spoken 10 additional times by each of the four talkers in two distinct recording sessions. Thus, a total of 390 words were used in each recognition test for each talker.

4.1 Recognition test results

The overall results of the evaluation tests are given in Table II,

Table II—Recognition accuracies as a function of the stationarity thresholds and the number of recognition passes for the four talkers

	(THU, THV)	Talker Number			
		1	2	3	4
Pass 1 Alone	$(-\infty, \infty)$	94.9	94.9	90.5	86.7
	$(-.3, .2)$	95.4	94.9	91.8	86.4
	$(0., 0)$	94.9	94.9	91.5	85.4
Pass 2 With Weight W	$(-\infty, \infty)$	96.7	95.6	94.1	87.2
	$(-.3, .2)$	96.4	95.6	94.9	88.5
	$(0., 0)$	95.6	95.4	94.4	86.4
Pass 2 With Subweight S	$(-\infty, \infty)$	95.6	95.9	95.4	87.2
	$(-.3, .2)$	95.4	95.4	96.2	87.9
	$(0., 0)$	95.4	95.6	95.1	87.2

which shows recognition accuracy as a function of stationarity thresholds, talker, and analysis condition. Three analysis conditions are shown, namely Pass 1 alone (no discriminant analysis), Pass 2 with weights, W , derived from single reference tokens, and Pass 2 with subweights, S , derived from all reference tokens.

The results of using Pass 1 alone show only a 0.4-percent improvement, on average, in recognition accuracy for the four talkers when comparing the old stationary model (where $\text{THU} = -\infty$, $\text{THV} = \infty$) with the new stationary model (where $\text{THU} = -0.3$, $\text{THV} = 0.2$).

The results of using Pass 2 with weights W show an average of 2.1-percent improvement in recognition accuracy for the four talkers over the old stationary model (when $\text{THU} = -0.3$ and $\text{THV} = 0.2$). When subweights S are used in Pass 2, the improvement in recognition accuracy is an average of 2 percent.

Table II also shows that when Pass 2 is used the recognition accuracy with stationarity thresholds set to $(-0.3, 0.2)$ is, on average, about 0.5 percent higher than with stationarity thresholds set to $(-\infty, \infty)$. This result indicates that the improved model provides a consistent recognition accuracy improvement of about 0.5 percent, with or without the second-pass weights.

V. DISCUSSION

The results presented in Section IV are both encouraging and discouraging. They are encouraging in that real improvements in recognition accuracy were obtained when a nonstationary analysis framework was used in place of the purely stationary framework used in earlier work. They are discouraging in that the average improvement resulting from the nonstationary model (0.5 percent) was considerably smaller than the average improvement resulting from the discrimination analysis of the second pass (1.6 percent).

There are several points worth noting that have bearing on the discussion and results of this paper. The first concerns the anticipated improvement in performance for the improved word-recognition model. If one carefully considers the sources for recognition errors with the alpha-digit vocabulary, it should become clear that the anticipated improvement resulting from the nonstationary analysis should be small unless some extra weighting is applied to the nonstationary regions. This is because words that are strongly affected by the nonstationary analysis (e.g., *p*, *d*, *t*, *k*, etc.) are easily confused with similar words in the vocabulary (e.g., *b*, *v*, *g*, *a*, etc.), and since the nonstationary regions are only a small subset of the word patterns, the improved analysis will be swamped out by the word-similarity regions. This is the original motivation for the discriminant analysis model used in the two-pass word recognizer.¹¹ Hence, the results of Section IV, which show a small (but consistent gain) for the improved analysis model and a somewhat larger gain for the discriminant model, are entirely consistent with the anticipated results given above.

A second point of note is that the implementation of the improved word model was more of a convenient one, rather than one that naturally followed from the theory. Thus, the short-time features were LPC coefficient sets derived from a short-time window. This implementation was straightforward and required only minimal modification of the recognizer structure. A more reasonable implementation of the short-time analysis in the model would have been something like a filter bank model, or a basilar membrane model. Such features would then have complemented the long-time LPC features and would have provided a better vehicle for testing and evaluating the improved model. The problem with using these alternative short-time feature sets is that there is no simple way of combining LPC and filter bank (or basilar membrane model) features and deriving from them a distance measure with good physical properties. The problem of combining LPC and energy features has already been investigated by Brown and Rabiner,²⁰ and it was shown that no simple metric existed even for such a simple case. The main point in the above discussion is that the small gain of the improved word model is more impressive when one considers the simplicity of the short-time analysis used to provide the performance gain.

The third point of note is the fact that the simple weighting derived from the robust training procedure seemed to provide the same performance improvement as the more sophisticated weighting obtained by using multiple tokens in obtaining the weights. The obvious conclusion to be drawn from the result is that the gain obtained from the second pass (which is due primarily to small regions of extreme spectral difference) is manifested in any pair of training tokens and that simple

smoothing (to eliminate statistical variability) is as good as using multiple tokens.

When one takes into consideration all of the above points, the results of Section IV provide a reasonable basis for believing that the improved word-recognition model is a reasonable one and that both the nonstationary analysis of the first pass, and the discrimination analysis of the second pass provide real performance gains.

VI. SUMMARY

An improved word-recognition model was proposed in which the standard long-time analysis features of the model are combined with a set of short-time analysis features. A stationarity index is also computed for each speech frame indicating which set of features (long-time or short-time) best characterized the current frame of speech. Appropriate modifications to the DTW algorithm were required to handle the enhanced analysis feature set. Also incorporated in the recognition model was a speaker-trained version of the discriminant analysis, two-pass model proposed by Rabiner and Wilpon.¹¹

An evaluation of the model based on an LPC implementation of both long-time and short-time feature sets showed the overall improved word model had from 1- to 5.7-percent improvement in recognition accuracy across four experienced users of speech recognition systems using an alpha-digit word vocabulary. On an average the nonstationary feature set alone led to a 0.5-percent improvement in accuracy, whereas the two-pass discriminant analysis alone led to a 1.6-percent average improvement in accuracy. The two improvements were almost independent and the overall recognizer had, on average, a 2.1-percent improvement in word accuracy.

The above results are considered encouraging and indicate that the improved model should be considered with alternative short-time feature sets.

REFERENCES

1. T. B. Martin, "Practical Applications of Voice Input to Machines," Proc. IEEE, 67 (April 1976), pp. 487-501.
2. N. R. Dixon and T. B. Martin, eds., *Automatic Speech and Speaker Recognition*, New York: IEEE Press, 1979.
3. W. Lea, ed., *Trends in Speech Recognition*, Englewood Cliffs, NJ: Prentice-Hall Inc., 1980.
4. D. R. Reddy, ed., *Speech Recognition*, New York: Academic Press, 1974.
5. L. R. Rabiner and S. E. Levinson, "Isolated and Connected Word Recognition—Theory and Selected Applications," IEEE Trans. on Commun., COM-29, No. 5 (May 1981), pp. 621-59.
6. G. M. White and R. B. Neely, "Speech Recognition Experiments with Linear Prediction, Bandpass Filtering, and Dynamic Programming," IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-24 (April 1976), pp. 183-8.
7. L. R. Rabiner, A. E. Rosenberg, J. G. Wilpon, and W. J. Keilin, "Isolated Word Recognition for Large Vocabularies," B.S.T.J., 61, No. 10 (December 1982).

8. B. Aldefeld, L. R. Rabiner, A. E. Rosenberg, and J. G. Wilpon, "Automated Directory Listing Retrieval System Based on Isolated Word Recognition," *Proc. IEEE*, 68 (November 1980), pp. 1364-79.
9. S. E. Levinson, "The Effects of Syntactic Analysis on Word Recognition Accuracy," *B.S.T.J.*, 57 (May-June 1977), pp. 1627-44.
10. L. R. Rabiner, J. G. Wilpon, and A. E. Rosenberg, "A Voice-Controlled Repertory-Dialer System," *B.S.T.J.*, 59 (September 1980), pp. 1153-63.
11. L. R. Rabiner and J. G. Wilpon, "A Two-Pass Pattern-Recognition Approach to Isolated Word Recognition," *B.S.T.J.*, 60, No. 5 (May-June 1981), pp. 739-66.
12. F. Itakura, "Minimum Prediction Residual Principle Applied to Speech Recognition," *IEEE Trans. Acoustics, Speech, and Signal Processing*, ASSP-23, No. 1 (February 1975), pp. 67-72.
13. J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech*, New York: Springer-Verlag, 1976.
14. P. V. de Souza, "Statistical Tests and Distance Measures for LPC Coefficients," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-25 (December 1977), pp. 554-9.
15. J. M. Tribolet, L. R. Rabiner, and M. M. Sondhi, "Statistical Properties of an LPC Distance Measure," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-27, No. 5 (October 1979), pp. 550-8.
16. H. Sakoe and S. Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-26 (February 1978), pp. 43-9.
17. L. R. Rabiner, A. E. Rosenberg, and S. E. Levinson, "Considerations in Dynamic Time Warping for Discrete Word Recognition," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-26 (December 1978), pp. 575-82.
18. C. S. Myers, L. R. Rabiner, and A. E. Rosenberg, "Performance Trade-offs in Dynamic Time Warping Algorithms for Isolated Word Recognition," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-28 (December 1980), pp. 622-35.
19. L. R. Rabiner and J. G. Wilpon, "A Simplified, Robust Training Procedure for Speaker Trained, Isolated Word Recognition Systems," *J. Acoust. Soc. Amer.*, 68, No. 5 (November 1980), pp. 1271-6.
20. M. K. Brown and L. R. Rabiner, "On the Use of Energy in LPC Based Recognition of Isolated Words," *B.S.T.J.*, 61, No. 10 (December 1982).

A Theoretical Model of Transient Heat Transfer in a Firestopped Cable Bundle

By P. B. GRIMADO

(Manuscript received April 7, 1982)

A mathematical model of heat transfer in a firestopped vertically oriented cable bundle is derived to assist in planning fire test experiments and to enumerate how changes in geometric and thermophysical properties affect the temperature rise in the cables when subjected to standard furnace fire tests. The analysis indicates that the primary heat transfer mode to the cable array is from the flow of hot furnace gases up through the void spaces between the individual cables. As expected, the most practical and effective way of reducing the heat transmission characteristics of a cable bundle is by tightly packing the firestop, which reduces the void space between cables and provides heat sinking to the cooler environs.

I. INTRODUCTION

In this paper a mathematical model of heat transfer in the cable bundle of a firestopped vertical cable assembly is developed to assist in planning experiments and to evaluate the relative influence of the geometrical and thermophysical properties of this portion of the firestopped configuration. A representation of this complex cable bundle geometry in terms of an approximate transient one-dimensional, lumped parameter model is obtained through heuristic arguments. This is accomplished in a systematic way by first deriving a simplified model for a single conductor wire and progressing up in scale, by averaging and lumping parameters, to arrive at a heat transfer model of a single cable. The heat transfer in a cable bundle is then treated using the single-cable model. Cable-to-cable heat transfer is handled through boundary conditions at the contact surfaces of the individual cables. A set of coupled transient one-dimensional equations results, with as many equations as cables in the array. The model is then exercised to compute transient temperature distributions within a

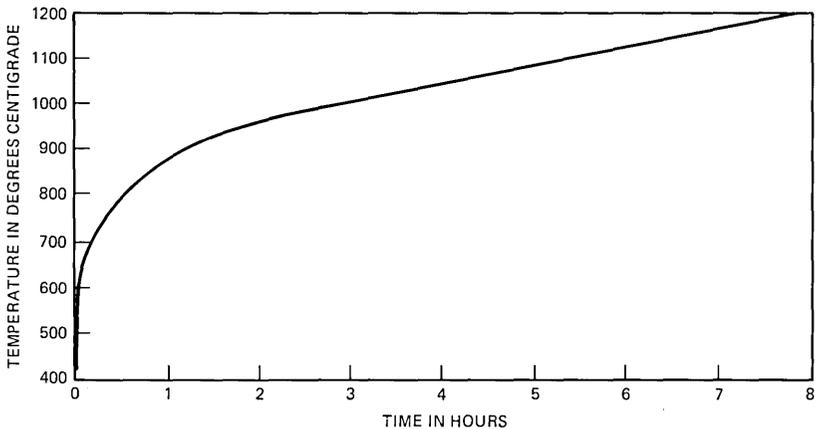


Fig. 1—ASTM E119 standard temperature vs time curve.

vertically oriented firestopped cable bundle when exposed to the ASTM E119* temperature variation shown in Fig. 1.

Properties of three generic-type cables (terminating, switchboard, and power cable) are used to suggest how physical and geometrical characteristics of the cables influence the effectiveness of a firestopped cable closure. It should be noted that an absolute evaluation of a firestop using the approximate heat transfer model developed here is beyond the capability of the model. This can only be attempted when considering potential nonlinear combustion modes of the polymeric cable materials, which is beyond the scope of this paper.

II. HEAT CONDUCTION MODEL OF A SINGLE CABLE

A cable consists of a core containing the insulated wire conductors and an outer protective sheath. In general, the cable core will contain a loose array of copper wires (which constitute 40 to 50 percent of the core) covered with a thin, polymeric insulating layer. The dominant path of heat conduction, because of their high conductivity, is in the longitudinal direction of the copper wires. Heat is also transferred radially, through the porous array of wires, by virtue of thermal radiation and heat conduction through contact points along the length of the wires, as shown in Fig. 2. The cable core cannot be considered a continuum because of the noncontiguous nature of the wires; hence, in developing the governing equation of heat conduction it is desirable to consider the individual wires as microstructural elements. This concept, as will be seen, permits interaction of the wires and leads, in

* Standard Methods of Fire Tests of Building Construction and Materials, American Society of Testing Materials.

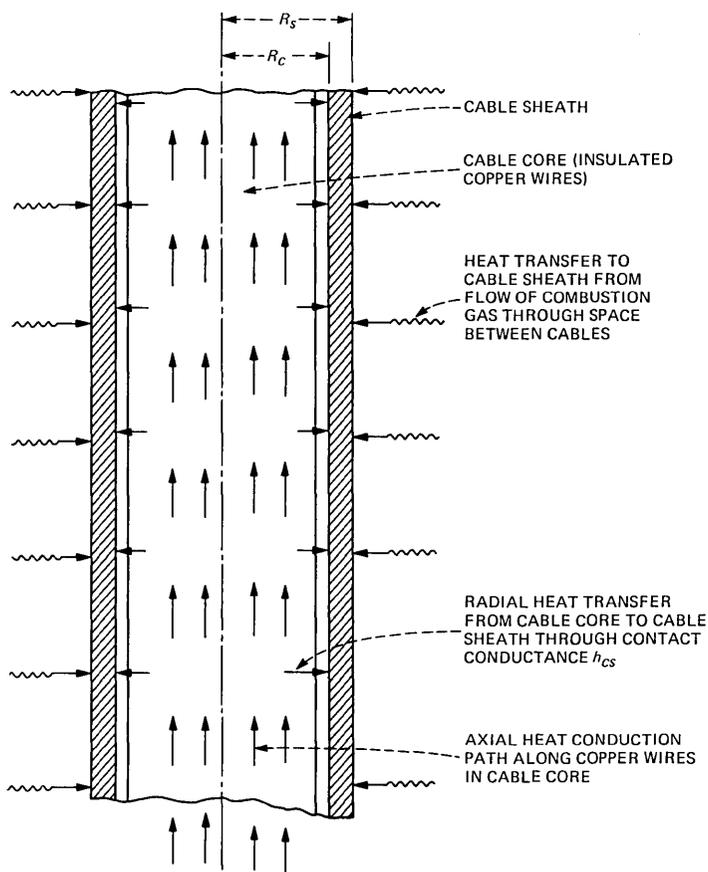


Fig. 2—Heat transfer modes in firestopped cables.

the limit, to an approximate continuum representation of heat conduction in the cable core. This development follows.

Consider the individual conductor wires as composite cylinders with copper wire radius r_c (see the appendix for a list of parameters) on which is affixed an insulating material with thickness $(r_d - r_c)$, as shown in Fig. 3. When radial symmetry prevails, Fourier heat conduction equations and boundary conditions for the conductor wire assume the following forms:

$$k_{cu} \left[\frac{\partial^2 T_{cu}}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial T_{cu}}{\partial r} \right) \right] = (\rho c)_{cu} \frac{\partial T_{cu}}{\partial t}, \quad \begin{array}{l} z > 0 \\ 0 \leq r \leq r_c, \\ t > 0 \end{array} \quad (1)$$

for the copper wire conductors, and

$$k_i \left[\frac{\partial^2 T_i}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial T_i}{\partial r} \right) \right] = (\rho c)_i \frac{\partial T_i}{\partial t}, \quad \begin{array}{l} z > 0 \\ r_c \leq r \leq r_d, \\ t > 0 \end{array} \quad (2)$$

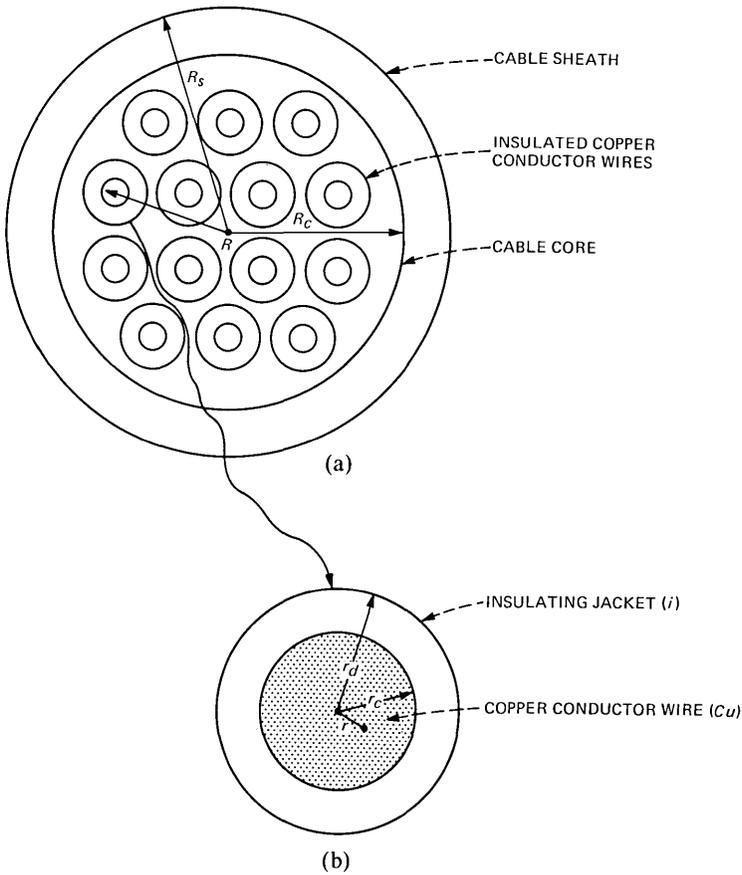


Fig. 3—(a) Macrocoordinate and (b) microcoordinate systems.

for the insulation, with boundary conditions:

$$T_{cu}(0, z, t) = \text{finite}, \quad (3)$$

$$T_{cu}(r_c, z, t) = T_i(r_c, z, t),^* \quad (4)$$

$$k_{cu} \frac{\partial T_{cu}}{\partial r}(r_c, z, t) = k_i \frac{\partial T_i}{\partial r}(r_c, z, t), \quad (5)$$

$$k_i \frac{\partial T_i}{\partial r}(r_d, z, t) = -h_i[T_i(r_d, z, t) - \hat{T}(z, t)], \quad (6)$$

and initial conditions:

$$T_{cu}(r, z, 0) = T_i(r, z, 0) = 0. \quad (7)$$

* A perfect contact between the copper wire and insulation is assumed, since the insulation is literally molded onto the wire.

In the above equations, $T(r, z, t)$ is the temperature, k is the thermal conductivity, ρc is the volumetric sensible heat capacity, z is the axial coordinate along the length of wire, r is the radial coordinate, t is the time, h_i is an effective linear heat transfer coefficient, and $\hat{T}(z, t)$ is a temperature to be assigned subsequently. The subscripts "cu" and "i" are assigned to quantities associated with the copper conductor and insulation, respectively. Since the diameter of the wire is much smaller than its length, it is convenient to express the temperature in the series

$$T_{cu}(r, z, t) = T_o(z, t) + T_{cu}^{(1)}(r, z, t) + \dots, \quad (8)$$

and

$$T_i(r, z, t) = T_o(z, t) + T_i^{(1)}(r, z, t) + \dots. \quad (9)$$

The leading term in eqs. (8) and (9) is the average temperature, i.e.,

$$T_o(z, t) = \frac{2}{r_d^2} \left[\int_0^{r_c} r T_{cu}(r, z, t) dr + \int_{r_c}^{r_d} r T_i(r, z, t) dr \right], \quad (10)$$

from which it follows that

$$\int_0^{r_c} r T_{cu}^{(n)}(r, z, t) dr + \int_{r_c}^{r_d} r T_i^{(n)}(r, z, t) dr = 0, \quad n = 1, 2, 3, \dots \quad (11)$$

Substituting series (8) and (9) into differential equations (1) and (2) produces a set of recurrent differential equations:

$$k_{cu} \left[\frac{\partial^2 T_{cu}^{(n-1)}}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial T_{cu}^{(n-1)}}{\partial r} \right) \right] = (\rho c)_{cu} \frac{\partial T_{cu}^{(n-1)}}{\partial t}, \quad n = 1, 2, \dots \quad (12)$$

$$k_i \left[\frac{\partial^2 T_i^{(n-1)}}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial T_i^{(n-1)}}{\partial r} \right) \right] = (\rho c)_i \frac{\partial T_i^{(n-1)}}{\partial t}. \quad n = 1, 2, \dots \quad (13)$$

All $T_{cu}^{(n)}(r, z, t)$ and $T_i^{(n)}(r, z, t)$, $n = 1, 2, 3, \dots$ are then expressible in terms of the average temperature, $T_o(z, t)$. Retaining terms up to first order, it follows from (12) and (13) that

$$T_{cu}(r, z, t) = T_o(z, t) - \frac{r^2}{4k_{cu}} L^{cu} + A(z, t) \ln r + B(z, t), \quad (14)$$

and

$$T_i(r, z, t) = T_o(z, t) - \frac{r^2}{4k_i} L^i + C(z, t) \ln r + D(z, t), \quad (15)$$

where

$$L^{cu} = k_{cu} \frac{\partial^2 T_o}{\partial z^2} - (\rho c)_{cu} \frac{\partial T_o}{\partial t}, L^i = k_i \frac{\partial^2 T_o}{\partial z^2} - (\rho c)_i \frac{\partial T_o}{\partial t}.$$

The variables of integration A, B, C, D are determined from boundary conditions, eqs. (3) through (5) and eq. (11). The final form of the temperature in the copper wire and insulation in terms of the average temperature $T_o(z, t)$ is

$$T_{cu}(r, z, t) = T_o(z, t) + \frac{1}{4k_{cu}} \left(r_c^2 - \frac{r_c^4}{2r_d^2} - r^2 \right) L^{cu} + \frac{r_c^2}{2k_i} \ln(r_d/r_c) (L^{cu} - L^i) + \frac{(r_d^2 - r_c^2)}{4k_i r_d^2} \left[-r_c^2 L^{cu} + \frac{(r_d^2 + r_c^2)}{2} L^i \right], \quad (16)$$

and

$$T_i(r, z, t) = T_o(z, t) + \frac{1}{k_i} \left(\frac{r_d^4 - r_c^4}{8r_d^2} + \frac{r_c^4}{4} - \frac{r^2}{4} \right) L^i + \frac{r_c^2 \ln(r/r_d)}{2k_i} (L^i - L^{cu}) - \left[\frac{r_c^2}{4k_i r_d^2} (r_d^2 - r_c^2) + \frac{r_c^4}{8k_{cu} r_d^2} \right] L^{cu}. \quad (17)$$

It now remains to obtain the differential equation for the average temperature, $T_o(z, t)$. This equation comes from the satisfaction of the boundary condition at $r = r_d$, as shown in eq. (6);

$$\left[\frac{1}{2} - \frac{r_c^2}{2r_d^2} + \frac{h_i}{8r_d^3 k_i} (r_d^2 - r_c^2)^2 \right] L^i + \left\{ \frac{r_c^2}{2r_d^2} + \frac{h_i}{8r_d^3} \left[\frac{2r_c^2(r_d^2 - r_c^2)}{k_i} + \frac{r_c^4}{k_{cu}} \right] \right\} L^{cu} = \frac{h_i}{r_d} [T_o(z, t) - \hat{T}(z, t)]. \quad (18)$$

Relating the variables in each conductor wire to those of neighboring ones in the limit leads to a continuum base from a discrete one. To arrive at this continuum description of the cable core consider the wires as embedded in a radially symmetric macrocoordinate system, as shown in Fig. 3. The center of, say, the n th conductor is located at R_n and eq. (18) is written at this point. In keeping with the assumption of radial symmetry, the temperature on the right of (18), $\hat{T}(z, t)$, is related to the average temperature of adjacent wires $T_o^{(n+1)}$ and $T_o^{(n-1)}$ located at $R_n + 2r_d$ and $R_n - 2r_d$, respectively. To determine $\hat{T}(z, t)$ the temperatures are weighted with respect to the location of the wire in the cable core. This has the effect of enforcing radial symmetry

$$\hat{T}(z, t) = \frac{1}{2R_n} [(R_n + r_d)T_o^{(n+1)} + (R_n - r_d)T_o^{(n-1)}]. \quad (19)$$

Thus, if D is used to denote the difference appearing on the right of (18), it can be written with the aid of (19) as

$$D = -\frac{1}{2R_n} [(R_n + r_d)T_o^{(n+1)} - 2R_n T_o^{(n)} + (R_n - r_d)T_o^{(n-1)}]. \quad (20)$$

When we expand $T_o^{(n+1)}$ and $T_o^{(n-1)}$ in a Taylor series about R_n , eq. (20) becomes

$$D = -\frac{2(r_d)^2}{R_n} \left\{ \frac{\partial}{\partial R_n} \left[R_n \frac{\partial T_o^{(n)}}{\partial R_n} \right] + 0(r_d) \right\}. \quad (21)$$

If we substitute (21) into (18) and consider $r_d \ll 1$ so that only the first term on the right of (21) is significant, the final form of the continuum representation of the cable core, after the subscript n is eliminated, is obtained as:

$$\left[1 - \frac{r_c^2}{r_d^2} + \frac{h_i(r_d^2 - r_c^2)^2}{4r_d^3 k_i} \right] L^i + \left\{ \frac{r_c^2}{r_d^2} + \frac{h_i}{4r_d^3} \left[\frac{2r_c^2(r_d^2 - r_c^2)}{k_i} + \frac{r_c^4}{k_{cu}} \right] \right\} L^{cu} + \frac{4h_i r_d}{R} \frac{\partial}{\partial R} \left(R \frac{\partial T_o}{\partial R} \right) = 0. \quad (22)$$

The quantity $(4h_i r_d)$ has the dimensions of conductivity and represents the effective radial conductivity of the cable core. The conductivity in the axial direction reflects the effect of the microstructure and is given as

$$(k_z)_{\text{eff}} = \left[1 - \frac{r_c^2}{r_d^2} + \frac{h_i(r_d^2 - r_c^2)^2}{4r_d^3 k_i} \right] k_i + \left\{ \frac{r_c^2}{r_d^2} + \frac{h_i}{4r_d^3} \left[\frac{2r_c^2(r_d^2 - r_c^2)}{k_i} + \frac{r_c^4}{k_{cu}} \right] \right\} k_{cu}, \quad (23a)$$

and the effective heat capacity as

$$(\rho c)_{\text{eff}} = \left[1 - \frac{r_c^2}{r_d^2} + \frac{h_i(r_d^2 - r_c^2)^2}{4r_d^3 k_i} \right] (\rho c)_i + \left\{ \frac{r_c^2}{r_d^2} + \frac{h_i}{4r_d^3} \left[\frac{2r_c^2}{k_i} (r_d^2 - r_c^2) + \frac{r_c^4}{k_{cu}} \right] \right\} (\rho c)_{cu}. \quad (23b)$$

Notice that when the conductor wires are assumed insulated, i.e., $h_i = 0$, there is no radial heat flow and the effective heat capacity and axial conductivity are given, as expected, by the law of mixtures.

Now that the heat conduction equation for the cable core has been determined, it remains to obtain the equation that governs the heat transfer in the entire cable including the outer sheath. This development parallels the procedure employed above. The only exception is that, since the cable core and sheath are not generally in perfect contact, the average temperatures in the cable core and sheath are assumed to be different. In many cables the cable core can be pushed

through the cable sheath with only a moderate pressure. Since the axial conductivity of the cable core is much larger than that of the cable sheath, as a simplification, it appears reasonable to assume that the axial heat conduction in the cable sheath can be neglected. Thus, the expression for the temperature in the cable core and sheath assumes the form

$$T_o(R, z, t) = T_o^{(0)}(z, t) + \frac{(2R^2 - R_c^2)}{8k_R^{(c)}} L^{(0)}(z, t), \quad 0 \leq R \leq R_c \quad (24a)$$

and

$$T_s(R, z, t) = T_s^{(0)}(z, t) + \frac{1}{4\kappa_s} \left[R^2 - \frac{(R_s^2 + R_c^2)}{2} \right] \frac{\partial T_s^{(0)}}{\partial t}(z, t) \\ + \frac{R_c^2}{2} \left[\ln(R/R_c) - \frac{R_s^2}{R_s^2 - R_c^2} \ln(R_s/R_c) + \frac{1}{2} \right] \left[L^{(0)}(z, t) - \frac{1}{\kappa_s} \frac{\partial T_s^{(0)}}{\partial t} \right], \\ R_c \leq R \leq R_s, \quad (24b)$$

where

$$L^{(0)}(z, t) = (\rho c)_{\text{eff}} \frac{\partial T_o^{(0)}}{\partial t} - (k_z)_{\text{eff}} \frac{\partial^2 T_o^{(0)}}{\partial z^2}$$

and $T_o^{(0)}(z, t)$, $T_s^{(0)}(z, t)$ are the average cable core and cable sheath temperatures, respectively. The other parameters are defined in the appendix. The average cable core and cable sheath temperatures given above are related through the imperfect heat transfer boundary condition at $R = R_c$, namely,

$$k_R^{(c)} \frac{\partial T_o}{\partial R} = -h_{cs}(T_o - T_s)|_{R=R_c}. \quad (25)$$

It follows that

$$\left\{ \frac{R_c}{2} + \frac{h_{cs}R_c^2}{8k_R^{(c)}} - \frac{h_{cs}R_c^2}{2k_s} \right. \\ \cdot \left. \left[\frac{-R_s^2}{R_s^2 - R_c^2} \ln(R_s/R_c) + \frac{1}{2} \right] \right\} L^{(0)}(z, t) + h_{cs}T_o^{(0)}(z, t) \\ = \left\{ \frac{h_{cs}(R_c^2 - R_s^2)}{8\kappa_s} - \frac{h_{cs}R_c^2}{2\kappa_s} \left[-\frac{R_s^2 \ln(R_s/R_c)}{R_s^2 - R_c^2} + \frac{1}{2} \right] \right\} \\ \cdot \frac{\partial T_s^{(0)}}{\partial t}(z, t) + h_{cs}T_s^{(0)}(z, t), \quad (26)$$

where h_{cs} is the heat transfer coefficient between the cable core and cable sheath.

In the following section eqs. (24b) and (26) are used to determine

the heat transfer in an array of sheathed cables that form a bundle. However, before this is accomplished it is worthwhile to briefly summarize the preceding developments.

The cable core is obviously not a continuum; consequently, a microstructural characterization starting with an individual conductor wire was used to derive the governing radially symmetric heat conduction equation for the cable core. This derivation produces an effective or averaged macroscopic continuum representation of the cable core that retains the physical characteristics of the conductor wire microstructure. It is, however, anisotropic, since the thermal conductivity in the axial and radial directions is not equal. After we obtain the governing equations for heat flow in the core, the equations that the cable sheath temperature satisfies [eqs. (24b) and (26)] were then determined. These were obtained, paralleling the preceding microstructural derivation, for heat flow predominantly in the axial direction. The radial heat transfer is introduced as a perturbation.

III. HEAT TRANSFER THROUGH A CABLE BUNDLE

The objective of this investigation is to analyze heat flow through firestopped cable bundles during fire tests. The ASTM E119 temperature-time history, Fig. 1, provides the fire environment. A typical firestopped cable bundle configuration and the above- and below-floor coordinate system is illustrated in Fig. 4.

As previously mentioned, cable penetrations vary in size and in the number and type of cables accommodated; to lend a degree of definiteness to the analysis it is convenient to consider a widely used arrangement. A square cable array containing nine cables, as illustrated in Fig. 5, is a mathematically manageable configuration, yet most of the heat transfer characteristics of larger cable arrays are maintained. From the point of view of symmetry, only three cables need be considered—the corner cable, to be designated hereafter as cable 1, the side cable, to be denoted as cable 2, and the center cable, to be called cable 3.

The section of each cable below the slab (see Fig. 4) is directly exposed in a furnace to a fire temperature of up to 1000 degrees centigrade and, therefore, after only a short time all polymeric insulating material is burned away. In the model this effect is approximated by assuming that only a loose array of independent copper wires projects below the slab. This section of the cable bundle extending up through and above the floor slab experiences a different thermal environment. A zone of decomposition of the polymeric insulating materials occurs and creeps upwards during the extended period of exposure to the below-floor fire. This zone of charred and expanded insulating material alters the temperature in the void spaces. In the

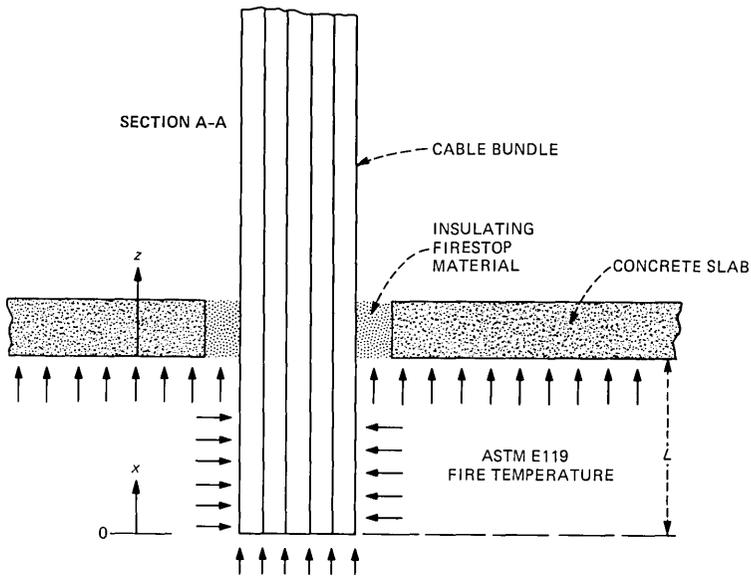
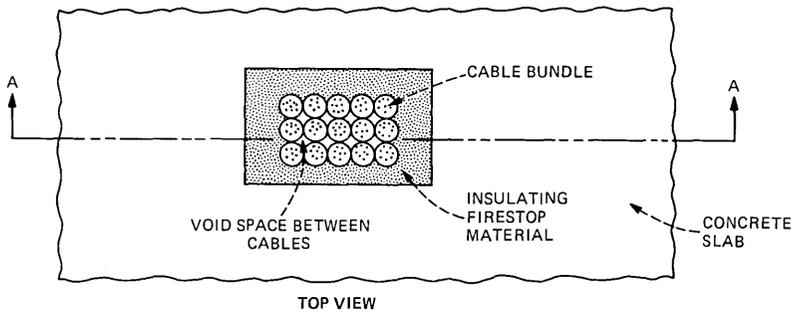


Fig. 4—Typical cable opening firestop test configuration.

present analysis the effect of this nonlinear phenomenon is ignored to render the analysis tractable. It is conservatively assumed here that the hot furnace gases move unimpeded in the void space between cables.

The average cable core and average cable sheath temperature for each of three cables is determined by performing a heat balance at the outer cable sheath surface $R = R_s$. This leads to the following three coupled equations:

$$k_s \frac{\partial T_s^{(1)}}{\partial R} (R_s, z, t) = 2A_c H_{cB} [T_s^{(2)} - T_s^{(1)}] + (3A_f + 2A_c) H_f [T_f^{(0)} - T_s^{(1)}]$$

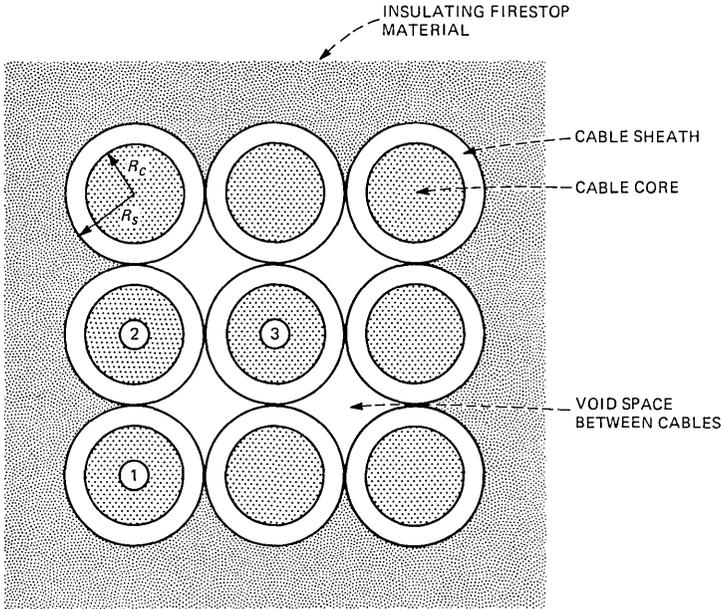


Fig. 5—Cable bundle configuration considered.

$$\begin{aligned}
 &+ A_f H_1 [T_f^{(1)} - T_s^{(1)}] + 2A_f H_r F_{12} [T_s^{(2)} - T_s^{(1)}] \\
 &+ A_f H_r F_{22} [T_s^{(3)} - T_s^{(1)}] |_{R=R_s} \quad (27a)
 \end{aligned}$$

$$\begin{aligned}
 k_s \frac{\partial T_s^{(2)}}{\partial R} (R_s, z, t) = &2A_c H_{cB} [T_s^{(1)} - T_s^{(2)}] + (A_c H_{cB} + 2A_f H_r F_{12}) \\
 &\cdot [T_s^{(3)} - T_s^{(2)}] + 2A_f H_1 [T_f^{(2)} - T_s^{(2)}] + (2A_f + A_c) H_f [T_f^{(0)} - T_s^{(2)}] \\
 &+ 2A_f H_r F_{12} [T_s^{(1)} - T_s^{(2)}] |_{R=R_s} \quad (27b)
 \end{aligned}$$

$$\begin{aligned}
 k_s \frac{\partial T_s^{(3)}}{\partial R} (R_s, z, t) = &4A_c H_{cB} [T_s^{(2)} - T_s^{(3)}] + 4A_f H_1 [T_f^{(1)} - T_s^{(3)}] \\
 &+ 8A_f H_r F_{12} [T_s^{(2)} - T_s^{(3)}] + 4A_f H_r F_{22} [T_s^{(1)} - T_s^{(3)}] |_{R=R_s}. \quad (27c)
 \end{aligned}$$

In the above equations the parameter A_c denotes the fraction of the cable sheath perimeter that is in contact with an adjacent cable, and A_f the portion exposed to the hot furnace gases that flow up through the interstices of the cable. It therefore follows that $4(A_c + A_f) = 1$. H_{cB} is the solid-contact conductance between the cables, H_f is the heat transfer coefficient between the cables and the firestop material, H_1 is the convective heat transfer coefficient between the cable and the hot furnace gases that flow up through the void spaces between cables,

and H_r is a linearized black body radiative heat transfer coefficient between the cables. F_{12} and F_{22} are radiation form factors between cable 3 and cable 2, and cable 1 and cable 3, respectively. The temperatures $T_f^{(0)}(z, t)$ and $T_f^{(1)}(z, t)$ are assumed known and represent the temperature distribution in the mineral wool insulating firestop material surrounding the cable bundle and the temperature of the furnace gases that move up through the void spaces between cables, respectively.

At this juncture, it is convenient to eliminate the independent time variable t from the equations by introducing the Laplace transformation

$$\bar{T}(z; p) = \int_0^\infty T(z, t)e^{-pt} dt. \quad (28)$$

In what follows, a bar over a variable indicates that the transformation (28) has been performed. Substituting eq. (24b) into (27) and using eq. (26) to eliminate the average cable sheath temperature, we produce a system of three coupled ordinary differential equations for the transformed average cable core temperatures $\bar{T}_{0,1}^{(0)}$, $\bar{T}_{0,2}^{(0)}$, and $\bar{T}_{0,3}^{(0)}$:

$$(\Gamma_1 D^2 - \Delta_1) \bar{T}_{0,1}^{(0)} - 2(\gamma_1 D^2 - \theta_1) \bar{T}_{0,2}^{(0)} - (\eta_1 D^2 - \epsilon_1) \bar{T}_{0,3}^{(0)} = -(h_{cs} + pE_1)[(3A_f + 2A_c)H_f \bar{T}_f^{(0)} + A_f H_1 \bar{T}_f^{(1)}], \quad (29a)$$

$$-2(\gamma_1 D^2 - \theta_1) \bar{T}_{0,1}^{(0)} + (\Gamma_2 D^2 - \Delta_2) \bar{T}_{0,2}^{(0)} - (\eta_2 D^2 - \epsilon_2) \bar{T}_{0,3}^{(0)} = -(h_{cs} + pE_1)[(2A_f + A_c)H_f \bar{T}_f^{(0)} + 2A_f H_1 \bar{T}_f^{(1)}], \quad (29b)$$

$$-4(\eta_1 D^2 - \epsilon_1) \bar{T}_{0,1}^{(0)} - (\gamma_3 D^2 - \theta_3) \bar{T}_{0,2}^{(0)} + (\Gamma_3 D^2 - \Delta_3) \bar{T}_{0,3}^{(0)} = -4(h_{cs} + pE_1)A_f H_1 \bar{T}_f^{(1)}, \quad (29c)$$

where the operator $D^2 = \frac{d^2}{dz^2}$ and the coefficients Γ , Δ , γ , θ , η , and ϵ are linear functions of the transform parameter p . After considerable algebra the system of differential equations (29) is solved for the three average cable core temperatures in the form

$$\bar{T}_0^{(0)} = C_1 \exp(-z\sqrt{m_1}) + C_2 \exp(-z\sqrt{m_2}) + C_3 \exp(-z\sqrt{m_3}) - (h_{cs} + pE_1)[R\bar{T}_f^{(0)} + H\bar{T}_f^{(1)}]. \quad (30)$$

The quantities m_1 , m_2 , and m_3 are functions of the Laplace transform parameter p and are evaluated as the roots of the characteristic cubic equation of the system of equations in (29). The coefficients of integration C_1 , C_2 , and C_3 are determined by matching the solution of the cable section below the slab. This is accomplished by enforcing tem-

perature and energy continuity for each of the three cables at the common boundary $Z = 0$ and $X = L$, as shown in Fig. 4. The cable sheath temperature evaluated at $R = R_s$ follows from (30) in the form

$$\bar{T}_s(R_s, z, p) = \left[\frac{h_{cs} + pB_1}{h_{cs} + pE_1} (1 + pA_1) + pD_1 \right] \bar{T}_0^{(0)} - (\kappa_z)_{\text{eff}} \left[\frac{B_1(1 + pA_1)}{(h_{cs} + pE_1)} + D_1 \right] \frac{d^2 \bar{T}_0^{(0)}}{dz^2}, \quad (31)$$

where A_1 , B_1 , D_1 and E_1 are constants.

The temperatures presented in this section are in the Laplace transform domain and must be inverted to the real-time regime. The functions, however, are too complicated to be inverted in closed form. A numerical procedure using the method of quadratures to obtain these inversions is discussed in Section IV.

IV. INVERSION OF LAPLACE TRANSFORM TEMPERATURE SOLUTION

The form of the transformed temperature solution given in eqs. (30) and (31) is much too complicated to use for obtaining a closed-form inversion formula. Consequently, we must resort to a numerical inversion procedure. Most of the methods that appear in the literature involve expanding the transformed function in a series that could then be inverted term by term using tabulated formulae. Littlewood and Zakian¹ suggest expanding in a series of Chebyshev polynomials, while Longman² proposes using the Pade table for the Taylor series expansion of the transformed function. Both of these methods were judged to be impractical because of the complexity of the functions to be inverted here.

The method that was finally adopted was developed by Talbot.³ The inversion of arbitrary transforms is accomplished by a method of quadrature along a special contour in the complex plane. The standard inversion formula for a transformed function $\bar{F}(p)$ involves performing the following integration in the complex plane:

$$F(t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \bar{F}(p) e^{pt} dp, \quad (32)$$

where p is considered a complex variable, $i = (-1)^{1/2}$, and γ is to the right of all the singularities of $\bar{F}(p)$. In Talbot's method the path of integration indicated by (32) is deformed to the path L shown in Fig. 6, the equation of which is

$$p = \alpha + i\theta, \quad (33)$$

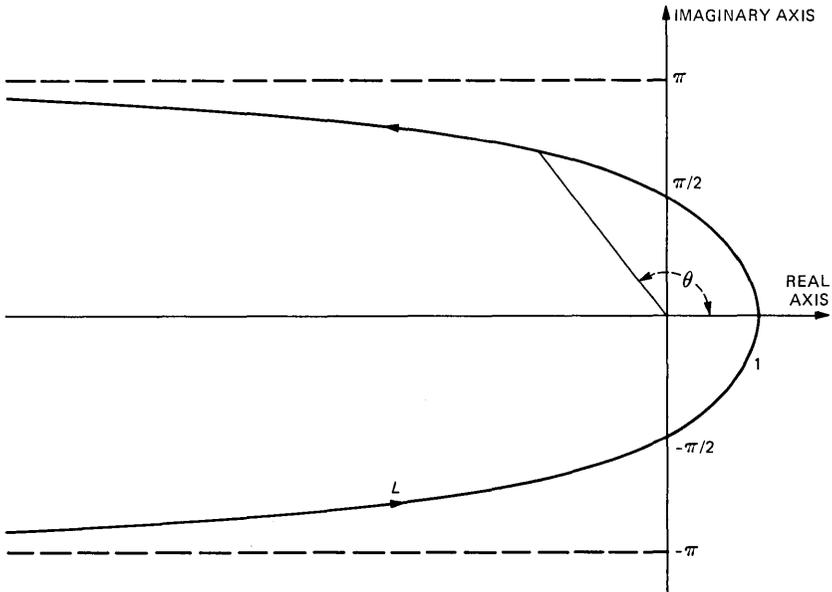


Fig. 6—Path L for Laplace transform inversion.

where $\alpha = \theta \cot \theta$.

Path L is equivalent to the standard integration path providing that
 (i) L encloses all singularities of $\bar{F}(p)$, and

(ii) $|\bar{F}(p)| \rightarrow 0$ uniformly in $Re p \leq 0$ as $|p| \rightarrow \infty$.

Condition (ii) holds for the functions considered here. Condition (i), in general, may not be satisfied by a given $\bar{F}(p)$; however, the modified function $\bar{F}(\lambda p + \sigma)$ for suitable scaling and shift factors λ and σ can be made to conform. In this regard, for the functions to be inverted here it can be argued that the singularities exhibited by eqs. (30) and (31) are located on the negative real axis; consequently, condition (i) is satisfied without resorting to scaling or shift factors. However, to accurately perform the inversions over large intervals of time (the temperatures are calculated for times up to two hours) a scaling factor $\lambda > 1$ is necessary. This scaling factor merely shifts the singularity along the negative real axis closer to the origin.

Once conditions (i) and (ii) are satisfied, the inversion formula (32), when taken about L assumes the form

$$F(t) = \frac{\lambda e^{\sigma\tau}}{2\pi i} \int_{-\pi}^{\pi} e^{p\tau} \bar{F}(\lambda p + \sigma) \frac{dp}{d\theta} d\theta, \quad \tau = \lambda t. \quad (34)$$

Finally, "1/3" Simpson quadrature with equal intervals π/n in the variable θ gives the approximation

$$F(t) \sim \frac{\lambda e^{\sigma t}}{3n} \{Re(H_0 + 4H_1 + 2H_2 + 4H_3 + \dots + 4H_{n-1})\}, \quad (35)$$

where

$$H_k = H(p)|_{\theta_k},$$

$$\theta_k = k\pi/n \quad (k = 0, \dots, n-1)$$

$$H(p) = e^{pr} \bar{F}(\lambda p + \sigma)(1 + i\beta),$$

and

$$\beta = [\theta - \frac{1}{2} \sin(2\theta)]/\sin^2\theta.$$

The symbol *Re* indicates that only the real part of the complex quantity is taken. Sufficient accuracy is obtained by a suitable choice of n , λ , and σ . The principles of choice are presented by Talbot³ and will not be discussed here.

A check of this technique was made by inverting the function given in Ref. 2. In all instances satisfactory results were obtained. For numerical inversions performed here, it was found sufficient to take $n = 20$, $\lambda = 8$, and $\sigma = 0$.

V. DISCUSSION AND SUMMARY

The concept of firestopping any type of penetration is synonymous with retarding the flow of heat from the fire side of the penetration to the unexposed side. The problem rests solely on the identification of procedures and materials to seal the space adjacent to the penetrant to meet certain standards. Cables by their very nature have excellent thermal conduction properties in the longitudinal direction. Very little can be done to prevent heat conduction up the array of copper wire conductors that make up the center core section of the cable. However, it appears reasonable to measure the efficiency of a firestop by the temperature rise on the free surfaces on the unexposed side of the penetration. If these surfaces are maintained at sufficiently low temperatures, ignition of combustibles that happen to be in direct contact or in close proximity cannot occur. In the case of cable penetrations the critical surfaces are the horizontal firestopping material surface on the unexposed side and the vertical cable sheath surfaces at the outer perimeter of the cable bundle, as shown in Fig. 4. In normal practice the firestop material is chosen to be a good thermal insulator and is applied with sufficient depth to preclude high temperatures at the top surface. Thus, the efficiency of the firestopped geometry will, in general, be determined more by the temperature rise on the cable sheath surface of the outer perimeter cables than by the firestop material. To ascertain the temperature at this critical location, a

theoretical transient heat transfer model of a firestopped cable bundle was developed. In addition to heat flow through the cable core, the model treats the lateral heat flow from the cable core to the cable sheath and convective heat transfer to the cable sheath from the flow of combustion gases through the interstices between the cables. The standard ASTM E119 temperature variation is applied at one end of the cable bundle and the temperature distribution of the individual cables is computed along the length for up to two hours. To quantify these effects, temperature distribution was computed at 10-minute intervals for two hours at the 20-cm and 30-cm firestopped depth for three different cable types—switchboard, terminating, and power cable. For each of these cable types the following parameters were varied:

- (i) R_s , cable radius
- (ii) A_c , fraction of cable sheath surface in contact with an adjacent cable
- (iii) h_{cs} , conductance between cable core and sheath
- (iv) H_f , conductance between cable bundle perimeter and firestop material
- (v) Δp , furnace pressure.

Some remarks concerning the inclusion of the furnace pressure as a parameter are in order. Furnace pressure influences the magnitude of the heat transmitted to the cable sheath—by the hot furnace gases that travel up through the space between cables—through the heat transfer coefficient H_1 , as shown in eq. (27). The value of H_1 is computed from standard empirical formulae⁴ once the gas flow velocity and flow-channel characteristics are known. The steady-state gas velocity is calculated from the furnace pressure assuming that the spaces between cables are independent flow channels.⁵ In general, H_1 increases with increased furnace gas pressure.

Some of the results of the analysis for a given type cable are given in Table I and Fig. 7. The calculated end-point temperature at $z = 30$ cm in Fig. 6 and $t = 2$ hours for the cable sheath and core are given in Table I, and a typical temporal temperature distribution is given in Fig. 7. The temperatures given in Table I and Fig. 7 should not be construed to be indicative of actual measured test values. As previously indicated, the model does not take into account nonlinear aspects of this obviously complex phenomenon, such as potential combustion modes and melting of the polymeric materials that in some instances could conceivably, for periods of time, constrict the void of spaces between cables and thereby reduce the flow of furnace gases. Nevertheless, the sensitivity of the cable temperatures to changes in the linear cable bundle parameters identified in the model, H_f , A_c , R_s , and Δp can be calculated and are presented in Table I.

Table I—Calculated temperatures at $z = 30$ cm and $t = 2$ hours for various values of the cable bundle parameters

Case	H_f ($W/cm^2-^{\circ}C$)	A_c	R_s (cm)	Δp (cm of Water)	Cable Sheath Temperature ($^{\circ}C$) Rise		Cable Core Temperature ($^{\circ}C$) Rise	
					Cable 1	Cable 3	Cable 1	Cable 3
1	0.034	0.1	1.27	1	115	164	199	246
2	0.017	0.1	1.27	1	224	265	293	332
3	0.017	0.15	1.27	1	170	190	248	270
4	0.017	0.1	0.635	1	93	115	135	154
5	0.017	0.1	1.27	0.13	102	115	188	204

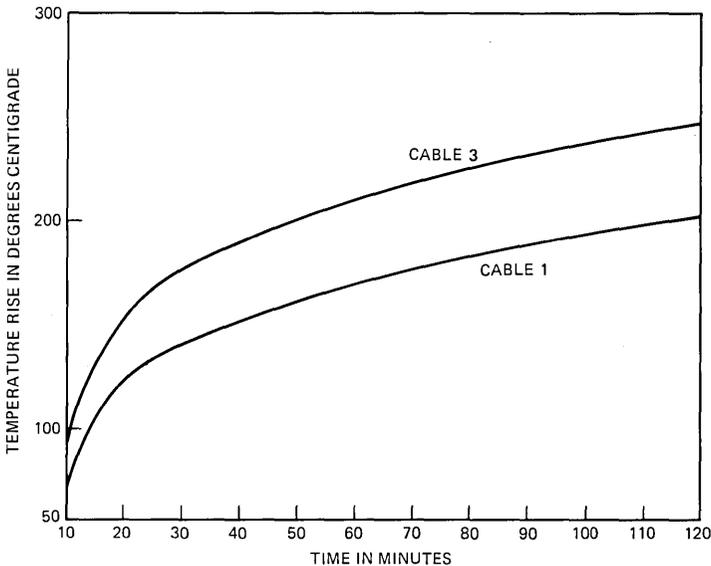


Fig. 7—Cable core temperature for Case 1 cable bundle parameters given in Table I.

It is clearly shown in Table I that the cable core temperature exceeds that of the cable sheath. A tightly packed firestop that exerts lateral pressure on the side of the cable array will provide sufficient heat sinking to reduce the temperature in the cable bundle. This physical effect is embodied in the contact conductance parameter, H_f . The larger the numerical value of this parameter, the tighter the firestop packing. Cases 1 and 2 of Table I show that increasing this parameter indeed results in a lowering of the cable sheath and cable core temperatures.

The same general result prevails when furnace overpressure Δp is reduced. This effectively reduces the heat transfer coefficient H_1 between the cable sheath and the hot furnace gases. This is demonstrated in Table I by comparing the end-point cable temperatures of

Cases 2 and 5. Cable temperature can also be reduced by increasing the contact surface between cables, as measured by an increase in the parameter A_c , suggesting a tighter and more compact cable bundle. The cable temperatures, as observed in Cases 2 and 3, are also lowered. This is primarily due to the resulting smaller void space between the cables and secondarily to the larger conduction path presented to the interior cables.

VI. CONCLUSIONS

The following trends were generally indicated:

(i) The cable core and cable sheath temperature is largest for the interior cable (cable 3 of Fig. 5) and the smallest for the corner cable (cable 1 of Fig. 5).

(ii) The temperature of cable core exceeds that of the cable sheath.

(iii) The primary heat transfer mode to the cable sheath is from the flow of hot combustion gases through the void space between cables.

(iv) Reducing the void space between the cables by tightly packing the cables and/or using smaller diameter cables impedes the flow of hot combustion gases and results in a significant reduction of the primary convective heat transmission mode.

(v) A tightly packed firestop capable of providing some heat sinking to the cooler environs is the most practical and effective method of reducing the heat transfer properties of the cable bundle.

(vi) The magnitude of the cable sheath temperature for similar size cables depends on the furnace gas pressure and, to a lesser degree, on the firestop depth.

VII. ACKNOWLEDGMENT

The comments and suggestions made by J. Simpson in the preparation of the text are appreciated. In addition, appreciation is expressed to W. Pferd for reviewing several drafts of the manuscript and suggesting valuable comments for its improvement.

REFERENCES

1. R. K. Littlewood and V. Zakian, "Numerical Evaluation of Fourier Integrals," *J. Inst. Maths. Applics.*, 18 (December 1976), pp. 333-9.
2. I. M. Longman, "Numerical Laplace Transform Inversion of a Function Arising in Viscoelasticity," *J. Comp. Phys.*, 10 (December 1972), pp. 224-31.
3. A. Talbot, "A New Method for Numerical Inversion of Laplace Transforms," *IEEE Proc.*, *ISCAS* (1976), pp. 130-3.
4. C. P. Kothandaraman and S. Subramanyan, *Heat and Mass Transfer Data Book*, New York: Halsted Press, 3rd edition, 1977.
5. R. G. Binder, *Fluid Mechanics*, Englewood Cliffs, N.J.: Prentice-Hall, 3rd edition, 1955.

APPENDIX

- A_c —Fraction of cable sheath surface in contact with an adjacent cable.
 A_f —Fraction of cable surface in contact with combustion gases, $4(A_c + A_f) = 1$.
 F_{12}, F_{22} —Radiation form factors between cables.
 H_1 —Heat transfer coefficient between cable and combustion gases.
 H_{cB} —Solid contact conductance between cables.
 H_f —Contact conductance between cable bundle perimeter and firestop.
 H_r —Linearized black body radiative heat transfer coefficient.
 h_{cs} —Conductance between cable core and cable sheath.
 h_i —Conductance between individual copper conductor wires.
 $k_R^{(c)} = 4h_i r_d$ —Effective radial thermal conductivity of cable core.
 k_{cu} —Thermal conductivity of copper wire.
 k_i —Thermal conductivity of wire insulation.
 k_s —Thermal conductivity of cable sheath.
 $(k_z)_{\text{eff}}$ —Effective axial thermal conductivity of cable core.
 $\kappa_{cu} = k_{cu}/(\rho c)_{cu}$ —Thermal diffusivity of copper wire.
 $\kappa_i = k_i/(\rho c)_i$ —Thermal diffusivity of wire insulation.
 $\kappa_s = k_s/(\rho c)_s$ —Thermal diffusivity of cable sheath.
 $(\kappa_z)_{\text{eff}} = (k_z)_{\text{eff}}/(\rho c)_{\text{eff}}$ —Effective axial thermal diffusivity of cable core.
 L —Length of cable below slab.
 R_c —Radius of cable core.
 R_s —Overall radius of cable.
 r_c —Radius of single copper conductor wire.
 r_d —Total radius of single conductor wire.
 Δp —Furnace pressure.
 $(\rho c)_{cu}$ —Sensible volumetric heat capacity of copper wire.
 $(\rho c)_{\text{eff}}$ —Effective sensible volumetric heat capacity of cable core.
 $(\rho c)_i$ —Sensible volumetric heat capacity of wire insulation.
 $(\rho c)_s$ —Sensible volumetric heat capacity of cable sheath.

Planning and Conducting Field-Tracking Studies

By S. J. AMSTER, G. G. BRUSH, and B. SAPERSTEIN

(Manuscript received April 5, 1982)

This paper reviews the major aspects of planning and conducting field-tracking studies, including: (i) establishing well-defined, realistic objectives; (ii) designing data collection and analysis procedures to meet the objectives; and (iii) ensuring the successful implementation of these procedures. The paper gives general guidelines on matching study objectives and procedures, as well as detailed information on sample size selection for some common field-study situations. Several studies recently conducted by Bell Laboratories Quality Assurance Center are used to illustrate the principles of field-study planning and implementation.

I. INTRODUCTION

It is the function of Bell Laboratories Quality Assurance Center (QAC) to provide assurance that telecommunication products purchased by the Bell Operating Companies (BOCs) are of satisfactory quality and perform as required. This assurance is provided through the three primary activities of the Quality Assurance effort:

(i) Quality inspection and auditing at manufacturing, repair, and installation locations.

(ii) Qualitative feedback gathered through informal contacts with BOC personnel and a more formal engineering complaint procedure.

(iii) Quantitative field-tracking studies of selected products and systems.

This paper discusses the third activity from both a historical and tutorial point of view. The authors relate some lessons and principles learned through field-tracking studies in the past and offer suggestions for those planning to conduct a field-tracking study (FTS) in the future.

Formal field-tracking studies were undertaken during the 1960s. The studies that will be described in this paper began in 1973 with Product Performance Surveys (PPS)¹ on Western Electric station sets. PPS are

designed to track field performance of the sets, identify problems quickly, quantify the extent of those problems so that economic corrective action can be taken, and assure that the "fixes" are effective. Typically, PPS on station sets are conducted concurrently in five or six BOC locations chosen to provide geographic and climatic diversity and good representation of a variety of set types. This permits approximately one million station sets to be tracked at any given time, and provides approximately 100,000 trouble events for recording and analysis each year.

PPS data on station sets have been instrumental in detecting and quantifying numerous field problems. Representative examples include a series of contact contamination problems in *Touch-Tone** dials, ringer failures in certain premium station sets, and lamp failures in key telephone sets.

The success of PPS has stimulated an increased effort into field studies of other products, such as PBX's, switching networks, channel bank equipment, switching machines—just about the entire range of telecommunications products purchased by the BOCs. Recently, this field-study effort has been extended to include selected general trade products manufactured by suppliers other than Western Electric. The remaining sections of this paper discuss principles learned by the authors in the process of conducting field-tracking studies and offer suggestions for those planning to conduct an FTS.

Section II discusses important considerations in planning an FTS; Section III discusses key steps in an FTS implementation program; Section IV is devoted to some illustrations from recent Quality Assurance Center studies.

II. PLANNING A FIELD-TRACKING STUDY

The principal steps involved in planning a successful FTS are:

- (i) Defining study objectives
- (ii) Planning data collection to meet those objectives
- (iii) Planning for successful data analysis.

2.1 *Defining study objectives*

Perhaps the single most important requirement for a successful FTS is a clear statement of purpose that has been agreed to by the concerned parties. A study will frequently have an impact on many different organizations through its implementation, interpretation, and the use of its results. The designer, the manufacturer, and the user all have legitimate concerns in a given FTS. Obtaining their understanding and agreement is an important, but not necessarily a simple, task.

* Registered service mark of AT&T.

Early in the planning of a study, small changes can easily be made to accommodate the needs of potential users. But care must be taken not to try to answer all questions with a single study. Setting precise objectives that simplify implementation can avoid many pitfalls. For example, taking all the data that are easily accessible may initially seem reasonable, since we certainly don't want to miss anything that might be important. But, trying to ensure that "too many" pieces and types of data are good invariably leads to a degraded level of data quality. The topic of data collection is discussed in detail in Section 2.2.2.

Frequently, objectives change as data are collected. This implies the need to provide for such changes initially and to monitor the flow of data to determine when such changes are appropriate. For example, a study that has the objective of comparing the performance of products from three suppliers may quickly show that one supplier is an obvious noncontender. Rules for dropping such a candidate could result in a more efficient use of resources.

Objectives can be classified² as:

- (i) Detecting problems
- (ii) Quantifying known problems
- (iii) Verifying quality audit information or reliability predictions
- (iv) Establishing problem causes
- (v) Measuring the impact of design or manufacturing change(s)
- (vi) Evaluating the product.

A study can involve aspects of several of these, but procedures must be matched to purposes. For example, some studies are intended primarily to find and make a preliminary evaluation of problems. Once a problem has been identified, a more detailed study can be used to better quantify its economic impact.

Early thinking about a proposed study may be clarified by the following list of objectives, stated in a statistical framework:

- (i) Point estimation (e.g., early failure rate)
- (ii) Interval estimation (e.g., confidence or prediction intervals)
- (iii) Comparisons (within study, with a standard or with results from a previous study)
- (iv) Model testing (e.g., decreasing failure rate)
- (v) Other information (previous list).

Failure to get agreement on specific objectives among all participants can easily lead to continuing disagreements regarding the implementation of the study and the interpretation of its results.

2.2 Planning data collection

Once the general objectives of a field study have been established, the work aimed at meeting those objectives begins with the planning of appropriate data collection procedures.

Most of this planning is aimed at answering the following questions: (i) *What* data will be collected? (ii) *How* will the data be collected? (iii) In what *study population* will the data be collected? and (iv) *How much* data (sample size) will be collected? Finding the appropriate answer to each of these questions for any given study is the key to its success. It is worthwhile examining each question separately and describing some of the answers that have been found appropriate in previous studies.

2.2.1 What data will be collected?

There are clearly many factors that will determine what data should be collected for any given field study. For purposes of this discussion, we assume that the study in question is directed at estimating the frequency of troubles occurring in a specified product population. This objective imposes the following minimum requirements on the data to be collected:

- (i) The data must include the size of the study population.
- (ii) The data must record or count every trouble "event" occurring in the study population during a specified time period, and must exclude or specifically identify events that are reported but occur outside the study population or specified time period.

Clearly, a field study satisfying only these minimum requirements will yield merely gross trouble rate information. However, there are a number of situations appropriate for such a minimal study.

First, for a larger, more detailed study, a preliminary estimate of the overall trouble rate is sometimes needed to determine the study population size. This topic will be further considered below, in the discussion on sample size (Section 2.2.4). Minimal data collection will usually suffice for such an estimate. Minimal data collection might also be appropriate after a detailed study to monitor the effectiveness of corrective actions that may have been taken in response to information obtained during the larger study.

A minimal program of data collection may also be justified in cases where the need for a larger, more detailed and more costly study must be demonstrated. Several tracking studies that we have conducted were operated in this way, with minimal trouble rate data collected until a need for more detailed information was indicated by observing higher than expected trouble rates.

For most field-tracking studies, however, minimal data collection falls short of what is needed in two important ways. First, since this approach provides no identification of the subpopulation in which any trouble occurs, it cannot yield specific trouble-rate estimates by subpopulations. Subpopulation, here, refers to a newly manufactured versus a repaired product, or to different manufacturing vintages of a

given product that may reside within a single overall study population. Second, because this approach provides no information on the nature of each trouble event, it cannot yield estimates of the frequency with which the product under study fails for specific reasons.

Information on subpopulations and trouble types makes up virtually all of the detailed data that must be collected for any study; and determining the level of detail for each is a principal objective of study planning.

As noted, subpopulation data would ordinarily include information on whether a piece of equipment in which a trouble occurred was newly manufactured or repaired, the date of manufacture or repair (vintage), service life, and additional descriptive information on the product, such as the issue or series number for a product that has undergone changes in design or manufacture. (Specifying series or issue numbers for circuit packs is an example of detailed product specifications used in tracking studies that are currently under way). Included, too, under the general heading of subpopulation information would be data on how or by whom the trouble was reported, e.g., customers or employees.

In almost all FTS situations the more detailed the data asked for, the more complicated and costly the collection process will have to be. Therefore, it is important to limit to the extent possible the level of detail in subpopulation data requested. The guiding principle in choosing which characteristics should be included in data collection is straightforward: Include only characteristics for which it will be both useful and worthwhile to obtain separate subpopulation trouble-rate estimates when all the data have been collected. Since almost any level of detail can be viewed as potentially useful, the key is to choose only those characteristics that produce "partitions" that will be worthwhile, i.e., that will yield subpopulations of sufficient size to permit making accurate trouble-rate estimates and comparisons. In other words, do not waste time and money partitioning the trouble data into subpopulations so small that the individual data are insufficient to yield accurate and, therefore, useful results.

In many studies it is important to determine precisely when in the life of the equipment each trouble occurs. In those cases deciding when the lifetime of a product starts (so-called "zero time") is of crucial importance. This is particularly true when early life failure rates are to be estimated. For example, does lifetime begin when units arrive, are inspected, are installed, or first operated? Dead-on-arrivals may show up as defective initially or later in time, depending on the type of failure, its effect on the system, the extent of failure detection, and the procedure for collecting the data.

Electronic hardware frequently exhibits a decreasing failure rate

during its early life. Here, failures tend to occur closer together during the early weeks of operation. Therefore, depending upon the "zero time" definition, much of the study's most useful data can be lost or misclassified. Particular care is required in defining zero time if units enter the study at different times, are turned on and off for testing, or are moved to different locations.

To relate a real-life incident, one of the authors was recently asked to analyze some data from a study where the objective was failure-rate estimation after six months of operation. But the records gave only the date of installation and failure. Plotting failures against time gave very strange results, solely because these units were turned on only intermittently and no record of actual operating time on each unit was available. In this case the ability to analyze important time-related failure characteristics was lost because of insufficient detail in the data collected.

Detailed data on the "nature" of troubles occurring during any study generally fall in one of two categories. The first category includes a description of the trouble symptoms, the particular portion or component of equipment in which the trouble was observed, and results of any detailed failure mode analyses performed on the failed components. The second category of detailed trouble information includes data on the particular circumstances or environmental conditions associated with any trouble. Whether equipment was observed to be initially defective or to fail in-service and usage conditions are examples of this second category. Below, we have listed some of the detailed items that may be included on the nature of subpopulations:

- (i) Product vintage (date of manufacture or repair)
- (ii) Source (new, repair, etc.)
- (iii) Length in service
- (iv) Issue, series number, or other product code identifiers.

Like the subpopulation information, the level of detail required on the nature of troubles can have a profound effect on the data collection process, including who will be involved in that process. We have listed the trouble types as follows:

- (i) Component or equipment subcode
- (ii) Trouble symptoms
- (iii) Repair analysis results
- (iv) Component failure mode analysis results
- (v) Precise time of failure.

Obtaining data on failure-mode analyses, for example, may require the participation of technical organizations not directly involved in the field tracking itself. This, in turn, imposes additional requirements on the flow of hardware and paper (trouble tickets, analysis results, etc.) for a given study. At the end of this section we will illustrate some of

these ideas with examples from recently conducted tracking studies. Now, we turn to a closer examination of the question, “How will the data be collected?”

2.2.2 How will the data be collected?

There are as many answers to this question as there are products to be studied. Our aim in this paper, therefore, is to identify goals and procedures common to all or most field-tracking situations.

Probably the best way to start this discussion is the same way it is best to start planning a data collection process—by identifying existing procedures for recording, collecting, and storing information on the field performance of the product under study. It is a rare product on which no information is recorded in the field or at a repair center. Planning data collection should ideally be viewed as a process of either supplementing or tailoring existing data sources to suit the needs of a particular FTS.

At this point it would be helpful to distinguish between data collection carried out in the field (i.e., where the product under study is used), and that carried out in repair locations, and to discuss each separately.

In most tracking studies, the collection of field failure data involves the use of a trouble ticket that must be completed by people responsible for maintaining the equipment under study. As noted, completion of existing trouble tickets is frequently a part of the regular maintenance routine, and substitution of a more detailed study ticket, or “piggybacking” of the study ticket on an existing form, is preferable to burdening maintenance people with a new and separate piece of paper. Whether or not a separate or modified existing form is used, there are a number of basic rules that govern the design of trouble tickets. First, the tickets should be kept as short and as simple as possible. Those are the obvious rules. Less obvious, but equally important, are the following: Wherever possible, the trouble tickets should be formatted in “modular” fashion, with separate sections devoted to different types of information—e.g., time and place of the trouble in one section, equipment description in another, trouble description in still another. The most frequently used modules should appear first and most prominently; less frequently used modules should appear later. The trouble ticket used in the station set Product Performance Survey (PPS) (Fig. 1) illustrates these ideas. The top of the ticket gives information on when and where a trouble occurred. That information is required for each trouble report. Next comes information on the nature of the trouble, also needed for each event. Data on the type of set or component involved in the trouble come next; however, these data are not needed if the equipment in question is returned with the

BELL SYSTEM	
PRODUCT PERFORMANCE SURVEY	
DATE _____	CRAFT I.D. _____
PHONE NO. _____	EXT. _____
TROUBLE CATEGORY	
WHEN DID APPARATUS FAIL? (CHECK ONE)	
<input type="checkbox"/> INITIALLY <input type="checkbox"/> IN-SERVICE	
TROUBLE REPORT _____	
CHECK IF ACTION DUE TO:	
<input type="checkbox"/> PREVENTIVE MAINTENANCE/ROUTINE	
<input type="checkbox"/> CUSTOMER DAMAGE	<input type="checkbox"/> LIGHTNING
<input type="checkbox"/> SHIPPING DAMAGE	<input type="checkbox"/> SUSPICION
IF COMPONENT REPLACED OR ADJUSTED — COMPLETE	
SET CODE _____	SET DATE _____
<input type="checkbox"/> C-STOCK/REISSUED	<input type="checkbox"/> NEW
<input type="checkbox"/> RAPID RECOVERY	<input type="checkbox"/> TELCO TURN-A-RND
IF ADJUSTMENT — COMPLETE	
COMPONENT CODE _____	COMP. DATE _____
ADJUSTMENT DESCRIPTION _____	
IF COIN APPARATUS — CHECK ONE:	
<input type="checkbox"/> ROTARY DIAL	<input type="checkbox"/> TOUCHTONE DIAL
OTHER COMMENTS MAY BE PUT ON TAG BACK	

Fig. 1—Station set Product Performance Survey trouble ticket.

trouble ticket. Finally, the last section of the ticket describes field adjustments, used only in those few cases where no hardware is returned along with the ticket.

As this last discussion of the station-set PPS implies, there is more to field data collection than the gathering of trouble tickets; there is frequently the gathering of failed hardware as well. The design of an effective, integrated hardware/trouble ticket data-flow system is as important as the design of the trouble ticket itself. The basic objectives of the data-flow system are:

- (i) To ensure that each piece of returned hardware reaches the designated repair or diagnostic location and, in many cases, the designated individual responsible for hardware analyses in the study; and
- (ii) To ensure that the information on the trouble tickets reaches the organization responsible for storing and analyzing the trouble data.

There are other important objectives, as well, primarily related to

assuring compliance with study procedures and ensuring that hardware analysis results may be uniquely identified with reported trouble events. We will discuss the issue of compliance later. The ability to associate hardware analysis results with trouble symptom reporting is important in tracing down the causes of No Trouble Found (NTF) returns (e.g., diagnostics problems). The use of serialized, multipart tickets is the prime vehicle for making such associations and will be illustrated below.

We have already noted that the burden imposed by an RTS on field personnel can be minimized by using existing reporting forms, whenever possible. For some products, the burden can be even further reduced by exploiting automatic data collection procedures. We include in this category fully automatic data collection, such as that associated with accessing maintenance channel output of software-controlled equipment, and semiautomatic data collection, such as that associated with accessing computerized administrative data on customer trouble reports where the initial entry of the data into the data base depends on action by customers or field personnel. Access of existing data sources such as these has become an increasingly prominent mode of data collection in field-tracking studies. Access of repair location data bases serves an analogous function for hardware-repair analysis data.

2.2.3 In what study population will the data be collected?

In choosing the study population it is important to explicitly define the limits of the inferences to be made from the study. Are the results to be applied to all units, all units made in a given period or under given conditions, or used in a particular fashion, etc.? If the members of the study population received special care, were hand-made, produced at one plant, etc., then conclusions beyond these boundaries depend upon engineering judgment more than upon statistical inference. Confidence intervals reflect variability only in the population actually sampled and not from other sources. For example, increasing the sample taken in one operating area gives no information regarding inter-area differences. When sampling is performed by first selecting K operating areas and then sampling only within these, the formulas appropriate are those used in cluster sampling.³ Here, the intra-area and inter-area variability are separated. Of course, looking at inter-area differences in detail can indicate important variables (maintenance procedures, environmental impact, etc.) that could be the focus of a follow-up study. Care must be taken before cause and effect relationships are assumed because of the multitudes of possible causes and interrelationships. As Cox relates:⁴

“If we wish to apply the conclusions to new conditions or units, some additional uncertainty is involved over and above the uncertainty measured by the standard error. The only exception . . . is when the units . . . are chosen from a well-defined population of units by a proper sampling procedure.”

And later,

“. . . it is important to recognize explicitly what are the restrictions on the conclusions of any particular experiment.”

In any tracking study there is a trade-off between more detailed conclusions regarding a smaller population and less detailed conclusions about a larger one. For example, a study may be aimed at determining whether a change in design has improved reliability in systems subject to certain load characteristics, or whether an overall reliability increase independent of load has occurred. A careful statement of objectives will greatly assist resolving such questions.

Once a population of interest has been defined and agreed upon, technical sampling questions can be addressed. There are certain population characteristics that require special attention. For example, if a small proportion of the units contribute a large proportion of the events under study, stratification and other specialized techniques may be required. Also, considerable gains in efficiency can sometimes be realized by the use of ratio or regression estimates. Here, known characteristics of products or systems under study are related to the characteristics of interest in the study.

2.2.4 How much data will be collected: sample size considerations

Selecting the appropriate number of units to be included in an FTS is very important. On the one hand, a sample size that is too large may add unnecessary expense to the study. On the other hand, a sample size that is too small may mean that any statistical test using study data may lack sufficient power to draw meaningful conclusions. Several authors^{5,6,7} have addressed this problem. Reference 5 took the theory of Refs. 7 and 8 and transformed it into usable curves; these curves will be discussed in general in this section and in detail in the appendix.

The parameters of interest in a field study are summarized in Table I. In cases A and D a sample size will be chosen to control the precision of the estimates within certain bounds. In the remaining cases

Table I—Parameters of interest in a field study

	Proportion	Rate
Estimating one parameter	Case A	Case D
Testing hypothesis about one parameter	Case B	Case E
Comparing two parameters	Case C	Case F

the sample size will be chosen to control the probability of making incorrect conclusions. If we assume that failures associated with a proportion occur according to a binomial model and that failures associated with a rate occur according to a Poisson model, it is possible to develop excellent sample sizing guidelines for each of the cases A through F. (A discussion of model selection and use is included in the next section.) Each case is discussed in detail, with examples, in the appendix.

2.3 Planning for successful data analysis

In this section, we consider both the data analysis, itself, and the data storage and retrieval procedures that make the analysis possible.

2.3.1 Model building and data analysis

It requires no lengthy argument to establish that the payoff from any field study comes only with the successful analysis of the data from that study. And in a very real sense, all of the detailed planning on data collection is aimed at ensuring that at the conclusion of the study it will be possible to carry out all of the data analyses appropriate to the study objectives.

In broad terms, there are three things that generally get done with field-tracking data. These are:

- (i) Estimating trouble or replacement rates, including the construction of confidence intervals, where appropriate and practical;
- (ii) Searching the data for anomalies—equipment types or vintages that stand out, or trouble causes that stand out; and
- (iii) Making comparisons of product performance among different types, or vintages, of equipment.

Each of these procedures requires careful planning and a close linkage between the setting of objectives, the design of the data collection process, and the data analysis itself.

During both planning and implementation of a study, the mechanism by which the study objectives, the actual data collection, and the data analysis are linked is the statistical “data model.” It is through the data model that the nondeterministic (stochastic) nature of the data is described, and through the model that statistical inferences on the questions of interest to the study are made.

As noted above, most field-tracking studies concern themselves with counts of events (failures, replacements, etc.). It is for this reason that the simplest and most frequently used models in field studies are the binomial and Poisson models.

The binomial model relates the number of events of interest (failures, say), X , to the total number of “trials” (opportunities for failure), N , through the expression:

$$\text{Probability } [X = k] = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k}, \quad k = 0, 1, \dots, N,$$

where p is the probability of a failure on an individual trial.

The Poisson model relates the number of events of interest, X , to the total amount of time during which those events can have occurred, t , through the expression:

$$\text{Probability } [X = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad k = 0, 1, \dots,$$

where λ is the rate at which the events occur in time.

Both models assume a uniform probability or intensity of occurrences—from trial-to-trial for the binomial, over time for the Poisson. For studies in which a model allowing for changing failure intensity seems appropriate (e.g., studies of equipment that may be subject to infant mortality), other models such as the Weibull and lognormal are commonly employed. Detailed information on the form and use of these models may be found in any one of several statistical/reliability texts⁹ and we will not attempt to describe them here.

None of the models mentioned thus far is equipped to handle data collected under changing study conditions (e.g., changing environment, age, study locations, etc.), or so-called “nuisance factors.”

To illustrate the problem of nuisance factors, suppose we wanted to compare the replacement of two types of equipment (called “old” and “new”), from a study in which the “old” equipment was observed, in one study location, while the “new” equipment was observed in that and other study locations. Here, the factor of interest is equipment type (old versus new); the nuisance factor is the difference that may exist between study locations, which could bias the comparison between the old and new equipment. It is at this point that the use of relatively sophisticated data-analytic techniques, employing tools such as the well-known linear (or log-linear) model, becomes necessary and worthwhile. These techniques allow for separating the effects (on replacement rates, for example) caused by equipment differences, study location differences, etc. and for getting at the factors of interest without ignoring potential biases introduced by the presence of nuisance factors. The use of linear models is well documented in both the statistical and engineering literatures.^{10,11} However, when confronted with an apparent need to make use of such techniques, the study designer and data analyst should seek the assistance of a statistician who is thoroughly familiar with the application of these techniques.

The use of any of the models mentioned above involves making some assumptions about the data. For example, as noted, use of the binomial or Poisson models assumes a uniform failure probability or

intensity. Use of a linear model generally involves some assumptions of independence between the way in which different factors affect the probability of equipment failure. If those assumptions are violated, the resulting data analysis can be invalid and, worse, misleading. For example, if the failure intensity changes with time (age) for a given type of equipment, use of the Poisson model in analyzing the data on that equipment could easily mask important information on both the short- and long-term reliability of the equipment. Invalid assumptions concerning the independence of various factors employed in a linear-model can mask or falsely create the impression of cause-and-effect relationships between various factors and the probability of failure. Rather than attempt to catalog all of the field-study conditions and assumptions associated with the use of any particular model, we will give some general guidelines on the choice and use of models in field-tracking studies.

Probably the simplest but most important rule to use in choosing a FTS model is “keep it simple.” The more complicated a model is, the more parameters it will use that must be estimated during the data analysis, and the more assumptions it will require to make that analysis valid. As this last discussion implies, there are two additional rules that are closely related to the simplicity rule:

(i) *Estimability*—Since data analysis, at its core, involves making statistical inferences about parameters in the model from the available data, it is essential that the model and the collection process be matched to ensure that the right data are available in sufficient quantities to make inferences about all the parameters of interest. This is a point we have already touched on in the discussion on data collection.

(ii) *Verifiability*—The assumptions implicit in the use of any model must be verifiable or the results of the FTS will remain open to doubt. In some cases, engineering judgment can be used to justify certain model assumptions. In all cases, every effort must be made to verify assumptions from the data—either during a procedural trial (see Section 3.2 below), or as the first step in the data analysis stage of the study. A wide variety of statistical techniques are available for testing the uniformity and independence assumptions typically encountered in FTS model use; these techniques should be applied with the advice and assistance of a trained statistician.

In summary, successful data analysis is dependent on the choice of an appropriate FTS model that is matched to both the actual study conditions and to the data collection procedures employed in the study.

2.3.2 Data storage and retrieval

With the exception of very small-scale studies, involving perhaps

fewer than 100 trouble events in all, computerized data storage is a great asset—if not a necessity—in permitting complete and timely analyses of field-tracking data. There are a number of data systems available [for example Data Management System (DMS)*, RAMIS®†, etc.] that lend themselves to constructing field-tracking data bases. Among the factors that must be considered are total eventual size, frequency of access required, and most important, flexibility of access—i.e., flexibility in retrieving and summarizing the data by one or more characteristics, such as equipment type or vintage, or type of trouble. It would be very difficult, for example, to compare the performances of different vintages of a given product if the data retrieval system did not permit easy, separate access to the trouble data for each vintage. On the other hand, it is important not to confuse a need for flexible data access with a need for an elaborate data retrieval system that turns out regular, detailed data summaries that display results in every conceivable way. The key is to retain flexibility without trying to preprogram every possible way of looking at the data.

III. FTS IMPLEMENTATION

In this section we briefly consider several topics in the actual implementation of an FTS:

- (i) Developing procedures and training personnel
- (ii) Assuring compliance with study procedures
- (iii) Conducting a procedural trial.

3.1 Developing procedures and training personnel

Based on mutually agreed-upon objectives, specific procedures and forms for data collection need to be developed. Determining the extent of automatic data retrieval, checking the validity of the inputs, deciding exactly what data are necessary, etc., are detailed questions that require resolution.

Unless rules are provided to meet contingencies, people tend to either make up their own rules or just get discouraged about participation in the study. Although all possibilities cannot be provided for, care should be taken to anticipate the most common “unusual” events. As a default, a space for “additional comments” or “other” on data forms will alert the data analyst to the fact that the specified categories were ambiguous, not mutually exclusive, not exhaustive, etc.

The training of the field personnel who will actually perform the data collection is a very important step. Hands-on teaching with real situations will prepare them for being on their own. Giving them an

* Data Management System, developed by Bell Laboratories.

† RAMIS is a trademark of Mathematica, Inc.

indication of the reasons for the study and how important their participation is can improve their morale and impact on the quality of the data collected. A specific procedure to provide continuing contact and periodic feedback of results can also be a strong positive stimulus.

3.2 Compliance

It is difficult to overemphasize the importance of monitoring compliance with tracking-study procedures. The basic output of any FTS is a measure of the reliability of the equipment under study. In order for that measure to be useful and unbiased (by differences in the completeness of reporting for different products, trouble types, etc.), all or substantially all of the trouble events experienced by the equipment must be reported. It is the function of compliance procedures to ensure that this is the case.

Basically, compliance can be checked in one of two ways. If an independent (of the FTS) count of trouble events for the equipment under study is available, compliance can be checked by comparing that count to the number of troubles reported through the study procedure. This method is used in the station set PPS, where administrative counts of customer trouble reports serve as the independent count of station troubles in any PPS study location. If no such count is available, but the equipment under study is located in a geographically small, reasonably well controlled setting, such as a central office, serializing of the equipment under study and periodic mapping of the office inventory—when compared to the reported troubles—can serve as an effective compliance check. With either procedure, the key to maintaining good compliance is fast feedback to the people responsible for providing the field data and their management about the degree to which study procedures are being followed. It is for this reason, principally, that some identity of the field person reporting the trouble is included on most field-tracking study tickets.

As noted earlier, in addition to field data collection, many tracking studies involve the collection of data—usually from failure-mode analyses—at repair locations and/or diagnostic laboratories. Reporting forms for such analyses will usually have to be tailored to the particular equipment under study. But some of the general principles that govern field data collection apply to the hardware failure analysis data as well. The flow of hardware and paper must be designed to ensure that (i) each piece of hardware returned can be accounted for and checked off against reported field troubles, and (ii) individual hardware analyses can be associated with reported field trouble symptoms.

3.3 Procedural Trial

Once study procedures and forms have, at least tentatively, been

developed, a trial is an excellent way to shake out unexpected problems. Here, an attempt is made to collect some actual data by people who will participate in the real study. Estimates of speed and accuracy of filling out forms, difficulties with interpreting procedures when faced with real situations, completeness of instructions, and potential usefulness of results are some of the possible outputs. If extensive revision of procedures, forms, etc., are required, a second trial may be necessary.

In addition to testing the data collection portion of the study, a trial of the data analysis methodology should also be made with simulated or actual data. It is useful to present possible conclusions, with their justification, to the users of the study results. Then, a comparison of their subjective impressions from the raw data with the quantitative results from the statistical analysis can be used to improve both. It is also at this point that model assumptions are to be verified or modified as needed.

IV. ILLUSTRATIONS

In this section, we briefly describe some recent field-tracking studies. Perhaps the longest running study is the Product Performance Survey on station sets, which we mentioned earlier in this paper. Figure 2 shows the flow of hardware and data in that study. The trouble ticket is shown in Fig. 1. Note the modularized design of the ticket described above. Analysis of returned equipment in this study is carried out by analysts in the Western Electric Quality Assurance organization who are dedicated to the study. These analysts encode the results of their analyses, as well as other information on the trouble tickets that accompany the returned hardware, for direct entry into a data base. Compliance is monitored by comparing the number of PPS trouble ticket returns to the total number of trouble reports tracked by administrative reporting systems in each study location.

A second example is illustrated in Figs. 3 and 4, which are the data-reporting form and a flow sheet, respectively, for the FTS of Northern

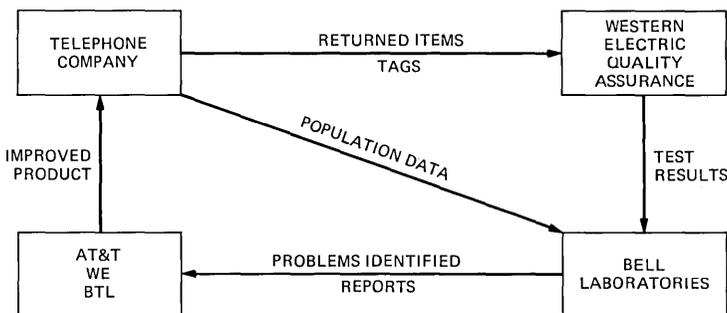
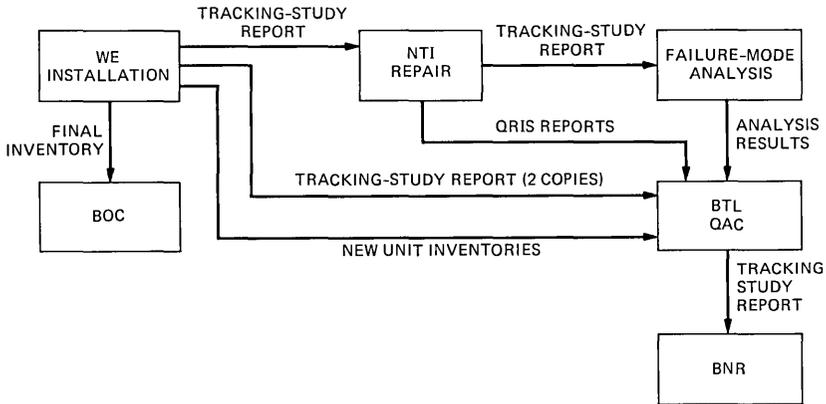


Fig. 2—Product Performance Survey data flow diagram.

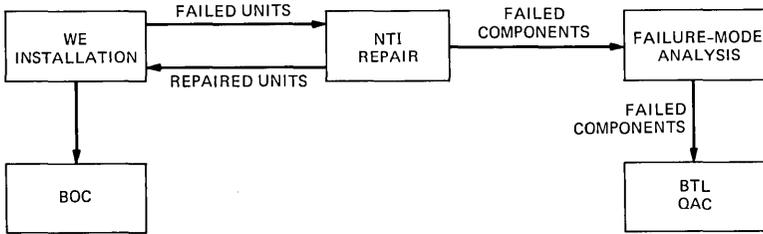
DMS-10 Tracking Study Report		AT&T/ Bell-Northern Research	
		M M D D Y Y	HR Min
Office _____	Start of Trouble/Activity	<input type="text"/>	<input type="text"/>
Report No. _____	Maintenance Org. Called ?	<input type="text"/>	<input type="text"/>
Generic Issue _____	Maintenance Org. Arrived ?	<input type="text"/>	<input type="text"/>
Craftsman _____	Trouble Cleared ?	<input type="text"/>	<input type="text"/>
REASON FOR REPORT			
A <input type="checkbox"/> Trouble B <input type="checkbox"/> Class A Change C <input type="checkbox"/> Service Disconnect Z <input type="checkbox"/> Other -- Specify			
DESCRIPTION OF TROUBLE			
A <input type="checkbox"/> Total System Outage		B <input type="checkbox"/> Subsystem Outage	
D <input type="checkbox"/> Excessive AUD Messages		E <input type="checkbox"/> Excessive BUG Messages	
G <input type="checkbox"/> Feature Problem (Specify)		H <input type="checkbox"/> CPF Message	
Further Description:		I <input type="checkbox"/> Initially Defective	
		J <input type="checkbox"/> Line/Trunk	
		Z <input type="checkbox"/> Other -- Specify	
MEANS OF TROUBLE DETECTION			
A <input type="checkbox"/> Customer Complaints B <input type="checkbox"/> System C <input type="checkbox"/> Referred In D <input type="checkbox"/> Alarm Given			
RECOVERY			
A <input type="checkbox"/> Automatic or B <input type="checkbox"/> Manual			
C <input type="checkbox"/> SYSLOAD D <input type="checkbox"/> SMART Recovery E <input type="checkbox"/> INITIALIZATION			
SPARING: F <input type="checkbox"/> CPU G <input type="checkbox"/> Memory H <input type="checkbox"/> Network I <input type="checkbox"/> PE Shelf J <input type="checkbox"/> Line CCT K <input type="checkbox"/> None			
HARDWARE MAINTENANCE ACTION			
A <input type="checkbox"/> Unit Reseated B <input type="checkbox"/> Unit Tested And Restored C <input type="checkbox"/> Unit Replaced			
Z <input type="checkbox"/> Other (Specify)			
UNIT IDENTIFICATION			
Unit Location: Slot <input type="text"/> <input type="text"/> Shelf <input type="text"/> <input type="text"/> Bay <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>			
Serial Numbers			
Unit Code & Series		Study No.	
NTI No.			
Reported Unit	<input type="text"/>	<input type="text"/>	<input type="text"/>
Replacement Unit	<input type="text"/>	<input type="text"/>	<input type="text"/>
COMMENTS			
INSTRUCTIONS			
1. Use a ball point pen and press firmly, you are making 4 copies.			
2. Attach TTY printouts and other information (TTY printouts before and after trouble).			
3. Report routing: copy 1-office file; copy 2-BTL/QAC; copy 3 -BNR via BTL/QAC copy 4 - with replaced hardware.			

Fig. 3—DMS-10 Tracking Study Report.

Telecom's DMS-10 switching office. The flow sheet illustrates a point discussed in Section II, namely, that numerous organizations are often involved in an FTS. Cooperative planning among organizations involved played an important role in making this study run smoothly and produce meaningful results. The report form shows a completely different set of data fields and possible responses than did the PPS trouble ticket. Just as trouble tickets are compared with local administrative data in the station set study, report forms for this FTS are



(a)



(b)

Fig. 4—DMS-10 switching system installation tracking study. (a) Routing of information. (b) Routing of study units.

compared with maintenance and outage data automatically collected from the switching machine's maintenance output channel.

V. CONCLUSIONS

In this paper we have discussed several important aspects of planning and conducting an FTS. We have shown how careful planning beforehand in the areas of data collection, population definition, sample size, and stating of objectives is essential. We have also discussed means of ensuring that the study is producing the required ongoing data. If properly planned and conducted, FTSs can and do play a key role in assuring the quality and reliability of telecommunication products.

REFERENCES

1. V. S. Hering and J. A. Hooke, "Product Performance Surveys: Field Tracking for Station Sets," Bell Lab Rec. (December 1978), special Quality Assurance issue.

2. W. A. Strong, unpublished work.
3. M. Hansen, W. Horowitz, and W. Madow, *Sample Survey Methods and Theory*, New York: John Wiley and Sons, Inc., 1953.
4. D. R. Cox, *Planning of Experiments*, New York: John Wiley and Sons, Inc., 1958.
5. G. G. Brush, unpublished work.
6. L. G. Lyon, unpublished work.
7. D. W. Swann, unpublished work.
8. A. Mace, *Sample Size Determination*, Melbourne, FL: Krieger Publishing Company, 1974.
9. J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, New York: John Wiley and Sons, Inc., 1982.
10. N. Draper and H. Smith, *Applied Regression Analysis*, New York: John Wiley and Sons, Inc., 1981.
11. J. Kalbfleisch, and R. Prentice, *The Statistical Analysis of Failure Time Data*, New York: John Wiley and Sons, Inc., 1980.
12. A. Duncan, *Quality Control and Industrial Statistics*, Homewood, IL: R. D. Irwin Inc., 1965.

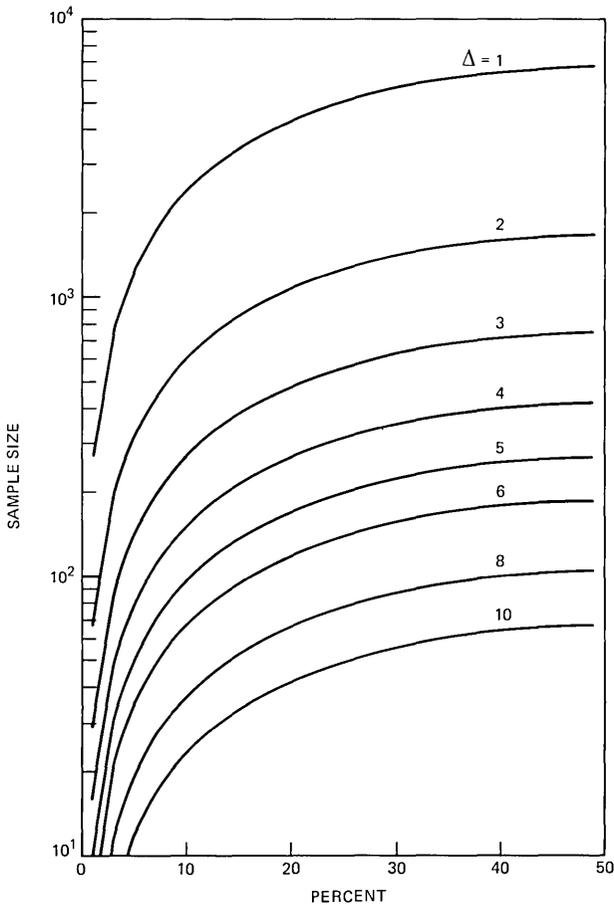


Fig. 5—Minimum sample sizes needed to generate 90-percent confidence intervals.

APPENDIX

Sample Size Selection

In this appendix we discuss in detail the six cases of sample size selection described in Section 2.2.4 of this article. These cases are:

- (i) Estimating a parameter
- (ii) Testing a hypothesis about one parameter
- (iii) Comparing two parameters for both proportion and rates.

Each case is discussed in turn below. The six cases are shown in Table I, Section 2.2.4.

A.1 Case A

In Case A we wish to have a sample size to control the precision of the estimate of a percentage within certain bounds. The estimation

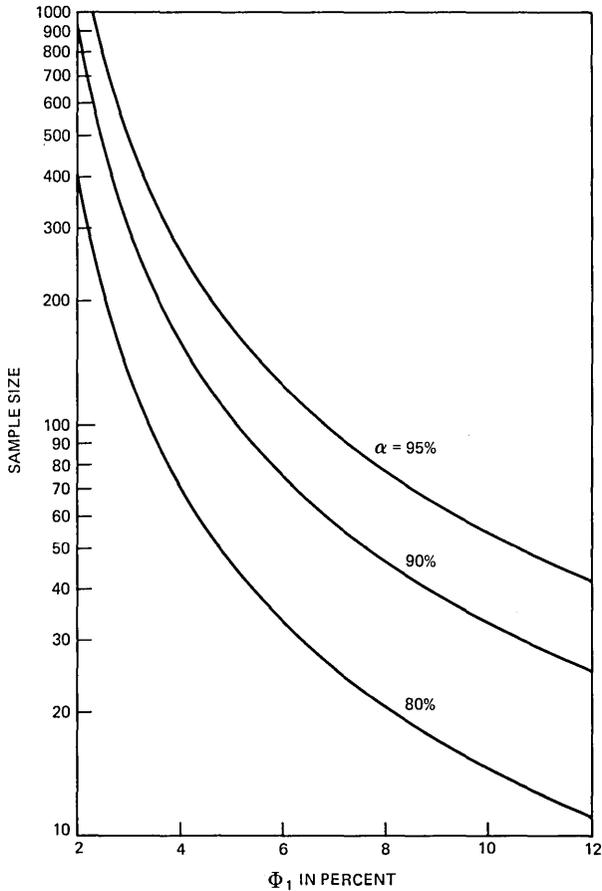


Fig. 6—Minimum sample sizes for $\Phi_0 = 1$ percent.

process is subject to imprecision; therefore, it is customary to express the estimate as an interval, say 2 to 6 percent, as opposed to a single point, say 4 percent. This interval is chosen so that if we were to repeat the process of data collection and interval construction, our intervals would cover the true, unknown percentage a very large proportion of the time. The shorter the interval, the more precise is our estimate. This interval will decrease in width as the sample size increases. We will then select the sample size before the FTS to obtain an anticipated width for our interval after the FTS. Figure 5 shows sample sizes necessary to generate 90-percent confidence intervals which are 2Δ wide. The sample size depends on the true percentage. The maximum sample size is required when the true percentage is 50 percent.

Example of Case A: Suppose we are only interested in estimating the percentage of units that are initially defective. We think that this percentage is less than 15 percent, and we want the estimated interval

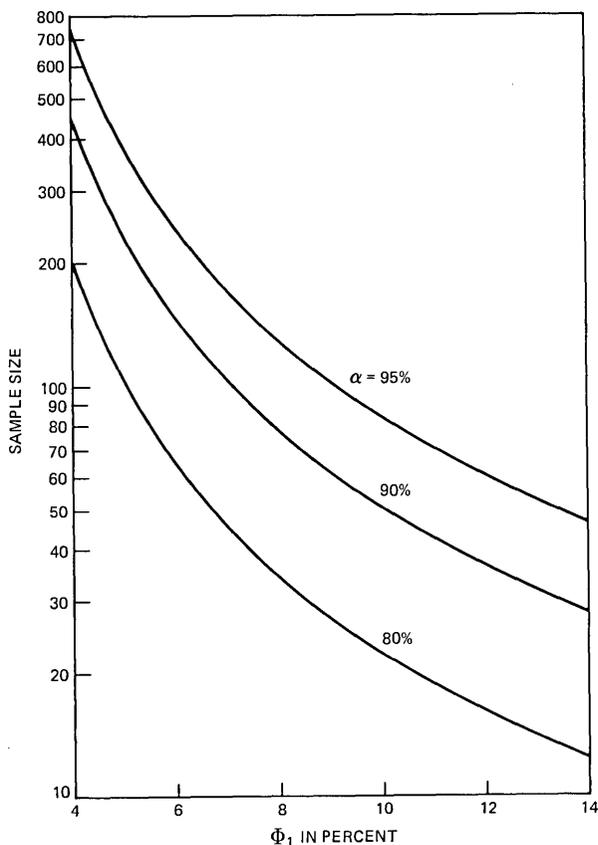


Fig. 7—Minimum sample sizes for $\Phi_0 = 2$ percent.

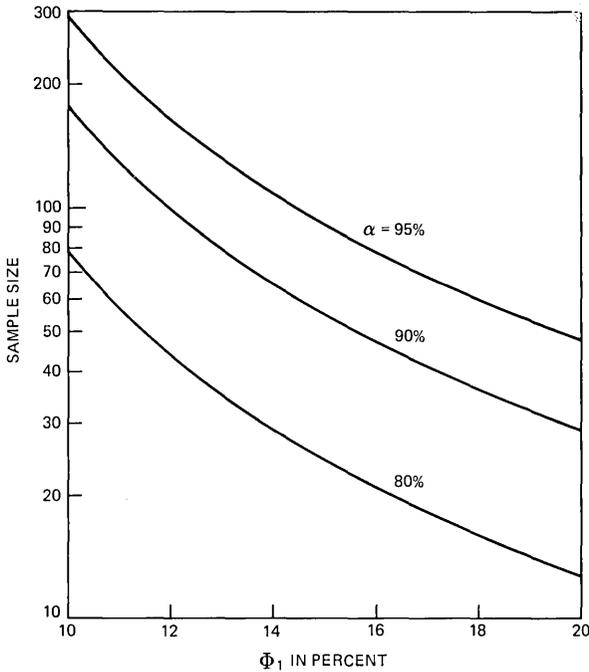


Fig. 8—Minimum sample sizes for $\Phi_0 = 5$ percent.

to be at most 6-percent wide. Therefore, $\Delta = 3$ and we see in Fig. 5 that a sample of size 400 is required. If we had no idea as to the true percentage we would use the maximum sample size for 50 percent, that is, 750. Note that the curves are symmetrical about 50 percent.

A.2 Case B

In Case B we wish to test the hypothesis that a proportion is less than or equal to Φ_0 . We will look at a sample of n units, and make one of the two decisions:

(i) If we see that a number of units less than or equal to c , the “acceptance number”, have the trait associated with the proportion, then we will accept the hypothesis that the proportion is less than or equal to Φ_0 .

(ii) If we see that more than c of the units have the trait, then we will reject the hypothesis in favor of the alternative that the proportion is greater than Φ_0 .

We wish to structure the test so that if the true value of the proportion is Φ_0 , we will make decision i a large portion of the time, and if the true value of the proportion is Φ_1 , we will make decision ii a large portion of the time. The reader more interested in acceptance

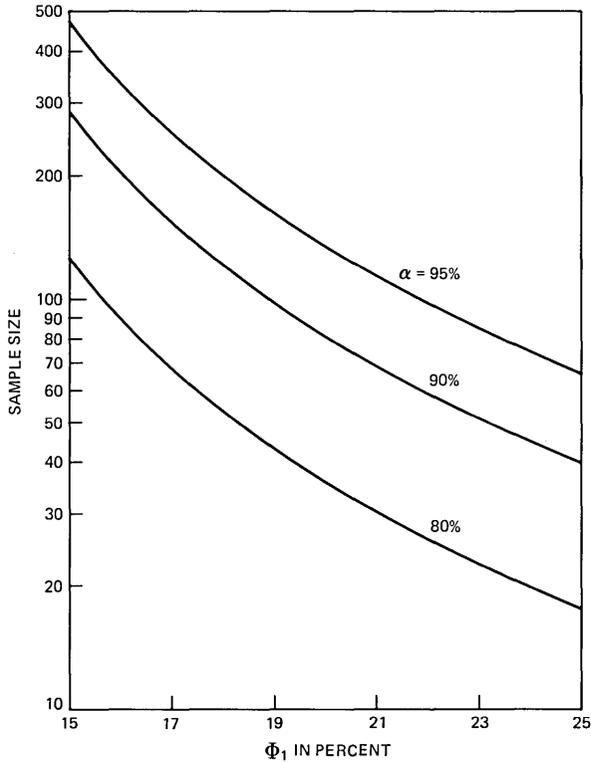


Fig. 9—Minimum sample sizes for $\Phi_0 = 10$ percent.

sampling plans, which is an example of such a situation, should refer to a specialized reference, e.g., Ref. 12.

Figures 6 through 9 show the required sample size for values of $\Phi_0 = 1, 2, 5,$ and 10 percent for 80% , 90% , and 95% confidence levels. As an example of the use of the curves, let $\Phi_0 = 1$ and $\Phi_1 = 5$ percent. We see in Fig. 6 that for a 90% confidence level, a sample size of 100 is needed.

A.3 Case C

This case deals with comparing two percentages, call them percentage *A* and percentage *B*. These percentages might be similar characteristics on competing products, or competing designs. For example, we might be interested in percentages of circuit packs that are dead-on-arrival from two suppliers. After the FTS we may arrive at one of three conclusions:

- (i) The two percentages are not significantly different
- (ii) Percentage *A* is larger than percentage *B*

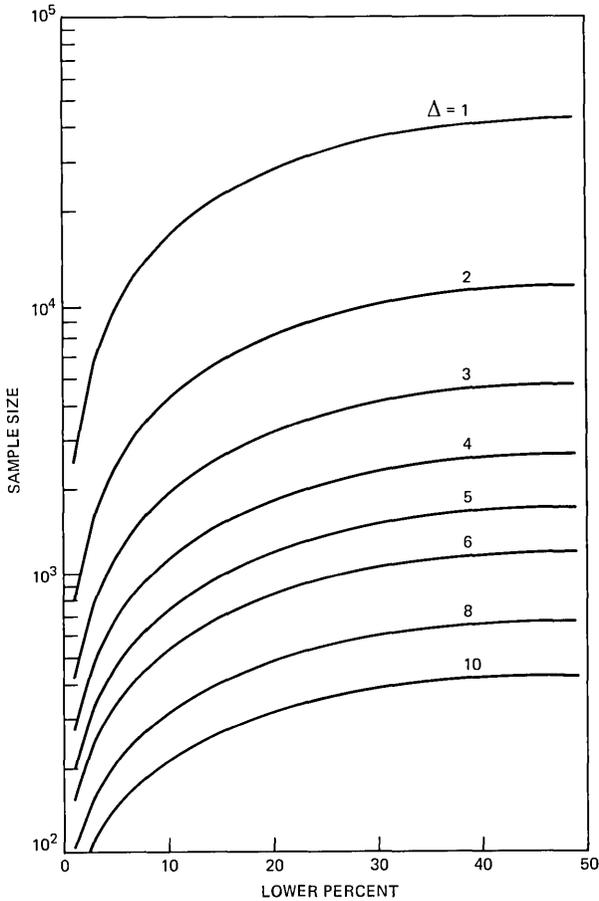


Fig. 10—Minimum sample sizes for comparing two proportions at the 90-percent confidence level.

(iii) Percentage B is larger than percentage A.

There are certain risks in arriving at incorrect conclusions. The risks decrease with increasing sample size. We wish to control, at a low level, the risk of not making conclusion (i) when percentages A and B are equal. And we wish to control, at a low level, the risk of not making conclusion (ii) when percentage A is Δ larger than percentage B [or, similarly, the risk of not making conclusion (iii) when percentage B is Δ larger than percentage A]. Figure 10 gives sample sizes necessary to accomplish this at the 90-percent confidence level.

Example of Case C: Suppose we wish to compare the percentages of plug-in units (from two suppliers) that fail during the warranty period. Further, we assume that the lower percentage will be less than 20 percent. We wish to have a high probability of concluding that the

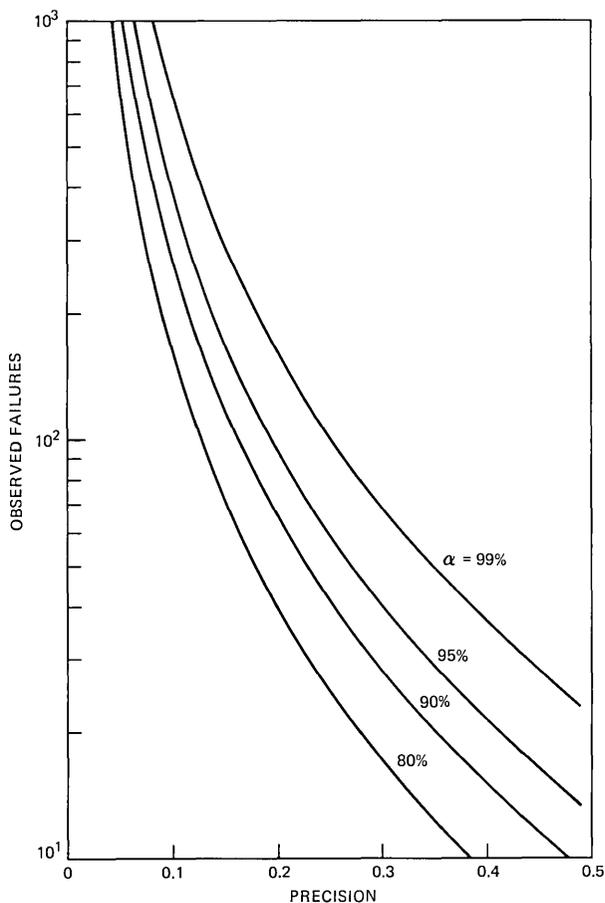


Fig. 11—Minimum observed failures for estimating a failure rate.

upper percentage is greater than the lower percentage when the upper percentage is 5 greater than the lower. For $\Delta = 5$ and a lower percent of 20, we need to look at 1300 units from each supplier. With no knowledge of the true percentages we would use the sample size for 50 percent, that is, 1700.

A.4 Case D

Cases D, E, and F deal with failure rates, as opposed to the percentages of Cases A, B, and C. (The results for Cases D, E, and F must be used subject to the cautions given at the end of this appendix.) Cases D, E, and F require the use of two curves. The first curve will tell us how many failures we need to see. The second curve will tell us how many units must be included in the FTS so that we are reasonably

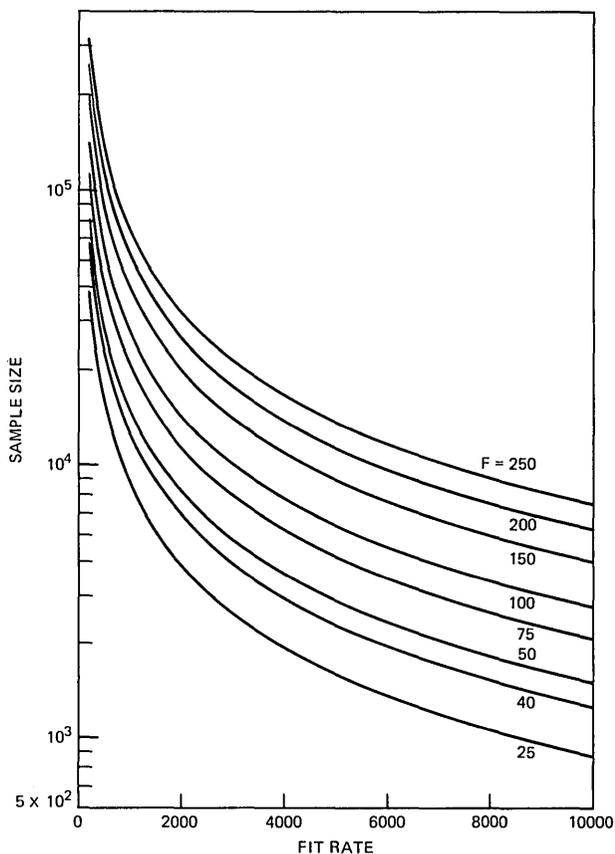


Fig. 12—Minimum sample sizes for failure-rate estimation (6-month interval).

certain that the failures occur in a prescribed time period. In Case A we measure the precision of our estimation by the width of the interval, expressed in *absolute* percentages. In Case D, we will measure the precision in terms of *relative* percentages. For example, if our interval is 1500 FITS* \pm 5 percent = 1500 \pm 75 FITS = (1425, 1575), then we will say that the precision is 5 percent. This interval corresponds to (1.25, 1.38) failures per 100 sets per year.

Example of Case D: Suppose that we wish to obtain a precision of 15 percent at the 90-percent confidence level in the estimate of the failure rate of a plug-in unit. In Fig. 11 at an abscissa of 0.15 (15 percent) we see that 120 failures must be observed. Suppose that the FTS is to last 12 months and that our reliability prediction gives us an

* FIT = Failures in 10^9 hours = 8.75×10^{-4} failures per 100 units per year.

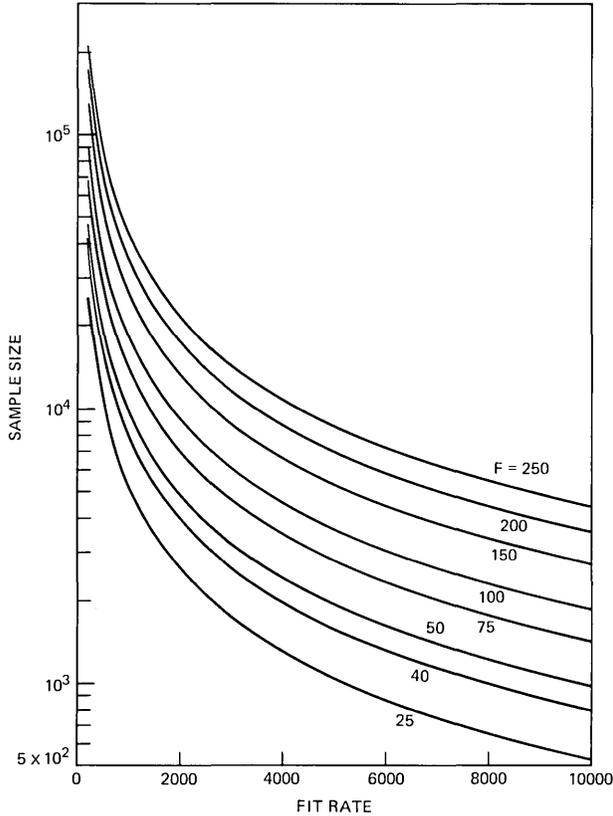


Fig. 13—Minimum sample sizes for failure-rate estimation (9-month interval).

estimated FIT rate of 2500. In Fig. 14 we see that about 7000 units need to be included in the study.

Figures 12 through 15 give required sample sizes for studies of lengths 6, 9, 12, and 18 months, and for FIT rates up to 10,000. If some other combination is needed, then the following formula should be used:

$$N = \frac{F + 1.645 \times \sqrt{F}}{(7.2 \times 10^{-7})\lambda T} + \frac{F}{2}, \quad (1)$$

where

F is the number of failures,

λ is the prior estimate of the failure rate in FITs (failures in 10^9 hours),

T is the number of months the study will last, and

N is the required sample size.

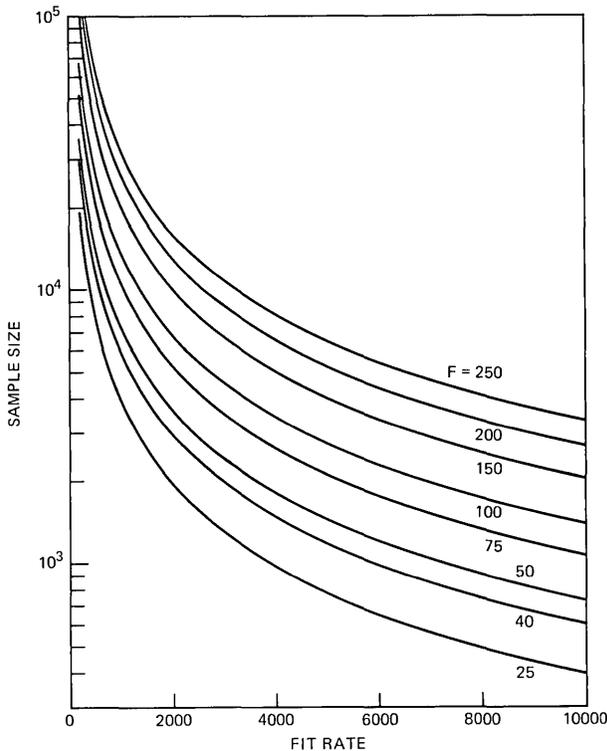


Fig. 14—Minimum sample sizes for failure-rate estimation (12-month interval).

This formula provides 95-percent confidence that the required number of failures will be observed.

A.5 Case E

In Case E we wish to test the hypothesis that a rate is less than or equal to a specified value, V_1 . Based upon the data observed, we will either

(i) Accept the hypothesis that the rate is less than or equal to V_1 , or

(ii) Reject the above hypothesis in favor of the alternative that the failure rate is greater than V_1 .

We wish to structure the test so that if the true value of the rate is V_1 , we make decision *i* with a high probability and if the true value of the rate is $(R)V_1$, we make decision *ii* with a high probability.

Example of Case E: Suppose we wish to check to see how a newly designed part has changed the reliability of a piece of equipment. We are satisfied with $R = 2$ and the 90-percent confidence level. Figure 16 shows that 15 failures must be observed.

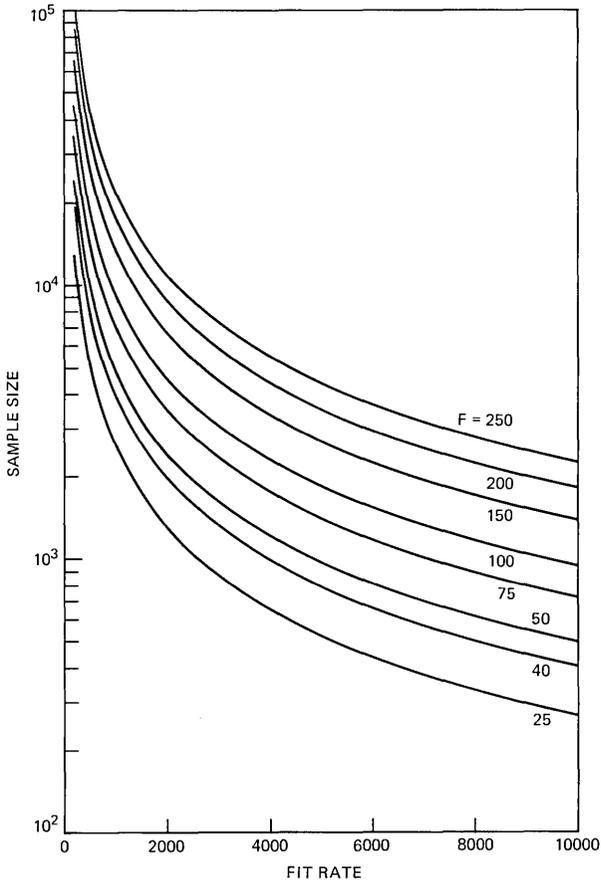


Fig. 15—Minimum sample sizes for failure-rate estimation (18-month interval).

A.6 Case F

Here we wish to compare failure rates of two competing products. At the end of the FTS we can arrive at one of three conclusions:

- (i) Failure rate A and failure rate B are not significantly different
- (ii) Failure rate A is larger than failure rate B
- (iii) Failure rate B is larger than failure rate A

Again there are risks of arriving at incorrect decisions. As we increase the sample sizes, we can decrease these risks. We wish to control, at a low level, the risk of not making conclusion (i) when failure rates A and B are equal. And we wish to control, at a low level, the risk of not making conclusion (ii) when failure rate A is R times as large as failure rate B (or similarly the risk of not making conclusion (iii) when failure rate B is R times as large as failure rate A).

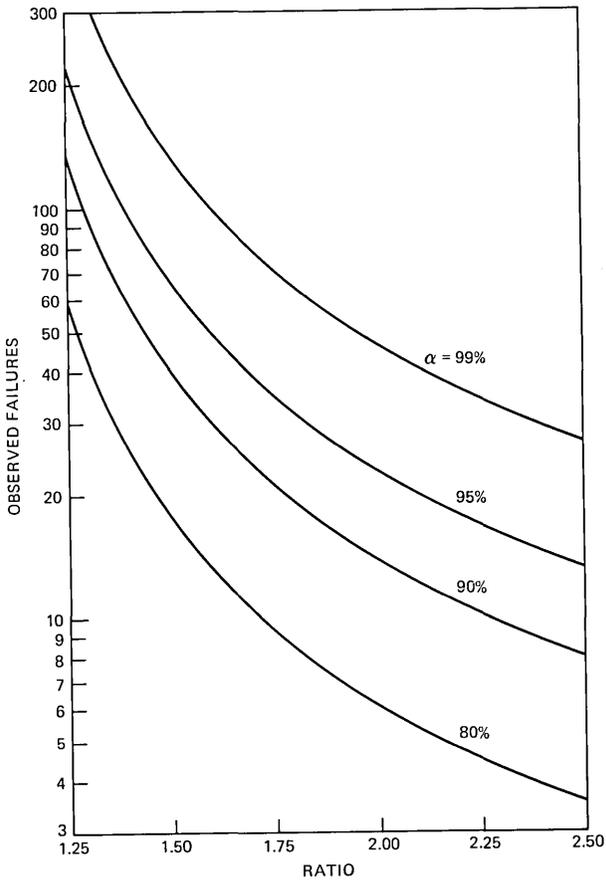


Fig. 16—Minimum observed failures for a hypothesis test (one failure rate).

Example of Case F: Suppose we wish to use the FTS to compare the failure rates of the channel units of two different suppliers. Suppose further that we wish to have a high chance (90-percent probability) of concluding that the larger failure rate is larger than the smaller failure rate when indeed the larger failure rate is twice the smaller. In Fig. 17 we see that we need to observe about 36 failures. If the study is to last 12 months and our reliability prediction yields an estimate of 6000 FITs, then Fig. 14 shows that a sample size of 900 is required to be 95 percent certain of observing the required number of failures. That is, we need 900 of each supplier's units in the study.

A.7 Cautions

In Cases D, E, and F, if the required number of failures is not observed in the nominal time period for the FTS, then the desired

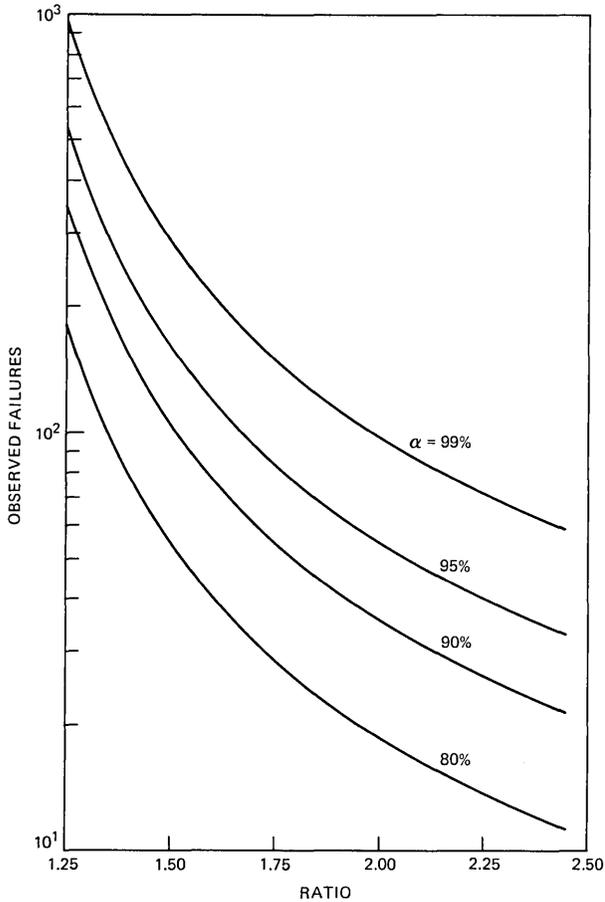


Fig. 17—Minimum observed failures comparing two failure rates.

precision will not be achieved. (This might occur if the reliability prediction is in error and yields a higher than actual FIT rate as a prediction. If the prediction is much higher than the actual, we will be incorrectly led to believe that the required number of failures will be observed in a shorter interval than is actually needed.) In this case it would be wise to extend the study period until the required number of failures is observed.

The theory developed for Cases D, E, and F requires that the failure rate be constant throughout the FTS. Even for very large sample sizes, the theory is sensitive to departures from this assumption. Therefore, if we know that the failure rate is high for one time period (e.g., early life) and low for a different time period (e.g., steady state), then we

must do a separate analysis on each period, as shown in the following example.

Assume that the early failure period is 3 months. Our reliability predictions indicate that the early failure rate will be about 10,000 FITs and that the steady-state failure rate will be about 4,000 FITs. We wish to obtain a precision of 0.25 at the 90-percent confidence level in estimating each of the failure rates in an FITs that we wish to finish in 6 months or less. What sample size is needed? Figure 11 shows that we need to observe 41 failures, that is, we must observe 41 failures in the early-life period (months 1 to 3) and 41 failures in the steady-state period (months 4 to 6). Use of eq. (1) shows that we need at least 2400 units for the early-life period and 5980 for the steady-state period. Since we need to satisfy both requirements we will need a sample size of 5980.

The example above illustrates another important point. If you want to use the FITs to estimate several characteristics, then go through the sample size analysis for each characteristic. The FITs will satisfy all requirements if it has the maximum of the required sample sizes.

In Cases B, D, E, and F, curves for several confidence levels are placed on one page. However, for Cases A and C, each confidence level would take a separate page, so only the 90-percent confidence level was given. For other confidence levels, see Ref. 8.

LETTER TO THE EDITOR

Comments on "Voice Storage in the Network—Perspective and History,"
by E. Nussbaum*

In a recent article E. Nussbaum discussed the FCC's rejection of AT&T's petition for waiver to allow the offering of Custom Calling Services II in the U.S. under the Computer Inquiry II decision. Unfortunately, references were not given to these decisions for the benefit of those readers who may wish to learn more about this apparent frustration of technology and the policy issues involved. The FCC rejection can be found in 88 FCC 2d 1. The Computer Inquiry II decision is given in 47 CFR 64.702, adopted in 77 FCC 2d 384 (Final Decision) on reconsideration, 84 FCC 2d 50, appeal pending *sub nom* CCIA vs. FCC, Case No. 80-471 (D.C. Cir. 1980).

Michael J. Marcus
Acting Chief
Technical Analysis Division
Office of Science & Technology
Federal Communications Commission

* B.S.T.J., 61, No. 5 (May-June 1982), pp. 811-13.

CONTRIBUTORS TO THIS ISSUE

Sigmund J. Amster, B.S., 1954, M.S., 1955, PH.D., 1962 (Mathematical Statistics), University of North Carolina; Bell Aircraft Corporation, 1955–56, General Analysis Corporation, 1959–60; Bell Laboratories, 1962—. Mr. Amster has worked in the areas of reliability and statistics with emphasis in mathematical models and experimental design. He is presently a consultant in the Quality Assurance Center.

David J. Bishop, B.S. (Physics), 1973, Syracuse University; M.S., 1977, Ph.D. (Physics), 1978, Cornell University; Bell Laboratories, 1978—. Mr. Bishop's research interests include low-temperature physics and acoustics.

Gary G. Brush, B.S. (Statistics and Computer Science), 1970, University of Delaware; M.S. (Statistics), 1972, Rutgers State University; Bell Laboratories, 1970—. Mr. Brush has been involved in a wide spectrum of quality assurance related activities including field studies, acceptance sampling, and evaluation of general trade products. He is currently in the Transmission Systems Quality Department, where he supervises a group responsible for quality engineering of transmission systems. Member, ASQC.

Bisrat Bulcha, B.S. (E.E.), 1971, Rensselaer Polytechnic Institute; M.S. (Operations Research), 1976, Syracuse University; L. M. Ericsson, Sweden, 1972–1973; Ethiopian Telecommunication Board, 1973–1974; Bell Laboratories, 1978—. At Bell Laboratories Mr. Bulcha has worked on economic studies of digital carrier in the loop plant.

Richard V. Cox, B.S. (Electrical Engineering), 1970, Rutgers University; M.A., and Ph.D., (Electrical Engineering), 1972 and 1974, respectively, Princeton University; The Aerospace Corporation, 1973–1977; Assistant Professor, Rutgers University, 1977–1979; Bell Laboratories, 1979—. Mr. Cox is a member of the Acoustics Research Department. His current research interests are in digital speech coding, analog speech encryption, and real-time speech processing systems.

Ronald E. Crochiere, B.S. (Electrical Engineering), 1967, Milwaukee School of Engineering, Milwaukee, Wisconsin; M.S. and Ph.D., (Electrical Engineering), 1968 and 1974, respectively, Massachusetts Institute of Technology, Cambridge, Massachusetts. Raytheon Co.,

1968-1970; Bell Laboratories, 1974—. At Raytheon Co. Mr. Crochiere was involved in the development of microwave phase shifters and side-band generators. In 1974 he joined the Acoustics Research Department of Bell Laboratories, where he has been involved in research activities in concepts of decimation and interpolation of digital signals, Fourier and filterbank methods of analysis/synthesis, sub-band and transform coding of speech, and the measurement of digital speech quality. In 1976 he received the IEEE ASSP paper award for his paper on decimation and interpolation of digital signals. Mr. Crochiere is an active member of the ASSP ADCOM committee and a past member of the ASSP Technical Committee on Digital Signal Processing. He has served for two years as a Technical Editor on digital signal processing for the ASSP transactions and for three years as the secretary-treasurer of the ASSP Society ADCOM. He is presently the vice-president of the ASSP Society.

Robert C. Dynes, B.Sc. (Physics), 1964, University of Western Ontario; M.Sc., 1965, Ph.D., 1968 (Physics), McMaster University; Bell Laboratories, 1968—. At Bell Laboratories Mr. Dynes' research interests are low-temperature physics, superconductivity, transport, and acoustics.

Philip B. Grimado, B.S., 1961 (Civil Engineering), City University of New York; M.S., 1962 (Applied Mechanics), and Ph.D., 1968 (Applied Mechanics), Columbia University, New York; Bell Laboratories, 1968—. Mr. Grimado's responsibilities include vulnerability studies of antiballistic missile systems, fire protection studies involving fire risk analyses, heat transfer calculations, development of standard fire testing methods for Bell System equipment, and development of algorithms for optimum control of building environmental equipment. He is currently completing a transient building response computer program for evaluating energy use characteristics of Bell System buildings.

James D. Johnston, B.S., 1975, M.S., 1976 (Electrical Engineering), Carnegie-Mellon University. Mr. Johnston has worked in the Acoustics Research Department of Bell Laboratories since June 1976, where he is a Member of Technical Staff. He has been involved in digital signal processing and speech encoding research, especially focusing on algorithms that can be adapted to real-time hardware. His interests include speech waveform encoding, VLSI and microprocessor-based hardware speech coding systems, real-time computing architecture, and microprocessor realizations of control systems for

acoustic and signal processing devices. Member, AES, and Contributor to IEEE-ASSP, IEEE-COMM, ASA, and IEEE Trans. on Comm.

Ronald Kaufman, B.A. (Mathematics/Statistics), 1973, University of Connecticut; M.A. (Operations Research), 1975, Ph.D. (Operations Research), 1977, Yale University; Bell Laboratories, 1977—. Mr. Kaufman has worked on various aspects of private network administration. He is currently in the Operations Systems Planning Center.

Adrian Kester, drs (Mathematics/Statistics), 1972, University of Amsterdam; M.Sc. (Business), 1976, Ph.D. (Statistics), 1978, University of Rochester; Bell Laboratories, 1977—. Mr. Kester has worked on various aspects of private network administration. He is currently in the Product Line Planning and Architecture Center.

Lynn E. Kodrich, B.S. (Mathematics), 1972, St. Peter's College; M.S. (Computer Science), 1975, Stevens Institute of Technology; Bell Laboratories, 1972—. Ms. Kodrich has been engaged in the development of software systems used for capacity expansion planning of the loop network. This includes an on-line Cumulative Fill Table program, an area transfer facility, and a conduit sizing algorithm. Since 1979, she has been involved with a software system for Digital Loop Carrier Planning.

David B. Luber, B.S. (E.E.), 1965, Worcester Polytechnic Institute; M.S. (E.E.), 1967, University of Pennsylvania; Ph.D. (S.E.), 1972, University of Pennsylvania; General Electric Company—Space Systems Organization, 1967–1969; Bell Laboratories, 1972—. Since joining Bell Laboratories, Mr. Luber has been active in developing method systems and mechanized planning tools for engineering the loop plant network. He is presently Supervisor of the Engineering Planning System Group in the Loop Engineering Operations department. Member, IEEE, Eta Kappa Nu, Tau Beta Pi; Associate Member, ORSA, Sigma Xi.

Martin H. Meyers, B.S. (EE), 1975, Worcester Polytechnic Institute, Worcester, MA; M.S.E.E., 1976, Ph.D., 1978, University of Massachusetts, Amherst, MA; Bell Laboratories, 1978—. Mr. Meyers initially worked on crosstalk modeling and analysis for digital paired cable repeater systems. Since 1981, he has been a member of the Digital Radio Department, where he has been evaluating the perform-

ance of various modulation, equalization, and combiner strategies for high-speed digital radio transmission over multipath fading channels.

William J. Mitchell, B.E. (Electrical Engineering), 1975, The Cooper Union for the Advancement of Science and Art; M.S. (Electrical Engineering and Computer Science), and (Electrical Engineer), 1977, M.I.T.; Bell Laboratories, Summer 1974, 1977—. Since joining Bell Laboratories, Mr. Mitchell has worked on the development of economic models and planning methods for new loop plant technologies, especially digital loop carrier systems. He has also developed computer aids and data bases to support this modeling effort and is presently working on models for new services and the loop plant needed to provide them. Member, Tau Beta Pi, Sigma Xi (associate), Eta Kappa Nu, IEEE (Communications, Computer, Industry Applications, and Power Engineering Societies).

Lawrence R. Rabiner, S.B. and S. M., 1964, Ph.D., 1967 (Electrical Engineering), Massachusetts Institute of Technology; Bell Laboratories, 1962—. From 1962 through 1964, Mr. Rabiner participated in the cooperative plan in electrical engineering at Bell Laboratories. He worked on digital circuitry, military communications problems, and problems in binaural hearing. Presently, he is engaged in research on speech communications and digital signal processing techniques. He is coauthor of *Theory and Application of Digital Signal Processing* (Prentice-Hall, 1978). Member, IEEE ASSP Technical Committee on Digital Signal Processing, IEEE ASSP Proceedings Editorial Board, Acoustical Society Executive Council, Eta Kappa Nu, Sigma Xi, Tau Beta Pi. Fellow, Acoustical Society of America, IEEE.

William D. Rummmler, B.S. in E.E., 1959, Drexel University; S.M., 1960, E.E., 1961, Sc.D., 1963, Massachusetts Institute of Technology; Bell Laboratories, 1963—. Mr. Rummmler initially worked on radar signal design and signal processing and associated measurement and estimation problems. He has also worked on single-sideband, long-haul radio systems. More recently, he has been involved with the problems of channel and performance characterization for high-speed digital radio systems. Member, IEEE, Tau Beta Pi, Eta Kappa Nu, Sigma Xi, Phi Kappa Phi.

Adel A. M. Saleh, B.Sc. degree (Electrical Engineering), University of Alexandria, Alexandria, Egypt, 1963; M.S., and Ph.D. degrees (Elec-

trical Engineering), Massachusetts Institute of Technology, Cambridge, in 1967 and 1970, respectively; Bell Laboratories, 1970—. From 1963 to 1965 Mr. Saleh worked as an instructor at the University of Alexandria. At Bell Laboratories he is engaged in research on microwave circuits, components, and communication systems. Senior Member, IEEE; member, Sigma Xi.

Bernard Saperstein, B.S. (Mathematics), 1964, CCNY; PH.D. (Operational Research), 1969, New York University; Bell Laboratories, 1969—. Mr. Saperstein's work has centered on applied research in reliability, quality control, and quality assurance. Since 1978, he has supervised groups responsible for field quality tracking of customer premises and switching network equipment in the Field Quality and Reliability Department.

Michael A. Schwartz, B.S. (Electrical Engineering), 1968, Rensselaer Polytechnic Institute; M.S. (Electrical Engineering), 1969, University of California—Berkeley; C&P Telephone Company of West Virginia, 1972–1974; Bell Laboratories, 1968–1972, 1974—. Since joining the Bell System, Mr. Schwartz has been involved in noise mitigation studies, cable test set development, subscriber carrier hardware development, application engineering and system planning, and operations planning for introduction of carrier systems into the subscriber network. He is presently Supervisor of the New Technology Operations Planning Group in the Customer Network Strategic Planning department. Member, Tau Beta Pi, IEEE.

Linda A. Seltzer, B.A., 1971, University of Pennsylvania; B.S.E.E., 1978, University of Colorado; M.S., 1980, California Institute of Technology. Ms. Seltzer is currently a graduate student at the University of California, San Diego. She has worked on multiport network studies for audio conferencing, on digital signal processing hardware, and on subjective quality of low bit rate speech. Member, IEEE, Acoustical Society of America, Tau Beta Pi, and Eta Kappa Nu.

Paul E. Sulewski, A.B. (Physics), 1982, Princeton University. Mr. Sulewski has recently completed his undergraduate thesis research for Princeton University, which involved examining several optoelectronic materials of current interest with acoustic microscopy. He will pursue his interests in solid state physics at Cornell University.

José M. Tribolet, B.S. (Electrical Engineering), 1972, Instituto Superior Técnico, Lisbon, Portugal; M.S., E.E., and Sc.D degrees (Electrical Engineering) from the Massachusetts Institute of Technology, Cambridge, in 1974, 1975, and 1977, respectively. From 1972 to 1977 Mr. Tribolet was a member of the Massachusetts Institute of Technology Research Laboratory of Electronics, where his research activities involved the application of homomorphic signal processing to speech and seismic data analysis. From 1977 to 1978 he was with the Acoustics Research Department, Bell Laboratories, Murray Hill, NJ, as a post-doctoral fellow, working on adaptive transform coding of speech. He is presently Full Professor of Electrical Engineering and Computer Science at the Instituto Superior Técnico, Lisbon, Portugal, where he directs the Research Institute in Systems Engineering and Computer Science (INESC). Mr. Tribolet was recently on sabbatical leave at the Acoustics Research Department, Bell Laboratories, Murray Hill, NJ, from July through December 1981, where he worked on speech recognition, coding, and scrambling. Member, Sigma Xi.

Jay G. Wilpon, B.S. (Mathematics), A.B. (Economics), cum laude, 1977, Lafayette College; Bell Laboratories, 1977—. At Bell Laboratories, Mr. Wilpon has been engaged in speech communications research and is presently concentrating on problems of speech recognition.

F. Nelson Woomeer, Jr., Electronic Technology, 1952, RCA Institutes; Bell Laboratories, 1952—. Mr. Woomeer has been involved in the development of fire control systems, including the Missile Site Data Processing System at Meck Island, Kwajalein Atoll in the Marshall Islands. On SAFEGUARD and AMPS he developed and maintained project integration plans. Since joining the Loop Engineering department he has developed and published the Outside Plant Planning Engineering and Construction Operation System Implementation and Application Guide and performed economic impact studies.

PAPERS BY BELL LABORATORIES AUTHORS

COMPUTING SCIENCES

- Benevit C. A., Cassard J. M., Dimmler K. J., Dumbri A. C., Mound M. G., Procyk F. J., Rosenzweig W., Yanof A. W., **256K Dynamic Random-Access Memory**. *ISSCC Diges* 25:76-77, 1982.
- Covington, W. O. et al., **A Test Data Synthesizer for Telephone Company Accounting Office Processing of Call Data**. *Software* 12(6):505-516, 1982.
- Joy D. C., **Microcomputer Control of a Stem**. *Ultramicros* 8(3):301-308, 1982.
- Levendel Y. H., Menon, P. R., **Test-Generation Algorithms for Computer Hardware Description Languages**. *IEEE Comput* 31(7):577-588, 1982.
- Lund J. C., Ordun M. R., Wojcik R. J., **Implementation of the Calling Card Service Capability—Application of a Software Methodology**. *IEEE Commun* 30(6):1297-1303, 1982.
- McConkey J., **Report on the Symposium on Small Computers in the Arts (Editorial)**. *Comput Mus* 6(2):61-64, 1982.
- Utter D. F., **Writing Requirements for Interfaces Between Computers**. *IEEE Commun* 30(6):1276-1280, 1982.
- White P. E., Feay M. R., **A Software-Development System for Reliable Applications**. *IEEE Commun* 30(6):1363-1368, 1982.
- Yannakakis M., **A Theory of Safe Locking Policies in Database-Systems**. *J ACM* 29(3):718-740, 1982.

ENGINEERING

- Alles D. S., **Accuracy in Lithography (Editorial)**. *Sol St Tech* 25(5):47, 1982.
- Anthony P. J. et al., **Device Characteristics of (Al, Ga)As Lasers With Ga(As, Sb) Active Layers**. *IEEE J Q El* 18(7):1094-1100, 1982.
- Bechtel J. H., Chraplyvy A. R., **Laser Diagnostics of Flames, Combustion Products, and Sprays**. *P IEEE* 70(6):658-677, 1982.
- Bergland G. D., Cheung R. C., **Special Issue on Communication Software—Prologue (Editorial)**. *IEEE Commun* 30(6):1265-1275, 1982.
- Beuscher H. J., **No. 5 ESS Maintenance Software**. *IEEE Commun* 30(6):1386-1392, 1982.
- Brown M. K., Rabiner L. R., **An Adaptive, Ordered, Graph Search Technique for Dynamic Time Warping for Isolated Word Recognition**. *IEEE Acoust* 30(4):535-544, 1982.
- Chu T. S., Legg W. E., **Gain of Corrugated Conical Horns**. *IEEE Antenn* 30(4):698-703, 1982.
- Cohen L. G., Stone J., **Interferometric Measurements of Minimum Dispersion Spectra in Short Lengths of Single-Mode Fiber**. *Electr Lett* 18(13):564-566, 1982.
- Cohen M. I., **The Special Issue on Laser Applications (Editorial)**. *P IEEE* 70(6):531-532, 1982.
- Cowan M. J., **Mask Limitations to VLSI—Overlay Accuracy and Linewidth Variation**. *Sol St Tech* 25(5):55-59, 1982.
- Davis P. C., Crigler F. W., Spires D. A., Aull D. W., **A Hybrid Integrated Trunk and Subscriber Line Interface**. *ISSCC Diges* 25:204+, 1982.
- Derr M. A., Wright C. E., Coriell A. S., **A Satellite System for Controlling Real-Time Experiments**. *Behav Res M* 14(2):264-271, 1982.
- Duncan T., Huen W. H., **Software Structure of No-5 ESS—A Distributed Telephone Switching-System**. *IEEE Commun* 30(6):1379-1385, 1982.
- Duttweiler D. L., **Adaptive Filter Performance With Nonlinearities in the Correlation Multiplier**. *IEEE Acoust* 30(4):578-586, 1982.
- Erickson R. E., Halfin S., Luss H., **Optimal Sizing of Records When Divided Messages Can Be Stored in Records of Different Sizes**. *Operat Res* 30(1):29-39, 1982.

- Fleischer P. E. et al., **A Single-Chip Dual-Tone and Dial-Pulse Signaling Receiver.** *ISSCC Diges* 25:212+, 1982.
- Gagnepain J. J., et al., **Special Issue on Piezoelectricity—Guest Editorial (Editorial).** *Ferroelectr* 43(1-2):R 3-R 6, 1982.
- Ginsparg, J. M., Gordon R. D., **Automatic Programming of Communications of Software via Nonprocedural Descriptions.** *IEEE Commun* 30(6):1343-1347, 1982.
- Gordon R. D., **Providing Multiple-Channel Communication Using The Experimental Digital Switch.** *IEEE Commun* 30(6):1409-1416, 1982.
- Greenstein L. J. et al., **Performance Comparisons Among Digital Radio Techniques Subjected to Multipath-Fading.** *IEEE Commun* 30(5):1184-1197, 1982.
- Helfand E., Pearson D. S., **Calculation of the Non-Linear Stress of Polymers in Oscillatory Shear Fields.** *J Pol Sc PP* 20(7):1249-1258, 1982.
- Honeyman P., **Testing Satisfaction of Functional-Dependencies.** *J ACM* 29(3):668-677, 1982.
- Hornbach B. H., **MML—CCITT Man-Machine Language.** *IEEE Commun* 30(6):1329-1336, 1982.
- Huppert D., Rand S. D., Reynolds A. H., Rentzepis P. M., **Intersystem Crossing and Predissociation of Haloaromatics.** *J Chem Phys* 77(3):1214-1224, 1982.
- Johnson R. E., Brown W. L., **Electronic Mechanisms for Sputtering of Condensed-Gas Solids by Energetic Ions.** *Nucl Instru* 198(1):103-118, 1982.
- Johnson R. E., Lanzerotti L. J., Brown W. L., **Planetary Applications of Ion-Induced Erosion of Condensed-Gas Frosts.** *Nucl Instru* 198(1):147-157, 1982.
- Kant R., **The Elastostatic Problem of an Infinite Strip Containing a Periodic Row of Line Cracks.** *Int J Eng S* 20(9):1057-1070, 1982.
- Kardontchik J. E., **Mode Reflectivity of Narrow Stripe-Geometry Double Heterostructure Lasers.** *IEEE J Q El* 18(8):1279-1286, 1982.
- Kardontchik J. E., **Far Field Asymmetry in Narrow Stripe Gain-Guided Lasers.** *IEEE J Q El* 18(8):1287-1294, 1982.
- Kodama Y., Hasegawa A., **Amplification and Reshaping of Optical Solitons in Glass-Fiber. 2.** *Optics Lett* 7(7):339-341, 1982.
- Kopp W., Fischer R., Thorne R. E., Su S. L., Drummond T. J., Morkoc H., Cho A. Y., **A New Al_{0.3}Ga_{0.7}As GaAs Modulation-Doped FET.** *Elec Dev L* 3(5):109-111, 1982.
- Krambeck R. H., Lee C. M., Law H. F. S., **High-Speed Compact Circuits With CMOS.** *IEEE J Soli* 17(3):614-619, 1982.
- Louzon T. J., **Tensile Property Improvements of Spinodal Cu-15Ni-8Sn by Two-Phase Heat-Treatment.** *J Eng Mater* 104(3):234-240, 1982.
- Mayo J. S., **Technology Needs of the Information Age.** *ISSCC Diges* 25:50-51, 1982.
- Mucha J. A., **Standard Addition Technique for Quantitative Trace Gas-Analysis Using Derivative Infrared Diode-Laser Spectroscopy.** *Appl Spectr* 36(4):393-400, 1982.
- Myers C. S., Levinson S. E., **Speaker Independent Connected Word Recognition Using a Syntax-Directed Dynamic-Programming Procedure.** *IEEE Acoust* 30(4):561-565, 1982.
- Nair V. N., **Some Extensions of ANOVA Techniques to Location-Scale Models.** *Comm Stat A* 11(14):1551-1570, 1982.
- Ogawa K., Vodhanel R. S., **Measurements of Mode Partition Noise of Laser-Diodes.** *IEEE J Q El* 18(7):1090-1093, 1982.
- Owen B., **PIN-GaAs FET Optical Receiver With a Wide Dynamic-Range.** *Electr Lett* 18(14):626-627, 1982.
- Paulikas G. A., Lanzerotti L. J., **Impact of Geospace on Terrestrial Technology (Review or Bibliography).** *Astro Aeron* 20(7-8):42+, 1982.
- Russell D. L., **Alternative Strategies for Managing Virtual Circuit Failure and Recovery in Local Area Computer-Networks.** *IEEE Commun* 30(6):1450-1454, 1982.
- Rustako A. J., **Rain Attenuation and Depolarization Over an Earth Space Path at 12 GHz—Experimental Results Using the CTS Beacon.** *IEEE Antenn* 30(4):720-725, 1982.
- Saunders T. E., **Wafer Flatness Utilizing the PIN-Recess Chuck.** *Sol St Tech* 25(5):73-76, 1982.

- Shakespeare W. J., **Laminar Boundary-Layer Near the Rotating End Wall of a Confined Vortex.** *J Fluid Eng* 104(2):171-177, 1982.
- Sinha A. K., **Metallization Technology for Very-Large-Scale Integration Circuits.** *Thin Sol Fi* 90(3):271-285, 1982.
- Smith E. R., Miller F. D., **Latent-Variable Models of Attributional Measurement.** *Pers Soc PS* 8(2):221-225, 1982.
- Smith G. E., Boyle W. S., **Citation Classic—Coupled Semiconductor-Devices.** *CC/Eng Tech* 1982(29):18, 1982.
- Stolen R. H., Bjorkholm J. E., **Parametric Amplification and Frequency-Conversion in Optical Fibers.** *IEEE J Q El* 18(7):1062-1072, 1982.
- Swartz R. G., Wooley B. A., Voshchenkov A. M., Archer V. D., Chin G. M., **A Monolithic Laser Driver for Optical Communications.** *ISSCC Diges* 25:162-163, 1982.
- Thompson J. S., Tewksbury S. K., **LSI Signal Processor Architecture for Telecommunications Applications.** *IEEE Acoust* 30(4):613-631, 1982.
- Thompson R. A., **Experimental Multiple-Channel Circuit-Switched Communications.** *IEEE Commun* 30(6):1399-1409, 1982.
- Whitt W., **Approximating a Point Process by a Renewal Process. 1. Two Basic Methods.** *Operat Res* 30(1):125-147, 1982.
- Williams G. F., **Wide-Dynamic-Range Fiber Optic Receivers.** *ISSCC Diges* 25:160-161, 1982.
- Wood O. R., Silfvast W. T., **Electron-Density and Energy Output Limits of Plasma-Recombination Lasers.** *Appl Phys L* 41(2):121-123, 1982.
- Yue O. C., **Maximum-Likelihood Combining for Noncoherent and Differentially Coherent Frequency-Hopping Multiple-Access Systems.** *IEEE Info T* 28(4):631-639, 1982.

MANAGEMENT/ECONOMICS

- Perry M. K., **Vertical Integration by Competitive Firms—Uncertainty and Diversification.** *S Econ J* 49(1):201-208, 1982.

MATHEMATICS

- Exoo G., Harary F., **The Smallest Cubic Multigraphs with Prescribed Bipartite, Block, Hamiltonian and Planar Properties.** *I J PA Math* 13(3):309-312, 1982.
- Fishburn P. C., **Aspects of Semi-Orders Within Interval Orders.** *Discr Math* 40(2-3):181-191, 1982.
- Fishburn P. C., **Dominant Strategies and Restricted Ballots With Variable Electorate.** *Math Soc Sc* 2(4):383-395, 1982.
- Kantor W. M., **Strongly Regular Graphs Defined by Spreads.** *ISR J Math* 41(4):298-312, 1982.
- Monma C. L., **Linear-Time Algorithms for Scheduling on Parallel Processors.** *Operat Res* 30(1):116-124, 1982.

PHYSICAL SCIENCES

- Allara D. L., Swalen J. D., **An Infrared Reflection Spectroscopy Study of Oriented Cadmium Arachidate Monolayer Films on Evaporated Silver.** *J Phys Chem* 86(14):2700-2704, 1982.
- Aspnes D. E., **High-Efficiency Concave-Grating Monochromator With Wavelength-Independent Focusing Characteristics.** *J Opt Soc* 72(8):1056-1061, 1982.
- Aspnes D. E., **Local-Field Effects and Effective-Medium Theory—A Microscopic Perspective.** *Am J Phys* 50(8):704-709, 1982.
- Aspnes D. E., Stucker H. J., **Peroxide Etch Chemistry on -Less-Than-100-Greater-Than-In_{0.53}Ga_{0.47}As).** *J Vac Sci T* 21(2):413-416, 1982.
- Brown W. L. et al., **Erosion and Molecule Formation in Condensed Gas Films by Electronic-Energy Loss of Fast Ions.** *Nucl Instru* 198(1):1-8, 1982.

- Burton R. H., Smolinsky G., **CCl₄ and Cl₂ Plasma-Etching of III-V-Semiconductors and the Role of Added O₂**. *J Elchem So* 129(7):1599-1604, 1982.
- Campbell J. C., Dentai A. G., **InP/InGaAs Heterojunction Photo-Transistor with Integrated Light-Emitting Diode**. *Appl Phys L* 41(2):192-193, 1982.
- Campbell J. C., Qua G. J., Copeland J. A., Dentai A. G., **Light-Activated Electroluminescent Switch With an Active Feedback-Circuit**. *J Appl Phys* 53(7):5182-5185, 1982.
- Carlson N. W., Geschwind S., Walker L. R., Devlin G. E., **Field-Induced Dispersion in the Amorphous Antiferromagnet**. *Phys Rev L* 49(2):165-168, 1982.
- Chabal Y. J., Hamann D. R., Rowe J. E., Schluter M., **Photoemission and Band-Structure Results for NiSi₂**. *Phys Rev B* 25(12):7598-7602, 1982.
- Chen C. H., Gibson J. M., Fleming R. M., **Microstructure in the Incommensurate and the Commensurate Charge-Density-Wave States of 2H-TASE2—A Direct Observation by Electron-Microscopy**. *Phys Rev B* 26(1):184-205, 1982.
- Chen C. Y., Cho A. Y., Bethea C. G., Garbinski P. A., **Bias-Free Selectively Doped Al_xGa_{1-x}As-GaAs Picosecond Photodetectors**. *Appl Phys L* 41(3):282-284, 1982.
- Chen C. Y., Cho A. Y., Cheng K. Y., Pearsall T. P., O'Connor P., Garbinski P. A., **Depletion Mode Modulation Doped Al_{0.48}In_{0.52}As-Ga_{0.47}In_{0.53}As Heterojunction Field-Effect Transistors**. *Elec Dev L* 3(6):152-155, 1982.
- Chen H. S., Fujiwara T., Waseda Y., **Structural Investigation of Boron-Containing Metallic Glasses**. *J Mater Sci* 17(5):1337-1347, 1982.
- Cox D. E., Cava R. J., McWhan D. B., Murphy D. W., **A Neutron Powder Diffraction Study of the Lithium Insertion Compound LiMoO₂ from 4-44OK**. *J Phys Ch S* 43(8):657-666, 1982.
- Cross M. C., Hohenberg P. C., Safran S. A., **Wave Number Selection in Rayleigh-Benard Convection—A Numerical Study**. *Physica D* 5(1):75-82, 1982.
- Cross M. C., **Boundary-Conditions on the Envelope Function of Convective Rolls Close to Onset**. *Phys Fluids* 25(6):936-941, 1982.
- Cullis A. G., Webber H. C., Chew N. G., Poate J. M., Baleri P., **Transitions to Defective Crystal and the Amorphous State Induced in Elemental Si by Laser Quenching**. *Phys Rev L* 49(3):219-222, 1982.
- Daniels P. G., Blythe P. A., Simpkins P. G., **Thermally Driven Cavity Flows in Porous-Media. 2. The Horizontal Boundary-Layer Structure**. *P Roy Soc A* 382(1782):135-154, 1982.
- Degani J., Leheny R. F., Nahory R. E., Shah J., **High-Field Transport Characteristics of Minority Electrons in p-In_{0.53}Ga_{0.47}As**. *Thin Sol Fi* 91(1):65-70, 1982.
- Digiuseppe M. A., Temkin H., Bonner W. A., **Large Area LPE Growth of InGaAsP InP Double Heterostructures on InP Preserved in a Phosphorus Ambient**. *J Cryst Gr* 58(1):279-284, 1982.
- DiSalvo F. J., **Some Remarkable Properties of Transition-Metal Chalcogenide Compounds**. *Ann Chim-Sc* 7(2-3):109-118, 1982.
- Duchovic R. J., Hase W. L., Schlegel H. B., Frisch M. J., Raghavachari K., **Abinitio Potential-Energy Curve for CH Bond-Dissociation in Methane**. *Chem P Lett* 89(2):120-125, 1982.
- Favier J. J., Wilson L. O., **A Test of the Boundary-Layer Model in Unsteady Czochralski Growth**. *J Cryst Gr* 58(1):103-110, 1982.
- Fork R. L., Shank C. V., Yen R. T., **Amplification of 70-fs Optical Pulses to Gigawatt Powers**. *Appl Phys L* 41(3):223-225, 1982.
- Franciosi A., Weaver J. H., O'Neill D. G., Chabal Y., Rowe J. E., Poate J. M., Bisi O., Calandra C., **Chemical Bonding at the Si-Metal Interface—Si-Ni and Si-Cr**. *J Vac Sci T* 21(2):624-627, 1982.
- Gallagher P. K., **Kinetics of the Dissociation of InP Under Vacuum**. *J Phys Chem* 86(16):3246-3250, 1982.
- Geballe T. H., Rowell J. M., **Vapor-Deposited Metastable Superconductors**. *Thin Sol Fi* 91(1):33-43, 1982.
- Gibbs H. M. et al., **Room-Temperature Excitonic Optical Bistability in a GaAs-GaAlAs Super-Lattice Etalon**. *Appl Phys L* 41(3):221-222, 1982.
- Gole J. L., English J. H., Bondybey V. E., **Laser Spectroscopy of Cooled Metal Clusters—Copper Dimer**. *J Phys Chem* 86(14):2560-2563, 1982.
- Gooden R., Mitchell J. W., **Photochemical Production, Infrared Absorptivity, and**

- Chemical-Reactions of Trichlorosilanol in Silicon Tetrachloride.** *J Elchem So* 129(7):1619-1622, 1982.
- Greywall D. S., Busch P. A., High-Precision Specific-Heat Measurements on Normal Liquid ^3He .** *Phys Rev L* 49(2):146-149, 1982.
- Griffiths J. E., Sunder W. A., Raman Study of the Phase-Transitions and Ionic Motions in Dioxygenyl Hexafluoroarsenate $\text{O}_2^+\text{AsF}_6^-$.** *J Chem Phys* 77(3):1087-1092, 1982.
- Gyorgy E. M., McWhan D. B., Dillon J. F., Walker L. R., Waszczak J. V., Magnetic-Behavior and Structure of Compositionally Modulated Cu-Ni Thin-Films.** *Phys Rev B* 25(11):6739-6747, 1982.
- Haight R., Feldman L. C., Atomic-Structure at the (111) Si-SiO₂ Interface.** *J Appl Phys* 53(7):4884-4887, 1982.
- Heller A., Chemical Control of Recombination at Grain Boundaries and Liquid Interfaces: Electrical Power and Hydrogen Generating Photoelectrochemical Cells.** *J Vac Sci T* 21(2):559-561, 1982.
- Jelinski L. W., Dumais J. J., Schilling F. C., Bovey F. A., Solid-State C-13 NMR-Studies of Polyester Thermoplastic Elastomers (Review or Bibliog.).** *ACS Symp S* 1982(191):345-360, 1982.
- Johnson P. D., Smith N. V., Inverse Photoemission Observation of an Unoccupied Surface-State on Pd(111).** *Phys Rev L* 49(4):290-292, 1982.
- Kastalsky A., Dingle R., Cheng K. Y., Cho A. Y., Two-Dimensional Electron-Gas at a Molecular-Beam Epitaxial-Grown, Selectively Doped, $\text{In}_{0.53}\text{Ga}_{0.47}\text{As-In}_{0.48}\text{Al}_{0.52}\text{As}$ Interface.** *Appl Phys L* 41(3):274-277, 1982.
- Kazarinov R. F., Henry C. H., Logan R. A., Longitudinal Mode Self-Stabilization in Semiconductor-Lasers.** *J Appl Phys* 53(7):4631-4644, 1982.
- Kazarinov R. F., Luryi S., Quantum Percolation and Quantization of Hall Resistance in Two-Dimensional Electron-Gas.** *Phys Rev B* 25(12):7626-7630, 1982.
- Kazarinov R. F., Luryi S., Majority Carrier Transistor Based on Voltage-Controlled Thermionic Emission.** *Appl Phys A* 28(3):151-160, 1982.
- Kelley D. F. et al., Spectroscopy and Nonradiative Relaxation of Propynal.** *J Chem Phys* 77(2):697-701, 1982.
- Kelso S. M., Calculation of Properties of the Electron-Hole Liquid in Uniaxially Stressed Ge and Si.** *Phys Rev B* 25(12):7631-7651, 1982.
- Klauder J. R., High-Temperature Expansion for Non-Renormalizable Quantum-Field Theories.** *Czech J Phys* 32(5):494-498, 1982.
- Kwei T. K., Phase-Separation in Segmented Polyurethanes.** *J Appl Poly* 27(8):2891-2899, 1982.
- Lamola A. A., Wrighton M. S., Recommended Standards for Reporting Photochemical Data.** *Pur A Chem* 54(6):1251-1256, 1982.
- Lee T. P., Burrus C. A., Copeland J. A., Dentai A. G., Marcuse D., Short-Cavity InGaAsP Injection-Lasers—Dependence of Mode Spectra and Single-Longitudinal-Node Power on Cavity Length.** *IEEE J Q El* 18(7):1101-1113, 1982.
- Levin R. M., Adams A. C., Low-Pressure Deposition of Phosphosilicate Glass-Films.** *J Elchem So* 129(7):1588-1592, 1982.
- Liao A. S. H., Tell B., Leheny R. F., Nahory R. E., DeWinter J. C., Martin R. J., An $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ p-Channel MOSFET With Plasma-Grown Native Oxide Insulated Gate.** *Elec Dev L* 3(6):158-160, 1982.
- Liao A. S. H., Tell B., Leheny R. F., Chang T. Y., $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ n-Channel Native Oxide Inversion Mode Field-Effect Transistor.** *Appl Phys L* 41(3):280-282, 1982.
- Lines M. E., Effect of Electric-Field Gradient Sign on Iron Mossbauer-Spectra in Amorphous Materials.** *J Phys Ch S* 43(8):723-730, 1982.
- Lundgren C. A., Rice C. E., Electrochemistry of Silane-Derivatized Iridium.** *ACS Symp S* 1982(192):197-204, 1982.
- Luther L. C., Lecraw R. C., Ermanis F., Bismuth Content and Growth-Induced Anisotropy in YLuBiLG.** *J Cryst Gr* 58(1):95-102, 1982.
- Mahajan S. et al., Manifestations of Melt-Carry-Over in InP and InGaAsP Layers Grown by Liquid-Phase Epitaxy.** *J Electrochem Soc* 129(3):106, 1982.
- Mahajan S., Keramidis V. G., Bonner W. A., The Influence of Intermittent Growth Procedures on Dislocation Densities in InP Epi-Layers.** *J Elchem So* 129(7):1556-1559, 1982.

- McIntyre J. D., Basu S., Peck W. F., Brown W. L., Augustyniak W. M., **Cation Insertion Reactions of Electrochromic Tungsten and Iridium Oxide-Films.** *Phys Rev B* 25(12):7242-7254, 1982.
- Merlin R., Pinczuk A., Beard W. T., Wood C. E. E., **Light-Scattering Study of Electrons Confined at Ge/GaAs Interfaces.** *J Vac Sci T* 21(2):516-518, 1982.
- Miller D. C., Pernell T. L., **Fluid-Flow Patterns in a Simulated Garnet Melt.** *J Cryst Gr* 58(1):253-260, 1982.
- Mills A. P., Wilson R. J., **Transmission of 1-6-Kev Positrons Through Thin Metal-Films.** *Phys Rev A* 26(1):490-500, 1982.
- Nakahara S., Felder E. C., **Generation of Grown-In Dislocations at Ni-Cu Misfit Boundaries During Three-Dimensional Nucleation and Growth.** *Thin Sol Fi* 91(2):111-122, 1982.
- Nassau K., Chadwick D. L., **A Survey of Some Germania-Containing Glass-Forming Compositions.** *Mater Res B* 17(6):715-719, 1982.
- Nelson D. F., **Transport of Broadening Charge Packets at Surfaces.** *J Appl Phys* 53(7):5013-5021, 1982.
- Patterson G. D., Stevens J. R., Carroll P. J., **Photon-Correlation Spectroscopy of Polystyrene Near the Glass-Transition Under High-Pressure.** *J Chem Phys* 77(2):622-624, 1982.
- Sinclair J. D., **Paper Extraction for Sampling Inorganic Salts on Surfaces (Review or Bibliog.).** *Analyt Chem* 54(9):1529-1533, 1982.
- Smith P. W., Maloney P. J., Ashkin A., **Use of a Liquid Suspension of Dielectric Spheres as an Artificial Kerr Medium.** *Optics Lett* 7(8):347-349, 1982.
- Sneddon L., Cross M. C., Fisher D. S., **Sliding Conductivity of Charge-Density Waves.** *Phys Rev L* 49(4):292-295, 1982.
- Swaminathan V. et al., **Doping of Ga_{1-x}Al_xAs Grown by LPE With Si and Ge.** *J Elchem So* 129(7):1563-1566, 1982.
- Swaminathan V., Zilko J. L., Tsang W. T., Wagner W. R., **Photo-Luminescence Study of Acceptors in Al_xGa_{1-x}As.** *J Appl Phys* 53(7):5163-5168, 1982.
- Swartz R. G., McFee J. H., Voshchenkov A. M., Finegan S. N., Archer V. D., O'Day P. J., **Insitu Low-Energy BF₃⁺ Ion Doping for Silicon Molecular-Beam Epitaxy.** *Elec Dev L* 3(5):138-140, 1982.
- Thiel F. A., **The Phase-Relations in the Cu, In, S System and the Growth of CuInS₂ Crystals from the Melt.** *J Elchem So* 129(7):1570-1571, 1982.
- Treacy M. M. J., Gibson J. M., **On the Detection of Point-Defects in Crystals Using High-Angle Diffuse-Scattering in the Stem.** *Inst Phys C*1982(61):263-266, 1982.
- Truhlar D. G., Duff J. W., Blais N. C., Tully J. C., Garrett B. C., **The Quenching of Na(3²P) by H₂-Interactions and Dynamics.** *J Chem Phys* 77(2):764-776, 1982.
- Weber T. A., Stillinger F. H., **Solvating Collisions Between Deuterons and Light-Water Octamers.** *Chem P Lett* 89(2):154-158, 1982.

SOCIAL AND LIFE SCIENCES

- Cassidy M. F., **Instructional-Technique—Davies, I. K. (Book Review).** *ECTJ* 30(2):120-122, 1982.
- Cassidy M. F., **Toward Integration—Education, Instructional Technology, and Semiotics.** *ECTJ* 30(2):75-89, 1982.
- Eisinger J., **Lead and Wine Gockel, Eberhard and the Colicapictonum.** *Med Hist* 26(3):279+, 1982.
- Fisher D. G., Rothkopf E. Z., **An Automated Eye-Movement Laboratory for On-Line Electrooculography.** *Behav Res M* 14(2):113-120, 1982.
- Ondrias M. R., Rousseau D. L., Shelnut J. A., Simon S. R., **Quaternary-Transformation-Induced Changes at the Heme in Deoxyheoglobins.** *Biochem* 21(14):3428-3437, 1982.
- Pregibon D., **Resistant Fits for Some Commonly Used Logistic-Models With Medical Applications.** *Biometrics* 38(2):485-498, 1982.
- Reed A. V., **The Radio Shack Color Computer in the Experimental-Psychology Laboratory—An Evaluation.** *Behav Res M* 14(2):109-112, 1982.

Rothkopf E. Z., **Principles of Instructional-Design, 2nd Edition**—Gagne R. M., Briggs L. J. (Book Review). Cont Psycho 27(7):552, 1982.

SPEECH AND ACOUSTICS

Weste N. H. E., Burr D. J., Ackland B. D., **A Systolic Processing Element for Speech Recognition**. ISSCC Diges 25:274-275, 1982.

CONTENTS, DECEMBER 1982

An Overview of PANACEA, a Software Package for Analyzing Markovian Queueing Networks

K. G. Ramakrishnan and D. Mitra

Special-Purpose Computer for Logic Simulation Using Distributed Processing

Y. H. Levendel, P. R. Menon, and S. H. Patel

The Limit of the Blocking as Offered Load Decreases With Fixed Peakedness

P. J. Burke

TV Bandwidth Compression Techniques Using Time-Companded Differentials and Their Applications to Satellite Transmissions

K. Y. Eng and B. G. Haskell

Sensitivity of Avalanche Photodetector Receivers for Long-Wavelength Optical Communications

R. G. Smith and S. R. Forrest

Transmitting Data on the Phase of Speech Signals

W. C. Wong, R. Steele, and C. S. Xydeas

On the Use of Energy in LPC-Based Recognition of Isolated Words

M. K. Brown and L. R. Rabiner

Isolated Word Recognition for Large Vocabularies

L. R. Rabiner, A. E. Rosenberg, J. G. Wilpon, and W. J. Keilin

Multitone Frequency-Hopped MFSK System for Mobile Radio

Uzi Timor

Design of Recovery Strategies for a Fault-Tolerant No. 4 Electronic Switching System

R. J. Willett

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION
(Act of August 12, 1970: Section 8685. Title 39, U.S. Code)

1. Title of publication: Bell System Technical Journal.
2. Date of filing: October, 1982.
3. Frequency of issue: Monthly, except combined May-June and July-August.
4. Publication office: Bell Laboratories, 101 J. F. Kennedy Pky, Short Hills, N. J. 07078.
5. Headquarters: American Telephone and Telegraph Co., 195 Broadway, New York, N. Y.
6. Editor and publisher: B. G. King, Editor, Bell Laboratories, 101 J. F. Kennedy Pky, Short Hills, N. J. 07078. Publisher, American Telephone and Telegraph Co., 195 Broadway, New York, N. Y.
7. Owner: American Telephone and Telegraph Co., 195 Broadway, New York, N. Y.
8. Bondholders, mortgages and other security holders: none.
9. Extent and nature of circulation:

	Average number of copies of each issue during preceding 12 months	Actual number of copies of single issue published nearest to filing date
A. Total no. copies printed (Net press run)	<u>12,397</u>	<u>12,428</u>
B. Paid circulation		
1. Sales through dealers and carriers, street vendors and counter sales	<u>2,658</u>	<u>2,634</u>
2. Mail subscriptions	<u>5,572</u>	<u>5,719</u>
C. Total paid circulation	<u>8,230</u>	<u>8,353</u>
D. Free distribution by mail, carrier or other means, samples, complimentary, and other free copies	<u>3,373</u>	<u>3,280</u>
E. Total distribution (sum of C and D)	<u>11,603</u>	<u>11,633</u>
F. Copies not distributed		
1. Office use, left over, unaccounted, spoiled after printing	<u>794</u>	<u>795</u>
2. Returns from news agents	None	None
G. Total (sum of E and F—should equal net press run shown in A)	<u>12,397</u>	<u>12,428</u>

I certify that the statements made by me above are correct and complete.

Bernard G. King, Editor

THE BELL SYSTEM TECHNICAL JOURNAL is abstracted or indexed by *Abstract Journal in Earthquake Engineering*, *Applied Mechanics Review*, *Applied Science & Technology Index*, *Chemical Abstracts*, *Computer Abstracts*, *Current Contents/Engineering, Technology & Applied Sciences*, *Current Index to Statistics*, *Current Papers in Electrical & Electronic Engineering*, *Current Papers on Computers & Control*, *Electronics & Communications Abstracts Journal*, *The Engineering Index*, *International Aerospace Abstracts*, *Journal of Current Laser Abstracts*, *Language and Language Behavior Abstracts*, *Mathematical Reviews*, *Science Abstracts (Series A, Physics Abstracts; Series B, Electrical and Electronic Abstracts; and Series C, Computer & Control Abstracts)*, *Science Citation Index*, *Sociological Abstracts*, *Social Welfare*, *Social Planning and Social Development*, and *Solid State Abstracts Journal*. Reproductions of the Journal by years are available in microform from University Microfilms, 300 N. Zeeb Road, Ann Arbor, Michigan 48106.



Bell System