

# THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XLV

DECEMBER 1966

NUMBER 10

Copyright © 1966, American Telephone and Telegraph Company

## The Effects of Transmission Delay in Four-Wire Teleconferencing\*

By A. G. VARTABEDIAN

(Manuscript received May 4, 1966)

*The effects of transmission delay upon the performance of a three-party teleconference were investigated using a problem-oriented task. The teleconference was simulated in the laboratory using 4-wire telephone sets interconnected to form a three-party conference network without echo sources or echo control devices. The two experimental conditions were characterized by (a) a network whose three legs provided transmission delays of 600-milliseconds, 300-milliseconds and no delay, and (b) a network each of whose three legs provided no transmission delay. It was found that: (i) time to complete the experimental task was 28 percent greater in the delay condition than in the no-delay condition, (ii) the error rate was less in the delay condition than in the no-delay condition, (iii) time-per-trial decreased with successive trials in both delay conditions; time-per-trial was less in the no-delay condition than in the delay condition, and (iv) no chairmanship pattern developed as a result of time delay in the network. Moreover, not one of the subjects reported having observed the existence of delay in the voice path. There were, however, more complaints of "talking together" in the delay condition.*

### I. INTRODUCTION

Two technological developments have recently entered the field of international communication. The first is the advent of the communi-

\* This paper is drawn from the author's thesis which has been accepted by the University of Pennsylvania in partial fulfillment of the requirements for the degree of Master of Science.

cation satellite. The second is the now widespread ability to establish teleconferences (telephone conference calls).<sup>\*</sup> Relatively little is directly known about either of these by itself, let alone how they may interact. This paper represents an attempt to systematically experiment with time-delayed communication in the context of conference telephony. The research examines the influence of the delay factor upon human teleconferencing performance on a given task and also explores the development of group structure in a network having several time delays.

In the present connection, the significant factor resulting from the use of a satellite is the relatively large amount of time delay introduced into the signal transmission. Riesz and Klemmer<sup>1</sup>, in a study of delayed conversation between two people, found that round-trip delays less than 600-milliseconds do not degrade the acceptability of the circuit. Mitchell<sup>2</sup> and Emling and Mitchell<sup>3</sup> have given the significant parameters of time delay for various types of satellites. Low orbit satellites (with a typical round-trip time delay of 100 milliseconds) and medium orbit satellites (with 190 milliseconds) should cause little trouble because of time delay. Hence, this paper investigates the effects on teleconferences resulting from signal delays introduced by a synchronous satellite (with a total round-trip time delay of 540 milliseconds).

We consider first the effects of time delay on the communication process. In a study of the length of reference phrases used in a conversation between two people to describe ambiguous figures over telephone circuits, Krauss and Weinheimer<sup>4, 5</sup> found that the length of reference phrases decreased on each successive occasion the figure was mentioned, eventually reaching a lower limit of one word. In a related but yet unpublished study, they found that when the circuit was degraded by the introduction of voice-operated devices, the mean number of words to describe the ambiguous figure was higher on the first occasion and decreased at a slower rate than when the telephone circuit was of standard quality.

Viewed collectively, these findings tend to indicate that a degraded circuit disrupts the communication process and that learning (evidenced by the decreasing length of reference phrases on successive references) is greatly slowed down. Hence, one might expect that time delay in a teleconference network could also degrade the communication and slow down the learning process.

---

<sup>\*</sup> In the context of this paper, a teleconference is a conference among at least three people using standard telephone sets.

In their study of naturally occurring conversations between two people, Riesz and Klemmer<sup>1</sup> found that subjects did not find round-trip delays of 600 and 1200 milliseconds objectionable. The criterion used to measure "objectionality" in that circuit was the number of times each of the subjects rejected the delayed circuit for a normal one. Riesz and Klemmer chose to deal with naturally occurring speech since other studies have shown that the subtle conversational difficulties produced by delay do not often occur in structured conversations.

As long as one person does all the talking, it is impossible to detect the presence of even very large delays. The natural speech simply arrives a few seconds late. Likewise, the presence of delay will be unnoticed in highly structured conversation, where it is agreed beforehand that one person will not begin talking until the other has stopped. There are simply longer pauses between talk spurts.

Delay plays an increasingly significant role in a nonstructured conversation. When the communication involves elements of information exchange, persuasion, or negotiation, the quality of the communication channel becomes more critical. For instance, in such teleconferenced conversations it is sometimes necessary to cut a person off for questioning or voicing objection. Consider a time-delayed conversation between two people, A and B where  $t$  is the one-way delay of the circuit. When A cuts into B's speech, he is heard by B  $t$  seconds later at which time B stops talking (assuming B has zero reaction time). A, however, continues to hear B for another  $t$  seconds. Hence, for period  $2t$  (the round-trip time delay of the circuit) both people are talking simultaneously. If there are more than two people in the conference and some (or all) of them are talking simultaneously, disorder may result.

An alternative to the naturally occurring speech technique in evaluating a communication system is the use of a problem solving task. The advantage of using a specified task over the naturally occurring speech technique is that the nature of the communication can be more closely controlled. Conceivably, naturally occurring speech could range from idle chit-chat to high-level negotiations with the characteristics of the conversation varying greatly even within a given class of teleconference. Consequently, it is desirable to test the teleconference facility using a task that simulates the pertinent characteristics of anticipated conferences.

We next turn to the aspects of group behavior which apply to teleconferences. Here a number of parameters are of importance such as

nature of conference, size of group, network configuration, and whether or not a chairman is designated. The nature of the communication will be the most important factor governing the proceedings of the conference. A large conference in which one member gives instructions or information to all others will certainly be quite different from a conference in which three of four heads of state attempt to negotiate a settlement to an impending crisis. Consequently, it is difficult to investigate teleconferences without first defining the nature (information exchange, lecture, problem solving, persuasion, negotiation, etc.) of the conference.

The size of the group is a particularly significant parameter in teleconferences. It is of even greater importance than in face-to-face conferences because of the greater difficulty in identifying participants due to the lack of nonverbal cues (Sinaiko<sup>6</sup>). A group size of three is the simplest form of teleconference possible. Here, the possibility of a deadlock is minimized since the minority is an isolated single person.

The third important parameter is the network configuration. Bavelas et al<sup>7</sup> have found that two basic configurations are of importance: one having a common circuit and one having a central control. Common circuit networks are those in which all participants can hear when any one of them talks. Central control networks are those over which all parties transmit to a single station which in turn can relay the message to all others. Note that the central control network lends itself naturally to strong chairmanship; while the chairman in a common circuit network must depend on rules and protocol to maintain order.

In a study on teleconferencing, Heise and Miller<sup>8</sup> found that an information collection task (completing a list of words; each subject having part of the list) was performed most rapidly using a common circuit network while a task requiring assembly plus coordination was performed most rapidly using a central control network with a chairman in charge. They also found that the differences between networks become more pronounced as "noise" is introduced into the communication channel. As "noise" they used white noise.

Sinaiko<sup>6</sup> found that in one teleconference, the chairman did not add to the effectiveness of the meeting because he had no means of enforcing his decisions. He found also that when using four conferees and acceptable circuits it was not difficult to consistently identify each voice. In a large teleconference of 12 participants in which the issue of chairmanship was deliberately left vague, a chairman did seem to emerge. When the conferees were asked whether a chairman emerged they generally agreed that the man who volunteered to call the roll

during the first few minutes of the conference was regarded as the chairman.

In a time-delayed multiparty teleconference in which a chairman has not been designated, it is of interest to determine whether or not any one position in the network would be favored for the emergence of a chairman. Consider the network depicted in Fig. 1 to be used in the present study. The three positions A, B, C are separated from a central point (telephone company conference operator) by one-way delays of 600, 0, and 300 milliseconds, respectively. The psychologically relevant factor here is the round-trip delay between any two given persons in the network.

One might anticipate (Bavelas<sup>9</sup> and Guetzkow and Simon<sup>10</sup>) that the person occupying the position of "relative centrality" would most likely emerge as chairman. The central position is that position which is closer than any other position to all other positions. Hence, using time delay as a measure of distance, this theory predicts that B will emerge as chairman. It is felt, however, that the delays depicted in the figure are insufficient to overshadow those personality factors which are thought to determine the normal development of group structure.

## II. PURPOSE

An experiment was run to investigate the subjective reaction to large time delays (experienced in synchronous satellite communication) in a three-party teleconference. The network simulated a teleconference whose three legs consisted of one satellite link, two satellite links, and no satellite link, respectively. Each satellite link had a one-way delay of 300-milliseconds. The two experimental conditions were specified by:

(a) Delay condition — a network, Fig. 1, in which one leg is delayed

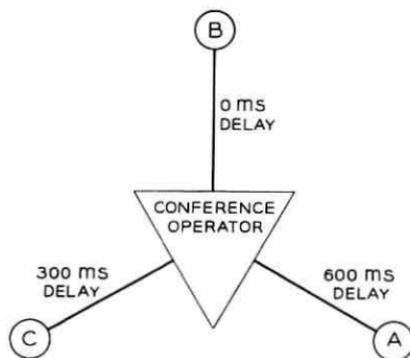


Fig. 1 — Network configuration.

600 milliseconds, another is delayed 300 milliseconds, and the third is undelayed.

(b) No-delay condition — a network in which each of the three legs is undelayed.

All other parameters remained the same between experimental conditions. No attempt was made to simulate echo or the effect of echo suppressors even though they would be expected to play an important role in actual satellite communication.

The two experimental conditions were tested using a task which required the subjects to communicate over the teleconference network about ambiguous figures (in the sense that they do not evoke a common reference phrase). Such ambiguous figures have been used previously by Krauss and Weinheimer to test circuit quality in circuits containing voice switching devices and in circuits containing delay. Subjects were faced with the task of describing the figures appearing on each of their stimulus cards and identifying the figures which were common to all three of them. There were 22 such trials.

With respect to the above defined task and the two experimental conditions it was hypothesized that:

(i) Time-on-task will be greater in the delay condition than in the no-delay condition.

(ii) There will be no difference in accuracy on task between the delay and the no-delay condition.

(iii) Time-per-trial will decrease with successive trials in both delay conditions. For each trial, time-per-trial will be greater in the delay condition than in the no-delay condition.

(iv) No one position in the network is likely to emerge as a seat for a chairman in the delay condition.

### III. METHOD

#### 3.1 Procedure

Upon arrival at the laboratory, subjects were introduced to each other. They were then given verbal instructions by the experimenter. Essentially they were told that they were to participate in a conference call as though each one of them were in a different country and that their voices would be transmitted to one another over simulated satellite circuits. They were told that the purpose of the study was to determine the effects of satellites on telephone communication. Subjects were not told whether or not their network contained delay.

After receiving instructions, subjects entered individual rooms and picked up telephones they found there. They were asked by the experimenter over the telephone to give their name, room number, and a listing of letters that were associated with the task. All subjects could hear the questions and answers. This was to identify the names with the voices and to give the subjects a "feel" for the circuit.

Subjects then began their task. If there were any misunderstandings, they were assisted by the experimenter. This rarely happened beyond the first minute of the task. After completing the task, subjects were individually interviewed by the experimenter and were then told not to discuss the details of the task with their friends since they might also be used as subjects later.

Experimental sessions lasted about 30 minutes and were spaced 45 minutes apart. Each group participated in only one delay condition.

### 3.2 *Experimental Task*

The three subjects, each in an acoustically isolated room, talked together over seemingly normal telephone sets. Before them on a table was a set of 22 cards each numbered, and mounted in a desk calendar holder. The holder kept the cards in order and presented only one card at a time.

Before entering their rooms, the subjects were told that on each card of their set were stamped five nondescript figures. They were told that two of those figures appeared on all of the cards while the other three figures appeared only on their own cards. Fig. 2 shows a sample stimulus card. By describing the figures to each other, they were to find out which were the two figures held in common. After agreeing on the first figure on card 1, each subject read the letter under it. They next located the second figure on that card, and each read the letter under it. They were then told to turn to card 2 and continue.

The two common figures were selected at random (using a table of random numbers) from a set of 11 figures. Any given figure appeared as a common figure twice in cards 1 through 11 and twice in cards 12 through 22. No figure was a common figure on both of two adjacent cards. The common figures were placed at random positions on the cards. Common figures appeared an equal number of times in each position on the cards of each set. The noncommon figures were selected from the remaining nine figures and placed randomly on the cards. The total number of appearances of all figures were equal for each set of cards.

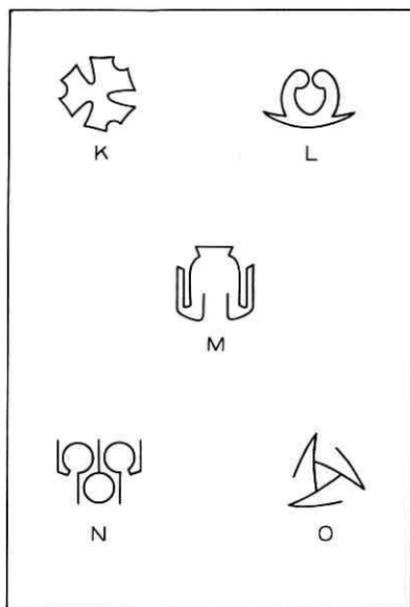


Fig. 2 — Sample stimulus card.

### 3.3 Apparatus

The experiment was run in the Customer Services Appraisal Laboratory of Bell Telephone Laboratories at Holmdel, New Jersey. The laboratory consists of five acoustically isolated rooms designed for psychological testing.

The network configuration for the experiment is given in Fig. 3. When the experiment was run in the no-delay condition, the delay units were physically removed from the circuit. The network employed 4-wire circuits providing isolation between receive and transmit paths. This was necessary to prevent echo and to interface with the magnetic disc delay units (Echo Vox Sr.) which are one-way devices. The delay units were calibrated for 600-milliseconds delay and 300-milliseconds delay, respectively.

Because of the need for 4-wire circuitry, the standard 500-telephone set was modified as in Fig. 4. Artificial sidetone was provided and loss and circuit noise were adjusted to values representative of standard telephone circuits.

The conference bridge was a standard 4-wire 6-branch bridge providing 19.5 dB net loss. Only four of the branches were used in the circuit — three for the participants and one for the monitor. The am-

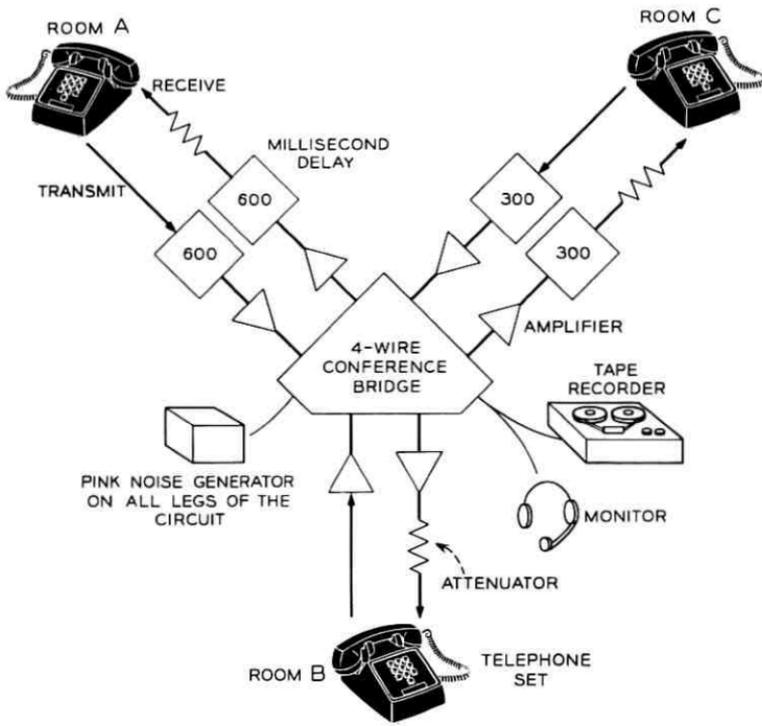


Fig. 3— Experimental configuration.

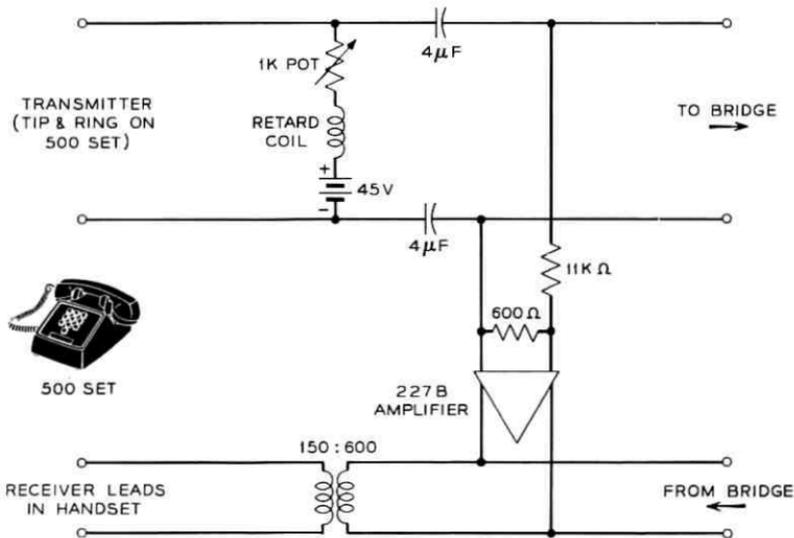


Fig. 4— Modification of 500 type telephone set.

plifiers at the bridge were adjusted to provide unity gain through the bridge in all directions.

### 3.4 Participants

The 36 subjects were employees of Bell Laboratories and were female clerks with roughly the same job classification. Their ages ranged from 18 through 25. All had at least a high school education. The subjects were unacquainted with the experimenter at the outset of the study and were volunteers recruited by members of the Appraisal Laboratory. Most of the subjects were at least vaguely acquainted with the other members of their experimental group.

### 3.5 Scoring

All measurements were made from the recorded tapes of the sessions. In all but one or two instances, the experimenter was able to identify the person talking from the tape. The tapes were first played to make an overall time measurement, to determine which participant spoke first, to determine which participant offered her letter first, to determine which of the two common figures were found first, to detect any errors, and to become familiar with the progress of the session. The tape was then replayed to obtain time-per-figure measurements. In addition, a questionnaire provided subjective data.

## IV. RESULTS

Table I presents the overall time each group spent on the experimental task. The average time to complete the task in the no-delay condition was 13.43 minutes while in the delay condition 17.23 minutes. A  $t$ -test was applied to these data to determine whether or not transmission delay had an effect on time to complete the task. The  $t$ -test ( $t_{10} = 3.81$ ,  $p < 0.01$ ) indicated that time-on-task was significantly greater in the delay condition.

*Hypothesis (ii)* postulated there would be no difference in accuracy on the experimental task between the delay and the no-delay condition. Accuracy was measured by noting every occurrence of an incorrect identification of a figure made by an individual. For any given figure, at most two errors could be made since the person describing the figure always correctly identified it by reciting the letter appearing under that figure. The data on accuracy are summarized in Table I.

It was possible for the individuals in any one group to make a total of 88 errors on the task. Note that the greatest number of errors

TABLE I — TIME-ON-TASK AND TOTAL ERRORS MADE BY INDIVIDUALS FOR EACH GROUP

Group	No-Delay	
	Time-on-Task (min.)	Errors
1	15.37	3
2	14.47	0
3	13.74	6
4	11.97	4
5	14.01	0
6	11.30	0
	Average: 13.48	Total: 13
	Delay	
7	17.43	0
8	18.58	0
9	16.72	1
10	15.07	1
11	20.02	0
12	15.58	0
	Average: 17.23	Total: 2

(made by Group 3) is 7 percent of this maximum. Of a possible total of 528 errors among all six groups in each experimental condition, 13 errors (2.5 percent of total possible) were made in the no-delay condition while two errors (0.4 percent of total possible) were made in the delay condition. A Poisson distribution test on these data indicates that the no-delay error rate is significantly (0.01 level) greater than the delay error rate (see Ref. 11).

*Hypothesis (iii)* stated that time-per-trial would decrease with successive trials in both delay conditions and that for each trial, time-per-trial would be greater in the delay condition than in the no-delay condition. The time measurements per trial are an average of the times for identifying the two figures on each of the 22 stimulus cards. Fig. 5 presents the time-per-trial averaged over the six groups in each circuit condition plotted on a log scale.

A 2-factor (22 trials  $\times$  2 circuit conditions) analysis of variance under a logarithmic transformation of the data was carried out. The logarithmic transformation was employed to reduce heterogeneity of variance and because an exponential-shaped function was found. The logarithms of time-per-trial and trial number fit a least squares linear regression model with 0.9 correlation coefficient.

As can be seen from Fig. 5, the time-per-trial decreases on successive trials for both circuit conditions. The analysis of variance showed

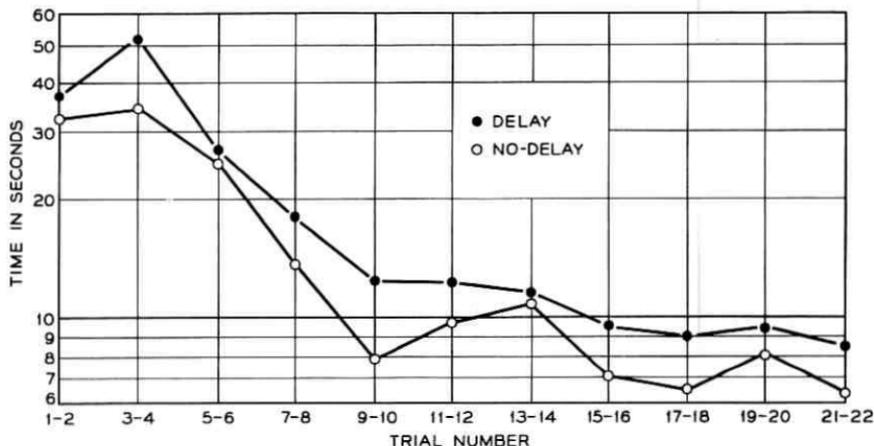


Fig. 5—Average time-per-trial (log scale) vs trial number (grouped 2 trials per data point).

that the difference in trial times was significant ( $F_{21, 21} = 30.8, p < 0.001$ ). Fig. 5 also shows that delay times are greater than no-delay times. The analysis of variance confirms this difference ( $F_{1, 21} = 27.5, p < 0.001$ ).

The sharp rise in Fig. 5 on the third trial is the result of the initial appearance of one ambiguous figure which all groups found exceptionally difficult to locate. Note that the curves begin to decrease much less after the eighth trial. By this time the subjects had converged on a single reference phrase for each figure. The twelfth trial marks the beginning of a random repetition of the first 11 trials. At this point the subjects have seen each of the 11 figures twice and now encounter them for the third time.

*Hypothesis (iv)* stated that no one position in the network is favored to emerge as a seat for a chairman in the delay condition. The hypothesis was tested using measurements of who was the first individual to begin describing a figure for each figure on the set of stimulus cards. This means of judging chairmanship was chosen because, to a rough approximation, chairmanship is determined by leadership which connotes the first to act. Also, during the interview the subjects generally stated that their criterion for judging chairmanship was based on who was the first person to speak up. For each group the number of times the subject in each position was first to describe a figure was noted. There were a total of 44 (the number of common figures on each set of stimulus cards) first responses for each group.

An analysis was carried out to test, on the basis of frequency of first responses, whether a chairman actually emerged in each of the 12 groups. Under the null hypothesis the probability of obtaining a chairman was assumed to be one-third. A chi-squared test was used to test the significance of the difference between the observed frequency and hypothesized frequency of one third. Table II presents the results of the chi-squared test ( $\chi^2$  with 2 *df*,  $p < 0.05$ ) indicating the position of the chairman in each of the 12 groups. From Table II the chi-squared test indicates that no chairmanship patterns are apparent in either circuit condition.

In addition, an analysis of variance of the data on number of first responses according to position in the circuit was made for the delay condition. The test resulted in failure to reject the null hypothesis at the 0.05 level that all positions in the network are equally likely for seating a chairman ( $F_{1,10} = 4.77$  ns).

A *questionnaire* consisting of five questions administered at the conclusion of the experiment provided qualitative data. To the first question, "Did you have any difficulty in understanding the other parties in the conversation?", all subjects in both delay conditions answered "NO". But in response to this question, five subjects in the delay condition added that members of their group often were talking at the same time. Only one subject in the no-delay condition complained of talking together.

TABLE II — INDICATED CHAIRMAN FOR EACH EXPERIMENTAL GROUP

Group	No-Delay	
	Chi-Squared <sup>a</sup>	Questionnaire <sup>b</sup>
1	C	C
2	A	None
3	B	B
4	B	B
5	None	C
6	A	A
	Delay	
7	C	None
8	C	None
9	None	A
10	B	C
11	B	A
12	C	C

<sup>a</sup> Chairman indicated by  $\chi^2$ -test on first person to respond measurements.

<sup>b</sup> Chairman indicated by majority decision rule on questionnaire responses.

The responses to question 2, "Did you find that a conference chairman seemed to emerge during the conversation?", are summarized in Table II. If at least two subjects of each group said that the same person was chairman, then that person is designated chairman of the group; otherwise no chairman is indicated.

To question 3, "Did you notice any difference between this circuit and the one you normally use here at the Labs?", 10 of the 18 people in the no-delay condition responded "NO". The remaining eight subjects who answered "YES" responded that the circuit was either "fuzzier", "clearer", or "farther" than their normal circuit. In the delay condition all but four subjects responded "YES" to this question. They commented that the circuit was either "fuzzier", "clearer", "nearer", or "farther" than their normal circuit. Of the 18 subjects in the delay condition, three in Position A, four in Position B, and one in Position C reported having observed a faint echo. The inadvertent echo was due to insufficient attenuation at the conference bridge.

To question 4, "Were you able to identify the other parties from their voices in the conversation?", all subjects in both circuit conditions responded "YES".

To question 5, "Do you have any other comments?", eight persons in each circuit condition responded "That was fun", indicating perhaps that delay can be as much fun as no-delay.

## V. DISCUSSION

As was hypothesized, the overall time to complete the experimental task was greater with time delay in the network. The average value of 17.23 minutes for the delayed circuit is 28 percent greater than the average value of 13.48 minutes for the nondelayed circuit. These overall average time values are composed of the 22 individual trial time measurements. As was cited in the previous section, these trial times and the trial numbers, both under a logarithmic transformation, were found to fit a straight line with rather high correlation, thus substantiating the exponential shaped function. The existence of the exponential function would tend to indicate an underlying model that is multiplicative rather than additive. What is meant by this is that any difference in no-delay and delay times would be properly expressed as a ratio rather than an additive difference. During any given trial the control of the circuit (as evidenced by the person talking) transferred many times among the three subjects. It is this behavior which is thought to be the cause of the multiplicative relationship between delay and no-delay trial times.

Although no quantitative data were collected on the components of trial time measurements, the experimenter observed a greater redundancy in communication with the presence of delay. More time was spent in giving descriptions and feeding back confirming information than in the no-delay case. There were also more questions asked of the person describing the figure. Descriptions did not seem to be any longer — there were simply more of them. In addition to greater redundancy, more time was spent in pauses and attempts to speak in the delay case than in the no-delay case. In answer to the first question on the questionnaire, five times as many subjects found difficulty because of talking together in the delay condition than in the no-delay condition.

The increased time-on-task in the delay condition may be thought of as a compensation for the greater difficulties encountered with the delayed network. The data on errors indicate that the increased amount of time-on-task experienced with the delayed circuit induced more reliable communication. The error rate of 2.5 percent for the no-delay network is significantly greater than 0.4 percent for the delayed network.

As hypothesized, the time-per-trial decreased with successive trials for each circuit condition. These time measurements include total elapsed time to select a figure, describe it, entertain all questions and comments, agree that the common figure has been located, and finally read off the letters that appear under the figure on the respective stimulus cards. The decreasing time-per-trial was observed to be a result of decreasing length of reference phrases for the figures as well as agreement among all three participants on a single reference phrase for each figure. A typical sequence of phrases for the ambiguous figure appearing above the letter O on the sample stimulus card (Fig. 2) is "... looks like three Vs connected in the middle with a triangle in the center," "... three Vs with the triangle in the center," "... three Vs."

It is clear from Fig. 5 and supported by the analysis that for each trial, trial times are less under the no-delay condition than the delay condition. Furthermore, Fig. 5 indicates that no-delay trial times drop sooner to a lower level than delay trial times. This graph would tend to indicate, as thought earlier, that learning is impaired by the introduction of time delay into the circuit. This, however, cannot be supported on the basis of data obtained in the present study.

As hypothesized, the delayed circuit imposed no tendency for chairmanship patterns to develop.

Finally, a noteworthy result of the experiment (from the questionnaire responses) was the fact that not one of the participants reported having observed the presence of a time-delayed circuit. Some com-

plained of increased "talking together" but this led no one to guess the underlying reason.

The results obtained here were based upon 4-wire circuits without echo sources or echo control devices. The possibility of echo and speech mutilation due to echo-suppressor action in commercial 2-wire circuits could add degradation beyond that encountered in this experiment.

#### VI. ACKNOWLEDGMENT

The author wishes to acknowledge the many helpful suggestions and invaluable guidance of his thesis supervisor Professor George E. Rowland of the Moore School of Electrical Engineering during the course of this study.

#### REFERENCES

1. Riesz, R. R. and Klemmer, E. T., Subjective Evaluation of Delay and Echo Suppressors in Telephone Communication, B.S.T.J., 42, 1963, pp. 2919-2941.
2. Mitchell, D., Orbiting Satellites for Data Transmission, Bell Laboratories Record, 43, 1965, pp. 16-23.
3. Emling, J. W. and Mitchell, D., The Effects of Time Delay and Echoes on Telephone Conversations, B.S.T.J., 42, 1963, pp. 2869-2891.
4. Krauss, R. M. and Weinheimer, S., Changes in Reference Phrases as a Function of Frequency of Usage in Social Interaction, Psychonomic Sci., 1, 1964, pp. 113-114.
5. Krauss, R. M. and Weinheimer, S., The Effect of Feedback on Changes in Reference Phrases, a Paper read at the annual meeting of the Psychonomic Society, Niagara Falls, Ontario, Canada, October 10, 1964.
6. Sinaiko, H. W., *Teleconferencing: Preliminary Experiments*, Arlington, Virginia, Institute For Defense Analyses, Research Paper P-108, 1963, pp. 15-17 and pp. 33-36.
7. Bavelas, A., et al, *Teleconferencing: Summary of a Preliminary Research Project*, Arlington, Virginia, Institute For Defense Analyses, Study S-138, 1964, pp. 6-8.
8. Heise, G. A. and Miller, G. A., Problem Solving by Small Groups Using Various Communication Nets, J. Abnorm. Soc. Psychol., 46, 1951, pp. 327-336.
9. Bavelas, A., Communication Patterns in Task Oriented Groups, J. Acoust. Soc. Amer., 22, 1950, pp. 725-730.
10. Guetzkow, H. and Simon, H. A., The Impact of Certain Communication Nets Upon Organization and Performance in Task-Oriented Groups, Manage. Sci., 1, 1955, pp. 233-250.
11. Pearson, E. S. and Hartley, H. O., Eds., *Biometrika Tables for Statisticians*, Cambridge, Cambridge University Press, 1958, p. 185.
12. Aircraft Armaments, Incorporated, *Teleconferencing: An Experimental Task*, Arlington, Virginia, Institute for Defense Analyses, Research Paper P-112, 1963, pp. 1-2.
13. Bales, R. F. and Borgatta, E. F., A Study of Group Size: Size of Group as a Factor in the Interaction Profile, In Hare, P. A., Borgatta, E. F., and Bales, R. F., (Eds), *Small Groups: Studies in Social Interaction*, New York, Knopf, 1955.
14. Hare, P. A., *Handbook of Small Group Research*, The Free Press of Glencoe, New York, 1962, pp. 208-287.

# Mutual Synchronization of Geographically Separated Oscillators

By A. GERSHO and B. J. KARAFIN

(Manuscript received July 27, 1966)

*A control scheme for synchronizing the frequencies of geographically separated oscillators connected by communication links consists of averaging the phases received at each station from remote oscillators, comparing the result with the local phase, and applying the filtered error signal as a correction to the local oscillator frequency. The system was studied by V. E. Beneš who found a sufficient condition for the stability of the system using advanced mathematical techniques. In this paper, the stability condition is derived (for a slightly more general control scheme) using only the transfer function concept of linear systems and some properties of determinants. A practical difficulty regarding the final frequency of the oscillators is discussed and a modification of the control scheme is shown to alleviate the difficulty. Also examined are the questions of sensitivity to parameter changes, the effect of jitter noise on the performance of the system, and the effect of failure of an oscillator or transmission link.*

## 1. INTRODUCTION

Consider a network of  $N$  geographically separated stations that are connected by directed communication links. A local clock, or oscillator, is situated at each station. The problem of synchronizing the frequencies of the oscillators is of considerable practical interest for continental pulse code modulation (PCM) systems.

The local oscillators have frequencies which may be altered in proportion to a control signal. In the absence of external control, each oscillator operates at a different frequency. The network is "connected" in the sense that from any station to any other station there is either a direct transmission link or an indirect path via one or more intermediate stations. A fixed time delay is associated with each transmission link.

In an important but unpublished paper, V. E. Beneš<sup>1</sup> has examined a linear control scheme in which each station receives the phases of

neighboring stations, i.e., those stations connected to it by direct transmission links. The phases are averaged and compared with the local phase; the error is filtered and applied as a correction to the frequency of the oscillator. Similar schemes were also proposed by Runyon.<sup>2</sup> Beneš has proved that under suitable conditions the system is stable, i.e., the oscillators asymptotically settle to a common frequency and the phase differences have finite asymptotic values. He also finds explicit formulae for the final frequency and asymptotic phase differences. To obtain these results, he resorted to the mathematical techniques of renewal theory and Tauberian theory. By assuming the stability of the system, as proved by Beneš, A. J. Goldstein<sup>3</sup> has rederived the expressions for final frequency and phase differences in a more direct manner. Bonomi, La Marche, and Varaiya<sup>4</sup> improved the treatment of the stability problem and suggested some avenues of approach for the study of transient response. In each case, the authors relied on the mathematical theory of Markoff chains and stochastic matrices.

M. Karnaugh<sup>5</sup> has formulated a more realistic and more sophisticated nonlinear control model. Broad stability conditions for this model are not yet known; however, certain special cases resemble the Beneš model.

In this paper, the stability conditions and the expression for final frequency for a slightly more general version of the Beneš model are derived in a simple manner using only the transfer function concept of linear systems and elementary properties of determinants. This approach permits a clearer intuitive understanding and should be readily comprehensible to the non-mathematician. The sensitivity of the system to parameter changes is also examined and certain questions regarding the final frequency of the oscillators are clarified.

In Section II we give a formulation of the problem and obtain the basic equations describing the system. In Section III certain crucial properties of the matrix of averaging coefficients are derived which result from the topological constraint that the network is connected. Stability is proved in Section IV and an expression for the final frequency is obtained. Section V considers some practical questions with regard to how the final frequency is related to the free-running frequencies of the oscillators. Section VI examines the questions of sensitivity to parameter changes, the effect of failure of an oscillator or transmission link, and the effect of jitter noise.

## II. FORMULATION

Let  $f_i$  be the frequency of the  $i$ th oscillator in the absence of external control, and  $r_i(t)$  the control signal applied to the  $i$ th oscillator at time  $t$ .

If  $p_i(t)$  denotes the total cyclical phase of the  $i$ th oscillator, then the actual frequency at time  $t$  is given by

$$\dot{p}_i(t) = f_i + r_i(t) \quad (1)$$

where the dot denotes the time derivative.

The control scheme at the  $i$ th oscillator is shown in Fig. 1. The phases of all neighboring stations are transmitted to the  $i$ th station. The transmission delay associated with the path from station  $j$  to station  $i$  is denoted as  $\tau_{ij}$ . Each phase received at station  $i$  is compared with the local phase; the differences are weighted with the nonnegative averaging coefficients  $a_{ij}$  and summed. The weighted sum of phase

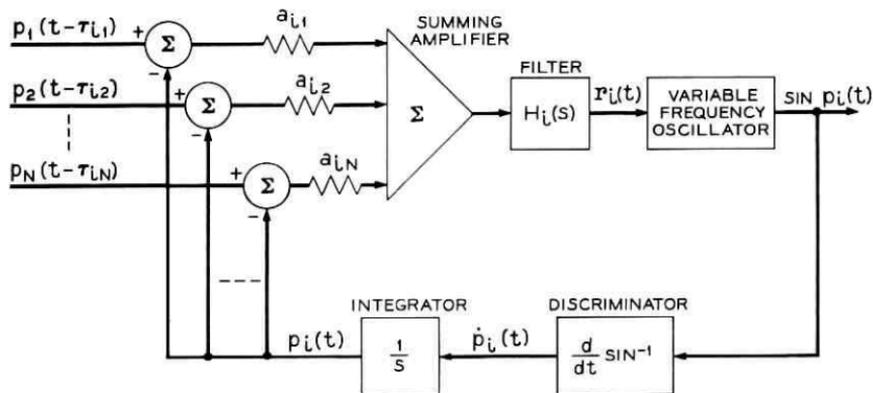


Fig. 1 — Station  $i$  of the phase averaging control system.

differences is applied to the filter with transfer function  $H_i(s)$ , and the filter output is the frequency correction term  $r_i(t)$ . Thus, we have

$$r_i(t) = h_i(t) * \sum_{j=1}^N a_{ij}[p_j(t - \tau_{ij}) - p_i(t)], \quad (2)$$

where  $h_i(t)$  is the impulse response of the filter at the  $i$ th station and the asterisk denotes convolution.

We assume the filters have three simple properties: (i) causality, i.e., the response at any instant does not depend on the future of the input, (ii) stability, in the sense that a bounded input always produces a bounded output, and (iii) positive dc gain, i.e.,

$$H_i(0) \equiv \lambda_i > 0. \quad (3)$$

Without loss in generality we may assume that the averaging coefficients sum to unity, i.e.,

$$\sum_{j=1}^N a_{ij} = 1. \quad (4)$$

Clearly a scaling of all the coefficients  $a_{ij}$  for fixed  $i$  is equivalent to a change in gain factor of the  $i$ th filter.

If there is no direct transmission path from the  $j$ th to the  $i$ th station, then the coefficient  $a_{ij}$  is presumed to be zero. Thus, the  $N \times N$  matrix,  $A$ , whose  $ij$ th element is  $a_{ij}$ , contains all the topological information about the network of communication links. In order that mutual synchronization be possible, it is certainly necessary that the network be connected so that from any station to any other station there is either a direct or indirect transmission path. The resulting properties of the averaging matrix  $A$  imposed by this topological constraint play a vital role in the proof of stability for the system.

In agreement with the Beneš model we consider the starting conditions where the oscillators are assumed to have been free-running for an indefinitely long time prior to  $t = 0$ , and at  $t = 0$  the control signals  $r_i$  are connected to the oscillators. Thus, we have

$$p_i(t) = f_i t + p_i(0), \quad t < 0 \quad (5)$$

where  $p_i(0)$  is the phase at  $t = 0$ , and from (1), (2), and (4), the frequency of the  $i$ th oscillator when the controls are operating is

$$\dot{p}_i(t) = h_i(t) * \sum_j a_{ij} [p_j(t - \tau_{ij}) - p_i(t)] + f_i, \quad t \geq 0. \quad (6)$$

Equations (5) and (6) for  $i = 1, 2, \dots, N$  completely characterize the behavior of the system under the particular starting conditions of interest. Taking the ordinary Laplace transform of (6), we obtain

$$sP_i = H_i \sum_j \hat{a}_{ij} P_j - H_i P_i + \frac{1}{s} f_i + p_i(0) + Q_i, \quad (7)$$

where

$$\hat{a}_{ij} = a_{ij} \exp(-s\tau_{ij}),$$

$$Q_i(s) = H_i(s) \sum_j \hat{a}_{ij} \int_{-\tau_{ij}}^0 p_j(t) \exp(-st) dt,$$

and  $P_i(s)$  is the Laplace transform of  $p_i(t)$ . The term  $Q_i(s)$  is the contribution to the  $i$ th oscillator frequency after  $t = 0$  due to the contents of the transmission links at  $t = 0$ . Using (5),  $Q_i(s)$  can be evaluated explicitly, but for our purposes it is sufficient to note that

$$sQ_i(s) \rightarrow 0 \quad \text{as} \quad s \rightarrow 0. \quad (8)$$

The transformed equations (7) can, in principle, be solved for the phases  $p_i(t)$  for  $t \geq 0$ . The desired stability information can be obtained

directly from these equations. We shall, however, obtain this information in a somewhat indirect but more profitable way by defining an associated linear time-invariant system with  $N$  inputs and  $N$  outputs.

Consider the same control arrangement for the  $N$  interconnected oscillators. Instead of the former starting conditions, suppose the control paths have always been connected and that each oscillator can be activated by an arbitrary frequency "input" as shown in Fig. 2. Then the actual frequency of the  $i$ th oscillator at time  $t$  is the sum of the basic frequency input  $v_i(t)$  and the correction component  $r_i(t)$  leaving the filter. The phases  $p_i(t)$  are considered the "outputs" of the linear system. When  $v_i(t) \equiv 0$  for each  $i$ , the system is at rest and all outputs  $p_i(t)$  are identically zero.

The importance of the associated linear system is that any desired starting conditions in the physical model can be treated by an equivalent set of inputs to the linear system. To clarify this, note that the system of Fig. 2 is characterized by the equations

$$\dot{p}_i(t) = h_i(t) * \sum_j a_{ij} [p_j(t - \tau_{ij}) - p_i(t)] + v_i(t), \quad -\infty < t < \infty. \quad (9)$$

Formally taking the exponential (two-sided Laplace) transform of (9) we obtain

$$sP_i = H_i \sum_j \hat{a}_{ij} P_j - H_i P_i + V_i, \quad (10)$$

where  $P_i(s)$  and  $V_i(s)$  are, respectively, the exponential transforms of  $p_i(t)$  and  $v_i(t)$ . Equation (10) implicitly characterizes the associated linear system whose inputs are  $v_i(t)$  and outputs  $p_i(t)$  as long as  $v_i(t)$  has an exponential transform. Comparing (7) and (10) we see that the

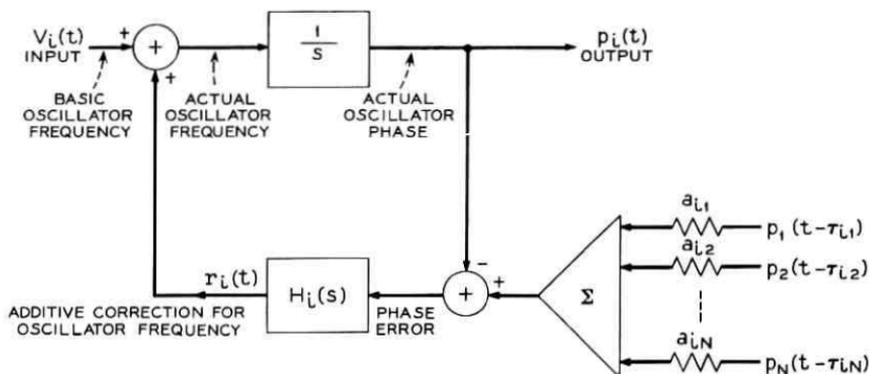


Fig. 2—Model of station  $i$  of the associated linear system.

phase responses of the physical model with the corresponding starting conditions will be the outputs of the associated linear system for  $t \geq 0$  if we select the inputs to be

$$V_i(s) = \frac{1}{s} f_i + p_i(0) + Q_i(s). \quad (11)$$

In the time domain, these inputs are

$$v_i(t) = f_i u(t) + p_i(0) \delta(t) + q_i(t), \quad (12)$$

where  $q_i(t)$  is the response of the filter  $H_i(s)$  to a time-limited input which begins at time  $t = -\max \tau_{ij}$  and ending at  $t = 0$ ,  $\delta(t)$  is the unit impulse function and  $u(t)$  is the unit step function. From (8) and the final value theorem it follows that  $q_i(t) \rightarrow 0$  as  $t \rightarrow \infty$ . It is important to note that the phase responses to the inputs (12) will be the same as the phase responses of the physical model only for  $t \geq 0$ . For  $t < 0$  the responses of the associated linear system do not correspond to the physical model.

Equation (10) may be expressed in the form

$$P_i = \beta_i(s) \sum_{j=1}^N a_{ij} P_j + \left( \frac{1}{s + H_i} \right) V_i, \quad (13)$$

where

$$\beta_i(s) = \frac{H_i(s)}{s + H_i(s)}. \quad (14)$$

The simplified model of the linear system, corresponding to (13), is shown in Fig. 3 where  $\beta_i(s)$  is the transfer function of the feedback configuration as shown. Thus, the operation of the  $i$ th station is to average the incoming phases, apply the average to the filter  $\beta_i(s)$ ,

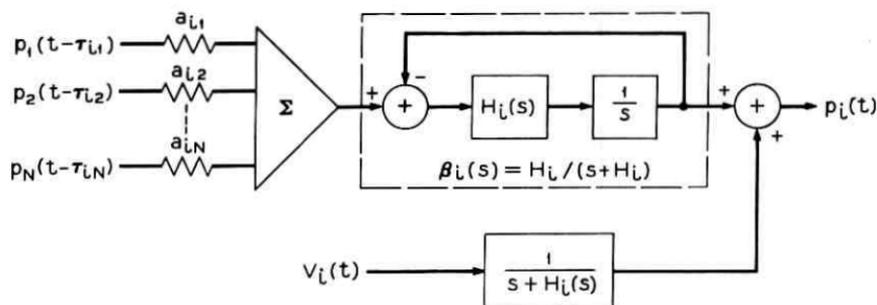


Fig. 3.—Simplified model of station  $i$  for the associated linear system.

and finally add a filtered input component to produce the phase response  $p_i(t)$ . We shall see in Section IV that the condition for stability of the system is simply that the filter  $\beta_i(s)$  have a gain less than unity for sinusoidal inputs.

Equations (13) for  $i = 1, 2, \dots, N$  can be formally solved for the phase responses with the help of matrix notation. Let  $B(s)$  be the  $N \times N$  matrix whose  $ij$ th component is

$$b_{ij}(s) = \beta_i(s) \hat{a}_{ij}$$

and let  $C(s)$  be the diagonal  $N \times N$  matrix whose  $ii$ th element is

$$c_{ii}(s) = \frac{1}{s + H_i(s)}.$$

Note that for  $s = 0$  we have  $c_{ii}(0) = 1/\lambda_i$  and  $B(0) = A$ , where  $\lambda_i$  is the dc gain of  $H_i(s)$  and  $A$  is the averaging matrix, both defined earlier. Let  $P(s)$  and  $V(s)$  be the  $N$  component column matrices whose  $i$ th elements are, respectively,  $P_i(s)$  and  $V_i(s)$ . Then (13) becomes

$$[I - B]P = CV \quad (15)$$

or

$$P = KV, \quad (16)$$

where

$$K(s) = [I - B(s)]^{-1} C(s) \quad (17)$$

is the matrix transfer function of the linear system. Thus, each element  $K_{ij}(s)$  of  $K(s)$  is the scalar transfer function relating the output  $p_i(t)$  to the input  $v_j(t)$  when all other inputs are zero. In Section IV we shall determine certain key properties of the singularities of  $K_{ij}(s)$ . In order to examine the behavior of  $I - B(s)$  in the neighborhood of  $s = 0$ , certain important properties of the averaging matrix  $A$  will be needed. In the next section these properties are derived.

### III. PROPERTIES OF THE AVERAGING MATRIX

As a result of (4), the averaging matrix  $A$  has row sums equal to one. From the requirement that the network be connected, certain restrictions are placed on which combinations of elements of  $A$  may be zero. These two characteristics of  $A$  imply certain essential properties of the matrix  $I - A$  where  $I$  is the identity matrix.

*Theorem 1: If  $A$  is the averaging matrix of a connected network of  $N$  stations then the matrix  $I - A$  has rank  $N - 1$ .*

*Proof:* Since the equation

$$(I - A)x = 0 \quad (18)$$

is satisfied by any vector  $x$  with all components equal, the matrix  $I - A$  is singular and must, therefore, have rank less than  $N$ . Suppose that its rank is less than  $N - 1$ . Then (18) has at least two nontrivial solutions that are linearly independent. Therefore, there exists a nontrivial solution,  $u$ , whose components are not all equal. Now let  $w$  be the solution vector with each component equal to the negative of the smallest component of  $u$ . Then  $y = u + w$  is a nontrivial solution of (18) with all components non-negative and at least one component equal to zero. Let  $\alpha = \{i_1, i_2, \dots, i_k\}$  be the set of indices for which  $y_i = 0$  and  $\beta = \{i_{k+1}, i_{k+2}, \dots, i_N\}$  the set of indices for which  $y_i$  is positive. Since  $y$  satisfies (18) we have

$$y_i - \sum_{j=1}^N a_{ij}y_j = 0 \quad i = 1, 2, \dots, N$$

and so

$$\sum_{j \in \beta} a_{ij}y_j = 0, \quad \text{for } i \in \alpha.$$

But this is only possible if

$$a_{ij} = 0, \quad i \in \alpha \quad \text{and} \quad j \in \beta,$$

which implies that there is no transmission path from any station with identifying index in  $\beta$  to any station with identifying index in  $\alpha$ . Consequently, the network of  $N$  stations is not connected, which is a contradiction. Therefore,  $I - A$  must have rank  $N - 1$ .

*Theorem 2: If  $A$  is the averaging matrix of a connected network, the cofactors of all the elements in any given row of  $I - A$  are equal and positive. Specifically, if  $M_{ij}$  is the cofactor of the  $ij$ th element of  $I - A$ , then*

$$M_{ij} = M_{ik} > 0$$

for  $i, j, k = 1, 2, \dots, N$ .

*Proof:* Since  $I - A$  has rank  $N - 1$ , the solutions of

$$(I - A)y = 0$$

satisfy<sup>6</sup>

$$\frac{y_j}{y_k} = \frac{M_{ij}}{M_{ik}}, \quad i = 1, 2, \dots, N.$$

But the only solutions  $y$  are those with all components equal. Therefore,

$$M_{ij} = M_{ik} \quad \text{all } i, j, k. \quad (19)$$

Let  $R(\epsilon) = I - \epsilon A$  and let  $M_{ij}(\epsilon)$  be the cofactor of the  $ij$ th element of  $R(\epsilon)$ . For  $0 \leq \epsilon < 1$ , each principal minor of  $R(\epsilon)$  is the determinant of a diagonally dominated matrix (see Appendix), so that

$$M_{ii}(\epsilon) \neq 0 \quad 0 \leq \epsilon < 1.$$

Since  $M_{ii}(0) = 1$ , it follows by continuity that  $M_{ii}(1) \geq 0$ . Hence, from (19)

$$M_{ij} = M_{ik} \geq 0 \quad \text{all } i, j, k. \quad (20)$$

Now  $(I - A)'$ , where the prime denotes the transpose, must also have rank  $N - 1$ . Thus, solutions of

$$(I - A)' z = 0 \quad (21)$$

satisfy

$$\frac{z_j}{z_k} = \frac{M_{ji}}{M_{ki}} \quad i = 1, 2, \dots, N. \quad (22)$$

Equations (20) and (22) imply that the nonzero components of  $z$  must have the same sign. Suppose a solution  $z$  of (21) has at least one component zero and nonzero components positive. Then the same argument used in Theorem 1 leads to the conclusion that the network is disconnected, which is a contradiction. Therefore, there is a solution  $z$  with all components positive and consequently (22) implies that all cofactors  $M_{ij}$  are positive, which completes the proof.

#### IV. ANALYSIS

With the help of the preceding results, we are now in a position to prove stability and determine the expression for final frequency. These results will be obtained under the assumption that  $\beta_i(s)$ , for each  $i$ , satisfies the condition

$$|\beta_i(j\omega)| < 1, \quad \omega \neq 0. \quad (23)$$

In Appendix B we show, with the help of the Nyquist criterion, that condition (23) implies the stronger condition

$$|\beta_i(s)| < 1 \text{ for } s \text{ in } \mathfrak{R}, \quad (24)$$

where  $\mathfrak{R}$  is the right half and imaginary axis of the  $s$  plane excluding the point  $s = 0$ .

In Section II we saw that the associated linear system is characterized by the matrix transfer function  $K(s)$  given by

$$K(s) = [I - B(s)]^{-1} C(s). \quad (25)$$

Now, since  $\beta_i = H_i/(s + H_i)$ , it follows from (24) that

$$c_{ii}(s) = 1/(s + H_i)$$

has no singularities in  $\mathfrak{R}$ . Furthermore, under condition (24) the matrix  $I - B(s)$  is diagonally dominated (see Appendix A) for all  $s$  in  $\mathfrak{R}$ . Thus, the determinant  $|I - B(s)|$  is nonzero for all  $s$  in  $\mathfrak{R}$ , and so we conclude that each component transfer function  $K_{ij}(s)$  is analytic in the region  $\mathfrak{R}$ .

At  $s = 0$ , the matrix  $I - B(s)$  reduces to  $I - A$  which is singular according to Theorem 1. Thus, the determinant  $|I - B(s)|$  has a zero at  $s = 0$ . To show that it is only a simple zero we find an asymptotic expression\* for the determinant in the neighborhood of  $s = 0$ . In the matrix  $I - B(s)$ , we replace the elements  $b_{ij}(s)$  by their asymptotic expressions

$$b_{ij}(s) \sim a_{ij} \left[ 1 - \left( \tau_{ij} + \frac{1}{\lambda_i} \right) s \right], \quad s \rightarrow 0$$

where we have used the relations  $\exp(-s\tau) \sim 1 - s\tau$  and  $H_i/(s + H_i) \sim 1 - s/\lambda_i$ . Without changing the value of the determinant, we may replace the first column by the sum of all the columns. The first column then becomes

$$s \left( \tau_1 + \frac{1}{\lambda_1} \right), s \left( \tau_2 + \frac{1}{\lambda_2} \right), \dots, s \left( \tau_n + \frac{1}{\lambda_n} \right),$$

where

$$\tau_i = \sum_{j=1}^N a_{ij} \tau_{ij} \quad (26)$$

is an average of the transmission delays of links arriving at the  $i$ th

\* The technique for finding the asymptotic expression is due to A. J. Goldstein.

station. Now we expand the determinant about the first column and obtain

$$|I - B(s)| \sim s \sum_{i=1}^N \left( \tau_i + \frac{1}{\lambda_i} \right) M_{i1}, \quad s \rightarrow 0, \quad (27)$$

where  $M_{ij}$  is the cofactor of the  $ij$ th element of  $I - B(0)$ , as defined in Section III.

Since  $M_{i1}$  is positive (from Theorem 2), it follows from (27) that  $|I - B(s)|$  has only a simple zero at  $s = 0$ . But  $c_{ii}(0) = 1/\lambda_i$  is finite, so that from (25) we conclude that each  $K_{ij}(s)$  has a simple pole at  $s = 0$ . Using (25) and the asymptotic expression (27) it follows that

$$K_{ij}(s) \sim \gamma_j/s, \quad s \rightarrow 0, \quad (28)$$

where

$$\gamma_j = \frac{M_{ji}/\lambda_j}{\sum_l \left( \tau_l + \frac{1}{\lambda_l} \right) M_{l1}}$$

Note that  $\gamma_j$  is positive and independent of  $i$  because  $M_{ji} = M_{j1} > 0$ , according to Theorem 2. Thus, letting

$$d_j = \frac{M_{j1}/\lambda_j}{\sum_l M_{l1}/\lambda_l}, \quad (29)$$

we have

$$\gamma_j = \frac{d_j}{1 + \sum_l \tau_l d_l \lambda_l} \quad (30)$$

with  $0 < d_j < 1$  and  $\sum d_j = 1$ .

We have, therefore, shown that each transfer function  $K_{ij}(s)$  is analytic in the right half and on the imaginary axis of the  $s$  plane except at  $s = 0$  where it has a simple pole with positive residue independent of  $i$ . The impulse response  $k_{ij}(t)$ , associated with  $K_{ij}(s)$ , will, therefore, consist of exponentially decaying sinusoids and a step function of height  $\gamma_j$ .

To determine the stability of the original model under the particular starting conditions, we examine the asymptotic behavior of the phase responses of the associated linear system when subjected to the inputs given by (11). From (16), we have

$$P_i(s) = \sum_{j=1}^N K_{ij}(s) \left[ \frac{1}{s} f_j + p_j(0) + Q_j(s) \right]. \quad (31)$$

From the known properties of  $K_{ij}(s)$ , it follows that the phase response  $p_i(t)$  for  $t > 0$  will be the sum of terms decaying exponentially to zero plus a term of the form  $ft + \eta_i$  where  $f$  and  $\eta_i$  are obtained by the residue theorem according to

$$f = \lim_{s \rightarrow 0} s^2 P_i(s)$$

and

$$\eta_i = \lim_{s \rightarrow 0} \frac{d}{ds} [s^2 P_i(s)].$$

The final frequency  $f$  of the  $i$ th oscillator is, therefore, given by

$$f = \sum_{j=1}^N \gamma_j f_j \quad (32)$$

which is independent of  $i$ . Thus, we have proved stability of the system since the frequency of each oscillator has been shown to asymptotically approach the common frequency  $f$  and the phase differences clearly approach finite values. From (30) the expression for the final frequency can be written as

$$f = \frac{\sum_j d_j f_j}{1 + \sum_i \tau_i d_i \lambda_i} \quad (33)$$

which, with the help of (29), shows the dependence of  $f$  on the delays  $\tau_{ij}$  and the dc gains  $\lambda_i$ .

#### V. REMARKS ON THE FINAL FREQUENCY

From the results of the preceding section it is clear that the final frequency can be below even the lowest oscillator free-running frequency. In fact, it is evident from (33) that the final frequency is a monotonically decreasing function of the system gain-delay products. Thus, the controls may bring the system to a frequency outside its practical operating range.

The final frequency reduction is a consequence of the fact that the frequency control of each station varies directly with the differences of total phase. The interstation delays introduce phase lags which drive down the frequency of each station. This point is made somewhat clearer by considering a system in which all the oscillators have the same frequency  $f$  and the same initial phase. When the controls are

applied at  $t = 0$  an average phase "error" ( $-f\tau_i$ ) is applied to the control path of each oscillator  $i$ . This "error" causes a simultaneous reduction in the oscillator frequencies from which the system never completely recovers.

As a conceptual solution to this difficulty, suppose the system of Fig. 1 is modified so that the local phase at the output of the integrator is passed through a delay line before being compared with the incoming phases. This local delay at station  $i$  is chosen equal to  $\tau_i$ , the average delay of links terminating at station  $i$ , as defined in (26).

In the previous model, the error signal was determined by a comparison of the local phase at the present time with the remote phases of earlier times. In this modified system, however, the comparison is made between phases which on the average occur at the same time. Thus, the undesired component of the error signal due to interstation delays is eliminated.

Using an argument which parallels the development of Sections II, III, and IV, it may readily be shown that the final frequency for the modified system is given by

$$f = \sum_j d_j f_j. \quad (34)$$

In contrast with the original system, it is evident that the final frequency of the modified system is always an average of the free-running frequencies.

The Beneš formulation (Fig. 1) may be viewed as a simplified abstraction of the more complex practical systems that have been proposed.<sup>2,5</sup> Both the Beneš formulation and the modified system contain a total phase comparator which is an impractical element. Karnaugh<sup>5</sup> has shown that an important linear subclass of the more realistic class of systems he has proposed obey equations of the same form (6) as in the Beneš model. This more realistic formulation also fits the linear system model with modified frequency "inputs" that depend in a different manner on the initial conditions. It is, therefore, subject to the stability condition (23).

Moreover, it has a different final frequency which approaches an average of the free-running frequencies as the interstation delays become large.

In short, the Beneš formulation was sufficient to provide the important stability criterion, but neglected factors affecting the final frequency. The linear system model developed here is general enough to be applicable to both the Beneš formulation and a linear subclass of the more realistic Karnaugh formulation.

## VI. SENSITIVITY AND RELATED QUESTIONS

Suppose that the system has been operating in a synchronized steady-state condition for a long time, and at some instant, say  $t = t_0$ , a sudden change is made in one or more parameters of the system. The subsequent phase responses are determined by considering the new associated linear system subjected to suitable inputs equivalent to the pertinent starting conditions. These inputs will have exactly the same form as (11) but now the term  $Q_i(s)$  will be evaluated using the past history of the phases given by

$$p_i(t) = f_s t + \eta_i \quad t < t_0,$$

where  $f_s$  is the synchronous frequency prior to the parameter change. If the stability condition (23) is satisfied for the new system and if the parameter change does not reduce any  $\lambda_i$  or  $a_{ij}$  to zero, the new system will also be stable. Consequently, after  $t = t_0$  the frequencies of the oscillators will asymptotically resynchronize to the new final frequency determined by (33) or (34) using the changed parameter values. From these arguments we can also deduce that the effect of a slowly time-varying parameter on the system operation is to cause a corresponding slowly varying synchronous frequency. By "slow" time variations we mean that the time for a noticeable change in a parameter value to occur is much longer than the time constants associated with the transient response of each  $K_{ij}(s)$ .

By similar arguments, it is easily seen that failure of a transmission link will lead to resynchronization if the remaining network is still connected. Also, in the case of oscillator failure, the remaining  $N - 1$  oscillators will resynchronize to a new frequency if the resulting network of  $N - 1$  stations is still connected after removal of all transmission links entering or leaving the inoperative station. In each case, the final frequency can be computed from (33) or (34) using the appropriate parameter values. To prove these results, the nonzero averaging weights can be rescaled so that  $A$  has row sums unity; the filter gains  $\lambda_i$  are assumed to be correspondingly rescaled. The characterizing equations for the new system then has the required form and so resynchronization will occur.

The effect of independent jitter noise on the frequency of each oscillator may be considered by including a noise term  $n_i(t)$  in each "input"  $v_i(t)$ . By superposition, the effect of noise can be considered separately. Thus, each phase response  $p_i(t)$  will consist of the response in the absence of noise plus a noise component whose power density spectrum is

$$\sum_{\nu=1}^N |K_{l\nu}(j\omega)|^2 S_{\nu}(\omega),$$

where  $S_{\nu}(\omega)$  is the power density spectrum of  $n_{\nu}(t)$ . Consequently, if the input noise jitter has zero mean and finite variance the output noise components will also have zero mean and finite variance. We conclude, therefore, that in the presence of noise jitter each oscillator will asymptotically have a common frequency with a random perturbation. The perturbations will be correlated but, in general, will not be identical. Furthermore, small jitter noise implies proportionately small perturbations.

#### VII. CONCLUDING REMARKS

We have seen that the transfer function approach has permitted a simple treatment of a rather complicated control system. Further studies regarding transient response or bounds on the size of perturbations due to jitter noise can be made for particular topological configurations by determining more information about the transfer functions  $K_{ij}(s)$  with the help of (25). The linear system approach together with the added generality of having different filters at each station has made it possible to consider the effect of parameter changes or oscillator failure on the behavior of the system.

#### VIII. ACKNOWLEDGMENT

We are indebted to M. B. Brilliant who pointed out an error in an earlier draft of this paper.

#### APPENDIX A

A square matrix  $A$  is said to be *diagonally dominated* when for each row the sum of the magnitudes of the off-diagonal elements is less than the magnitude of the diagonal element, i.e.,

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad \text{each } i.$$

*Theorem:* If  $A$  is diagonally dominated it is nonsingular.

*Proof:* Suppose the contrary. Then there exists a nontrivial solution  $\{x_i\}$  satisfying

$$\sum_j a_{ij}x_j = 0 \quad \text{each } i.$$

Let  $r$  be one of the indices for which  $|x_i|$  is a maximum. Then

$$a_{rr}x_r = -\sum_{k \neq r} a_{rk}x_k$$

so that

$$|a_{rr}| |x_r| \leq \sum_{k \neq r} |a_{rk}| |x_k| \leq \sum_{k \neq r} |a_{rk}| |x_r|$$

which is a contradiction. Hence, the theorem is proved.

## APPENDIX B

### *On the Boundedness of $\beta_i(s)$*

*Theorem: If  $\beta_i(s)$  is bounded by unity on the  $j\omega$  axis then it is bounded by unity in the entire right-half plane.*

*Proof:* Since

$$\beta_i(s) = \frac{H_i(s)/s}{1 + H_i(s)/s},$$

the condition  $|\beta_i(j\omega)| < 1$  is equivalent to

$$|A| < |1 + A \exp(i\varphi)|, \quad (35)$$

where

$$A \exp(i\varphi) = H_i(j\omega)/j\omega.$$

But (35) is equivalent to

$$A \cos \varphi > -\frac{1}{2}$$

so that the locus of  $H(j\omega)/j\omega$ , as  $\omega$  increases from  $-\infty$  to  $\infty$ , cannot encircle the point  $-1$ . Hence, by the Nyquist stability criterion,  $\beta_i(s)$  is analytic in the right-half plane. Furthermore,  $\beta_i(\infty) = 0$ . Thus, it follows that  $|\beta_i(s)| < 1$  in the right-half plane according to the maximum modulus theorem.

## REFERENCES

1. Beneš, V. E., unpublished work, 1959.
2. Runyon, J. P., Reciprocal Timing of Time Division Switching Centers, U.S. Patent No. 3,050,586, August 21, 1962.
3. Goldstein, A. J., unpublished work, 1963.
4. Bonomi, M. J., LaMarche, R. E., and Varaiya, P. P., unpublished work, 1963.
5. Karnaugh, M., A Model for the Organic Synchronization of Communication Systems, B.S.T.J., this issue, pp. 1705-1735.
6. Guillemin, E. A., *Mathematics of Circuit Analysis*, John Wiley & Sons, Inc., New York, 1949, p. 17.

# A Model for the Organic Synchronization of Communications Systems

By M. KARNAUGH

(Manuscript received August 16, 1966)

*Organic synchronization is a method for the mutual synchronization of a set of geographically separated clocks. It is applicable to pulse code modulation (PCM) communications networks and to other systems which have similar requirements for synchronism.*

*After a brief review and history of the problem, a model for organic synchronization is developed. A control-independent study of possible equilibrium solutions is then carried out. A special class of controls is shown to provide asymptotic stability in the limiting case of zero delays. This result leads heuristically to the synthesis of a broad class of nonlinear controls. With these controls, the systems are represented by families of nonlinear differential-functional equations. This model provides a basis for the simulation of organic synchronization. Broad conditions which are mathematically sufficient for the stability of the nonlinear systems are not yet known. The final frequencies of a linear subclass of organic systems, known to be stable, are examined.*

## I. INTRODUCTION

The timing of the switching actions at each switching center of a pulse code modulated (PCM) communications system is governed by a device called the "local clock." It may consist of a cyclic counter driven by an oscillator. Each cycle of the counter is then one clock cycle.

In a geographically widespread PCM system, the local clocks may be either autonomous or synchronized. This choice should be made with the best possible knowledge of the available technology, as well as consideration of its functional and economic consequences. The choice is clearly a rather basic one, and it may have long term effects upon the evolution of the system.

The time-multiplexed PCM signals arriving at any locality may have arbitrary, and usually scattered, points of origin. Some of them require

decoding into a common analog form. In particular, they may be voice signals. A homogeneous, time-multiplexed set of such signals is easily decoded by a common digital-to-analog converter, provided that the transmitted samples have been generated synchronously. A nonsynchronous alternative is to insert extra digits into the signals in order to permit multiplex transmission. Additional equipment is needed to remove these digits and smooth the timing of the demultiplexed samples before or after decoding them.

This paper is only one of a number of studies of system synchronization, and it does not provide a complete solution to the problems touched upon. After a very brief review of some past work in this field, I shall go back to fundamentals to derive a model for organic synchronization. Following this, the sections entitled "Equilibrium Points", "Reduced System Equations", "Controls: Qualitative Discussion", and "System with Zero Delays" provide background for the synthesis of a family of controls which is introduced under the heading, "A Family of Realizable Organic Systems".

The question of the final frequencies of certain linear organic systems is then taken up. Finally, some remarks are made to clarify the stability problem.

## II. HISTORY

The synchronization of PCM networks has long been a subject of interest. The question of synchronizing switching centers, in addition to the transmission links, arose in 1956, when the PCM telephone switching experiment, later named Essex,<sup>1</sup> was planned.

The term "organic synchronization", which seems to have been introduced by V. E. Beneš,<sup>2</sup> will be used herein for systems fitting the model to be derived in later sections. The systems treated by Beneš, excepting a certain minor idealization, form a subclass of these systems. This same subclass of systems is discussed in a patent<sup>3</sup> by J. P. Runyon.

Beneš<sup>2</sup> has demonstrated asymptotic stability for his systems, which are linear, under quite interesting conditions. He has also given formulas for the asymptotic system frequency and for the asymptotic relative phases of the oscillators. A. J. Goldstein<sup>4</sup> has given simplified derivations of these formulas.

An alternative mutual synchronization method, called "frequency averaging", has been treated by Beneš and Goldstein.<sup>5</sup> Frequency averaging systems, while stable, lack a frequency determining element. Each oscillator puts out the average of the frequencies received from its neighbors, and the system frequency will wander in the presence of

noise. Because of this feature, it does not seem to be very practical, unless it is combined with other techniques.

The transmission of a synchronizing signal from a master oscillator to all other oscillators, which are locked to this signal, is perhaps the simplest approach to synchronization. However, such a system is vulnerable to failure of the master oscillator or failure of a transmission link. Means for mitigating this weakness have been proposed by G. P. Darwin and R. C. Prim.<sup>6</sup> They equip the system with automatic means to reorganize itself in the event of a failure. Unfortunately, this adds considerable complexity to the basically simple method.

Further discussion will be limited to organic systems for synchronization.

B. J. Karafin<sup>7</sup> has carried out some digital computer simulations of organic synchronization of small systems. A. Gersho and B. J. Karafin<sup>8</sup> have simplified the proof of asymptotic stability for Beneš' systems. C. J. Candy and M. Karnaugh<sup>9</sup> have studied organic systems of up to four switching centers by means of an analog simulator. M. B. Brilliant has also studied linear organic systems<sup>10</sup> and has computed transient responses of certain large linear systems.<sup>11</sup>

Linear systems with zero delays have also been studied at the University of Tokyo by T. Saito, H. Fujisaki and H. Inose.<sup>12</sup>

### III. THE SYNCHRONIZED NETWORK

Consider a set of  $N \geq 2$  geographically separated pulse code switching centers, interconnected by directed pulse transmission links, as illustrated in Fig. 1 for the case  $N = 4$ .

All possible links need not be physically provided. The cases of greatest interest are those in which there is a directed path from any center

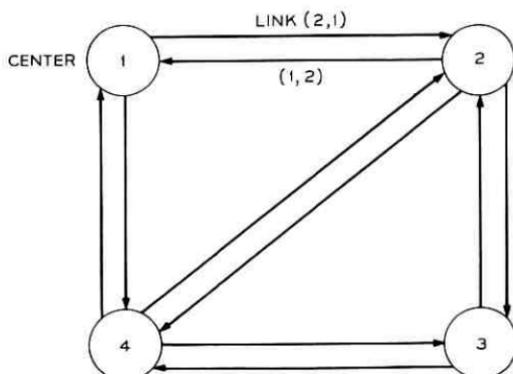


Fig. 1 — A sample network.

to any other center, possibly by way of some intermediate centers. Systems with this property will be called "connected systems".

A connected system of  $N$  centers must have at least  $N$  links, because at least one link must terminate at each center. When the centers are connected in a unidirectional loop, there are exactly  $N$  links. The maximum possible number of links, assuming no duplications, is realized when every ordered pair  $(i, j)$  of distinct centers is connected by a link to  $i$  from  $j$ . This number is  $N(N - 1)$ . The correspondence between the ordered pair  $(i, j)$  and the direction to  $i$  from  $j$  is a convention which will be followed consistently.

An important component at each center is the local clock which determines the timing of all switching actions at that center. The messages from all other centers arrive in the form of framed pulse codes. These are pulse codes divided into sequences containing equal numbers of digits by means of periodically introduced framing digits. In order for the pulse codes to be correctly processed, a correct phase relationship must exist between the arriving framed code and the local clock.

The desired phase relations are realized by providing a certain amount of buffer storage for each incoming link.<sup>13</sup> Such equipment is illustrated in Fig. 2. The arriving digits are stored in a cyclically addressed discrete memory. They are read out of the memory under control of the local clock and of a circuit which monitors the appearances of the framing digits, so as to be correctly phased.

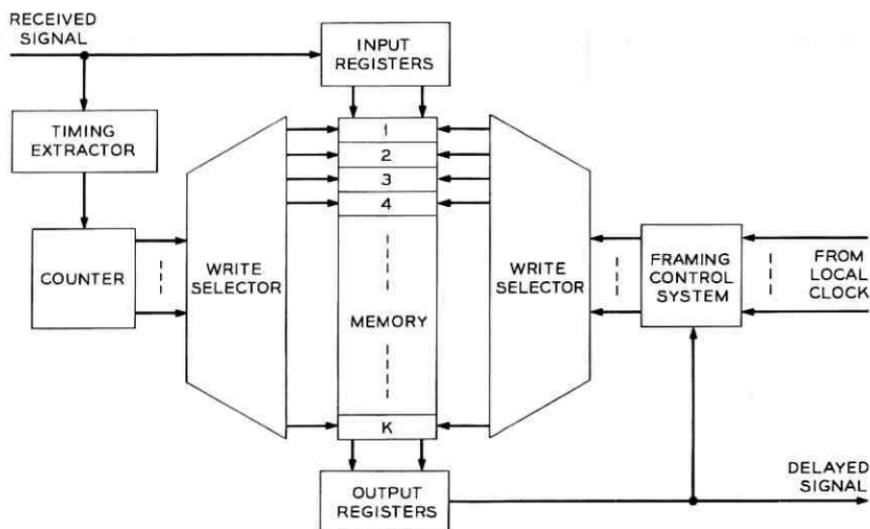


Fig. 2 — Buffer memory.

Other forms of buffer memory which incorporate variable delay lines have also been proposed. These may be acceptable and less costly in some cases.

Under favorable conditions, the arriving signals can be correctly phased by means of the buffers. However, unless the oscillator frequencies are properly controlled, their phase differences will wander beyond any bound. Then, some of the buffer stores will fill up or be emptied, causing erroneous codes to appear at their outputs. It is the primary object of the controls to avoid such malfunctions. The system will be considered to be operating synchronously when no information is being lost in this fashion.

I shall adopt the point of view that it is desired to keep the buffer memories just half full, in which case the system would not be unduly vulnerable to transient disturbances. Controls will be sought which tend to inhibit large deviations from the desired condition. We shall see that these deviations cannot, in general, be reduced to zero. The oscillator control signals will be derived from them.

It should be noted that the transmission delays between centers are variable over some small fractions of their center values. These delays will depend upon the environmental conditions of the propagating media and on message-induced jitter at pulse repeaters.<sup>14</sup> The buffer memories must mop up the delay variations as well as the effects of phase wander of the oscillators.

#### IV. NOTATION

The single subscripts  $i, j, k, \dots$ , refer to the various centers and to the oscillators located at these centers. Their range is the integers, 1, 2,  $\dots, N$ . When one of them appears in a statement or equation with no other qualification, the statement or equation holds over the whole range.

It has already been pointed out that the ordered pair  $(j, i)$  designates the link to center  $j$  from center  $i$ . When a statement or equation contains a pair of subscripts with no other qualification, it holds for all pairs  $(j, i)$  which designate existing links.

The set of all existing links will be called  $R$ . Thus,  $(j, i) \in R$  means there is a link to  $j$  from  $i$  in the system.

Similarly,  $R_i$  is the set of links terminating (i.e., receiving) at center  $i$ , and  $S_i$  is the set of links originating (i.e., sending) at center  $i$ . Thus,

$$R = \bigcup_{i=1}^N R_i = \bigcup_{i=1}^N S_i = \bigcup_R \{(j, i)\}.$$

Let  $M$  be the number of links in the system. We have seen that

$$N \leq M \leq N(N - 1).$$

System controls will be supposed in effect for  $t \geq 0$ . Prior history of the system provides the initial condition. Statements about functions of  $t$  with no other qualification will hold for  $t \geq 0$ .

Occasionally, vectors will be used having  $N$  singly subscripted components or else  $M$  doubly subscripted components. For example, the delay vector  $\tau$  has the  $M$  components  $\tau_{ji}$ ,  $(j, i) \in R$ . It will be clear which vector space is meant in each case.

#### V. PHASE, FREQUENCY, AND DELAY

The local clocks will emit coherent signals. That is, for time intervals which are very long compared to one period of the clock, the output will be approximately periodic. Under these conditions, many formally different definitions of instantaneous frequency will be in good numerical agreement. I shall simply postulate the existence of such continuous functions,  $f_i(t)$ .

Phases of the oscillators are defined to be

$$p_i(t) = p_i(0) + \int_0^t f_i(s) ds \quad (1)$$

in cycles, and

$$f_i = p_i' \quad (2)$$

The principal value of the phase is

$$\varphi_i = p_i \text{ modulo } 1 \quad (3)$$

and has the range  $0 \leq \varphi_i < 1$ .

The initial condition for the phases will be

$$p_i(0) = \varphi_i(0). \quad (4)$$

The values,  $\varphi_i(0)$ , are observables of the physical system. In fact, the switching actions at center  $i$  are timed according to  $\varphi_i(t)$ .

If there is a transmission link to center  $j$  from center  $i$ , the signals transmitted therein will be subject to a time delay  $\tau_{ji}(t)$ . If a pulse is launched from center  $i$  at a time  $t_1$  and received at center  $j$  at time  $t_2$ , then the delay is defined to be

$$\tau_{ji}(t_2) = t_2 - t_1. \quad (5)$$

Analogously, the phase of the signal received at time  $t$  by center  $j$

from center  $i$  is defined to be

$$p_{ji}(t) = p_i[t - \tau_{ji}(t)]. \quad (6)$$

It should be recognized that the principal phase of a timing wave recovered from the received pulse code would, in practice, only approximate  $\varphi_{ji}(t)$ . Errors of a few percent of one pulse period are common in pulse repeaters.<sup>14</sup> This corresponds to a fraction of one percent of a typical frame period.

The frequency of the received signal is obtained by differentiation of (6).

$$f_{ji}(t) = [1 - \tau_{ji}'(t)] \cdot f_i[t - \tau_{ji}(t)]. \quad (7)$$

This equation displays the Doppler shift of frequency caused by variation of  $\tau_{ji}$ .

The clock frequency at the  $i$ th center may be represented in the idealized form

$$f_i(t) = F + E_i + g_i(t) + \eta_i(t). \quad (8)$$

$F$  is the mean center frequency of all clocks in the system, averaged over the time during which the system is observed.  $E_i$  is the incremental center frequency of the  $i$ th clock, also time averaged. By definition,

$$\sum_{i=1}^N E_i = 0. \quad (9)$$

The time function  $g_i(t)$  is the contribution of the system controls, while  $\eta_i(t)$  is a random noise with zero mean. There will be a symmetrical bound on  $g_i(t)$ ,

$$|g_i(t)| \leq G, \quad (10)$$

which is supposed to be larger than the other frequency deviations. This is necessary if the controls are to bring all oscillators to the same average frequency. Under realistic conditions of operation,

$$G > \max_i |E_i| + \max_{t \geq 0} \sigma \left( F \int_t^{t+1/F} \eta_i(s) ds \right) \quad (11)$$

where  $\sigma(\cdot)$  is the standard deviation.

## VI. FUNCTION OF THE BUFFER MEMORY

The principal phase of the signal arriving at center  $j$  from center  $i$  will usually not agree with that of the oscillator at  $j$ . The purpose of

the buffer memory in this link is to delay the signal by an additional time,  $d_{ji}(t)$ , so that

$$\varphi_j(t) = \varphi_{ji}[t - d_{ji}(t)]. \quad (12)$$

In view of (6), this can be written

$$\varphi_j(t) = \varphi_i(t_{ji}), \quad (13)$$

where

$$t_{ji} = t - d_{ji}(t) - \tau_{ji}[t - d_{ji}(t)]. \quad (14)$$

Taking the right-hand derivative of (13), and because this derivative of the principal phase is always equal to the derivative of the phase, we get

$$f_j(t) = [1 - d_{ji}'(t)] \cdot \{1 - \tau_{ji}'[t - d_{ji}(t)]\} \cdot f_i(t_{ji}), \quad (15)$$

where physical considerations make it clear that

$$|d_{ji}'|, \quad |\tau_{ji}'| \ll 1.$$

Equation (15) implies the dependence of  $d_{ji}$  on  $f_i$ ,  $f_j$ , and  $\tau_{ji}$ .

Matters can be simplified by shifting attention from the delays in the buffer memories to the numbers of frames, i.e., clock cycles, they contain. The number of cycles in the  $(j,i)$  buffer at time  $t$  is:

$$y_{ji}(t) = p_{ji}(t) - p_{ji}[t - d_{ji}(t)] \quad (16)$$

and

$$y_{ji}'(t) = f_{ji}(t) - [1 - d_{ji}'(t)] \cdot f_{ji}[t - d_{ji}(t)].$$

Using (7) and (15), this can be put in the form

$$y_{ji}'(t) = f_{ji}(t) - f_j(t), \quad (17)$$

which equates the rate of accumulation of cycles to the rate of arrival minus the rate of removal. In terms of the oscillator frequencies,

$$y_{ji}'(t) = [1 - \tau_{ji}'(t)] \cdot f_i[t - \tau_{ji}(t)] - f_j(t). \quad (18)$$

Suppose the  $(i,i)$  buffer has a capacity of  $2D_{ji}$  cycles. The normalized state of this buffer is defined to be

$$x_{ji}(t) = D_{ji}^{-1} \cdot [y_{ji}(t) - D_{ji}], \quad (19)$$

which is the fractional deviation of its contents from half its capacity. In terms of this variable, (18) becomes

$$x_{ji}'(t) = D_{ji}^{-1} \cdot [1 - \tau_{ji}'(t)] \cdot f_i[t - \tau_{ji}(t)] - D_{ji}^{-1} \cdot f_j(t). \quad (20)$$

This equation explicitly relates the derivative of the buffer memory state variable,  $\mathbf{x}(t)$ , to the delays and clock frequencies in the system.

I have remarked that the frequencies should be controlled so as to prevent any buffer memory from emptying or filling. For example, if buffer  $(j, i)$  is nearly empty, then we desire  $f_i > f_j$  until the situation is sufficiently corrected. On the other hand, if the  $(j, i)$  buffer is nearly full, then the inequality is reversed. However, things are complicated by the fact that all buffers associated with the  $i$ th center, that is, those in links of the set  $R_i \cup S_i$ , are affected by a change in  $f_i$ .

The system is said to malfunction whenever

$$\max_{(j,i) \in R} |x_{ji}(t)| > 1,$$

that is, whenever the buffer memory state vector leaves the "unit cube".

Defining the nonnegative scalar,

$$r_\infty(\mathbf{x}) = \max_{(j,i) \in R} |x_{ji}|, \quad (21)$$

we see that the system is in a permitted buffer state when

$$r_\infty(\mathbf{x}) \leq 1. \quad (22)$$

## VII. EQUILIBRIUM POINTS

Suppose the system is so controlled that an equilibrium solution to the system equations is possible. That is, in the absence of disturbances,

$$\mathbf{x}'(t) = 0 \quad \text{for } t \geq 0$$

and there exists a constant  $f$  such that

$$f_i(t) = f \quad \text{for } t \geq 0, \quad i = 1, 2, \dots, N.$$

If the system, in or near this state, is disturbed by a change in the network configuration, noise in the oscillators, or changes in some of the delays, then variations in the state of the buffer memories will result. To minimize the chance of malfunction under such disturbances, it seems reasonable to seek an equilibrium in which the buffer state is, in some sense, near  $\mathbf{x} = 0$ . That is, the buffers are nearly half full.

I shall begin along these lines by seeking the set of equilibrium points which can be reached from arbitrary initial conditions and under any controls whatever. The situation of asymptotic equilibrium to be considered is as follows.

$$(i) \quad \boldsymbol{\eta}(t) = 0 \quad \text{for } t \geq 0.$$

$$(ii) \quad \tau_{ji}(t) = \tau_{ji}(0) > 0 \quad \text{for } t \geq 0, \quad (j, i) \in R.$$

$$(iii) \lim_{t \rightarrow \infty} f_i(t) = f \quad \text{for } i = 1, 2, \dots, N.$$

$$(iv) \lim_{t \rightarrow \infty} [p_i(t) - p(t)] = \bar{q}_i \quad \text{for } i = 1, 2, \dots, N,$$

where

$$p(t) = \frac{1}{N} \sum_{i=1}^N p_i(t), \quad |\bar{q}_i| < \infty.$$

Let  $\mathbf{x} = \lim_{t \rightarrow \infty} \mathbf{x}(t)$ . The locus of attainable points  $\mathbf{x}$  will be examined under the above conditions.

With the delay vector  $\boldsymbol{\tau}$  assumed to be constant, (20) has the simple form,

$$x_{ji}' = D_{ji}^{-1} \cdot f_i(t - \tau_{ji}) - D_{ji}^{-1} \cdot f_j(t).$$

Therefore,

$$\begin{aligned} x_{ji} &= x_{ji}(0) + \int_0^{\infty} x_{ji}'(t) dt \\ &= x_{ji}(0) + D_{ji}^{-1} \int_0^{\infty} [f_i(t - \tau_{ji}) - f_j(t)] dt. \end{aligned}$$

The integral may be evaluated, using (1) and the conditions (iii) and (iv) of asymptotic stability. The result is,

$$x_{ji} = D_{ji}^{-1}(\bar{q}_i - \bar{q}_j - \tau_{ji}f) + B_{ji}, \quad (23)$$

where

$$B_{ji} = x_{ji}(0) + D_{ji}^{-1}[p_j(0) - p_i(-\tau_{ji})] \quad (24)$$

is a constant which depends upon the initial condition. Equations (23) express the set of buffer memory equilibrium states attainable from a given initial condition in terms of the asymptotic phase differences and the asymptotic system frequency. This set depends upon the initial condition through the parameters,  $B_{ji}$

It is shown in Appendix A that the set of phase differences

$$\{(\bar{q}_i - \bar{q}_j) \mid (j, i) \in R\}$$

contains exactly  $(N - 1)$  linearly independent elements. There are  $M$  components of the asymptotic buffer state vector  $\mathbf{x}$ , with  $M \geq N$ . Therefore, if  $f$  were an unconstrained real variable, we see that, as a function of the phase differences and  $f$ ,  $\mathbf{x}$  would range over an  $N$ -dimensional linear manifold of its  $M$ -dimensional space.

However, the system frequency  $f$  must be near  $F$ , the average center frequency of the clocks. More specifically,

$$F + \max_i E_i - G < f < F + \min_i E_i + G.$$

This inequality requires that the bound on the magnitude of the frequency control be large enough to reduce the highest frequency below the system frequency and to raise the lowest frequency above the system frequency. Thus, all clocks can be brought to a common frequency, even in the presence of noise. Nevertheless,  $G \ll F$  in cases of practical interest.

In a typical application,  $G = 10^{-6} F$ , so the domain of allowed values of  $f$  is a very narrow interval. It may be said that the range of  $\mathbf{x}$  is a neighborhood of an  $(N - 1)$ -dimensional linear manifold. The distance of this manifold from the origin is determined by the initial condition.

In cases of practical interest  $\tau_{ji}G \ll 1$ , so that

$$p_j(0) - p_i(-\tau_{ji}) \cong \tau_{ji}f + p_j(0) - p_i(0).$$

In such cases, (23) has the approximation

$$x_{ji} \cong x_{ji}(0) + D_{ji}^{-1}[\tilde{q}_i - \tilde{q}_j + p_j(0) - p_i(0)]. \quad (25)$$

From this it is clear that  $x_{ji} \cong x_{ji}(0)$  if  $\tilde{q}_i - \tilde{q}_j$  equals  $p_i(0) - p_j(0)$ . Therefore, an asymptotic state vector  $\mathbf{x}$  which is, in some sense, small is attainable when the initial state vector  $\mathbf{x}(0)$  is small in the same sense.

#### VIII. REDUCED SYSTEM EQUATIONS

The trajectory of the buffer memory state vector is of central importance to this work. However, the system controls operate directly upon the clock frequencies. For this reason, it will be convenient to shift attention from the  $M$  equations (20) to an  $N$ -dimensional vector equation for the frequencies. This equation, and its component equations, will be called the "reduced" system equations because  $N \leq M$ .

I shall proceed under the assumptions of no frequency noise and constant delays;

$$\boldsymbol{\eta}(t) = 0 \quad (26)$$

$$\boldsymbol{\tau}(t) = \boldsymbol{\tau}(0). \quad (27)$$

Then, (20) can be integrated to the form

$$x_{ji}(t) = D_{ji}^{-1}[p_i(t - \tau_{ji}) - p_j(t)] + B_{ji}, \quad (28)$$

where, as before,  $B_{ji}$  depends upon the initial condition and is defined by (24).

Now a simple change of variables is made.

$$q_i(t) = p_i(t) - ft. \quad (29)$$

In vector notation, this is,

$$\mathbf{q}(t) = \mathbf{p}(t) - ft \mathbf{1}_N, \quad (30)$$

where  $\mathbf{1}_N$  is an  $N$ -dimensional column vector with unit coordinates.

If the system is asymptotically stable at the frequency  $f$ , then  $\mathbf{q}'(t)$  will go asymptotically to zero. Therefore, given any controls, it will be of interest to see whether this condition is realized for any value of  $f$  in the allowed domain.

Substitution of (29) in (28) gives

$$x_{ji}(t) = D_{ji}^{-1}[q_i(t - \tau_{ji}) - q_j(t) - \tau_{ji}f] + B_{ji}. \quad (31)$$

In view of (2), (8), (26), and (30),

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \mathbf{g}(t), \quad (32)$$

where  $\mathbf{g}(t)$  is the increment to the clock frequencies under system control. In particular, let

$$\mathbf{g}(t) = \mathbf{\Gamma}\{t, \mathbf{x}(\cdot)\}. \quad (33)$$

Each component of  $\mathbf{\Gamma}$  must have a realizable dependence upon the buffer memory state vector trajectory. The problem of control synthesis is precisely the problem of finding a suitable form for  $\mathbf{\Gamma}$ .

The reduced system equations are the differential-functional equations,

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \mathbf{\Gamma}\{t, \mathbf{x}[\mathbf{q}(\cdot), f, \mathbf{B}]\}, \quad (34)$$

with  $\mathbf{x}$  defined by (31). The parameter,  $\mathbf{B}$ , which depends upon the initialization of the system, enters (34) as a parameter in the controls. It is, therefore, not surprising that the system's trajectory and its final state, if any, depend upon its initialization.  $\mathbf{B}$  is not entirely arbitrary. This can be seen by applying the condition for phase agreement, (12) at  $t = 0$ . Some manipulation of (16), (6), (12), (19), and (24) shows that

$$B_{ji} = D_{ji}^{-1}K_{ji} - 1, \quad (35)$$

where  $K_{ji}$  is an integer, and

$$K_{ji} = y_{ji}(0) + p_j(0) - p_i(-\tau_{ji}).$$

Given the initial phases, the initial condition can be changed only by integral changes in the numbers of cycles  $\mathbf{y}(0)$  stored in the buffer memories.

#### IX. CONTROLS: QUALITATIVE DISCUSSION

Loosely speaking, it is desired to control the system so as to keep the buffer memory state vector small, in some appropriate sense. More specifically, the vector should be kept away from the faces of the unit cube.

These qualitative considerations will be made more concrete by defining a class of real valued functions  $r(\mathbf{x})$  of the buffer memory state vector, which will be called "penalty functions". Each such function will have the following properties:

- (i)  $r(\mathbf{0}) = 0$ .
- (ii)  $r(-\mathbf{x}) = r(\mathbf{x})$ .
- (iii)  $r(\mathbf{x})$  has a continuous gradient  $\nabla r(\mathbf{x})$ .
- (iv)  $r(\mathbf{x})$  is strictly convex; that is, for any two distinct vectors  $\mathbf{x}_1, \mathbf{x}_2$  and any real number  $\lambda$  in the open interval  $(0, 1)$ ,

$$r[\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2] < \lambda r(\mathbf{x}_1) + (1 - \lambda)r(\mathbf{x}_2).$$

- (v) For any  $\mathbf{x}$  such that

$$|\mathbf{x}| = 1, \quad \lim_{s \rightarrow \infty} \frac{s}{r(s\mathbf{x})} = 0.$$

These functions will have a unique minimum at the origin and will go to infinity uniformly on all rays from the origin. When properly chosen, their convex surfaces of constant value may closely approximate the cube surfaces having equal values of  $\max_R |x_{ij}|$ . This latter quantity, however, does not have a continuous gradient.

The attainable equilibrium points have been shown to lie in a neighborhood of an  $(N - 1)$ -dimensional linear manifold. The infimum of the values of  $r(\mathbf{x})$  over this set is realized at a unique point in its closure. This point is either the origin or else the point of tangency with a surface of constant  $r$ . After selecting a suitable penalty function, controls will be sought which bring the buffer state vector near this point.

In attempting to reach this objective, a subclass of penalty functions having the simple form

$$r(\mathbf{x}) = \sum_R u(x_{ij}) \quad (36)$$

will be employed. The function  $u(\cdot)$  must have the properties, (i) through (v), of a penalty function on a one-dimensional real space. A simple example of such a function is

$$u_n(x) = x^{2n}, \quad n = 1, 2, \dots$$

Fig. 3 illustrates the surfaces,

$$\begin{aligned} r_\infty &\equiv \max(x_{12}, x_{21}) = 1 \\ r_1 &\equiv x_{12}^2 + x_{21}^2 = 1 \\ r_2 &\equiv x_{12}^4 + x_{21}^4 = 1 \\ r_3 &\equiv x_{12}^6 + x_{21}^6 = 1. \end{aligned}$$

The last three are penalty functions of the type defined in (36), for a system having just two links.

#### X. SYSTEM WITH ZERO DELAYS

The family of systems under consideration will have widely varying nonnegative delays for the transmission links. In many cases of interest, the product of maximum loop delay and control bandwidth may be very small compared to unity. In such cases, the extrapolation to zero delays

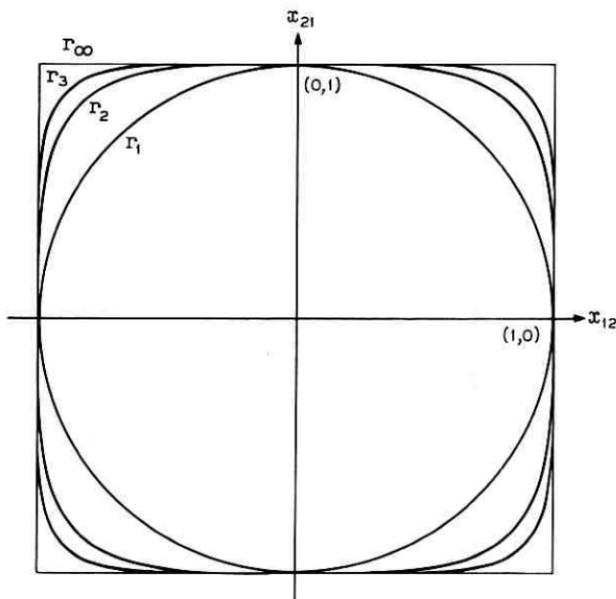


Fig. 3 — Curves of unit value.

may be a useful and illuminating exercise. One such system will be treated here as a step toward the synthesis of a family of controls.

In the zero delay case, let the controls  $\Gamma$  be simply a function of  $\mathbf{x}$ . This results in a system of ordinary, autonomous differential equations.

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \Gamma(\mathbf{x}) \quad (37)$$

in which  $\mathbf{x}$  is now defined by

$$x_{ji}(t) = D_{ji}^{-1}[q_i(t) - q_j(t)] + B_{ji}. \quad (38)$$

The synthesis of  $\Gamma(\mathbf{x})$  will be based upon a penalty function which takes the form given in (36). One of the desired properties of the system (37) is that it come to rest near the attainable minimum of  $r(\mathbf{x})$ . It is therefore reasonable to try controls which have a component along the negative gradient of  $r(\mathbf{x})$ .

Let  $\nabla_q r(\mathbf{x})$  be the column vector whose  $i$ th component is

$$[\nabla_q r(\mathbf{x})]_i = \frac{\partial}{\partial q_i} r(\mathbf{x}) \quad (39)$$

and let  $A$  be any  $N \times N$  positive definite matrix. The controls to be considered here are of the form

$$\Gamma(\mathbf{x}) = -A \nabla_q r(\mathbf{x}). \quad (40)$$

Thus, we are assured that\*

$$[-\nabla_q r(\mathbf{x})]^T \cdot \Gamma(\mathbf{x}) \geq 0 \quad (41)$$

with equality if and only if  $\nabla_q r(\mathbf{x}) = 0$ .

Make the linear change of variables,

$$\mathbf{w}(t) = A^{-1}\mathbf{q}(t). \quad (42)$$

Then, the system equations are

$$\mathbf{w}'(t) = A^{-1}[(F - f)\mathbf{1}_N + \mathbf{E}] - \nabla_q r(\mathbf{x}). \quad (43)$$

The inverse of a positive definite matrix is also positive definite. Therefore,

$$\mathbf{1}_N^T A^{-1} \mathbf{1}_N > 0$$

and it is possible to choose  $f$  such that

$$\mathbf{1}_N^T A^{-1} [(F - f)\mathbf{1}_N + \mathbf{E}] = 0. \quad (44)$$

\* The superscript  $T$  will indicate the transpose of a vector or matrix.

It is easy to verify that

$$\mathbf{1}_N^T \nabla_{\mathbf{q}} r(\mathbf{x}) = 0 \quad (45)$$

in the zero delay system. Having chosen  $f$  to satisfy (44), we find that the trajectory of  $\mathbf{w}$  must lie in a linear manifold which is orthogonal to  $\mathbf{1}_N$ . It will now be shown that the system exhibits global asymptotic stability when  $\mathbf{w}$  is restricted to this linear manifold. That is, from any starting point in the manifold, the system will ultimately come to rest at a unique point in the manifold.

Consider the following function:

$$L(\mathbf{w}) = -\mathbf{w}^T A^T A^{-1} [(F - f)\mathbf{1}_N + \mathbf{E}] + r(\mathbf{x}). \quad (46)$$

Because of some of the properties of  $r(\mathbf{x})$  and the connectedness of the systems under study, it can be shown that  $L(\mathbf{w})$  has a unique minimum in every linear manifold orthogonal to  $\mathbf{1}_N$ . Also,  $\nabla_{\mathbf{w}} L(\mathbf{w})$  is zero only at this minimum. The proofs are given in Appendix B.

The time rate of change of  $L(\mathbf{w})$  is

$$\dot{L}(\mathbf{w}) = [\mathbf{w}'(t)]^T \nabla_{\mathbf{w}} L(\mathbf{w}).$$

But

$$\begin{aligned} \nabla_{\mathbf{w}} L(\mathbf{w}) &= -A^T A^{-1} [(F - f)\mathbf{1}_N + \mathbf{E}] + A^T \nabla_{\mathbf{q}} r(\mathbf{x}) \\ &= -A^T \mathbf{w}'(t). \end{aligned} \quad (47)$$

Therefore,

$$\dot{L}(\mathbf{w}) = -[\mathbf{w}'(t)]^T A^T \mathbf{w}'(t) = -[w'(t)]^T A w'(t). \quad (48)$$

In view of the hypothesis that  $A$  is positive definite,

$$\dot{L}(\mathbf{w}) \leq 0,$$

with equality if and only if  $w'(t) = 0$ . When this occurs,

$$-A^T \mathbf{w}'(t) = \nabla_{\mathbf{w}} L(\mathbf{w}) = 0$$

and the system is at the minimum of  $L(\mathbf{w})$ . Thus,  $L(\mathbf{w})$  is a Liapunov function for the system<sup>15,16</sup> and the system is globally asymptotically stable in the restricted sense mentioned above.

Inasmuch as

$$\mathbf{q}'(t) = A \mathbf{w}'(t),$$

$\mathbf{w}'(t) = 0$  implies  $\mathbf{q}'(t) = 0$  and the system of (37) and (40) is also globally asymptotically stable in the linear manifold of its motion.

It is apparent from the definition (30) of  $\mathbf{q}(t)$  that the system has

the common frequency  $f$  when  $\mathbf{q}'(t) = 0$ . The Liapunov function

$$L(\mathbf{w}) = \hat{L}(\mathbf{q}) = -\mathbf{q}^T A^{-1}[(f - f)\mathbf{1}_N + \mathbf{E}] + r[\mathbf{x}(\mathbf{q})]$$

has its minimum at the equilibrium point,  $\mathbf{q}_{\min}$ , where

$$\nabla_{\mathbf{q}} \hat{L}(\mathbf{q}_{\min}) = -A^{-1}[(F - f)\mathbf{1} + \mathbf{E}] + \nabla_{\mathbf{q}^r}[\mathbf{x}(\mathbf{q}_{\min})] = 0.$$

The equilibrium point is seen to be offset from that point at which  $\nabla_{\mathbf{q}^r} = 0$  when the mistuning  $\mathbf{E}$  of the clocks does not vanish. This should not be surprising, because the controls must compensate for the frequency differences.

Suppose the  $(N - 1)$ -dimensional subspace orthogonal to  $\mathbf{1}_N$  is invariant under  $A^{-1}$ . Then the requirement (44) reduces to  $f = F$ , and the system frequency is the average of the clock center frequencies. It can be shown that when the matrix  $A$  has the above property, its column sums are all equal. In particular, if a diagonal matrix has this property, it is the identity matrix, multiplied by a positive scalar.

In this section, I have considered global asymptotic stability, rather than trajectories within the unit cube. Attainment of a suitably bounded trajectory will depend upon the  $A$  matrix and the initialization of the system.

#### XI. A FAMILY OF REALIZABLE ORGANIC SYSTEMS

In the last section, the controls

$$\Gamma(\mathbf{x}) = -A \nabla_{\mathbf{q}^r}(\mathbf{x}) = -A \nabla_{\mathbf{q}} \sum_R u(x_{ji})$$

were shown to stabilize the system of (37) with zero delays. A family of controls will now be synthesized so as to be realizable and practical for systems having positive delay.

Equations (38) show that

$$\frac{\partial}{\partial q_k} u(x_{ji}) = D_{ji}^{-1}(\delta_{ik} - \delta_{jk})u'(x_{ji}) \quad (49)$$

using the Kronecker  $\delta$  notation. Thus, when the matrix  $A$  is diagonal, the controls for the clock at center  $i$  depend only upon the buffer memory states in links terminating at center  $i$  or originating at center  $i$ . This is a very desirable simplification, and there seems to be no merit in employing more complicated forms. A more general type of control having this property is

$$\Gamma_k(\mathbf{x}) = \sum_{R_k} a_{kj} u'(x_{kj}) - \sum_{S_k} b_{jk} u'(x_{jk}), \quad (50)$$

where the signs have been chosen to agree with the earlier model when  $a_{kj}$ ,  $b_{jk}$  are positive. The development is quite heuristic at this point, because a proof of stability for the system having the controls (50) with arbitrary positive coefficients is lacking.

The final model is based upon a modified form of the controls (50).

- (i) To achieve realizability, a delay  $\Delta_{kj}$  must be imposed on the argument of a control signal from the  $(j,k)$ th buffer memory to center  $k$ .
- (ii) The controlled frequency deviations must be limited. For this purpose, I introduce the limiter function  $\rho(\cdot)$  such that

$$\begin{aligned} \rho(x) &= x, & |x| &\leq 1 \\ &= 1, & x &> 1 \\ &= -1, & x &< -1. \end{aligned}$$

- (iii) For the purpose of reducing system bandwidth, a filter with impulse  $h(t)$  and unit dc response may be employed. Let  $*$  indicate convolution.

$$\Gamma_i\{t, \mathbf{x}(\cdot)\} = G\rho\left(h(t) * \left\{ \sum_{R_i} a_{ij} u'[x_{ij}(t)] - \sum_{S_i} b_{ji} u'[x_{ji}(t - \Delta_{ij})] \right\}\right). \quad (51)$$

The complete system equations are

$$\mathbf{q}'(t) = (F - f)\mathbf{1}_N + \mathbf{E} + \Gamma\{t, \mathbf{x}(\cdot)\} + \mathbf{n}(t) \quad (52)$$

$$x_{ji}(t) = D_{ji}^{-1}\{q_i[t - \tau_{ji}(t)] - q_j(t) - \tau_{ji}(t)f\} + B_{ji} \quad (53)$$

$$B_{ji} = x_{ji}(0) + D_{ji}^{-1}\{p_j(0) - p_i[-\tau_{ji}(0)]\}. \quad (54)$$

Equations (53) and (54) have been obtained by integration of (20).

The definition of the delay  $\Delta_{ij}$  in (51) will depend upon the manner of transmission of the control signal. When the state  $x_{ji}$  is transmitted to center  $i$  via link  $(i,j)$ ,

$$\Delta_{ij} = \Delta_{ij}(t) = d_{ij}(t) + \tau_{ij}[t - d_{ij}(t)]. \quad (55)$$

This form is particularly awkward because the buffer memory delay,  $d_{ij}(t)$ , is not one of the canonical variables. However, it can be very closely approximated as follows:

$$d_{ij}(t) \cong D_{ij}[x_{ij}(t) + 1]F. \quad (56)$$

A simpler but cruder approximation is

$$d_{ij}(t) \cong D_{ij}F. \quad (57)$$

In the special case of (51), for which  $b_{ji} = 0$ , this complication does not arise.

## XII. COMPARISON WITH THE SYSTEMS OF BENEŠ

V. E. Beneš<sup>2</sup> has analyzed a class of linear systems having delays and filters. In my notation, these systems obey the equations

$$p_i'(t) = F + E_i + Gh(t) * \sum_{R_i} \tilde{a}_{ij} [p_j(t - \tau_{ij}) - p_i(t)] + \eta_i(t). \quad (58)$$

He considers the delays to be fixed, the systems to be connected systems, and adds the constraints

$$\tilde{a}_{ij} > 0 \quad \text{for } (i, j) \in R_i$$

$$\sum_{R_i} \tilde{a}_{ij} = 1.$$

Then, assuming the noise  $\eta(t)$  to be bounded and to go asymptotically to zero, he finds a sufficient condition for global asymptotic stability. This asymptotic stability is defined by

$$\lim_{t \rightarrow \infty} p_i'(t) = f, \quad i = 1, 2, \dots, N$$

$$\left| \lim_{t \rightarrow \infty} [p_i(t) - p_N(t)] \right| < \infty, \quad i = 1, 2, \dots, (N - 1).$$

Beneš sufficient condition is that

$$G > 0$$

$$\left| \frac{GH(i\omega)}{i\omega + GH(i\omega)} \right| < 1 \quad \text{for all } \omega \neq 0.$$

$H(s)$  is the Laplace transform of the filters' impulse response,  $h(t)$ ;  $\omega$  is real radian frequency,  $i$  is the imaginary unit, and it is assumed that  $H(0) = 1$ , as before.

This condition is stricter than that needed to stabilize an ordinary phase controlled oscillator, but it is not too difficult to satisfy. It is also quite remarkable in its independence of the system graph and its delays.

Beneš also gives formulas for the final system frequency and phase differences. These have been rederived more simply by Goldstein,<sup>4</sup> using the final value theorem.

A very direct approach to the final values is to insert them in (58), replacing the convolution with  $Gh(t)$  by multiplication with  $G$ .

Let

$$\lim_{t \rightarrow \infty} p_i'(t) = f$$

$$\lim_{t \rightarrow \infty} [p_i(t) - p_N(t)] = \bar{p}_i$$

and note that

$$\lim_{t \rightarrow \infty} [p_i(t - \tau_{ij}) - p_i(t)] = -\tau_{ij}f.$$

Making the appropriate changes in (58),

$$f = F + E_i + G \sum_{R_i} \bar{a}_{ij}[\bar{p}_j - \bar{p}_i - \tau_{ij}f].$$

Because  $\bar{p}_N = 0$  by definition, we now have a set of  $N$  linear equations in the  $N$  variables  $f, \bar{p}_1, \dots, \bar{p}_{N-1}$ . Using the notation,

$$\tau_i = \sum_{R_i} \bar{a}_{ij}\tau_{ij},$$

for an average of the delays in the links to the  $i$ th center, the equations assume the simpler form,

$$(1 + G\tau_i)f + G\bar{p}_i - G \sum_{R_i} \bar{a}_{ij}\bar{p}_j = F + E_i. \quad (59)$$

They have been shown by Goldstein to give the following solution for  $f$

$$f = \frac{F + \sum b_i E_i}{1 + G \sum b_i \tau_i}. \quad (60)$$

Here  $b_i, i = 1, 2, \dots, N$ , depends only upon the averaging coefficients,  $\bar{a}_{ij}, (i, j) \in R$ , and

$$b_i \geq 0$$

$$\sum b_i = 1.$$

A glance at (60) shows that the final system frequency is monotone decreasing with the product of the dc gain,  $G$ , and an average of all delays in the system. This effect has caused some dismay, but it results from an unrealistic model.

Let us go back to the family of organic systems defined by (51) through (54) and (29). These will be specialized in such a way as to obtain a class of linear systems analogous to that of Beneš. The following steps must be taken.

(i) Eliminate the limiter,  $\rho(\cdot)$ , from the controls, (51).

(ii) Let  $u(x_{ij}) = \frac{1}{2}x_{ij}^2$ , so  $u'(x_{ij}) = x_{ij}$ .

(iii) Use the phase variables,

$$p_i(t) = q_i(t) + ft.$$

(iv) Set  $b_{ji} = 0$ , for all  $(j, i)$ , in (51).

(v) Make the identification

$$\tilde{a}_{ij} = a_{ij}D_{ij}^{-1}$$

and impose the constraint,

$$\sum_{R_i} \tilde{a}_{ij} = 1.$$

(vi) Assume the delays to be constant.

With these changes,

$$\begin{aligned} p_i'(t) = F + E_i + Gh(t) * \sum_{R_i} \tilde{a}_{ij} [p_j(t - \tau_{ij}) - p_i(t)] \\ + G \sum_{R_i} \tilde{a}_{ij} [D_{ij}x_{ij}(0) + p_i(0) - p_j(-\tau_{ij})] + \eta_i(t). \end{aligned} \quad (61)$$

This set of system equations differs from those of Beneš, (58), only in the addition of a term which is constant in time, but which depends upon the initial condition. It may be considered to be a modification of the mistuning,  $E_i$ , in the treatment of the stability problem. Therefore, the proof given by Beneš of global asymptotic stability under his sufficient condition also applies to (61).

Now let us derive the equations for final values. In doing this, note that the  $j$ th oscillator has the natural frequency,  $F + E_j$ , for  $t < 0$ , while its frequency has the final value  $f$ . Therefore,

$$\begin{aligned} p_j(-\tau_{ij}) &= p_j(0) - \tau_{ij}(F + E_j) \\ \lim_{t \rightarrow \infty} [p_j(t - \tau_{ij}) - p_j(t)] &= -\tau_{ij}f. \end{aligned}$$

Now, proceeding as before, (61) leads to

$$\begin{aligned} f = F + E_i + G \sum_{R_i} \tilde{a}_{ij} \cdot (\tilde{p}_j - \tilde{p}_i - \tau_{ij}f) \\ + G \sum_{R_i} \tilde{a}_{ij} [D_{ij}x_{ij}(0) + p_i(0) - p_j(0) + \tau_{ij} \cdot (F + E_j)] \end{aligned}$$

Putting this in a form analogous to (59),

$$\begin{aligned} (1 + G\tau_i)f + G\tilde{p}_i - G \sum_{R_i} \tilde{a}_{ij}\tilde{p}_j = (1 + G\tau_i)F + E_i \\ + G \sum_{R_i} \tilde{a}_{ij} [\tau_{ij}E_j + D_{ij}x_{ij}(0) + p_i(0) - p_j(0)]. \end{aligned} \quad (62)$$

The solution for  $f$ , analogous to (60), is

$$f = F + \sum_R c_{ij} E_j + \frac{G \sum_R b_i \tilde{a}_{ij} D_{ij} x_{ij}(0)}{1 + G \sum b_i \tau_i} \quad (63)$$

where the coefficients  $c$  are averaging coefficients defined by

$$c_{ij} = \frac{b_j \tilde{a}_{ji} + G b_i \tilde{a}_{ij} \tau_{ij}}{\sum_R (b_j \tilde{a}_{ji} + G b_i \tilde{a}_{ij} \tau_{ij})} \quad (64)$$

and the coefficients  $b$  are the same as before.

The first two terms on the right of (63) give an average of the individual clock frequencies. The last term depends on the initial condition, but it goes to zero as the system delays become large. Thus, the behavior indicated by (60) does not really occur in our model for organic systems.

### XIII. SOME REMARKS ON THE STABILITY PROBLEM

The mathematical problem of stability is not yet satisfactorily solved for general organic systems. Two special families of organic systems are now known to be globally asymptotically stable. These are certain nonlinear systems with zero delays and certain linear systems with delays and filters. These very special cases nourish the hope that broader sufficient conditions for stability can be found.

The present section will be devoted to redefinition of the stability problem and a discussion of some necessary conditions for stability.

We have seen that the system will malfunction whenever the  $M$ -dimensional buffer state vector  $\mathbf{x}(t)$  leaves the unit cube. This leads to the following practical definition of stability.

*Definition:* For any positive  $\varepsilon$ , trajectory  $\mathbf{x}(t)$  is  $\varepsilon$ -stable if

$$\max_i |x_i(t)| \leq \varepsilon \quad \text{for } 0 \leq t < \infty.$$

*Definition:* A trajectory is stable if there is an  $\varepsilon < 1$  for which it is  $\varepsilon$ -stable.

The trajectory of an undisturbed organic system will depend upon the system parameters and the initial condition. Therefore, the domain of system stability must be defined in a space having the following coordinates, which appear in (51) through (54).  $G$ ,  $h(\cdot)$ ,  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $u(\cdot)$ ,  $\mathbf{\Delta}$ ,  $\mathbf{D}$ ,  $\mathbf{x}(0)$ ,  $\mathbf{E}$ ,  $\boldsymbol{\tau}$  and  $\mathbf{p}(t)$  for  $t \leq 0$ . The initialization of the filter states must also be defined.

Several necessary conditions should be kept in mind when testing for stability. First, the asymptotic frequency of the system must lie within the controllable range,

$$F + E_i - G \leq f \leq F + E_i + G, \quad i = 1, 2, \dots, N.$$

Second, the limit points of the buffer state trajectory must lie within the unit cube. Third, the system must be "connected," in some sense.

The connectedness of the systems deserves further discussion. I have required that the organic systems must be connected in the following sense: that there must exist a directed transmission path from each center to every other center. Beneš has used the same condition. On the other hand, the stability of the special nonlinear systems which I have treated depends upon a weaker condition. Namely, that the nonoriented graph having a branch corresponding to each link must be connected. This condition is used in Appendix B, which is essential to the proof.

An important difference between the systems treated here and those of Beneš, is that the former have the additional control coefficients,  $b_{ij} \geq 0$  for  $(i,j) \in R$ . Thus, the state  $x_{ij}(t)$  of a particular buffer may exert a control over the frequency of the sending center,  $j$ , as well as over the receiving center,  $i$ .

Intuition suggests a necessary condition for stability based upon the control coefficients,  $a_{ij}$ ,  $b_{ij}$ , which appear in the family of equations, (51). Inasmuch as negative coefficients tend to make the systems unstable, these are assumed to be either positive or zero.

Consider a "control graph" with nodes numbered 1, 2,  $\dots$ ,  $N$ . Let a directed branch exist from node  $j$  to node  $i$  if and only if  $a_{ij} + b_{ji} > 0$ . This condition permits the frequency at center  $i$  to be influenced by its phase relative to that at center  $j$ . Then a necessary condition for system stability is that the control graph shall have a node from which directed paths exist to all other nodes.

Under this "weak" condition, some parts of the system may simply be "slaves" of another part of the system. The "strong" condition that there exist a directed path from each node to each other node precludes this possibility. However, it should be understood that the condition satisfied by the control graph need not be satisfied by the graph whose branches correspond to transmission links. This is the case because each link may give rise to two oppositely directed branches of the control graph. On the other hand, a connected system may lack stability when too few of the control coefficients are positive.

A simple example is provided by the system shown in Fig. 4. The digital transmission links appear in the "system graph" Fig. 4(a). When

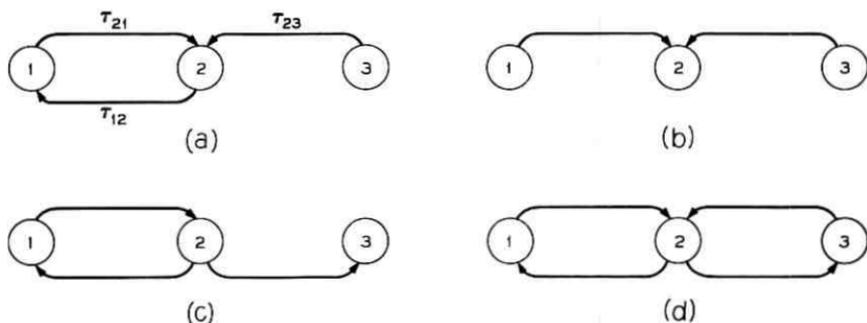


Fig. 4—A system graph and three control graphs. (a) system graph; (b) inadequately connected control graph; (c) weakly connected control graph; (d) strongly connected control graph.

the only positive coefficients are  $a_{21}$ ,  $a_{23}$ , the control graph of Fig. 4(b) results. This fails to satisfy even the weak condition. It is clear that no means exists for bringing centers 1 and 3 to a common frequency. The control graph in Fig. 4(c) results when  $a_{12}$ ,  $a_{21}$ ,  $b_{23}$  are the positive coefficients. It must be assumed that a *separate* means for transmitting the narrow band signal,  $b_{23}$ , to center 3 exists. This control graph satisfies the weak condition. If the system is stable, center 3 will be a slave to the rest of the system. It will have no influence upon the frequency trajectories of centers 1 and 2. If  $a_{12}$ ,  $a_{21}$ ,  $a_{23}$ ,  $b_{23}$  are the positive coefficients, then the control graph of Fig. 4(d) results. This one obeys the strong condition. When the control graph is the same as the system graph, Fig. 4(a), it is weakly connected. In this case, if the system is stable, center 3 determines the common frequency.

#### XIV. SUMMARY

A class of systems for the mutual synchronization of spatially separated oscillators has been synthesized and a mathematical model for these systems has been presented. The model may be said to be physically realizable in that real systems can be built whose function will very closely approximate the behavior of the model. While no such system hardware has been presented here, a simple hardware analog has been built.<sup>9</sup>

These systems, called "organic synchronization systems," have a possible application to continental or worldwide PCM communications.

A re-examination of the systems treated by Beneš in the light of the newly derived organic model indicates that

- (i) his stability proof does apply to a particular class of linearized organic systems, and

- (ii) that his formula for the final system frequency must be modified; the suitably modified formula no longer displays a monotone decreasing final frequency with increasing system delays.

## APPENDIX A

*Asymptotic Phase Differences*

The set of all differences of the form,  $(\bar{q}_i - \bar{q}_j)$ , can have at most  $(N - 1)$  linearly independent elements. To verify this, consider the  $(N - 1)$  elements

$$(\bar{q}_N - \bar{q}_{N-1}), (\bar{q}_{N-1} - \bar{q}_{N-2}), \dots, (\bar{q}_2 - \bar{q}_1).$$

Suppose  $i > j$ . Then

$$(\bar{q}_i - \bar{q}_j) = (\bar{q}_i - \bar{q}_{i-1}) + (\bar{q}_{i-1} - \bar{q}_{i-2}) + \dots + (\bar{q}_{j+1} - \bar{q}_j)$$

while

$$(\bar{q}_j - \bar{q}_i) = -(\bar{q}_i - \bar{q}_j).$$

Hence, any other element of the set of differences can be represented as a linear combination of the selected  $(N - 1)$  elements.

When the directed graph, which corresponds to the synchronizing network, is connected the set of differences

$$\{(\bar{q}_i - \bar{q}_j) \mid (j, i) \in R\}$$

has at least  $(N - 1)$  linearly independent elements. Actually, only the weak, i.e., nonoriented, sense of connectedness of necessary for the proof.

*Theorem:* Let  $G$  be a directed graph with  $N$  vertices such that the corresponding nonoriented graph is connected. Associate the  $N$  independent real variables,  $\bar{q}_1, \bar{q}_2, \dots, \bar{q}_N$ , one to one with the correspondingly indexed vertices. Associate the difference  $(\bar{q}_i - \bar{q}_j)$  with the edge from vertex  $i$  to vertex  $j$ , for each edge in  $G$ . Let  $T$  be any complete tree of  $G$ .

Then, there is a set of  $(N - 1)$  linearly independent differences associated with edges of  $T$ .

*Proof:* We shall proceed by induction. The theorem clearly is true when  $N = 2$ .

Suppose the theorem to be true for  $N = L - 1$ . Now consider  $G$  to have  $L$  vertices. Since it is connected, it contains a complete tree  $T$ , which will have  $L - 1$  edges, but all  $L$  vertices. It follows that not all vertices can have more than one edge of  $T$  incident on them. Let vertex  $j$  be an end vertex. Then, only one of the differences associated with the

edges of  $\tau$  contains  $\bar{q}_j$ . It follows that this difference, say  $\pm(\bar{q}_i - \bar{q}_j)$ , is linearly independent of the remaining  $(L - 2)$  differences. But if vertex  $j$  and the edge incident on it are removed from  $T$ , a tree having  $(L - 1)$  vertices remains. By hypothesis, this contains  $(L - 2)$  edges whose associated differences are linearly independent. Hence, there are  $(L - 1)$  linearly independent differences associated with the edges of  $T$ , and the proof is complete.

*Corollary:* The differences are unchanged when the average of the  $\bar{q}$ 's is subtracted from each of them. Therefore, the theorem applies even when the  $N$  real variables are constrained to have zero sum.

## APPENDIX B

### Properties of $L(\mathbf{w})$

Certain preliminaries concerning convex functions will be necessary here. While they are familiar to mathematicians, other readers may find the following review helpful.

All sets to be considered here are subsets of finite dimensional real linear spaces. All functions will be defined on such sets.

*Definition 1:* A set of points  $X$  is convex if every point of the set

$$\{\lambda \mathbf{x}_2 + (1 - \lambda) \mathbf{x}_1 : \mathbf{x}_1, \mathbf{x}_2 \in X \text{ and } 0 \leq \lambda \leq 1\}$$

is also a point of  $X$ .

*Definition 2:* A real valued function  $f(\cdot)$  defined on a convex set  $X$  is a convex function on  $X$  if

$$f[\lambda \mathbf{x}_2 + (1 - \lambda) \mathbf{x}_1] \leq \lambda f(\mathbf{x}_2) + (1 - \lambda) f(\mathbf{x}_1) \quad (65)$$

whenever  $\mathbf{x}_1, \mathbf{x}_2 \in X$  and  $0 \leq \lambda \leq 1$ .

A convex function  $f(\cdot)$  is *strictly convex* if the equality in (65) implies that  $\lambda = 0$  or  $\lambda = 1$  or  $\mathbf{x}_1 = \mathbf{x}_2$ .

*Theorem 1:* If two convex functions are defined on the same convex set, their sum is a convex function on that set. If one of the functions is strictly convex, then the sum is strictly convex.

*Theorem 2:* If  $f(\cdot)$  is a convex function on a convex set  $X$ , and if  $\alpha \geq 0$ , then  $\alpha f(\cdot)$  is a convex function on  $X$ . If  $f(\cdot)$  is strictly convex, and if  $\alpha > 0$ , then  $\alpha f(\cdot)$  is strictly convex.

*Theorem 3:* If  $\mathbf{c}$  is a fixed vector in an  $n$ -dimensional space, and  $\mathbf{x}$  is a

variable vector in the same space, and  $d$  is a real number, then  $(\mathbf{c}^T \cdot \mathbf{x} + d)$  is a convex function on the entire space.

*Theorem 4:* Let  $f(\cdot)$  be a convex function on an  $n$ -dimensional space. Then the sets of points

$$\{\mathbf{x}: f(\mathbf{x}) \leq d\}, \quad \{\mathbf{x}: f(\mathbf{x}) < d\}$$

are convex subsets of the space for every real number,  $d$ .

The above theorems are elementary consequences of the definitions, 1 and 2.

*Theorem 5:* Let  $f(\cdot)$  be a differentiable convex function defined on a convex set  $X$ . Let  $\mathbf{x}_1, \mathbf{x}_2$  be distinct points of  $X$ . Then the directional derivative,  $(d/d\lambda) f[\lambda\mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1]$ , is an increasing function of  $\lambda$  in  $(0,1)$ .

*Proof:* Select any  $\lambda_0 > 0, \delta > 0$  such that  $\lambda_0 + \delta < 1$ . Let

$$\mathbf{x}_0 = \lambda_0\mathbf{x}_2 + (1 - \lambda_0)\mathbf{x}_1, \quad \mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1.$$

Then

$$(\lambda_0 + \delta)\mathbf{x}_2 + (1 - \lambda_0 - \delta)\mathbf{x}_1 = \mathbf{x}_0 + \delta\mathbf{y}.$$

Now select an  $\varepsilon$  such that  $\varepsilon > 0$  and  $\lambda_0 + \delta + \varepsilon < 1$  and apply (65) twice as follows:

$$\begin{aligned} f(\mathbf{x}_0 + \delta\mathbf{y}) &= f\left[\frac{\delta}{\delta + \varepsilon}(\mathbf{x}_0 + \delta\mathbf{y} + \varepsilon\mathbf{y}) + \frac{\varepsilon}{\delta + \varepsilon}\mathbf{x}_0\right] \\ &\leq \frac{\delta}{\delta + \varepsilon}f(\mathbf{x}_0 + \delta\mathbf{y} + \varepsilon\mathbf{y}) + \frac{\varepsilon}{\delta + \varepsilon}f(\mathbf{x}_0) \\ f(\mathbf{x}_0 + \varepsilon\mathbf{y}) &\leq \frac{\varepsilon}{\delta + \varepsilon}f(\mathbf{x}_0 + \delta\mathbf{y} + \varepsilon\mathbf{y}) + \frac{\delta}{\delta + \varepsilon}f(\mathbf{x}_0). \end{aligned}$$

Adding these inequalities and rearranging terms,

$$f(\mathbf{x}_0 + \delta\mathbf{y} + \varepsilon\mathbf{y}) - f(\mathbf{x}_0 + \delta\mathbf{y}) \geq f(\mathbf{x}_0 + \varepsilon\mathbf{y}) - f(\mathbf{x}_0).$$

Dividing both members by  $\varepsilon$  and taking the limits as  $\varepsilon \rightarrow 0$  yields the desired result,

$$\left. \frac{d}{d\lambda} f[\lambda\mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1] \right|_{\lambda_0 + \delta} \geq \left. \frac{d}{d\lambda} f[\lambda\mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1] \right|_{\lambda_0}.$$

*Corollary:* If  $f(\cdot)$  is strictly convex on the convex set  $X$ , and if  $\mathbf{x}_1, \mathbf{x}_2$  are distinct points of  $X$ , then  $(d/d\lambda) f[\lambda\mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1]$  is a strictly increasing function of  $\lambda$  in  $(0,1)$ .

*Proof:* Suppose that the equality holds in Theorem 5. If the directional derivative is nondecreasing in  $[\lambda_0, (\lambda_0 + \delta)]$  and takes equal values at the end points, then it must be constant on this interval. It follows that  $f(\cdot)$  varies linearly on the line segment from

$$\mathbf{x} = \lambda_0 \mathbf{x}_2 + (1 - \lambda_0) \mathbf{x}_1 \quad \text{to} \quad \mathbf{x} = (\lambda_0 + \delta) \mathbf{x}_2 + (1 - \lambda_0 - \delta) \mathbf{x}_1.$$

This contradicts the hypothesis that  $f(\cdot)$  is strictly convex on  $X$ , and the proof is complete.

*Theorem 6:* Let  $f(\cdot)$  be a strictly convex differentiable function defined on a convex set  $D$ . Let  $C$  be a closed and bounded convex set in the interior of  $D$ . Then,

- (i)  $f(\cdot)$  assumes its minimum value over  $C$  at a unique point of  $C$ ,
- (ii)  $f(\cdot)$  has a vanishing gradient at no more than one point of  $C$ , and
- (iii)  $f(\cdot)$  assumes its minimum value over  $C$  at an interior point of  $C$  if and only if the gradient of  $f(\cdot)$  vanishes at that point.

*Proof:*

- (i) The hypotheses imply that  $f(\cdot)$  is a continuous function and that  $C$  is a compact set. It follows that  $f(\cdot)$  assumes its minimum value over  $C$ ,  $f_{\min}$ , at some point of  $C$ . Now suppose that  $\mathbf{x}_1, \mathbf{x}_2$  are distinct points of  $C$  such that

$$f(\mathbf{x}_1) = f(\mathbf{x}_2) = f_{\min}.$$

Then the strict convexity of  $f(\cdot)$  implies that

$$f\left(\frac{\mathbf{x}_1}{2} + \frac{\mathbf{x}_2}{2}\right) < f_{\min}.$$

This is a contradiction of the hypothesis that  $f_{\min}$  is the least value of  $f(\cdot)$  over  $C$ .

- (ii) Suppose the gradient of  $f$  vanishes at two distinct points of  $C$ ,  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . Consider the directional derivative of  $f(\cdot)$  along  $(\mathbf{x}_2 - \mathbf{x}_1)$ . By hypothesis, this derivative vanishes at  $\mathbf{x}_1$ . Because  $f(\cdot)$  is strictly convex, it is a strictly increasing function of position along  $(\mathbf{x}_2 - \mathbf{x}_1)$ . Therefore, it is greater than zero at  $\mathbf{x}_2$ , which contradicts the hypothesis that the gradient vanishes at  $\mathbf{x}_2$ .
- (iii) Suppose  $f(\mathbf{x}_0) = f_{\min}$  and  $\mathbf{x}_0$  is an interior point of  $C$ . All points in a neighborhood of  $\mathbf{x}_0$  are in  $C$ . If the gradient of  $f(\cdot)$  does not vanish at  $\mathbf{x}_0$ , then there are points in this neighborhood, along the negative gradient from  $\mathbf{x}_0$ , at which  $f(\cdot)$  assumes smaller values. This contradicts the hypothesis that  $f_{\min}$  is the least value of  $f(\cdot)$  in  $C$ .

Now suppose that the gradient of  $f(\cdot)$  vanishes at  $\mathbf{x}_0$ , an interior point of  $C$ . Let  $\mathbf{x}_1$  be any other point of  $C$  and consider the directional derivative of  $f(\cdot)$  along  $(\mathbf{x}_1 - \mathbf{x}_0)$ . This vanishes at  $\mathbf{x}_0$  by hypothesis. The strict convexity of  $f(\cdot)$  implies that it is strictly increasing from  $\mathbf{x}_0$  to  $\mathbf{x}_1$ . Therefore,  $f(\mathbf{x}_1) > f(\mathbf{x}_0)$  and  $f(\cdot)$  has its minimum over  $C$  at  $\mathbf{x}_0$ .

This completes the proof.

In considering the properties of  $L(\mathbf{w})$ , an additional result will be needed concerning the penalty function  $r(\cdot)$ . Its properties are numbered (i) through (v). Note that the convergence in (v) was not assumed to be uniform over the unit sphere. This property will now be deduced.

Because  $r(\cdot)$  is a strictly convex function on the  $M$ -dimensional vector space, it is also strictly convex on any subspace. In particular,  $r(s\mathbf{x}_1)$  is a strictly convex function of  $s$  for any  $\mathbf{x}_1$  on the unit sphere.

Then, for any real number  $P$ , however large,

$$e(s, \mathbf{x}_1) \equiv r(s\mathbf{x}_1) - sP$$

is a strictly convex function of  $s$ . Let  $E_s$  be the set of vectors,  $\mathbf{x}$ , on the unit sphere for which

$$e(s, \mathbf{x}) \leq 0.$$

For any fixed  $s$ ,  $E_s$  is a closed subset of the unit sphere because  $e(s, \mathbf{x})$  is continuous. It follows that  $E_s$  is a compact set.

The corollary to Theorem 5 tells us that along any ray from the origin, i.e., for  $s$  going from zero to infinity, the derivative of  $e(s, \mathbf{x}_1)$  is strictly monotone increasing. From this it can be seen that  $\mathbf{x}_1 \notin E_{s_0}$  implies that  $\mathbf{x}_1 \notin E_s$  for  $s > s_0$ . Therefore, the sets  $E_s$  decrease as  $s$  increases from zero to infinity.

The intersection of a class of compact, decreasing, nonempty sets is nonempty. Therefore, if

$$\bigcap_{s=0}^{\infty} E_s = 0,$$

it is clear that there exists an  $s_0$  such that  $E_s = 0$  for  $s \geq s_0$ . In this case,

$$r(s\mathbf{x}) > sP \quad \text{for all } s \geq s_0(P)$$

independent of  $\mathbf{x}$  on the unit sphere.

On the other hand, if

$$\bigcap_{s=0}^{\infty} E_s \neq 0,$$

then there exists a unit vector  $\mathbf{x}_1$ , such that

$$r(s\mathbf{x}_1) \leq sP \quad \text{for all } s \text{ in } (0, \infty).$$

This contradicts hypothesis (v) concerning the penalty functions.

Thus, we have seen that  $r(s\mathbf{x})/s$  goes to infinity with  $s$  uniformly for all  $\mathbf{x}$  on the unit sphere.

Now consider the function  $L(\mathbf{w})$ , defined by (46), with  $\mathbf{w}$  restricted to an  $(N - 1)$ -dimensional linear manifold orthogonal to  $\mathbf{1}_N$ . Let

$$\mathbf{w} = \mathbf{v} + \alpha \mathbf{1}_N$$

where  $\alpha$  is a real number and  $\mathbf{v}$  is restricted to the  $(N - 1)$ -dimensional linear subspace orthogonal to  $\mathbf{1}_N$ .

In view of (38) and (42),  $L(\mathbf{w})$  can be put in the form,

$$L(\mathbf{w}) = \bar{L}(\mathbf{v}) = \mathbf{c}^T \cdot \mathbf{v} + d + r(KA\mathbf{v} + \mathbf{b}). \quad (66)$$

Here,  $\mathbf{c}$  is a fixed  $N$ -dimensional vector,  $d$  is a scalar constant, and  $\mathbf{b}$  is a fixed  $M$ -dimensional vector, with  $M \geq N$ . The fixed  $N \times N$  matrix  $A$  is positive definite. It can be seen from the discussion in Appendix A and (38) that the fixed  $M \times N$  matrix  $K$  is of rank  $(N - 1)$  for connected systems. Its null space is spanned by  $\mathbf{1}_N$ .

Using the convexity of  $r(\cdot)$ ,

$$\begin{aligned} r(\tfrac{1}{2}KA\mathbf{v}) &\leq \tfrac{1}{2}r(-\mathbf{b}) + \tfrac{1}{2}r(KA\mathbf{v} + \mathbf{b}) \\ r(KA\mathbf{v} + \mathbf{b}) &\geq 2r(\tfrac{1}{2}KA\mathbf{v}) - r(-\mathbf{b}). \end{aligned}$$

Using this in (66),

$$L(\mathbf{w}) = \bar{L}(\mathbf{v}) \geq 2r(\tfrac{1}{2}KA\mathbf{v}) + \mathbf{c}^T \cdot \mathbf{v} + d - r(-\mathbf{b}).$$

The right-hand member is dominated by its first term as  $|\mathbf{v}|$  becomes large, uniformly over the subspace orthogonal to  $\mathbf{1}_N$ . Therefore, we can find a sphere of sufficiently large radius so that

$$\bar{L}(\mathbf{v}) > \bar{L}(0)$$

for all  $\mathbf{v}$  on its surface. Then the minimum value of  $\bar{L}(\mathbf{v})$  over this sphere is not assumed on the boundary.

Now it will be shown that  $\bar{L}(\mathbf{v})$  is a strictly convex function of  $\mathbf{v}$  on the subspace orthogonal to  $\mathbf{1}_N$ . Let

$$\begin{aligned} \mathbf{x}_1 &= KA\mathbf{v}_1 + \mathbf{b} \\ \mathbf{x}_2 &= KA\mathbf{v}_2 + \mathbf{b}. \end{aligned}$$

The properties of  $K$  and  $A$  are such that  $\mathbf{v}_1 \neq \mathbf{v}_2$  implies  $\mathbf{x}_1 \neq \mathbf{x}_2$ . This

permits us to apply the strict convexity condition,

$$\begin{aligned} r[K\lambda\mathbf{v}_2 + KA(1 - \lambda)\mathbf{v}_1 + \mathbf{b}] \\ &= r[\lambda\mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1] < \lambda r(\mathbf{x}_2) + (1 - \lambda)r(\mathbf{x}_1) \\ &= \lambda r(KA\mathbf{v}_2 + \mathbf{b}) + (1 - \lambda)r(KA\mathbf{v}_1 + \mathbf{b}). \end{aligned}$$

This is sufficient to establish the strict convexity of  $\bar{L}(\mathbf{v})$ .

Inasmuch as the strictly convex function  $\bar{L}(\mathbf{v})$  takes on its minimum value over every large sphere at an interior point, its gradient vanishes uniquely at that point.

The above statement applies to the restriction of  $\bar{L}(\mathbf{v})$  to an  $(N - 1)$ -dimensional space. However, we know that  $\mathbf{w}'(t)$  vanishes along  $\mathbf{l}_N$ . Equation (47) then shows that

$$\nabla_{\mathbf{w}}L(\mathbf{w}) = -A^T\mathbf{w}'(t)$$

vanishes along a direction which is not orthogonal to  $\mathbf{l}_N$ . It follows that the unrestricted gradient of  $L(\mathbf{w})$  vanishes at a unique point of every linear manifold orthogonal to  $\mathbf{l}_N$ .

#### REFERENCES

1. Vaughan, H. E., Research Model for Time-Separation Integrated Communication, B.S.T.J., 38, July, 1959, pp. 909-932.
2. Beneš, V. E., unpublished work.
3. Runyon, J. P., Reciprocal Timing of Time Division Switching Centers, U. S. Patent No. 3,050,586, August 21, 1962.
4. Goldstein, A. J., unpublished work.
5. Beneš, V. E. and Goldstein, A. J., unpublished work.
6. Darwin, G. P. and Prim, R. C., Synchronization in a System of Interconnected Units, U. S. Patent No. 2,986,723, May, 1961.
7. Karafin, B. J., unpublished work.
8. Gersho, A. and Karafin, B. J., Mutual Synchronization of Geographically Separated Oscillators, B.S.T.J., this issue, pp. 1689-1704.
9. Candy, C. J. N. and Karnaug, M., to be published.
10. Brilliant, M. B., The Determination of Frequency in Systems of Mutually Synchronized Oscillators, B.S.T.J., this issue, pp. 1737-1748.
11. Brilliant, M. B., Dynamic Response of Systems of Mutually Synchronized Oscillators, B.S.T.J., to be published.
12. Saito, T., Fujisaki, H., and Inose, H., unpublished work.
13. Byrne, C. J., Karnaug, M., and Scattaglia, J. V., Retiming of Digital Signals with a Local Clock, 1960 Northeast Electronic Res. and Eng. Meeting, NEREM Record, November, 1960.
14. Byrne, C. J., Karafin, B. J., and Robinson, D. B., Jr., Systematic Jitter in a Chain of Digital Regenerators, B.S.T.J., 42, November, 1963, pp. 2679-2714.
15. LaSalle, J. and Lefschetz, S., *Stability by Liapunov's Direct Method*, Academic Press, New York, 1961.
16. Hahn, W., *Theory and Application of Liapunov's Direct Method*, Prentice-Hall, Englewood Cliffs, New Jersey, 1963.



# The Determination of Frequency in Systems of Mutually Synchronized Oscillators

By M. B. BRILLIANT

(Manuscript received September 7, 1966)

*The synchronization of large systems of geographically separated oscillators is of considerable practical interest for pulse code modulation (PCM) switching. This study examines the factors that determine the frequency at which such a system operates, considering both the procedure by which it is set up and the topology of system interconnections. A necessary and sufficient connectivity condition is established.*

## I. INTRODUCTION

The synchronization of large systems of geographically separated oscillators is of considerable practical interest for pulse code modulation (PCM) switching. Synchronization could be achieved by establishing a single master oscillator, with every other oscillator slaved either directly to the master or to another oscillator that is slaved directly or indirectly to the master. However, the system would then be vulnerable to failure of a single link or a single oscillator. An alternative called "mutual synchronization" would permit the oscillators to determine the system frequency jointly and to exchange synchronization information over redundant paths. However, the complexity of the system raises questions concerning the factors that determine the system frequency as well as system stability and dynamic response.

A broad sufficient condition for the stability of mutually synchronized systems was first established by Beneš.<sup>1</sup> This condition has recently been rederived by a different method, for a slightly more general system, by Gersho and Karafin.<sup>2</sup> The model used in both these studies was oversimplified so that it gave a paradoxical result for the system frequency at equilibrium. A model that corrected this oversimplification, by considering the received signal phases observed at each oscillator station at the initial moment when all oscillator controls are put into operation, was first devised by Runyon.<sup>3</sup> A corrected model based on the same

principle, but differing in detail, has recently been independently derived by Karbaugh.<sup>4</sup> In all these studies, it was assumed that the system was so interconnected that every oscillator transmitted timing information either directly or indirectly to every other oscillator. This condition has been generally assumed to be necessary for mutual synchronization, and has been proved sufficient by Gersho and Karafin.<sup>2</sup>

This paper will generalize the foregoing results in two ways that appear to be significant for practical applications. In the first place, a mathematical model will be described that allows the synchronized system to be set up by less drastic methods than the simultaneous closure of all control paths at  $t = 0$ . In the second place, a weaker connectivity condition, which is satisfied by systems in which only some of the oscillators participate in frequency determination, will be proved necessary and sufficient for synchronization.

The practical consequence of these generalizations is that a system with a single master oscillator can be regarded as a special case within the general class of mutually synchronized systems, and a locked oscillator synchronized to a remote source can be regarded as a special case of an oscillator station in a mutual synchronization system. Between the extremes of a system with no slaves and a system in which all stations but one are slaves, a variety of hierarchical organizations may be envisioned. However, the description of particular configurations is beyond the scope of this article. The model developed here also provides the flexibility by which new stations can be added to an existing system, and the system frequency can be adjusted after synchronism has been established.

## II. THE MATHEMATICAL MODEL

The system is assumed to consist of  $N$  oscillators, or "clocks," numbered  $i = 1, \dots, N$ . Each oscillator has its own free-running frequency  $f_i$ , at which it would operate in the absence of a control input. Each oscillator accepts a control input that causes its frequency to deviate from the free-running frequency by an amount proportional to the control input. For concreteness, the control input will be referred to as a voltage, although it may, in practice, take other forms. Thus, the instantaneous frequency of the  $i$ th oscillator, which will be expressed simply as the rate of change of phase  $p_i'(t)$ , will, in general, be different from the free-running frequency  $f_i$ .

In applications to switched PCM networks, each oscillator controls the timing of a digital signal which is assumed to be organized with a fixed number of pulses per frame. It will be convenient to measure phase

in frames of the digital signal, and frequency in frames per second. Each station sends a digital signal, controlled by its local oscillator, to a number of other stations, and this signal conveys timing information. To simplify the description, it will be assumed that all these signals are sent in the same phase, and this will be taken to define the phase of the local clock. However, the model could easily be adapted to the case in which each signal is sent in some arbitrary but fixed phase with respect to the local clock.

The transmission delay from the  $j$ th station to the  $i$ th will be designated as  $\tau_{ij}$ . Thus, the phase of the signal received at the  $i$ th station from the  $j$ th is  $p_j(t - \tau_{ij})$ . The phase is defined principally by a regular pattern of framing pulses. The pulses between the framing pulses carry information, and are therefore different in successive frames. Since successive frames are distinguishable, the cyclic ambiguity inherent in the measurement of the phase of sinusoidal signals is not inherent in the digital case.

Thus it is possible to measure, at the  $i$ th station, the phase difference  $p_j(t - \tau_{ij}) - p_i(t)$  between the received signal from the  $j$ th station and the local clock. This phase difference will be called the "observed phase" of the  $j$ th signal at the  $i$ th station. In the Beneš<sup>1</sup> model, used also by Gersho and Karafin,<sup>2</sup> the control voltage at the oscillator consists only of components proportional to the observed phases. However, as Gersho and Karafin<sup>2</sup> pointed out, if all the clocks are in phase all the observed phases will be negative, and every clock will be made to run slower than its free-running frequency. In the present model, a fixed reference phase  $r_{ij}$  will be subtracted from each observed phase, this reference phase preferably being equal to the phase difference one would expect to observe. If the observed phase of each signal is equal to the reference phase, no control voltage is applied to the oscillator, which then runs at its free-running frequency.

Historically,<sup>1</sup> the concept of mutual synchronization evolved in terms of phase averaging. Thus, the observed phases of the received signals were respectively multiplied by nonnegative averaging coefficients  $a_{ij}$ ,

$$\sum_{j=1}^N a_{ij} = 1, \quad (1)$$

to form an average phase difference between the local clock and the signals received from its neighbors. The average observed phase may then be multiplied by a nonnegative factor  $\lambda_i$ , having the dimensions of inverse time, to determine the frequency displacement of the local clock. This basic notation has been continued in subsequent studies and will

be used here. Thus, in the present model, the system equations must be

$$p_i'(t) = f_i + \lambda_i \sum_{j=1}^N a_{ij}[p_j(t - \tau_{ij}) - p_i(t) - r_{ij}], \quad (2)$$

$$i = 1, \dots, N.$$

The reference phases can be absorbed into the free-running frequency term by defining a reference frequency

$$v_i = f_i - \lambda_i \sum_{j=1}^N a_{ij}r_{ij}, \quad i = 1, \dots, N. \quad (3)$$

The system equations can now be written as

$$p_i'(t) = v_i + \lambda_i \sum_{j=1}^N a_{ij}[p_j(t - \tau_{ij}) - p_i(t)], \quad i = 1, \dots, N. \quad (4)$$

These equations have formally reverted to those of the Beneš model,<sup>1,2</sup> in which the reference phases do not appear. However, while the equations are the same, their application is different, since  $v_i$  is not the free-running frequency, but is normally greater than the free-running frequency, because the reference phases  $r_{ij}$  are normally negative. The reference phases may, in fact, be identified with the initial-condition terms of Runyon<sup>3</sup> and Karnaugh,<sup>4</sup> so that (4) covers their models as well as the Beneš model.

The dynamic response of the system can be modified by using a filter in each control system. Multiplication by  $\lambda_i$  is then replaced by convolution with the impulse response  $h_i(t)$  of a filter whose zero-frequency gain is

$$\int_0^{\infty} h_i(t) dt = \lambda_i, \quad i = 1, \dots, N. \quad (5)$$

This has been done in all the referenced studies. Gersho and Karafin<sup>2</sup> also added a variable term to  $v_i$ , replacing it formally by  $v_i(t)$ , to represent the effects of transient disturbances. The system equations therefore become, in the most general form to be used here,

$$p_i'(t) = v_i(t) + h_i(t) * \sum_{j=1}^N a_{ij}[p_j(t - \tau_{ij}) - p_i(t)], \quad (6)$$

$$i = 1, \dots, N.$$

where the asterisk (\*) denotes convolution. Neither of these changes affects the equilibrium frequency.

It has been assumed that the filter gains and averaging coefficients are

all nonnegative,

$$\lambda_i \geq 0, \quad (7)$$

$$a_{ij} \geq 0. \quad (8)$$

The connectivity of the network depends on which of these coefficients are zero. If the  $i$ th station does not receive from the  $j$ th,  $a_{ij}$  is zero, except when the  $i$ th station does not receive from any other station, in which case (1) forbids all  $a_{ij}$  to be zero and  $\lambda_i$  must be zero, and the  $a_{ij}$  are then arbitrary. It is understood that if the  $i$ th station in fact receives a digital signal from the  $j$ th, but uses it only as a medium of communication and does not use its observed phase in controlling its clock, it will be said that the  $i$ th station "does not receive from" the  $j$ th.

### III. THE INITIATION OF SYNCHRONOUS OPERATION

Previous studies have assumed that the system is placed in synchronous operation at  $t = 0$  by simultaneously closing all the switches at each station that connect the control voltages to the oscillators. It is assumed that before  $t = 0$  all oscillators are operating at their free-running frequencies, and have been running for a sufficiently long time so that, in spite of transmission delays, all stations are receiving signals on all links by the time the switches are closed. Closure of each switch will, in general, cause an immediate change of frequency at every station, and prediction of the frequency at which the system finally will settle down would be a matter of practical importance.

In practice it may be preferable to assemble the system in more leisurely fashion — one station at a time — checking for proper operation after each station is connected before connecting the next one. One might, for example, realize the reference phases  $r_{ij}$  as manually controlled bias voltages. When a new station is to be connected into the system, the first connection will be made at the new station, from one of the phase detectors to the input of the clock control filter. This connection will synchronize the new station with the system as a slave station, and adjustment of the corresponding reference phase can be used to establish any desired phase relation between it and the rest of the system. When each subsequent connection is made from a phase detector to a clock control filter, the associated reference phase is adjusted so as to null the voltage across the switch at the moment when it is closed. There is then no discontinuous change in frequency at any time during the connection process.

If the system were built up in this way, starting from one station as the

initial system, and if there were no drifts in either free-running frequencies or transmission delays, the final equilibrium frequency of the system would be the free-running frequency of the first station. In this case the equilibrium frequency could be predicted without any calculation. In any case, the system frequency can be deliberately changed after initiation of synchronous operation by adjusting the bias voltages.

The equation for system frequency is still useful as a means of predicting the effects of drifts in the free-running frequencies and the transmission delays. However, serious questions can in principle arise with regard to the applicability of the general system equations (6). These equations represent a system with invariant connectivity, represented by invariant averaging coefficients  $a_{ij}$  and gains  $\lambda_i$ , while the actual system connectivity has been a function of time. Karnough<sup>4</sup> and Gersho and Karafin<sup>2</sup> have answered these questions for their models under the particular initiation procedure they assumed. The answer will now be extended to cover the present model for arbitrary initiation procedure.

I shall take the point of view that there is some specifiable moment  $t_0$  at which the system has been completely assembled, so that the  $a_{ij}$  and  $\lambda_i$  are invariant for  $t \geq t_0$ , and that we need only to predict the future behavior of the system, for specified disturbances in  $v_i(t)$  and drifts in  $\tau_{ij}$ , having full knowledge of the past behavior of the system. For the purpose of this discussion, if the transmission delays  $\tau_{ij}$  are to be allowed to change, they should be considered as having been written as  $\tau_{ij}(t)$ .

The system equations (6) are actually integrodifferential equations, since the convolution symbol (\*) implies an integration. The initial conditions on which the solution of this equation depends are the entire history of the phase variables  $p_i(t)$ , to the extent that this history determines the state of the filters. The output of each filter for  $t \geq t_0$  can be considered as the sum of two components: a transient term determined by the state of the filter, which is in turn determined by the input for  $t < t_0$ , and a term representing the response to inputs for  $t \geq t_0$ . The transient terms can be calculated from the known filter inputs for  $t < t_0$  and included in the  $v_i(t)$  terms. Equations (6), with these terms included in  $v_i(t)$ , with the filters considered quiescent at  $t = t_0$ , and with the correct initial values of  $p_i(t_0)$ , will then give a correct description of the behavior of the system for  $t \geq t_0$ .

This argument is included here only to establish the validity of (6) in principle. In practical calculations, estimates of the effects of transient disturbances would normally assume an equilibrium state as the initial condition.

## IV. EQUILIBRIUM STATES

Gersho and Karafin<sup>2</sup> determined the equilibrium frequency of the system as a limiting value derived by means of the final value theorem for Laplace transforms. Karnaug<sup>4</sup> used a simpler method, claiming for it only heuristic value. The following approach claims rigorous validity for the simpler method.

The first step in the analysis of the system will be the determination of its equilibrium states, without regard for whether they are stable or unstable equilibria. These can be determined by assuming that the system has been placed in some state, and that it will not change state spontaneously; any state that satisfies these conditions is an equilibrium state. We determine in this step whether the equilibrium state is unique. The second step is to determine whether the system can respond to any transient excitation with components that do not approach zero with increasing time; this step determines the stability of the system. The linearity of the system now implies that if the equilibrium state is unique, and the transient response approaches zero, the system will always approach the equilibrium state in the absence of a disturbance.

It will, in fact, be found that the equilibrium state is not unique, because the system equations include the phases only in phase difference terms, and an arbitrary common constant can be added to every phase variable without changing the phase differences. There is, therefore, a continuum of equilibrium states, all of which are equivalent for practical purposes in that they have the same phase differences and the same system frequency. Because of this equivalence, the system will be considered stable if, after a transient disturbance, it approaches any equilibrium state, not necessarily the one it occupied before the disturbance. This requires only that the transient components of the phase differences approach zero, while the transient components of the phases may approach arbitrary limits.

This section will deal only with the first step: the identification, including determination of conditions for existence, of equilibrium states. The stability of the equilibrium states can be assured by the sufficient condition studied by Gersho and Karafin;<sup>2</sup> their proof remains valid under the weaker connectivity condition shown here to be necessary and sufficient, the statement that at least one  $M_{i1}$  is positive sufficing to replace their statement that all  $M_{i1}$  are positive.

The only equilibrium states to be considered here are those in which all clock frequencies are constant at a common value; if they are con-

stant, but at different values, synchronism has not been established. It will also be required that the existence of such a state should not depend on the values of the free-running frequencies; if it does, the system is not self-synchronizing, but is synchronous only if the clocks are adjusted by means external to the system.

The instantaneous frequencies  $p_i'(t)$  are, therefore, set equal to an equilibrium frequency denoted simply by  $f$ . Then

$$p_i(t) = ft + P_i, \quad i = 1, \dots, N. \quad (9)$$

Equation (4), which suffices even in the most general linear case represented by (6) for the description of the steady state, becomes

$$f = v_i + \lambda_i \sum_{j=1}^N a_{ij}(P_j - P_i - f\tau_{ij}), \quad i = 1, \dots, N, \quad (10)$$

or, in more symmetrical form,

$$\sum_{j=1}^N \lambda_i(\delta_{ij} - a_{ij})P_j = v_i - f(1 + \lambda_i\tau_i), \quad i = 1, \dots, N, \quad (11)$$

where

$$\tau_i = \sum_{j=1}^N a_{ij}\tau_{ij} \quad (12)$$

and  $\delta_{ij}$  is the Kronecker delta, equal to unity for  $i = j$  and zero otherwise. The set of equations (11) looks as though it could be solved for the  $P_j$  in terms of arbitrary  $v_i$  and  $f$ , but it cannot, because the matrix of coefficients on the left is singular. This will be stated and proved as a theorem.

*Theorem I:* Let  $L$  denote the diagonal matrix with diagonal elements  $\lambda_i$ , let  $A$  denote the averaging matrix with elements  $a_{ij}$ , and let  $I$  denote the identity matrix. Then the matrix  $L(I - A)$ , with elements  $\lambda_i(\delta_{ij} - a_{ij})$ , has rank less than its order  $N$ .

*Proof:* By (1), the sum of the elements in any row is zero. Therefore, the sum of all columns is a column of zeros. Therefore the matrix is singular and its rank is less than its order, Q.E.D.

It is advantageous at this point to choose one  $P_i$  arbitrarily as a reference for the others. With  $P_1$  as reference, we change to the phase difference variables

$$Q_j = P_j - P_1, \quad j = 2, \dots, N. \quad (13)$$

The equations (11) then become

$$\sum_{j=2}^N \lambda_i(\delta_{ij} - a_{ij})Q_j = v_i - f(1 + \lambda_i\tau_i), \quad i = 1, \dots, N. \quad (14)$$

If the term in  $f$  is transposed to the left side, we get the set of  $N$  equations

$$f(1 + \lambda_i\tau_i) + \sum_{j=2}^N \lambda_i(\delta_{ij} - a_{ij})Q_j = v_i, \quad i = 1, \dots, N, \quad (15)$$

which we expect to be able to solve for the  $N$  unknowns  $f$  and  $Q_j, j = 2, \dots, N$ , for arbitrary  $v_i$ .

If we formally solve for  $f$  by determinants, and expand each determinant in terms of the elements of the first column and their cofactors, the result is

$$f = \frac{\sum_{i=1}^N b_i v_i}{\sum_{i=1}^N b_i(1 + \lambda_i\tau_i)}, \quad (16)$$

where  $b_i$  is the cofactor (signed minor) of the element in the first column,  $i$ th row, of the matrix  $L(I - A)$ . The following theorem shows that the arbitrary choice of  $P_1$  as the reference for phase differences, and the definition of  $b_i$  in terms of the first column, makes no difference in the result.

*Theorem II: Let  $M_{ij}$  be the cofactor of the  $(i, j)$ th element of  $L(I - A)$ ; then  $M_{ij} = M_{ik}$  for all  $i, j, k$ , that is, all cofactors of elements in the same row are equal, and hence  $M_{ij} = b_i$  for all  $i, j=1, \dots, N$ .*

*Proof:* If the rank of  $L(I - A)$  is less than  $N - 1$  then all  $M_{ij}$  are zero and the theorem is satisfied. If the rank is  $N - 1$ , the matrix equation  $L(I - A)x = 0$ , where  $x$  is an  $N$  element column matrix, has only one independent solution. It is known from (1) that a solution exists in which all components are equal, and this must now be true of any solution. It can also be shown that the cofactors of any single row of the matrix  $L(I - A)$  must be a solution (see, for example, Guillemin<sup>6</sup>), hence all cofactors of elements in a row must be equal,  $M_{ij} = M_{ik}$ , hence  $M_{ij} = M_{i1} = b_i$ , Q.E.D.

Since  $b_i$  can now be defined without reference to any particular column, the single-index notation is justified. Since we expect that increasing the free-running frequency of any oscillator will never decrease the equilibrium frequency  $f$ , we should expect all the  $b_i$  to be nonnegative. The following theorem verifies this expectation.

*Theorem III:* The cofactors of elements of the matrix  $L(I - A)$  are nonnegative,  $b_i \geq 0$ ,  $i = 1, \dots, N$ .

*Proof:* If a matrix is diagonally dominated, i.e., if every diagonal element is greater in magnitude than the sum of the magnitudes of all other elements in the same row, it is easily shown (Appendix I, Gersho and Karafin<sup>2</sup>) that it must be nonsingular. Consider the matrix  $L(I - \varepsilon A)$ ,  $0 \leq \varepsilon \leq 1$ . The cofactors of its diagonal elements are continuous functions of  $\varepsilon$ . For  $\varepsilon = 0$  they are all unity, hence positive. For  $0 < \varepsilon < 1$  the cofactors are the determinants of diagonally dominated submatrices, hence nonzero, so that they cannot pass through zero, and must remain positive. Hence, as  $\varepsilon \rightarrow 1$  they cannot approach negative limits. But as  $\varepsilon \rightarrow 1$  they approach the values  $b_i$ , hence  $b_i \geq 0$ , Q.E.D.

The formal solution (16) is valid if and only if the matrix of coefficients on the left side of (15) is nonsingular; that is, if and only if the denominator of (16) is nonzero. But, since  $b_i$ ,  $\lambda_i$ , and  $\tau_i$  are all nonnegative, this is equivalent to the condition that at least one  $b_i$  be positive. The following definitions and theorems relate this algebraic condition to the connectivity properties of the network.

*Definition:* The  $j$ th station is said to send to the  $i$ th, or equivalently, the  $i$ th station is said to receive from the  $j$ th, if  $\lambda_i a_{ij}$  is positive.

*Definition:* The  $j$ th station is said to send directly or indirectly to the  $i$ th, or equivalently, the  $i$ th station is said to receive directly or indirectly from the  $j$ th, if there exists a chain (ordered set) of stations such that the first is station  $j$ , the second receives from  $j$ , each receives from the one before, and the last is station  $i$ .

*Theorem IV:* If the  $k$ th station does not transmit directly or indirectly to all other stations then  $b_k = 0$ .

*Proof:* Let  $A_{kk}$  be the submatrix formed by deleting the  $k$ th row and column of  $L(I - A)$ . Let  $S_k$  be the set of indices of all stations that do not receive directly or indirectly from the  $k$ th. By hypothesis  $S_k$  is nonempty; choose  $i \in S_k$ . By the definition of  $S_k$ ,  $\lambda_i a_{ij}$  is zero if  $j$  is not in  $S_k$ , hence, from (1),

$$\sum_{j \in S_k} \lambda_i a_{ij} = \lambda_i, \quad i \in S_k. \quad (17)$$

Let  $B_k$  be the square submatrix of  $A_{kk}$  consisting of all elements whose row and column indices are both in  $S_k$ . Then (17) shows that  $B_k$  can be written in the form  $L'(I - A')$ , where  $A'$  is an averaging matrix satisfying (1). Hence, by Theorem I,  $B_k$  is singular, and the rows of  $B_k$  are

linearly dependent. But since the  $i$ th row of  $A_{kk}$ , for all  $i \in S_k$ , is the  $i$ th row of  $B_k$  augmented with zeros, the same linear dependence holds among the rows of  $A_{kk}$ . Hence,  $A_{kk}$  is singular; hence its determinant, which is  $b_k$ , is zero, Q.E.D.

*Theorem V: If  $b_k$  is zero, then the  $k$ th station does not transmit directly or indirectly to every other station.*

*Proof:* Let  $A_{kk}$  be defined as in the proof of Theorem IV. By hypothesis,  $A_{kk}$  is singular; hence there exists a column matrix  $x$ , with elements  $x_i$  not all zero, such that  $A_{kk}x = 0$ , or equivalently

$$\sum_{j \neq k} \lambda_i a_{ij} x_j = \lambda_i x_i, \quad i \neq k. \quad (18)$$

Let  $M$  be the magnitude of the  $x_i$  having the largest magnitude. Let  $S_k$  be defined now as the set of all indices  $i$  for which  $|x_i| = M$ ; obviously  $S_k$  is nonempty. Now (18) implies

$$\left| \sum_{j \neq k} \lambda_i a_{ij} x_j \right| = \lambda_i M, \quad i \in S_k. \quad (19)$$

Now (1) and  $|x_j| \leq M$  imply that this can be true only if  $\lambda_i a_{ik} = 0$  and  $|x_j| = M$  whenever  $\lambda_i a_{ij} > 0$ . Hence, for all  $i \in S_k$ ,  $\lambda_i a_{ij} = 0$  except when  $j \in S_k$ , and thus the  $i$ th station cannot receive directly or indirectly from the  $k$ th. Since  $S_k$  is nonempty, the  $k$ th station does not transmit directly or indirectly to all stations, Q.E.D.

It follows from these theorems that the formal solution (16) is valid for the set of equations (15) if and only if there is at least one station that transmits directly or indirectly to all other stations.

If there is no such station, the matrix of coefficients on the left side of (15) is singular, and the set of equations has either no solution or an infinity of solutions, depending on the values of the  $v_i$ . Since a solution defines an equilibrium state in which all oscillators run at the same frequency, this means that the oscillators will run at the same frequency only if their free-running frequencies are appropriately adjusted; that is, the system is not self-synchronizing.

If there is only one station that transmits directly or indirectly to all others, that station is the master, setting the frequency for the whole system. A single master receives from no other station, since any station that transmitted to it would thereby transmit indirectly to all other stations. Thus, a station can become a master simply by the loss of all inputs from other stations. However, if two stations lose all their inputs, the system fails to synchronize, since neither station sends directly or indirectly to every other.

If more than one station sends directly or indirectly to all others, these stations are mutually synchronized, and jointly establish the system frequency. Any station that does not send directly or indirectly to every other station is in effect a slave station.

#### V. SUMMARY AND CONCLUSIONS

The process for initiation of synchronous operation described in Section III is not necessarily recommended as the best possible. It is intended as a constructive existence proof, showing that there exists a method of setting up a synchronized system of geographically separated clocks that will lead to a final frequency that can be determined in advance. The second part of that section shows, in perhaps unnecessary detail, that the behavior of a system, once it has been set up, can be determined without considering how it was set up, so that it is not necessary to specify the set-up procedure before studying its steady-state or dynamic behavior.

Under these circumstances the equation for equilibrium frequency developed in Section IV plays no part in the process of setting the system in synchronism and adjusting it to run at the desired frequency. It serves to identify the factors that affect the final frequency and indicate the quantitative effect of each factor, and as such would appear to find its greatest usefulness in the design and control of the configuration of system interconnections.

The connectivity condition evolved in Section IV permits the inclusion of single-master systems in the same general class as completely mutually synchronized systems. It is suggested that these two types are in fact opposite extremes of a more general class in which the most useful configurations may have some intermediate hierarchical form.

#### REFERENCES

1. Beneš, V. E., unpublished memorandum, 1959.
2. Gersho, A. and Karafin, B. J., Mutual Synchronization of Geographically Separated Oscillators, B.S.T.J., this issue, pp. 1673-1704.
3. Runyon, J. P., unpublished memorandum, 1959.
4. Karnaug, M., A Model for the Organic Synchronization of Communications Systems, B.S.T.J., this issue, pp. 1705-1735.
5. Guillemin, E. A., *The Mathematics of Circuit Analysis*, John Wiley & Sons, New York, 1949, pp. 17, 103.

# Intermodulation Noise in FM Systems Due to Transmission Deviations and AM / PM Conversion\*

By T. G. CROSS

(Manuscript received August 2, 1966)

*Two noise contributors in FM systems are: (i) intermodulation noise due to transmission deviations; and (ii) intermodulation noise due to transmission deviations and AM/PM conversion, designated AM/PM intermodulation noise. Expressions for the second- and third-order AM/PM intermodulation noise are derived in terms of transmission medium coefficients and a continuous pre-emphasis characteristic, with the unpre-emphasized baseband signal being simulated by white Gaussian noise. These expressions have been programmed on a digital computer and representative noise responses and properties of AM/PM intermodulation noise were obtained. General properties and characteristics for the two noise contributors are documented in parallel for comparative purposes. It was found that AM/PM intermodulation noise can be a significant noise contributor in FM systems.*

## I. INTRODUCTION

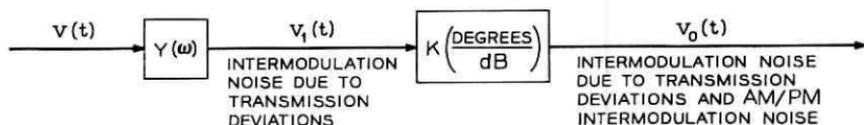
Intermodulation noise is produced whenever a phase modulated signal is passed through a linear transmission medium whose amplitude and phase characteristics are nonlinear functions of frequency. The output signal from this medium is both envelope and phase modulated, with the phase modulation being a distorted replica of the input phase function. The envelope modulation and phase modulation functions are similar in that both consist of first (linear), second-, third-, and higher-order functions of the input phase function. They differ in that the coefficients of the terms making up the two modulating functions are related in different ways to the transmission medium characteristic.

The distortion terms higher than first order, in the output phase

---

\* Portions of this paper were presented at the 1966 IEEE International Communications Conference in Philadelphia, Pa., June 16, 1966.

modulating function, produce intermodulation noise. This source of noise has been the subject of much work over the past ten to twenty years. The envelope distortion terms directly produce no degrading effects in linear systems. However, when the linear transmission medium is followed by a device that converts envelope variations at its input to phase variations at its output then a different noise-generating mechanism exists. This latter source of noise will be designated as "AM/PM intermodulation noise" to distinguish it from the intermodulation noise produced directly by transmission deviations.\* The two phenomena are illustrated in Fig. 1 which depicts the two-step process involved in the



$$v(t) = \text{EXP}^i[\omega_c t + \varphi(t)]$$

$$v_i(t) = \text{EXP}^i a(t) \text{EXP}^i[\omega_c t + \varphi_0(t)]$$

$$v_o(t) = \text{EXP}^i a_i(t) \text{EXP}^i[\omega_c t + \varphi_0(t) + K a(t)]$$

WHERE

$\varphi(t)$  = PHASE MODULATING FUNCTION DUE TO MULTICHANNEL SIGNAL

$\varphi_0(t)$  =  $\varphi(t)$  + PHASE DISTORTION TERMS

$k = 0.1516 K$  = PHASE MODULATION INDEX IN RADIANS DIVIDED BY THE AMPLITUDE MODULATION INDEX

$K$  = AM/PM CONVERSION CONSTANT MEASURED IN  $\frac{\text{DEGREES}}{\text{dB}}$  ASSUMING  $a(t) \ll 1$

$a_i(t) \neq a(t)$  IN GENERAL

$Y(\omega)$  = TRANSMISSION MEDIUM WITH TRANSMISSION DEVIATIONS

Fig. 1 — Model illustrating sources of intermodulation noise due to transmission deviations and AM/PM conversion.

AM/PM intermodulation noise generation. The AM/PM converter will be characterized by the constant  $K$  which has the dimension of degrees/dB and can be interpreted as the peak phase change at the output for a 1-dB change in envelope at the input. In reality, this  $K$  may be a function of a number of quantities, e.g., carrier drive power, frequency (carrier and/or baseband), bias levels, or may even be complex. However, many presently developed broadband radio systems use TWT amplifiers as power output tubes which are often the major source of AM/PM conversion within a radio repeater. These tubes, when driven at moder-

\* Transmission deviations are defined as any deviation in the gain and phase characteristics from the ideal characteristics of constant gain and linear phase for all frequency components of the FM wave.

ate, essentially constant input power level and biased from well controlled sources, are adequately characterized for small envelope fluctuations by a constant  $K$  degrees/dB.<sup>1</sup>

Both noise phenomena are of prime interest in frequency modulated systems. Intermodulation noise due to transmission deviations is of interest because it is a recognized significant noise source. AM/PM intermodulation noise is of interest because of the basic lack of knowledge which has existed on this subject. Due to this deficiency, the AM/PM phenomenon has become the underlying scapegoat for many system problems that appear to be unexplainable using existing system knowledge.

The purpose of this paper is two-fold: (i) to present the mathematical development and ensuing solution for the problem of AM/PM intermodulation noise in FM systems; and (ii) to provide enough general information about the two noise contributors considered in this paper such that one can analyze a system's performance and/or set system requirements with some degree of confidence without having to necessarily utilize the associated digital computer programs.

The analysis to follow considers a linear transmission medium, with generalized transmission deviations, followed by an AM/PM converting device. The baseband signal is simulated by a Gaussian distributed band of noise with flat power density spectrum which is pre-emphasized by a continuous pre-emphasis function before the FM process. The end result of the treatment is the signal-to-noise ratio for second- and third-order AM/PM intermodulation noise. The mathematical framework for this paper is derived from a recent paper which treated the subject of intermodulation noise due to an imperfect transmission medium.<sup>2</sup> Certain facets of that work will be included here for the sake of continuity.

## II. THEORY FOR AM/PM INTERMODULATION NOISE

### 2.1 General Development

Consider the system model shown in Fig. 1 where an FM signal is put into a linear transmission medium followed by an AM/PM converting device. The transfer function of the transmission medium is

$$Y(\omega) = \exp [-\alpha(\omega) - i\beta(\omega)] \quad (1)$$

and the impulse response is

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) \exp(i\omega x) d\omega. \quad (2)$$

The FM signal input to the transmission medium is

$$v(t) = \exp \{i[\omega_c t + \varphi(t)]\} \quad (3)$$

and the output signal is

$$v_1(t) = \exp [a(t)] \exp \{i[\omega_c t + \varphi_o(t)]\}, \quad (4)$$

where  $\omega_c$  is the carrier frequency and  $\varphi(t)$  is the phase modulating signal. Since  $Y(\omega)$  is a linear system, the input and output can be related by

$$v_1(t) = \int_{-\infty}^{\infty} v(t-x)g(x) dx. \quad (5)$$

Substituting (3) and (4) in (5) gives

$$\exp [a(t)] \exp [i\varphi_o(t)] = \int_{-\infty}^{\infty} \exp [i\varphi(t-x) - i\omega_c x]g(x) dx. \quad (6)$$

The output function,  $\varphi_o(t)$ , was the subject of a previous paper<sup>2</sup> and will not be considered further here. Our prime objective is to determine the envelope variation in terms of its functional relationship to the phase modulating signal,  $\varphi(t)$ . It follows from (6) that

$$a(t) = \text{Re} \ln \int_{-\infty}^{\infty} \exp [i\varphi(t-x) - i\omega_c x]g(x) dx. \quad (7)$$

It can be shown that<sup>2</sup>

$$\begin{aligned} a(t) = & -\alpha(f_c) + m_{1i}\varphi' - \frac{m_{2i}}{2!}\varphi'' + \frac{m_{3i}}{3!}\varphi''' - \frac{m_{4i}}{4!}\varphi'''' + \dots \\ & + \frac{l_{2r}}{2}\varphi'\varphi'' - \frac{l_{3r}}{6}\varphi'\varphi''' - \frac{\lambda_{2r}}{2}\varphi'^2 - \frac{l_{5r}}{8}\varphi''^2 + \dots \\ & + \frac{l_{1i}}{4}\varphi'^2\varphi'' - \frac{\lambda_{3i}}{6}\varphi'^3 + \dots + \frac{\lambda_{4r}}{24}\varphi'^4 \dots, \end{aligned} \quad (8)$$

where the subscripts  $r$  and  $i$  denote the real and imaginary parts of the corresponding coefficients, and the prime notation indicates the derivative with respect to time. The argument of the phase functions in (8) is  $t - t_d$  where  $t_d$  is an arbitrary delay.<sup>2</sup> The moments,  $m_n$ , in (8) are related to the transmission medium by

$$m_n = \frac{(-1)^n}{Y(\omega_c)} \left[ \frac{d^n}{d(i\omega)^n} Y(\omega_c + \omega) \exp(i\omega t_d) \right]_{\omega=0} \quad (9)$$

and the  $l$  and  $\lambda$  coefficients are defined as follows:

$$l_1 = m_4 - 2m_1m_3 - m_2^2 + 2m_1^2m_2$$

$$l_2 = m_3 - m_1m_2$$

$$l_3 = m_4 - m_1m_3$$

$$l_5 = m_4 - m_2^2$$

$$\lambda_2 = m_2 - m_1^2$$

$$\lambda_3 = m_3 - 3m_1m_2 + 2m_1^3$$

$$\lambda_4 = m_4 - 4m_1m_3 - 3m_2^2 + 12m_1^2m_2 - 6m_1^4.$$

As an example, we have

$$l_{1i} = m_{4i} - 2m_{2r}m_{2i} - 2m_{1i}m_{3r} - 2m_{1i}^2m_{2i}$$

since  $m_{1r} = 0$  (Appendix I of Ref. 2).

For the following transmission medium

$$Y(\omega + \omega_c) = [1 + g_1\omega + g_2\omega^2 + g_3\omega^3 + g_4\omega^4 + \sum_{j=1}^N u_j \cos(P_j\omega + \theta_j)] \exp \left\{ i[b_2\omega^2 + b_3\omega^3 + b_4\omega^4 + \sum_{j=1}^N \nu_j \sin(q_j\omega + \sigma_j)] \right\} \quad (10)$$

the moments  $m_n$  given by (9) have been evaluated and expressed in terms of the transmission deviations in Appendix I of Ref. 2.

For the analysis to follow, the transmission deviations in (10) are limited to values typically encountered in broadband radio relay systems. However, the ripple type transmission deviations must have ripple periods greater than approximately twice the top baseband frequency. These restrictions are dictated by the limited number of terms of  $a(t)$  which are to be considered.

Referring once again to Fig. 1, we see that when the output signal from  $Y(\omega)$  passes through the AM/PM converter the envelope perturbation given by  $a(t)$  is converted into a phase perturbation, given by  $k a(t)$ . The  $k$  coefficient is related to  $K$  (degrees/dB) as follows: the envelope distortion term expressed in dB is

$$20 \log \frac{\exp [a(t)]}{1} \text{ dB} = 8.686 a(t) \text{ dB}$$

so the phase distortion, due to envelope perturbations, after AM/PM conversion is

$$8.686K a(t) \text{ degrees} = 0.1516 K a(t) \text{ radians.}$$

We will let

$$k = 0.1516 K \text{ radians.}$$

Hence, the phase distortion function, due to envelope variations, after the AM/PM converting device is

$$k a(t) \text{ radians,} \quad (11)$$

where  $a(t)$  is given by (8).

The analysis up to here has been perfectly general (except for the assumption that the AM/PM process may be represented by a constant factor). The terms in (8) consist of first- (linear), second-, third- and higher-order functions of the input phase modulating signal,  $\varphi(t)$ . The linear terms produce baseband amplitude distortion which we shall not concern ourselves with in this paper. Also, terms higher than third order will not be considered. This is not an undue restriction because the prime contributors of intermodulation type noise in broadband systems are second- and third-order phase distortion terms. Therefore, neglecting linear, fourth-, and higher-order terms in (8) gives\*

$$k a(t) = k \left[ \frac{l_{2r}}{2} \varphi' \varphi'' - \frac{l_{3r}}{6} \varphi' \varphi''' - \frac{\lambda_{2r}}{2} \varphi'^2 - \frac{l_{5r}}{8} \varphi''^2 \right] \\ + k \left[ \frac{l_{1i}}{4} \varphi'^2 \varphi'' - \frac{\lambda_{3i}}{6} \varphi'^3 \right] \text{ radians.} \quad (12)$$

Using the relationships

$$\frac{d}{dt} \varphi'^2 = 2\varphi' \varphi''$$

$$\frac{d^2}{dt^2} \varphi'^2 = 2\varphi' \varphi''' + 2\varphi''^2$$

$$\frac{d}{dt} \varphi'^3 = 3\varphi'^2 \varphi''$$

in (12) gives

$$k a(t) = \theta_2(t) + \theta_3(t) = \theta_T(t) \text{ radians,} \quad (13)$$

where

\* It should be noted that additional second- and third-order terms exist which are not shown in (8) nor included in (12). These additional terms are considered to be negligible for the transmission deviation constraints previously mentioned.

$$\theta_2(t) = k \left[ -\frac{\lambda_{2r}}{2} + \frac{l_{2r}}{4} \frac{d}{dt} - \frac{l_{3r}}{12} \frac{d^2}{dt^2} \right] \varphi'^2 + k \left[ \frac{l_{4r}}{24} \right] \varphi''^2 \quad (14)$$

with

$$l_{4r} = 4l_{3r} - 3l_{5r} \quad (15)$$

and

$$\theta_3(t) = k \left[ -\frac{\lambda_{3i}}{6} + \frac{l_{1i}}{12} \frac{d}{dt} \right] \varphi'^3 \quad (16)$$

In the Appendix it is shown that the second-order distortion,  $\theta_2(t)$ , and the third-order distortion,  $\theta_3(t)$ , are uncorrelated. Hence, the total AM/PM intermodulation noise power density spectrum, considering only second- and third-order distortions, is the sum of the two individual noise power density spectra.

## 2.2 Second-Order Noise Power Density Spectrum

In this section we will derive the equation for the second-order AM/PM intermodulation noise power density spectrum. The time representation for the second-order phase distortion due to AM/PM conversion was derived in the previous section and is

$$\theta_2(t) = k \left[ -\frac{\lambda_{2r}}{2} + \frac{l_{2r}}{4} \frac{d}{dt} - \frac{l_{3r}}{12} \frac{d^2}{dt^2} \right] \varphi'^2 + k \left[ \frac{l_{4r}}{24} \right] \varphi''^2 \quad (17)$$

The terms in brackets are operators on their respective functions, so (17) can be represented by the block diagram shown in Fig. 2, where

$$\frac{1}{k} G_1(\omega) = \left[ \frac{l_{3r}}{12} \omega^2 - \frac{\lambda_{2r}}{2} \right] + i \left[ \frac{l_{2r}}{4} \omega \right] \quad (18)$$

and

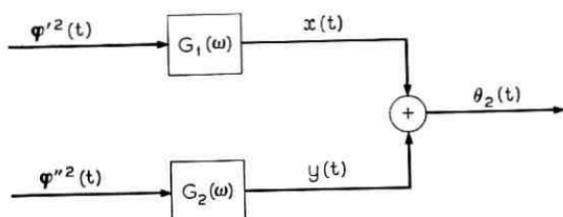


Fig. 2 — Second-order noise block diagram.

$$\frac{1}{k} G_2(\omega) = \frac{1}{24} l_{4r}. \quad (19)$$

It can easily be shown, using the relationship for the cross-correlation of linearly transformed random functions,<sup>3</sup> that the power density spectrum of  $\theta_2(t)$  is

$$S_{\theta_2}(\omega) = G_1(-\omega) G_1(\omega) S_{\varphi'^2}(\omega) + G_1(-\omega) G_2(\omega) S_{\varphi'^2\varphi''^2}(\omega) \\ + G_2(-\omega) G_1(\omega) S_{\varphi''^2\varphi'^2}(\omega) + G_2(-\omega) G_2(\omega) S_{\varphi''^2}(\omega) \quad (20)$$

where, for example,  $S_{\varphi'^2\varphi''^2}(\omega)$  is the cross-power density spectrum of  $\varphi'^2(t)$  and  $\varphi''^2(t)$ . As in Ref. 2, (20) can be expressed as

$$S_{\theta_2}(\omega) = 2 |G_1(\omega)|^2 \mathfrak{F}[R_{\varphi'^2}(\tau)] + 2 |G_2(\omega)|^2 \mathfrak{F}[R_{\varphi''^2}(\tau)] \\ + 2[G_1(-\omega)G_2(\omega) + G_2(-\omega)G_1(\omega)] \mathfrak{F}[R_{\varphi'\varphi''^2}(\tau)], \quad (21)$$

where  $G_1(\omega)$  and  $G_2(\omega)$  are given by (18) and (19), respectively, and  $\mathfrak{F}$  stands for the Fourier transform.

Now, redefining the transfer functions given in (18) and (19) we can write

$$\frac{1}{k^2} S_{\theta_2}(\omega) = 2 |G_1(\omega)|^2 \mathfrak{F}[R_{\varphi'^2}(\tau)] + 2 |G_2(\omega)|^2 \mathfrak{F}[R_{\varphi''^2}(\tau)] \\ + 2[G_1(-\omega)G_2(\omega) + G_2(-\omega)G_1(\omega)] \mathfrak{F}[R_{\varphi'\varphi''^2}(\tau)] \quad (22)$$

where now

$$G_1(\omega) = \left[ \frac{l_{3r}}{12} \omega^2 - \frac{\lambda_{2r}}{2} \right] + i \left[ \frac{l_{2r}}{4} \omega \right] \quad (23)$$

$$G_2(\omega) = \frac{1}{24} l_{4r}. \quad (24)$$

Equation (22) is the second-order AM/PM intermodulation noise power density spectrum weighted by the AM/PM conversion parameter. The ability to pull the  $k$  out of the calculation provides great flexibility.

### 2.3 Third-Order Noise Power Density Spectrum

The time representation for the third-order phase distortion due to AM/PM conversion is, from (16),

$$\theta_3(t) = k \left[ -\frac{\lambda_{3i}}{6} + \frac{l_{1i}}{12} \frac{d}{dt} \right] \varphi'^3 \quad (25)$$

which can be represented by the block diagram in Fig. 3 where

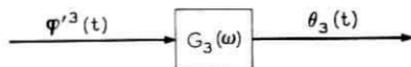


Fig. 3 — Third-order noise block diagram.

$$\frac{1}{k} G_3(\omega) = \left[ -\frac{\lambda_{3i}}{6} \right] + i \left[ \frac{l_{1i}}{12} \omega \right]. \quad (26)$$

It follows that the third-order AM/PM intermodulation noise power density spectrum is

$$S_{\theta_3}(\omega) = |G_3(\omega)|^2 S_{\varphi'^3}(\omega).$$

It can be shown that<sup>6</sup>

$$S_{\varphi'^3}(\omega) = 6 \mathfrak{F}[R_{\varphi'}^3(\tau)] + 9R_{\varphi'}^2(0) S_{\varphi'}(\omega), \quad (27)$$

which can be written

$$S_{\varphi'^3}(\omega) = 6\mathfrak{F}[R_{\varphi'}^3(\tau)] \quad (28)$$

since  $9 R_{\varphi'}^2(0) S_{\varphi'}(\omega)$  is a scaled power density spectrum of the input FM signal and hence can be neglected since it does not contribute to the distortion.\* Therefore,

$$S_{\theta_3}(\omega) = 6 |G_3(\omega)|^2 \mathfrak{F}[R_{\varphi'}^3(\tau)] \quad (29)$$

where  $G_3(\omega)$  is given by (26). Redefining the transfer function we have

$$\frac{1}{k^2} S_{\theta_3}(\omega) = 6 |G_3(\omega)|^2 \mathfrak{F}[R_{\varphi'}^3(\tau)] \quad (30)$$

where now

$$G_3(\omega) = \left[ -\frac{\lambda_{3i}}{6} \right] + i \left[ \frac{l_{1i}}{12} \omega \right]. \quad (31)$$

Hence, (30) gives the third-order AM/PM intermodulation noise power density spectrum weighted by the AM/PM conversion parameter.

A quantity of interest in engineering problems is the signal-to-noise ratio. Thus, we now characterize the simulated multichannel baseband signal.

#### 2.4 Pre-Emphasized Signal Power Density Spectrum

The basic block diagram arrangement for a typical signal transmission path is shown in Fig. 4. The unpre-emphasized baseband signal is ob-

\* This term causes baseband amplitude distortion instead of intermodulation noise.

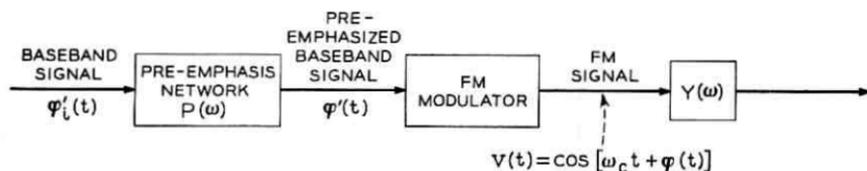


Fig. 4—Typical signal transmission path.

tained from the frequency division multiplex terminals, directly or via a transmission facility, and is pre-emphasized prior to being applied to a FM modulator. The output of the FM modulator is consistent with  $v(t)$  shown in Fig. 1. Assume that the unpre-emphasized baseband signal has a Gaussian distribution and a flat power density spectrum,  $P_o$ , between  $-f_b$  and  $f_b$ , where  $f_b$  is the top baseband frequency. The output power density spectrum from the pre-emphasis network is

$$S_{\varphi'}(\omega) = P_o |P(\omega)|^2, \quad |f| \leq f_b \quad (32)$$

where  $P(\omega)$  is the transfer function of the pre-emphasis network. Letting

$$|P(\omega)|^2 = a_0 + a_2 f^2 + a_4 f^4 + a_6 f^6, \quad |f| \leq f_b, \quad (33)$$

where the  $a$ 's are real constants, we have

$$S_{\varphi'}(\omega) = P_o [a_0 + a_2 f^2 + a_4 f^4 + a_6 f^6], \quad |f| \leq f_b. \quad (34)$$

It can easily be shown that<sup>2</sup>

$$P_o = \frac{(2\pi\sigma)^2}{2f_b \left( a_0 + \frac{a_2 f_b^2}{3} + \frac{a_4 f_b^4}{5} + \frac{a_6 f_b^6}{7} \right)}, \quad (\text{rad/sec})^2/\text{Hz} \quad (35)$$

where  $\sigma$  = rms frequency deviation, in Hz, due to the baseband signal, and  $f_b$  is in Hz. Equation (34) gives the power density spectrum of the pre-emphasized baseband signal in terms of the coefficients of a continuous pre-emphasis characteristic, and in terms of system parameters,  $\sigma$  and  $f_b$ .

### 2.5 Signal-to-Noise Ratio

We are now in a position to express the signal-to-noise ratio for second- and third-order AM/PM intermodulation noise. The expressions given in (22) and (30) are for PM distortions so we convert them to FM distortions by multiplying by  $\omega^2$ . Hence, the signal-to-noise ratios can be

expressed as

$$\left[ 10 \log \frac{S_{\varphi'}(\omega)}{\omega^2 S_{\theta_2}(\omega)} \right]_{\text{2nd order}} = 10 \log \frac{S_{\varphi'}(\omega)}{\left(\frac{\omega}{k}\right)^2 S_{\theta_2}(\omega)} - 20 \log k \quad (36)$$

and

$$\left[ 10 \log \frac{S_{\varphi'}(\omega)}{\omega^2 S_{\theta_3}(\omega)} \right]_{\text{3rd order}} = 10 \log \frac{S_{\varphi'}(\omega)}{\left(\frac{\omega}{k}\right)^2 S_{\theta_3}(\omega)} - 20 \log k, \quad (37)$$

where  $S_{\varphi'}(\omega)$  is given by (34),  $1/k^2 S_{\theta_2}(\omega)$  is given by (22), and  $1/k^2 S_{\theta_3}(\omega)$  is given by (30). A digital computer program has been written which will evaluate (36) and (37) for any values of the transmission deviation coefficients, pre-emphasis coefficients, rms frequency deviation due to the baseband signal, top baseband frequency, and AM/PM conversion factor. The derivations of  $\mathfrak{F}[R_{\varphi'}^2(\tau)]$ ,  $\mathfrak{F}[R_{\varphi'\varphi'}^2(\tau)]$ ,  $\mathfrak{F}[R_{\varphi'}^2(\tau)]$ , and  $\mathfrak{F}[R_{\varphi'}^3(\tau)]$  in a form applicable to a digital computer program are given in Appendix II of Ref. 2.

### III. NOISE PROPERTIES AND CHARACTERISTICS

The previous material provided the mathematical treatment of AM/PM intermodulation noise. In this section we will document the various properties of both AM/PM intermodulation noise and intermodulation noise due to transmission deviations.\* Also, the characteristics of these two noise phenomena will be explored by utilizing a representative system model. Both noise contributors are treated in parallel throughout the section for comparison purposes. The results are presented in three discrete modes: (i) properties which are true in general; (ii) properties which are approximately true; and (iii) characteristics which are derived from a representative system model. The theoretical treatment previously presented was for a transmission medium given by (10). In this section we will confine our analysis to the power series transmission deviations in (10). This is done for two reasons: (i) the properties of the two noise phenomena can be concisely documented for power series transmission deviations; and (ii) the gain and phase ripple properties need more analysis as well as mathematical treatment in order to fully characterize the effects of ripples in the transmission medium.

\*The information for this latter noise contributor was obtained from Ref. 2, which gives it implicitly, as well as from the associated digital computer program.

## 3.1 General Properties

Equation (8) of this paper and (23) of Ref. 2 have been expressed in terms of the transmission deviations and tabulated as shown in Table I. This table is an extension of Table 21-1 of Ref. 4.

TABLE I—AMPLITUDE AND PHASE MODULATION CAUSED BY TRANSMISSION DEVIATIONS

Type of transmission deviation	Resulting amplitude modulation, $a(t)$	Resulting phase modulation, $\varphi_o(t) - \varphi(t)$
Linear gain, $g_1$	$g_1\varphi' - \frac{1}{2}g_1^2\varphi'^2 + \frac{1}{3}g_1^3\varphi'^3$	$-g_2\varphi'' + g_2^2\varphi'^2\varphi''$
Parabolic gain, $g_2$	$g_2\varphi'^2 + \frac{1}{2}g_2^2\varphi''^2$	$-3g_3g_3'\varphi''$
Cubic gain, $g_3$	$-g_3\varphi'' + g_3\varphi'^3$	$g_4\varphi'' + g_4\varphi'^2\varphi''$
Quartic gain, $g_4$	$-4g_4\varphi'' + 3g_4\varphi''^2$	$-2b_2^2\varphi'' + b_2\varphi'^2$
Parabolic phase, $b_2$	$b_2\varphi'' + 2b_2^2\varphi'' + b_2^2\varphi''^2$	$+2b_2^2\varphi''^2\varphi''$
Cubic phase, $b_3$	$3b_3\varphi''$	$-b_3\varphi'' + b_3\varphi'^3$
Quartic phase, $b_4$	$-b_4\varphi'' + 6b_4\varphi''^2\varphi''$	$-4b_4\varphi'' + 3b_4\varphi''^2$
Interaction terms	$-[g_1b_3 + g_2b_2]\varphi'' + [2g_1b_2]\varphi'' + g_1g_2\varphi''$	$-g_1b_2\varphi'' + g_1g_2\varphi''$
	$-g_1g_2\varphi'' + [3g_1b_3 + 4g_2b_2 - 2g_1^2b_2]\varphi''$	$-4(g_1b_3 + g_2b_2)\varphi''$
		$+ (3g_1g_3 - g_1^2g_2)\varphi''$

Input signal =  $\exp\{i[\omega_c t + \varphi(t)]\}$ ; output signal =  $\exp\{a(t)\} \exp\{i[\omega_c t + \varphi_o(t)]\}$ ; transmission medium transfer function =  $Y(\omega + \omega_c) = [1 + g_1\omega + g_2\omega^2 + g_3\omega^3 + g_4\omega^4] \exp\{i[b_2\omega^2 + b_3\omega^3 + b_4\omega^4]\}$ .

The argument of all the amplitude and phase functions is  $t$ .

The order of the noise produced by different transmission deviations (e.g.,  $g_1$ ,  $b_2$ ) are given in Table II for intermodulation noise due to transmission deviations and for AM/PM intermodulation noise. Two rules of thumb can be stated. For intermodulation noise due to transmission deviations the rule is:

*Even-order gain and delay transmission deviations cause odd-order noise.*

TABLE II — ORDER OF NOISE

Transmission deviation	Intermodulation noise	
	Due to transmission deviations	Due to AM/PM conversion
Linear gain ( $g_1$ ):	No noise	*Second and third
Parabolic gain ( $g_2$ ):	Third	Second
Cubic gain ( $g_3$ ):	Second	Third
Quartic gain ( $g_4$ ):	Third	Second
Linear delay ( $b_2$ ):	*Second and third	Second
Parabolic delay ( $b_3$ ):	Third	Second
Cubic delay ( $b_4$ ):	Second	Third

\* Indicates predominant component of the two possible.

*Odd-order gain and delay transmission deviations cause even-order noise.*  
For AM/PM intermodulation noise the rule is, for those transmission deviations that cause significant relative noise (will become apparent later),

*Even-order gain and delay transmission deviations cause even-order noise.*

*Odd-order gain and delay transmission deviations cause odd-order noise.*

The two types of intermodulation noise are related to the magnitude of the transmission deviation coefficient by the relationships shown in Table III. Once a noise response is obtained for a particular system and transmission deviation coefficient value, then the system noise for any other coefficient value typically encountered in transmission systems can be easily predicted.

### 3.2 Approximate Properties

The variation in the top message channel noise, for both noise contributors, with number of channels, assuming the peak frequency deviation remains constant as the number of message channels increase, is shown in Table IV for the different transmission deviations. These approximate relationships yield results with an error of <1 dB for smooth pre-emphasis functions typically used in broadband radio systems.

The assumptions used were that the peak frequency deviation remained constant, and that a typical frequency division multiplex plan was used. The rms frequency deviation, due to the baseband signal,

TABLE III — VARIATION OF RELATIVE NOISE WITH TRANSMISSION DEVIATION COEFFICIENT VALUE

Transmission deviation	Intermodulation noise	
	Due to transmission deviations	Due to AM/PM conversion
Linear gain ( $g_1$ ):	No noise	$*40 \log  g_1'/g_1 , 60 \log  g_1'/g_1 $
Parabolic gain ( $g_2$ ):	$40 \log  g_2'/g_2 $	$\leq 20 \log  g_2'/g_2 $ (approximation error $< \frac{1}{2}$ dB)
Cubic gain ( $g_3$ ):	$20 \log  g_3'/g_3 $	$20 \log  g_3'/g_3 $
Quartic gain ( $g_4$ ):	$20 \log  g_4'/g_4 $	$20 \log  g_4'/g_4 $
Linear delay ( $b_2$ ):	$*20 \log  b_2'/b_2 , 40 \log  b_2'/b_2 $	$40 \log  b_2'/b_2 $
Parabolic delay ( $b_3$ ):	$20 \log  b_3'/b_3 $	$20 \log  b_3'/b_3 $
Cubic delay ( $b_4$ ):	$20 \log  b_4'/b_4 $	$20 \log  b_4'/b_4 $

Where the prime (') notation depicts the terminal value and the unprimed notation indicates the initial value.

\* Indicates predominant component of the two possible.

TABLE IV — VARIATION OF TOP CHANNEL NOISE WITH NUMBER OF MESSAGE CHANNELS

Transmission deviation	Intermodulation noise	
	Due to transmission deviations	Due to AM/PM conversion
Linear gain ( $g_1$ ):	No noise	$\cong 20 \log N'/N$
Parabolic gain ( $g_2$ ):	Relative top channel noise increase $\cong 41 \log N'/N$	$\cong 21 \log N'/N$
Cubic gain ( $g_3$ ):	Relative top channel noise increase $\cong 39 \log N'/N$	$\cong 21 \log N'/N$
Quartic gain ( $g_4$ ):	Relative top channel noise increase $\cong 41 \log N'/N$	$\cong 57 \log N'/N$
Linear delay ( $b_2$ ):	Relative top channel noise increase $\cong 21 \log N'/N$	$\cong 58 \log N'/N$
Parabolic delay ( $b_3$ ):	Relative top channel noise increase $\cong 23 \log N'/N$	$\cong 39 \log N'/N$
Cubic delay ( $b_4$ ):	Relative top channel noise increase $\cong 58 \log N'/N$	$\cong 40 \log N'/N$

Where  $N'$  = increased number of channels;  $N$  = initial number of channels.

was allowed to change, accordingly, as the number of message channels increased.

### 3.3 Noise Characteristics

#### 3.3.1 Representative System Model

As a system model, we will use the following system parameters:

$N$  = number of message channels = 1200

$f_b$  = top baseband frequency = 5.772 MHz

$\Delta F$  = peak frequency deviation = 4 MHz

$\sigma$  = rms frequency deviation due to the multichannel baseband signal = 0.771 MHz.

The pre-emphasis characteristic is shown in Fig. 5 and can be expressed by

$$|P(\omega)|^2 = 0.9989 + 3.5839 \times 10^{-1} f^2 - 5.0245 \times 10^{-3} f^4 + 3.894 \times 10^{-5} f^6,$$

where  $f$  is in MHz.

#### 3.3.2 Noise Response for the Individual Transmission Deviations

It is instructive to show the individual noise responses on a comparative basis. This can be done by letting all gain transmission deviations

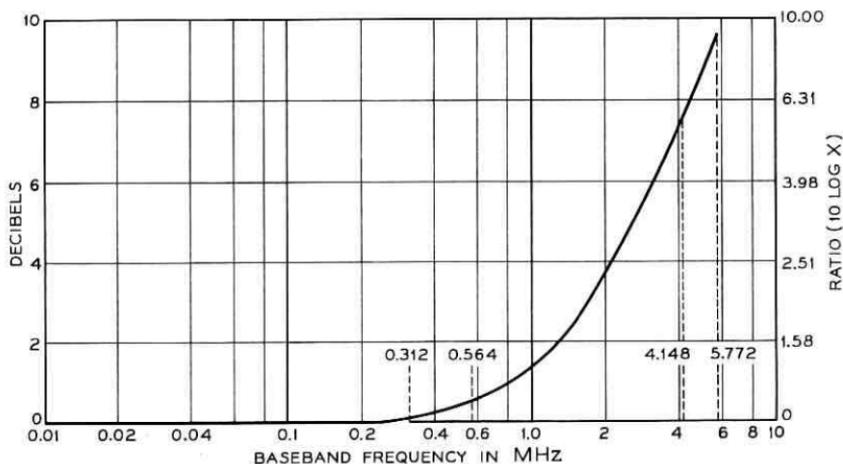


Fig. 5 — Pre-emphasis characteristic.

have 1-dB distortion, relative to the carrier, at 10 MHz away from the carrier. Also, we let all delay transmission deviations have 1 nanosecond (ns) distortion, relative to the carrier, at 10 MHz away from the carrier. This allows us to directly compare the noise contributions of the different gain and phase transmission deviations, respectively, and also allows for some sort of pseudo comparison between a 1-dB gain distortion and a 1-ns delay distortion. The intermodulation noise response, due to transmission deviations, for the different transmission deviations are shown in Fig. 6. Similarly, the AM/PM intermodulation noise responses are shown in Fig. 7. Note that the responses in Fig. 7 are for  $k = 1.0$  radian or a 6.6 degrees/dB AM/PM conversion device. For any other value of  $k$ , say  $k_1$ , we raise or lower the responses according to  $20 \log k_1$ , as indicated by (36) and (37).

It is interesting to note that linear delay is an important contributor to intermodulation noise, due to transmission deviations, but is not a significant AM/PM intermodulation noise contributor. Also, we observe that parabolic gain is a large relative contributor for AM/PM intermodulation noise but is a negligible relative contributor for intermodulation noise due to transmission deviations. As a side point, we point out that parabolic gain is also a significant source of baseband amplitude distortion.

The phase transmission deviation noise responses in Figs. 6 and 7 are of particular interest because the values used in these two figures are realistic even for an equalized system; this is not the case for the

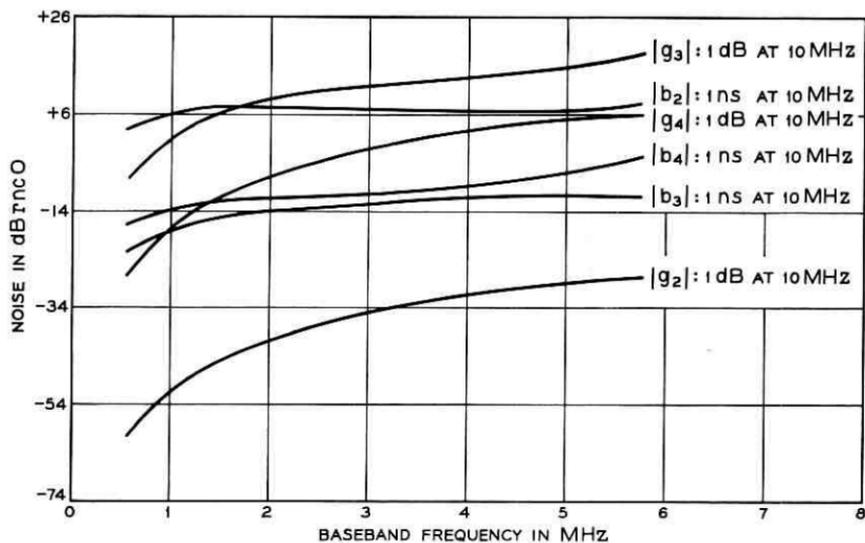


Fig. 6—Intermodulation noise due to transmission deviations.

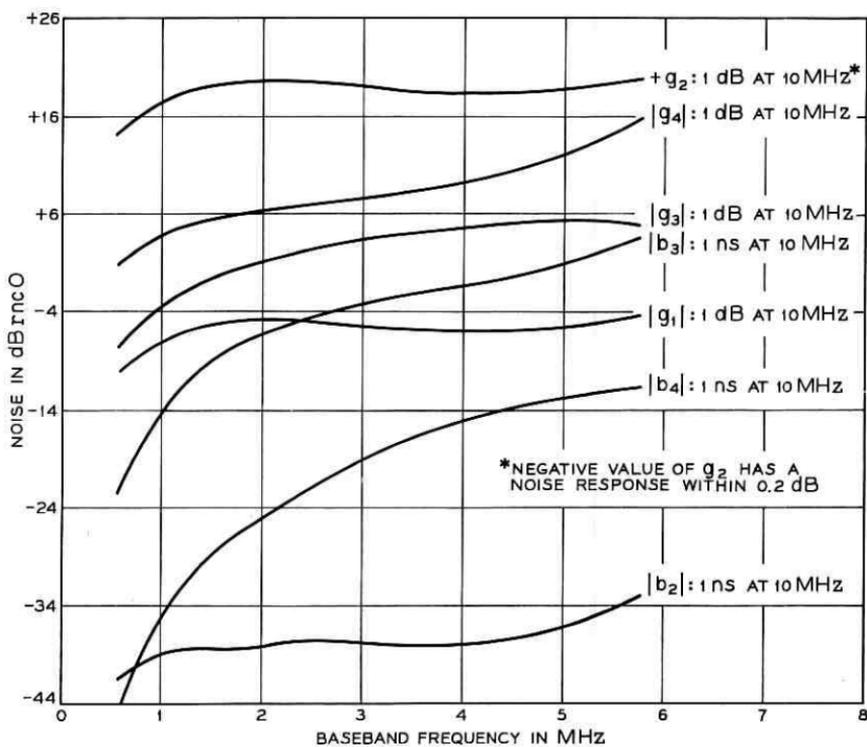


Fig. 7—AM/PM intermodulation noise ( $k = 1.0$  radian).

gain transmission deviation values used so one need not be unduly alarmed at first inspection of the noise responses shown. However, the gain deviation noise responses are of interest in order to determine which types of gain deviations a particular noise source is most sensitive to.

Comparison of Figs. 6 and 7 and Table II show that for all transmission deviations of significance, the order of the noise distortions are different for the two noise phenomena. Coupling this with the results of the Appendix shows that for a given transmission deviation, the two noise responses are uncorrelated.

Of great significance and importance is the parabolic delay AM/PM intermodulation noise response. We see from Fig. 7 that this particular delay deviation is, by far, the largest contributor of AM/PM intermodulation noise compared to the other delay terms. The importance of this finding lies in the fact that TWT amplifiers, when used as output power tubes in broadband radio systems, are separated from transmitter modulators (used to go from IF to RF) by band pass filters which may possess large amounts of parabolic delay. Hence, we have large parabolic delay distortion prior to an important AM/PM conversion device. The noise impairment due to this typical system arrangement will be examined in a later section.

Another point of interest is the noise response for linear gain. We see that linear gain is not a significant AM/PM intermodulation noise contributor. This is useful knowledge because in the past system requirements for linear gain have been set based on speculated AM/PM intermodulation noise impairments, as well as on derivable baseband amplitude distortion due to linear gain and AM/PM conversion.

### 3.3.3 *Effects of Interaction Terms*

Referring back to Table I we note the row marked interaction terms. By the form of the terms involved it is obvious why they are so named. If one were to evaluate (22) and (30) in terms of the transmission deviations explicitly, he would find that over 80 percent of the terms are interaction terms. To examine the effects of these interaction terms we compare the response we would get if we combined the curves shown in Fig. 6, for example, on a power basis with the response we would obtain by using all the transmission deviations at once, i.e., by taking into account the interaction terms. There are a large number of possibilities that could be examined, but to put the problem in perspective the analysis considered only the cases shown in Table V. The results for the two noise phenomena are shown in Figs. 8 and 9. The responses

TABLE V — CASES CONSIDERED IN THE STUDY OF INTERACTION EFFECTS

Case	Condition*
1	Power addition of noise responses due to individual transmission deviations (all $g$ 's and $b$ 's positive)
2	Noise response under the condition: $g_1$ negative and all other $g$ 's and $b$ 's positive
3	Noise response under the condition: $g_2$ negative and all other $g$ 's and $b$ 's positive
4	Noise response under the condition: $g_3$ negative and all other $g$ 's and $b$ 's positive
5	Noise response under the condition: $g_4$ negative and all other $g$ 's and $b$ 's positive
6	Noise response under the condition: $b_2$ negative and all other $g$ 's and $b$ 's positive
7	Noise response under the condition: $b_3$ negative and all other $g$ 's and $b$ 's positive
8	Noise response under the condition: $b_4$ negative and all other $g$ 's and $b$ 's positive
9	Noise response under the condition: all $g$ 's and $b$ 's positive
10	Noise response under the condition: all $g$ 's and $b$ 's negative

$$\text{Where } \begin{array}{l} |g_1| = 1 \text{ dB at 10 MHz} \\ |g_2| = 1 \text{ dB at 10 MHz} \\ |g_3| = 1 \text{ dB at 10 MHz} \\ |g_4| = 1 \text{ dB at 10 MHz} \\ |b_2| = 1 \text{ ns at 10 MHz} \\ |b_3| = 1 \text{ ns at 10 MHz} \\ |b_4| = 1 \text{ ns at 10 MHz} \end{array}$$

\* All the conditions take into account the effects of interaction terms except for case 1.

shown in Fig. 8 are rewarding from a systems analysis standpoint because it indicates that the interaction components for intermodulation noise, due to transmission deviations, do not significantly perturb the noise response obtained by adding up the individual transmission deviation noise responses on a power basis. Hence, for this noise source, a system analyst could set requirements based on power addition of the individual noise responses and be fairly confident that the actual system noise response, due to transmission deviations, will be within a dB of that response.

We see from Fig. 9 that the above desirable property does not hold for AM/PM intermodulation noise. The responses shown in Fig. 9 deviate significant amounts from the power addition response (case 1) by mere shifts of signs, the greatest departures occurring for parabolic and quartic gain distortion which are, in their own right, the largest relative noise contributors as evident from Fig. 7. The relative tendencies indicated in Fig. 9 also occur when typical equalized repeater trans-

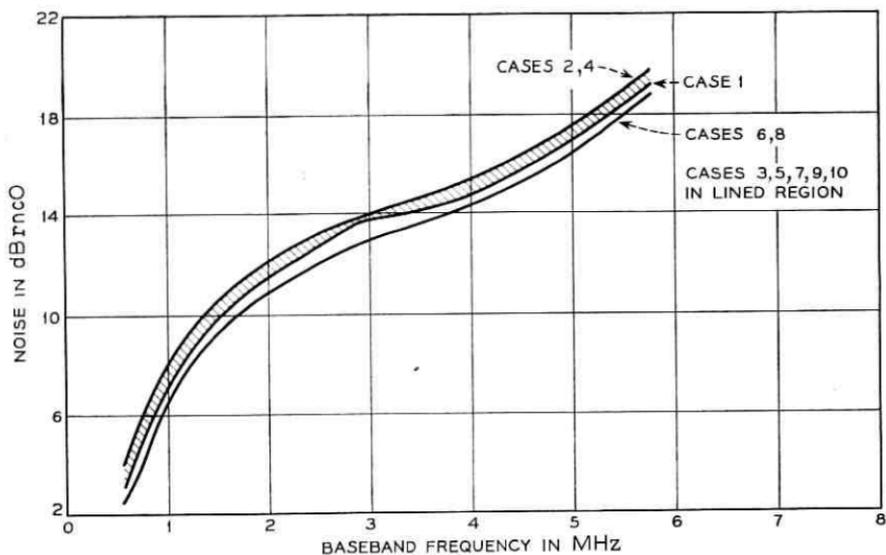


Fig. 8—Intermodulation noise due to transmission deviations—effects of interaction terms (refer to Table V for case listing).

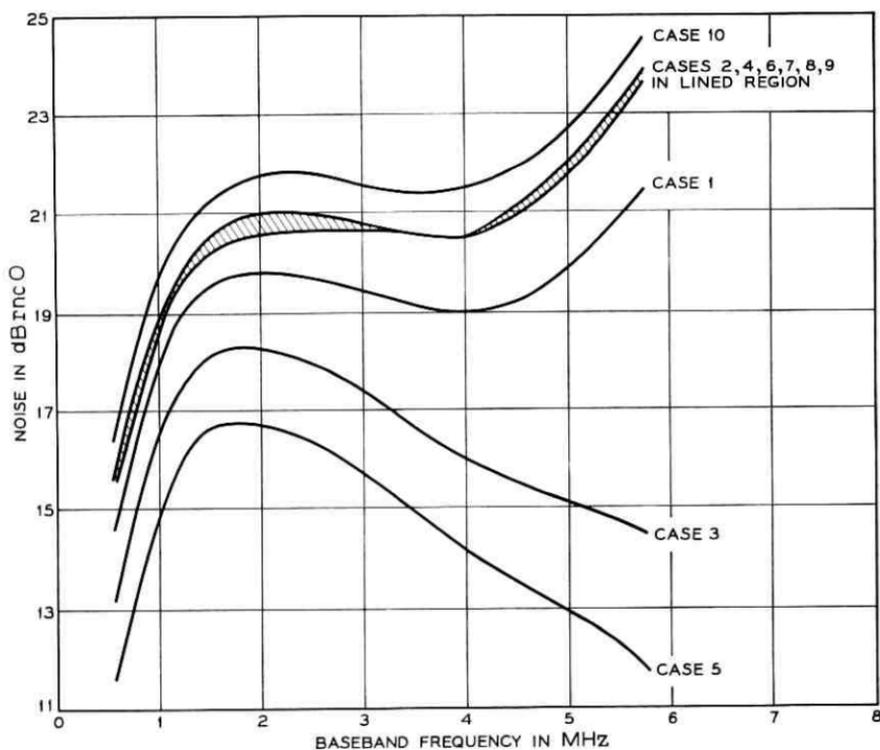


Fig. 9—AM/PM intermodulation noise—effects of interaction terms ( $k = 1.0$  radian) (refer to Table V for case listing).

mission deviation values are used. In fact, least squares approximations on equalized repeaters, RF band pass filters, etc., yield values for parabolic and quartic gain which are either, or both, negative, i.e., a loss with increasing frequency. Hence, even under practical situations one cannot, in general, expect power addition of the individual transmission deviation AM/PM intermodulation noise responses to yield representative AM/PM intermodulation noise performance.

### 3.3.4 Noise Response for a Representative Radio System Repeater

The results up to this point utilized representative system parameters, but normalized values for the transmission deviations were used. Of interest is the predicted noise response for a typical situation, i.e., making use of values typically encountered in practice. We will now use the representative gain and delay responses shown in Fig. 10 for an unequalized and equalized radio repeater. The predicted intermodulation noise responses, due to transmission deviations, are shown in Fig. 11. It is obvious that the equalization has greatly improved the system's noise response.

To examine the AM/PM intermodulation noise we take note of the previously mentioned fact that the TWT has a band pass filter (whose gain and delay responses are given in Fig. 10) preceding it. The AM/PM intermodulation noise due to the band pass filter and the TWT amplifier (assuming 2.5 degrees/dB) is also shown in Fig. 11.

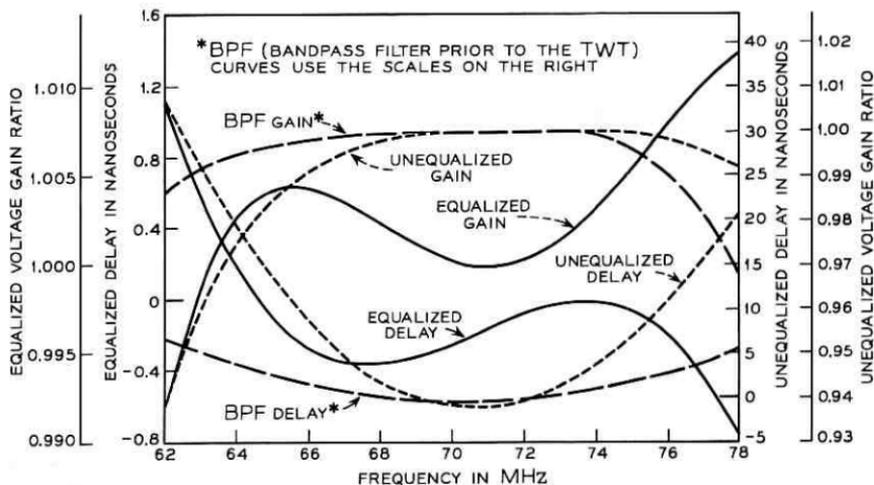


Fig. 10 — Gain and delay characteristics.

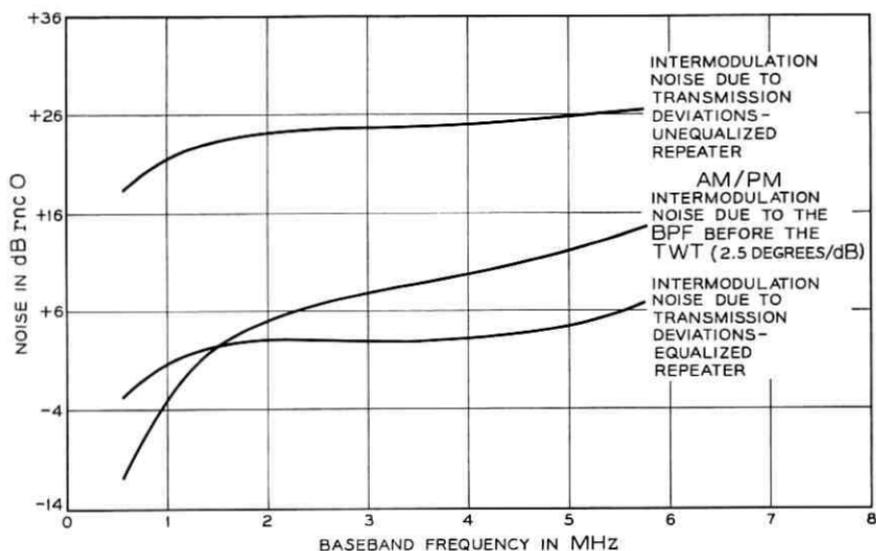


Fig. 11 — Representative radio system noise response.

We see from Fig. 11 that the AM/PM intermodulation noise in the top channel is much larger than the intermodulation noise due to transmission deviations for an equalized repeater. The repeater equalizer is designed to correct for gain and delay shapes obtained from measurements which do not recognize the AM/PM conversion phenomenon. Hence, repeater equalizers based on such measurements, even though effective for reducing intermodulation noise due to transmission deviations, prove ineffective for AM/PM intermodulation noise which occurs as indicated. In other words, the system has an AM/PM intermodulation noise floor which is transparent to external gain and delay measurements.

The transmission deviation of the band pass filter which is the major noise contributor is the parabolic delay term. Hence, to reduce the AM/PM intermodulation noise one must devise some method of correcting for this transmission deviation. Two means of equalizing the band pass filter are: (i) pre-equalization at IF prior to up-converting in the transmitter modulator; and (ii) microwave equalization directly before or after the bandpass filter. The first method may not yield perfect correction because up-converters using varactor diodes possess AM/PM conversion characteristics, in some cases 1 degree/dB. In essence, it would effectively be like trading noise due to 2.5 degrees/dB for noise due to 1.0 degrees/dB or an 8-dB improvement in the ideal

case, i.e., no compression in the up-converter and a perfect inverse band pass filter characteristic. However, an improvement anywhere near this value would greatly reduce the effects of AM/PM conversion.

#### IV. CONCLUSIONS

Two noise contributors in FM systems are: (i) intermodulation noise due to transmission deviations; and (ii) intermodulation noise due to transmission deviations and AM/PM conversion. This latter source of noise is designated "AM/PM intermodulation noise" in this paper. Analysis was carried out in order to predict the second- and third-order AM/PM intermodulation noise for the transmission medium given in (10) and a continuous pre-emphasis function. Flat Gaussian noise was used to simulate the unpre-emphasized baseband signal so the results are consistent with the laboratory system tests using "noise loading". Expressions were derived which specify the signal-to-noise ratio in terms of system parameters, transmission deviations, pre-emphasis characteristics and AM/PM conversion parameter. The latter parameter, assumed to be a real constant, was separated from the body of the calculations so that the resulting noise responses could be easily altered for any value of AM/PM conversion.

The paper presented general noise properties and characteristics for the two noise contributors. This material was presented in parallel, for the two noise contributors, for comparison purposes. The order of the noise component for different transmission deviations was given so that one would know if a given transmission deviation causes second- or third-order noise. The variation of the relative noise with transmission deviation coefficient value was given so that a system analyst can determine the relative detriment to a system response that would result from a change in a given transmission deviation. Another useful result was the variation of top channel noise with number of message channels. This would be of use in the case where one is interested in increasing a system's message channel capacity.

Noise responses were given using a representative radio system model. It was found when all gain transmission deviations had the same distortion and when all delay transmission deviations had the same distortion that: (i) for intermodulation noise due to transmission deviations the cubic and quartic gain terms created the greatest top channel noise due to gain transmission deviations, and that linear delay created the greatest top channel noise due to delay transmission deviations; and (ii) for AM/PM intermodulation noise the parabolic and quartic gain

terms created the greatest top channel noise due to gain transmission deviations, and that parabolic delay created the greatest top channel noise due to delay transmission deviations. The effects of interaction terms were examined. It was found that interaction terms do not significantly perturb the noise response from that of the case of power addition of the individual noise responses for intermodulation noise due to transmission deviations. However, this desirable property did not hold for AM/PM intermodulation noise which says that power addition of the individual noise responses may be in gross error; in other words, interaction terms must be considered when evaluating AM/PM intermodulation noise.

The intermodulation noise due to both noise contributors was predicted for a representative radio system repeater. It was observed that the AM/PM intermodulation noise due to the band pass filter preceding the TWT amplifier created more top channel noise than that due solely to the equalized transmission characteristic, i.e., intermodulation noise due to transmission deviations. Possible correction methods were given.

A point of interest, is that the two noise contributors considered in this paper are correlated so that combining the two spectra together assuming random addition, i.e., power addition, is not sufficient in general. The significance of this correlation is presently being examined and will be reported on in a later paper.

#### V. ACKNOWLEDGMENT

The author wishes to express his gratitude to Miss J. D. Witkowski, of Bell Telephone Laboratories, for programming the AM/PM intermodulation noise equations on a digital computer, and to M. Liou, of Bell Telephone Laboratories, for his motivation and interest in this paper.

#### APPENDIX

##### *Uncorrelated Second- and Third-Order Distortions*

We will show here that the second-order distortion,  $\theta_2(t)$ , and the third-order distortion,  $\theta_3(t)$ , are uncorrelated. Consider

$$\theta_T(t) = \theta_2(t) + \theta_3(t). \quad (38)$$

Now the autocorrelation function of  $\theta_T(t)$  is

$$R_{\theta_T}(\tau) = R_{\theta_2}(\tau) + R_{\theta_3}(\tau) + R_{\theta_2\theta_3}(\tau) + R_{\theta_3\theta_2}(\tau) \quad (39)$$

where, e.g.,  $R_{\theta_2\theta_3}(\tau)$  is the cross-correlation function of  $\theta_2(t)$  and  $\theta_3(t)$ . Taking the Fourier transform of (39) gives

$$S_{\theta_T}(\omega) = S_{\theta_2}(\omega) + S_{\theta_3}(\omega) + S_{\theta_2\theta_3}(\omega) + S_{\theta_3\theta_2}(\omega). \quad (40)$$

From Fig. 2, we have

$$\theta_2(t) = x(t) + y(t) \quad (41)$$

so it follows that

$$R_{\theta_2\theta_3}(\tau) = R_{x\theta_3}(\tau) + R_{y\theta_3}(\tau). \quad (42)$$

Referring to Figs. 2 and 3 we have, using the relationship for the cross-correlation of linearly transformed random functions,<sup>3</sup>

$$S_{\theta_2\theta_3}(\omega) = G_1(-\omega) G_3(\omega) S_{\varphi'2\varphi'3}(\omega) + G_2(-\omega) G_3(\omega) S_{\varphi''2\varphi'3}(\omega). \quad (43)$$

Now we can write

$$S_{\varphi'2\varphi'3}(\omega) = \mathfrak{F} [R_{\varphi'2\varphi'3}(\tau)] = \mathfrak{F} [\text{ave} (\varphi'^2\varphi'^3)] \quad (44)$$

and

$$S_{\varphi''2\varphi'3}(\omega) = \mathfrak{F} [R_{\varphi''2\varphi'3}(\tau)] = \mathfrak{F} [\text{ave} (\varphi''2\varphi'^3)], \quad (45)$$

where  $\mathfrak{F}$  stands for the Fourier Transform.

The phase modulating signal,  $\varphi(t)$  represents the multichannel message load and so for a large number of talkers  $\varphi(t)$  is Gaussian with zero mean.<sup>5</sup> It follows that derivatives of  $\varphi(t)$  are Gaussian with zero mean. It can be shown that<sup>6</sup>

$$\text{ave} [x_1^{r_1} \cdots x_n^{r_n}] = 0, \quad \sum_{i=1}^n r_i \text{ odd} \quad (46)$$

where  $x_1 \cdots x_n$  are Gaussian random variables with zero mean, and  $r_1 \cdots r_n$  are any set of integers. Hence, letting

$$x_1 = \varphi'$$

$$x_2 = \varphi'$$

in (44), and letting

$$x_1 = \varphi''$$

$$x_2 = \varphi'$$

in (45) gives, using (46),

$$S_{\theta_2\theta_3}(\omega) = 0.$$

Similarly,

$$S_{\theta_2\theta_2}(\omega) = 0.$$

Hence,  $\theta_2(t)$  and  $\theta_3(t)$  are uncorrelated so

$$S_{\theta_T}(\omega) = S_{\theta_2}(\omega) + S_{\theta_3}(\omega).$$

#### REFERENCES

1. Laico, J. P., McDowell, H. L., and Moster, C. R., Minimum-Power Traveling-Wave Tube for 6000-mc Radio Relay, *B.S.T.J.*, 35, November, 1956, pp. 1285-1346.
2. Liou, M. L., Noise in an FM System Due to an Imperfect Linear Transducer, *B.S.T.J.*, 45, November, 1966, pp. 1537-1561.
3. Lee, Y. W., *Statistical Theory of Communication*, John Wiley and Sons, Inc., 1960, pp. 348-51.
4. Members of the Technical Staff, Bell Telephone Laboratories, *Transmission Systems for Communications*, Revised Third Edition, 1964.
5. Holbrook, B. D. and Dixon, J. T., Load Rating Theory for Multichannel Amplifiers, *B.S.T.J.*, 18, October, 1939.
6. Laning, J. H., Jr. and Battin, R. H., *Random Processes in Automatic Control*, McGraw-Hill Book Co., N. Y., 1956, pp. 82-85.



# Synthesis of Band-Limited Orthogonal Signals for Multichannel Data Transmission

By ROBERT W. CHANG

(Manuscript received August 4, 1966)

*This paper presents a principle of orthogonal multiplexing for transmitting a number of data messages simultaneously through a linear band-limited transmission medium at a maximum data rate without interchannel and intersymbol interferences. A general method is given for synthesizing an infinite number of classes of band-limited orthogonal time functions in a limited frequency band. Stated in practical terms, the method permits the synthesis of a large class of practical transmitting filter characteristics for an arbitrarily given amplitude characteristic of the transmission medium. Rectangular-shaped ideal filters are not required. The synthesis procedure is convenient. Furthermore, the amplitude and the phase characteristics of the transmitting filters can be synthesized independently. Adaptive correlation reception can be used for data processing, since the received signals remain orthogonal no matter what the phase distortion is in the transmission medium. The system provides the same signal distance protection against channel noises as if the signals of each channel were transmitted through an independent medium and intersymbol interference in each channel were eliminated by reducing data rate.*

## 1. INTRODUCTION

In data transmission, it is common practice to operate a number of AM data channels through a single band-limited transmission medium. The system designer is faced with the problem of maximizing the overall data rate, and minimizing interchannel and intersymbol interferences. In certain applications, the channels may operate on equally spaced center frequencies and transmit at the same data rate, and the signaling intervals of different channels can be synchronized. For these applications, orthogonal multiplexing techniques can be considered. Several

orthogonal-multiplexed systems developed<sup>1,2</sup> in the past use special sets of time-limited orthogonal signals. These signals have widely spread frequency spectra (e.g., a  $\sin x/x$  spectrum). Consequently, when these signals are transmitted through a band-limited transmission medium at a data rate equivalent to that proposed in this paper, certain portions of the signal spectrum will be cut off and interferences will take place. For instance, the interferences because of band-limitation have been computed<sup>3</sup> for a system using time-limited orthogonal sine and cosine functions.

This paper shows that by using a new class of band-limited orthogonal signals, the AM channels can transmit through a linear band-limited transmission medium at a maximum possible data rate without inter-channel and intersymbol interferences. A general method is given for synthesizing an infinite number of classes of band-limited orthogonal time functions in a limited frequency band. This method permits one to synthesize a large class of transmitting filter characteristics for arbitrarily given amplitude and phase characteristics of the transmission medium. The synthesis procedure is convenient. Furthermore, the amplitude and the phase characteristics of the transmitting filters can be synthesized independently, i.e., the amplitude characteristics need not be altered when the phase characteristics are changed, and vice versa. The system can be used to transmit not only binary digits (as in Ref. 1) or  $m$ -ary digits (as in Ref. 2), but also real numbers, such as time samples of analog information sources. As will be shown, the system satisfies the following requirements.

(i) The transmitting filters have gradual cutoff amplitude characteristics. Perpendicular cutoffs and linear phases are not required.

(ii) The data rate per channel is  $2f_s$  bauds,\* where  $f_s$  is the center frequency difference between two adjacent channels. Overall data rate of the system is  $[N/(N + 1)] R_{\max}$ , where  $N$  is the total number of channels and  $R_{\max}$ , which equals two times the overall baseband bandwidth, is the Nyquist rate for which unrealizable rectangular filters with perpendicular cutoffs and linear phases are required. Thus, as  $N$  increases, the overall data rate of the system approaches the theoretical maximum rate  $R_{\max}$ , yet rectangular filtering is not required.

(iii) When transmitting filters are designed for an arbitrary given amplitude characteristic of the transmission medium, the received signals remain orthogonal for all phase characteristics of the transmission medium. Thus, the system (orthogonal transmission plus adaptive

\* The speed in bauds is equal to the number of signal digits transmitted in one second.

correlation reception) eliminates interchannel and intersymbol interferences for all phase characteristics of the transmission medium.

(iv) The distance in signal space between any two sets of received signals is the same as if the signals of each AM channel were transmitted through an independent medium and intersymbol interference in each channel were eliminated by reducing data rate. The same distance protection is therefore provided against channel noises (impulse and Gaussian noise). For instance, for band-limited white Gaussian noise, the receiver receives each of the overlapping signals with the same probability of error as if only that signal were transmitted. The distances in signal space are also independent of the phase characteristics of the transmitting filters and the transmission medium.

(v) When signaling intervals of different channels are not synchronized, at least half of the channels can transmit simultaneously without interchannel or intersymbol interference.

## II. ORTHOGONAL MULTIPLEXING USING BAND-LIMITED SIGNALS

Consider  $N$  AM data channels sharing a single linear transmission medium which has an impulse response  $h(t)$  and a transfer function  $H(f) \exp [J\eta(f)]$  (see Fig. 1).<sup>\*</sup>  $H(f)$  and  $\eta(f)$  will be referred to, respectively, as the amplitude and the phase characteristic of the transmission medium.

Since this analysis treats only transmission media having linear properties, the question of performance on real channels subject to such impairments as nonlinear distortion and carrier frequency offset is not considered here. Such considerations are subjects of studies beyond the scope of the present paper.

In deriving the following results, it is not necessary to assume that the transmitting filters and data processors operate in baseband. However, this assumption will be made since in practice signal shaping and data processing are usually performed in baseband. Carrier modulation and demodulation (included in the transmission medium) can be performed by standard techniques and need not be specified here.

Consider a single channel first (say, the  $i$ th channel). Let  $b_0, b_1, b_2, \dots$ , be a sequence of  $m$ -ary ( $m \geq 2$ ) signal digits or a sequence of real numbers to be transmitted over the  $i$ th channel. As is well known,<sup>4</sup>  $b_0, b_1, b_2, \dots$  can be assumed to be represented by impulses with proportional heights. These impulses are applied to the  $i$ th transmitting filter at the rate of one impulse every  $T$  seconds (data rate per channel

<sup>\*</sup>  $J$  denotes the imaginary number  $\sqrt{-1}$ , while  $j$  is used as an index.

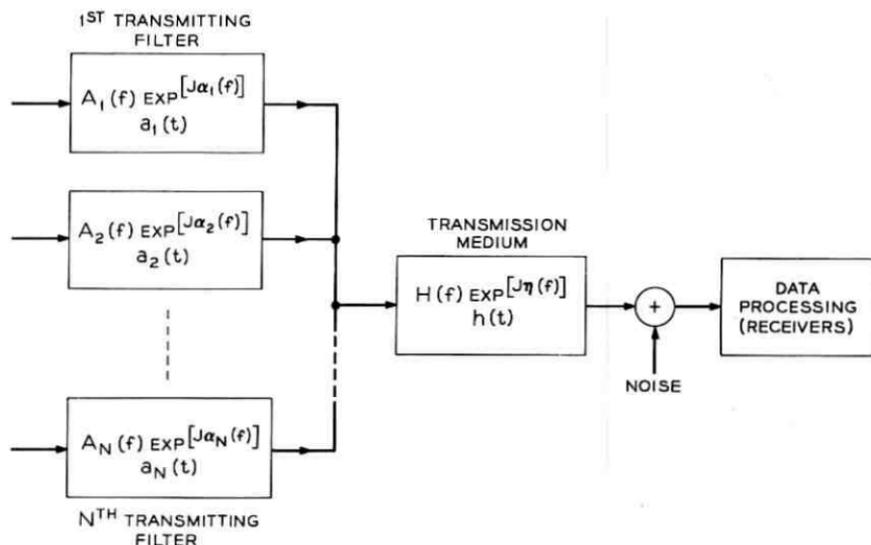


Fig. 1 —  $N$  data channels transmitting over one transmission medium.

equals  $1/T$  bauds). Let  $a_i(t)$  be the impulse response of the  $i$ th transmitting filter, then the  $i$ th transmitting filter transmits a sequence of signals as

$$b_0 \cdot a_i(t) + b_1 \cdot a_i(t - T) + b_2 \cdot a_i(t - 2T) + \dots$$

The received signals at the output of the transmission medium are

$$b_0 \cdot u_i(t) + b_1 \cdot u_i(t - T) + b_2 \cdot u_i(t - 2T) + \dots,$$

where

$$u_i(t) = \int_{-\infty}^{\infty} h(t - \tau) a_i(\tau) d\tau.$$

These received signals overlap in time, but they are orthogonal if

$$\int_{-\infty}^{\infty} u_i(t) u_i(t - kT) dt = 0, \quad k = \pm 1, \pm 2, \dots \quad (1)$$

As is well known, orthogonal signals can be separated at the receiver by correlation techniques;\* hence, intersymbol interference in the  $i$ th channel can be eliminated if (1) is satisfied.

Next consider interchannel interference. Let  $c_0, c_1, c_2, \dots$  be the

\* Correlation reception and its adaptive feature will be briefly discussed in Appendix C.

$m$ -ary signal digits or real numbers transmitted over the  $j$ th channel which has impulse response  $a_j(t)$ . It has been assumed in Section I that the channels transmit at the same data rate and that the signaling intervals of different channels are synchronized, hence the  $j$ th transmitting filter transmits a sequence of signals,

$$c_0 \cdot a_j(t) + c_1 \cdot a_j(t - T) + c_2 \cdot a_j(t - 2T) + \dots$$

The received signals at the output of the transmission medium are

$$c_0 \cdot u_j(t) + c_1 \cdot u_j(t - T) + c_2 \cdot u_j(t - 2T) + \dots$$

These received signals overlap with the received signals of the  $i$ th channel, but they are mutually orthogonal (no interchannel interference) if

$$\int_{-\infty}^{\infty} u_i(t)u_j(t - kT) dt = 0, \quad k = 0, \pm 1, \pm 2, \dots \quad (2)$$

Thus, intersymbol and interchannel interferences can be simultaneously eliminated if the transmitting filters can be designed (i.e., if the transmitted signals can be designed) such that (1) is satisfied for all  $i$  and (2) is satisfied for all  $i$  and  $j$  ( $i \neq j$ ).

Denote  $U_i(f) \exp [J\mu_i(f)]$  as the Fourier transform of  $u_i(t)$ . One can rewrite (1) as

$$\int_{-\infty}^{\infty} U_i^2(f) \exp (-J2\pi f k T) df = 0$$

$$k = \pm 1, \pm 2, \dots$$

$$i = 1, 2, \dots, N,$$
(3)

and rewrite (2) as

$$\int_{-\infty}^{\infty} U_i(f) \exp [J\mu_i(f)] U_j(f) \exp [-J\mu_j(f)]$$

$$\cdot \exp [-J2\pi f k T] df = 0$$

$$k = 0, \pm 1, \pm 2, \dots$$

$$i, j = 1, 2, \dots, N$$

$$i \neq j.$$
(4)

Let  $A_i(f) \exp [J\alpha_i(f)]$  be the Fourier transform of  $a_i(t)$ . The transfer function of the transmission medium is  $H(f) \exp [J\eta(f)]$ . Equation (3) becomes

$$\int_{-\infty}^{\infty} A_i^2(f) H^2(f) \exp(-J2\pi f k T) df = 0$$

$$k = \pm 1, \pm 2, \dots$$

$$i = 1, 2, \dots, N,$$
(5)

or

$$\int_0^{\infty} A_i^2(f) H^2(f) \cos 2\pi f k T df = 0$$

$$k = 1, 2, 3, \dots$$

$$i = 1, 2, \dots, N.$$
(6)

Equation (4) becomes

$$\int_{-\infty}^{\infty} A_i(f) A_j(f) H^2(f) \exp\{J[\alpha_i(f) - \alpha_j(f) - 2\pi f k T]\} df = 0$$

$$k = 0, \pm 1, \pm 2, \dots$$

$$i, j = 1, 2, \dots, N$$

$$i \neq j.$$
(7)

Writing (7) in real and imaginary parts and comparing parts for  $k = 1, 2, \dots$  and  $k = -1, -2, \dots$ , it is seen that (7) holds if and only if

$$\int_0^{\infty} A_i(f) A_j(f) H^2(f) \cos[\alpha_i(f) - \alpha_j(f)] \cos 2\pi f k T df = 0, \quad (8)$$

and

$$\int_0^{\infty} A_i(f) A_j(f) H^2(f) \sin[\alpha_i(f) - \alpha_j(f)] \sin 2\pi f k T df = 0, \quad (9)$$

where

$$k = 0, 1, 2, \dots$$

$$i, j = 1, 2, \dots, N$$

$$i \neq j.$$

It will be recalled that the transmitting filters and the data processors operate in baseband. Let  $f_i, i = 1, 2, \dots, N$ , denote the equally spaced baseband center frequencies of the  $N$  independent channels. One can choose

$$f_i = (h + \frac{1}{2})f_s, \quad (10)$$

where  $h$  is any positive integer (including zero), and  $f_s$  is the difference between center frequencies of two adjacent channels. Thus,

$$f_i = f_1 + (i - 1)f_s = (h + i - \frac{1}{2})f_s. \tag{11}$$

Carrier modulation will translate the baseband signals to a given frequency band for transmission.

Each AM data channel transmits at the data rate  $2f_s$  bauds. Hence,

$$T = \frac{1}{2f_s} \text{ seconds.} \tag{12}$$

For a given amplitude characteristic  $H(f)$  of the transmission medium, band-limited transmitting filters can be designed (i.e., band-limited transmitted signals can be designed) such that (6), (8), (9), and (12) are simultaneously satisfied (no intersymbol and interchannel interference for a data rate of  $2f_s$  bauds per channel). In addition, the five requirements in Section I are also satisfied. A general method of designing these transmitting filters is given in the following theorem.

*Theorem: For a given  $H(f)$ , let  $A_i(f)$ ,  $i = 1, 2, \dots, N$ , be shaped such that*

$$\begin{aligned} A_i^2(f)H^2(f) &= C_i + Q_i(f) > 0, & f_i - f_s < f < f_i + f_s \\ &= 0, & f < f_i - f_s, \quad f > f_i + f_s, \end{aligned} \tag{13}$$

where  $C_i$  is an arbitrary constant and  $Q_i(f)$  is a shaping function having odd symmetries about  $f_i + (f_s/2)$  and  $f_i - (f_s/2)$ , i.e.,

$$Q_i \left[ \left( f_i + \frac{f_s}{2} \right) + f' \right] = -Q_i \left[ \left( f_i + \frac{f_s}{2} \right) - f' \right], \quad 0 < f' < \frac{f_s}{2}, \tag{14}$$

$$Q_i \left[ \left( f_i - \frac{f_s}{2} \right) + f' \right] = -Q_i \left[ \left( f_i - \frac{f_s}{2} \right) - f' \right], \quad 0 < f' < \frac{f_s}{2}. \tag{15}$$

Furthermore, the function  $[C_i + Q_i(f)] \cdot [C_{i+1} + Q_{i+1}(f)]$  is an even function about  $f_i + (f_s/2)$ , i.e.,

$$\begin{aligned} & \left[ C_i + Q_i \left( f_i + \frac{f_s}{2} + f' \right) \right] \left[ C_{i+1} + Q_{i+1} \left( f_i + \frac{f_s}{2} + f' \right) \right] \\ &= \left[ C_i + Q_i \left( f_i + \frac{f_s}{2} - f' \right) \right] \left[ C_{i+1} + Q_{i+1} \left( f_i + \frac{f_s}{2} - f' \right) \right] \tag{16} \\ & \quad 0 < f' < \frac{f_s}{2} \end{aligned}$$

$$i = 1, 2, \dots, N - 1.$$

Let the phase characteristic  $\alpha_i(f)$ ,  $i = 1, 2, \dots, N$ , be shaped such that

$$\alpha_i(f) - \alpha_{i+1}(f) = \pm \frac{\pi}{2} + \gamma_i(f), \quad f_i < f < f_i + f_s \quad (17)$$

$$i = 1, 2, \dots, N - 1,$$

where  $\gamma_i(f)$  is an arbitrary phase function with odd symmetry about  $f_i + (f_s/2)$ .

If  $A_i(f)$  and  $\alpha_i(f)$  are shaped as in (13) through (17) and  $f_1$  is set according to (10), then (6), (8), (9), and (12) are simultaneously satisfied (no intersymbol or interchannel interference for a data rate of  $2f_s$  bauds per channel). Furthermore, the five requirements in Section I are also satisfied.

The proof of this theorem will be broken down into two parts. The first part [showing that (6), (8), (9), and (12) are simultaneously satisfied] will be given in Appendix A. The second part (showing that the five requirements in Section I are satisfied) will be given in Section III following a discussion of the various choices of the shaping functions and transmitting filter characteristics.

### III. TRANSMITTING FILTER CHARACTERISTICS

Consider first the shaping of the amplitude characteristics  $A_i(f)$  of the transmitting filters. Equations (13), (14), and (15) in the theorem can be easily satisfied. Equation (16) can be satisfied in many ways. For instance, a simple, practical way to satisfy (16) is stated in the following corollary.

*Corollary 1: Under the simplifying condition that*

- (i)  $C_i$  should be the same for all  $i$
- (ii)  $Q_i(f)$ ,  $i = 1, 2, \dots, N$ , should be identically shaped, i.e.,

$$Q_{i+1}(f) = Q_i(f - f_s), \quad i = 1, 2, \dots, N - 1, \quad (18)$$

(16) holds when  $Q_i(f)$  is an even function about  $f_i$ , i.e.,

$$Q_i(f_i + f') = Q_i(f_i - f'), \quad 0 < f' < f_s. \quad (19)$$

The proof of this corollary is straightforward and need not be given here. Two examples are given for illustration purpose. The first example is illustrated in Fig. 2 where  $Q_i(f)$  is chosen to be

$$Q_i(f) = \frac{1}{2} \cos \pi \frac{f - f_i}{f_s}, \quad f_i - f_s < f < f_i + f_s \quad i = 1, 2, \dots, N.$$

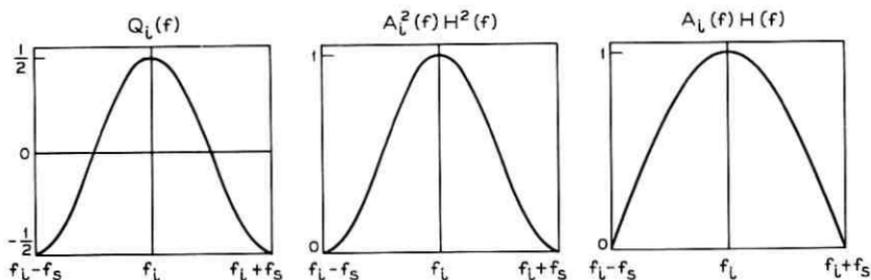


Fig. 2—First example of shaping the amplitude characteristic  $A_i(f)$  of the transmitting filters.

This simple choice satisfies (14), (15), (18), and (19). Let  $C_i$  be  $\frac{1}{2}$  for all  $i$ , then (14), (15), and (16) are all satisfied. From (13)

$$\begin{aligned} A_i^2(f)H^2(f) &= C_i + Q_i(f) \\ &= \frac{1}{2} + \frac{1}{2} \cos \pi \frac{f - f_i}{f_s}, \end{aligned}$$

and

$$A_i(f)H(f) = \cos \pi \frac{f - f_i}{2f_s}, \quad f_i - f_s < f < f_i + f_s$$

$$i = 1, 2, \dots, N.$$

This  $A_i(f)H(f)$  is similar to the amplitude characteristic of a standard duobinary filter (except shift in center frequency). The second example is illustrated in Fig. 3 where  $Q_i(f)$  is chosen such that  $A_i(f)H(f)$  has a shape similar to that of a multiple tuned circuit. It can be seen from these two examples that there is a great deal of freedom in choosing

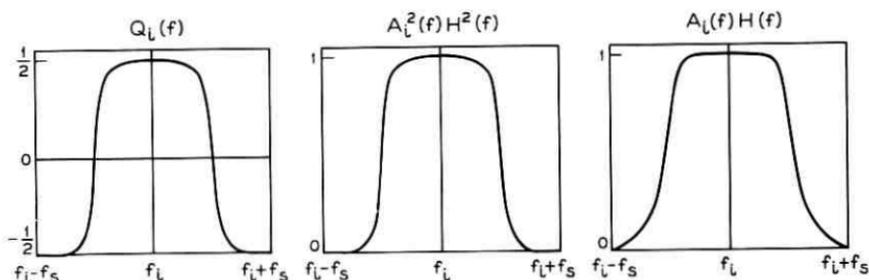


Fig. 3—Second example of shaping the amplitude characteristic  $A_i(f)$  of the transmitting filters.

the shaping function  $Q_i(f)$ . Consequently,  $A_i(f)H(f)$  can be easily shaped into various standard forms.  $A_i(f)$  would have the same shape as  $A_i(f)H(f)$ , if  $H(f)$  is flat in the frequency band  $f_i - f_s$  to  $f_i + f_s$  of the  $i$ th channel. If  $H(f)$  is not flat in this individual band,  $A_i(f)$  can be obtained from  $A_i(f)H(f)/H(f)$ , provided that  $H(f) \neq 0$  for any  $f$  in the band.

It is also noted from the preceding that if  $C_i$  is chosen to be the same for all  $i$  and if  $Q_i(f)$ ,  $i = 1, 2, \dots, N$ , are chosen to be identically shaped (i.e., identical in shape except shifts in center frequencies), then  $A_i(f)H(f)$ ,  $i = 1, 2, \dots, N$ , will also be identically shaped. Consequently,  $A_i(f)$ ,  $i = 1, 2, \dots, N$ , will be identically shaped if  $H(f)$  is flat or is made flat. An advantage of having identically shaped filter characteristics is that each filter can be realized by using an identical shaping filter plus frequency translation.

$H(f)$  can be made flat by using a single compensating network which compensates the variation of  $H(f)$  over the entire band. As an alternative, note that  $A_i(f)$  exists only from  $f_i - f_s$  to  $f_i + f_s$ . Hence, for the  $i$ th receiver, the integration limits of (6), (8), and (9) can be changed to  $f_i - f_s$  and  $f_i + f_s$ . Therefore, the signal at the  $i$ th receiver only has to satisfy the theorem in the limited frequency band  $f_i - f_s$  to  $f_i + f_s$ . This permits one to design the transmitting filters for flat  $H(f)$  and then compensate the variation of  $H(f)$  individually at the receivers, i.e., use an individual network at the  $i$ th receiver to compensate only for the variation of  $H(f)$  in the limited frequency band  $f_i - f_s$  to  $f_i + f_s$ .

Finally, note that if the channels are narrow, each channel will usually be approximately flat. In these cases, one may design the transmitting filters for flat  $H(f)$  without using compensating networks. This design should lead to only small distortion.

Consider next the shaping of the phase characteristics  $\alpha_i(f)$  of the transmitting filters. It is only required in the theorem that (17) be satisfied. However, if it is desired to have identically shaped transmitting filter characteristics, one may consider a simple method such as that in the following corollary.

*Corollary 2: Under the simplifying condition that  $\alpha_i(f)$ ,  $i = 1, 2, \dots, N$ , be identically shaped, i.e.,*

$$\alpha_{i+1}(f) = \alpha_i(f - f_s), \quad i = 1, 2, \dots, N - 1 \quad (20)$$

(17) holds when

$$\alpha_i(f) = h\pi \frac{f - f_i}{2f_s} + \varphi_0 + \sum_m \varphi_m \cos m\pi \frac{f - f_i}{f_s} + \sum_n \psi_n \sin n\pi \frac{f - f_i}{f_s} \quad (21)$$

$$m = 1, 2, 3, 4, 5, \dots$$

$$n = 2, 4, 6, \dots$$

$$f_i - f_s < f < f_i + f_s,$$

where  $h$  is an arbitrary odd integer and the other coefficients ( $\varphi_0$ ;  $\varphi_m$ ,  $m = 1, 2, 3, 4, 5, \dots$ ;  $\psi_n$ ,  $n = 2, 4, 6, \dots$ ) can all be chosen arbitrarily.

This corollary is proven in Appendix B. Note that if the index  $n$  in the corollary were not required to be even,  $\alpha_i(f)$  would be completely arbitrary (a Fourier series with arbitrary coefficients). This shows that there is a great deal of freedom in shaping  $\alpha_i(f)$  even if the additional constraint of identical shaping is introduced (three-fourths of the Fourier coefficients can be chosen arbitrarily). The linear term  $h\pi[(f - f_i)/2f_s]$  is introduced not only to give the term  $\pm\pi/2$  in (17), but also because a linear component is usually present in filter phase characteristics. A simple example is given in Fig. 4 to illustrate (21). For clarity, the arbitrary Fourier coefficients are all set to zero, except  $\psi_2$ , and  $h$  is set to  $-1$ .

As can be seen in the theorem, the requirement on  $\alpha_i(f)$  is independent of the requirements on  $A_i(f)$ . Hence, the amplitude and the phase characteristics of the transmitting filters can be synthesized independently. This gives even more freedom in designing the transmitting filters.

A simple set of  $A_i(f)$  and  $\alpha_i(f)$  is sketched in Fig. 5 for three adjacent channels. This illustrates that the frequency spectrum of each channel is limited and overlaps only with that of the adjacent channel.  $H(f)$  is assumed flat and the transmitting filters are identically shaped. As mentioned previously, these filters can be realized either by different networks or simply by using identical shaping filters plus frequency translations.

Now consider the five requirements in Section I. The first requirement is satisfied since the transmitting filters designed are of standard forms (see the examples in Figs. 2 and 3). Perpendicular cutoffs and linear phase characteristics are not required.

As for the second requirement, it can be seen from Fig. 5(a) that the overall baseband bandwidth of  $N$  channels is  $(N + 1)f_s$ . Since data rate

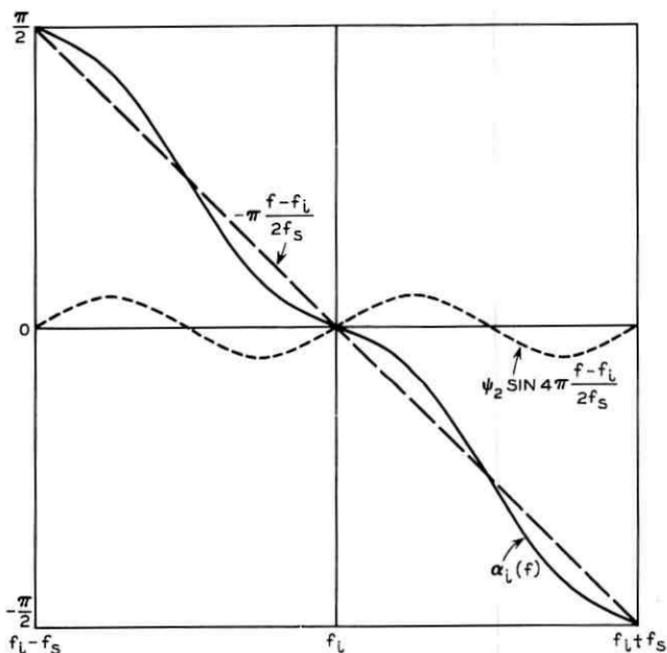


Fig. 4—An example illustrating (21) ( $h = -1$ ,  $\psi_2 \neq 0$ , all other coefficients set to zero for clarity).

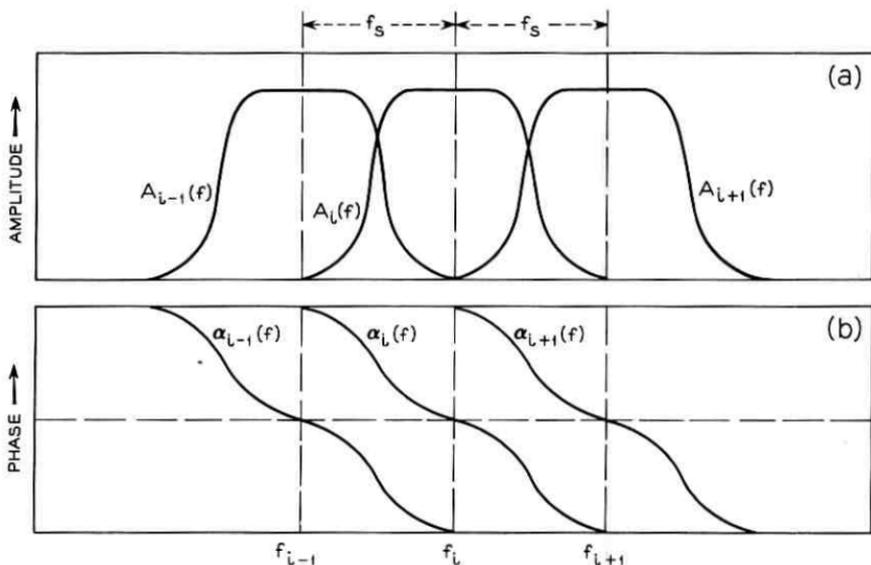


Fig. 5—Example of transmitting filter characteristics for orthogonal multiplexing data transmission.

per channel is  $2f_s$  bauds, the overall data rate of  $N$  channels is  $2f_s N$  bauds. Hence,

$$\begin{aligned} \text{overall data rate} &= \frac{2N}{N+1} \times \text{overall baseband bandwidth} \\ &= \frac{N}{N+1} R_{\max}, \end{aligned}$$

where  $R_{\max}$ , which equals two times overall baseband bandwidth, is the Nyquist rate for which unrealizable filters with perpendicular cutoffs and linear phases are required. Thus, for moderate values of  $N$ , the overall data rate of the orthogonal multiplexing data transmission system is close to the Nyquist rate, yet rectangular filtering is not required. This satisfies requirement (ii).

Now consider the third requirement. As has been shown, the received pulses are orthogonal if (6), (8), and (9) are simultaneously satisfied. Note that the phase characteristic  $\eta(f)$  of the transmission medium does not enter into these equations. Hence, the received signals will remain orthogonal for all  $\eta(f)$ , and adaptive correlation reception (see Appendix C) can be used no matter what the phase distortion is in the transmission medium. Also note that so far as each receiver is concerned, the phase characteristics of the networks in each receiver (including the bandpass filter at the input of each receiver) can be considered as part of  $\eta(f)$ , and hence has no effect on the orthogonality of the received signals.

In the case of the fourth requirement, let

$$b_k^i, \quad k = 0, 1, 2, \dots; \quad i = 1, 2, \dots, N,$$

and

$$c_k^i, \quad k = 0, 1, 2, \dots; \quad i = 1, 2, \dots, N$$

be two arbitrary distinct sets of  $m$ -ary signal digits or real numbers to be transmitted by the  $N$  AM channels. The distance in signal space between the two sets of received signals

$$\sum_i \sum_k b_k^i \cdot u_i(t - kT)$$

and

$$\sum_i \sum_k c_k^i \cdot u_i(t - kT)$$

is

$$d = \left[ \int_{-\infty}^{\infty} \left[ \sum_i \sum_k b_k^i \cdot u_i(t - kT) - \sum_i \sum_k c_k^i \cdot u_i(t - kT) \right]^2 dt \right]^{\frac{1}{2}}.$$

In an ideal case where interchannel and intersymbol interferences are eliminated by transmitting the signals of each channel through an independent medium and slowing down data rate such that the received signals in each channel do not overlap, the distance  $d$  can be written as

$$d_{\text{ideal}} = \left[ \sum_i \sum_k \int_{-\infty}^{\infty} (b_k^i - c_k^i)^2 u_i^2(t - kT) dt \right]^{\frac{1}{2}}$$

In this study, the  $N$  channels transmit over the same transmission medium at the maximum data rate  $T = 1/2f_s$ . If the transmitting filters were not properly designed, the distance  $d$  could be much less than  $d_{\text{ideal}}$  and the system would be much more vulnerable to channel noises (impulse and Gaussian noises). However, if the transmitting filters are designed in accordance with the theorem in Section II, the received signals will be orthogonal and  $d = d_{\text{ideal}}$ . Thus, the distance between any two sets of received signals is preserved and the same distance protection is provided against channel noises. For instance, since  $d = d_{\text{ideal}}$ , it follows from maximum likelihood detection principle that for band-limited white Gaussian noise and  $m$ -ary transmission the receiver will receive each of the overlapping signals with the same probability of error as if only that signal is transmitted.

Note further that  $d_{\text{ideal}}$  can be written as

$$d_{\text{ideal}} = \left[ \sum_i \sum_k (b_k^i - c_k^i)^2 \int_{-\infty}^{\infty} A_i^2(f) H^2(f) df \right]^{\frac{1}{2}}$$

Thus,  $d_{\text{ideal}}$  is independent of the phase characteristics  $\alpha_i(f)$  of the transmitting filters and the phase characteristic  $\eta(f)$  of the transmission medium. Since  $d = d_{\text{ideal}}$ , it follows that  $d$  is also independent of  $\alpha_i(f)$  and  $\eta(f)$  and the same distance protection is provided against channel noises for all  $\alpha_i(f)$  and  $\eta(f)$ .

Finally, consider the fifth requirement. It is assumed in this paper that signaling intervals of different channels are synchronized. However, it is interesting to point out that the frequency spectra of alternate channels (for instance,  $i = 1, 3, 5, \dots$ ) do not overlap (see Fig. 5). Hence, if one uses only the odd- or the even-numbered channels, one can transmit without interchannel and intersymbol interferences and without synchronization among signaling intervals of different channels.\* The overall data rate becomes  $\frac{1}{2} R_{\text{max}}$  for all  $N$ . A very attractive feature is obtained in that the transmitting filters may now have arbitrary phase

\* For instance, signal digits are applied to the  $i$ th transmitting filter at  $0, T, 2T, 3T, \dots$ , while signal digits are applied to the  $(i+2)$ th transmitting filter at  $\tau, T + \tau, 2T + \tau, 3T + \tau, \dots$ , where  $\tau$  is an unknown constant.

characteristics  $\alpha_i(f)$ . [This is because  $\alpha_i(f)$  is not involved in (6) and intersymbol interference is eliminated for all  $\alpha_i(f)$ .] Thus, only the amplitude characteristics of the transmitting filters need to be designed as in the theorem and the transmitting filters can be implemented very easily.

Another case of interest is where part of the channels are synchronized. As a simple example, assume that there are five channels and that channel 1 is synchronized with channel 2; channel 4 is synchronized with channel 5; while channel 3 cannot be synchronized with other channels. If the amplitude characteristics of the five channels plus the phase characteristics of channels 1, 2, 4, and 5 are designed as in the theorem, one can transmit simultaneously through channels, 1, 2, 4, and 5 or simultaneously through channels 1, 3, and 5 without interchannel or intersymbol interferences. The overall data rate is then between  $\frac{1}{2}R_{\max}$  and  $(N/N + 1)R_{\max}$ .

#### IV. CONCLUSION

This paper presents a principle of orthogonal multiplexing for transmitting  $N(N \geq 2)$  AM data channels simultaneously through a linear band-limited transmission medium. The channels operate on equally spaced center frequencies and transmit at the same data rate with signaling intervals synchronized. Each channel can transmit binary digits,  $m$ -ary digits, or real numbers. By limiting and stacking the frequency spectrums of the channels in a proper manner, an overall data rate of

$$\frac{2N}{N + 1} \times \text{overall baseband bandwidth} \quad \text{bauds}$$

is obtained which approaches the Nyquist rate when  $N$  is large. Interchannel and intersymbol interferences are eliminated by a new method of synthesizing the transmitting filter characteristics (i.e., designing band-limited orthogonal signals). The method permits one to synthesize a large class of transmitting filter characteristics in a very convenient manner. The amplitude and the phase characteristics can be synthesized independently. The transmitting filter characteristics obtained are practical in that

- (i) The amplitude characteristics may have gradual rolloffs, and the phase characteristics need not be linear.
- (ii) The transmitting filters may be identically shaped and can be realized simply by identical shaping filters plus frequency translations.

It is noted that the principle presented in this paper uses band-limited orthogonal signals as opposed to other orthogonal multiplexing schemes using nonband-limited orthogonal signals. The chief advantage of using band-limited signals is that (as mentioned in Section I) these signals can be transmitted through a band-limited transmission medium at a maximum data rate without interchannel and intersymbol interferences. Other advantages of using band-limited signals over methods using nonband-limited signals are

- (i) Permitting the use of a narrowband bandpass filter at the input of each receiver (see Appendix C) to reject noises and signals outside the band of interest. This is particularly important in suppressing impulse noises and in preventing overloading the front ends of the receivers.
- (ii) Permitting unsynchronized operations at data rates between  $\frac{1}{2}R_{\max}$  and  $(N/N + 1)R_{\max}$ .

It has been shown that the received signals remain orthogonal for all phase characteristics of the transmission medium; hence, adaptive correlation reception can be used to separate the received signals no matter what the phase distortion is in the transmission medium. These correlators adapt not only to the phase distortions in the system (including transmission medium, bandpass receiving filters, etc.), but also (see Appendix C) to the phase difference between modulation and demodulation carriers (easing synchronization requirements).

#### V. ACKNOWLEDGMENTS

I wish to thank A. B. Brown, Jr., R. A. Gibby, and J. W. Smith for stimulating discussions and helpful suggestions.

#### APPENDIX A

In this appendix, it will be proven that if the transmitting filters  $A_i(f) \exp [j\alpha_i(f)]$ ,  $i = 1, 2, \dots, N$ , are shaped as in the theorem in Section II and  $f_1$  is set according to (10), then equations (6), (8), (9), and (12) are simultaneously satisfied.

First consider (6). From (13)

$$\int_0^{\infty} A_i^2(f) H^2(f) \cos 2\pi f k T df = \int_{f_i - f_s}^{f_i + f_s} [C_i + Q_i(f)] \cos 2\pi f k T df. \quad (22)$$

Since  $T = 1/2f_s$ , one has

$$\begin{aligned}
 \int_{f_i-f_s}^{f_i+f_s} C_i \cos 2\pi f k T \, df &= \frac{C_i}{2\pi k T} [\sin 2\pi(f_i + f_s)kT \\
 &\quad - \sin 2\pi(f_i - f_s)kT] \\
 &= \frac{C_i}{\pi k T} \sin \pi k \cos 2\pi f_i k T \qquad (23) \\
 &= 0, \quad k = 1, 2, 3, \dots \\
 &\quad i = 1, 2, 3, \dots, N.
 \end{aligned}$$

Since  $f_i = (h + i - \frac{1}{2})f_s$ , one has

$$\begin{aligned}
 2\pi \left( f_i - \frac{f_s}{2} \right) k T &= 2\pi(h + i - 1)f_s k \frac{1}{2f_s} \\
 &= (h + i - 1)k\pi.
 \end{aligned} \qquad (24)$$

Hence,  $\cos 2\pi f k T$  is an even function about  $f_i - (f_s/2)$ . This, together with the fact that  $Q_i(f)$  is an odd function about  $f_i - (f_s/2)$  [see (15)], gives

$$\begin{aligned}
 \int_{f_i-f_s}^{f_i} Q_i(f) \cos 2\pi f k T \, df &= 0, \quad k = 1, 2, 3, \dots \\
 &\quad i = 1, 2, \dots, N.
 \end{aligned} \qquad (25)$$

Similarly, one can show

$$\begin{aligned}
 \int_{f_i}^{f_i+f_s} Q_i(f) \cos 2\pi f k T \, df &= 0, \quad k = 1, 2, 3, \dots \\
 &\quad i = 1, 2, \dots, N.
 \end{aligned} \qquad (26)$$

Substituting (23), (25), and (26) into (22) gives

$$\begin{aligned}
 \int_0^\infty A_i^2(f) H^2(f) \cos 2\pi f k T \, df &= 0, \quad k = 1, 2, 3, \dots \\
 &\quad i = 1, 2, 3, \dots, N.
 \end{aligned}$$

Thus, (6) is satisfied and intersymbol interference is eliminated.

Next consider interchannel interference and (8) and (9). From (13),

$$A_i(f)H(f) = 0, \quad f < f_i - f_s, f > f_i + f_s,$$

so

$$A_i(f)A_j(f)H^2(f) = 0 \quad \text{for } j = i \pm 2, i \pm 3, i \pm 4, \dots,$$

or

$$\int_0^{\infty} A_i(f)A_j(f)H^2(f) \cos [\alpha_i(f) - \alpha_j(f)] \cos 2\pi f k T df = 0$$

$$\int_0^{\infty} A_i(f)A_j(f)H^2(f) \sin [\alpha_i(f) - \alpha_j(f)] \sin 2\pi f k T df = 0 \quad (27)$$

$$k = 0, 1, 2, \dots$$

$$i = 1, 2, 3, \dots, N$$

$$j = i \pm 2, i \pm 3, i \pm 4, \dots$$

Equation (27) shows that (8) and (9) are satisfied for  $j = i \pm 2, i \pm 3, i \pm 4, \dots$ . It remains to show that (8) and (9) hold for  $j = i \pm 1$ . Consider  $j = i + 1$ . It is seen from (13) that

$$A_i(f)A_{i+1}(f)H^2(f) = [C_i + Q_i(f)]^{\frac{1}{2}}[C_{i+1} + Q_{i+1}(f)]^{\frac{1}{2}}, \quad f_i < f < f_i + f_s$$

$$= 0, \quad f < f_i, f > f_i + f_s. \quad (28)$$

One can write from (17) and (28)

$$\int_0^{\infty} A_i(f)A_{i+1}(f)H^2(f) \cos [\alpha_i(f) - \alpha_{i+1}(f)] \cos 2\pi f k T df$$

$$= \int_{f_i}^{f_i+f_s} [C_i + Q_i(f)]^{\frac{1}{2}}[C_{i+1} + Q_{i+1}(f)]^{\frac{1}{2}} \cos \left[ \pm \frac{\pi}{2} + \gamma_i(f) \right]$$

$$\cdot \cos 2\pi f k T df \quad (29)$$

$$k = 0, 1, 2, \dots$$

$$i = 1, 2, \dots, N.$$

It is required in the theorem that

$$[C_i + Q_i(f)]^{\frac{1}{2}}[C_{i+1} + Q_{i+1}(f)]^{\frac{1}{2}}$$

be an even function about  $f_i + (f_s/2)$ . Furthermore,  $\cos [\pm(\pi/2) + \gamma_i(f)]$  and  $\cos 2\pi f k T$  are, respectively, odd and even functions about  $f_i + (f_s/2)$ . Hence, from (29)

$$\int_0^{\infty} A_i(f)A_{i+1}(f)H^2(f) \cos [\alpha_i(f) - \alpha_{i+1}(f)] \cos 2\pi f k T df = 0$$

$$k = 0, 1, 2, \dots \quad (30)$$

$$i = 1, 2, \dots, N.$$

Equation (30) shows that (8) is satisfied for  $j = i + 1$ . In a similar manner, one can show that (8) holds for  $j = i - 1$  and that (9) holds for  $j = i \pm 1$ . These, together with (27), prove that (8) and (9) hold for all  $k, i,$  and  $j$ .

APPENDIX B

*Proof of Corollary 2*

From (20) and (21)

$$\begin{aligned} \alpha_{i+1}(f) &= \alpha_i(f - f_s) \\ &= h\pi \frac{f - f_s - f_i}{2f_s} + \varphi_0 \\ &\quad + \sum_m \varphi_m \cos m\pi \frac{f - f_s - f_i}{f_s} \\ &\quad + \sum_n \psi_n \sin n\pi \frac{f - f_s - f_i}{f_s} \end{aligned} \tag{31}$$

$$m = 1, 2, 3, 4, 5, \dots$$

$$n = 2, 4, 6, \dots$$

$$f_i < f < f_i + 2f_s.$$

For  $f_i < f < f_i + f_s$ , one has from (21) and (31)

$$\begin{aligned} \alpha_i(f) - \alpha_{i+1}(f) &= \frac{h\pi}{2f_s} [f - f_i - (f - f_s - f_i)] \\ &\quad + \sum_m \varphi_m \left[ \cos m\pi \frac{f - f_i}{f_s} \right. \\ &\quad \left. - \cos m\pi \frac{f - f_s - f_i}{f_s} \right] \\ &\quad + \sum_n \psi_n \left[ \sin n\pi \frac{f - f_i}{f_s} \right. \\ &\quad \left. - \sin n\pi \frac{f - f_s - f_i}{f_s} \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{h\pi}{2} + \sum_m \varphi_m \left[ -2 \sin \frac{1}{2} \left( m\pi \frac{2f - 2f_i - f_s}{f_s} \right) \sin \frac{m\pi}{2} \right] \\
&\quad + \sum_n \psi_n \left[ 2 \sin \frac{n\pi}{2} \cos \frac{1}{2} \left( n\pi \frac{2f - 2f_i - f_s}{f_s} \right) \right] \\
&= \frac{h\pi}{2} - 2 \sum_l \varphi_l \sin \frac{l\pi}{2} \sin \frac{l\pi(2f - 2f_i - f_s)}{2f_s} \\
&\qquad\qquad\qquad l = 1, 3, 5, \dots \\
&\qquad\qquad\qquad h = \pm 1, \pm 3, \dots
\end{aligned} \tag{32}$$

Since  $\sin [l\pi(2f - 2f_i - f_s)/2f_s]$  is an odd function about  $f = f_i + (f_s/2)$ , (32) is equivalent to (17) and corollary 2 is proven.

#### APPENDIX C

This appendix briefly describes a possible receiver structure for receiving the multichannel orthogonal signals.

The receiver of a single channel (say, the fifth channel) is shown in Fig. 6(a). When viewed at point B toward the transmitter, the channels have amplitude characteristics as shown in Fig. 6(b). The bandpass filter at the input of the fifth receiver has a passband from  $f_5 - f_s$  to  $f_5 + f_s$  [Fig. 6(c)]. This filter serves the important purpose of rejecting noises and signals outside the band of interest. Sharp impulse noises with broad frequency spectra are greatly attenuated by this filter. Signals in other channels are rejected to prevent overloading and cross modulation.

The product device translates the frequency spectra further toward the origin so that the signal can be represented by a minimum number of accurate time samples and the adaptive correlator can operate in digital fashion. The transmitter can transmit a reference frequency  $f_s$  or a known multiple of  $f_s$  to the receivers for deriving the signals  $\cos [2\pi(i - 1)f_s t + \theta_i]$  for the product devices. It is important to note that the transmitter can lock this frequency  $f_s$  to the data rate  $2f_s$  so that the arbitrary phase angle  $\theta_i$  is time invariant and can be taken into account by adaptive correlation. Furthermore, the receiver can also derive the sampling rate  $2f_s$  from this reference frequency.

When observed at point D, the channels have amplitude characteristics as shown in Fig. 6(d). Note that the fifth channel now has a center frequency at  $1.5f_s$  [satisfying (11)] and an undistorted amplitude characteristic; hence, the signals in channel 5 remain orthogonal. The overlapping frequency spectra between channel 5 and channels 4 and 6 remain undistorted, and the phase differences  $\alpha_4(f) - \alpha_5(f)$  and  $\alpha_5(f) - \alpha_6(f)$  are unchanged; therefore, the signals in channels 4 and 6 remain

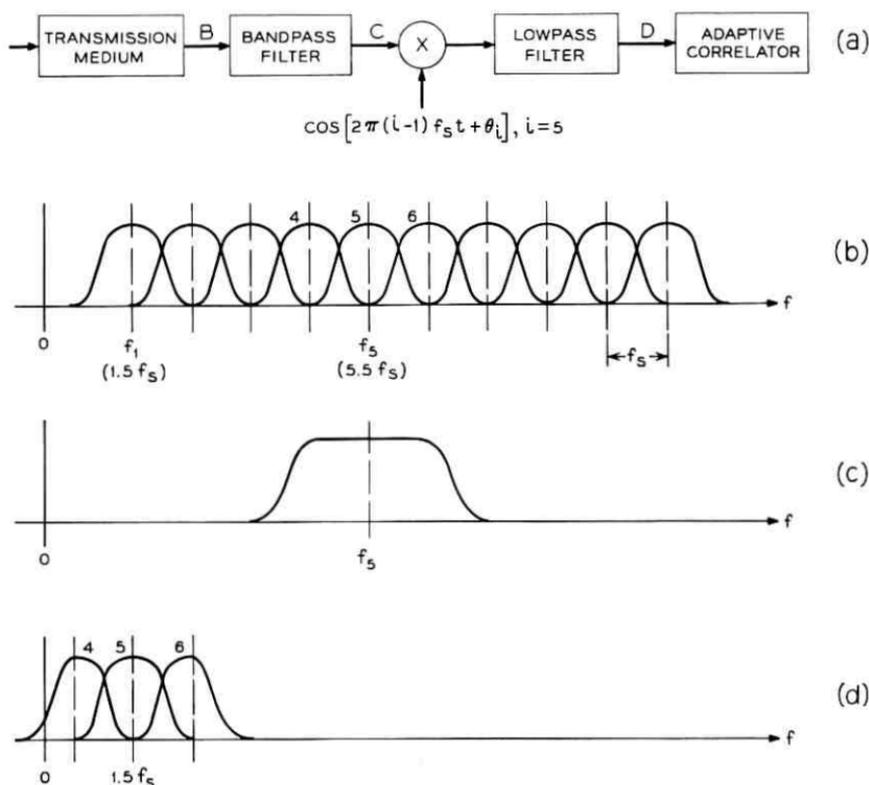


Fig. 6 — Reception of the signals in channel 5.

orthogonal to those in channel 5. Other channels produce no interference since their spectra do not overlap with that of channel 5.

Let  $b_0 u(t), b_1 u(t - T), b_2 u(t - 2T), \dots$  be the signals in channel 5 at point D, where  $b_0, b_1, b_2, \dots$  are the information digits. These signals can be represented by vectors of time samples as

$$b_0 \underline{u}_0, b_1 \underline{u}_1, b_2 \underline{u}_2, \dots$$

Since  $u(t), u(t - T), \dots$  differ only in time origin, it is only necessary to learn  $\underline{u}_0$  for correlation purposes. The received signal at point D can be written as

$$\sum_n b_n \underline{u}_n + \underline{v},$$

where  $\underline{v}$  represents the sum of the signals in other channels. From discussions in the preceding paragraph

$$\begin{aligned} \underline{u}_k' \underline{u}_j &= \lambda & k &= j \\ &= 0 & k &\neq j \\ \underline{u}_k' \underline{v} &= 0. \end{aligned}$$

Thus, the adaptive correlator can learn the vector  $\underline{u}_0$  prior to data transmission and then correlate the received signal with  $\underline{u}_k$ ,  $k = 0, 1, 2, \dots$  to obtain the information digits  $b_k$ ,  $k = 0, 1, 2, \dots$ .

In order to describe the operation more clearly we assume that the signal at point D is fed to a delay line tapped at  $T/3$ -second intervals (signal at D is band-limited between 0 and  $3f_s$ ). Assuming that  $u(t)$  is essentially time-limited to  $mT$  seconds for all possible phase characteristics of the transmission system, then  $3m$  taps are sufficient. The  $i$ th tap is connected to a gain control  $G_i$ . In the training period prior to data transmission, the  $i$ th tap is also connected to a sampler  $s_i$ . In the training period, the transmitter transmits a series of identical test pulses at  $t = 0, lT, 2lT, \dots$ . The integer  $l$  is chosen large enough such that the received test pulses  $u(t), u(t - lT), u(t - 2lT), \dots$  do not overlap. The sampler  $s_i$  samples at  $t = \tau, lT + \tau, 2lT + \tau, \dots$ . The only requirement on  $\tau$  is that  $u(t)$  should be approximately centered on the tapped delay line at  $t = \tau$ . The output of  $s_i$  (without noise) is a series of samples each representing the  $i$ th time sample  $u_i$  of  $u(t)$ . Since noise is always present, these samples are passed through a network (probably a simple RC circuit) such that the output  $\hat{u}_i$  of this network is an estimate of  $u_i$ .  $\hat{u}_i$  is in the form of a voltage or current and hence can be used to set the gain control  $G_i$  of the  $i$ th tap. Thus, at the end of the training period, the gain controls of the successive taps are set according to the magnitudes of the successive time samples of  $u(t)$ .

During data transmission, the transmitter transmits the information digits  $b_0, b_1, b_2, \dots$  sequentially at  $t = 0, T, 2T, \dots$ . A sampler at the receiver samples the sum of the outputs of all the tap gain controls at  $t = \tau, T + \tau, 2T + \tau, \dots$  to recover  $b_0, b_1, b_2, \dots$ . The time delay  $\tau$  remains the same as in the training period. The data transmission operates in real time.

#### REFERENCES

1. Mosier, R. R., A Data Transmission System Using Pulse Phase Modulation, IRE Convention Record of First National Convention on Military Electronics, June 17-19, 1957, Washington, D. C.
2. *Dynamic Error-Free Transmission*, General Dynamics/Electronics—Rochester, N. Y.
3. Harmuth, H. F., On the Transmission of Information by Orthogonal Time Functions, AIEE Trans., pt. I (Communication and Electronics), July, 1960, pp. 248-255.
4. Bennett, W. R. and Davey, J. R., *Data Transmission*, McGraw-Hill Book Company, Inc., New York, 1965, p. 65.

# Avalanche Region of IMPATT Diodes

By H. K. GUMMEL and D. L. SCHARFETTER

(Manuscript received August 10, 1966)

*The avalanche region of an IMPATT (IMPact ionization Avalanche Transit Time) diode under small signal conditions is characterized by the fraction of the total alternating current that is carried by holes and electrons in their respective drift spaces and by a residual impedance. The current fractions are roughly in phase with the total current below, and nearly 180° out of phase above a resonance frequency that is proportional to the square root of direct current density.*

*This paper extends the calculations of Gilden and Hines for the current fraction to include phase shifts in the avalanche region so that extended avalanche regions can be considered. Realistic values ( $\alpha \neq \beta$  in Si) for the ionization coefficients are used. Results of detailed numerical calculations for the current fractions as a function of frequency and direct current density are presented.*

*For typical frequencies and current densities, the residual impedance is negligible and hole and electron current fractions are equal. The avalanche region at a given frequency and current density is then characterized by one complex number and the admittance of a diode containing the avalanche region and adjacent drift regions is easily calculated. Plots showing the admittance as a function of frequency and current density for typical structures are given.*

*It is found that an optimal exponential growth rate of oscillations is obtained when the current density is such that the resonance frequency is about equal to one half the reciprocal transit time through the longest drift region. If the assumption is made that conditions giving the largest small-signal exponential growth rate give the best large-signal performance, an optimum Read-diode design is obtained for which the avalanche region width is a substantial fraction ( $\approx \frac{1}{3}$ ) of the drift region width.*

## I. INTRODUCTION

This paper considers the avalanche region of IMPATT\* diodes,<sup>1,2,3</sup> especially of Read<sup>1</sup>-type diodes in which the avalanche region is localized.

\* IMPact ionization Avalanche Transit Time.

Recent theoretical<sup>4</sup> and experimental<sup>5</sup> results for the impedance of Read diodes show considerable structure in the current and frequency dependence. This paper attempts to enhance the understanding of the small-signal negative resistance of IMPATT diodes by isolating the role of the avalanche region. Using the calculated parameters by which the avalanche region is characterized, the admittance for typical IMPATT diodes is calculated and exponential growth rate for oscillations is studied. An application of the results to the design of Read diode oscillators is made.

At a given angular frequency  $\omega$  and ac terminal current density  $i_{tot}$  through the diode, the avalanche region can be characterized by three complex numbers:  $Z_r$ ,  $F_h$ ,  $F_e$  defined below. In Fig. 1, to the right of the avalanche region the small-signal particle current consists of a plane wave of holes. If we assume a constant drift velocity, then the magnitude is constant and the phase changes linearly with distance. Let  $A_h$  be the complex amplitude, extrapolated to a phase reference plane at  $x_0$  in the avalanche region. Then we define the *hole current fraction*  $F_h$

$$F_h = A_h/i_{tot}. \quad (1)$$

Similarly, we define the amplitude of the electron current density, extrapolated to  $x_0$ , to be  $A_e$  and we define the *electron current fraction*  $F_e$

$$F_e = A_e/i_{tot}. \quad (2)$$

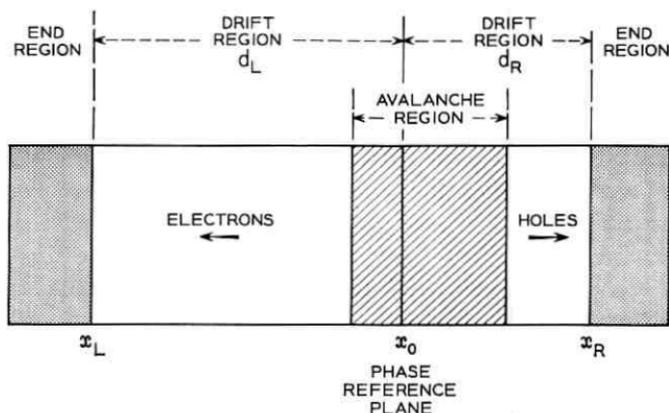


Fig. 1—Schematic of IMPATT diode.

In the limit as the frequency goes to zero, all the current through the device is particle current and  $F_h$  and  $F_e$  are of equal magnitude. The phase reference plane is chosen so that at low frequencies  $F_h$  and  $F_e$  are of equal phase also. The phase reference plane is approximately located at the point where the hole and electron currents are equal.

Assume for a moment that the avalanche region were infinitely thin. From the fact that the total current equals the particle current plus displacement current we find that the ac field per total ac current, to the left of  $x_0$  is

$$\frac{E}{i_{\text{tot}}} = \frac{1 - i_e/i_{\text{tot}}}{i\omega\epsilon} = \frac{1 - F_e \exp [i\omega(x - x_0)/v_e]}{i\omega\epsilon}, \quad (3)$$

where  $i_e$  is the electron current density and  $v_e$  the electron velocity. The impedance,  $Z_{L0}$  between  $x_L$  and  $x_0$  (see Fig. 1) is obtained by integration of (3).

$$Z_{L0} = \frac{1}{i\omega C_L} \left[ 1 - F_e \frac{1 - \exp(-i\omega\tau_L)}{i\omega\tau_L} \right], \quad (4)$$

where  $\tau_L$  is the transit time between  $x_0$  and  $x_L$

$$\tau_L = \frac{x_0 - x_L}{v_e} \quad (5)$$

and  $C_L$  is the geometric capacitance per unit area,

$$C_L = \frac{\epsilon}{x_0 - x_L}. \quad (6)$$

A corresponding expression obtains for  $Z_{0R}$ , the impedance between  $x_0$  and  $x_R$ .

In deriving (4) it was assumed that the electron current is of the form

$$i_e = F_e \exp [i\omega(x - x_0)/v_e]. \quad (7)$$

This is true only outside of the avalanche region. The difference between the actual and asymptotic hole and electron current densities gives rise to an additional, *residual* impedance  $Z_r$ . The total impedance, exclusive of end-region resistances, is then

$$Z_{\text{tot}} = Z_r + Z_{L0} + Z_{0R}. \quad (8)$$

Typically, the residual impedance term is small in comparison to the other terms. Also, for narrow structures and at sufficiently low frequencies such that  $v/\omega$  is large compared to the width of the avalanche

region, the hole and electron current fractions are equal in value and will be denoted by  $F$  without subscript.

In Section II we consider, following Gilden and Hines,<sup>6</sup> the consequences of a simplified model of the avalanche zone, as embodied in Read's (13), (Ref. 1) and derive an expression for  $F$  from it

$$F = \frac{1}{1 - \left(\frac{\omega}{\omega_c}\right)^2}. \quad (9)$$

Here the quantity  $f_c = \omega_c/2\pi$ , which we shall call the *critical frequency*, is real, independent of operating frequency  $\omega/2\pi$  and is proportional to the square root of the dc current density.

Section III presents a treatment that is rigorous except for the idealizing assumption that hole and electron velocities are constant and equal in value. Numerical results for several structures are given in Section IV, and it will be shown that for frequencies  $f$  and direct current densities of interest  $F$  can be represented as

$$F = \frac{(jf - B)}{(jf - A)(jf - \bar{A})} \left(\frac{A\bar{A}}{-B}\right), \quad (10)$$

where  $B$  is a zero on the real axis and is nearly independent of current density and  $A$  and  $\bar{A}$  are complex conjugate poles that traverse a parabolic path in the complex frequency plane as the current density is changed.

In Section V a qualitative description of the change in negative resistance characteristics is given as a transition is made from p-i-n to p-n to "Read" diode. Admittance as function of frequency and current density is shown for two specific structures. Values for the exponential growth rate factor  $g (= -1/2Q)$ , maximized with respect to frequency, are shown as a function of current density. It is found that an optimal growth rate is obtained when the current density is such that the resonance frequency is about equal to one half the reciprocal transit time through the longest drift region.

In Section VI we explore the consequences of the assumption that best large signal performance is obtained for the same conditions that yield largest small signal growth rate. This assumption in conjunction with a constraint expressing drift-region output limitations leads to Read-diode designs in which the width of the avalanche region is a substantial fraction ( $\approx \frac{1}{3}$ ) of the drift region.

## II. SIMPLIFIED MODEL

The current fraction  $F$  can be deduced<sup>6</sup> from the simplified theory of Read.<sup>1</sup> His (13) for a thin avalanche region

$$\frac{\tau_1}{2} \frac{dI_0}{dt} = I_0 \left( \int \alpha dx - 1 \right) + I_s \quad (11)$$

states that the time derivative of the particle current density through the avalanche region,  $I_0$ , equals the product of  $I_0$  and a field dependent actor  $h(E)$

$$h(E) = \frac{2}{\tau_1} \left( \int \alpha dx - 1 \right) \quad (12)$$

if the saturation current  $I_s$  can be neglected, as is usually the case. Under dc conditions  $h(E) = 0$ . If  $i_p$  and  $e$  represent small signal ac components at frequency  $\omega$  of particle current density and electric field in the avalanche region, then

$$i\omega i_p = I_0 \frac{\partial h}{\partial E} e. \quad (13)$$

The total alternating current density  $i_{tot}$  equals the sum of particle current density and displacement current density,

$$i_{tot} = i_p + i\omega\epsilon e \quad (14)$$

or, with the definition

$$F = i_p / i_{tot}, \quad (15)$$

$$\frac{e}{i_{tot}} = \frac{1}{i\omega\epsilon} (1 - F). \quad (16)$$

Division of (13) by  $i_{tot}$  and substitution of  $e$  from (16) yields

$$F = \frac{1}{1 - \left( \frac{\omega}{\omega_{crit}} \right)^2} \quad (17)$$

with

$$\omega_{crit} = 2\pi f_{crit} = \sqrt{I_0 \frac{\partial h}{\partial E} / \epsilon}. \quad (18)$$

Equations (17) and (18) predict a pole in the impedance or a zero in the admittance at a critical frequency  $f_{crit}$  that is proportional to the square root of the current density. Experimental measurements<sup>7</sup> and

numerical calculations<sup>8</sup> have shown that in p-n diodes and p-i-n diodes the admittance goes through a minimum, but not through zero, at frequencies approximately proportional to the square root of current density. This indicates that the simplified theory based on (11) is roughly valid but not accurate in detail. In the next Section a treatment is presented that does not make the assumptions inherent in (11) and that is, therefore, applicable to wide as well as narrow avalanche regions.

### III. IMPROVED MODEL

In this section we give a more rigorous treatment of the avalanche zone. The present treatment is facilitated greatly by the idealizing assumption that hole and electron velocities are constant and equal in value. If this assumption is not made the problem is still tractable and a method of solution was obtained.<sup>9</sup> However, it is felt that a somewhat idealized treatment with the attendant reduction in complexity and, hopefully, gain in physical insight, is worthwhile. This is especially the case since experimental values for particle velocity vs electric field are available only for prebreakdown fields. Though electron velocities in silicon, the material of present greatest interest, can be extrapolated reasonably well into the breakdown region, considerable uncertainty prevails about hole velocities. However, since in silicon the electrons ionize much more strongly (roughly a factor of 10 more) than holes, it is felt that the results of calculations for the current fractions  $F$  are not affected significantly by the choice of hole velocity. Thus, the assumption that the hole velocity equals the electron velocity appears adequate. The value of  $10^7$  cm/sec is used for the velocity  $v$ .

For the ionization coefficients  $\alpha$  and  $\beta$  of holes and electrons as function of electric field  $E$  the expressions

$$\alpha = 1.8 \times 10^7 \text{ (cm}^{-1}\text{)} \exp(-3.2 \times 10^6 \text{ (V/cm)}/E)$$

$$\beta = 2.4 \times 10^6 \text{ (cm}^{-1}\text{)} \exp(-1.6 \times 10^6 \text{ (V/cm)}/E)$$

are used. The numerical values refer to silicon at room temperature and are based on the work of Lee et al.<sup>10</sup>

Let  $I_h$  and  $I_e$  be the hole and electron current densities. Then the continuity equations state

$$\frac{1}{v} \frac{\partial}{\partial t} I_h = -I_h' + \alpha I_h + \beta I_e \quad (19)$$

$$\frac{1}{v} \frac{\partial}{\partial t} I_e = I_e' + \alpha I_h + \beta I_e \quad (20)$$

where primes denote spatial derivatives. Here we assume that holes move in the  $+x$  direction and electrons in the  $-x$  direction corresponding to a positive electric field as in an n-p structure. We introduce the total electric current associated with particle motion

$$I_{\Sigma} = I_e + I_h \quad (21)$$

and the difference in currents

$$I_{\Delta} = I_e - I_h. \quad (22)$$

Addition of (19) and (20) yields

$$\frac{1}{v} \frac{\partial}{\partial t} I_{\Sigma} = I_{\Delta}' + (\beta + \alpha) I_{\Sigma} + (\beta - \alpha) I_{\Delta}. \quad (23)$$

From Poisson's equation

$$I_{\Delta} = -v\epsilon E' + qvN_D, \quad (24)$$

where  $E$  is the electric field and  $qN_D$  the net impurity space charge. We shall now again consider small ac quantities at frequency  $\omega$ , denoted by lower case symbols, superimposed on dc quantities and we denote by  $i_{\text{tot}}$  the total ac current. From the continuity of total current we obtain

$$i_{\text{tot}} = i_{\Sigma} + i\omega\epsilon e. \quad (25)$$

We define

$$k = i\omega/v \quad (26)$$

and normalize the ac electric field to the total alternating current

$$z = \frac{v\epsilon}{i_{\text{tot}}} e. \quad (27)$$

Then (25) can be written

$$\frac{i_{\Sigma}}{i_{\text{tot}}} = 1 - kz. \quad (28)$$

We introduce the derivative with respect to field of the last two terms of (23)

$$H = \frac{1}{v\epsilon} \left[ I_{\Sigma} \frac{\partial}{\partial E} (\beta + \alpha) + I_{\Delta} \frac{\partial}{\partial E} (\beta - \alpha) \right]. \quad (29)$$

Then the small signal ac version of (23) is

$$-z'' + (\alpha + \beta - k)(1 - kz) - (\beta - \alpha)z' + Hz = 0 \quad (30)$$

or

$$[D^2 - k^2 + (\beta + \alpha)k + (\beta - \alpha)D - H]z = \alpha + B - k, \quad (31)$$

where  $D$  is the spatial derivative operator, and where  $\alpha$ ,  $\beta$ , and  $H$  are evaluated for the average, or dc, field as a function of distance.

Equation (31) is a second-order differential equation for the complex quantity  $z$ . We postulate that to the left of the avalanche zone all particle current is electron current, i.e.,

$$i_z = \text{plane wave moving to left} \quad (32)$$

and that to the right of the avalanche zone all particle current is hole current

$$i_z = \text{plane wave moving to right.} \quad (33)$$

We do not specify that the magnitude of the ac electron current on the left equals that of the hole current on the right. Since we differ in this point from previous treatments, some discussion may be warranted.

Holes and electrons are generated in pairs. One might, therefore, be led to conclude that the magnitudes of hole and electron particle currents must be equal. This is true for the dc or average currents but not generally so for the ac currents. Consider that a periodic generation rate  $g(x)$  is given. We allow  $g$  to be complex to represent the variation in phase with distance. The continuity equations (19) and (20) for the small signal case may be written

$$(k + D)i_h = g \quad (34)$$

$$(k - D)i_e = g. \quad (35)$$

With  $g$  considered given, these are first-order differential equations for  $i_h$  and  $i_e$ . With the boundary conditions (32) and (33) the solutions are

$$i_h(x) = \int_{x_L}^x g(s) \exp[-k(x-s)] ds \quad (36)$$

$$i_e(x) = \int_x^{x_R} g(s) \exp[+k(x-s)] ds. \quad (37)$$

Since  $k$  is purely imaginary and  $g$  is complex, it is seen that  $i_h(x_R)$  and  $i_e(x_L)$  need not be the same. For narrow structures and at low frequencies, the difference in magnitude is negligible. This, however, is the result of the calculations and not imposed as a constraint.

We specify, then, that outside the avalanche region the particle current consists of a plane wave moving to the left or right.

$$\frac{i_{\Sigma}}{i_{\text{tot}}} = A_R \exp(-kx) \quad \text{right} \quad (38)$$

$$\frac{i_{\Sigma}}{i_{\text{tot}}} = A_L \exp(+kx). \quad \text{left} \quad (39)$$

The amplitudes  $A_R$  and  $A_L$  are as yet undetermined. By (28) the electric field outside the avalanche region is

$$z = \frac{1}{k} [1 - A_R \exp(-kx)] \quad \text{right} \quad (40)$$

$$z = \frac{1}{k} [1 - A_L \exp(+kx)] \quad \text{left} \quad (41)$$

and we obtain the boundary conditions

$$z' = 1 - kz \quad \text{right} \quad (42)$$

$$z' = -(1 - kz) \quad \text{left} \quad (43)$$

in accordance with (22) and (24).

These boundary conditions specify completely the solution of (31). If the solution for  $z$  is carried over the entire depletion region, then the integral over  $z$  gives the total impedance, exclusive of resistive losses in the end regions. However, in order to isolate the role of the avalanche region and to bring the impedance into the form (8) we proceed as follows: If a solution for  $z$  has been obtained at some sufficiently low frequency, then the magnitudes of  $A_R$  and  $A_L$  in (38) and (39) are equal. A value  $x_0$  exists at which the right hand side of (38) equals the right hand side of (39); i.e., at which the asymptotic particle currents, extrapolated into the avalanche zone, are equal in phase as well as magnitude. We call  $x_0$  the phase reference point and define the extrapolated particle currents at this point as the current fractions  $F_e$  and  $F_h$ ;

$$F_h = A_R \exp(-kx_0) \quad (44)$$

$$F_e = A_L \exp(+kx_0). \quad (45)$$

We define the asymptotic particle current  $i_a$ ,

$$i_a/i_{\text{tot}} = \begin{cases} F_h \exp[-k(x - x_0)] & x \geq x_0 \\ F_e \exp[+k(x - x_0)] & x < x_0 \end{cases} \quad (46)$$

and the residual particle current  $i_r$ , as the difference between the actual particle current  $i_{\text{tot}}(1 - kz)$  and the asymptotic particle current

$$\frac{i_r}{i_{tot}} = 1 - kz - \frac{i_a}{i_{tot}}. \quad (47)$$

Note that by construction  $i_r$  vanishes outside the avalanche zone and is thus independent of the dimensions of the drift zones, as are  $F_e$  and  $F_h$ . We can now solve (47) for  $z$ ,

$$z = \frac{1}{k} \left[ 1 - \frac{i_r}{i_{tot}} - \begin{cases} F_h \exp[-k(x - x_0)] & x \geq x_0 \\ F_e \exp[+k(x - x_0)] & x < x_0. \end{cases} \right] \quad (48)$$

The integral of  $z$ , multiplied by  $1/v\epsilon = k/(i\omega\epsilon)$  is the total impedance

$$Z = Z_r + \frac{x_0 - x_L}{i\omega\epsilon} \left[ 1 - F_e \frac{\exp(i\omega\tau_L)}{i\omega\tau_L} \right] + \frac{x_R - x_0}{i\omega\epsilon} \left[ 1 - F_h \frac{1 - \exp(i\omega\tau_R)}{i\omega\tau_R} \right] \quad (49)$$

$$= Z_r + Z_{L0} + Z_{0R}, \quad (50)$$

with

$$Z_r = -\frac{1}{i\omega\epsilon} \int \frac{i_r}{i_{tot}} dx. \quad (51)$$

We shall use the symbol  $F$  without subscript to refer to either  $F_e$  or  $F_h$  if their difference is negligible or the distinction unimportant.

#### IV. NUMERICAL RESULTS

This section contains numerical results for the following structures in silicon:

- D3 error function complement p-diffusion from a surface concentration of  $10^{20} \text{ cm}^{-3}$  into a  $3 \times 10^{16} \text{ cm}^{-3}$  n-type substrate with a junction depth of 3 microns.
- D12 similar to D3, but with a 12-micron junction depth.
- L22 a linearly graded junction with a concentration gradient of  $10^{22}$  impurities/ $\text{cm}^4$ .
- C0.1 Constant field avalanche zone (field =  $5.71 \times 10^5$  volts/cm) of 0.1-micron width, surrounded by regions of sufficiently low field that negligible avalanching takes place there.
- C1 Similar to C0.1, but with 1-micron width; field =  $3.57 \times 10^5$  volts/cm.
- C5 Similar to C0.1, but with 5-micron width; field =  $2.79 \times 10^5$  volts/cm.

C10 Similar to C0.1, but with 10-micron width; field =  $2.52 \times 10^5$  volts/cm.

For each desired direct current density  $I_{\Sigma}$ , self-consistent values of dc electric field  $E$ , difference current density  $I_{\Delta}$ , ionization coefficients  $\alpha$  and  $\beta$ , and derivative quantity  $H$ , (29), were computed as described in Appendix A. Using these quantities, the differential equation (31), subject to boundary conditions (42) and (43), was solved numerically for  $k$ -values corresponding to frequencies of interest.

Figs. 2 and 3 show the results for the quantity  $1 - 1/F = (-\text{displacement current/particle current})$  for direct current densities of 100 and 1000 amps/cm<sup>2</sup>. According to the simplified model of Section II,  $1 - 1/F$  should have a real part varying as the square of frequency and a vanishing imaginary part. As Figs. 2 and 3 show, the square law is obeyed quite well at low frequencies and current densities. Only for the wide structures and/or at high frequencies and current densities does  $F$  deviate from the square law and do  $F_e$  and  $F_h$  deviate from each other appreciably. The frequency  $f_{\text{crit}}$  at which  $(1 - 1/F)_{\text{real}}$  is unity

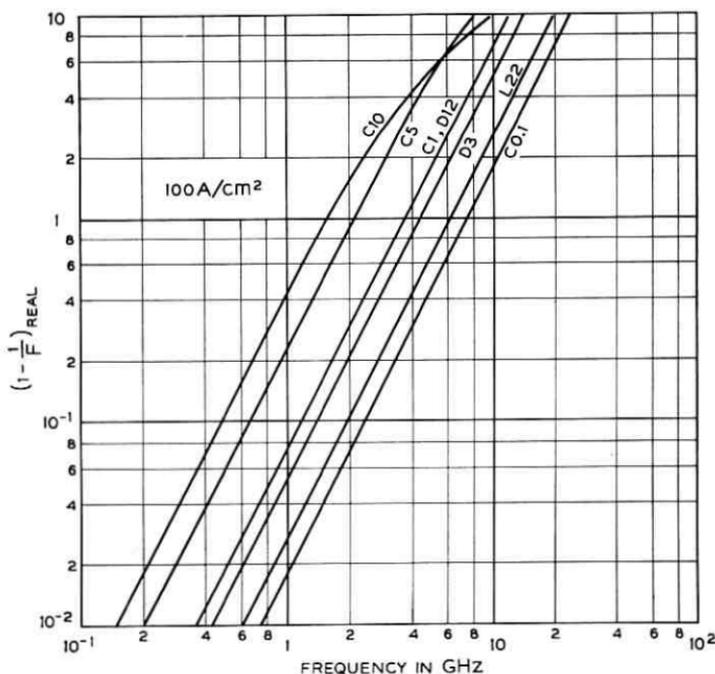


Fig. 2 — Real part of  $1 - 1/F$  as function of frequency for various structures. Current density 100 amps/cm<sup>2</sup>.

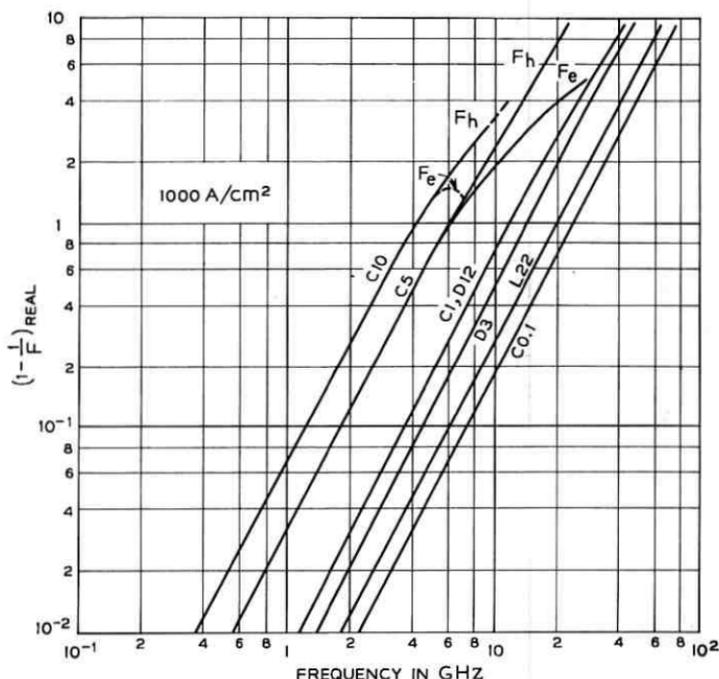


Fig. 3—Real part of  $1 - 1/F$  as function of frequency for various structures. Current density 1000 amps/cm<sup>2</sup>.

is in good agreement with values obtained from the simple theory of Section II.

The absolute value of the imaginary part of  $1 - 1/F$  at a current density of 100 amp/cm<sup>2</sup> is plotted in Fig. 4 as a function of frequency  $f$  for the various structures. The current dependence of the diffused diode D3 is shown in Fig. 5; that for the other structures is similar. The vertical line on each curve to the right of the maximum indicates the frequency at which the real part goes to zero. At frequencies 15 to 25 percent higher,  $(1 - 1/F)_{\text{imag}}$  goes through zero, being positive for lower and negative for higher frequencies. Asymptotically, at low frequencies  $(1 - 1/F)_{\text{imag}}$  is proportional to  $+f$  and at higher frequencies to  $-f^3$ . The similarity of the curves is striking and suggests that a fit with a few parameters ought to be possible. The most meaningful expansion, and one working well here, is to consider  $F$  as an analytical function of a complex frequency variable and to expand it in terms of the poles and zeroes nearest the frequency domain of interest. At current densities of interest and for frequencies  $f$  from zero to a few times

$f_{crit}$ , the numerical results can be represented well by two poles and a zero;

$$F = \frac{(if - B)}{(if - A)(if - \bar{A})} \cdot \frac{A\bar{A}}{-B}, \quad (52)$$

where  $\bar{\sim}$  denotes the complex conjugate.  $A$  is the complex resonance frequency

$$A = f_a + if_r. \quad (53)$$

We denote

$$B = -f_c.$$

We utilize the following inequalities which hold for current densities of interest:

$$f_a \ll f_r \quad (54)$$

$$f_r \ll f_c \quad (55)$$

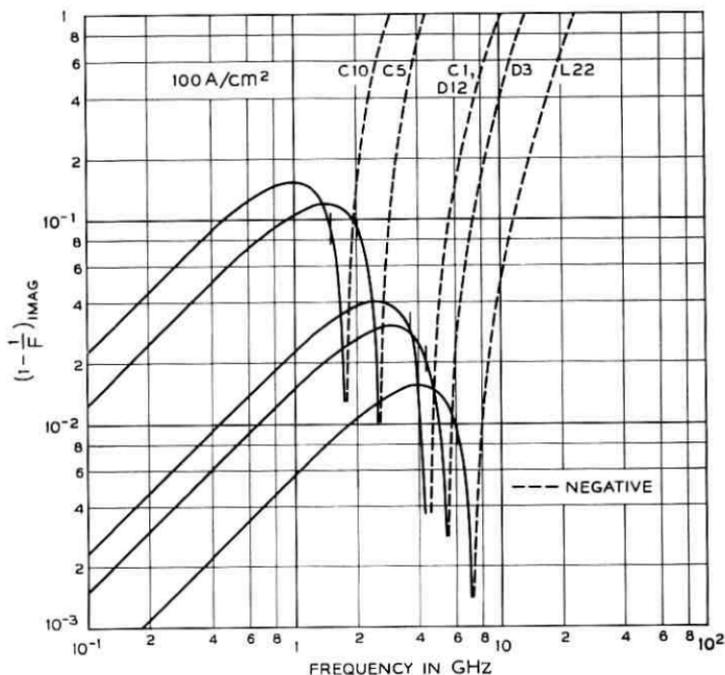


Fig. 4—Imaginary part of  $1 - 1/F$  as function of frequency for various structures. Current density 100 amps/cm<sup>2</sup>. The vertical bars indicate  $f_{crit}$ .

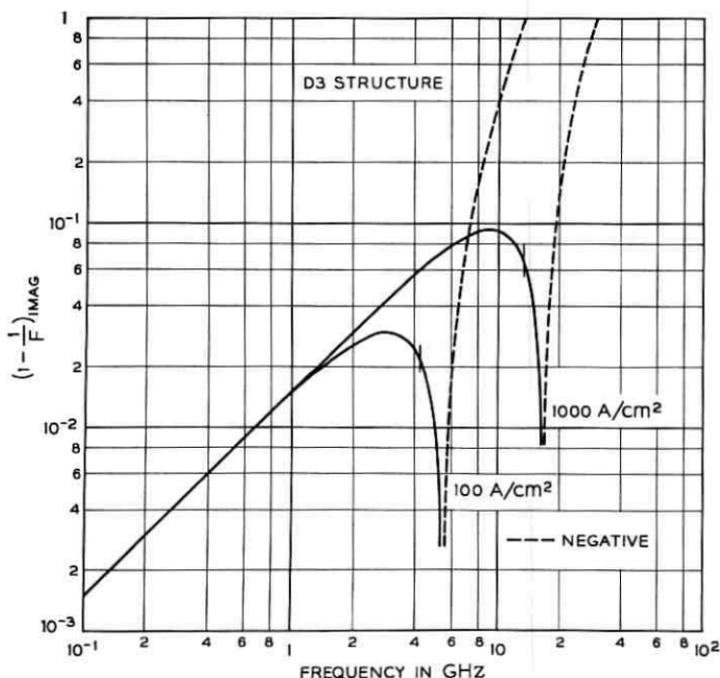


Fig. 5—Imaginary part of  $1 - 1/F$  as a function of current density for diode D3 at 100 and 1000 amps/cm<sup>2</sup>.

and we consider frequencies small compared to  $f_c$ . Then the reciprocal of (52) can be written

$$\frac{1}{F} = 1 - \frac{f^2}{f_r^2} - jf \left( \frac{1}{f_c} + \frac{2f_a}{f_r^2} - \frac{f^2}{f_c f_r^2} \right). \quad (56)$$

Equation (56) has the frequency dependence embodied in Figs. 2 through 5, i.e., the real part of  $1 - 1/F$  changes as frequency squared and the imaginary part is the difference of a linear and cubic term in frequency. The low-frequency asymptote of the imaginary part is

$$(1 - 1/F)_{\text{imag}} = f/f_s, \quad (57)$$

with

$$f_s = \frac{f_c}{1 + 2f_a f_c / f_r^2}. \quad (58)$$

The frequency,  $f_{\text{crit}}$ , at which the real part of  $1/F$  vanishes equals  $f_r$ :

$$f_{\text{crit}} = f_r \quad (59)$$

and the frequency  $f_i$  at which the imaginary part of  $1/F$  vanishes is

$$f_i = \sqrt{f_r^2 + 2f_a f_c}. \quad (60)$$

Now let us consider the dependence of  $F$  on the direct current density  $I$ . Three relations hold approximately for the curves of Figs. 2 through 5.

The critical frequency  $f_{\text{crit}}$  is proportional  
to the square root of direct current density. (61)

The low-frequency, linear asymptote of  $(1 - 1/F)_{\text{imag}}$   
is independent of current density. (62)

The ratio of the frequencies  $f_{\text{crit}}$  and  $f_i$  at which  
the real and imaginary parts of  $1/F$  cross zero  
is independent of current density. (63)

We show now that these relations lead to the following current density dependences:

$$f_a \propto I \quad (64)$$

$$f_r \propto \sqrt{I} \quad (65)$$

$$f_c \text{ independent of } I. \quad (66)$$

Equation (65) follows directly from (59) and (61). Taking the ratio of (60) and (59) and using (63) we find that

$$f_a f_c / f_r^2 \text{ is independent of } I. \quad (67)$$

Use of (62), (57), (58), and (67) yields (66). Finally, from (65), (66), and (67) we obtain (64). Thus, as long as the inequalities (54) and (55) and relations (61) through (63) hold, the current density dependence of  $F$  can be expressed as follows (see Fig. 6): *With increasing current density the poles  $A$  and  $\tilde{A}$  of  $F$  move through parabolas, starting at the origin, while the zero at  $-f_c$  is independent of current density.* Where  $F_e$  and  $F_h$  deviate from each other, the representation (52) is still usable. The poles  $A$  and  $\tilde{A}$  are the same, but  $F_e$  and  $F_h$  have separate zeros at  $-f_{ce}$  and  $-f_{ch}$ .

The above discussion shows how the poles and zero can be obtained from calculated results of  $F$  as a function of frequency and current density. In Appendix B a more direct way of obtaining the poles as solutions of an eigen-value problem is presented. The results are shown in Fig. 6, where the poles are plotted in the complex frequency plane with current density varying along the curves. The zeroes are listed in the insert. Only positive imaginary values are shown. Symmetry exists about the

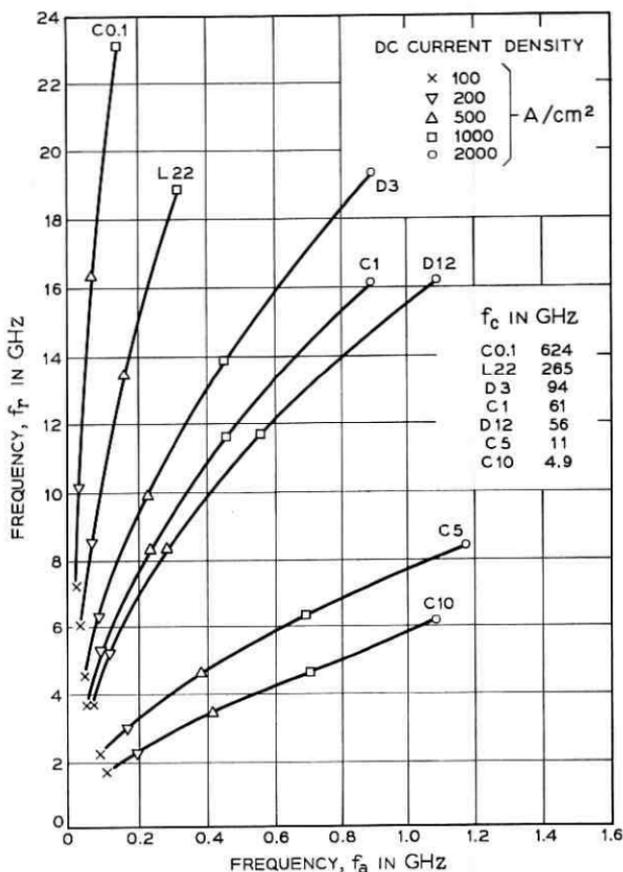


Fig. 6—Locus of poles of  $F$  as function of current density in the complex frequency plane. Note unequal scales of real and imaginary axes.

real axis. Note that the scales of real and imaginary axes differ by a factor of 10.

#### V. NEGATIVE RESISTANCE AND MAXIMUM GROWTH FACTOR

In this section, we consider the interaction of the avalanche and drift regions in causing negative resistance. The most important aspect of the current fraction  $F$  is the rapid change of phase, with respect to total current, from near zero below, to more than  $180^\circ$  above, the resonance frequency. This behavior follows from (52) and is illustrated for structure D3 in Fig. 7 where the phase  $\varphi_F$  of  $F$  is shown for current densities of 100 and 1000  $A/cm^2$ . At the higher current density the transition is

more gradual. Likewise, for wider structures the transition is more gradual than for narrower structures.

Let us consider a diode in which the hole drift region is negligible and for which the residual impedance  $Z_r$  is negligible so that the dominant term in the impedance is

$$Z_{L0} = \frac{1}{i\omega C_L} \left[ 1 - F \frac{1 - \exp(-i\omega\tau_L)}{i\omega\tau_L} \right]. \quad (68)$$

We want to explore under what conditions the small signal impedance has a negative real part. This will occur when the quantity

$$F \frac{1 - \exp(-i\omega\tau_L)}{i\omega\tau_L} \equiv FG \quad (69)$$

has a phase angle between 0 and 180°.  $\varphi_G$ , the negative of the phase of  $G$ , is plotted in Fig. 8 as a function of frequency. We define as drift frequency,  $f_{\text{drift}}$ , the frequency at which the phase  $\varphi_G$  is 90°,

$$f_{\text{drift}} = \frac{v}{2(x_0 - x_L)} = \frac{1}{2\tau_L}. \quad (70)$$

The singularity at twice the drift frequency occurs where  $G$  goes to zero and the phase is indeterminate. If the operating frequency is below  $1/\tau_L$  we refer to the diode as operating in the  $\pi$  mode. The mode-

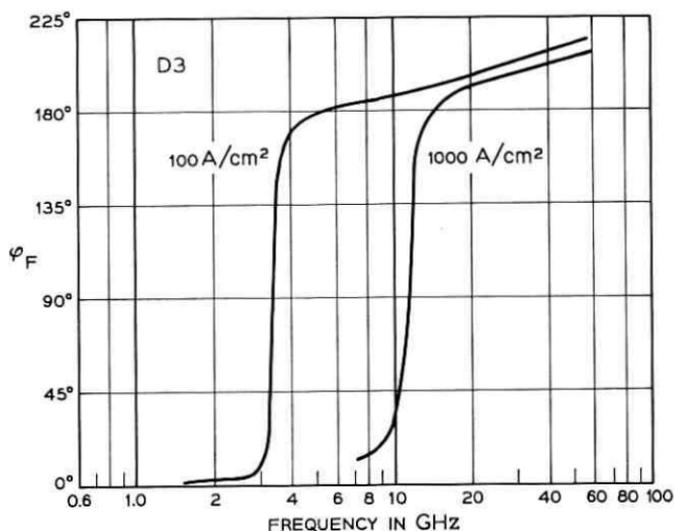


Fig. 7 — Phase of  $F$  for diode D3 at 100 and 1000 amps/cm<sup>2</sup>.

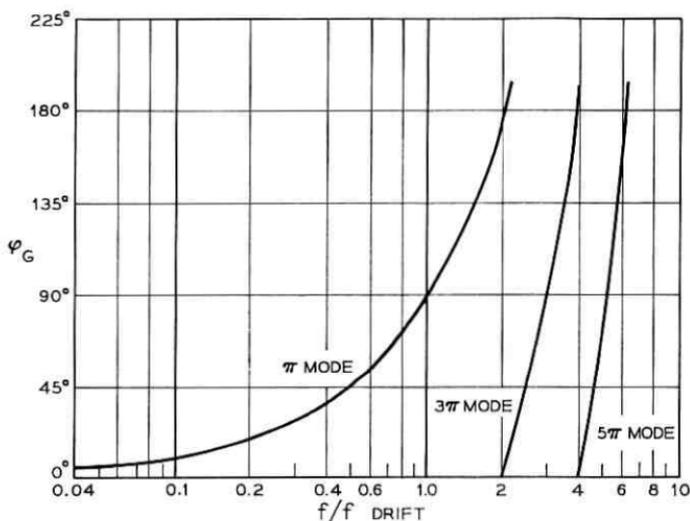


Fig. 8 — Negative of drift delay phase.

designation for higher frequencies is given in Fig. 8. The condition on the phases for negative resistance can now be stated

$$0 < \varphi_F - \varphi_G < 180^\circ \quad (71)$$

and can easily be visualized if Figs. 7 and 8 are combined. This has been done in Fig. 9 where a drift frequency of 60 GHz appropriate for the depletion layer of diode D3 has been used. Starting with the 100 A/cm<sup>2</sup> curve for  $\varphi_F$ , we see that negative resistance prevails over a wide frequency range, as indicated by the shaded region. Near  $f = 2f_{\text{drift}}$  the magnitude of  $G$  is small, hence the negative resistance contribution of  $FG$  is small and is likely to be outweighed by parasitic end resistance. At frequencies beyond  $2f_{\text{drift}}$  negative resistance occurs in higher-order modes. At frequencies much higher than  $f_{\text{crit}}$  the phase  $\varphi_F$  depends on frequency in a complicated way and the present calculations lose accuracy. As the current density is increased, the  $\varphi_F$  curve moves to the right with some softening of the transition. The lowest frequency at which negative resistance sets in increases also, but a wide range of negative resistance continues to exist. The behavior thus far described is typical of a single-diffused p-n junction diode.

Next, consider a typical "Read" structure, i.e., one in which the drift space is much wider than the avalanche region. We consider the same avalanche zone as before but let the drift space be, say 8 microns

with a drift frequency of 6.25 GHz, as shown in Fig. 10. We now see that the frequency range for negative resistance in the  $\pi$  mode is much narrower; at 1000 A/cm<sup>2</sup> there is no  $\pi$ -mode negative resistance for this structure.

Now let us consider structures with narrow drift regions. The extreme case for which the drift regions coincide with the avalanche region is represented by p-i-n diodes.<sup>8</sup> All three terms in (8) contribute, but the conclusions reached from a study of  $Z_{L0}$  alone are qualitatively correct. Negative resistance is obtained for a wide range of frequencies. For a consideration of the low-frequency behavior we expand  $F$  to first order in  $f$

$$F \approx 1 + i \frac{f}{f_s} \quad (72)$$

(cf. (57) and (58)). With

$$G \approx 1 - \frac{i\omega\tau_L}{2} = 1 - i \frac{\pi f}{2f_{\text{drift}}} \quad (73)$$

(68) yields

$$R_{L0} = \frac{1}{4f_{\text{drift}} C_L} \left( 1 - \frac{2f_{\text{drift}}}{\pi f_s} \right). \quad (74)$$

Thus, if

$$f_{\text{drift}} < \frac{\pi}{2} f_s, \quad (75)$$

negative resistance prevails from  $\pi$  mode frequencies down to dc. The case of the uniform-field avalanching plasma has been discussed by Misawa.<sup>8</sup>

For an interpretation of the frequency  $f_s$ , consider the change in electric field  $\delta E$ , per change in total current in the limit as the frequency goes to zero, as shown in Fig. 11, for diode D3 at 100 A/cm<sup>2</sup>. The space charge associated with the additional current causes  $\delta E$  to be roughly hyperbolically shaped with the apex at the point where hole and electron currents are equal. The placement of the  $\delta E$  curve with respect to zero is such that reduction of ionization in the center where  $\delta E$  is negative is compensated by an increase in the adjacent regions. Outside the avalanche zone  $\delta E$  changes linearly. At points A and B, separated by distance  $d_{AB}$  the asymptotes to  $\delta E$  outside the avalanche zone cross zero. From a consideration of (28) through (32) it can be seen that  $1/f_s$  is half the transit time between points A and B,

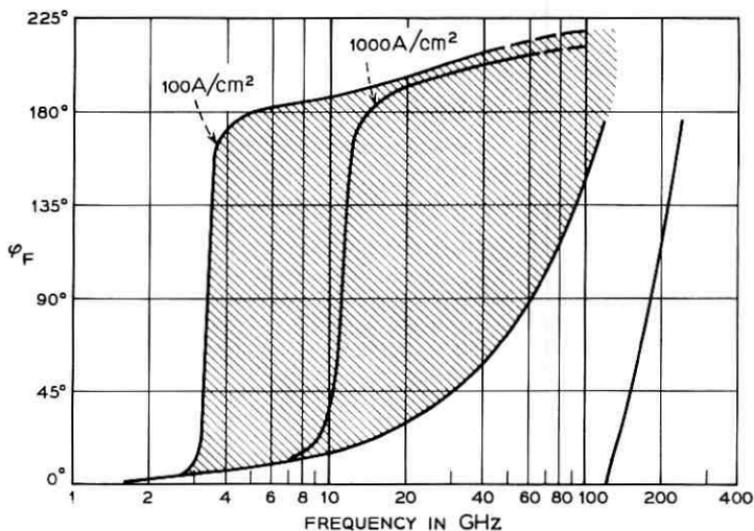


Fig. 9—Superposition of Figs. 8 and 9 appropriate for p-n junction case.

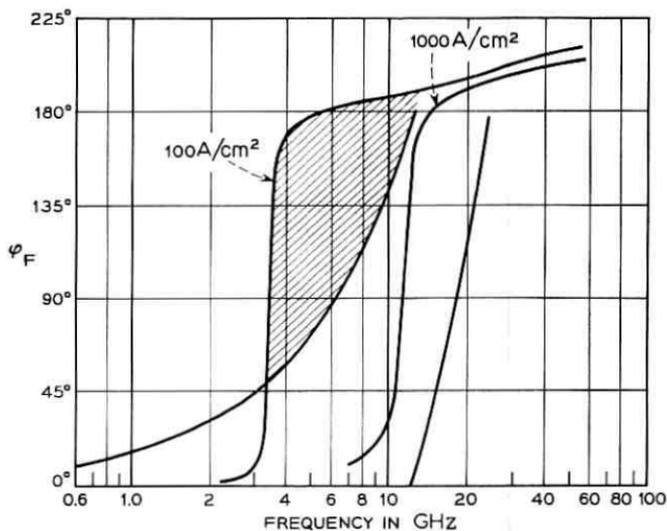


Fig. 10—Superposition of Figs. 8 and 9 appropriate for "Read" diode case.

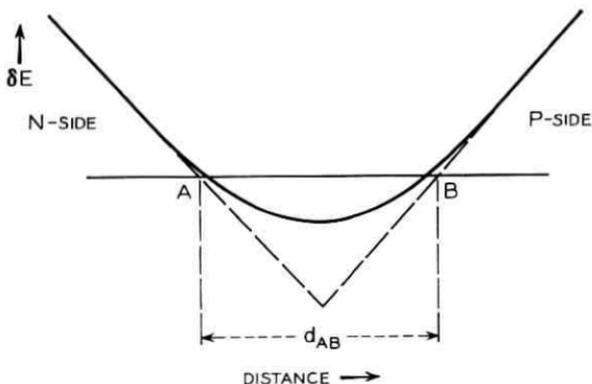


Fig. 11—Incremental field at low frequencies for p-n diode.

$$f_s = \frac{d_{AB}}{2v} \quad (76)$$

If the avalanche zone is narrow compared to the drift region the integral over  $\delta E$  is positive and so is the dc incremental resistance.

For a p-i-n diode (C1 at  $100 \text{ A/cm}^2$ ) the incremental field  $\delta E$  is as shown in Fig. 12. The asymmetry results from the inequality of hole and electron ionization coefficients. It is seen by inspection that the integral over  $\delta E$  is negative and also that the transit time between A and B is less than  $\pi/2$  times the transit time through the depletion layer (this is the condition corresponding to (75) when both left and right drift regions are considered), hence the resistance is negative. It is to be noted, however, that if equal hole and ionization rates had been used in this calculation, zero incremental dc resistance would have been obtained.

Up to this point we have considered the phase conditions for the occurrence of negative resistance. Now let us consider the actual resistance or, more conveniently, the admittance  $Y$ . Fig. 13 shows plots of the imaginary vs real part of  $Y$ , with frequency varying along each curve, for diode D3, area =  $10^{-4} \text{ cm}^2$ , at various current densities. Fig. 14 shows a corresponding plot for a Read structure R1 consisting of avalanche region C1 (1 micron wide, constant field) and an adjacent electron drift space 9 microns wide. The diode area is also taken as  $10^{-4} \text{ cm}^2$ . The admittances in Figs. 13 and 14 are the reciprocal total impedance, but differ insignificantly from what would have been obtained with neglect of the residual impedance.

Now let us consider the quality factor  $Q$  or the related *growth factor*  $g$ ,

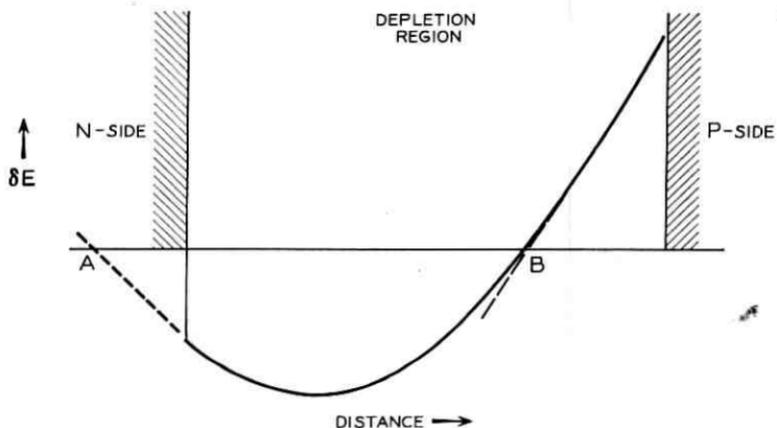


Fig. 12—Incremental field at low frequencies for p-i-n diode.

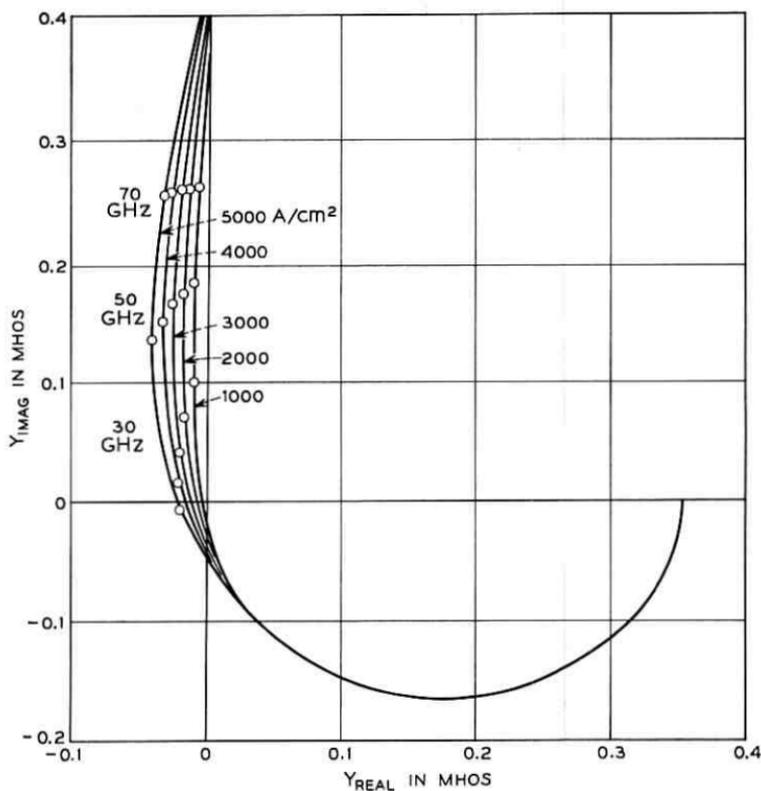


Fig. 13—Admittance of diode D3. Selected frequencies, in GHz, are marked off on each curve.

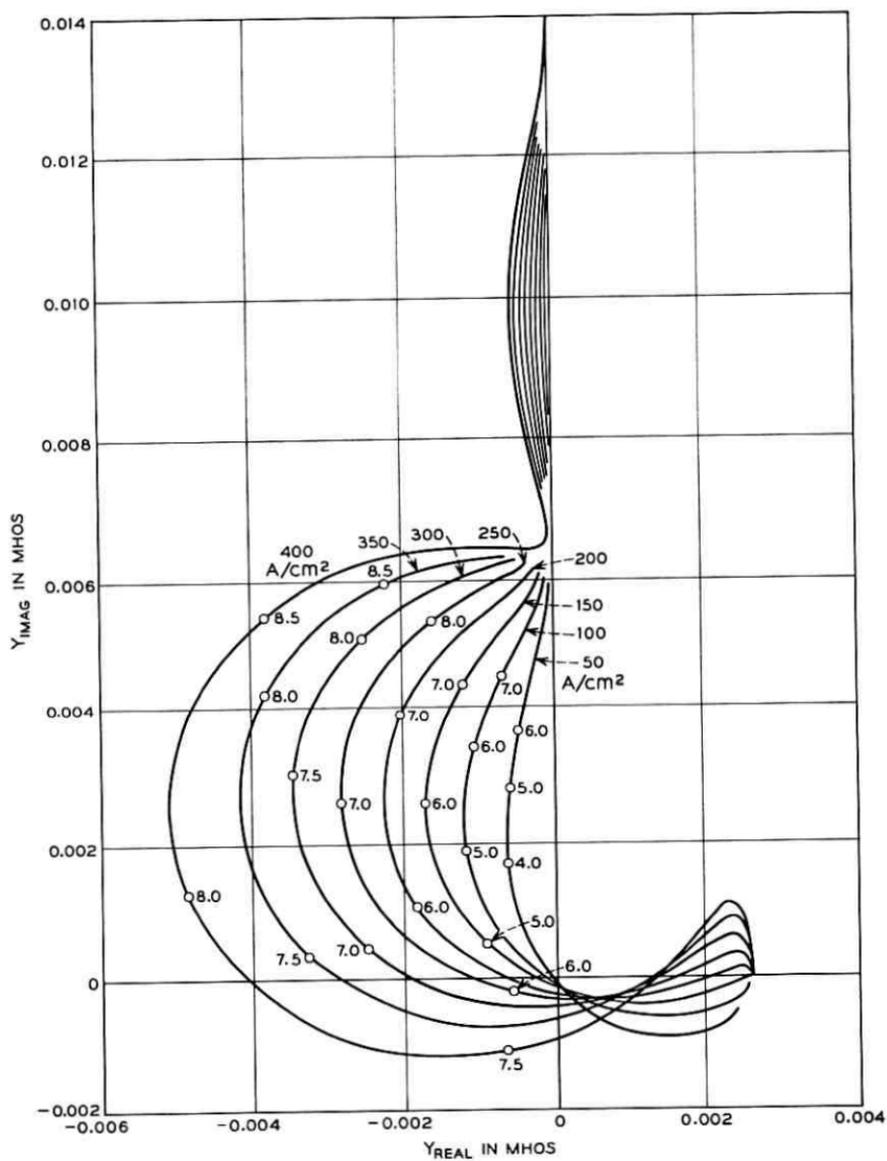


Fig. 14—Admittance of Read diode R1. Selected frequencies, in GHz, are marked off on each curve.

$$g = -\frac{1}{2Q} \quad (77)$$

which has the following meaning: If a diode having negative resistance is imbedded in a lossless circuit in which it oscillates at frequency  $f = \omega/2\pi$ , then the absolute values of voltage and current amplitudes vary with time as  $\exp(g\omega t)$ . It is convenient to introduce the complex frequency variable  $s$

$$s = (g + i)f. \quad (78)$$

From (4), (52), and (53) it can be inferred that under open-circuit conditions the diode will oscillate at a value

$$s = A = f_a + if_r$$

and, therefore, the open-circuit growth factor is

$$g_{\text{open}} = f_a/f_r.$$

According to the current-dependence of Fig. 6,  $g_{\text{open}}$  is proportional to the square root of the current density. Note that under open-circuit conditions  $g$  depends on the avalanche structure only and is independent of the drift regions. This is not the case under more general circuit conditions.

If the admittance  $Y$  is an analytic function of frequency, as is physically reasonable and as we shall assume to be the case here, then the admittance is defined also for complex frequencies. The largest growth rate  $g(f)$  for reactive circuit imbedding at a given frequency  $f$  is that  $g$  for which

$$Y[(g + i)f] + \frac{1}{2\pi(g + i)fL} = 0 \quad (79)$$

if  $Y$  is capacitive; or

$$Y[(g + i)f] + 2\pi(g + i)fC = 0 \quad (80)$$

if  $Y$  is inductive. Here  $L$  or  $C$  are chosen so as to resonate the imaginary part of  $Y$ . The quantity  $g$  thus obtained is related to the quality factor  $Q$ , as conventionally defined, by (77). So far,  $g$  is defined at a given frequency  $f$  and current density. We define as  $g_{\text{max}}$  the maximum of  $g$  (or maxima, where relative maxima exist) with respect to frequency at a given current density. Finally, we define as  $g_{\text{opt}}$  the maximum of  $g_{\text{max}}$  with respect to current density. We define by  $f_{\text{max}}$  the frequency (ies) at which  $g_{\text{max}}$  occurs. Figs. 15 and 16 show  $g_{\text{max}}$  and  $f_{\text{max}}$  as a function

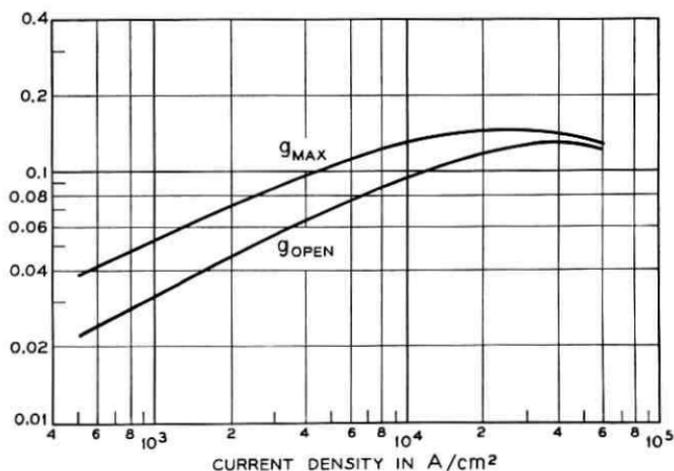


Fig. 15—Maximum growth rate and open-circuit growth rate of diode D3.

of current density for diode D3 and Figs. 17 and 18 for diode R1. Also shown for comparison are  $f_r$  and  $g_{open}$ . Included in Figs. 16 and 18 are horizontal lines marked  $f_{drift}$  representing the drift frequency of the longest (= electron) drift region. An important result contained in Figs. 15 to 18 is that  $g_{opt}$  is obtained for current densities at which the resonance frequency  $f_r$  is about equal to the drift frequency and that

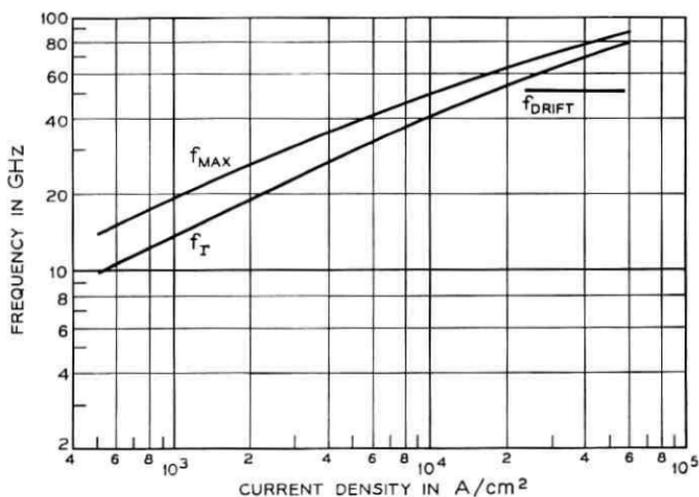


Fig. 16—Frequency at which  $g_{max}$  occurs and  $f_r$  for diode D3.

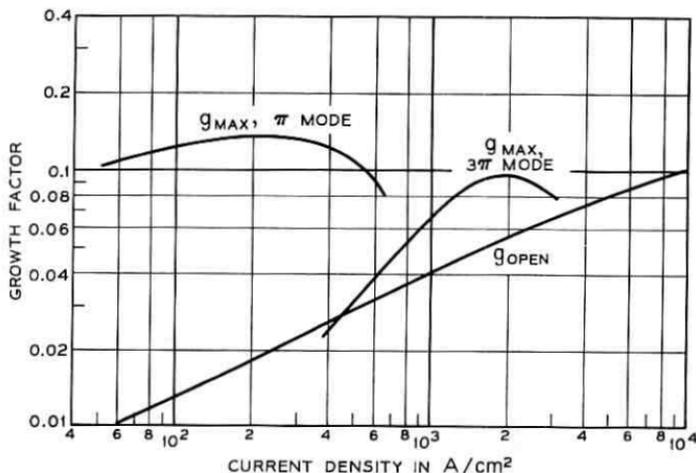


Fig. 17—Maximum growth rate and open circuit growth rate for Read diode R1.

$f_{\max}$  at this current density is some 20 to 30 percent above  $f_r$ . For the diffused diode D3 the drift frequency is high ( $\approx 60$  GHz) and therefore,  $g_{\text{opt}}$  is obtained at a high current density,  $3 \times 10^4$  amps/cm<sup>2</sup>. This current density is so high that a substantial widening of the avalanche region, with respect to its low-current-density configuration, has taken place. As a consequence,  $f_c$  now is no longer independent of current density, but rather decreases with current density. Deviations from the low-current-density relations are also seen in the break-away from the line of slope  $\frac{1}{2}$  for  $g_{\text{open}}$  in Fig. 15.

For the Read diode R1 the largest  $g_{\text{opt}}$ , occurring in the  $\pi$  mode, is obtained at 200 amps/cm<sup>2</sup>. The values of  $g_{\text{max}}$  in the  $n\pi$  mode fall off rapidly as the frequency approaches  $(n + 1)f_{\text{drift}}$ . For a Read-type structure, i.e., one having a substantial drift space in which negligible avalanche multiplication takes place,  $g_{\text{opt}}$  is near 0.13 and is only weakly dependent of the detailed structure of the avalanche region. On the other hand, for diodes in which the avalanche region occupies all of, or a substantial fraction of, the depletion region  $g_{\text{opt}}$  is closely related to the value of  $g_{\text{open}}$  and is larger than 0.13, as is the case for D3.

## VI. LARGE SIGNAL DESIGN CONSIDERATIONS FOR READ DIODES

In Section V we obtained the result that the optimum growth rate under small-signal conditions occurs when  $f_{\text{drift}} \approx f_r$ . This condition

provides a relation between current density, drift region width and avalanche region width (via  $f_r$ ). It is plausible to assume that the same relation should be approximately satisfied for best large-signal performance. Preliminary large-signal results\* have shown that a direct correspondence between the large-signal and small-signal properties does not exist and that, for example, large-signal self-sustained oscillation can be obtained for frequency-current-density combinations for which the small-signal resistance is positive. Nevertheless, as a point of departure for a large-signal design we choose a structure in which avalanche region, drift region, and current density are so related that  $g_{opt}$  is obtained at

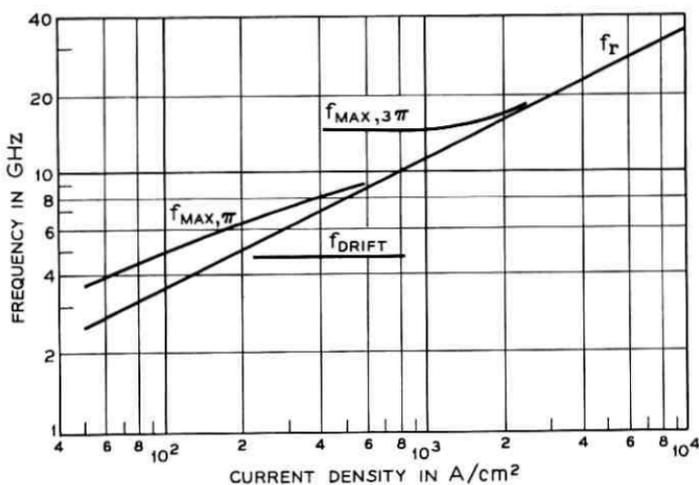


Fig. 18 — Frequency at which  $g_{max}$  occurs and  $f_r$  for diode R1.

the desired frequency under small signal conditions. This provides one constraint.

Resonance frequency  $f_r$  as a function of current density for the various structures is shown in Fig. 19. Suppose a 5-GHz operating frequency is wanted. We would need a resonance frequency about 25 percent lower or 4 GHz. Thus, structure C1 at 120 amps/cm<sup>2</sup>, or C5 at 400 amps/cm<sup>2</sup>, etc. could be used. The drift region would be 12.5 microns wide.

Another constraint between average current density and frequency is imposed by the output limitations of the drift space. The following discussion applies to Read diodes having a constant-field drift region only. As shown by Read,<sup>1</sup> under large-signal conditions the carrier current through the drift region is carried in the form of charge pulses. If

\* D. L. Scharfetter and H. K. Gummel, work in progress.

$f$  is the frequency of operation and  $I_{DC}$  the average current density, then the charge per pulse is

$$Q = \frac{I_{DC}}{f}. \quad (81)$$

This pulse causes a change in electric field of

$$\Delta E = \frac{I_{DC}}{(\epsilon f)}. \quad (82)$$

Let  $E_{\text{drift}}$  be the field in the drift region at the onset of breakdown; we shall use a value of  $1.5 \times 10^5$  V/cm which is reasonable for silicon.  $\Delta E$  must not be larger than approximately  $\frac{1}{2} E_{\text{drift}}$ . The factor of  $\frac{1}{2}$  takes into account that in the desired mode of operation the terminal voltage reaches its minimum while the charge moves through the drift region and that therefore an additional lowering of the drift field below  $E_{\text{drift}}$  occurs. Thus, we have the condition

$$I_{DC} < \frac{1}{2} \epsilon E_{\text{drift}} f. \quad (83)$$

This constraint is shown in Fig. 19 by the line of slope 1. Only the region to the upper left is allowed. For a diode to have a large power capability the highest feasible current density should be used. Thus, a design cor-

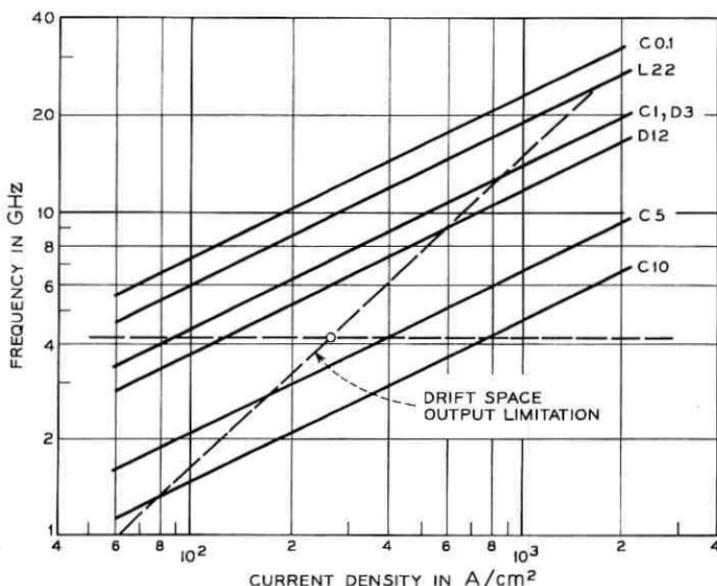


Fig. 19— Illustration of Read diode design considerations.

responding to the circle appears most promising. The diode would operate at a current density of 300 amps/cm<sup>2</sup> and have an avalanche region 4 microns wide. Note that the avalanche region width would be about  $\frac{1}{3}$  of the drift region width. This is in contrast to Read's proposal that the avalanche region should be very narrow in comparison to the drift region. However, the design suggested here is based on a combination of small-signal dynamics and large-signal output limitations and may have to be modified when large-signal results become available.

#### APPENDIX A

This appendix describes the steps by which self-consistent dc solutions for electric field  $E$  and difference current density  $I_{\Delta}$  are obtained for a given terminal current density  $I_{\Sigma}$ . The electric field is the sum of a component  $E_D(x)$ , related to the space charge of the impurities and a component

$$E_C(x) = (\epsilon v / I_{\Sigma}) z(x),$$

where  $z' = I_{\Delta} / I_{\Sigma}$ , related to the space charge of the mobile carriers.

(i) Select a point  $x_1$  in the avalanche region and assign a trial value  $E_1$  for the electric field at this point. Conveniently,  $x_1$  may be the location of the metallurgical junction, but the choice is not critical.

(ii) Choose a trial function for the difference current  $I_{\Delta}$ . A convenient, though crude, choice is

$$I_{\Delta} = I_{\Sigma}, \quad x \leq x_1 \quad (84)$$

$$I_{\Delta} = -I_{\Sigma}, \quad x > x_1. \quad (85)$$

(iii) Compute the electric field  $E$

$$E(x) = E_1 + \frac{1}{\epsilon} \int_{x_1}^x \left[ \frac{I_{\Delta}(s)}{v} + qN_D(s) \right] ds. \quad (86)$$

Extend the integral up to boundary points  $x_L$  and  $x_R$  to the left and right of  $x_1$ , i.e., to points where the field reaches the value required to carry the current through the not swept-out semiconductor region:

$$E(x_L) = I_{\Sigma} / [\mu_n | N_D(x_L) |] \quad (87)$$

$$E(x_R) = I_{\Sigma} / [\mu_p | N_D(x_R) |]. \quad (88)$$

Define

$$E_D(x) = E_1 + \frac{q}{\epsilon} \int_{x_1}^x N_D(s) ds \quad (89)$$

Use

$$z_0(x) = \int_{x_1}^x I_{\Delta}(s)/I_{\Sigma} ds \quad (90)$$

as trial function for  $z$ . The beginning of the next step, (iv), is redundant, but is required for the iteration loop.

(iv) Set  $E(x) = E_D(x) + (\epsilon v/I_{\Sigma})z_0(x)$ , and  $I_{\Delta} = I_{\Sigma}z_0'$ . Using these values, evaluate  $\alpha$ ,  $\beta$ , and  $H$ , (29).

(v) Solve

$$[D^2 + (\beta - \alpha)D - H]z = \alpha + \beta - Hz_0 \quad (91)$$

for  $z$  with boundary conditions  $z' = -1$  at left boundary and  $z = +1$  at right boundary.

(vi) Make  $z$  the new trial solution  $z_0$ .

(vii) Repeat steps (iv) through (vi) until  $z$  and  $z_0$  differ negligibly. If the final field at  $x_1$  differs appreciably from the trial field  $E_1$  the boundaries are incorrectly defined, and steps (i) and following should be repeated with the current value  $E(x_1)$  as trial field  $E_1$ . This time use  $I_{\Delta} = z'I_{\Sigma}$  instead of (84) and (85) as trial function for  $I$  in step (ii).

(viii) When the conditions of step (vii) are satisfied, i.e., when  $z \approx z_0$  and  $E(x_1) \approx E_1$ , then  $E$ ,  $I_{\Delta} = z'I_{\Sigma}$ ,  $\alpha$ ,  $\beta$ , and  $H$  are the desired self-consistent dc quantities.

#### APPENDIX B

In this appendix we show how the complex resonance frequency  $A = f_a + if_r$  at which the impedance goes to infinity can be obtained directly from (31), rather than by curve fitting of results obtained at a set of imaginary  $s$ -values  $s = 0 + if$ , or real frequencies. At resonance the normalized field  $z$  goes to infinity and therefore the right-hand side of (31) is negligible compared to the left-hand side, i.e., we have the homogeneous equation

$$[D^2 - k^2 + (\alpha + \beta)k + (\alpha - \beta)D - H]z = 0. \quad (92)$$

Solutions of (92) with boundary conditions (42) and (43) exist only for special values of  $k$ . The quantity  $k(k - \alpha - \beta)$  may be considered the eigenvalue of (92). Of the possible  $k$  values we select the complex conjugate pair of smallest absolute value. Let  $k_0$  be the  $k$  with positive imaginary part. Then

$$vk_0 = 2\pi(f_a + if_r) \quad (93)$$

is the desired resonance frequency.

Numerical values of  $k$  may be obtained as follows. The differential equation (92) is replaced by a set of appropriate difference equations for the values  $z_n$  of  $z$  at a set of meshpoints. This set of equations may be written in matrix form

$$M \cdot Z = 0, \quad (94)$$

where  $Z$  is a vector having the  $z_n$  for its elements. The matrix  $M$ , being derived from a second-order differential equation, is a "tridiagonal matrix", i.e., it has nonvanishing elements only on the main diagonal and the two adjacent diagonals. Now  $k$  must be chosen so that the determinant of  $M$  vanishes. If  $k_1$  and  $k_2$  are trial values and the corresponding values of the determinant are  $D_1$  and  $D_2$ , then one may choose as next trial value

$$k_3 = - \frac{D_2(k_2 - k_1)}{D_2 - D_1} \quad (95)$$

and thus iteratively approach the desired value of  $k_0$ .

The frequency  $f_c$  of the zero is chosen so that  $F$  has the proper phase in the neighborhood of the resonance frequency. Let  $F$  be the current fraction at the complex frequency  $s_1 = f_a + if_1$ , where  $f_1$  is near  $f_r$ , then

$$f_c = f_1 \frac{F_{\text{real}}}{F_{\text{imag}}} - f_a. \quad (96)$$

#### REFERENCES

1. Read, W. T., A Proposed High-Frequency Negative-Resistance Diode, B.S.T.J., 37, March, 1958, pp. 401-466.
2. Johnston, R. L., DeLoach, B. C., Jr., and Cohen, B. G., A Silicon Diode Microwave Oscillator, B.S.T.J., 44, February, 1965, pp. 369-372.
3. Lee, C. A., Batdorf, R. L., Wiegmann, W., and Kaminski, G., The Read Diode-An Avalanche Transit-Time, Negative Resistance Oscillator, Appl. Phys. Letters, March, 1965, pp. 89-91.
4. Misawa, T., private communication.
5. Josenhans, J. G., private communication.
6. Gilden, M. and Hines, M. E., Electronic Tuning Effects in the Read Microwave Avalanche Diode, IEEE Trans. on Electron Devices, January, 1966, pp. 169-175.
7. Josenhans, J. G. and Misawa, T., Experimental Characterization of a Negative Resistance Avalanche Diode, IEEE Trans. on Electron Devices, January, 1966, pp. 206-208.
8. Misawa, T., Negative Resistance in p-n Junctions Under Avalanche Breakdown Conditions, Parts I & II, IEEE Trans. on Electron Devices, January, 1966, pp. 137-151.
9. Scharfetter, D. L. and Gummel, H. K., Calculation of Large Signal Transients in Avalanche Diode Oscillators International Electron Devices Meeting, Washington, D. C., October 20-22, 1965, Paper 2.5.
10. Lee, C. A., Logan, R. A., Batdorf, R. L., Kleimack, J. J., and Wiegmann, W., Ionization Rates of Holes and Electrons in Silicon, Phys. Rev., 134, May, 1964, pp. A761-A773.



# Some Examples of Comparisons of Connecting Networks

By V. E. BENEŠ

(Manuscript received August 17, 1966)

*In the theory of telephone traffic it is of interest to compare the performance of connecting networks, as measured by the probability of blocking, when they are subjected to the same traffic sources. The question arises whether there are examples of pairs of networks, with the same number of cross-points, whose respective graphs of loss as a function of offered load cross each other. The existence of such examples would establish the principle that some network configurations are inherently more efficient at some traffic levels than at others, so that the "excellence" of a network is not necessarily a purely combinatorial notion independent of offered traffic. Examples of the above phenomenon are exhibited which do not involve only very small networks.*

## I. INTRODUCTION

In the theory of telephone traffic it is of interest to compare the performance of connecting networks, as measured by the probability of blocking, when they are subjected to the same traffic sources. Naturally, there are cases in which the result of this comparison is independent of the calling rate  $\lambda$ .<sup>1</sup> In this connection, H. O. Pollak has raised the question whether there are examples of pairs of networks, with the same number of crosspoints, the first of which is better than the second at one value of  $\lambda$ , while the second is better than the first at another value of  $\lambda$ .

The existence of such examples would establish the principle that some network configurations (in particular, some switch sizes) are inherently more efficient at some traffic levels than at others, so that the "excellence" of a network is not necessarily a purely combinatorial notion independent of offered traffic. We shall exhibit examples of the above phenomenon which are nontrivial in that they do not involve only very small networks.

## II. PRELIMINARIES

The notations and conventions of Refs. 2 and 3 will be used. We shall need machinery for studying the probability of blocking at very high values of the traffic  $\lambda$ ; this is provided by the natural expansion of the equilibrium state probabilities in inverse powers of  $\lambda$ :

*Lemma:* The state probabilities  $\{p_x, x \in S\}$  can be expanded in a power series

$$p_x = \sum_{m=0}^{\infty} d_m(x) \lambda^{-m} \quad (1)$$

valid for  $\lambda$  real and sufficiently large. With  $w = \max_{x \in S} |x|$ , the coefficients  $d_m(x)$  have the property

$$d_m(x) = 0 \quad \text{for} \quad 0 \leq m < w - |x|, \quad (2)$$

and the numbers  $d_{w-|x|}(x)$  satisfy

$$\sum_{|x|=w} d_0(x) = 1 \quad (3)$$

$$s(y) d_{w-|y|}(y) = \sum_{z \in A_y} d_{w-|z|}(z), \quad |y| < w,$$

$$d_{w-|x|}(x) \geq 0. \quad (4)$$

*Proof:*  $p_x(\lambda)$  is a rational function of  $\lambda$ , and so has an expansion of the form (1) if  $\lambda$  is large enough. Substitution of (1) into the equilibrium condition gives these equations for the coefficients  $d(\cdot)$ : (No unblocked call is rejected.)

$$|x| d_{m-1}(x) + s(x) d_m(x) = \sum_{y \in A_x} d_{m-1}(y) + \sum_{y \in B_x} d_m(y) r_{yx}.$$

It follows at once that if 0 = zero state (with no calls up), then  $d_0(0) = 0$ , and

$$s(x) d_0(x) = \sum_{y \in B_x} d_0(y) r_{yx},$$

so that  $d_0(x) = 0$  unless  $x$  is maximal in the natural partial ordering of states.

Thus, if  $x$  is not maximal then

$$d_m(x) = 0 \quad \text{for} \quad |x| < w - m$$

holds for  $m = 0$ . Assume that it holds for some  $m - 1 \geq 0$ . For  $x$  not maximal,  $s(x) > 0$  and

$$s(x)d_m(x) = -|x|d_{m-1}(x) + \sum_{y \in A_x} d_{m-1}(y) + \sum_{y \in B_x} d_m(y)r_{yx}.$$

If  $|x| < w - m$ , then  $d_{m-1}(x) = 0$  and  $y \in A_x$  implies  $d_{m-1}(y) = 0$ , both by the induction hypothesis. Thus,  $d_m(x)$  is expressible as a constant times  $d_m(0)$ . But

$$s(0)d_m(0) = \sum_{|y|=1} d_{m-1}(y) = 0,$$

by the induction hypothesis.

If  $x$  is maximal with  $|x| < w - m$ , then

$$|x|d_m(x) = \sum_{y \in B_x} d_{m+1}(y)r_{yx}, \quad m \geq 0.$$

But  $y \in B_x$  implies  $|y| = |x| - 1 < w - m - 1$ , and so  $d_{m+1}(y) = 0$ . (1) and (2) imply (3) and (4).

The formula

$$\sum_{|x|=w} d_0(x) = 1,$$

follows from

$$p_x = \sum_{m=w-|x|} d_m(x)\lambda^{-m}$$

and  $\sum_{x \in S} p_x = 1$  by letting  $\lambda \rightarrow \infty$ .

It follows from the lemma just proved that for sufficiently high values of the traffic parameter  $\lambda$ , the probability of blocking has the form

$$\Pr\{\text{bl}\} = \frac{\sum_{|x|=k} \beta_x d_{w-k}(x)}{\sum_{|x| \geq k} \alpha_x \sum_{j=w-|x|}^{w-k} d_j(x)\lambda^{w-k-j}} + o(1), \quad \lambda \rightarrow \infty,$$

where  $k$  is the greatest integer such that some states with  $k$  calls in progress have blocked calls ( $\beta_x > 0$ ). In particular, we see that

$$\lim_{\lambda \rightarrow \infty} \Pr\{\text{bl}\} = \begin{cases} 0 & \text{if } k < w \\ 1 & \text{if } k = w. \end{cases}$$

### III. COMPARISONS

The examples to be studied are the networks A and B in Figs. 1 and 2, respectively. Both are three-stage networks of the type due to C. Clos,<sup>4</sup> each with  $nr$  inlets (outlets). We show that there are values of

$m$ ,  $n$ , and  $r$  such that (i)  $A$  and  $B$  have very nearly the same number of crosspoints, and (ii)  $A$  has lower blocking than  $B$  at all sufficiently low values of the traffic  $\lambda$ , while  $B$  has lower blocking than  $A$  at all sufficiently high values of  $\lambda$ . The calculations forming this comparison will be carried out in the traffic model of Chapter 8 of Ref. 2; familiarity with this model is assumed.

In  $A$ , at least  $m$  calls must be in progress in order for there to be any blocking. Hence,<sup>5</sup>

$$\Pr\{\text{bl}\}_A = \kappa\lambda^m + o(\lambda^m), \quad \lambda \rightarrow 0.$$

In  $B$  there is a least integer  $k \geq 0$  such that

$$\Pr\{\text{bl}\}_B = c\lambda^k + o(\lambda^k), \quad \lambda \rightarrow 0$$

with  $c > 0$ . We shall show that  $k \leq r + 1$ , independently of the routing used to run  $B$ . It has been shown<sup>5</sup> that the probability  $p_x$  of a state in the model of Ref. 2 is of the form

$$p_x = p_0 \sum_{\pi} \lambda^{\frac{1}{2}[\ell(\pi) + |x|]} \prod_{y \in \pi} \frac{1}{|y| + \lambda s(y)}, \quad (5)$$

where the sum is over paths  $\pi$  on  $(S, \leq)$  permitted by the routing rule in use starting at 0 and ending at  $x$ , the product is along the path, and  $\ell(\pi)$  is the path-length.

In  $B$  it takes  $r$  calls in progress to block a call. Choose an outer switch on each side of  $B$  and consider a sequence of  $r$  attempted calls, each of which is from one of these switches to the other, together with one

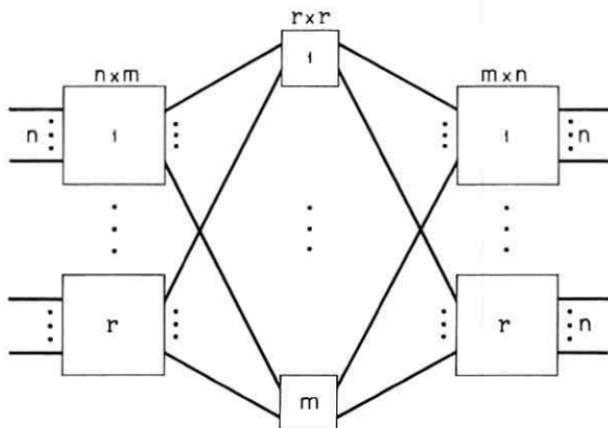
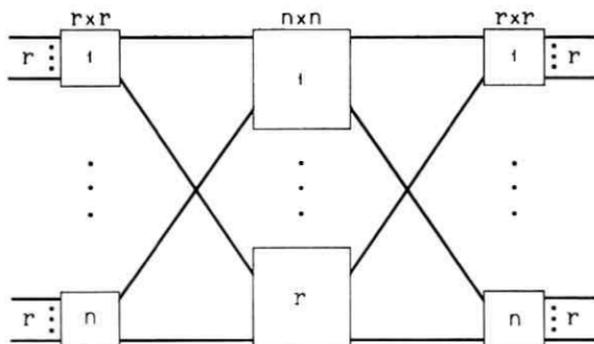


Fig. 1 — Network  $A$ :  $2mnr + mr^2$  crosspoints.

Fig. 2 — Network B:  $2nr^2 + nr^2$  crosspoints.

more call  $c$ . The last call,  $c$ , will have to go on one of the  $r$  middle switches each of which already has exactly one call. (Fig. 3, upper half.) If now the *other* call on the switch carrying  $c$  hangs up (Fig. 3, lower half), we will have reached a blocking state from 0 with positive probability along a path  $\pi$  of length  $l(\pi) = r + 2$ . Since the blocking state reached has  $r$  calls in progress, there is a contribution in formula (5) of the form

$$c\lambda^{r+1}, \quad c > 0.$$

It follows that if  $m > r + 1$ , then

$$\Pr\{\text{bl}\}_A < \Pr\{\text{bl}\}_B$$

for all  $\lambda$  sufficiently small.

Now take  $n > m$ , so that Lemma 1 gives

$$\begin{aligned} \Pr\{\text{bl}\}_A &= \frac{\sum_{|x|=nr} \beta_x d_0(x)}{\sum_{|x|=nr} \alpha_x d_0(x)} + o(1) \\ &= 1 + o(1), \quad \lambda \rightarrow \infty. \end{aligned}$$

At the same time, it can be seen that in network B,  $\beta_x = 0$  for  $|x| > nr - 2$ , so that

$$\Pr\{\text{bl}\}_B = \frac{\sum_{|x|=nr-2} \beta_x d_2(x)}{\sum_{|x| \geq nr-2} \alpha_x \sum_{j=nr-|x|}^2 d_j(x) \lambda^{2-j}} + o(1).$$

For  $|x| = nr$ ,

$$|x| d_0(x) = \sum_{y \in B_x} d_1(y) r_{yx}.$$

Thus,

$$\begin{aligned} nr &= \sum_{|x|=nr} |x| d_0(x) = \sum_{|x|=nr} \sum_{y \in S} d_1(y) r_{yx} = \sum_{|y|=nr-1} d_1(y) s(y) \\ &= \sum_{|y|=nr-1} d_1(y), \end{aligned}$$

and since  $\alpha_x = 0$  for  $|x| = nr$ , and  $\alpha_y = s(y) = 1$  for  $|y| = nr - 1$ .

$$\Pr\{\text{bl}\}_B = \frac{\sum_{|z|=nr-2} \beta_z d_2(z)}{\sum_{|x|=nr-2} \alpha_x d_2(x) + \lambda \sum_{|y|=nr-1} d_1(y)} + o(1).$$

The leading term is  $< 1$ , and so for all  $\lambda$  sufficiently large

$$\Pr\{\text{bl}\}_B < \Pr\{\text{bl}\}_A.$$

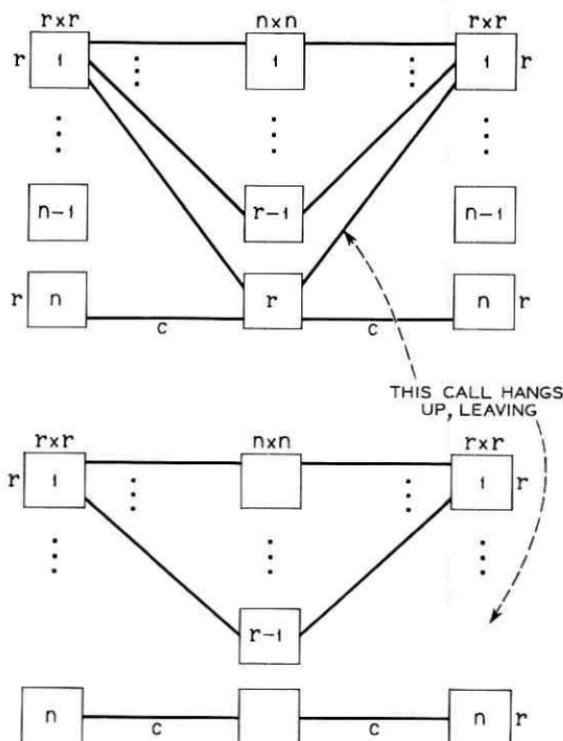


Fig. 3—Blocking state of A reached in  $r + 2$  steps (only links in use indicated).

It remains to show that there are values of  $m, n, r$  such that  $n > m > r + 1$  for which the number of crosspoints of  $A$  is very nearly equal to that of  $B$ . Picking  $m = r + 2$ , the condition for equality is that

$$n = 2 + (4 + 2r + r^2)^{\frac{1}{2}}.$$

With  $[t]$  the integer part of  $t$ , we pick  $n$  as

$$2 + [(4 + 2r + r^2)^{\frac{1}{2}}] > 2 + (1 + 2r + r^2)^{\frac{1}{2}} = r + 3 > m.$$

With this choice of  $n$   $A$  actually has more crosspoints than  $B$  and yet gives higher blocking at large values of  $\lambda$  than  $B$  does.

#### REFERENCES

1. Beneš, V. E., On Comparing Connecting Networks, to appear in J. Combinatorial Theor.
2. Beneš, V. E., *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, New York, 1965.
3. Beneš, V. E., Programming and Control Problems Arising from Optimal Routing in Telephone Networks, abstract in SIAM J. Control, 4, February, 1966, pp. 6-18; B.S.T.J., 45, November, 1966, pp. 1373-1438.
4. See Chapter 3 of Ref. 2.
5. See Chapter 8 of Ref. 2.



## Contributors to This Issue

VÁCLAV E. BENEŠ, A. B., 1950, Harvard College; M.A. and Ph.D., 1953, Princeton University; Bell Telephone Laboratories, 1953—. Mr. Beneš has been engaged in mathematical research on stochastic processes, traffic theory, and servomechanisms. In 1959–60 he was visiting lecturer in mathematics at Dartmouth College. He is the author of *General Stochastic Process in the Theory of Queues* (Addison-Wesley, 1963), and of *Mathematical Theory of Connecting Networks and Telephone Traffic* (Academic Press, 1965). Member, American Mathematical Society, Association for Symbolic Logic, Institute of Mathematical Statistics, SIAM, Mind Association, Phi Beta Kappa.

MARTIN B. BRILLIANT, B.A., 1955, Washington and Jefferson College; S.B., S.M., 1955, Sc.D., 1958, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1955 and 1966—. Mr. Brilliant has also held positions with the Air Force Cambridge Research Center; National Company, Inc.; Hazeltine Research Corporation; University of Kansas; and Booz·Allen Applied Research, Inc. At Bell Telephone Laboratories, he worked in 1955 on a transistor pulse generator for the Electronic Central Office. He is now concerned with systems engineering problems in the switching of time-multiplexed signals. Member, Sigma Xi, IEEE, AAAS.

ROBERT W. CHANG, B.S.E.E., 1955, National Taiwan University; M.S.E.E., 1960, North Carolina State College; Ph.D., 1965, Purdue University; Bell Telephone Laboratories, 1965—. Mr. Chang has been concerned with problems in data transmission and communication theory. Member, Eta Kappa Nu, Sigma Xi, Phi Kappa Phi, IEEE.

THOMAS G. CROSS, B.S.E.E., 1963, California State Polytechnic College; M.S.E.E., 1965, Northeastern University; Bell Telephone Laboratories, 1963—. Mr. Cross has been involved in a group responsible for the system analysis of the new TD-3 Long-Haul Radio System, and is currently involved in system planning for microwave radio relay systems.

A. GERSHO, B.S., 1960, Massachusetts Institute of Technology; M.S., 1961, Ph.D., 1963, Cornell University; Bell Telephone Laboratories, 1963—. Dr. Gersho is now Assistant Professor of Electrical Engineering at the City College of the City University of New York. He has been engaged in research in automatic and adaptive equalization for digital communications, synchronization of remote clocks, and synthesis of distributed networks.

HERMANN K. GUMMEL, Dipl. Phys., 1952, Philipps University (Marburg, Germany); M.S., 1952, and Ph.D., 1957, Syracuse University; Bell Telephone Laboratories, 1957—. Mr. Gummel's work has been in semiconductor electronics. He presently supervises a group responsible for analysis of device performance. Member, American Physical Society, Sigma Xi.

BARRY J. KARAFIN, B.S.E.E., 1961, University of Pennsylvania; M.E.E., 1963, New York University; Bell Telephone Laboratories, 1961—. Mr. Karafin has been concerned with timing problems in digital communication systems. He has also worked on computer languages for use in engineering simulation studies. Currently, he is concerned with problems of man-computer interaction in the development of transmission systems. Member, ACM, Simulation Councils, Tau Beta Pi.

MAURICE KARNAUGH, B.S., 1948, The City College of New York; M.S., 1950, and Ph.D., 1952, Yale University; Bell Telephone Laboratories, 1952-1966. Mr. Karnough has engaged in research on the synthesis of digital data systems, logic circuits, telephone switching networks, and pulse code modulation techniques. As Head, Systems Research Department, he has recently been concerned with special problems in analog-to-digital conversion, system simulation, and nonlinear signal processing. Member, IEEE, Sigma Xi, Phi Beta Kappa.

D. L. SCHARFETTER, B.S., 1960, M.S., 1961, and Ph.D., 1962, Carnegie Institute of Technology; Bell Telephone Laboratories, 1962—. Mr. Scharfetter has engaged in research on metal-semiconductor diodes, junction diodes and transistors, and microwave diode oscillators. Member, IEEE, Sigma Xi, Tau Beta Pi, Eta Kappa Nu, Pi Mu Epsilon.

ALLEN G. VARTABEDIAN, B.S.E. (Sci. Engr.) and B.S.E. (Math.), 1964, University of Michigan; M.S.E., 1965, University of Pennsylvania;

Bell Telephone Laboratories, 1964—. Mr. Vartabedian has been engaged in data switching studies in system engineering for the No. 1 ESS adapted for data features. Currently, he is involved in investigating input/output devices and algorithms used for communication between man and computing machines related to business information systems. Member, Tau Beta Pi.



## B.S.T.J. BRIEFS

### Display of Holograms in White Light

By C. B. BURCKHARDT

(Manuscript received September 7, 1966)

This paper describes a new method for displaying holograms in white light. The method gives reasonably good reconstructions although certain image defects are inherent in the method. It differs from previously reported methods of white light reconstruction<sup>1,2</sup> in that the whole spectrum is used for reconstruction and therefore black and white reconstructions can be obtained. The method does not depend on the volume properties of the photographic emulsion.

The basic arrangement is shown in Fig. 1. The white light illuminates a hologram which had been formed with a plane off-axis reference beam.<sup>3</sup> Behind the hologram there is a Venetian blind structure which blocks off the direct light but lets through the diffracted beam. The diffracted beam is diffracted a second time at a plane transmission grating which can be formed photographically with two plane beams. The angle between the two beams which form the plane grating has to be equal to the mean angle between the subject beam and reference beam used to form the hologram.

The reconstruction resulting from the configuration of Fig. 1 will now be explained. Intuitively, one can say that there is a large color dispersion at the first hologram because it can be considered a high spatial frequency diffraction grating. Since the light is diffracted in the opposite direction by the second grating this color dispersion is compensated. In order to be more quantitative, assume that during the formation of the hologram the subject beam  $A_s$  on the photographic plate is given by

$$A_s = a(x,y) \exp(j\omega_s x), \quad (1)$$

where  $\omega_s$  is the mean radian frequency of the subject beam and the center of the spatial frequency spectrum of  $a(x,y)$  is at zero. Assume that the reference beam  $R$  is given by

$$R = B \exp(-j\omega_r x), \quad (2)$$

where  $B$  is the amplitude and  $-\omega_r$  is the radian frequency of the reference beam. The photographic plate will record the intensity<sup>3</sup> and for the virtual image term we obtain

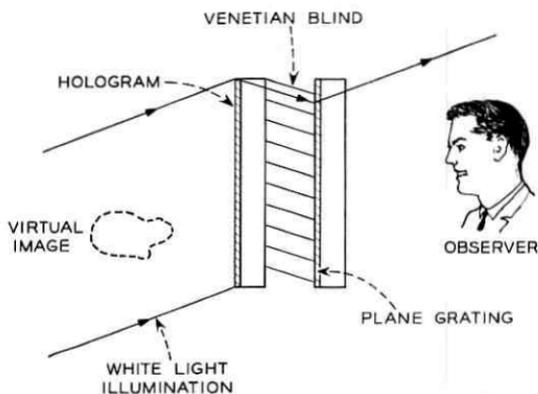


Fig. 1—White light display of a hologram.

$$A_s \cdot R^* = a(x,y)B \exp(j[\omega_s + \omega_r]x), \quad (3)$$

where the star means complex conjugate.

The plane grating is formed by two plane wave beams which can be expressed as  $G \exp(j\omega_s x)$  and  $G \exp(-j\omega_r x)$  and for the intensity  $\chi\chi^*$  on the plate we will obtain

$$\chi\chi^* = 2G^2 + G^2 \exp(j[\omega_s + \omega_r]x) + G^2 \exp(-j[\omega_s + \omega_r]x). \quad (4)$$

We now assume that the two plates are developed in such a way that their amplitude transmittance is proportional to the intensity during exposure. Since the hologram and the plane grating are spaced only a short distance behind one another, to get the amplitude transmittance corresponding to the virtual image we can multiply the amplitude transmittance of the hologram,  $A_s \cdot R^*$ , with that of the plane grating  $\chi\chi^*$ . We are now interested in the term which is given by the product of (3) with the last term of (4). This product,

$$a(x,y)BG^2, \quad (5)$$

is equal to the subject beam term, (1), translated to a center frequency of zero. This term, therefore, represents the reconstruction of a virtual image in the direction of the illuminating beam. It is possible to make a single hologram where the virtual image term has the form of (5). One chooses a subject beam as given in (1) and a reference beam  $B \exp(j\omega_s x)$ , i.e., the reference beam has the same mean direction as the subject beam. This is the on-axis hologram which has the well-known disadvantage that the direct beam, the real image, and the virtual image fall onto each other. The configuration of Fig. 1 does not have this disadvantage, with respect

to the virtual image; it is, however, in a sense "equivalent" to the on-axis hologram. Therefore, the following explanations will be in terms of the equivalent on-axis hologram.

In the on-axis hologram each object point forms its own on-axis Fresnel zone plate on the photographic plate. Upon illumination each zone plate forms a divergent spherical wavefront and therefore, a virtual image point. Since the focal length of the zone plate is inversely proportional to the wavelength of the illuminating light, the virtual image points for the different colors are staggered in depth. It is important to note that one does not perceive this difference in depth if one looks at the virtual image point through the center of its zone plate and if the eye has sufficient depth of field as is usually the case. For a particular point of observation, a region of image points will be approximately "on-axis" and this region will be in sharpest focus and have minimum color. As the eye is moved different regions of image points will come into sharp focus.

It is possible to use to advantage a spherically converging reference and illuminating beam of the same curvature. A little thought will show that by placing the eye at the focal point of the reference beam one can look at all the virtual image points through the centers of their respective zone plates. From this point one therefore sees an image which could be called "quasi-achromatic". (This can also be shown by using imaging formulas.<sup>4</sup> They show that for the point mentioned, the ratio between the eye-to-image-point distance and lateral image magnification is independent of wavelength.) If one moves the eye away from this point the image starts to blur.

Experimentally, it was found that the best reconstructions were obtained by using a slightly convergent reference beam and viewing the reconstruction from a point in front of the focus of the reference beam. This is probably so, because one then has more tolerance with respect to movement of the eye.

It is, of course, possible to place the plane grating in front of the hologram in the configuration of Fig. 1. Particularly bright reconstructions are obtained by using the first diffracted order of a blazed reflection grating for illuminating the hologram. The reconstruction can then be easily viewed against a background of ordinary room light.

Fig. 2 shows a photograph of the virtual image of a white light reconstruction. The lens used to form a convergent reference and illuminating beam has a focal length of 48 cm and was placed 18 cm in front of the hologram plate. The image-forming photographic lens was placed at the focal point of the reference lens where the image is quasi-

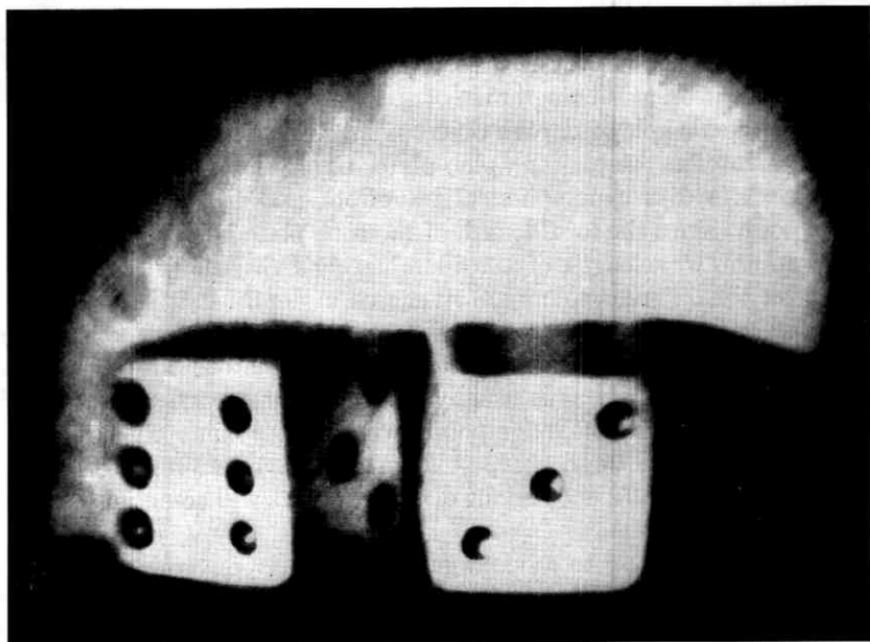


Fig. 2 — Photograph of the virtual image with the image-forming lens at the quasi-achromatic point.

achromatic. The photographic lens had an aperture of 6.5 mm and a focal length of 17 cm. The distance between subject and plate during the formation of the hologram was 18 cm and the angle between the subject beam and reference beam was  $22^\circ$ . The angle between the two plane beams used to form the photographic grating was also  $22^\circ$ . The Venetian blind structure is not visible in Fig. 2 because it was out of focus. A zirconium arc lamp was used as white light source.

The author would like to acknowledge the very competent experimental assistance of E. T. Doherty.

#### REFERENCES

1. Stroke, G. W. and Labeyrie, A. E., White-Light Reconstruction of Holographic Images Using the Lippman-Bragg Diffraction Effect, *Phys. Letters*, *20*, March, 1966, 368-370.
2. Lin, L. H., Pennington, K. S., Stroke, G. W., and Labeyrie, A. E., Multicolor Holographic Image Reconstruction with White-Light Illumination, *B.S.T.J.*, *45*, April, 1966, pp. 659-661.
3. Leith, E. N. and Upatnieks, J., Wavefront Reconstruction with Continuous-Tone Objects, *J. Opt. Soc. Amer.*, *53*, December, 1963, pp. 1377-1381.
4. Meier, R. W., Magnification and Third-Order Aberrations in Holography, *J. Opt. Soc. Amer.*, *55*, August, 1965, pp. 987-992.

## Approximation of the Error Probability in a Regenerative Repeater with Quantized Feedback

By **M. R. AARON** and **M. K. SIMON**

(Manuscript received September 21, 1966)

### I. INTRODUCTION

Recently, Zador<sup>1</sup> gave a clever functional iteration procedure for determining the error probability in a binary regenerative repeater with quantized feedback. Unfortunately, quantitative results for the long pulse sequences of interest are difficult to come by due to the prohibitive amount of computer time required to carry out the iterations. We have found a simple approximate procedure that breaks the computational bottleneck in all cases of practical interest. The crux of our approach is the approximation of the functional iteration by a difference equation. For clarity, we use only a few terms of a Taylor series in establishing the difference equation approximation. More terms can be used to obtain a better approximation if needed.

### II. RECAP OF ZADOR'S WORK

In Ref. 1, Zador shows that the  $k$ th iterate of the transformation  $Uf(x)$  — denoted by  $U^k f(x) = U^{k-1}[Uf(x)]$  — when evaluated at  $x = 0$  yields the average bit error probability  $p(k)$ , for the last bit in a random sequence of  $k + 1$  bits processed by a regenerative repeater with quantized feedback. The transformation  $Uf(x)$  is given by [Zador's (14)]

$$Uf(x) = p_1(x)f(rx - a) + p_2(x)f(rx) + p_3(x)f(rx + a), \quad (1)$$

where

$$\begin{aligned} p_1(x) &= p[N(-g_0 - x)] \\ p_2(x) &= 1 - p_1(x) - p_3(x) \\ p_3(x) &= q[1 - N(g_0 - x)] \\ f(x) &= p_1(x) + p_3(x) \end{aligned} \quad (2)$$

$$N(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp(-y^2/2\sigma^2) dy$$

$$0 < r < 1$$

$$a = -2g_1.$$

In (2) above,  $r = \exp(-T/T_c) = \exp(-b)$  is the decrement of the

simple RC low-pass filter in the feedback path of the regenerative repeater;  $T_c$  is the time constant of this filter or equivalently the reciprocal of the 3-dB point on the high-pass filter in the transmission line preceding the repeater. The quantity  $g_0$  is the peak value of pulse response of the medium in response to the transmission of a +1, and  $g_1 < 0$  is the value of the isolated pulse response of the medium one time slot ( $T$ ) away from the peak. For convenience, we assume equally likely transmitted pulses,  $p = q = \frac{1}{2}$ .

### III. APPROXIMATION

If we (i) substitute for  $p_2(x)$  from (2) into (1), and (ii) expand  $f(rx \pm a)$  in a Taylor's Series about  $rx$ , and (iii) retain terms in the series through  $a^2$  we get the approximation

$$Uf(x) = f(rx) + a \left. \frac{df(y)}{dy} \right|_{y=rx} [p_3(x) - p_1(x)] + \frac{a^2}{2} \left. \frac{d^2f(y)}{dy^2} \right|_{y=rx} [p_3(x) + p_1(x)]. \quad (3)$$

When we note that

$$\left. \frac{d^n f(y)}{dy^n} \right|_{y=0} = 0 \quad \text{for } n \text{ odd} \\ p_3(0) = p_1(0) \quad (4)$$

$$f(0) = p_3(0) + p_1(0) = p(0) = \frac{1}{2} \left[ 1 - \operatorname{erfc} \left( \frac{g_0}{\sigma\sqrt{2}} \right) \right],$$

then

$$p(1) = p(0) \left[ 1 + \frac{a^2}{2} \left. \frac{d^2f(y)}{dy^2} \right|_{y=0} \right]. \quad (5)$$

Proceeding through the functional iteration defined by (3), retaining terms involving  $a^2$  and lower, and using (4) we obtain the difference equation

$$p(k) = p(k-1) + p(0)KZ^{(2k-1)}, \quad (6)$$

where  $Z = g_0/\sigma$  is the peak signal to rms noise ratio for an isolated pulse, and

$$K = \frac{Z}{2\sqrt{2\pi}} \left( \frac{a}{\sigma} \right)^2 \exp \left( -\frac{1}{2}Z^2 \right). \quad (7)$$

Equation (6) is easily solved to give

$$p(k) = p(0) \left[ 1 + K \sum_{j=1}^k r^{2(j-1)} \right]. \quad (8)$$

Then

$$\lim_{k \rightarrow \infty} p(k) = p(0) \left[ 1 + \frac{K}{1 - r^2} \right]. \quad (9)$$

The bracketed term in (8) or (9) gives the enhancement of error probability due to errors made on all previous pulses.

To compare results obtained by application of (9) with those obtained experimentally by R. D. Howson,<sup>2</sup> we take  $b = \frac{1}{4}$ ,  $g_0 = \exp(-\frac{1}{8})$ , and  $g_1 = -g_0(1 - \exp(-\frac{1}{4}))$ . Over a wide range of S/N ratios of interest, analysis based upon (9) predicts about a 0.6-dB S/N penalty of this system over the ideal case of no low frequency cutoff. Agreement with experiment is excellent. It should be noted that the enhancement term in (9) is very close to unity and the S/N penalty is due essentially to the reduction of the pulse peak by the low-frequency cut-off.

In a future paper we will show (i) how Zador's approach can be extended to a wider class of systems, and (ii) how the approximation given herein can be used and improved when necessary, to arrive at meaningful quantitative results.

#### REFERENCES

1. Zador, P. L., Error Probabilities in Data System Pulse Regenerator with DC Restoration, B.S.T.J., 45, July, 1966, pp. 979-984.
2. Flavin, M. A., Gerrish, A. M., Howson, R. D., and Salz, J., unpublished work.

## Application of Automatic Transversal Filters to the Problem of Echo Suppression

By F. K. BECKER and H. R. RUDIN

(Manuscript received October 6, 1966)

Long-haul voice communication has long been subject to the problem of returned echo. The advent of synchronous satellite communication introduces increased delay as a degrading factor in the overall quality of two-way conversation. This compounds the problem in that the echo

Equation (6) is easily solved to give

$$p(k) = p(0) \left[ 1 + K \sum_{j=1}^k r^{2(j-1)} \right]. \quad (8)$$

Then

$$\lim_{k \rightarrow \infty} p(k) = p(0) \left[ 1 + \frac{K}{1 - r^2} \right]. \quad (9)$$

The bracketed term in (8) or (9) gives the enhancement of error probability due to errors made on all previous pulses.

To compare results obtained by application of (9) with those obtained experimentally by R. D. Howson,<sup>2</sup> we take  $b = \frac{1}{4}$ ,  $g_0 = \exp(-\frac{1}{8})$ , and  $g_1 = -g_0(1 - \exp(-\frac{1}{4}))$ . Over a wide range of S/N ratios of interest, analysis based upon (9) predicts about a 0.6-dB S/N penalty of this system over the ideal case of no low frequency cutoff. Agreement with experiment is excellent. It should be noted that the enhancement term in (9) is very close to unity and the S/N penalty is due essentially to the reduction of the pulse peak by the low-frequency cut-off.

In a future paper we will show (i) how Zador's approach can be extended to a wider class of systems, and (ii) how the approximation given herein can be used and improved when necessary, to arrive at meaningful quantitative results.

#### REFERENCES

1. Zador, P. L., Error Probabilities in Data System Pulse Regenerator with DC Restoration, B.S.T.J., 45, July, 1966, pp. 979-984.
2. Flavin, M. A., Gerrish, A. M., Howson, R. D., and Salz, J., unpublished work.

## Application of Automatic Transversal Filters to the Problem of Echo Suppression

By F. K. BECKER and H. R. RUDIN

(Manuscript received October 6, 1966)

Long-haul voice communication has long been subject to the problem of returned echo. The advent of synchronous satellite communication introduces increased delay as a degrading factor in the overall quality of two-way conversation. This compounds the problem in that the echo

must be further attenuated if its annoyance level is not to be increased as the delay of the returned echo increases.

At the present, the problem of returned echo is alleviated by the insertion of attenuation into the path of the weaker signal, i.e., the echo path.<sup>1</sup> Given the larger delay inherent in synchronous satellite communication, a better technique is wanted. One attractive scheme is the use of a transversal filter to synthesize a replica of the echo, which is then subtracted from the actual echo so that the two signals cancel. An algorithm which allows this synthesis to be carried out automatically was discovered by B. F. Logan and the late J. L. Kelly, Jr.

A scheme for simulating echoes and the technique for suppressing them are shown in block form in Fig. 1. The person using the handset at the left of the figure would experience echo; no echo is simulated for the person using the handset on the right. The echo, which would normally be caused by an improperly balanced hybrid, was instead simulated using various linear networks. A tape recorder simulated the long delay re-

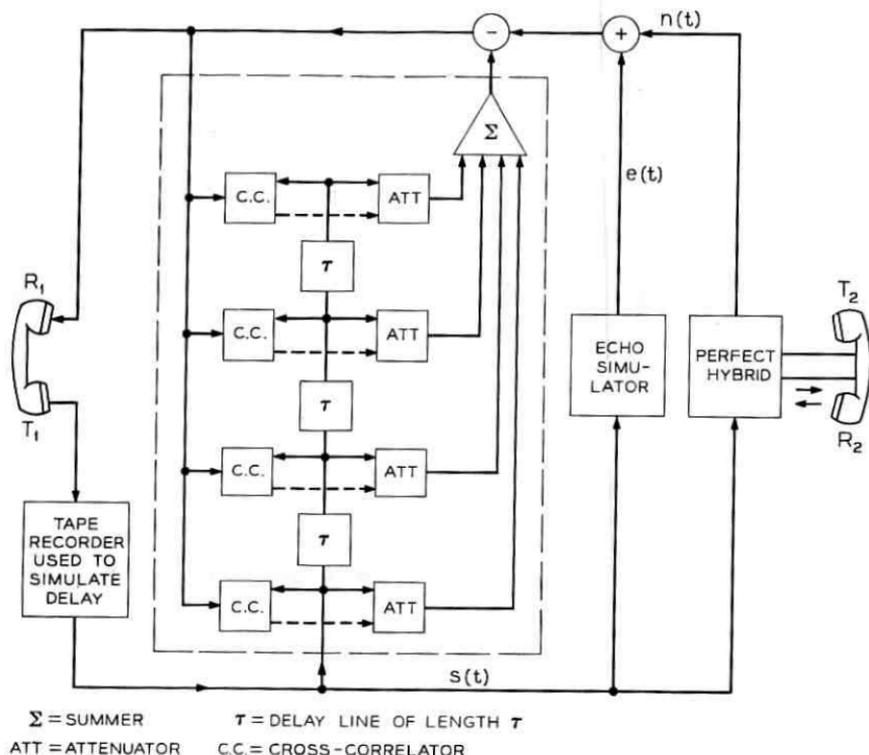


Fig. 1—System for demonstration of automatic transversal echo suppressor.

sulting from satellite transmission. The automatic transversal filter is enclosed in dashed lines; its input is the delayed speech from  $T_1$  and its function is the minimization of the echo delivered to  $R_1$ . The signal from  $T_2$  merely appears as noise,  $n(t)$ , in this minimization procedure.

Mathematically, the transversal filter's function is the minimization of the mean-squared value of the echo delivered to  $R_1$ . This is achieved through the approximation of  $e(t)$  by the sum of the weighted, delayed versions of  $s(t)$ . Thus, the canceler strives to minimize

$$I = \int_{-\infty}^{\infty} \left[ e(t) - \sum_{j=0}^N c_j s(t - j\tau) \right]^2 dt, \quad (1)$$

where  $c_j$  is the weighting (attenuator setting) associated with the  $j$ th tap and  $\tau$  is the tap spacing, usually the reciprocal of twice the highest frequency of interest.

The attenuator settings (which may be positive or negative) can be calculated by partial differentiation of  $I$  by the various  $c_j$ 's. Specifically,

$$\frac{\partial I}{\partial c_k} = -2 \int_{-\infty}^{\infty} \left[ e(t) - \sum_{j=0}^N c_j s(t - j\tau) \right] [s(t - k\tau)] dt. \quad (2)$$

Note that the second equation states that the partial derivative of  $I$  with respect to the tap gain  $c_k$  is given by the cross-correlation of the signal at the  $k$ th tap on the delay line with the signal delivered to  $R_1$ . The optimum settings for all the tap weighting coefficients occur when all the partial derivatives are zero.

Assuming a reasonable spectrum for the signal  $s(t)$ , it can be shown that the integral  $I$  is a convex function of the tap gains. Given this fact, the information contained in the various partial derivatives is sufficient to point the way toward the unique minimum of  $I$ .

An experimental implementation of the echo canceler is built around a general-purpose automatic equalizer intended for the reduction of linear distortion in communication channels.<sup>2</sup> The attenuators are digitally-controlled, resistive ladder-networks. In the implementation, the information obtained by cross-correlation is used to increase or decrease the attenuator setting by a constant increment. The cross-correlation coefficients are then recalculated and the attenuators again changed. The attenuators are permitted to change their setting only when the  $s(t)$  signal exceeds a threshold. The attenuators have infinite memory and so retain their setting during long lapses in the speech originating at  $T_1$ . A simple RC low-pass filter provides a sufficient approximation to the integration indicated in the second equation.

The signal originating at  $T_2$ , despite the fact that it may well be several times the size of the echo,  $e(t)$ , produces only small perturbations in the attenuator settings. This is a result of the powerful cross-correlation detection used to set the attenuators. In the experimental implementation described, a second speaker at  $T_2$  did not perceptibly degrade echo suppression.

Another feature of the system is that it is inherently adaptive. If the characteristics of the transmission channel should change, the scheme automatically makes the necessary modifications as long as the speaker level is above the threshold. This is true provided that the change in the channel characteristics occurs at a very slow rate.

Early results from this implementation indicate that echo suppression of some 20 dB is attainable. Further evaluation is necessary to accurately predict behavior on real channels. The settling time for this experimental echo canceller was in the order of two seconds. The settling time is dependent on the echo-to-interfering-noise ratio.

An accompanying brief by A. J. Presti and M. M. Sondhi describes a different implementation of an echo canceller based upon similar principles.

There are a number of engineering problems which must be solved before adaptive echo cancellation becomes a practical reality. One of these is the long, distortion-free delay which the echo canceler must supply. The magnitude of this delay depends on the echo delay. Another problem arises when there are several reflection points (hybrids) in the echo path. A third problem is that the apparatus tracks a change in echo path only if the change is slow and the signal threshold is exceeded. Given the solution to these problems, however, the future of this technique of echo cancellation is a promising one.

#### REFERENCES

1. Holman, E. W. and Suhocki, V. P., A New Echo Suppressor, Bell Laboratories Record, April, 1966, pp. 139-142.
2. Lucky, R. W. and Rudin, H. R., Generalized Automatic Equalization for Communication Channels, Proc. IEEE, March, 1966, pp. 439-440.

## A Self-Adaptive Echo Canceller

By **M. M. SONDHI** and **A. J. PRESTI**

(Manuscript received October 14, 1966)

Conventional echo suppressors combat echoes generated at hybrid junctions in long distance telephone connections by interrupting the return path according to some decision scheme based upon the relative levels of the outgoing and return signals. In this brief, a new device is described for *cancelling* the echo without interrupting the return path. We call this device an echo canceller to distinguish it from conventional echo suppressors. It generates a replica of the echo (which is then subtracted from the return signal) by synthesizing a linear approximation to the echo transmission path. It is self-adapting in that it automatically tracks variations in the echo path which may arise during a telephone conversation (e.g., connection or disconnection of extension phones, etc.).

A schematic of such a self-adapting echo canceller is shown in Fig. 1. It is based upon an idea originally proposed by J. L. Kelly, Jr. and B. F. Logan, and incorporates modifications which simplify and improve the implementation and performance.

With reference to Fig. 1, let  $x(t)$  be the input speech signal and  $y(t)$  the return signal. The return signal consists of an echo  $z(t)$  (which is the result of convolving  $x(t)$  with the impulse response  $h(t)$  of the echo path) and a noise  $n(t)$  (which may include a second speech signal). An estimate  $\hat{z}(t)$  of  $z(t)$  is subtracted from the return signal and the error signal  $e(t) = y(t) - \hat{z}(t)$  is continuously used to improve this estimate. The signal  $\hat{z}(t)$  is given by a linear expansion with time varying coefficients,  $g_i(t)$ . Thus,

$$\hat{z}(t) = \sum_{i=1}^N g_i(t)x_i(t), \quad (1)$$

where

$$x_i(t) = \int_0^t w_i(\tau)x(t - \tau) d\tau, \quad (2)$$

and the  $w_i(\tau)$  ( $i = 1, 2, \dots$ ) form a complete set of orthogonal functions. The dynamic behavior of the system of Fig. 1 is then governed by the set of equations

$$\frac{d}{dt} g_i(t) = KF[e(t)]x_i(t) \quad i = 1, \dots, N. \quad (3)$$

Here  $K$  is a positive constant and the function  $F[e]$  is chosen to be an odd

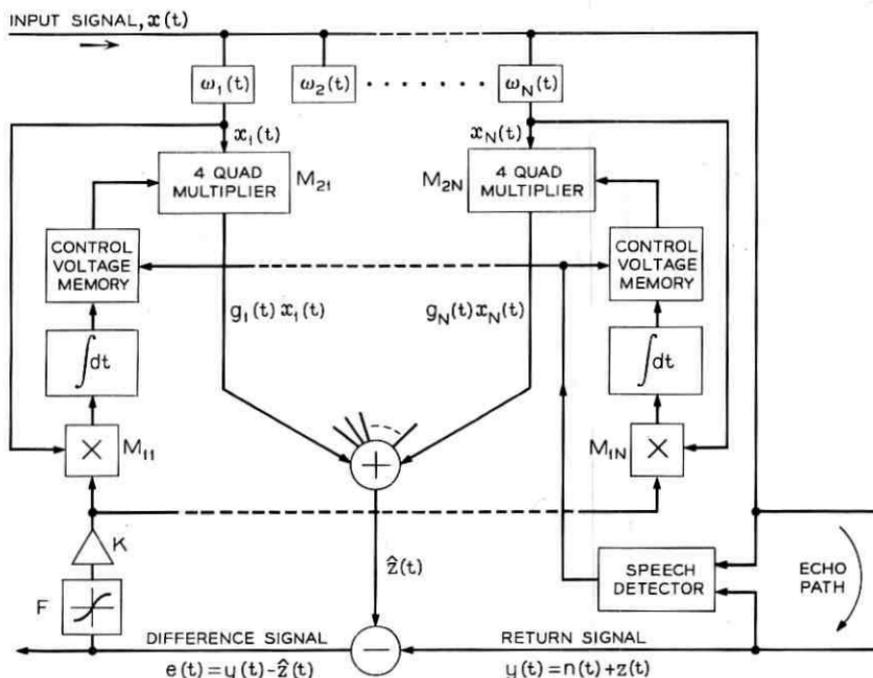


Fig. 1—Schematic of echo canceller.

and non-decreasing function of the error  $e$ . It can then be shown that if  $h(t)$  can be well approximated by the expansion

$$h(t) \cong \sum_1^N h_i w_i(t), \quad (4)$$

where the  $h_i$  are constant or slowly varying, then  $g_i(t)$  converges to  $h_i$  ( $i = 1, \dots, N$ ) in the limit as  $t \rightarrow \infty$ . That is,  $\hat{z}(t)$  approaches  $z(t)$ . The quantity

$$\sum_{i=1}^N (g_i(t) - h_i)^2$$

converges monotonically to zero if  $n(t) \equiv 0$  and if the expansion given by (4) is exact. However, convergence takes place even in the presence of relatively large amounts of noise. The amount of noise that can be tolerated decreases as the speed of convergence increases. The speed of convergence depends upon the constant  $K$ , the choice of function  $F$ , and the level and statistical properties of the signal  $x(t)$ . The proof of convergence, estimates of convergence rate, factors affecting choice of  $F$ ,

and the set  $w_i(t)$  as well as results of computer simulations of such echo-cancellers are the subject of a paper under preparation.

A prototype of such a system has been implemented. In this implementation the  $x_i(t)$  are obtained from taps on a delay line so that  $\omega_i(t) = \delta(t - iT)$ , where  $T$  is approximately 0.1 msec and  $N$  in (4) is 50. The function  $F$  has been chosen to be an infinite clipper. This allows the multipliers  $M_{1i}$  in Fig. 1 to be replaced by simple transistor switches. The integrators also pose no particular problem and simple operational amplifiers were found satisfactory. However, the multipliers  $M_{2i}$  must satisfy more stringent requirements. Their outputs  $g_i(t)x_i(t)$  must be strictly proportional to the  $x_i(t)$ , while strict proportionality to the  $g_i(t)$  is of secondary importance. A four quadrant multiplier was designed using current-controlled photo-resistors in a feedback circuit. This circuit exhibits a nonlinearity of less than 10 percent with respect to the  $g_i(t)$  and less than 0.5 percent with respect to the  $x_i(t)$  over a 50-dB dynamic range.

During time intervals when the input  $x(t)$  is zero, there is no corrective feedback and it is important that the gain settings  $g_i(t)$  be unaffected by drift in the integrators. It is also desirable (though not absolutely essential) that the feedback path be opened when the noise  $n(t)$  (which, as previously noted, may include a second speech signal) is considerably larger than the echo. The box designated as speech detector in Fig. 1 achieves this dual objective. It disconnects the operational amplifiers from the integrating capacitors whenever

$$\langle |x(t)| \rangle - \langle |y(t)| \rangle < \varepsilon$$

where  $\varepsilon$  is a predetermined positive threshold and  $\langle \rangle$  indicates time averaging for about 0.5 sec.

Further details of the implementation will be described in a forthcoming publication. The results of computer simulations and tests on the prototype may be summarized as follows:

(i) The system converges in about 0.5 second for normal speech levels. This convergence time increases to about 5 seconds for a speech signal 20 dB lower.

(ii) When the echo canceller and the echo path were simulated on the computer, cancellations of 60 dB and higher were easily achieved. However, in the case of echoes generated on laboratory simulated telephone connections, the cancellation was about 20–25 dB on both the computer simulation and the prototype.

(iii) The system converges in the presence of noise which is up to 8 or

10 dB higher than the echo. After the system has converged, however, even a much larger noise does not appreciably degrade the cancellation.

The implementation must be extensively tested in a variety of telephone circuits before the merits of the proposed system can be fully evaluated.

An accompanying brief by F. K. Becker and H. R. Rudin describes a different instrumentation of an echo canceller based upon similar principles.