

THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XL

SEPTEMBER 1961

NUMBER 5

Copyright 1961, American Telephone and Telegraph Company

A 150-mc Personal Radio Signaling System

By DOREN MITCHELL and K. G. VAN WYNEN

(Manuscript received February 23, 1961)

An experimental 150-mc personal signaling system has been set up in New York City to evaluate over-all technical performance and explore subscriber reactions to the system. The system includes pocket receivers equipped with tuned reeds and central office arrangements adapted for direct customer dialing. This paper describes the system in over-all terms, and tells how it was engineered. It also compares this system with the 35-mc systems and discusses traffic and radio transmission problems.

I. INTRODUCTION

An experimental trial of a 150-mc personal radio signaling system known as the Bellboy system was started in New York City in October 1960. The service given to customers is very similar to that provided by 35-mc telephone company systems now operating in several smaller cities.^{1,2,3} The new system, however, has a much greater traffic capacity, uses receivers of greater selectivity and sensitivity, and includes direct customer dialing.

Basically, this system extends the telephone bell to the customer's pocket. Each customer carries a pocket radio receiver containing reeds tuned to specific voice frequencies. This allows him to be selectively signaled. Three transmitters are installed in the lower part of Manhattan to give coverage to practically all locations below 59th Street. Central office arrangements are provided that accept calls that are dialed in and, in response, cause the specific signal of the customer who is being called to be radiated three times on the air.

When this signal reaches the customer's set, it causes the set to emit an alerting signal. The customer now knows that his office has received an important call and he goes to the nearest telephone and calls back to get his message.

The systems now in service at 35 mc give, in general, this same kind of service. As will be pointed out later in this paper, however, the limitations of the 35-mc system, which include receivers of lower selectivity and sensitivity and slower signaling speed, severely limit its usage in big cities. In addition, there is a more fundamental problem arising from the limited number of frequency allocations available at 35 mc, preventing growth of 35-mc systems to more than a few hundred customers in many small cities. These were the main reasons for proceeding with the trial of the 150-mc system. Some comparisons of the environments of the two systems are given in Section III.

An interesting general feature of this experimental trial is that some of the receivers were provided with replaceable batteries and some with rechargeable batteries. The latter type of receiver requires that the battery be recharged every night in a desk-top unit provided for that purpose, but this removes the necessity of supplying a fresh battery to the customer every few weeks.

In the early planning of this project it was apparent that there were two major technical problems. One was the need for a highly sensitive pocket receiver with a code-responsive device. The other was the need for more accurate knowledge of radio transmission loss into buildings.

A very sensitive, stable, and selective receiver has been developed by Bell Telephone Laboratories to operate in the 150-mc range. This receiver will be described in more detail in a forthcoming paper, but some of its characteristics are covered later in this paper. A receiver capable of giving service at 35 mc has also been developed by an outside supplier.⁴

In order to obtain knowledge of radio transmission loss into buildings an extensive test program was set up early in this project. These measurements have been covered in an earlier paper by Rice.⁵

II. OVER-ALL SYSTEM PLAN

Operation of the Bellboy system can be understood by tracing the path of a call as it progresses through the various parts of the system that are shown in Fig. 1. When there is an urgent need to reach a customer who is out of the office, his secretary dials his number, which might be, for example, 225-4654, and listens. She hears a momentary ringback, followed by a warbling tone, which signifies that the number has been received, checked, and stored in the central office equipment.

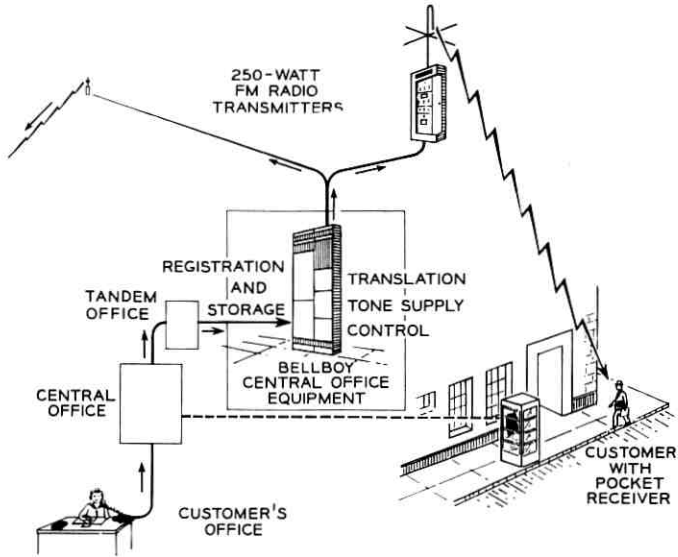


Fig. 1 — Over-all plan of operation.

She may now hang up, but if instead she listens a few seconds longer she will hear a verbal statement that her Bellboy system number will be signaled. Within a period ranging from a few seconds to about one minute, the signal is put on the air and the customer's set emits a steady beep. The customer now depresses a button to stop the beep, goes to the nearest telephone (in this case, a pay station), and calls his office to get the message.

2.1 Central Office Arrangements

What actually happened right after the secretary dialed was typical of many dialed telephone calls. The first three digits of the number, 225, correspond to a central office designation and normally would represent the one which serves the called party with incoming service. In the Bellboy system the counterpart of the terminating central office is actually the J-1 terminal which couples the wire plant to the radio portion of the system. All originating offices are wired to recognize the code 225 and set up a proper route through a tandem office that has the ability to accept digits in any form. The tandem office then forwards the digits as dial pulses, since the present J-1 is arranged to accept only this type of signal. Fig. 2 illustrates the functions included in the J-1 terminal.

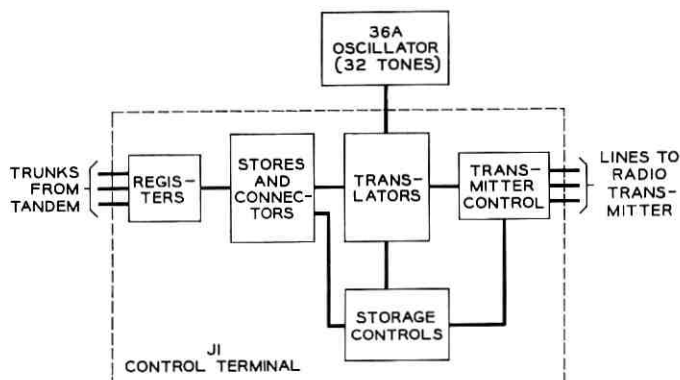


Fig. 2 — Central office arrangements.

When the last four digits, 4654, are spilled into the trunk they are received by dial-pulse registers. A number-checking circuit then verifies that the number is actually within the number blocks containing the valid codes. A connecting circuit next transfers the four digits from the register to a reed switch store on a two-out-of-five basis.

This number, along with those due to additional calls, stays in storage until called for by the storage control circuit, which occurs when the radio transmitters are available. At this time the translator uses the information in the stores to connect three of the 32 tone generators in the 36A oscillator to the transmitter control circuit. There the tones are amplified and fed to the three radio transmitters in the trial system. All transmitters are turned on, and the presence of the modulating tones and RF carrier is verified by "sniffer" circuits at each one. A timing circuit then maintains the modulation for about 0.5 second, and then inserts a pause of 0.5 second before the next signal. Each individual code signal is sent three times to guard against dead spots.

In principle there are 4960 possible codes when three are chosen at a time from a universe of 32. In this system the translation from the four-digit decimal number assigned to each receiver to the three tones chosen from 32 is made more simple and economical by using only 3200 of the possible codes. More codes could be provided by introducing a more complicated translator or by increasing the number of tones.

Bellboy system numbers are served on the basis of groups that arrive during time intervals which are slightly over one minute long. This arrangement permits an economical design, but results in minor delays, which can range from one second to somewhat over a minute, with an

average value of about one-half minute. A four-digit number spilled into the terminal is transmitted immediately if there is no other number in storage. That number retains exclusive use of the radio transmitters until it has been radiated the third time about 60 seconds later. Meanwhile, arriving numbers must wait for service. Then they are radiated in sequence, each taking about one second to process. After the third transmission, each number is cleared from storage and, when the last to arrive has been cleared, the next waiting group is served in similar fashion.

Fig. 3 is a time diagram illustrating the foregoing for a moderately loaded system. It is assumed that at time $t = 0$ the control circuit is just starting to serve calls 1, 2, and 3 in group A. The third transmission will be completed 63 seconds later. During that interval calls 4 through 7 arrive in storage and wait there for intervals ranging from 62 to 6 seconds. At 64 seconds after $t = 0$ call 4 is served. The figure also shows the arrival and service of group C. It will be clear that the maximum time interval between the arrival of a call in storage and the completion of the third transmission is $(120 + n)$ seconds, where n is number of calls in the immediately preceding group.

Each code that is transmitted comprises three audio tones, each being applied at a voltage that produces a carrier deviation of 1.3 kc. The

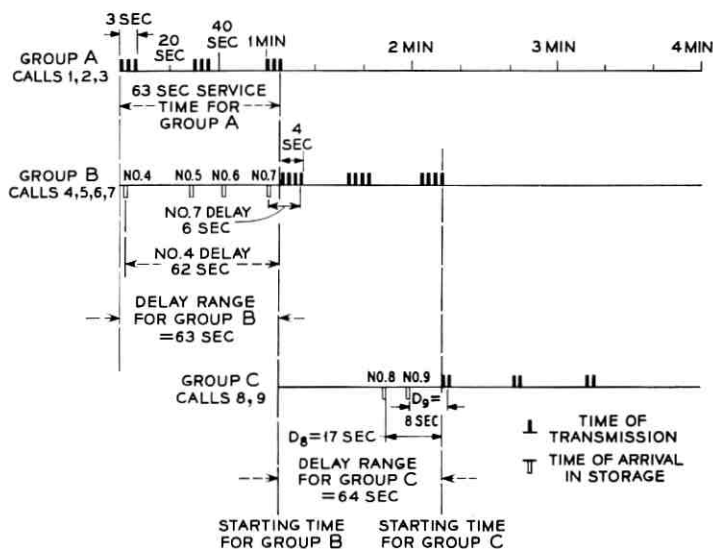


Fig. 3 — Time diagram illustrating Bellboy system calls (not to exact scale).

tone sources are crystal-controlled transistorized square-wave generators. They produce odd multiples of 7.5 cps, beginning at 517.5 cps and ending at 997.5. These square waves are then filtered to produce sine waves. Two frequency groups, a low and a high, are formed by omitting the center frequency, 757.5 cps. In the present system each code consists of two frequencies from one group and one frequency from the other. It will be clear that second-order products generated in the overall system fall outside the present range of tones and would always fall between tones if the range were extended by adding more odd harmonics of 7.5 cps. Although some third-order products can fall on operating frequencies, they will be too low in level to introduce errors.

2.2 Pocket Receiver

The pocket receiver will be covered in some detail in a forthcoming paper, but a brief general description of it will be of interest. It is a highly sensitive and selective radio receiver using ten transistors and three vibrating reeds. Two views of a receiver are shown in Fig. 4. Reeds are available for each of 32 frequencies produced by the 32-tone generator, and any three may be inserted into any receiver.

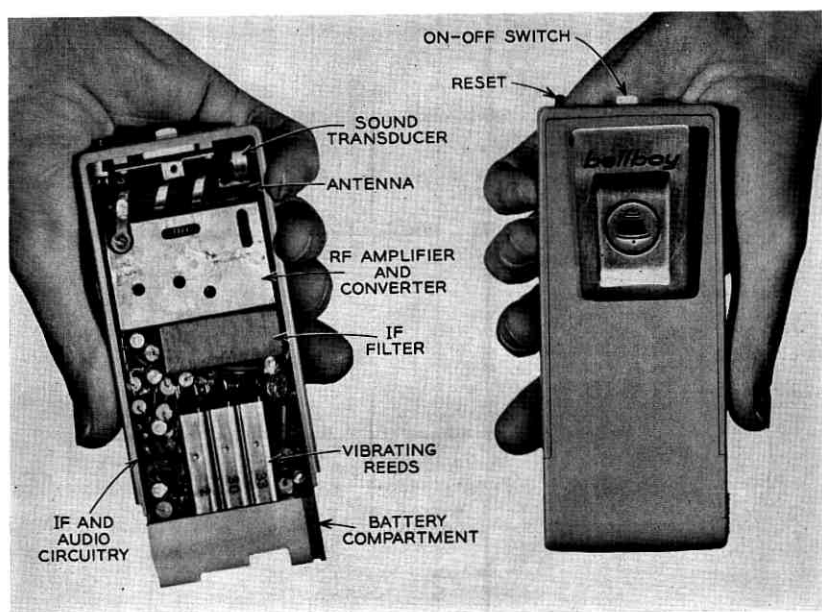


Fig. 4 — 150-mc Bellboy pocket receiver.

The receiver proper consists of two stages of RF amplification at 150 mc, an oscillator with diode mixer, and three stages of IF. Six transistors are used for these functions, another is used for the limiter, and two for the amplifier to drive the reeds. A final transistor provides audio signaling tone to a small transducer which produces the beep. An interesting feature of the receiver is that it uses only one frequency conversion with the IF at about 5.5 kc. This feature has allowed an economical receiver design and provides very satisfactory operation for reed signaling.

A subsequent paper will describe the reeds in more detail. Each is a highly stable tuning-fork structure with a bandwidth of one cycle, similar to those used in the trial of reed signaling in mobile radio at Richmond, Virginia, some years ago.^{6,7} The new miniature reeds,⁸ however, occupy only about one-seventh the space, have about one-sixth the weight, and require one-fifth the power of those used at Richmond.

2.3 *Integrated Operation*

The operation of the terminal described above assumed that the radio channel was assigned solely to signaling, as was the case for the trial. The terminal, however, has been designed so that it can operate on a channel also used for mobile telephone service. In this "integration," use of the radio channel is alternated between the two services, with mobile telephone service having priority. Bellboy service numbers which have been transmitted once may be required to wait more than the normal interval if a telephone call originates before the second transmission. This can result in intervals of variable duration between first and second or second and third transmission. As this period of time increases, the customer might become confused as to whether he was receiving a repetition or a new call. For this reason, a timing feature is provided in the control circuit to clear from storage any calls which have been served whenever this interval becomes excessively large.

III. COMPARISON OF 35-MC AND 150-MC ENVIRONMENT

It is well known that long-range "skip" interference has necessitated dividing the North American continent into zones, in each of which only one or two 35-mc common carrier frequency allocations are made. The nature of skip interference is such that, although the frequency allocation may be reused at distances up to a few hundred miles, it cannot be reused at distances from roughly 1000 to 3000 miles. Fig. 5 indicates this well-known zoning plan as it is used for 35-mc mobile telephone service. In general, this means that a transmitter of reasonably high power, such

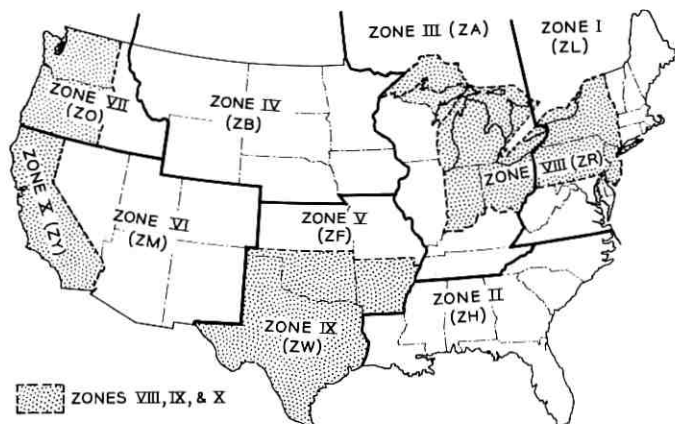


Fig. 5 — Common carrier frequency zoning plan for 30-44 mc mobile telephone service.

as 250 watts, must be assigned only to its own zone frequency. Otherwise, occasional skip interference from its radiation will cause serious interference to mobile telephone receivers in distant zones. It is rather significant and of interest, however, that such interference, which amounts to free-space transmission, is usually no more than about 20 db above a usable signal.

All Bellboy receivers are at least 20 db less sensitive than mobile telephone receivers because of their small antennas. Thus at 35 mc there would probably be no interference problem into Bellboy receivers from a distant zone. On the other hand, it would not be allowable to transmit at 250 watts on a channel assigned to another zone because of interference into mobile radio receivers.

Reducing transmitter power 20 db or more does not appear to be a practical solution, since even with 250-watt transmitters and high-performance receivers the range into modern steel-reinforced buildings is only about one and one-half miles, and into big-city streets it is about five miles. (Section V covers these matters in more detail.)

These factors sum up to the conclusion that in the majority of cases it is only practical to give Bellboy service at 35 mc on the regular zone channel in any one area. This means that if it is desired to give such service in an area already having highway mobile telephone service, it is necessary to integrate Bellboy with the mobile service. This can be readily done, but it places a severe limitation on the possible number of customers and the grade of service.

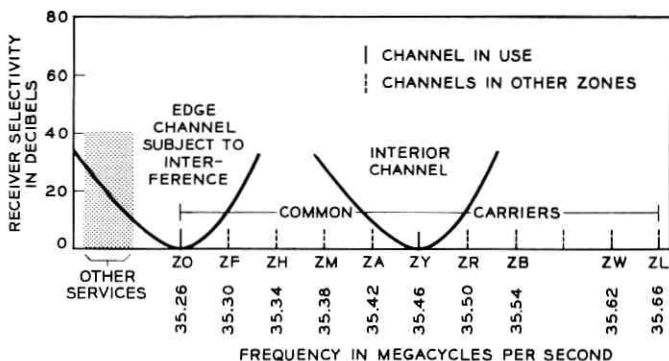


Fig. 6 — Factors in channel usage at 35 mc.

However, there is an unexpected dividend from this restriction to using only the zone channel, which is illustrated in Fig. 6. In most zones, one of the interior channels is used, so that there will be no high-powered radiation in that zone on any nearby channels. This allows the use of a relatively nonselective receiver. A problem may arise, however, if there is potential interference from nearby channels assigned to other than common carrier services. One example is the use of an edge channel, as shown in Fig. 6.

These factors permit the satisfactory use of a super-regenerative receiver to give service in many small to moderate-sized cities throughout the country. About 20 systems are now in service. Such a receiver is naturally substantially simpler and therefore cheaper than a more selective receiver. As a matter of fact, it is apparent from Fig. 6 that in most situations there is no incentive to use a more selective receiver, as long as the present allocation plan is used in the 35-mc frequency range.

On the other hand, it is also apparent that, in order to provide adequate service in the areas where there are large numbers of potential customers, a system is needed that can operate where there are many more channels available. Fig. 7 indicates the situation in the 150-mc band. Here there are 11 channels assigned to common carrier usage and, since there is no long-range interference problem, all these channels are available in any area.

In some highly settled areas it may be necessary to use almost all of these channels for mobile telephones, thereby requiring that Bellboy service be given on an integrated basis. In many areas, even the thickly settled ones, however, it will be possible to set aside at least one clear channel for Bellboy. As pointed out in Section IV, one clear channel pro-

vides a large customer capacity. In addition, the 150-mc system has a 4-to-1 speed advantage due to the use of simultaneous tones instead of the sequential system used in the 35-mc receiver.

These advantages, however, are obtained at a considerable cost. As indicated in Fig. 7, the receiver selectivity must be good enough to allow adjacent-channel operation. This, of course, also implies a high degree of frequency stability.

IV. TRAFFIC CONSIDERATIONS

The central office equipment in the J-1 terminal is designed on a flexible basis, so only the equipment needed for a particular situation will be installed. When fully equipped, the system will accommodate a maximum of 3200 four-digit numbers. These fall in the following number groups:

1000-1999,
2100-2699,
3000-3999,
4100-4699.

The number of Bellboy customers, N , that can be served from a single radio channel depends on the amount of traffic which they generate. Since the holding time, h , of each call is fixed at approximately three seconds by design, the calling rate during the busy hour, c , becomes the controlling factor. Early experience with the 35-mc system indicates only 0.4 call per customer per day and approximately 0.1 call per hour in the

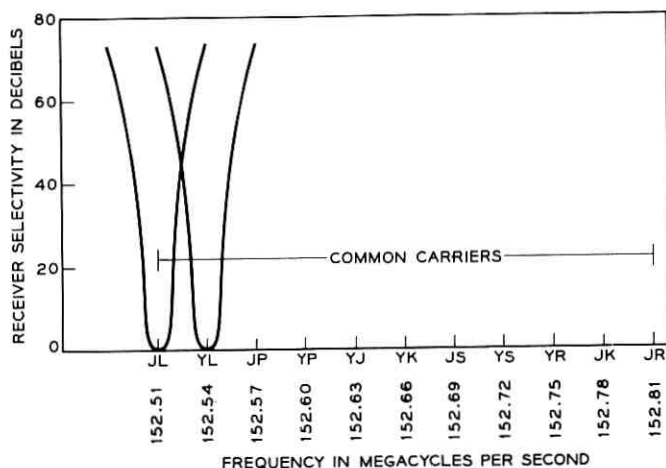


Fig. 7 — Factors in channel usage at 150 mc.

busy hour. However, in the New York trial the corresponding figures were about three times larger. The following expression can be used to find the RF channel capacity:

$$N = \frac{3600}{ch} E,$$

where E is an efficiency factor. Taking $E = 0.8$, $c = 0.3$, and $h = 3$ seconds, it is found that

$$N = \frac{(3600)(0.8)}{(0.3)(3)} = 3200.$$

It thus appears that for this calling rate the radio channel would be fully loaded by the 3200 codes included in the present system. It is possible that the calling rate will vary from one city to another, depending on the characteristics of the principle users of the system.

The number of registers and storage units needed is also affected by the calling rate. Registers can be installed on a unit basis, while stores come in blocks of 10. The present maximum is 10 registers and 40 stores.

The holding time for an incoming trunk and register circuit depends somewhat on the customer. In cases where the call comes through tandem and the customer hangs up promptly upon hearing the warble tone signifying acceptance of the call, the holding time, h_r , may be as short as 7 seconds. In cities where the customer dials directly from a step-by-step office and waits for the complete voice announcement the holding time may increase to about 12 seconds. Taking the lower figure of 7 seconds and assuming $c = 0.2$ and $N = 3200$, the load offered the registers in the busy hour would be

$$a_r = \frac{ch_r N}{3600} = \frac{(0.2)(7)(3200)}{3600} = 1.24.$$

This average figure, which is derived by assuming that all calls come in at a uniform rate, is defined as the load in *erlangs*. Standard tables are available⁹ that show the grade of service for a given load when various numbers of the given device or facility are available.

From such a table it is apparent that the provision of five registers would insure a good grade of service, in which the probability of being delayed due to lack of a trunk and register would be less than 0.01. Assumption of a 12-second holding time would increase the load to 2.1 erlangs. This would require seven registers for approximately the same grade of service.

In order to develop the holding time for a storage unit, it is necessary to recognize that the basic time interval during which new calls arrive in

storage has a duration equal to the maximum service time for the immediately preceding group of calls. This is 60 seconds plus the time needed for the third transmission of the n old calls being serviced. Since each transmission of a call requires one second, the basic acceptance interval is $(60 + n)$ seconds, during which n' new calls arrive in storage. Since it is equally likely that new calls may appear at any time in this interval, the average value of the acceptance interval is thus $(60 + n)/2$ seconds. Each new call has its third transmission completed in 60 seconds, and so the average call transmitting time is

$$60 + \frac{n'}{2}.$$

The average storage holding time is the sum of these two times, i.e.,

$$90 + \frac{n}{2} + \frac{n'}{2},$$

or about $90 + \bar{n}$, where \bar{n} is assumed to be the average n .

It is of interest to develop the load factor for stores but, in order to do so, we must determine \bar{n} for various calling rates. The average number of stores needed, \bar{n} , depends on the calling rate and holding time, as before. Substituting the value $(3t + \bar{n})$ for the holding time, where t is the interval in seconds between successive transmissions, it is found that

$$\bar{n} = \frac{cN(3t + \bar{n})}{3600},$$

which is the same as the load expressed in erlangs. Solving gives

$$a_s = \bar{n} = \frac{3tcN}{3600 - cN} \text{ erlangs.}$$

This is not a rigorous solution, but is believed to be a reasonably close approximation.

The following table shows the load offered to the storage system for 3200 customers and various values of calling rate c and time between signals t :

| Calling rate in busy hour, c | Load offered, a_s , in erlangs | | |
|-----------------------------------|----------------------------------|------------------|------------------|
| | $t = 20$ seconds | $t = 25$ seconds | $t = 30$ seconds |
| 0.30 | 21.7 | 27.3 | 32.5 |
| 0.20 | 13.0 | 16.3 | 20.8 |
| 0.10 | 5.9 | 7.4 | 9.4 |

Reference to delay curves shows that, with 3200 customers, 40 stores will be sufficient if the calling rate is 0.2, and be nearly sufficient for a calling rate of 0.3 to assure a delay probability of 0.01. If desirable, the load could be reduced by reducing the interval t between transmissions. The figure of 30 seconds has been used during the trial, but further experience may indicate that this can be reduced somewhat. It should also be noted that the present system is more than adequate for a calling rate of 0.1, which has been found in systems operating in other cities. A more rigorous analysis of this case is beyond the scope of this paper.

The tone generators are designed on an individual basis, each unit being a replaceable printed circuit board. In small installations it is possible to substitute a resistance for some of the generators and equip only those required for the N codes in the initial system. The number needed is approximately $2(N^2 + 1)$ generators. A new installation having 300 customers could start by equipping 16 generators. As growth occurs additional units can be inserted in the mounting rack.

A study of the calling habits of customers in the New York trial was carried out by slaving a 20-pen recorder to the relays controlling connection of the tone generators. These recordings furnished the data for the estimates that were used earlier for the average calling rates during the five-day working week. There is some calling on Saturday and Sunday, but at a greatly reduced rate.

V. RADIO TRANSMISSION CONSIDERATIONS

The key element in transmission performance of the Bellboy system is the pocket receiver. The 150-mc receiver can trigger on an electrical input signal of about one microvolt and is thus comparable in electrical performance to receivers now being used for mobile telephone. In addition, its selectivity is better than 80 db at a frequency 30 kc away from the carrier, which is also comparable to mobile telephone requirements.

Any pocket receiver, however, suffers a very substantial loss in effective sensitivity due to the fact that the antenna is small compared to a half wavelength. The antenna in this receiver is a ferrite stick, which has been found to give about the best performance of any tested so far. Even this antenna, however, has a loss of over 20 db compared to a half-wave dipole.

One of the problems in engineering this system was to find a reliable and simple method of measuring the small receivers. Two types of measurement are important. One is to determine sensitivity in free space and the other to determine sensitivity when the receiver is carried in the normal manner near a person's body.

Early tests were made in an open field at Holmdel, New Jersey, to determine free-space sensitivity, and these figures were then compared to typical sensitivity on the body. Sensitivity is defined as the field in decibels above 1 microvolt per meter required to just trigger the receiver.

Fig. 8 shows the relative response of a typical receiver when worn in a man's outer coat pocket as compared to its sensitivity in free space. As might be expected, there is quite a large variation depending on which way the receiver or the man with a receiver is facing relative to the transmitter. It is of interest that in some orientations the man provides gain. Apparently, a ferrite stick couples electrically to some extent to the body, thus effectively using the body as an antenna. In other orientations, however, the body provides shielding. It has been assumed for engineering purposes that the average sensitivity of a receiver would be used for estimating range. This seems reasonable, since it appears equally probable that a person may be facing in any direction when he is signaled. It is of interest that the average sensitivity of this receiver is 35.8 db on the body and 35 db in free space.

It is apparent that measuring receivers in an open field is a clumsy and time-consuming process, so experiments were made with various methods of measuring receivers in the laboratory. The most compact arrangement was a jig in which the receiver was placed in a plastic structure so that it coupled closely to a fixed pick-up coil. Although this method was satisfactory for rough checking, it was found that its accuracy left something to be desired. For this reason another method was worked out and is shown in Fig. 9. This involved the use of a short section of waveguide which is big enough to transmit 150 mc with very little attenuation. Al-

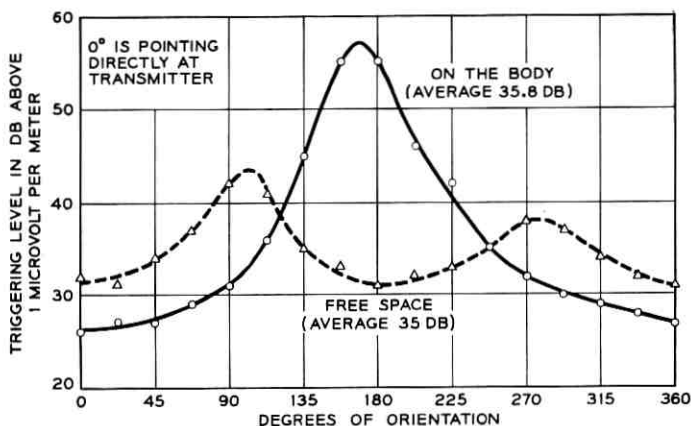


Fig. 8 — Relative response of receiver in free space and on the body.

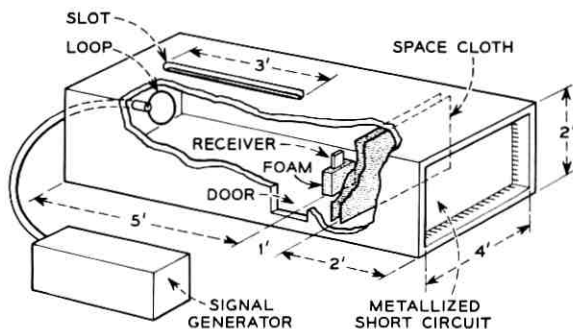


Fig. 9 — Method of measuring receivers using 150-mc waveguide "bear trap."

though this arrangement, which was nicknamed the "bear trap," was somewhat bulky, it could be set up in a moderate-sized room and was found to give reproducible results.

As indicated, the receiver was placed in a piece of foam plastic about four feet away from the transmitting loop and in the center of the waveguide. The combination of space cloth and metallic termination at the right turned out to be quite effective in holding down the standing wave ratio, which was about 1.2. It thus turned out that the receiver was in a fairly uniform vertically polarized electric field, and the measurements were not overly critical as to receiver position. Fig. 10 shows a correlation of free-space measurements in an open field and measurements in the "bear trap" for about 30 receivers.

Using the figure of 36 db sensitivity, it is possible to predict range using the technique described by Rice.⁵

The following table summarizes the figures used in this estimate in the same manner as those shown in Table II of the Rice paper:

| | |
|---|----------|
| <i>A</i> — Receiver sensitivity above 1 microvolt per meter: | 36 db |
| <i>B</i> — Corresponding minimum usable power in a whip antenna: | -116 dbw |
| <i>C</i> — Radio transmitter power (dipole): | 24 dbw |
| <i>D</i> — Maximum allowable path loss (<i>C</i> - <i>B</i>): | 140 db |
| <i>E</i> — Building loss for 78 per cent reliability: | 31 db |
| <i>F</i> — Equivalent maximum allowable median loss in streets (<i>D</i> - <i>E</i>): | 109 db |

For convenience, Fig. 11 is reproduced from the Rice paper,⁵ showing the building loss in db referred to median field in the streets. Fig. 12, reproduced from the same paper, indicates path loss between a half-wave dipole and a whip in the streets. (Most of this material originally ap-

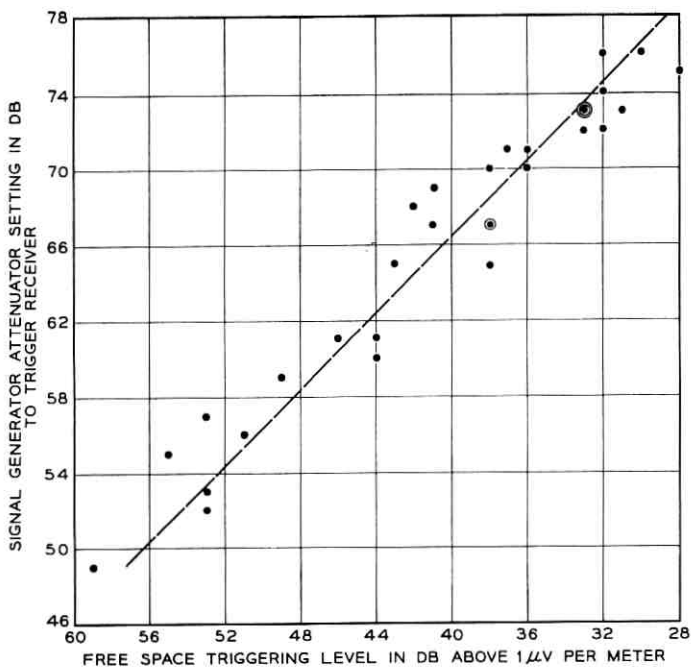


Fig. 10 — Bellboy system receiver measurements in the "bear trap" vs. measurements in free space.

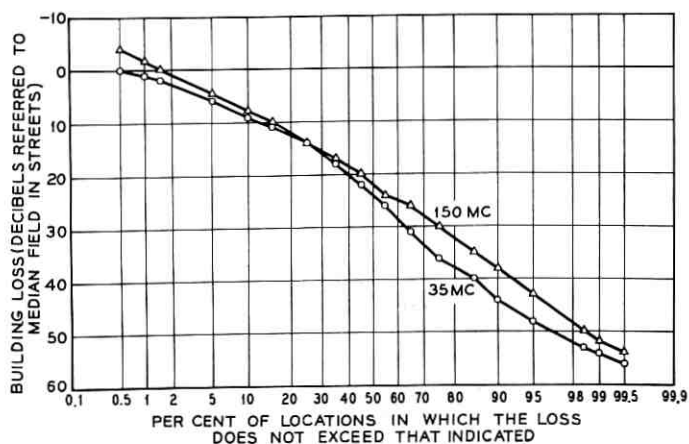


Fig. 11 — Over-all distribution of building losses at 35 and 150 mc.

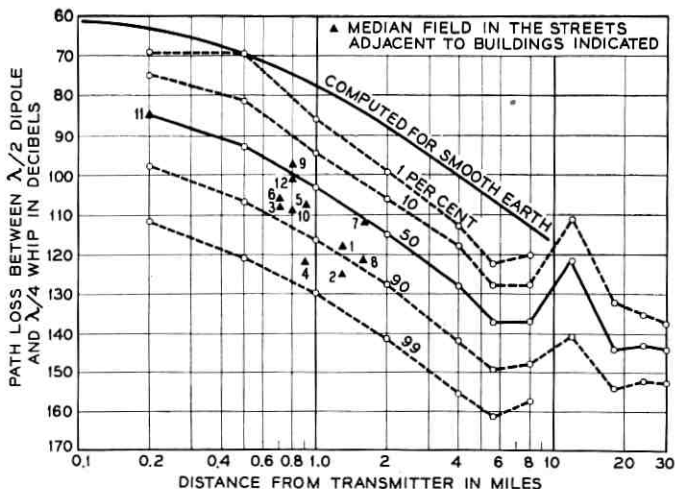


Fig. 12 — Measured path loss at 150 mc between a half-wave dipole and a quarter-wave whip in Manhattan, the Bronx, and suburbs. Antenna heights were 450 feet for transmitter, 6 feet for receiver.

peared in a paper by Young.¹⁰) From Fig. 12 it is apparent that 109 db median path loss in the streets corresponds to a range of about 1.6 miles. Thus, we would estimate a little over a mile and one-half range into buildings with this receiver, with a 78 per cent chance of success in one try.

Similarly, by interpolating on Fig. 12 we would estimate a range of about four and one-half miles in the streets for the path loss to be less than 140 db for 78 per cent of the cases.

The use of 78 per cent reliability for one try ties in with the planned reliability of Bellboy system service of 0.99 or better. Thus, by engineering a system so that the probability of receiving a single transmission is 0.78, the probability that failure will occur three successive times is $(0.22)^3$ or one per cent. Thus the probability that at least one transmission will be successfully received is $1 - (0.2)^3$ or 99 per cent.

Our experience has indicated that this estimate of one and one-half miles range into buildings and four and one-half miles in the streets is reasonably correct. Obviously, several transmitters will be needed to cover a big city.

It has been found possible to transmit simultaneously from transmitters with overlapping coverage provided radio frequency is controlled to an accuracy of one part per million, i.e., 10^{-6} . This holds the operating frequency of 152.780 mc to ± 153 cps, and keeps beat notes in the receiver below the lowest tone in the coding system.

Laboratory tests have shown that simultaneous reception of two signals produces a slight gain in receiving provided the phases of the modulating tone differ by less than 70° . This is easy to assure, and modulation of the three transmitters in the trial was well within this limit. It has not been possible to discover any effect, either positive or negative, in field testing. If, however, the phases of the modulation at two transmitters differ by more than 90 degrees there is a definite penalty. When the difference is 130 degrees the receiver will not operate if the two signal strengths are within 4 db of each other. At 180 degrees the corresponding figure is 6 db. These test results were obtained when the frequency difference between the two carriers ranged from 20 to 300 cycles per second.

Fig. 13 shows the locations of the three 250-watt transmitters in lower Manhattan. These give coverage in every street location tested below 59th Street and in practically all buildings. As might be expected, however, there are a few dead spots. These occur particularly in the lower floors and the interior portions of large buildings which have considerable metal in their construction. One example is the interior of the lobby of the Waldorf-Astoria Hotel, while another is part of Grand Central Terminal.

One customer also found that he could not receive signals while he was in the hold of a ship docked in Brooklyn. This does not seem surpris-

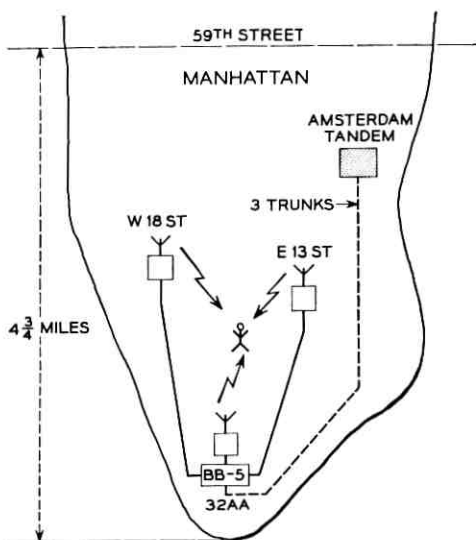


Fig. 13 — Experimental trial of 150-mc Bellboy service in Manhattan.

ing, since a ship interior is probably nearly the equivalent of a shielded room.

In general, it appears that the estimated range figures can be used for most commercial situations. It may turn out, however, that additional low-power transmitters will be needed to cover certain important buildings such as hospitals or large office buildings.

Transmitters would have to be located fairly close together in places like downtown Manhattan, but could be spread out considerably in residential areas. The short range, however, is not entirely a disadvantage. It will allow a channel frequency used exclusively for Bellboy service to be reused at much shorter distances than is the case with mobile telephone.

A potential interference problem exists, however, from the use of satellite transmitters. This comes about when a mobile telephone customer on an adjacent channel drives his car in the near vicinity of a satellite transmitter. In this case, the mobile customer's receiver may be subjected to a strong field compared to the field he receives on the desired channel from a distant fixed transmitter. It appears, however, that trouble can be avoided on this score.

In the first place, the fixed transmitter used on the adjacent mobile channels can be coordinated with the Bellboy system so that it will always be turned on if Bellboy transmitters are radiating. This will guarantee that there is at least some field present near every satellite transmitter. A further improvement in this situation can be effected by the use of directive antennas for the satellites. These may be stacked dipoles or some similar arrangements providing directivity in the vertical plane, so that excessively strong fields just below the satellite transmitters would be avoided.

REFERENCES

1. Strack, W., Pocket Radio Signaling, Bell Lab. Rec., **36**, 1958, p. 9.
2. Kraus, C. R., Citywide Personal Signaling at Allentown and Bethlehem, Pa., A.I.E.E. Summer Meeting, Buffalo, N. Y., June 22-27, 1958.
3. Guernsey, E. D., and Monk, N., Personal Signaling, A New Telephone Service, Wescon meeting, Los Angeles, August 22, 1958.
4. Young, J. W., Page Master Receiver and Modulator Equipment, A.I.E.E. Summer Meeting, Buffalo, N. Y., June 22-27, 1958.
5. Rice, L. P., Radio Transmission into Buildings on 35 and 150 mc, B.S.T.J., **38**, 1959, p. 197.
6. Pruden, H. M., and Hoth, D. F., Vibrating Reed Signaling for Mobile Radio, Trans. A.I.E.E., **68**, Pt. 1, 1949, p. 387.
7. Keller, A. C., and Bostwick, L. G., Vibrating Reed Selectors for Mobile Radio Systems, Trans. A.I.E.E., **68**, Pt. 1, 1949, p. 383.
8. Bostwick L. G., and Guncelle, R. L., U. S. Patent No. 2,877,319.
9. Palm, C., *Table of Erlang Loss Formula*, L. M. Ericsson Co., Stockholm, 1954.
10. Young, W. R., Jr., Comparison of Mobile Radio Transmission on 150, 450, 900, and 3700 mc. B.S.T.J., **31**, 1952, p. 1068.

The Effect of Driving Electrode Shape on the Electrical Properties of Piezoelectric Crystals

By J. A. LEWIS

(Manuscript received February 23, 1961)

In the present paper general formulas for the electrical admittance of a piezoelectric crystal, in terms of its resonant frequencies and static and motional capacitances, are derived and applied to the investigation of the effect of electrode shape on the spectrum of resonances and the capacitance ratio of the crystal. Particular attention is given in two cases of practical importance, namely, small piezoelectric coupling and thin crystal plates.

I. INTRODUCTION

Piezoelectric crystals are often used as circuit elements in filters and oscillators. Fig. 1 shows a typical admittance curve for such a crystal and Fig. 2 shows the corresponding equivalent circuit. At very low frequencies the crystal behaves like a capacitor, with a capacitance approximately equal to the static capacitance between the driving electrodes. Due to the piezoelectric effect, an applied alternating electric field causes the crystal to vibrate and, at certain natural frequencies of free vibration, it is driven into mechanical resonance by the applied voltage.

In the neighborhood of such natural frequencies, the admittance of the crystal is closely approximated by the simple equivalent circuit of Fig. 3. This is the equivalent circuit commonly used in the applications. It is a good approximation over a frequency range proportional to the spacing between the resonant frequency, at which the admittance is infinite (in the absence of dissipation), and the antiresonant frequency, at which the admittance vanishes.

For small electromechanical coupling this spacing is proportional to the capacitance ratio C_n/C_0 . It is desirable to make this ratio as large as possible. Bechmann and Parsons¹ have shown how this may be done in various simple cases.

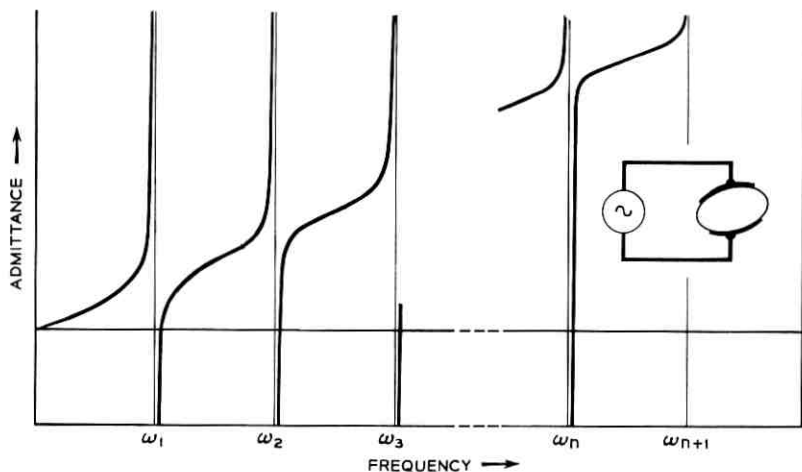


Fig. 1 — The equivalent electrical admittance of a piezoelectric crystal.

At high frequencies, where many resonances may occur within a narrow frequency range, the simple equivalent circuit of Fig. 3 may cease to be applicable. However, it is always possible, at least in principle, to find an electrode configuration which does not excite one or more of these resonances. A simple example is a symmetric electrode configuration, exciting only symmetric modes of free vibration. Vomer² has shown theoretically and experimentally what the appropriate electrode shape is for the longitudinal vibrations of a piezoelectric bar.

In the following, we consider both the question of capacitance ratio maximization and that of resonance suppression in some detail from the theoretical point of view. The investigation is divided into three parts.

In Sections II through VI we consider the general problem of steady vibrations of a piezoelectric body. The principal tools used in these sections are the piezoelectric analogs of various integral theorems of classical elasticity, found, for instance, in Love.³ We prove, for example, that the material particle displacements corresponding to two different

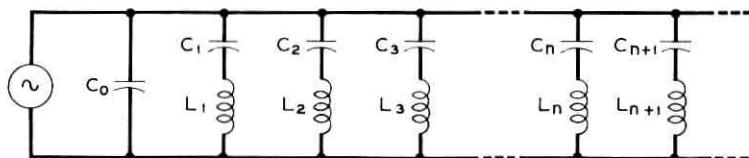


Fig. 2 — The equivalent electrical circuit of a piezoelectric crystal.

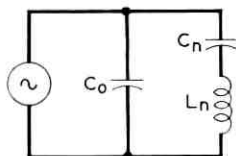


Fig. 3 — The equivalent electrical circuit near resonance.

modes of free vibration of a piezoelectric body are orthogonal just as they are for an ordinary elastic body.

The principal results of these sections are general expressions for the equivalent electrical admittance of any piezoelectric body and for the motional capacitances. The latter expression indicates how the electrode configuration must be chosen to suppress a given resonance.

Section VII is devoted to the case of small piezoelectric coupling, in which, to a first approximation, we may separate the mechanical and electrical problems completely, thus making it possible to obtain more explicit results than in the general case. A sample result states that the capacitance ratio C_n/C_0 is equal to the square of that portion of the input electrical energy which goes into exciting the n th mode, divided by the product of the electrostatic energy of the driving field and the strain energy of this mode of free vibration. As incidental by-products of our calculations, we obtain upper and lower bounds on the change in resonant frequency produced by a change in electrode configuration and an upper bound on the capacitance ratio.

Finally, in Sections VIII and IX, we make use of the simplifications possible when the piezoelectric body is a thin plate, in particular obtaining an explicit relation for the electrode configuration maximizing the capacitance ratio. As a simple although somewhat artificial example, we consider the low-frequency longitudinal vibrations of a piezoelectric bar (the example treated by Vormer²) in some detail, using this example to point out some of the advantages, as well as the difficulties, in the practical application of the foregoing techniques.

Before commencing our discussion of these problems, a few general remarks may be appropriate. In order to apply the techniques presently proposed, considerable detailed information about the modes of free vibration is required. Until recently, such information was available only for a few special cases. However, at least for crystal plates, it appears that such information may be obtained by the approximate theoretical techniques developed by Mindlin and his co-workers* over the

* The output of Mindlin and his coworkers in this field is so extensive that a complete bibliography is impossible here. Refs. 4, 5, and 6 are most closely connected with our work.

last decade. Experimental methods using an electric probe which may be moved over the surface of the crystal, such as the methods developed by Van Dyke⁷ and by Koga, Fukuyo, and Rhodes,⁸ may also yield the desired information. Conversely, measurement of the electrical admittance of a given crystal for various electrode configurations and driving frequencies may yield considerable information about mode shapes.

II. THE BASIC EQUATIONS

The equations governing the steady vibrations of a piezoelectric crystal at angular frequency ω may be written in the form

$$T_{ij,j} + \rho\omega^2 u_i = 0, \quad (1)$$

$$D_{i,i} = 0, \quad (2)$$

where T_{ij} is the stress tensor, ρ the mass density, u_i the material particle displacement vector, and D_i the electric displacement vector. Here and in the following, we use Cartesian tensor notation (see, for example, Jeffries⁹), in which commas denote differentiation with respect to the Cartesian coordinates (x_1, x_2, x_3) and a repeated subscript indicates summation over all possible values of that subscript. Thus, for example, (2) is just the usual quasistatic electric field equation, stating that the divergence of the electric displacement vector vanishes. The symbols used for stress, strain, etc., are those used by Mason.¹⁰

In the case of a piezoelectric medium the stress T_{ij} and electric displacement D_i are given in terms of the strain S_{ij} and electric field E_i by the linear anisotropic constitutive relations

$$T_{ij} = c_{ijkl}^E S_{km} - e_{kij} E_k, \quad (3)$$

$$D_i = \epsilon_{ik}^S E_k + e_{ikm} S_{km}, \quad (4)$$

where S_{ij} and E_i are given in terms of the particle displacement u_i and electrical potential V by the relations

$$S_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}), \quad (5)$$

$$E_i = -V_{,i}. \quad (6)$$

The elastic constants c_{ijkl}^E , piezoelectric constants e_{kij} , and dielectric constants ϵ_{ij}^S satisfy the symmetry relations

$$c_{ijkl}^E = c_{jikm}^E = c_{ijmk}^E = c_{kmij}^E,$$

$$e_{kij} = e_{kji},$$

$$\epsilon_{ij}^S = \epsilon_{ji}^S.$$

We shall also have occasion to use the alternate constitutive relations

$$T_{ij} = c_{ijkm}^D S_{km} - h_{kij} D_k, \quad (7)$$

$$E_i = \beta_{ik}^S D_k - h_{ikm} S_{km}, \quad (8)$$

$$S_{ij} = s_{ijkm}^E T_{km} + d_{kij} E_k, \quad (9)$$

$$D_i = \epsilon_{ik}^T E_k + d_{ikm} T_{km}, \quad (10)$$

where

$$\beta_{ik}^S \epsilon_{jk}^S = \delta_{ij} = \begin{cases} 1, & \text{for } i = j, \\ 0, & \text{for } i \neq j, \end{cases}$$

$$h_{ikm} = \beta_{ij}^S e_{jkm},$$

$$c_{ijkm}^D = c_{ijkm}^E + h_{pij} e_{pkm},$$

$$s_{ipkq}^E c_{jpmq}^E = \delta_{ij} \delta_{km},$$

$$d_{kij} = e_{kpq} s_{ijpq}^E,$$

$$\epsilon_{ij}^T = \epsilon_{ij}^S + d_{ipq} e_{j pq}.$$

The magnitude of the piezoelectric effect is specified by the piezoelectric coupling coefficient k , where k^2 may be given by any one of the three dimensionless ratios, $e^2/c\epsilon$, $h^2/c\beta$, $d^2/s\epsilon$, with e , c , ϵ , etc., being representative values of the corresponding material constants. For all real piezoelectric materials k is small compared with unity.

III. BOUNDARY CONDITIONS

The particle displacement u_i and electric potential V , satisfying the preceding equations in a piezoelectric body B , are completely determined by the specification of certain conditions on its surface S (see Fig. 4). We shall always assume that the body is supported in such a fashion that its surface is free of tractions, i.e.,

$$T_{ij} n_j = 0 \quad (11)$$

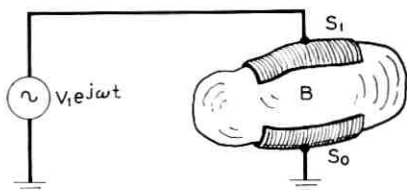


Fig. 4 — Piezoelectric body with driving electrodes.

on S , where n_i is the outward normal vector. This assumption is not essential; we could consider just as well the case of a rigid (or compliant) support. It does, however, make the subsequent algebra somewhat simpler.

The body is assumed to be driven by an alternating voltage of constant amplitude V_1 , applied between electrodes S_0 and S_1 , plated on its surface. Thus

$$V = \begin{cases} V_1, & \text{on } S_1, \\ 0, & \text{on } S_0. \end{cases} \quad (12)$$

The remaining surface is assumed to be free of plating. On this portion of the surface, then, one has two conditions requiring that the potential and normal electric displacement be equal to the potential and normal electric displacement in the external field. We may take this external leakage field into account formally by simply requiring that the equations of the preceding section hold throughout all space (except on S_0 and S_1), with the stress, strain, and piezoelectric constants vanishing identically outside B and the dielectric constant outside B being that of free space. In the following we shall indicate that an integral is to be taken over all space by using B_∞ in place of B . To complete the set of boundary conditions, we assume that V vanishes at large distance from B . The external field is included here only for formal logical completeness; in almost all practical problems it is of negligible importance.

IV. INTEGRAL RELATIONS FOR PIEZOELECTRIC BODIES

Suppose that u_i and V are any (suitably continuous) vector and scalar functions, defined on $B + S$. Note that, by our previous convention, V is actually defined throughout all space. We shall assume that V is continuous across S and zero at infinity. In general, of course, u_i and V will not satisfy the equations of the preceding section. However, by introducing suitable volume and surface forces, F_i , T_i , and volume and surface charges, q_B , q_S , we may construct a boundary value problem satisfied by these functions.

First we calculate strains and electric field components, S_{ij} , E_i , from (5) and (6) and stresses and electric displacement components from the constitutive relations, (3) and (4). The required distributions of volume force and charge are then given by the relations

$$\begin{aligned} \rho F_i &= - (T_{ij,j} + \rho \omega^2 u_i), \\ q_B &= D_{i,i}, \end{aligned}$$

where ω is an arbitrary positive constant. Similarly, the distributions of surface force and charge on S are

$$\begin{aligned} T_i &= T_{ij}n_j, \\ q_s &= [D_i]n_i, \end{aligned}$$

where n_i is the outward normal to S and $[D_i]$ is the jump in the electric displacement vector across S , given by

$$[D_i] = (D_i)_{\text{ext}} - (D_i)_B.$$

Now suppose we have any two pairs of such functions (u_i', V') , (u_i'', V'') and calculate the corresponding stresses, electric displacements, etc. Then the divergence theorem yields

$$\int_B u_i' T_{ij,j}'' dB = \int_S u_i' T_{ij}n_j dS - \int_B S_{ij}' T_{ij}'' dB, \quad (13)$$

$$\int_{B_\infty} V' D_{i,i}'' dB = - \int_S V' [D_i''] n_i dS + \int_{B_\infty} E_i' D_i'' dB, \quad (14)$$

or, in terms of the equivalent forces and charges,

$$\int_B \rho u_i' (F_i'' + \omega''^2 u_i'') dB = - \int_S u_i' T_i'' dS + \int_B S_{ij}' T_{ij}'' dB \quad (15)$$

$$\int_{B_\infty} V' q_B'' dB = - \int_S V' q_s'' dS + \int_{B_\infty} E_i' D_i'' dB. \quad (16)$$

For example, for the forced vibrations previously considered, with $(u_i', V') = (u_i'', V'') = (u_i, V)$, $\omega'' = \omega$, $F_i'' = T_i'' = q_B'' = 0$, addition and rearrangement of (15) and (16) yields

$$\int_{B_\infty} (c_{ijkl}^E S_{ij} S_{kl} + \epsilon_{ik}^S E_i E_k - \rho \omega^2 u_i u_i) dB = V_1 Q_1, \quad (17)$$

where V_1 is constant and the total charge on S_1 is

$$Q_1 = \int_{S_1} q_s dS = \int_{S_1} [D_i] n_i dS. \quad (18)$$

Note that, by our previous convention, c_{ijkl}^E and u_i vanish outside B . Equation (17) essentially states that the sum of the potential energy, made up of the strain energy and electrostatic energy, and the kinetic energy is equal to the energy passing into B through S_1 due to the applied voltage V_1 .

For free vibrations, with $(u_i, V) = (u_i^n, V^n, \omega_n)$, (17) yields the

Rayleigh quotient for the natural frequency ω_n , i.e.,

$$\omega_n^2 = \frac{\int_{B_\infty} (c_{ijkl}^E S_{ij}^n S_{km}^n + \epsilon_{ik}^S E_i^n E_k^n) dB}{\int_B \rho u_i^n u_i^n dB} \quad (19)$$

This relation suggests that the effect of piezoelectricity is to increase the natural frequencies over the values they would have in an ordinary elastic body with the elastic constants c_{ijkl}^E . In Section VII we shall prove that this is the case for small piezoelectric coupling coefficient k .

If we subtract (16) from (15) and rearrange terms, we obtain

$$\begin{aligned} \int_{B_\infty} [\rho u_i' (F_i'' + \omega''^2 u_i'') - V' q_B''] dB + \int_S (u_i' T_i'' - V' q_s'') dS \\ = \int_{B_\infty} (S_{ij}' T_{ij}'' - E_i' D_i'') dB. \end{aligned}$$

This equation still holds when we interchange primed and double-primed quantities. Furthermore

$$S_{ij}'' T_{ij}' - E_i'' D_i' = S_{ij}' T_{ij}'' - E_i' D_i'',$$

and thus

$$\begin{aligned} \int_{B_\infty} [\rho u_i' (F_i'' + \omega''^2 u_i'') - V' q_B''] dB + \int_S (u_i' T_i'' - V' q_s'') dS \\ = \int_{B_\infty} [\rho u_i'' (F_i' + \omega'^2 u_i') - V'' q_B'] dB + \int_S (u_i'' T_i' - V'' q_s') dS. \end{aligned} \quad (20)$$

This is the analytical form of the so-called reciprocal theorem (see Ref. 3, Chapter VII, p. 174) for a piezoelectric body. In words it may be stated as follows:

The difference between the mechanical and electrical work done by the forces (including kinetic reactions) and charges of the first set, acting over the displacements and potential produced by the second set, is equal to the difference between the mechanical and electrical work done by the forces and charges of the second set, acting over the displacements and potential produced by the first.

V. EXPANSION IN NORMAL MODES

We now return to the problem of forced vibrations, described in Sections I and II. The equations and boundary conditions of those sections

completely determine a unique particle displacement u_i and potential V , except at certain natural frequencies $\omega_n (n = 1, 2, \dots)$ where a non-trivial solution (u_i^n, V^n) of the free vibration problem (with $V_1 = 0$) exists. If we denote the solution of the static problem (with $\omega = 0$) by (u_i^s, V^s) , we may satisfy all of the preceding equations and boundary conditions, except (1), by setting

$$u_i = u_i^s + \sum_{n=1}^{\infty} a_n u_i^n, \quad (21)$$

$$V = V^s + \sum_{n=1}^{\infty} a_n V^n, \quad (22)$$

where the a_n 's are to be determined to satisfy (1). Since

$$T_{ij,j}^s = T_{ij,j}^n + \rho \omega_n^2 u_i^n = 0,$$

in this case (1) is satisfied if

$$\rho \sum_{n=1}^{\infty} (\omega_n^2 - \omega^2) a_n u_i^n = \rho \omega^2 u_i^s. \quad (23)$$

Now suppose we set (u_i', V') = (u_i^m, V^m) , (u_i'', V'') = (u_i^n, V^n) , with $\omega_m \neq \omega_n$, in (20). We obtain

$$\rho(\omega_m^2 - \omega_n^2) \int_B u_i^m u_i^n dB = 0, \quad (24)$$

so that displacements corresponding to two different natural frequencies are orthogonal, just as they are in ordinary elasticity. Thus, multiplying (23) by u_i^k and integrating over B , we obtain

$$\rho(\omega_k^2 - \omega^2) a_k \int_B u_i^k u_i^k dB = \rho \omega^2 \int_B u_i^s u_i^k dB.$$

Again applying (20), this time with (u_i', V') = (u_i^k, V^k) , (u_i'', V'') = (u_i^s, V^s) , we obtain

$$\rho \omega_k^2 \int_B u_i^s u_i^k dB = \int_{S_e} V^s [D_i^k] n_i dS = V_1 \int_{S_1} [D_i^k] n_i dS.$$

Thus

$$a_k = \frac{\omega^2 V_1 \int_{S_1} [D_i^k] n_i dS}{\rho \omega_k^2 (\omega_k^2 - \omega^2)}. \quad (25)$$

For future reference, we also note the identities

$$\rho\omega_n^2 \int_B u_i^n u_i^n dB = \int_B S_{ij}^n T_{ij}^n dB, \quad (26)$$

$$\int_{S_1} V_1 [D_i^n] n_i dS = \int_B e_{kij} (S_{ij}^s E_k^n - S_{ij}^n E_k^s) dB, \quad (27)$$

obtained by direct application of the divergence theorem.

VI. THE EQUIVALENT ELECTRICAL ADMITTANCE

Using the apparatus developed in the previous sections, it is a simple matter to obtain a general expression for the admittance of the crystal body. The admittance $Y(\omega)$ is the ratio of the total input current I_1 into the crystal to the voltage V_1 across its terminals. The current is the rate of increase of the total charge Q_1 on the electrode S_1 , where

$$Q_1 = \int_{S_1} [D_i] n_i dS,$$

so that

$$Y(\omega) = I_1/V_1 = j\omega Q_1/V_1 = j\omega V_1^{-1} \int_{S_1} [D_i] n_i dS.$$

Substituting

$$[D_i] = [D_i^s] + \sum_{n=1}^{\infty} a_n [D_i^n],$$

we obtain

$$Y(\omega) = j\omega \left(C_s + \sum_{n=1}^{\infty} \frac{\omega^2 C_n}{\omega_n^2 - \omega^2} \right), \quad (28)$$

where the static capacitance C_s is given by

$$C_s = V_1^{-1} \int_{S_1} [D_i^s] n_i dS, \quad (29)$$

and the motional capacitances C_n by

$$C_n = \frac{\left(\int_{S_1} [D_i^n] n_i dS \right)^2}{\rho\omega_n^2 \int_B u_i^n u_i^n dB}. \quad (30)$$

If we set

$$C_0 = C_s - \sum_{n=1}^{\infty} C_n,$$

(28) may also be written in the form

$$Y(\omega) = j\omega \left(C_0 + \sum_{n=1}^{\infty} \frac{C_n}{1 - \omega^2/\omega_n^2} \right). \quad (31)$$

Fig. 2 shows the corresponding equivalent circuit with

$$L_n = 1/\omega_n^2 C_n, \quad n = 1, 2, \dots,$$

and Fig. 1 shows the corresponding behavior of the admittance with frequency. Since $C_n \ll C_s$ for real piezoelectric materials, the admittance is very nearly equal to $j\omega C_s$, except in the vicinity of a resonant frequency, where the admittance is that of a simple series-resonant circuit, shunted by the capacitance C_0 . The corresponding antiresonant frequency ω_n' is given approximately by

$$\frac{\omega_n'^2}{\omega_n^2} = \frac{1 + C_n}{C_0},$$

or, since $C_n/C_0 \ll 1$,

$$\frac{\Delta\omega_n}{\omega_n} = \frac{\omega_n' - \omega_n}{\omega_n} = \frac{C_n}{2C_0}. \quad (32)$$

All of these considerations hold when the resonant frequencies are not too closely spaced, specifically when $(\omega_{n+1} - \omega_n)/\omega_n$ is large compared with C_n/C_0 . At high frequencies, where many resonances may occur in a narrow frequency band, the region of applicability of the above simple single-resonance circuit may be very small, and the simple relation for the antiresonant frequency ω_n' given by (32) may not be an adequate approximation. In order to circumvent this difficulty, we may choose the shape of the electrodes S_1 and S_0 so that resonances in the vicinity of a desired resonance in this frequency range are not excited, i.e., so that the corresponding C_n 's vanish, or

$$\int_{S_1} [D_i^n] n_i dS = 0. \quad (33)$$

This is a necessary and sufficient condition that the n th resonance not be excited by the given plating shape.

In order to make use of this condition, the surface charge $[D_i^n] n_i$ must be determined, either theoretically, perhaps by using an approximate theory of the type applied so successfully by Mindlin and his co-workers,^{4,5,6} or experimentally using a point electrical probe as developed by Van Dyke⁷ or Koga, Fukuyo, and Rhodes.⁸ In the case of high-frequency vibrations, one would expect that the resulting electrode shape would be quite complex. Clearly the choice of electrode shape suppressing a

finite number of resonances is not unique. For these and other even more pressing reasons, it is not clear whether the use of the condition is of any particular practical importance.

Besides affecting the value of C_n , the choice of electrode shape also affects the magnitude of the static capacitance C_s (or C_0). Since C_0 shunts the series-resonant elements, it is desirable to make its value as small as possible, i.e., to make the capacitance ratio C_n/C_0 as large as possible at the desired resonance. Finally, because of piezoelectric coupling, the values of the resonant frequencies themselves depend on the electrode shape, although only to second order in the piezoelectric coupling coefficient k . In order to obtain more concrete information concerning these effects, in the next section we consider the case of small piezoelectric coupling.

VII. SMALL PIEZOELECTRIC COUPLING

In this section we shall obtain expressions for the static capacitance, the capacitance ratios, and the shift in natural frequencies due to a change in electrode shape, valid for small piezoelectric coupling coefficient k . With an eye to the applications, these expressions should be the most convenient forms for application to the most common practical case, namely, a crystal plate.

First of all, it is clear that, to first order in k , the static capacitance C_s is simply the ordinary electrostatic capacitance in the absence of any piezoelectric effect. Thus, to first order, the static potential V^s satisfies the equations

$$\begin{aligned} D_i^s &= \epsilon_{ik}^s E_k^s = -\epsilon_{ik}^s V_{,k}^s, \\ D_{i,i}^s &= 0, \end{aligned}$$

in B , and

$$V^s = \begin{cases} V_1, & \text{on } S_1, \\ 0, & \text{on } S_0. \end{cases}$$

(In this instance, one should recall our previously agreed upon convention concerning the treatment of the external region.) The static capacitance is then given by

$$C_s = V_1^{-1} \int_{S_1} [D_i^s] n_i dS = V_1^{-2} \int_{B_\infty} \epsilon_{ij}^s E_i E_j^s dB. \quad (34)$$

Next we consider the solution of the free vibration problem for small coupling. To first order in k the particle displacement u_i^n and natural

frequency ω_n are those of the purely elastic problem, and the potential V^n and electric displacement D_i^n are of first order in k . Finally, the change in natural frequency is of second order in k . Thus we set

$$\begin{aligned}u_i &= u_i^0 + u_i'', \\V &= V', \\ \omega^2 &= \omega_0^2(1 + \nu''),\end{aligned}$$

where the quantities with index "0" are of zero order in k , those with single primes first order, and those with double primes second order in k . We have also temporarily dropped the index "n" in order to reduce the number of indices present. The governing equations and boundary conditions then become

$$\begin{aligned}T_{ij,j}^0 + \rho\omega_0^2 u_i^0 &= 0, & \text{in } B, \\T_{ij}^0 n_j &= 0, & \text{on } S, \\D_{i,i}' &= 0, & \text{in } B_\infty, \\V' &= 0, & \text{on } S_e = S_0 + S_1, \\T_{ij,j}'' + \rho\omega_0^2(u_i'' + \nu'' u_i^0) &= 0, & \text{in } B, \\T_{ij}'' n_j &= 0, & \text{on } S,\end{aligned}$$

where the T_{ij} 's and D_i 's remain to be defined.

The definitions of these quantities, and thus u_i^0 , ω_0 , V' , u_i'' , and ν'' depend on the choice of constitutive relations used. For example, if we use (3) and (4) for the stresses and electric displacements in terms of the strains and electric fields, we have

$$\begin{aligned}T_{ij}^0 &= c_{ijkm}^E S_{km}^0, \\D_i' &= \epsilon_{ik}^S E_k' + e_{ikm} S_{km}^0, \\T_{ij}'' &= c_{ijkm}^E S_{km}'' - e_{kij} E_k',\end{aligned}$$

whereas if we use (7) and (8) we have

$$\begin{aligned}T_{ij}^0 &= c_{ijkm}^D S_{km}^0, \\E_i' &= \beta_{ik}^S D_k' - h_{ikm} S_{km}^0, \\T_{ij}'' &= c_{ijkm}^D S_{km}'' - h_{kij} D_k'.\end{aligned}$$

Use of either of these sets of relations should give the same results (to the order considered) for the potential V' , the electric displacement D_i' , and the natural frequency $\omega^2 = \omega_0^2(1 + \nu'')$. We may apply the

reciprocal theorem, (20), to obtain a simple expression for ν'' in either case. We set $(u_i', V') = (u_i^0, 0)$, $(u_i'', V'') = (u_i'', V')$ to obtain the expressions

$$\nu_E = \int_{B_\infty} \epsilon_{ij}^S E_i' E_j' dB / \rho \omega_E^2 \int_B u_i^0 u_i^0 dB, \quad (35)$$

$$\nu_D = - \int_{B_\infty} \beta_{ij}^S D_i' D_j' dB / \rho \omega_D^2 \int_B u_i^0 u_i^0 dB, \quad (36)$$

where ω_E and ω_D are the natural frequencies found using the elastic constants c_{ijkm}^E and c_{ijkm}^D , respectively, and ν_E and ν_D are the corresponding frequency shifts due to the piezoelectric effect. Since, to second order,

$$\omega^2 = \omega_E^2(1 + \nu_E) = \omega_D^2(1 + \nu_D)$$

and $\nu_E \geq 0$, $\nu_D \leq 0$, we must have

$$\omega_E \leq \omega \leq \omega_D, \quad (37)$$

with equality only if ν_E or ν_D vanish. According to (35) and (36), this can only happen if E_i' or D_i' vanishes identically, i.e., if the divergence of the vector $e_{ikm} S_{km}^0$ vanishes in the former case, or if the curl of the vector $h_{ikm} S_{km}^0$ vanishes in the latter case. Thus, in general, neither of these bounds will be attained. We can, however, prove that the lower bound ω_E is most nearly attained by a completely plated crystal and the upper bound ω_D by a completely unplated crystal. Furthermore, making the electrodes larger always decreases the resonant frequencies. These results, which are not of central importance in our present considerations, may be obtained by the application of Schwarz's inequality and Green's identity to (35).

We have now obtained expressions for the static capacitance and the shift in resonant frequency caused by the piezoelectric effect. We now derive a simple and symmetric expression for the capacitance ratio C_n/C_s , using (26) and (27). We have

$$\int_{S_1} V_1 [D_i^n] n_i dS = \int_B e_{kij} (S_{ij}^s E_k^n - S_{ij}^n E_k^s) dB = - \int_B e_{kij} S_{ij}^n E_k^s dB,$$

neglecting terms of higher order in k . Also, from (26),

$$\rho \omega_n^2 \int_B u_i^n u_i^n dB = \int_B S_{ij}^n T_{ij}^n dB,$$

which, in the present case, is either the strain energy in terms of c_{ijkm}^E or in terms of c_{ijkm}^D , depending on our choice of constitutive relations.

To first order in k the two expressions are equal, of course. Finally, using (34), we obtain

$$\frac{C_n}{C_s} = \frac{\left(\int_B c_{kij} E_k^s S_{ij}^n dB \right)^2}{\left(\int_B \epsilon_{ij}^s E_i^s E_j^s dB \right) \left(\int_B c_{ijk}^E S_{ij}^n S_{km}^n dB \right)}. \quad (38)$$

Direct calculation, again using Schwarz's inequality, leads to the upper bound

$$\frac{C_n}{C_s} \leq \frac{\omega_D^2 - \omega_E^2}{\omega_E^2}. \quad (39)$$

Again we will not dwell on this incidental result, only remarking that it provides a general upper bound on a quantity which we usually wish to make as large as possible. An alternate form of (38), in terms of stresses, is

$$\frac{C_n}{C_s} = \frac{\left(\int_B d_{kij} E_k^s T_{ij}^n dB \right)^2}{\left(\int_B \epsilon_{ij}^T E_i^s E_j^s dB \right) \left(\int_B s_{ijk}^E T_{ij}^n T_{km}^n dB \right)}. \quad (40)$$

VIII. CRYSTAL PLATES

In this section we consider a thin piezoelectric crystal plate, driven by symmetrically placed electrodes on its face ($x_2 = 0$, $x_2 = h$), as sketched schematically in Fig. 5. In this case, if the thickness h is small compared

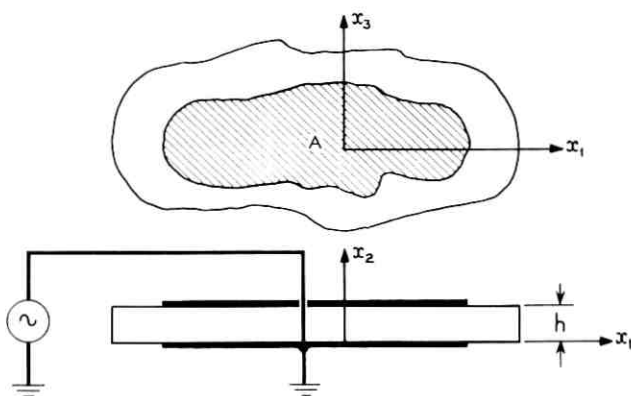


Fig. 5 — The crystal plate.

with the diameter of the electrodes, the static capacitance is given approximately by

$$C_s = \epsilon_{22}^S A/h, \quad (41)$$

where A is the electrode area. To the same approximation, the static field is given by

$$E_1^s = E_3^s = 0, \\ E_2^s = \begin{cases} -V_1/h, & \text{for } (x_1, x_3) \text{ in } A, \\ 0, & \text{for } (x_1, x_3) \text{ outside } A. \end{cases}$$

(We now neglect the external field completely.) Thus we have

$$\int_B e_{kij} E_k^s S_{ij}^n dB = -V_1 \iint_A e_{2ij} \bar{S}_{ij}^n(x_1, x_3) dx_1 dx_3,$$

where

$$\bar{S}_{ij}^n(x_1, x_3) = h^{-1} \int_0^h S_{ij}^n(x_1, x_2, x_3) dx_2,$$

or, in terms of average stresses,

$$\int_B d_{kij} E_k^s T_{ij}^n dB = -V_1 \iint_A d_{2ij} \bar{T}_{ij}^n(x_1, x_3) dx_1 dx_3,$$

where

$$\bar{T}_{ij}^n(x_1, x_3) = h^{-1} \int_0^h T_{ij}^n(x_1, x_2, x_3) dx_2.$$

With

$$\int_B \epsilon_{ij}^S E_i^s E_j^s dB = V_1^2 C_s = \epsilon_{22}^E A V_1^2 / h,$$

the capacitance ratio is given by

$$\frac{C_n}{C_s} = \left(\frac{h}{\epsilon_{22}^S A} \right) \left[\iint_A p_n(x_1, x_3) dx_1 dx_3 \right]^2, \quad (42)$$

where

$$p_n = \frac{e_{2ij} \bar{S}_{ij}^n}{\left(\int_B c_{ijkn}^E S_{ij}^n S_{kn}^n dB \right)^{1/2}}, \quad (43)$$

in terms of strains, and

$$p_n = \frac{d_{2ij} T_{ij}^n}{\left(\int_B \delta_{ijkm} E T_{ij}^n T_{km}^n dB \right)^{1/2}}, \quad (44)$$

in terms of stresses. Note that the integrals in the denominators of these two expressions for p_n are simply normalization factors. We retain them only to keep the dimensions straight.

Equation (42) gives the capacitance ratio in a form which is particularly easy to study for given p_n . We wish to know either how to choose A to make C_n/C_s zero, in order to suppress a given resonance, or how to choose A to make C_n/C_s as large as possible, to increase the useful bandwidth of the crystal.

Clearly the former problem has many, many solutions. If, for example, p_n is both positive and negative over the plate area, as will be the case unless we are dealing with some sort of fundamental mode, we need only distribute the electrode area A so that the integrals over the positive and negative portions are equal. The apparent difficulty when p_n is everywhere positive (or negative) can be circumvented by a simple artifice. We obtained (42) by assuming at the outset that voltages of the same polarity were applied between all portions of the top and bottom electrodes. If instead we imagine A to be divided into two parts, A^+ and A^- , with voltages of opposite polarities applied to these parts, then

$$\iint_A p_n dx_1 dx_3 = \iint_{A^+} p_n dx_1 dx_3 - \iint_{A^-} p_n dx_1 dx_3,$$

which clearly can be made to vanish by choosing A^+ and A^- so that the corresponding integrals are equal.

The problem of maximizing the capacitance ratio is mostly visualized simply as a geometric problem. First of all, note that in this case we may replace p_n by its absolute value in the integral in (42), for in regions where p_n is negative we may assume that the polarity of the driving voltage has been reversed. This obviously will always increase the capacitance ratio over the value it has with unreversed polarity in such regions. Now we imagine the surface $x_2 = |p_n(x_1, x_3)|$ to be erected above the plane $x_2 = 0$. The integral over $|p_n|$ can then be interpreted as the volume between this surface and the base plane, cut out by a cylinder of cross section A . Denoting this volume by V , we then must choose A to make V^2/A as large as possible. This problem has a very

simple solution, although one which is not particularly well adapted to calculation.

First of all, we observe that, for *fixed* area, to maximize the volume we should choose A to be the region bounded by the appropriate level curve of $|p_n|$. Any other choice of A , giving the same area, must consist of this region less a subregion plus an external region having the same area. But the value of $|p_n|$ for points outside the level curve is smaller than for points inside, so that the integral can only be decreased by this alternate choice.

Let us denote the area bounded by the level curve $|p_n(x_1, x_3)| = p = \text{a constant}$ by A_p and the corresponding volume by V_p . We must now choose A_p so that V_p^2/A_p is maximized, i.e., so that

$$\frac{d(V_p^2/A_p)}{dA_p} = 0.$$

Now the change in V_p due to an increment ΔA_p in A_p is $p\Delta A_p$, since the height of the surface above the base plane is constant around its boundary, and thus

$$\frac{d(V_p^2/A_p)}{dA_p} = \frac{V_p(2pA_p - V_p)}{A_p^2},$$

which vanishes when $pA_p = \frac{1}{2}V_p$, i.e., when the volumes above and below a plane through the level curve are equal. Explicitly, to maximize the capacity ratio we have the condition

$$pA_p = \frac{1}{2} \iint_{A_p} |p_n(x_1, x_3)| dx_1 dx_3, \quad (45)$$

where A_p is the total area of the electrode on either plate face, bounded by the level curve $|p_n(x_1, x_3)| = p$, and p_n is given by (43) or (44). We also assume that the electrode polarity is positive when p_n is positive and negative when p_n is negative. Since A_p must of course be less than or equal to the plate area, the maximum capacitance ratio given by (45) may not actually be attained. If this is the case, a plate wholly covered with electrodes gives the largest possible value to the capacitance ratio. It may also happen that the maximizing area A_p may be partly bounded by a level curve and partly bounded by the plate edge.

Bechmann and Parsons¹ have considered the excitation of piezoelectric bars and plates with partially applied electrodes and have obtained results similar to the above for the capacitance ratio, both theoretically and experimentally.

IX. THE LONGITUDINAL VIBRATIONS OF A PIEZOELECTRIC BAR

To illustrate the foregoing results, we now consider a concrete example. To keep the analysis simple, we treat the example discussed by Vomer² of the low-frequency longitudinal vibrations of the piezoelectric bar sketched in Fig. 6. We assume that $h \ll a \ll L$ and restrict our discussion to vibrations of wavelength large compared with a . Then we may assume that all the stresses T_{ij} are zero except T_{11} , which we take to be a function of x_1 only. Thus, for free vibrations,

$$T_{11} = T_{11}^n(x_1),$$

satisfying the equation of motion

$$T_{11,1}^n + \rho\omega_n^2 u_1^n = 0,$$

and the stress-strain relation

$$S_{11}^n = u_{1,1}^n = s_{1111}^E T_{11}^n.$$

Eliminating u_1^n between these two equations leads to the equation

$$T_{11,11}^n + (\omega_n/c)^2 T_{11}^n = 0,$$

where $c^2 = 1/\rho s_{1111}^E$. Together with the boundary conditions

$$T_{11}^n(0) = T_{11}^n(L) = 0,$$

this equation yields

$$T_{11}^n(x_1) = \sin n\pi x_1/L, \quad \omega_n = n\pi c/L.$$

If we assume that the width of the electrode A at $x = x_1/L$ is $af(x)$, as in Fig. 6, we find that the capacitance ratio is given by

$$\frac{C_n}{C_s} = 2k^2 \left(\frac{A_0}{A} \right) \left[\int_0^1 f(x) \sin n\pi x dx \right]^2, \quad (46)$$

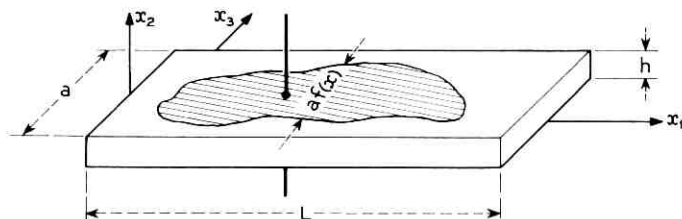


Fig. 6 — The crystal bar.

where k is the piezoelectric coupling coefficient, here given by

$$k^2 = \frac{(d_{211})^2}{s_{1111}^E \epsilon_{22}^S},$$

$A_0 = aL$ is the bar-face area, and A is the electrode area, given by

$$A = A_0 \int_0^1 |f(x)| dx.$$

We must assume that $0 \leq |f(x)| \leq 1$, and we may take into account portions of electrode of reversed polarity by assuming that $f(x)$ may be both positive and negative.

For the wholly covered bar ($f \equiv 1$), for example, (46) gives

$$\frac{C_n}{C_s} = \begin{cases} 4k^2/n^2, & \text{for } n = 1, 3, \dots, \\ 0, & \text{for } n = 2, 4, \dots. \end{cases}$$

(Clearly an even excitation, such as the above, cannot excite the odd modes.) This already provides an example of resonance suppression of a very trivial nature. In the present case, within the limitations of the theory, we can actually find an electrode shape which suppresses all resonances except one, for the T_{11}^n 's are orthogonal. Thus, if we set

$$f(x) = \sin m\pi x,$$

we find

$$\frac{C_n}{C_s} = \begin{cases} \pi k^2/4, & \text{for } n = m, \\ 0, & \text{otherwise.} \end{cases}$$

Of course, since the application of the present simple theory is restricted to low frequencies, we cannot expect the above to be valid for large n and m . On the other hand, the low-frequency resonances are sufficiently widely spaced so that resonance suppression is of no particular practical importance in this case. Thus the present example must be taken as illustrative, rather than practical.

In the present case of one-dimensional vibrations, the level lines of $|p_n|$, considered in the previous section, are straight lines $x_1 = \text{a constant}$. Thus the electrode shape maximizing the capacitance ratio consists simply of a sequence of rectangular bands of suitable widths and polarities, extending across the width of the bar. To determine the optimum width, it suffices to consider the fundamental mode ($n = 1$). For convenience we shift the origin of coordinates to the center of the plate so that

$$T_{11}^1(x_1) = \cos(\pi x_1/L) = \cos \pi x,$$

and assume that the electrode band covers the interval $|x| \leq x_0$. We must choose x_0 according to (45),

$$2a \int_0^{x_0} \cos \pi x dx = 2a(2x_0 \cos \pi x_0),$$

or

$$\tan \pi x_0 = 2\pi x_0,$$

which is satisfied by $x_0 = 0.371$, which is the value found also by Bechmann and Parsons¹ in this case, i.e., 74.2 per cent of the bar length covered by electrodes. This maximizes the capacitance ratio C_1/C_s for the fundamental mode. The capacitance ratio for overtone modes may be maximized similarly, since the overtone modes may be imagined to be made up of a set of fundamental modes for n bars, $1/n$ in length, set end to end.

For the fundamental mode, this maximum capacitance ratio is $0.909k^2$ compared with the values $0.811k^2$ and $0.786k^2$ for a completely covered bar and for a bar with sinusoidal plating which suppresses all overtones. Thus we may increase the capacitance ratio about 10 per cent over its value for the wholly covered bar by the above technique.

A difficulty which we do not encounter with the fundamental mode shows up when we consider the excitation of the first overtone ($n = 2$).

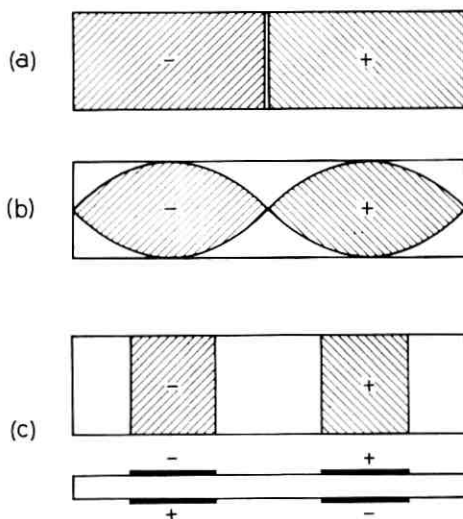


Fig. 7 — Various electrode configurations for excitation of first overtone in crystal bar.

In order to excite this mode, we must use electrodes having odd polarity about the middle of the plate, as shown in Fig. 7. For example, we may use complete (split) electrodes, as in Fig. 7(a), sinusoidal electrodes, as in Fig. 7(b), or band electrodes, as in Fig. 7(c). Especially in the first case, there will be a substantial contribution to the total static capacitance from the capacitance between adjoining electrodes of opposite polarity. This contribution has been completely neglected in our previous calculations. To make it small compared with the ordinary plate capacitance, proportional to electrode area, the spacing between adjoining electrodes of opposite polarity should be large compared with the plate thickness. This consideration makes the band electrodes [Fig. 7(c)], which maximize the capacitance ratio, particularly attractive.

The presence of a substantial electrical field component, parallel to the bar faces, also changes the form of the function p_n and thus the motional capacitance. At high frequencies, with a complicated resonance spectrum, this field component may excite unwanted resonances. In the present case of low-frequency longitudinal vibrations, this effect is probably not important.

X. ACKNOWLEDGMENT

The author would like to express his debt to W. P. Mason, at whose suggestion this work was initiated, and to Professor R. D. Mindlin of Columbia University for many illuminating discussions.

REFERENCES

1. Bechmann, R., and Parsons, P. L., General Post Office Engineering Report No. 9, in *Piezoelectricity*, Her Majesty's Stationery Office, London, 1957, p. 293.
2. Vorner, J. J., Crystal-Plates Without Overtones, Proc. I.R.E., **39**, 1951, p. 1086.
3. Love, A. E. H., *The Mathematical Theory of Elasticity*, 4th ed., Dover Publications, New York, 1944.
4. Mindlin, R. D., and Deresiewicz, H., Thickness-Shear Vibrations of Piezoelectric Crystal Plates with Incomplete Electrodes, J. Appl. Phys., **25**, 1954, pp. 21-24, 25-27.
5. U.S. Army Signal Corps, *Investigations in the Mathematical Theory of Vibrations of Anisotropic Bodies*, Final Report, 1956, Contract DA-36-039 sc-64687.
6. U.S. Army Signal Corps, *An Introduction to the Mathematical Theory of Vibrations of Elastic Plates*, Final Report, 1955, Contract DA-36039 sc-56772.
7. Van Dyke, K. S., Proc. 10th Ann. Symp. on Frequency Control, Signal Corps Engineering Laboratories, Fort Monmouth, N. J., 1956, pp. 1-9.
8. Koga, I., Fukuyo, H., and Rhodes, J. E., Modes of Vibration of Quartz Crystal Resonators Investigated by Means of the Proper Method, Proc. 13th Ann. Symp. on Frequency Control, Signal Corps Engineering Laboratories, Fort Monmouth, N. J., 1959, pp. 54-70.
9. Jeffries, H., *Cartesian Tensors*, Cambridge Univ. Press, Cambridge, 1931.
10. Mason, W. P., First and Second Order Equations for Piezoelectric Crystals Expressed in Tensor Form, B.S.T.J., **26**, 1947, p. 80.

Asymptotic Behavior of General Queues with One Server

By V. E. BENEŠ

(Manuscript received January 12, 1961)

The asymptotic behavior of the virtual delay in a single-server queue with order of arrival service is studied, under no restrictions on the stochastic nature of the arrival process and the service times, except for a weak stationarity condition. This behavior is shown to be governed by a functional equation closely analogous to the "fundamental equation" of branching processes, already used in special queueing models. A generalization of the Pollaczek-Khinchin delay formula is derived for the case in which delays do not build up.

I. INTRODUCTION

There is a queue in front of one server; waiting customers are served in order of arrival, and no defections from the queue occur. The delays incurred in such a system can be described by a stochastic process $\{W(t), t \geq 0\}$, the virtual waiting time, defined as the time a customer would have to wait for service if he arrived at time t . Stochastic processes of this kind have been studied.^{1,2}

In the present work we examine the asymptotic behavior of the probabilities $\Pr\{W(t) \leq w\}$, as t becomes large, for a class of processes satisfying a weak condition of stationarity.

It is customary to define the process $W(\cdot)$ in terms of the arrival epoch t_k and the service time S_k of the k th arriving customer, for $k = 1, 2, \dots$. We can, however, describe the service times and the arrival epochs simultaneously by a single function $K(\cdot)$, defined for $t \geq 0$, left-continuous, nondecreasing, and constant between successive jumps. The locations of the jumps are the epochs of arrivals, and the magnitudes are the service times. With $K(0) = W(0)$ for convenience, $K(t)$ can be interpreted as the "work load" submitted to the server during $[0, t)$, and $W(t)$ as the work remaining to be done at t .

Formally, $W(\cdot)$ is defined in terms of $K(u)$ for $0 \leq u < t$ by the equation

$$W(t) = K(t) - t + \int_0^t U(-W(u)) du, \quad (1)$$

where $U(\cdot)$ is the unit step function, i.e., $U(x) = 1$ for $x \geq 0$, and $U(x) = 0$ otherwise. Equation (1) may be interpreted as follows:

$$\begin{aligned} & \text{work remaining at } t = \text{load offered up to } t \\ & \quad - \text{elapsed time} \\ & \quad + \text{total time that server was idle} \\ & \quad \quad \text{in } (0, t). \end{aligned} \quad (2)$$

The relationship of $W(\cdot)$ and $K(\cdot)$ has previously been described and illustrated.^{1,2} A principal result of the latter paper (and the only one needed for reading the present one) was a formula for the distribution of $W(t)$ in terms of the functions

$$\Pr\{K(t) \leq u\},$$

$$R(t, u, w) = \Pr\{K(t) - K(u) - t + u \leq w \mid W(u) = 0\}.$$

The reference to $W(\cdot)$ in the condition is not circular because, as shown,² the events

$$\{W(u) = 0\}, \quad \left\{ \sup_{0 < y < u} [K(u) - K(y) - u + y] \leq 0 \right\}$$

differ by at most a set of measure zero. The formula for $\text{distr}\{W(t)\}$ is

$$\begin{aligned} \Pr\{W(t) \leq w\} &= \Pr\{K(t) - t \leq w\} \\ &\quad - \frac{\partial}{\partial w} \int_0^t R(t, u, w) \Pr\{W(u) = 0\} du, \quad w > 0, \end{aligned} \quad (3)$$

or, in an integrated form,

$$\begin{aligned} \int_0^w \Pr\{K(t) \leq u\} du &= \int_0^{t+w} \Pr\{K(t) \leq u\} du \\ &\quad - \int_0^t R(t, u, w) \Pr\{W(u) = 0\} du. \end{aligned} \quad (4)$$

For $-t \leq w \leq 0$, the chance $\Pr\{W(u) = 0\}$ that the server be idle at time u satisfies the Volterra equation

$$\begin{aligned} \int_0^{t+w} R(t, u, w) \Pr\{W(u) = 0\} du &= \int_0^{t+w} \Pr\{K(t) \leq u\} du, \\ &= E\{\max[0, w - K(t) + t]\}. \end{aligned} \quad (5)$$

These results can be cast into a form more useful for the present endeavor by use of Laplace-Stieltjes transforms. Theorem 5 of the author's paper² implies that for $\text{Re}(s) > 0$

$$E\{e^{-sW(t)}\} = E\{e^{-sK(t)+st}\} - s \int_0^t E\{e^{-s[K(t)-K(u)-t+u]} | W(u) = 0\} \Pr\{W(u) = 0\} du. \quad (6)$$

This formula for the Laplace-Stieltjes transform of the distribution of $W(t)$ can be obtained directly from (4) and (5) by taking the Laplace-Stieltjes transform with respect to w .

II. BASIC ASSUMPTIONS

Unless the input process $K(\cdot)$ has some asymptotic temporal uniformity, one cannot expect $W(t)$ to have a limiting distribution as $t \rightarrow \infty$. We shall assume that the increments of $K(\cdot)$ are *weakly stationary* in the sense³ that the basic kernel $R(t, u, w)$ of (3) depends only on $(t - u)$ and w , i.e.,

$$\Pr\{K(t) - K(u) - t + u \leq w | W(u) = 0\} = R(t - u, w). \quad (7)$$

The word "weak" refers to the dependence of the kernel only on the difference of its two temporal arguments.

The stationarity condition (7) turns (5) into an equation of *convolution* type. Our strategy will be to work with (5), find conditions under which $\Pr\{W(u) = 0\}$ converges as $u \rightarrow \infty$, and then use Abelian arguments on (4) and (6) to obtain the asymptotic behavior of

$$\Pr\{W(t) \leq w\} \quad \text{as } t \rightarrow \infty.$$

A similar approach has been used in two prior papers by the author.^{3,4}

We shall only consider the case in which

$$W(0) = 0, \quad (8)$$

with probability one. This "start empty" condition is in a sense unavoidable in practice, and it leads to simple, elegant results. Equations (7) and (8) imply that for $t, y \geq 0$

$$\begin{aligned} E\{e^{-sK(t)+st}\} &= E\{e^{-sK(t)+st} | W(0) = 0\} \\ &= E\{e^{-sK(t+y)+sK(y)+sy} | W(y) = 0\} \\ &= \int_0^\infty e^{-sw} R(t, dw). \end{aligned} \quad (9)$$

The condition $\{W(\cdot) = 0\}$ will therefore be omitted from now on in all probabilities and expectations.

The following functions occur frequently in the discussion and merit abbreviations:

$$\begin{aligned} E\{\max [0, t - K(t)]\} &= F(t), \\ \Pr\{K(t) \leq t\} &= R(t, 0) = R(t), \\ \Pr\{W(t) = 0\} &= P(t). \end{aligned}$$

In this notation, the Volterra equation (5) for $w = 0$ can be restated as

$$F(t) = \int_0^t R(t-u)P(u) du. \quad (10)$$

The Laplace-Stieltjes transform of a function is denoted by the corresponding *lower-case* Greek letter; e.g.,

$$\varphi(s) = \int_0^\infty e^{-st} dF(t), \quad \operatorname{Re}(s) > 0.$$

The ordinary Laplace transform is denoted by the corresponding *capital* Greek letter; e.g.,

$$\Pi(\tau) = \int_0^\infty e^{-\tau t} P(t) dt.$$

From (1) defining $W(\cdot)$ it is evident that the mean delay is

$$E\{W(t)\} = E\{K(t)\} - t + \int_0^t \Pr\{W(u) = 0\} du.$$

Also, we have

$$\begin{aligned} E\{K(t)\} &= \int_0^\infty \Pr\{K(t) > u\} du \\ &= \int_0^t [1 - \Pr\{K(t) \leq u\}] du + \int_t^\infty \Pr\{K(t) > u\} du \\ &= t - F(t) + \int_t^\infty \Pr\{K(t) > u\} du, \end{aligned}$$

since

$$E\{\max(0, K(t) - t)\} = \int_0^t \Pr\{K(t) \leq u\} du.$$

In addition to the stationarity condition (7) we use the following

assumptions (since some theorems do not depend on all the assumptions to be listed, the relevant hypotheses are repeated when a result is stated):

- i. $\limsup_{t \rightarrow \infty} t^{-1} E\{K(t)\} < \infty$;
- ii. $\liminf_{t \rightarrow \infty} t^{-1} \log \Pr\{K(t) = 0\} > -\infty$;
- iii. There exists a neighborhood N of the positive real axis and a $T > 0$ such that $E\{e^{-sK(t)}\} \neq 0$ for $s \in N$ and $t > T$.

Hypothesis iii will only be needed when the system is recurrent-null; i.e., the service-factor is unity.

III. SUMMARY

Basic assumptions and preliminary lemmata occupy Sections II and IV, respectively, the principal result being the existence of a solution $s(\tau)$ of the functional equation

$$\tau - s(\tau) = a(s(\tau)), \quad \tau \geq 0,$$

where, for real $s \geq 0$,

$$a(s) = \limsup_{t \rightarrow \infty} t^{-1} \log E\{e^{-sK(t)} \mid W(0) = 0\}.$$

This functional equation is closely analogous to the "fundamental equation" of branching processes. In Section V it is shown that the Laplace transform of $\Pr\{W(\cdot) = 0\}$ is just $1/s(\tau)$, whence it follows (Section VI) that the asymptotic value of $\Pr\{W(t) = 0\}$ as $t \rightarrow \infty$ (if any) can be determined from the limit of $a(s)/s$ as $s \rightarrow 0$.

The existence of a limiting value $\Pr\{W(\infty) = 0\}$ is investigated in Sections VII and VIII, by Mercerian and Tauberian methods respectively, applied to (5), for $w = 0$. Section IX contains limit theorems for $\Pr\{W(\cdot) \leq w\}$, $w > 0$; these are obtained readily by Abelian methods from (4) or (6), once the convergence of $\Pr\{W(t) = 0\}$ as $t \rightarrow \infty$ has been ascertained.

IV. PRELIMINARY RESULTS

Let $s \geq 0$, and define $a(s)$ to be the abscissa of convergence of the Laplace transform*

$$\Psi(\tau, s) = \int_0^{\infty} e^{-\tau t} E\{e^{-sK(t)}\} dt.$$

Obviously $a(0) = 0$, $a(s) \leq 0$ for $s \geq 0$, and $a(\cdot)$ is monotone decreasing in $s \geq 0$. It is not obvious that $a(s) \neq -\infty$, at present.

* We assume throughout that $E\{e^{-sK(t)}\}$ is a measurable function of t for s -values under discussion.

Lemma 1: The abscissa $a(s)$ of convergence of the integral

$$\Psi(\tau, s) = \int_0^{\infty} e^{-\tau t} E\{e^{-sK(t)}\} dt \quad (11)$$

is given by the formula

$$a(s) = \limsup_{t \rightarrow \infty} t^{-1} \log E\{e^{-sK(t)}\}, \quad s \geq 0.$$

Proof: Let $s > 0$ be fixed and set $E\{e^{-sK(t)}\} = \psi(t)$. The function $\psi(\cdot)$ is monotone decreasing, and satisfies $0 \leq \psi \leq 1$. Hence, by Theorem 2.4d of Widder,⁵ the abscissa of convergence of the integral

$$\int_0^{\infty} e^{-\tau t} d\psi(t) \quad (12)$$

is

$$\limsup_{t \rightarrow \infty} t^{-1} \log |\psi(t)| \leq 0.$$

However,

$$\int_{0-}^T e^{-\tau t} d\psi(t) = e^{-\tau T} \psi(T) + \tau \int_0^T e^{-\tau t} \psi(t) dt,$$

so that (11) and (12) have the same abscissa of convergence.

Lemma 2: For $s, h > 0$ and $u \geq 0$

$$E\{e^{-sK(u)}\} E\{e^{-hK(u)}\} \leq E\{e^{-(s+h)K(u)}\}.$$

Proof: Let $K(u) = \xi$. Since x^{s+h} is convex in x^s and in x^h , it follows from Jensen's theorem that

$$0 \leq E\{e^{-s\xi}\} \leq E^{s/(s+h)}\{e^{-(s+h)\xi}\},$$

$$0 \leq E\{e^{-h\xi}\} \leq E^{h/(s+h)}\{e^{-(s+h)\xi}\}.$$

Multiplying these inequalities together gives the result.

Lemma 3: If

$$\limsup_{t \rightarrow \infty} E\{K(u)/u\} < \beta,$$

then, given $\epsilon > 0$, there exists h so small and t so large that

$$E\{e^{-hK(u)}\} \geq e^{-\epsilon u}$$

for all $u > t$.

Proof: Choose h so that $h\beta < \epsilon$, and t so that

$$E\{K(u)\} < \beta u$$

for all $u > t$. Then, by Jensen's theorem, for $u > t$,

$$E\{e^{-hK(u)}\} \geq e^{-hE\{K(u)\}} \geq e^{-\epsilon u}.$$

Lemma 4: If

$$\limsup_{t \rightarrow \infty} E\{K(u)/u\} < \infty,$$

then for given $\epsilon > 0$ and any $s > 0$, there exist h and t so that

$$t^{-1} \log E\{e^{-sK(u)}\} - t^{-1} \log E\{e^{-(s+h)K(u)}\} < \epsilon$$

for all $u > t$.

Proof: Lemmata 2 and 3.

Lemma 5: $a(\cdot)$ is a continuous function of s , $s \geq 0$.

Proof: Set $t^{-1} \log E\{e^{-sK(t)}\} = g_s(t)$. Let $\epsilon > 0$ be given, and find, by Lemma 4, an h and a t so that $u > t$ implies

$$0 \leq g_s(u) - g_{s+h}(u) \leq \frac{\epsilon}{2}. \quad (13)$$

There exists a sequence $u_n \rightarrow \infty$ along which

$$a(s) - \frac{\epsilon}{4} \leq g_s(u_n),$$

$$g_{s+h}(u_n) \leq a(s+h) + \frac{\epsilon}{4},$$

for all n sufficiently large; this is because

$$a(s) = \limsup_{t \rightarrow \infty} g_s(t).$$

For such n , then, since $a(\cdot)$ is monotone decreasing and (13) holds,

$$0 \leq a(s) - a(s+h) \leq g_s(u_n) - g_{s+h}(u_n) + \frac{\epsilon}{2},$$

$$0 \leq a(s) - a(s+h) \leq \epsilon,$$

so that $a(\cdot)$ is right-continuous. A mirror image of this argument proves $a(\cdot)$ left-continuous.

Lemma 6: $a(\cdot)$ is convex in $s \geq 0$.

Proof: For each $t > 0$, $E\{e^{-sK(t)}\}$ is a completely monotonic function of s , and so (Widder,⁵ p. 167) is logarithmically convex; indeed, in $s \geq 0$ we have

$$\frac{1}{t} \frac{\partial^2}{\partial s^2} \log E\{e^{-sK(t)}\} = \frac{\partial^2}{\partial s^2} g_t(s) \geq 0.$$

This implies that for $s, h \geq 0$

$$\frac{1}{2}g_t(s) + \frac{1}{2}g_t(s+h) \geq g_t(s + \frac{1}{2}h).$$

Taking the lim sup as $t \rightarrow \infty$, we find that $a(\cdot)$ is convex.

Lemma 7: For $s, h > 0$,

$$\frac{a(s)}{s} \leq \frac{a(s+h)}{s+h}, \quad (14)$$

and

$$\rho = -\lim_{s \rightarrow 0} \frac{a(s)}{s} > 0$$

exists as a finite or infinite limit.

Proof: If ξ is a positive variate then

$$E^{1/r}\{\xi^r\}$$

is a nondecreasing function of r . (See Loève,⁶ p. 156.) Choosing $\xi = e^{-K(t)}$, we have

$$\begin{aligned} E^{1/s}\{e^{-sK(t)}\} &\leq E^{1/(s+h)}\{e^{-(s+h)K(t)}\}, \\ s^{-1}g_t(s) &\leq (s+h)^{-1}g_t(s+h). \end{aligned}$$

But

$$a(s) = \limsup_{t \rightarrow \infty} g_t(s),$$

so (14) is true.

Theorem 1: If

$$\liminf_{t \rightarrow \infty} t^{-1} \log \Pr\{K(t) = 0\} > -\infty \quad (16)$$

then for each $\tau \geq 0$ there is at least one number $s(\tau)$ such that

$$\tau - s(\tau) = a(s(\tau)),$$

and

$$\begin{cases} \tau - s \geq a(s) & \text{for } s \leq s(\tau). \\ \tau - s \leq a(s) & \text{for } s \geq s(\tau) \end{cases} \quad (17)$$

Proof: The hypothesis implies that there is a constant $c > 0$ such that

$$\Pr\{K(t) \leq 0\} \geq e^{-ct}$$

for all sufficiently large t . However, this implies that

$$E\{e^{-sK(t)}\} \geq e^{-ct}$$

for all t large enough, and so $a(s) \geq -c$ for all $s \geq 0$. Since $a(\cdot)$ is continuous it must intersect the line $\tau - s$ at least once, and the relations (17) follow from the convexity of $a(\cdot)$.

This section ends with some preliminary analytical results for $F(\cdot)$ and $P(\cdot)$.

Lemma 8: If $R(\cdot)$ is of bounded variation in every finite interval, then

$$F'(t) = d/dt F(t)$$

exists almost everywhere.

Proof: The hypothesis implies that the transform

$$\rho(s) = \int_{0-}^{\infty} e^{-st} dR(t)$$

exists in $\text{Re}(s) > 0$. From (10) we find

$$\varphi(s) = \rho(s) \int_0^{\infty} e^{-st} P(t) dt = \rho(s)\Pi(s),$$

but $\Pi(\cdot)$ is the Laplace-Stieltjes transform of an absolutely continuous function; hence $\varphi(\cdot)$ is also.

In order to use various Tauberian theorems it will usually be necessary to impose "Tauberian" conditions on the oscillations of $P(\cdot)$. We recall (Widder,⁵ p. 209) that a function $f(\cdot)$ is *slowly oscillating* if

$$\lim_{\substack{y-x \rightarrow 0 \\ y \rightarrow \infty}} f(y) - f(x) = 0,$$

and $f(\cdot)$ is *slowly decreasing* if

$$\liminf_{y-x \rightarrow 0} f(y) - f(x) \geq 0$$

as $x \rightarrow \infty$, $y = y(x) > x$.

Lemma 9: If F' is continuous uniformly in $t \geq 0$, and $R(\cdot)$ is of bounded total variation, with the form

$$R = U + H_1 + H_2$$

where U is the unit step at zero, H_1 is absolutely continuous, and

$$\int |dH_2| = \lambda < 1,$$

then $P(\cdot)$ is slowly oscillating.

*Proof:** Since F' exists we may differentiate (10) to find that

* The procedure of this proof is patterned after that of Karlin.⁷

$$F'(t) = \int_{0-}^t P(t-u) dR(u).$$

This is equivalent to the renewal equation

$$P(t) = F'(t) + \int_{0-}^t P(t-u) d_u\{U - R\},$$

or, iterating n times,

$$P(t) = \int_{0-}^t P(t-u) d_u\{L + M\} + \int_{0-}^t F'(t-u) dN(u),$$

where L is absolutely continuous,

$$\int |dM| < \lambda^n, \text{ and } \int |dN| < \infty.$$

Then

$$\begin{aligned} |P(t+\epsilon) - P(t)| &\leq \int_{-\infty}^{\infty} |L'(t+\epsilon) - L'(t)| dt + \lambda^n \\ &\quad + \int_t^{t+\epsilon} |F'(t+\epsilon-u)| |dN(u)| \\ &\quad + \int_0^{\infty} |F'(t+\epsilon-u) - F'(t-u)| |dN(u)|. \end{aligned}$$

The first term goes to zero with $\epsilon \rightarrow 0$, by a known result of Lebesgue (Wiener,⁸ p. 14). The second term vanishes as n increases. The third term approaches zero as $t \rightarrow \infty$, since F' is bounded. The fourth term goes to zero with $\epsilon \rightarrow 0$ by the uniform continuity of F' .^{*} Hence

$$\lim_{\substack{\epsilon \rightarrow 0 \\ t \rightarrow \infty}} |P(t+\epsilon) - P(t)| = 0.$$

The hypothesis that F' is continuous can be replaced by the condition that $F'(t)$ approach a limit as $t \rightarrow \infty$. Also the uniform continuity of F' could be replaced by a weaker but more complicated asymptotic condition.

V. THE TRANSFORM $\Pi(\cdot)$ OF $\Pr\{W(\cdot) = 0\}$

Theorem 2: Let s be real and positive, and let

$$\int_0^{\infty} e^{-\tau t} P(t) dt = \Pi(\tau);$$

^{*} It is assumed that $F'(t) = 0$ for $t < 0$.

then (16) implies

$$s\Pi(s + a(s)) \equiv 1, \quad s > 0. \quad (18)$$

Proof: Equation (6) implies that

$$\int_0^{\infty} e^{-\tau t} E\{e^{-sW(t)}\} dt = \Psi(\tau, s)[1 - s\Pi(\tau)].$$

The left-hand side has no singularities in $\text{Re}(\tau) > 0$. But by Widder⁵ (p. 58, Theorem 5b), $\Psi(\tau, s)$ is singular at $\tau = s + a(s)$, and so (18) holds.

Theorem 3: If $\tau > 0$, and (16) holds, then

$$\Pi(\tau) = \frac{1}{s(\tau)},$$

where $s(\tau)$ satisfies (and is determined uniquely by) the equation

$$\tau - s(\tau) = a(s(\tau)). \quad (19)$$

Proof: For given τ , there exists at least one $s(\tau) > 0$ satisfying (19). By Theorem 2, any such $s(\tau)$ has the property

$$\begin{aligned} \frac{1}{s(\tau)} &= \Pi(s(\tau) + a(s(\tau))), \\ &= \Pi(\tau), \end{aligned}$$

because of (18). Since $\Pi(\tau)$ is a strictly monotone function of real τ , being of the form

$$\Pi(\tau) = \int_0^{\infty} e^{-\tau t} P(t) dt, \quad P(\cdot) \geq 0,$$

the solution $s(\tau)$ of (19) is unique, and strictly monotone increasing in τ .

When $-a(s)/s \rightarrow 1$ as $s \rightarrow 0$, we need some additional properties of $a(\cdot)$. For each t , there is a neighborhood N_t of the positive real axis in which

$$E\{e^{-sK(t)}\}$$

has no zeros. However, we shall need the condition iii of Section II that there exist one neighborhood N of the positive real axis in which

$$E\{e^{-sK(t)}\} \neq 0$$

for all t sufficiently large, say $t > T$. Define, for $z \in N$,

$$g_z(t) = t^{-1} \log |E\{e^{-zK(t)}\}|.$$

Then $g_{(\cdot)}(t)$ is harmonic in N for $t > T$, and

$$\begin{aligned} |g_z(t)| &= |t^{-1} \log |E\{e^{-zK(t)}\}|| \\ &\leq t^{-1} |\log \Pr\{K(t) = 0\}| < \infty, \end{aligned}$$

so that $|g_z(t)| \leq c$ for $z \in N$, where c is the constant of Theorem 1.

Theorem 4: If

$$\liminf_{t \rightarrow \infty} t^{-1} \log \Pr\{K(t) = 0\} > -\infty \quad (16)$$

and N is a neighborhood of the positive real axis such that

$$E\{e^{-zK(t)}\} \neq 0, \quad z \in N$$

for all sufficiently large t , then $a(\cdot)$ can be extended to be harmonic in N .

Proof: For $t > T$ define a class V_t of functions on N as follows: V_t is the smallest class containing all $g_{(\cdot)}(t')$ for $t' > t$, and closed under the operations

$$(a) \quad v_1, v_2 \rightarrow \max(v_1, v_2),$$

$$(b) \quad v \rightarrow v_{\text{mod}} = \begin{cases} v & \text{outside } \Delta \\ Pv & \text{inside } \Delta \end{cases}$$

where Δ is a disc with $\bar{\Delta} \subset N$, and P is the Poisson integral operator on the disc Δ . These operations preserve the property of being bounded by the constant c of Theorem 1. Now let

$$u_t(z) = \sup_{v \in V_t} v(z).$$

By a standard argument (Ahlfors,⁹ p. 197), the function $u_t(\cdot)$ is harmonic in N , and clearly

$$u_t(z) \geq \sup_{t_1 > t} g_z(t_1).$$

We prove the reverse inequality:

$$u_t(z) \leq \sup_{t_1 > t} g_z(t_1).$$

It is enough to show that the operations (a) and (b) preserve this property. For (a) this is obvious; for (b), we observe that

$$Pv(z) = \frac{1}{2\pi} \int_0^{2\pi} v(z + re^{i\theta}) d\theta \leq \sup_{t_1 > t} g_z(t_1).$$

Hence for real z

$$\begin{aligned} a(z) &= \inf_t \sup_{t_1 > t} g_z(t_1), \\ &= \lim_{t \rightarrow \infty} u_t(z). \end{aligned}$$

The functions $u_t(\cdot)$ are harmonic and monotone decreasing in t . Hence by Harnack's principle they either tend to $-\infty$ or to a harmonic function. The first alternative is ruled out by the inequality $|g_z(t)| \leq c$. Hence $a(\cdot)$ is harmonic in N .

Theorem 5: Under the conditions of Theorem 4, let

$$\rho = \lim_{s \rightarrow 0} -\frac{a(s)}{s}.$$

If $\rho > 1$, there exists a largest root $\zeta > 0$ of the equation $\zeta = a(\zeta)$, and $s(\tau) \rightarrow \zeta$ as $\tau \rightarrow 0$. If $\rho \leq 1$, then $s(\tau) \rightarrow 0$ as $\tau \rightarrow 0$.

Proof: $s(\cdot)$ is monotone, nonnegative, and nonincreasing, so $s(0+) \geq 0$ exists. If $\rho > 1$, then $a(x) < -x$ for some $x > 0$. However, for $\tau > 0$,

$$a(s(\tau)) = \tau - s(\tau) > -s(\tau),$$

so $s(0+) > x > 0$. Since $s(\cdot)$ is continuous we may let $\tau \rightarrow 0$ to find $\zeta = a(\zeta) = s(0+) > 0$.

Next, if $\rho < 1$, suppose that $s(0+) > 0$. Then

$$\frac{a(s(0+))}{s(0+)} = -1,$$

which is impossible, since

$$\frac{a(s)}{s} \rightarrow -\rho > -1 \quad \text{as } s \rightarrow 0,$$

in a monotone decreasing manner.

The case $\rho = 1$ requires a special argument. If $s(0+) > 0$, then again

$$\frac{a(s(0+))}{s(0+)} = -1$$

and we must have

$$\frac{a(s)}{s} \equiv 1, \quad 0 \leq s \leq s(0+).$$

However, since $a(\cdot)$ is harmonic in a neighborhood N of the positive

real axis, this would imply, by the Cauchy-Riemann relations, that $a(s) = s$ throughout N ; but $a(s) \geq -c$ for real s . Hence $s(0+) = 0$.

The hypothesis about the existence of the neighborhood N is needed only for the case $\rho = 1$.

Theorem 6: Suppose that $P(\infty) = \lim P(t)$ as $t \rightarrow \infty$ exists. Then

$$P(\infty) = \begin{cases} 0 & \text{if } \rho \geq 1, \\ 1 - \rho & \text{if } \rho < 1. \end{cases}$$

Proof: Since $\tau - s(\tau) = a(s(\tau))$, we find that

$$\tau \Pi(\tau) = \frac{\tau}{s(\tau)} = 1 - \frac{a(s(\tau))}{s(\tau)}.$$

Then if $\rho \leq 1$, the result follows from the previous theorem and a standard Abelian theorem for the Laplace transform. If $\rho > 1$ then $P(\cdot) \in L$, and if $P(\infty)$ exists it must be zero.

VI. ASYMPTOTIC RELATIONSHIPS

We now prove some lemmata that exhibit some of the basic asymptotic relationships between ρ , $a(\cdot)$, $F(\cdot)$, $R(\cdot)$, etc.

Lemma 10:

$$1 - \rho = 1 + a'(0) \geq \limsup_{t \rightarrow \infty} 1 - E\{K(t)/t\}.$$

Proof: By Jensen's theorem

$$\begin{aligned} E\{e^{-sK(t)}\} &\geq e^{-sE\{K(t)\}}, \\ t^{-1} \log E\{e^{-sK(t)}\} &\geq -sE\{K(t)/t\}. \end{aligned}$$

Definition:

$$Q(t) = \int_t^\infty \Pr\{K(t) > u\} du = E\{\max[0, K(t) - t]\}$$

Lemma 11: If $Q(t) = o(t)$ as $t \rightarrow \infty$, then

- i. $\rho \leq 1$;
- ii. If $d/dt E\{K(t)\}$ also exists and approaches a limit as $t \rightarrow \infty$, and if $Q(\cdot)$ can be differentiated in the usual way, then $R(t) \rightarrow 1$ as $t \rightarrow \infty$.

Proof:

$$0 \leq t^{-1}F(t) \leq 1$$

and

$$t^{-1}F(t) = 1 - E\{K(t)/t\} + Q(t)/t,$$

so that

$$\liminf_{t \rightarrow \infty} t^{-1}F(t) = \liminf_{t \rightarrow \infty} 1 - E\{K(t)/t\}.$$

Then i follows by the previous lemma. To prove ii note that

$$F'(t) = 1 - \frac{d}{dt} E\{K(t)\} + 1 - R(t) + \int_t^\infty \frac{\partial}{\partial t} \Pr\{K(t) > u\} du.$$

The last term is positive, so

$$R(t) \geq 1 + 1 - \frac{d}{dt} E\{K(t)\} - F'(t),$$

$$R(t) \rightarrow 1.$$

Lemma 12: If $R(t) \rightarrow R(\infty) > 0$ as $t \rightarrow \infty$, then $\rho \leq 1$.

Proof:

$$\begin{aligned} E\{e^{-sK(t)+st}\} &\geq E\{e^{s \min\{0, t-K(t)\}}\}, \\ &\geq \Pr\{K(t) \leq t\}. \end{aligned}$$

So

$$E\{e^{-sK(t)}\} \geq R(t)e^{-st}.$$

Hence also

$$1 + \frac{a(s)}{s} \geq \frac{1}{s} \limsup_{t \rightarrow \infty} t^{-1} \log R(t).$$

If $R(\infty) > 0$, the lim sup is zero, so that

$$1 + \frac{a(s)}{s} \geq 0, \quad \rho \leq 1.$$

Lemma 13: If

$$\limsup_{t \rightarrow \infty} t^{-1} \log R(t) = 0,$$

then $\rho \leq 1$. If this lim sup is less than 0, then $R(\cdot)$ is integrable and $R(\infty) = 0$. Hence also $P(\infty) = 0$ and $\rho \geq 1$.

Lemma 14: If

$$\liminf_{t \rightarrow \infty} Q(t) = 0,$$

then $\rho \leq 1$.

Proof:

$$E\{\min [0, t - K(t)]\} = E\{t - K(t)\} - F(t),$$

so

$$\begin{aligned} 1 + \frac{a(s)}{s} &\geq \limsup_{t \rightarrow \infty} \left[1 - \frac{E\{K(t)\}}{t} - F(t) \right], \\ &\geq \limsup_{t \rightarrow \infty} [-Q(t)], \\ &\geq - \liminf_{t \rightarrow \infty} Q(t). \end{aligned}$$

VII. CONVERGENCE OF $\Pr\{W(t) = 0\}$: MERCERIAN METHODS

A Mercerian theorem (see Pitt,¹⁰ p. 94) is one which, for example, enables us to study the asymptotic behavior of $P(\cdot)$ directly from that of the convolution

$$F(t) = \int_0^t P(t-u)R(u) du,$$

i.e., that of $F(\cdot)$, without the intervention of "Tauberian" hypotheses on the oscillations of $P(\cdot)$. Tauberian methods usually require $P(\cdot)$ to be a slowly decreasing function; such methods are considered in Section VIII.

In what follows the norm symbol (when used) refers to the total variation of a function of bounded total variation. The subscripts "dis" and "sing" (applied to a function symbol) denote the discontinuous and singular components, respectively.

Theorem 7: If $R(\cdot)$ is of bounded variation in any finite interval, and for some $\sigma \geq 0$,

$$\inf_{\tau} |\rho(\sigma - i\tau)| > 0, \quad (20)$$

$$\| (R_{\sigma})_{\text{sing}} \| < \inf_{\tau} |\rho_{\text{dis}}(\sigma - i\tau)|, \quad (21)$$

then $P(\cdot)$ can be represented by the inversion formula

$$\int_{-\infty}^t F'(t-u)e^{\sigma u} dG_{\sigma}(u), \quad \int_{-\infty}^{\infty} |dG_{\sigma}| < \infty,$$

provided that

$$R_{\sigma}(t) = \int_{-\infty}^t e^{-\sigma u} dR(u) \quad (22)$$

is of bounded total variation.

Proof: The hypotheses (20) through (22), together with Theorem 1 of Wiener and Pitt,¹¹ imply that there exists a function $G_\sigma(\cdot)$ of bounded total variation whose Fourier-Stieltjes transform is

$$[\rho(\sigma - i\tau)]^{-1}, \quad \sigma \text{ fixed.}$$

We then solve the equation

$$e^{-\sigma t}F(t) = \int_0^t e^{-\sigma(t-u)}P(t-u)e^{-\sigma u}R(u) du$$

by Laplace transforms, obtaining

$$\int_0^\infty e^{-(p+\sigma)t}P(t) dt = \frac{\varphi(p + \sigma)}{\rho(p + \sigma)}.$$

However, by Lemma 8, F' exists and has Laplace transform $\varphi(\cdot)$. Therefore

$$e^{-\sigma t}P(t) = \int_{-\infty}^t F'(t-u)e^{-\sigma(t-u)} dG_\sigma(u),$$

and division by $e^{-\sigma t}$ completes the proof.

This result provides another way of proving that $P(\cdot)$ is slowly oscillating, as follows:

Lemma 15: If the hypotheses of Theorem 7 are true with $\sigma = 0$, and if $F'(\cdot)$ is uniformly continuous and bounded, then $P(\cdot)$ is slowly oscillating.

Proof: Let b be a bound on F' . Then

$$\begin{aligned} |P(t + \epsilon) - P(t)| &\leq \int_{-\infty}^\infty |F'(t + \epsilon - u) - F'(t - u)| |dG_0(u)| \\ &\leq 2b \int_{|u| > T} |dG_0(u)| + \int_{-T}^T |F'(t + \epsilon - u) \\ &\quad - F'(t - u)| |dG_0(u)|. \end{aligned}$$

Choose first T large, then ϵ small, using the uniform continuity of F' to let $t \rightarrow \infty$.

Theorem 8: If the hypotheses of Theorem 7 are true with $\sigma = 0$, if F' is bounded below, and for some constant A

$$\varphi(s) \sim \frac{A}{s} \quad \text{as } s \rightarrow 0+, \tag{23}$$

then

$$P(t) \xrightarrow{(c,1)} \frac{A}{R(\infty)} \quad \text{as } t \rightarrow \infty.$$

Proof: The hypotheses imply that $P(\cdot)$ can be written as

$$P(u) = \int_{-\infty}^u F'(u-y) dG_0(y)$$

with $G_0(\cdot)$ of bounded total variation; then also

$$t^{-1} \int_0^t P(u) du = \int_0^t \frac{F(t-u)}{t-u} \frac{t-u}{t} dG_0(u).$$

There exists a constant $b > 0$ such that

$$F(t) + bt$$

is nondecreasing in $t \geq 0$; hence by a known Tauberian result (Widder,⁵ p. 197, Theorem 4.6), condition (23) implies that

$$\frac{F(t)}{t} \rightarrow A \geq 0 \quad \text{as } t \rightarrow \infty,$$

the convergence being bounded, since $0 \leq F(t) \leq t$. It follows that

$$\begin{aligned} t^{-1} \int_0^t P(u) du &\rightarrow A[G_0(+\infty) - G_0(-\infty)] \\ &\rightarrow \frac{A}{R(\infty)}. \end{aligned}$$

The condition (20) of Theorem 7 guarantees that $R(\infty) > 0$.

Theorem 9: If the hypotheses of Theorem 7 are true with $\sigma = 0$, and $F' \rightarrow F'(\infty)$ boundedly as $t \rightarrow \infty$, then

$$P(t) \rightarrow \frac{F'(\infty)}{R(\infty)} \quad \text{as } t \rightarrow \infty.$$

Proof: The hypotheses imply that

$$P(t) = \int_{-\infty}^t F'(t-u) dG_0(u),$$

where $G_0(\cdot)$ is of bounded total variation. The result then follows from a known Abelian lemma, e.g., Lemma 1 of Smith.¹²

Theorem 10: If the function $R(\cdot) - U(\cdot)$ is absolutely continuous, if $F'(\infty)$ exists, and if $\rho(s)$ has no zeros in $\text{Re}(s) \geq 0$, then

$$P(t) \rightarrow \frac{F'(\infty)}{R(\infty)} \quad \text{as } t \rightarrow \infty.$$

Proof: Set $[R(\cdot) - U(\cdot)]' = k(\cdot)$, so that (10) implies (by differentiation) that

$$\begin{aligned} F'(t) &= \int_{0-}^{\infty} P(t-u) dR(u) = P(t) + \int_0^t P(t-u) d_u(R-U) \\ &= P(t) + \int_0^t P(t-u)k(u) du. \end{aligned}$$

Then, as $t \rightarrow \infty$,

$$P(t) + \int_0^t P(t-u)k(u) du \rightarrow F'(\infty).$$

By Theorem XVII of Paley and Wiener,¹³ this, together with the conditions that $P(\cdot)$ be bounded and $\rho(s) \neq 0$ in $\text{Re}(s) \geq 0$, implies that, as $t \rightarrow \infty$,

$$P(t) \rightarrow \frac{F'(\infty)}{1 + \int_0^{\infty} k(u) du} = \frac{F'(\infty)}{R(\infty)},$$

since for $t > 0$

$$R(t) = 1 + \int_0^t k(u) du.$$

By a theorem of Pitt¹⁰ (p. 115), the restriction that $\rho(s) \neq 0$ for $\text{Re}(s) \geq 0$ can be weakened to $\rho(i\tau) \neq 0$ for real τ .

VIII. CONVERGENCE OF $\text{Pr}\{W(t) = 0\}$: TAUBERIAN METHODS

The addition of "Tauberian" conditions on $P(\cdot)$,* such as the property of slow decrease, makes it possible to change or weaken the hypotheses on $R(\cdot)$ and $F(\cdot)$ necessary to ensure the convergence of $P(\cdot)$. The next result shows how the convergence of $F'(\cdot)$ in Theorem 9 can be relaxed.

Theorem 11: If

i. For some integrable function $k(\cdot)$ with a nonvanishing Fourier transform $K(\cdot)$, and some number A ,

$$\int_0^{\infty} k(t-u) dF(u) \rightarrow A \int_{-\infty}^{\infty} k(u) du \quad \text{as } t \rightarrow \infty;$$

- ii. $P(\cdot)$ is slowly decreasing;
 iii. $R(\cdot)$ is of bounded total variation with

* Such conditions were briefly studied in Section IV.

$$\begin{aligned} \rho(-i\tau) &\neq 0, \\ \|R_{\text{dis}}\| &< \inf_{\tau} |\rho_{\text{dis}}(-i\tau)|; \end{aligned}$$

then

$$P(t) \rightarrow \frac{A}{R(\infty)} \quad \text{as } t \rightarrow \infty.$$

Proof: The conditions on $\rho(\cdot)$ imply, by Theorem 1 of Wiener and Pitt,¹¹ that

$$\frac{K(\tau)}{\rho(-i\tau)}$$

is the nonvanishing Fourier transform of an integrable function $x(\cdot)$ such that

$$\int_0^{\infty} k(t-u) dF(u) = \int_0^{\infty} x(t-u)P(u) du.$$

The theorem to be proved then follows from Pitt's form of Wiener's fundamental Tauberian theorem (see Theorem 10a, p. 211 of Widder⁵).

Instead of imposing conditions on $F(\cdot)$ and $R(\cdot)$, one can place them on the Laplace-Stieltjes transforms $\varphi(s)$ and $\rho(s)$. We shall use the following Tauberian result:

Theorem 12: Let $\gamma(s)$ be the Laplace transform, convergent in $\text{Re}(s) > 0$, of a bounded function $C(\cdot) \geq 0$, vanishing for negative argument, and slowly decreasing; i.e.,

$$\liminf_{\substack{\epsilon \rightarrow 0 \\ t \rightarrow \infty}} C(t + \epsilon) - C(t) \geq 0.$$

Suppose that, for some function $g(\cdot)$, constant c , and $s = \sigma + i\tau$,

$$\lim_{\sigma \rightarrow 0+} \gamma(s) - s^{-1}c = g(\tau) \quad (24)$$

exists uniformly in every finite interval, $-a \leq \tau \leq +a$. Then

$$C(t) \rightarrow c \quad \text{as } t \rightarrow \infty.$$

Proof: This is a variant of Ikehara's theorem, Widder⁵ (p. 233). Define, for each $\lambda > 0$,

$$K_{\lambda}(x) \begin{cases} = 2\lambda(1 - |x/2\lambda|) & |x| \leq 2\lambda, \\ = 0 & |x| > 2\lambda. \end{cases} \quad (25)$$

This is the Fourier transform of the function

$$k_\lambda(x) = 2\lambda(2\pi)^{-\frac{1}{2}} \left(\frac{\sin \lambda x}{\lambda x} \right)^2.$$

For $\epsilon > 0$, set

$$\begin{aligned} I_\lambda(x) &= (2\pi)^{-\frac{1}{2}} \int_0^\infty k_\lambda(x-u)[C(u) - c]e^{-\epsilon u} du \\ &= \frac{1}{2\pi} \int_0^\infty [C(u) - c]e^{-\epsilon u} du \int_{-2\lambda}^{2\lambda} K_\lambda(y)e^{-iy(x-u)} dy \\ &= \frac{1}{2\pi} \int_{-2\lambda}^{2\lambda} K_\lambda(y)e^{-iyx} dy \int_0^\infty [C(u) - c]e^{-t(\epsilon+iy)} dt \\ &= \frac{1}{2\pi} \int_{-2\lambda}^{2\lambda} K_\lambda(y)e^{-iyx} [\gamma(\epsilon+iy) - (\epsilon+iy)^{-1}c] dy. \end{aligned}$$

By (24) we may take the limit under the integral as $\epsilon \rightarrow 0$, to obtain

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} I_\lambda(x) &= (2\pi)^{-1} \int_{-2\lambda}^{2\lambda} K_\lambda(x)e^{-ixy}g(y) dy \\ &= (2\pi)^{-\frac{1}{2}} \int_0^\infty k_\lambda(x-u)[C(u) - c] du, \end{aligned}$$

the last identity following from Widder⁵ (p. 183, Corollary 1c). The function $K_\lambda(\cdot)g(\cdot)$ belongs to L_1 , and so, by the Riemann-Lebesgue lemma,

$$(2\pi)^{-\frac{1}{2}} \int_0^\infty k_\lambda(x-u)[C(u) - c] du \rightarrow 0 \quad \text{as } x \rightarrow \infty.$$

Since $C(\cdot)$ is bounded and slowly decreasing, we conclude from Widder⁵ (p. 209, Theorem 9), that

$$C(u) \rightarrow c \quad \text{as } u \rightarrow \infty.$$

The theorem just proved implies directly the following consequence:

Corollary: Suppose that

- i. For some number L and function $g(\cdot)$

$$\frac{\varphi(s)}{\rho(s)} - \frac{L}{s} \rightarrow g(\tau) \quad \text{as } \sigma \rightarrow 0, \quad (s = \sigma + i\tau) \quad (26)$$

uniformly in every finite interval;

- ii. $P(\cdot)$ is slowly decreasing.

Then $P(t) \rightarrow L$ as $t \rightarrow \infty$.

Some simple conditions on $F(\cdot)$ and $R(\cdot)$ which ensure that hypothesis i of the corollary obtains are described in the next result:

Theorem 13: If

i. For some number L the function

$$F(t) - L \int_0^t R(u) du \quad (27)$$

(assumed to vanish for $t < 0$) is of bounded total variation;

ii. $R(\cdot)$ is of bounded total variation, and

$$\rho(-i\tau) \neq 0, \quad \|R_{\text{sing}}\| < \inf_{\tau} |\rho_{\text{dis}}(-i\tau)|;$$

then there exists a $\mu(\cdot)$ such that

$$\frac{\varphi(s)}{\rho(s)} - \frac{L}{s} = \mu(s) \rightarrow \mu(\tau) \quad \text{as } \sigma \rightarrow 0$$

uniformly in τ , and $\mu(\tau)$ is the Fourier-Stieltjes transform of a function of bounded total variation.

Proof: The Laplace-Stieltjes transform of (27) is

$$\varphi(s) - \frac{L}{s} \rho(s), \quad \text{Re}(s) \geq 0.$$

The condition ii ensures, by the Wiener-Pitt theorem, that $[\rho(-i\tau)]^{-1}$ is the Fourier-Stieltjes transform of a function of bounded total variation. It follows that $\mu(-i\tau)$ is also. But $\mu(\cdot)$ is the transform of $P(\cdot) - L$, and so extends analytically into $\text{Re}(s) > 0$ with the representation

$$\mu(s) = \int_0^{\infty} e^{-su} dG(u), \quad \|G\| < \infty.$$

Then

$$\begin{aligned} |\mu(s) - \mu(\tau)| &\leq \int_0^{\infty} |1 - e^{-\sigma u}| |dG(u)| \\ &\leq (1 - e^{-\sigma T}) \int_0^T |dG(u)| + \int_T^{\infty} |dG|. \end{aligned}$$

Let $\epsilon > 0$ be given, and pick T so large that the second term on the right is less than $\epsilon/2$; then pick σ so small that the first term is also. This proves the result.

The condition i of Theorem 13 is satisfied, for example, if $F'(\infty)$ exists and both of

- ii. for each $\tau \neq 0$, the function $E\{e^{i\tau[K(u)-u]}\}$ belongs to $L_1(0, \infty)$ and vanishes at ∞ (as a function of u);
 iii. the function

$$\lambda(\tau) = \Pr\{W(\infty) = 0\} \cdot i\tau \int_0^\infty E\{e^{i\tau[K(u)-u]}\} du \quad (30)$$

is continuous at $\tau = 0$, with $\lambda(0) = 1$;
 then there is a distribution function

$$L(\cdot) = \Pr\{W(\infty) \leq \cdot\}$$

such that

$$\lim_{t \rightarrow \infty} E\{e^{i\tau W(t)}\} = \lambda(\tau), \quad (31)$$

$$\lim_{k \rightarrow \infty} \Pr\{W(t) \leq w\} = L(w),$$

at continuity points of $L(\cdot)$, at least.

Proof: A direct application of Lemma 1 of Smith¹² to the real and imaginary parts of (6) with $s = -i\tau$, τ real, proves (31). The rest is a consequence of the standard continuity theorem for characteristic functions. Formula (30) is a generalization of the Pollaczek-Khinchin formula; this may be verified by letting $K(\cdot)$ be a compound Poisson process.

If $\Pr\{W(\infty) = 0\} = 0$, and condition ii of Theorem 16 is fulfilled, then a similar argument shows that

$$E\{e^{i\tau W(t)}\} \rightarrow 0 \quad \text{as } t \rightarrow \infty$$

for all $\tau \neq 0$, so no limiting distribution exists.

Alternatively, one can apply similar Abelian methods directly to formula (3). For convenience we shall use, instead of (3), the w -integrated form ($w \geq 0$)

$$\begin{aligned} \int_0^w \Pr\{W(t) \leq u\} du &= \int_0^{t+w} \Pr\{K(t) \leq u\} du \\ &\quad - \int_0^t R(t-u, w) \Pr\{W(u) = 0\} du, \end{aligned} \quad (4)$$

in which $R(t, u, w)$ has been replaced by $R(t-u, w)$, in accordance with the stationarity condition (7).

Theorem 17: If

i. For each $w > 0$

$$\int_0^w [R(u,w) - R(u,0)] du = r(w) < \infty;$$

ii. $R(t) \rightarrow 1$ as $t \rightarrow \infty$;

iii. $P(t) = \Pr\{W(t) = 0\} \rightarrow \Pr\{W(\infty) = 0\}$ as $t \rightarrow \infty$;

then, for $w > 0$,

$$\lim_{t \rightarrow \infty} \int_0^w \Pr\{W(t) \leq u\} du = w - r(w) \Pr\{W(\infty) = 0\}.$$

Proof: By subtracting the case $w = 0$ of (4) from the case $w > 0$, we find

$$\begin{aligned} \int_0^w \Pr\{W(t) \leq u\} du &= \int_t^{t+w} \Pr\{K(t) \leq u\} du \\ &\quad - \int_0^t [R(t-u,w) - R(t-u,0)]P(u) du. \end{aligned} \quad (32)$$

If $u \geq t$, then

$$\Pr\{K(t) \leq u\} \geq \Pr\{K(t) \leq t\} = R(t).$$

Hence, by ii,

$$\lim_{t \rightarrow \infty} \int_t^{t+w} \Pr\{K(t) \leq u\} du = w.$$

By i. and Lemma 1 of Smith,¹² the last term of (32) approaches

$$r(w) \Pr\{W(\infty) = 0\}.$$

To clarify the meaning of the condition i of Theorem 17, we note that

$$R(t,w) - R(t,0) = \Pr\{0 < K(t) - t \leq w\},$$

$$r(w) = E\{\text{amount of time that } 0 < K(t) - t \leq w\}.$$

Other limit theorems can readily be obtained, e.g., for the nonintegrated equation (3), as soon as suitable differentiability conditions are imposed. Since no new principle is involved, we shall leave the matter here.

REFERENCES

1. Beneš, V. E., General Stochastic Processes in Traffic Systems with One Server, B.S.T.J., **39**, 1960, p. 127.
2. Beneš, V. E., Combinatory Methods and Stochastic Kolmogorov Equations in the Theory of Queues with One Server, Trans. Am. Math. Soc., **94**, 1960, p. 282.

3. Beneš, V. E., Weakly Markov Queues. Transactions of Second Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, held at Liblice, June 1-6, 1959; Czech. Acad. Sci., Prague, 1960.
4. Beneš, V. E., A "Renewal" Limit Theorem for General Stochastic Processes, to be published.
5. Widder, D. V., *The Laplace Transform*, Princeton Univ. Press, Princeton, 1946.
6. Loève, M., *Probability Theory*, D. Van Nostrand Co., New York, 1955.
7. Karlin, S., On the Renewal Equation, *Pac. J. Math.*, **5**, 1955, p. 229.
8. Wiener, N., *The Fourier Integral and Certain of Its Applications*, Cambridge Univ. Press, Cambridge, 1933 (reprinted by Dover Publications, New York).
9. Ahlfors, L. V., *Complex Analysis*, McGraw-Hill, New York, 1953.
10. Pitt, H. R., *Tauberian Theorems*, Oxford Univ. Press, Oxford, 1958.
11. Wiener, N., and Pitt, H. R., On Absolutely Convergent Fourier-Stieltjes Transforms, *Duke Math J.*, **4**, 1938, p. 420.
12. Smith, W., Asymptotic Renewal Theorems, *Proc. Roy. Soc. Edinburgh*, **64A**, 1954, p. 9.
13. Paley, R. E. A. C., and Wiener, N., *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications, XIX, New York, 1934.
14. Pitt, H. R., General Tauberian Theorems (II), *J.*, London Math. Soc., **15**, 1940, p. 97.

A Nonlinear Integral Equation from the Theory of Servomechanisms

By V. E. BENEŠ

(Manuscript received March 28, 1961)

*The equation $x(t) = s(t) - k * F(x)(t)$, where $s(\cdot)$ is a given signal, $F(\cdot)$ is a nonlinear function, $k(\cdot)$ is the response of a linear system, and $*$ denotes convolution, describes a general class of servomechanisms. Properties of a solution $x(\cdot)$ can be established by finding a fixed point in a specific set of a function space, using Schauder's theorem.*

I. INTRODUCTION

A general class of nonlinear servomechanisms is described by the integral equation

$$x(t) = s(t) - \int_{-\infty}^{\infty} k(t-u)F(x(u)) du, \quad -\infty < t < \infty, \quad (1)$$

where $s(\cdot)$ is an input signal, $k(\cdot)$ is an impulse response function, and $F(\cdot)$ is a nonlinear function. The equation (1) represents the system diagram of Fig. 1, with $F(\cdot)$ as above, and with $K(\cdot)$ the transfer function corresponding to $k(\cdot)$. We assume that $F(\cdot)$ satisfies the uniform Lipschitz condition

$$|F(x) - F(y)| \leq \beta |x - y|,$$

and that $F(0) = 0$.

A classical method for studying nonlinear servomechanisms like that of Fig. 1 is to specify exactly the nonlinear element $F(\cdot)$, to assume that the response $k(\cdot)$ is the Green's function of a differential operator of low order, and to use some sort of phase-plane analysis. This method has two theoretical disadvantages: it lacks generality, and, when applied, it tends to give more information than is needed; thus it provides detailed knowledge about a restricted class of cases.

In this paper we shall use a method that has the opposite characteristics: it provides a small amount of highly relevant information about a

large class of cases. We shall exemplify the use of Schauder's fixed point theorem for studying solutions $x(\cdot)$ of (1) without specifying either $k(\cdot)$ or $F(\cdot)$ in detail. We establish definite properties of $x(\cdot)$ by finding a fixed point (corresponding to a solution of the equation) in a specific set of a function space. Since the function space and the set can be chosen in many ways, depending in part on what properties of $x(\cdot)$ are of interest, such a method can be used for a wide class of problems. The theory in the sequel is therefore restricted to sample results for the function space L_2 of square-integrable functions, and is to be regarded only as a particular example of the method described above.

II. FUNCTIONS OF FINITE ENERGY

In many situations it is desirable that the convolution term

$$\int_{-\infty}^{\infty} k(t-u)F(x(u)) du$$

follow the input signal $s(\cdot)$. The error in this approximation is then $x(\cdot)$ itself. It is then reasonable to work in the space L_2 of real, square-integrable functions, i.e., functions of finite energy. Accordingly, we assume that $k(\cdot)$ and $s(\cdot)$ are in L_2 , and we seek to bound the energy of a solution $x(\cdot)$ of (1).

Now the functions of L_2 cannot assume values appreciably different from zero on sets of arbitrarily large measure. Hence they may be viewed physically as pulses. By restricting $s(\cdot)$ and the solution $x(\cdot)$ to L_2 we are therefore studying the response of the system of Fig. 1 to certain pulses of finite energy. We shall be particularly interested in finding out how much of the energy of $x(\cdot)$ lies outside a given time interval.

The norm symbol $\|\cdot\|$ is used to denote the square root of the energy of a function. Thus for $x(\cdot)$ in L_2 ,

$$\|x\| = \left(\int_{-\infty}^{\infty} |x(t)|^2 dt \right)^{\frac{1}{2}},$$

and a sequence of functions $\{x_n(\cdot), n \geq 0\}$ is said to converge to a function $x(\cdot)$ in L_2 -norm if $\|x - x_n\|$ approaches zero with increasing n .

III. HYPOTHESES AND PRELIMINARY RESULTS

If $x(\cdot)$ is a function of L_2 , we let

$$Tx(\omega) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{i\omega t} x(t) dt$$

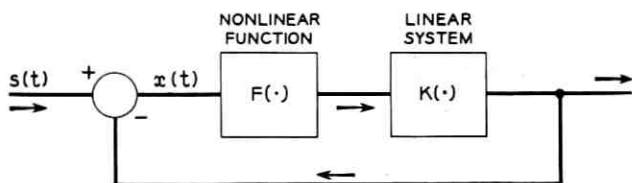


Fig. 1 — System diagram of servomechanism.

denote its Fourier transform; for $k(\cdot) \in L_2$, we reserve the special notation

$$Tk(\omega) = K(\omega).$$

The operator H on L_2 is defined by the condition

$$Hx(t) = \int_{-\infty}^{\infty} k(t-u)F(x(u)) du.$$

Lemma 1: If $K(\cdot)$ is bounded in ω , then H is a continuous transformation of L_2 into itself.

Proof: For $x(\cdot) \in L_2$ the Lipschitz condition on $F(\cdot)$ yields $\|F(x)\| \leq \beta \|x\|$, so that $F(x(\cdot)) \in L_2$. It is a known result that the convolution of two L_2 functions belongs to L_2 . Hence $Hx \in L_2$. Also, by the Parseval relations,

$$\begin{aligned} \|Hx - Hy\|^2 &= \int_{-\infty}^{\infty} |K(\omega)|^2 |TF(x) - TF(y)|^2 d\omega \\ &\leq \sup_{\omega} |K(\omega)|^2 \|F(x) - F(y)\|^2 \\ &\leq \beta \sup_{\omega} |K(\omega)|^2 \|x - y\|^2, \end{aligned}$$

which shows that H is continuous.

Now let $w(\cdot)$ be a given non-negative function of L_2 , and let S be the set of all $x(\cdot)$ in L_2 such that

$$|x(t)| \leq w(t), \quad \text{almost everywhere.} \quad (2)$$

Lemma 2: S is closed and convex.

Proof: Let $x_n(\cdot) \in S$ be a sequence of functions approaching $x(\cdot)$ in L_2 . Then for $\epsilon > 0$ and $\mu(\cdot) =$ Lebesgue measure,

$$\begin{aligned} \|x - x_n\|^2 &\geq \int_{|x_n - x| > \epsilon} |x_n - x|^2 dt \\ &\geq \epsilon^2 \mu\{t: |x_n(t) - x(t)| > \epsilon\}. \end{aligned}$$

However,

$$\begin{aligned} |x(t)| &\leq |x(t) - x_n(t)| + |x_n(t)| \\ &\leq |x(t) - x_n(t)| + w(t). \end{aligned}$$

Hence $|x(t)| - w(t) > \epsilon$ implies $|x(t) - x_n(t)| > \epsilon$ and

$$\mu\{t: |x(t) - x_n(t)| > \epsilon\} \geq \mu\{t: |x(t)| - w(t) > \epsilon\}.$$

Letting n approach infinity on the left, we find that

$$\{t: |x(t)| - w(t) > \epsilon\}$$

has measure zero for each $\epsilon > 0$. Hence almost everywhere

$$|x(t)| \leq w(t),$$

and so S is closed. The convexity of S is obvious.

We denote by B the subset of functions $x(\cdot)$ of L_2 which are "band-limited" to the frequency interval $(-\Omega, \Omega)$, i.e., representable as

$$x(t) = (2\pi)^{-\frac{1}{2}} \int_{-\Omega}^{\Omega} e^{i\omega t} T x(\omega) d\omega.$$

The physical interpretation of membership in B is of course that the sinusoidal oscillations into which a function is decomposed by the Fourier transform are restricted in frequency to the interval $(-\Omega, \Omega)$; i.e., $T x(\omega) = 0$ for $|\omega| > \Omega$.

The input signal $s(\cdot)$, and the response $k(\cdot)$ will be assumed to belong to B . If we define the operator J on L_2 by

$$Jx(t) = s(t) - Hx(t),$$

then the range of J is a subset of B . It follows that any solution of (1), i.e., any fixed point of J , will belong to B as long as $s(\cdot)$ and $k(\cdot)$ do so. Such a "band-limiting" restriction is natural physically, because of the known attenuation at high frequencies characteristic of physical circuits, and it will have an important mathematical role in finding fixed points of J . In particular, we note that $JS \subset B$.

To obtain a bound on the amount of energy that a solution $x(\cdot)$ has outside a given interval, we shall suppose that the non-negative function $w(\cdot)$ of L_2 , used in the definition of S , satisfies the integral inequality

$$|s(t)| + \beta \int_{-\infty}^{\infty} |k(t-u)| w(u) du \leq w(t). \quad (3)$$

This inequality may be thought of as defining an associated *linear*

problem; it will be used to ensure that $Jx(\cdot)$ belongs to S if $x(\cdot)$ does. The nonlinear function $F(\cdot)$ enters formula (3) only via its Lipschitz constant (of order 1) β .

Lemma 3: If (3) holds, and $F(0) = 0$, then $JS \subset S$.

Proof: Let $x(\cdot)$ belong to S . Then

$$\begin{aligned} |Jx(t)| &\leq |s(t)| + \int_{-\infty}^{\infty} |k(t-u)| |F(x(u))| du \\ &\leq |s(t)| + \beta \int_{-\infty}^{\infty} |k(t-u)| w(u) du \\ &\leq w(t). \end{aligned}$$

Our preliminaries are completed by

Lemma 4: $S \cap B$ is compact in L_2 .

Proof: Let $E = \|w\|^2$. The functions of $S \cap B$ are (uniformly) equicontinuous with modulus

$$\left(\frac{2\Omega E}{\pi}\right)^{\frac{1}{2}} \left(1 - \frac{\sin \Omega \epsilon}{\Omega \epsilon}\right)^{\frac{1}{2}}.$$

This follows from the inequalities:

$$\begin{aligned} |x(t + \epsilon) - x(t)| &\leq (2\pi)^{-\frac{1}{2}} \int_{-\Omega}^{\Omega} |e^{i\omega\epsilon} - 1| |Tx(\omega)| d\omega \\ &\leq (2\pi)^{-\frac{1}{2}} \left(\int_{-\Omega}^{\Omega} |e^{i\omega\epsilon} - 1|^2 d\omega\right)^{\frac{1}{2}} \|x\| \\ &\leq \left(\frac{2E}{\pi} \int_{-\Omega}^{\Omega} (1 - \cos \omega\epsilon) d\omega\right)^{\frac{1}{2}} \\ &\leq \left(\frac{2\Omega E}{\pi}\right)^{\frac{1}{2}} \left(1 - \frac{\sin \Omega \epsilon}{\Omega \epsilon}\right)^{\frac{1}{2}}, \end{aligned}$$

the last bound on the right being independent of t and $x(\cdot)$. Also, the inequalities

$$\begin{aligned} |x(t)| &\leq (2\pi)^{-\frac{1}{2}} \int_{-\Omega}^{\Omega} |Tx(\omega)| d\omega \\ &\leq (2\pi)^{-\frac{1}{2}} (2\Omega)^{\frac{1}{2}} \|x\| \\ &\leq \left(\frac{\Omega E}{\pi}\right)^{\frac{1}{2}}, \end{aligned}$$

show that the functions of $S \cap B$ are uniformly bounded.

Since both S and B are closed sets, it suffices (to prove Lemma 4) to show that $S \cap B$ is sequentially compact. Let $x_n(\cdot) \in S \cap B$ be an arbitrary sequence of functions. The $x_n(\cdot)$ are uniformly bounded and uniformly equicontinuous. By a standard diagonal argument using the σ -compactness of the real line, we can select a subsequence $x_m(\cdot)$ which converges to a function $x(\cdot)$ uniformly on any compact set. We have

$$\begin{aligned} |x(t)| &\leq |x(t) - x_m(t)| + |x_m(t)|, \\ \int_{-t}^t |x(u)|^2 du &\leq \int_{-t}^t |x(u) - x_m(u)|^2 du \\ &\quad + 2 \int_{-t}^t |x_m(u)| |x(u) - x_m(u)| du + \int_{-t}^t |x_m(u)|^2 du. \end{aligned}$$

For each fixed t , the first two terms on the right of the last inequality approach zero as m becomes large, and the third term is at most $\|w\|^2 = E$ uniformly in t . Hence $\|x\|^2 \leq E$ and $x(\cdot) \in L_2$. Using Minkowski's inequality, we find

$$\begin{aligned} \|x - x_m\| &\leq \left(\int_{|u| > t} |x - x_m|^2 du \right)^{\frac{1}{2}} + \left(\int_{|u| \leq t} |x - x_m|^2 du \right)^{\frac{1}{2}} \\ &\leq \left(\int_{|u| > t} |x(u)|^2 du \right)^{\frac{1}{2}} + \left(\int_{|u| > t} |x_m(u)|^2 du \right)^{\frac{1}{2}} \\ &\quad + \left(\int_{|u| \leq t} |x - x_m|^2 du \right)^{\frac{1}{2}}. \end{aligned}$$

The first two terms on the right can be made arbitrarily small by a large enough choice of t , uniformly in m ; for t fixed, the third term goes to zero as $m \rightarrow \infty$. Hence the $x_m(\cdot)$ converge to $x(\cdot)$ in L_2 , which proves Lemma 4.

IV. PRINCIPAL RESULTS FOR GENERAL $F(\cdot)$

Theorem 1: Let $s(\cdot)$ and $k(\cdot)$ belong to B , with $K(\cdot)$ bounded, let $F(0) = 0$, and let the integral inequality (3) obtain. Then there exists a solution $x(\cdot)$ of (1) in the set $S \cap B$, with the properties

$$|x(t)| \leq w(t) \quad (\text{and so } \|x\| \leq \|w\|),$$

$$|x(t)| \leq \left(\frac{\Omega}{\pi} \right)^{\frac{1}{2}} \|x\|,$$

$$x(t) = (2\pi)^{-\frac{1}{2}} \int_{-\Omega}^{\Omega} e^{i\omega t} T x(\omega) d\omega.$$

Proof: J is a continuous mapping of the closed convex set S into its compact subset $S \cap B$. By the "strong form" of Schauder's theorem,^{1,2} there exists a point $x(\cdot)$ in $S \cap B$ such that $x = Jx$. The properties listed above are immediate consequences of belonging to $S \cap B$.

The following slight modification of Theorem 1 involves no new principle:

Extension: If, in addition to the hypotheses of Theorem 1,

$$\beta \sup_{\omega} |K(\omega)| < 1, \quad (4)$$

then to the conclusion of Theorem 1 can be added

$$\|x\| \leq \frac{\|s\|}{1 - \beta \sup_{\omega} |K(\omega)|}.$$

Proof: Let a denote the bound on the right of the last inequality. Then the intersection Q of S with the closed ball of radius a is closed and convex. With condition (4), and $x(\cdot) \in Q$, the inequalities

$$\begin{aligned} \|Jx\| &\leq \|s\| + \left(\int_{-\Omega}^{\Omega} |K(\omega)|^2 |TF(x)|^2 d\omega \right)^{\frac{1}{2}} \\ &\leq \|s\| + \sup_{\omega} |K(\omega)| \|F(x)\| \\ &\leq \|s\| + \beta \sup_{\omega} |K(\omega)| \|x\| \end{aligned}$$

show that $JQ \subset Q \cap B$. Since the topology is Hausdorff, $Q \cap B$ is a closed subset of the compact set $S \cap B$, so it is compact. The result follows from Schauder's theorem.

V. PRELIMINARIES FOR $F(\cdot)$ NEARLY LINEAR AT THE ORIGIN

It is clear that stronger assumptions concerning the nonlinear function $F(\cdot)$ are necessary if we are to obtain results that make the energy of $x(\cdot)$ less than that of $s(\cdot)$. A particularly important case is one in which

$$F(t) = t + o(t), \quad \text{as } t \rightarrow 0;$$

that is, $F(\cdot)$ is linear near the origin.

Let $F(\cdot)$ have the form [where $n(k)$ are integers, $n(1) = 1$]

$$F(t) = \sum_{k=1}^{\infty} t^{n(k)} (-1)^{k-1} f_k, \quad (5)$$

with $f_k > 0$, $f_1 = 1$, $n(k+1) > n(k)$, the series converging for $|t| < \rho$, where

$$\rho^{-1} = \limsup_{k \rightarrow \infty} |f_k|^{1/k}.$$

Suppose also that there is a number a , $0 < a < \rho$, such that

$$a^{n(k+1)-n(k)} < \frac{f_k}{f_{k+1}}, \quad \text{for } k \geq 1. \quad (6)$$

Then $|t| < a$ implies that $F(t)$ has the sign of t and

$$|F(t) - t| \leq f_2 |t|^{n(2)}, \quad (7)$$

for then $F(\cdot)$ is represented by a power series of alternating sign whose terms are monotone in magnitude.

Since we are comparing (1) to a linearized version of (1) obtained by setting $F(t) = t$, we shall need the solution of the resulting linearized equation: this is a function $y(\cdot)$ defined by its Fourier transform

$$Ty(\omega) = \frac{Ts(\omega)}{1 + K(\omega)}.$$

Similarly, the closed-loop transfer function of the linearized loop is the Fourier transform

$$Tz(\omega) = \frac{K(\omega)}{1 + K(\omega)}$$

of a function $z(\cdot)$. These definitions will be justified in the theorem to be proved.

By dint of our stronger assumptions on $F(\cdot)$, we can use a different integral inequality from (3). We assume instead that there exists a real non-negative function $v(\cdot) \in L_2$ such that

$$|y(t)| + f_2 \int_{-\infty}^{\infty} v^{n(2)}(t-u) |z(u)| du \leq v(t). \quad (8)$$

With this inequality playing the role of (3), the method used to prove Theorem 1 can be applied almost without modification.

However, since the integral inequality (8) is nonlinear in $v(\cdot)$, we shall digress a little and give a sufficient condition for its validity. One way to do this is to find a non-negative $v(\cdot) \in L_2$ that satisfies (8) *with equality*, i.e., is a solution of the nonlinear equation

$$\begin{aligned} v(t) &= |y(t)| + f_2 \int_{-\infty}^{\infty} v^{n(2)}(t-u) |z(u)| du \\ &= Mv(t). \end{aligned} \quad (9)$$

We shall show how the classical contraction principle for complete metric

spaces can be used to find a solution $v(\cdot)$ of (9), i.e., a fixed point of M . Such a result is exemplified by

Lemma 5: If for

$$b = f_2 \int_{-\infty}^{\infty} |z(u)| du < \infty, \quad \alpha = \left(\frac{1}{bn(2)} \right)^{1/n(2)-1}$$

we have for some $\delta > 0$

$$\sup_u |y(u)| < \alpha \left(1 - \frac{1}{n(2)} \right) - \delta, \quad (10)$$

then the map M is contracting on the closed set Y of $x(\cdot) \in L_2$ such that

$$\begin{aligned} x(\cdot) &\geq 0 \\ \text{ess sup}_u x(u) &\leq \alpha - \delta. \end{aligned}$$

Proof: Consider the equation for $a > 0$,

$$\sup_u |y(u)| + ba^{n(2)} = a. \quad (11)$$

The left-hand side has unity slope at the point $a = \alpha$, and the inequality (10) implies that at this point the left-hand side is less than the right. Hence (11) has two roots in $a > 0$, and, for $x \in Y$,

$$\begin{aligned} \sup_t |Mx(t)| &\leq \sup_u |y(u)| + f_2 \int_{-\infty}^{\infty} x^{n(2)}(t-u) |z(u)| du \\ &\leq \sup_u |y(u)| + b\alpha^{n(2)} \\ &\leq \alpha - \delta. \end{aligned}$$

Thus $MY \subset Y$. To show that Y is a closed set we recall that convergence in L_2 implies convergence in measure. Let $x_n \in Y$ converge to x in L_2 ; then, as $n \rightarrow \infty$,

$$\mu\{t: |x_n(t) - x(t)| \geq \epsilon\} \rightarrow 0$$

for each $\epsilon > 0$. However, almost everywhere we have

$$-|x(t) - x_n(t)| \leq x(t) \leq |x_n(t) - x(t)| + \alpha - \delta,$$

and so

$$\mu\{t: x(t) < -\epsilon\} \cup \mu\{t: x(t) \geq \alpha - \delta + \epsilon\} \leq \mu\{t: |x_n(t) - x(t)| \geq \epsilon\},$$

where $\mu(\cdot)$ denotes Lebesgue measure. Letting $\epsilon \rightarrow 0$, we find

$$\operatorname{ess\,sup}_u x(t) \leq \alpha - \delta,$$

$$\operatorname{ess\,inf}_u x(t) \geq 0.$$

To show that M is contracting on Y , let $x(\cdot)$ and $y(\cdot)$ be arbitrary functions in Y . Then

$$\begin{aligned} \|Mx - My\| &\leq f_2 \left(\int_{-\infty}^{\infty} |Tz|^2 |Tx^{n(2)} - Ty^{n(2)}|^2 d\omega \right)^{\frac{1}{2}} \\ &\leq f_2 \sup_{\omega} |Tz| \|x^{n(2)} - y^{n(2)}\|. \end{aligned}$$

However, on Y

$$\begin{aligned} |x^{n(2)} - y^{n(2)}| &\leq |x - y| n(2)(\alpha - \delta)^{n(2)-1}, \\ \|x^{n(2)} - y^{n(2)}\| &\leq \|x - y\| n(2)(\alpha - \delta)^{n(2)-1}, \end{aligned}$$

and so, since $z \in L_1$,

$$\begin{aligned} \|Mx - My\| &\leq f_2 \sup_{\omega} |Tz| n(2)(\alpha - \delta)^{n(2)-1} \|x - y\| \\ &\leq bn(2)(\alpha - \delta)^{n(2)-1} \|x - y\|. \end{aligned}$$

But $bn(2)(\alpha - \delta)^{n(2)-1} < 1$, so M is contracting on Y .

Lemma 5 implies, by the contraction principle, that there exists a unique solution $v(\cdot)$ of (9) in the set Y , obtainable as the limit of successive approximations starting at any point of Y .

VI. PRINCIPAL RESULTS FOR $F(\cdot)$ NEARLY LINEAR AT THE ORIGIN

Let R be the set of functions $x(\cdot)$ of L_2 that satisfy the condition

$$|x(t)| \leq v(t), \quad \text{almost everywhere,}$$

where $v(\cdot)$ is the function in the inequality (8). The argument of Lemma 2 shows that R is closed and convex, and that of Lemma 4 shows that $R \cap B$ is compact.

Theorem 2: If $a > 0$ and $F(\cdot)$ have the properties (5) and (6), and if $k(\cdot)$ and $s(\cdot)$ both belong to B , with $k(\cdot) \in L_1$ and $K(\omega) \neq -1$, and if (8) holds with

$$\|v\|^2 < \frac{a^2 \pi}{\Omega}, \quad (12)$$

then a solution $x(\cdot)$ of (1) exists in B with the properties

$$|x(t)| \leq v(t),$$

$$|x(t)| \leq \left(\frac{\Omega}{\pi}\right)^{\frac{1}{2}} \|x\|.$$

Proof: Since $K(\omega)$ is continuous, and tends to zero at ∞ , it must be bounded away from -1 ; hence by the Wiener-Lévy theorem,^{3,4}

$$[1 + K(\omega)]^{-1}$$

is the Fourier transform of an integrable function $g(\cdot)$, and so

$$Ts(\omega)[1 + K(\omega)]^{-1}$$

is the Fourier transform of a function $y(\cdot)$ of $L_2 \cap B$, and also

$$K(\omega)[1 + K(\omega)]^{-1}$$

is the Fourier transform of a function $z(\cdot)$ of $L_1 \cap L_2 \cap B$.

We now write (1) as

$$x(t) + \int_{-\infty}^{\infty} k(t-u)x(u) du =$$

$$s(t) - \int_{-\infty}^{\infty} k(t-u)[F(x(u)) - x(u)] du.$$

Taking Fourier transforms gives

$$Tx = Ty - Tz T[F(x) - x].$$

We shall therefore consider the equivalent equation

$$x(t) = y(t) - \int_{-\infty}^{\infty} z(t-u)[F(x(u)) - x(u)] du,$$

$$= Gx(t).$$

This is of exactly the same form as (1); in particular, G is a continuous map. To apply Schauder's theorem it remains to verify that $GR \subset R$. For $x(\cdot) \in R$,

$$|Gx(t)| \leq |y(t)| + \int_{-\infty}^{\infty} |F(x(t-u)) - x(t-u)| |z(u)| du.$$

But $|x(\cdot)| \leq a$, by (12); so (7) gives

$$|F(x(t-u)) - x(t-u)| \leq f_2 |x(t-u)|^{n(2)},$$

$$\leq f_2 |v^{n(2)}(t-u)|, \text{ almost everywhere.}$$

Hence (8) implies that $|Gx(t)| \leq v(t)$.

The energy of the solution of the linearized equation with the input signal $s(\cdot)$ is

$$\left(\int_{-\infty}^{\infty} \left| \frac{Ts(\omega)}{1 + K(\omega)} \right|^2 d\omega \right)^{\frac{1}{2}} = \|y\|.$$

The gain of the closed linearized loop at the frequency ω is

$$\left| \frac{K(\omega)}{1 + K(\omega)} \right|. \quad (13)$$

It is reasonable to expect that, if the function $F(\cdot)$ is close to being linear, then the solution $x(\cdot)$ will have an energy close to that of the linear solution $y(\cdot)$, in the sense that, for some constant ξ that approaches unity as $F(\cdot)$ becomes linear, we have

$$\|x\| \leq \xi \|y\|.$$

A precise form of this intuitive idea, depending on the linearized loop gain (13), is given in

Theorem 3: If, in addition to the hypotheses of Theorem 2, it is true that

$$c = a^{n(2)-2} f_2 \sup_{\omega} \left| \frac{K(\omega)}{1 + K(\omega)} \right| < 1, \quad (14)$$

then to the conclusion of Theorem 2 may be added

$$\|x\| \leq \min \left(\frac{\|y\|}{1 - c}, \|v\| \right).$$

Proof: The intersection V of R which has the closed ball of radius $\|y\| / (1 - c)$ is closed and convex. With condition (14), and $x(\cdot) \in V$, the inequalities

$$\begin{aligned} \|Gx\| &\leq \|y\| + \left(\int_{-\infty}^{\infty} |Tz|^2 |T[F(x) - x]|^2 d\omega \right)^{\frac{1}{2}} \\ &\leq \|y\| + \sup_{\omega} |Tz| \|F(x) - x\| \\ &\leq \|y\| + c \|x\| \leq \frac{\|y\|}{(1 - c)} \end{aligned}$$

show that $GV \subset V \cap B$. Also, $V \cap B$ is a closed subset of the compact set $R \cap B$. So the result follows from Schauder's theorem.

The condition (14) used in Theorem 3 relates the maximum gain of the linearized loop with the second nonzero coefficient f_2 in the expansion of

$F(t)$ around the origin, and with the power $n(2)$ associated with this coefficient.

VII. ACKNOWLEDGMENT

The author is indebted to H. Landau for reading an early draft and making helpful suggestions.

REFERENCES

1. Schauder, J., Der Fixpunktsatz in Funktionalraum, *Studia Mathematica*, **2**, 1930, pp. 171-181.
2. Bers, L., Introduction to Topology, Lecture XV, New York Univ., New York, 1954-5.
3. Pitt, H. R., *Tauberian Theorems*, Oxford Univ. Press, Oxford, 1958, p. 97.
4. Paley, R. E. A. C., and Wiener, N., *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications, XIX, New York, 1934, p. 63.

The Observed 33 to 90 kmc Attenuation of Two-Inch Improved Waveguide

By A. P. KING and G. D. MANDEVILLE

(Manuscript received February 16, 1961)

The TE_{01} losses in long two-inch inside diameter circular waveguide lines were measured over the frequency range of 33 to 90 kmc using shuttle-pulse techniques. Experimental results obtained with two types of waveguide—one solid-wall copper and one helix—are presented. Loss curves for both types are given, along with the theoretical loss curve. Statistical results on TE_{01} loss variations due to mode conversion in the solid wall guide are also given. A new method for making shuttle-pulse loss measurements with a single backward-wave oscillator is described.

I. INTRODUCTION

The conditions favorable for low TE_{01} transmission loss in waveguide occur when the operating frequency is remote from cutoff and many other modes can consequently coexist. Two-inch inside diameter copper pipe can, for example, propagate 80 modes at 33 kmc, nearly 300 modes at 60 kmc, and about 700 modes at 90 kmc. When so many modes can be propagated, the presence of imperfections in the waveguide will introduce coupling between the TE_{01} and other modes. The spurious modes which are excited at such imperfections increase the TE_{01} loss, and a distribution of such imperfections can produce mode conversion-reconversion effects^{1,2} which cause the TE_{01} loss to fluctuate in an unpredictable manner with frequency.

Imperfections which occur in a long line can be divided into two categories: joints between sections and imperfections in the pipe itself. Discontinuities at the joint comprise tilt, offset, and change in diameter. The imperfections of the pipe are of a more distributed nature and comprise changes in diameter, ellipticity, and the more or less random deviations of the axis from a straight line. It has been shown by experience that with

care both the joints and cross-section tolerances of the pipe can be controlled so that increases in the TE_{01} losses from these sources are quite small. An analytic study and experiments by Rowe and Warters³ have shown that the variations in axial straightness contribute appreciably to increased losses and fluctuations of the TE_{01} mode. The theory shows that these mode conversion-reconversion effects occur when the period of these variations equals the beat wavelength.* The beat wavelengths of the TE_{01} mode with TE_{12} and with the TE_{11} modes over the 33 to 90 kmc band fall in the 1- to 5-foot range. In a waveguide, when the periods of components of the axial variations relative to a straight line also fall in the 1- to 5-foot range, the resulting interaction causes the average TE_{01} loss to increase and the actual loss to fluctuate with frequency.

It is the purpose of this paper to present the experimental results obtained with two types of waveguide made with very precise tolerances. One of the lines employed conventional solid-copper-wall waveguide, the other was an all-helix waveguide. The measured data cover a 33 to 90 kmc frequency band. A brief description of the measuring technique which was employed for these measurements is included.

II. DESCRIPTION OF LINES

The TE_{01} loss measurements were made on straight two-inch inside diameter lines laid in a straight steel conduit. The individual sections were about 15 feet long and were joined together with special threaded couplings⁴ which minimized offset and tilt at the joint. "O" rings between the coupling and the pipe were employed to seal the line. It was important to pump out the line to remove the oxygen and water vapor in order to avoid added losses due to absorption effects.⁵ The presence of atmospheric oxygen at standard pressure increases the TE_{01} line loss by a factor of 20 at 60 kmc. The presence of water vapor increases the losses to a lesser extent over the 50 to 90 kmc band, but the increase rises rapidly with frequency above 100 kmc.⁶

The solid-wall waveguide comprised oxygen-free high-conductivity copper of commercial manufacture to very close tolerances. The helix waveguide⁷ line (Fig. 1) was constructed in our laboratory by winding No. 37 wire on a precision mandrel. Over the helix winding a layer of glass fibers, coated with conductive material, was wound. This layer provided the necessary loss to attenuate the spurious modes. These elements were bonded together and to the outer steel shell with epoxy resin.

* The beat wavelength, λ_B , of mode 1 with mode 2 is defined as $1/\lambda_B = (1/\lambda_1) - (1/\lambda_2)$, where λ_1 and λ_2 are the guide wavelengths of the individual modes.

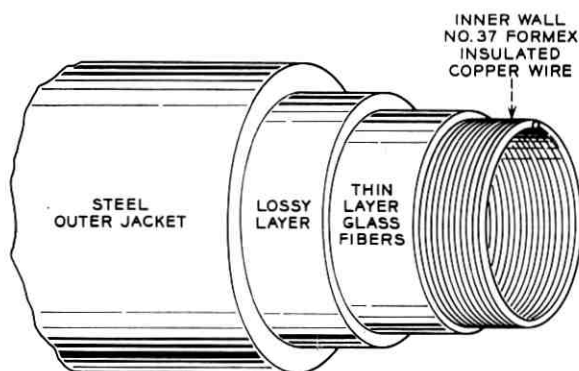


Fig. 1 — Helix waveguide line.

III. MEASURING TECHNIQUE

The TE_{01} losses were measured by the shuttle-pulse method (Fig. 2). This was accomplished by placing an iris at the input and a short at the far end of the section of line to be measured. The iris is a plane traverse metallic mesh with uniformly spaced holes which transmits only a small part of the incident pulse and decouples the measuring equipment from the line without introducing spurious modes. In the measuring equipment the iris was proportioned to transmit approximately 1 per cent of the incident energy. Thus the received signal pulses from the test section were reduced in level by twice this amount or about -40 db relative to P_0 . While this power loss tends to place a lower limit on the power of the signal source, a level in the 1 to 5 milliwatt range has been found to be adequate.

The shuttle-pulse measuring technique with an iris has several attractive features. It separates the input waveguide circuits which have large losses from the low loss waveguide being measured. With a line whose one-trip loss is a fraction of a decibel, the shuttle-pulse method permits the loss to accumulate for a number of trips, thus permitting an improved measurement accuracy. The results presented herein were based on measurements of 10 to 30 round-trips.

The 33 to 38 kmc measurements were made with a two-klystron set-up. The signal pulse was obtained from one of the klystrons by introducing a short pulse in the repeller circuit so that oscillations occurred only during the pulse interval. The other tube provided the continuous-wave beating oscillator signal.

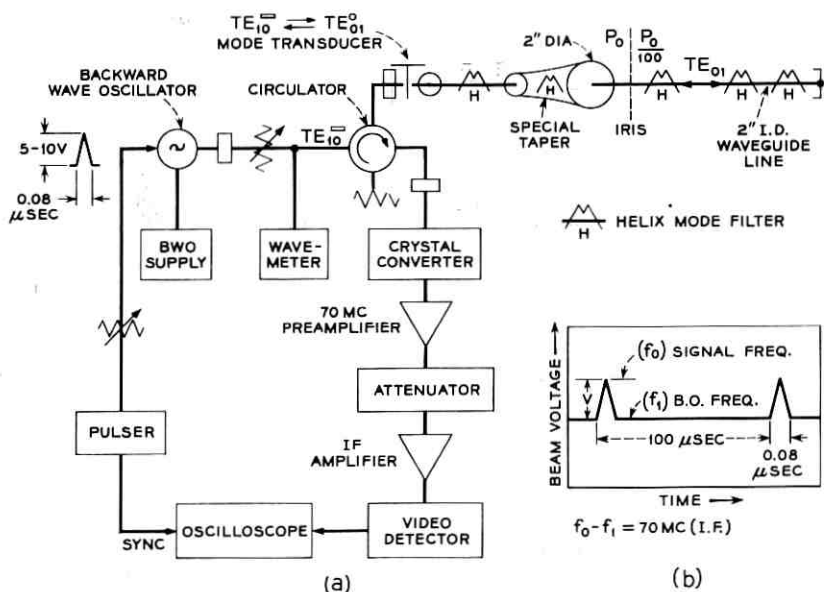


Fig. 2 — (a) Measuring circuit and (b) modulation scheme for shuttle-pulse method of measuring TE_{01} losses.

For the 41 to 90 kmc band backward-wave oscillators were used. A single oscillator tube was employed in a special measuring circuit [Fig. 2(a)] proposed by Ring⁸ to provide both the transmitted signal pulse and the receiving beating oscillator power. The modulation scheme required to do this is shown graphically in Fig. 2(b). Here the modulating pulse amplitude of a few volts is adjusted to a voltage v such that at the top part of the pulse the oscillator frequency is shifted to frequency f_0 . The steady-state unmodulated frequency of the oscillator is f_1 . The intermediate frequency is $f_0 - f_1 = 70$ mc. Thus, after the pulse, the beam voltage is restored to the steady-state condition and supplies continuous beating oscillator signal between signal pulses. The 0.08-microsecond pulse width is dictated by the bandwidth of the receiving circuitry; the 100-microsecond repetition rate permits the energy at the signal frequency f_0 to decay before a new pulse is generated. The beating oscillator power impressed upon the converter is the reflection of the forward power from the front face of the iris; it is higher than the pulses from the section of line under test by a factor of twice the iris coupling, which is about 40 db. This condition ensures a good signal linearity.

Ring's circuit is an important advance in measurement techniques.

The advantages are that the frequency of a backward-wave oscillator can be closely set and the single-tube pulse method provides more convenient and faster measurements.

The waveguide components mounted between the signal source, the crystal converter, and the TE_{01} wave transducer were in rectangular waveguide and were of more or less conventional form. The wave transducer has been described elsewhere.⁹ A helix waveguide mode filter was connected at the TE_{01} output of the transducer to minimize spurious mode impurities and a special taper¹⁰ matched the transducer to the two-inch waveguide line. The taper section was designed to minimize conversion to higher-order TE_{0n} modes, especially the TE_{02} mode. The iris, which has already been described, was placed at the input of the section of line to be measured.

Three helix waveguide sections 10 to 15 feet long, which served as spurious-mode filters, were employed in the solid-copper test line. Two of these mode filters, one at each end of the line, were necessary in order to eliminate accumulative mode conversion-reconversion effects with successive trips and to ensure that each trip had the same TE_{01} transmission characteristic and loss. The third mode filter was placed approximately in the middle of the line. Its purpose was to reduce the continuous length of solid-copper waveguide to half the length, from 400 to 200 feet for the line described. The presence of this additional mode filter had a negligible effect upon the average TE_{01} loss, but the shorter run of solid-copper waveguide reduced the length over which continuous mode interaction could occur and resulted in smaller variations in TE_{01} loss with frequency.

IV. EXPERIMENTAL RESULTS

The observed data were plotted in terms of the one-way transmission loss as a function of frequency. The measurements were made on a round-trip basis and have been corrected for end losses. These include both the end loss and coupling loss at the iris and the loss due to the short at the end of the line.

The measured data for an all-helix line, a solid-wall-copper waveguide line, and the theoretical loss for solid-wall copper are plotted in Fig. 3 in db per mile as a function of frequency.

The smooth dashed line curve in Fig. 3 represents the loss which is based on the measurements of helix lines varying in length from 415 feet to over 1500 feet. These measured data have been extrapolated so that the ordinate values express the loss in db per mile. The theoretical loss curve is also in db per mile. At the lowest frequency, 33 kmc, the meas-

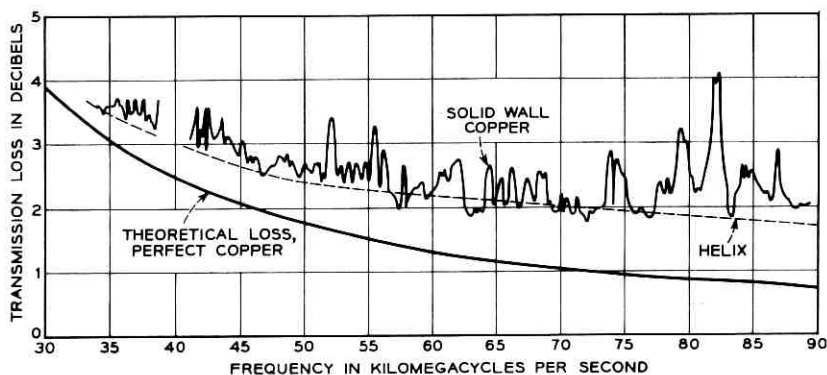


Fig. 3 — TE_{01} loss versus frequency.

ured loss is 3.7 db per mile; at 60 kmc, the loss is 2.2 db per mile; and at 90 kmc, the highest frequency measured, the loss is 1.7 db per mile. At 55 kmc, the helix loss is in good agreement with values determined on individual sections with a Q -measuring set.¹¹ The ratios of the measured to the theoretical attenuation at these frequencies are 1.11, 1.67, and 2.3 respectively. Thus, while the measured loss of this line relative to the theoretical value increased with frequency, the actual loss was still decreasing at the highest frequency measured. It is, therefore, presumed that the transmission loss of the helix line will be less than 1.7 db at frequencies above 90 kmc. The smoothness of this transmission loss curve is the result of the continuous mode-filtering action of helix waveguide.

The measured data obtained with a solid-wall-copper waveguide 443 feet long are shown by the upper solid line curve of Fig. 3. The loss is plotted on a relative basis for reasons which are mentioned below. The fluctuations in the TE_{01} loss which occurred in this line are due to mode conversion-reconversion effects that resulted from imperfect waveguide geometry. These variations in loss were relatively small at the lower frequencies, but in the vicinity of 80 kmc and above there were undesirably high loss peaks.

For similar lines with the same mode filter spacing the average loss increases linearly with length. The results for the 443-foot line may also be extended by using the analysis of random imperfections³ to give the statistical properties of a much longer line. Thus, for a longer line with the same mode-filter spacing, the rms fluctuations in TE_{01} loss will increase by the factor \sqrt{n} , where n represents the number of 443-foot sections which are required for the longer line. Both the average loss

TABLE I

| | Frequency | | | | | |
|--|-----------|-----------|-----------|-----------|-----------|-----------|
| | 33-38 kmc | 41-50 kmc | 50-60 kmc | 60-70 kmc | 70-80 kmc | 80-90 kmc |
| Helix line average loss, db per mile | 3.38 | 2.75 | 2.29 | 2.11 | 1.92 | 1.79 |
| Solid-wall copper average loss, db per mile | 3.53 | 2.90 | 2.49 | 2.25 | 2.21 | 2.42 |
| Solid-wall copper rms variation in db for line 1 mile long | 0.050 | 0.043 | 0.078 | 0.074 | 0.118 | 0.141 |

and the rms fluctuations of the solid wall are computed from the measured data of Fig. 3 for a line one mile long in Table I. The average loss of the helix line is included. These data are for 10 kmc frequency bands above 50 kmc, and for the 33 to 38 and the 41 to 50 kmc bands below 50 kmc. The data of Table I were prepared by C. W. Curry.

V. CONCLUSIONS

The observed TE_{01} transmission losses over a frequency band of nearly 3-to-1 (33 to 90 kmc) show low transmission losses. At the lower frequencies both lines show an increasing attenuation with decreasing frequency which approaches the theoretical value at 33 kmc. The helix line shows a decreasing loss as the frequency increases, and is still decreasing at 90 kmc, the highest frequency of measurement. Above 70 kmc the solid-wall-copper line exhibits increasing loss with frequency. This is due to mode conversion-reconversion effects. Of the two lines, the helix waveguide has a lower attenuation and a smoother transmission characteristic.

REFERENCES

1. Miller, S. E., Waveguide as a Communication Medium B.S.T.J., **33**, 1954., p. 1209.
2. Morgan, S. P., Mode Conversion Losses in Transmission of Circular Electric Waves Through Slightly Non-Cylindrical Guides, J. Appl. Phys., **21**, 1950, p. 329.
3. Rowe, H. E., and Warters, W. D., Transmission Deviations in Waveguide Due to Mode Conversion, Proc. I.E.E., **106**, Pt. B, Supp. 13, 1959, p. 30.
4. King, A. P., Observed 5-6 mm Attenuation for Circular Electric Wave in Various Pipes, B.S.T.J., **35**, 1956, p. 1116.
5. Crawford, A. B., and Hogg, D. C., Measurement of Atmospheric Attenuation at Millimeter Wavelengths, B.S.T.J. **35**, 1956, p. 907.
6. Hogg, D. C., and Mumford, W. W., The Effective Noise Temperature of the Sky, Microwave J. **3**, March 1960, p. 80.

7. Beck, A. C., and Rose, C. F. P., Waveguide for Circular Electric Mode Transmission, Proc. I.E.E., **106**, Pt. B, Supp. 13, 1959, p. 159.
8. Ring, D. H., unpublished data.
9. Miller, S. E., and Beck, A. C., Low Loss Waveguide Transmission, Proc. I.R.E. **41**, 1953, p. 354.
10. Unger, H. G., Circular Waveguide Taper of Improved Design, B.S.T.J., **37**, 1958, p. 899.
11. Young, J. A., Resonant Cavity Measurements of Circular Electric Waveguide Characteristics, Proc. I.E.E., **106**, Pt. B, Supp. 13, 1959, p. 62.

The Effect of Rain and Water Vapor on Sky Noise at Centimeter Wavelengths

By D. C. HOGG and R. A. SEMPLAK

(Manuscript received March 7, 1961)

Measurements of sky noise temperature at a frequency of 6.0 kmc have been made for various conditions of absolute humidity and precipitation. For an antenna beam position 5° above the horizon, the sky noise temperature was found to increase about 15°K from winter to summer due to the change in absolute humidity. During periods of rain, with the antenna beam pointed to the zenith, sky noise temperatures as high as 130°K have been observed, compared with the usual background value of 3° due to oxygen and water vapor. Theoretical calculations of sky noise temperature for typical dry and humid summer days are presented.

I. INTRODUCTION

This study is directed toward evaluating the variability in the noise level of sensitive ground-based microwave receivers to be used in radar or space communications.¹ The sensitivity of such receivers is limited by the antenna noise received from the sky due to emission of radiation by the earth's atmosphere and by noise sources in space. The sun and moon and the galaxy are serious noise contributors in space; however, since the sun and moon are seldom in the antenna beam and since galaxy noise is small at centimeter wavelengths, sky noise is due mainly to emission by the earth's atmosphere. The atmosphere contains oxygen and water vapor, both of which absorb and therefore emit radio waves, especially in the microwave band. The oxygen content is relatively constant in time; it therefore produces a sensibly constant background noise. However, the water vapor content and the noise it produces change with location and season. Condensed water, in the form of rain, is a strong absorber of microwaves and contributes a significant amount of noise, whose level changes with time.

A brief theoretical treatment outlining the amount of noise to be

expected from the atmosphere is followed by a discussion of experimental data obtained at a frequency of 6 kmc from a receiving site, at an elevation of 370 feet above sea level, located at Crawford Hill, Holmdel, New Jersey.

II. THEORETICAL DISCUSSION

The noise received from a horizontally stratified atmosphere by the infinitely sharp beam of an idealized antenna is given by $P_a = kT_s B$, where k is Boltzmann's constant, B is bandwidth, and T_s is effective sky temperature. If there is no precipitation, oxygen and water vapor are the two most significant contributors of noise, and the sky noise temperature is

$$T_s = \int_0^{\infty} (\alpha_1 + \alpha_2) T \exp \left[- \int_0^r (\alpha_1 + \alpha_2) dr \right] dr, \quad (1)$$

with T being the actual temperature at a given point P in the atmosphere defined by the distance r along the antenna beam and an angle φ , as shown in Fig. 1; α_1 is the power absorption coefficient due to oxygen at point P and α_2 is that due to water vapor, both being functions of frequency, temperature, and pressure. The oxygen absorption coefficient α_1 is due to the magnetic dipole resonance of the oxygen molecule at frequencies near 60 kmc. Pressure-broadening causes the skirts of these resonance lines to extend through the centimeter wave region, and the absorption contributes significantly to antenna noise. The water vapor absorption coefficient α_2 consists of two terms, absorption due to the electric dipole resonance of the water molecule near 22.5 kmc and that due to the skirts of rotational lines in the infrared region.

Assuming a perfectly dry atmosphere, $\alpha_2 = 0$, and the effective sky noise temperature due to oxygen alone can be calculated using (1). These values are plotted as the lower curves of the groups of curves shown in Fig. 2; they are based upon the temperature and pressure dependence of Fig. 1(a). The four sets of curves in Fig. 2 correspond to various beam directions, ϕ , of the antenna. The central curves in the figures are also calculated for standard temperature and pressure conditions, but in this case, α_2 is not zero* since the atmosphere is assumed to contain water

* In these calculations a line-broadening constant of 0.3 cm^{-1} is used for the infrared lines; this choice is based on absorption measurements in the millimeter wavelength region. It has recently been proposed³ that mutual coupling between O_2 and H_2O molecules in air may cause additional absorption at millimeter wavelengths and that the broadening constant should be 0.1 cm^{-1} , as it is for the 22.5 kmc line. This value would decrease considerably the noise temperature due to water vapor in the calculated curves of Fig. 2. Nevertheless, the measured data indicated in Fig. 6 appear to justify the use of 0.3 cm^{-1} as the line-broadening constant.

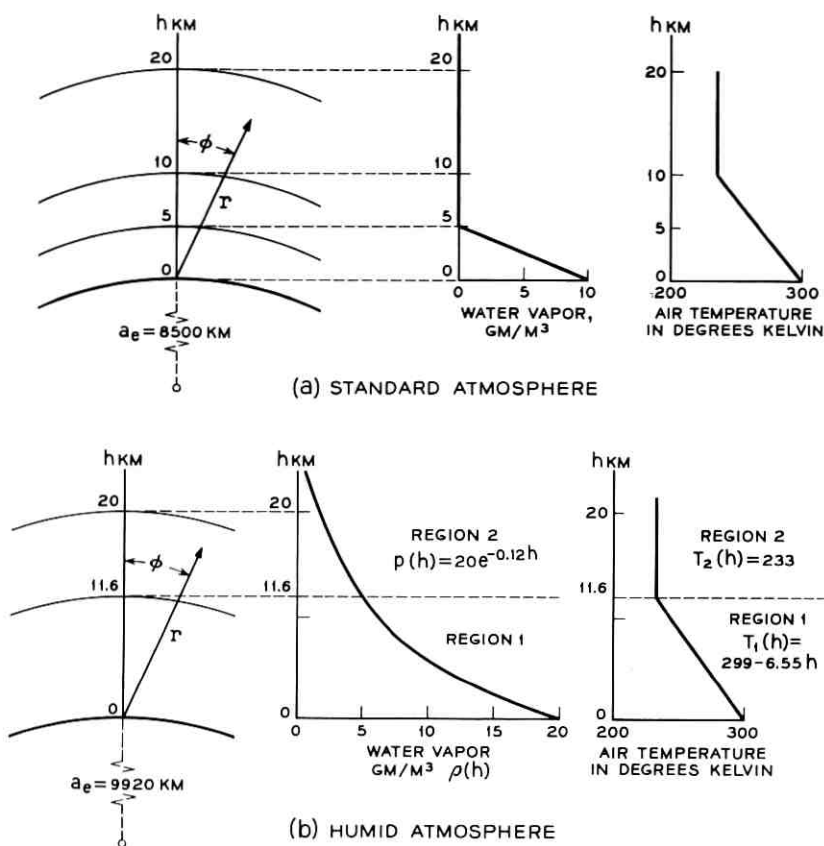


Fig. 1 — (a) Assumed model of standard summer atmosphere; (b) model of very humid summer atmosphere based on meteorological data from Newark Air Service.

vapor of density 10 grams per cubic meter at the surface of the earth, which decreases linearly to zero at an altitude of 5 kilometers, as shown in Fig. 1(a). This condition is representative of average summer conditions in temperate zones. In computing both the lower and central curves, the standard value $\frac{4}{3}$ was used for the normalized radio radius of the earth.

The upper curves in Fig. 2 were calculated for an atmosphere of a typical humid summer day. This model atmosphere, shown in Fig. 1(b), is based upon meteorological data for August 8, 1960.* The normalized radio radius calculated from these data is $\frac{3}{2}$ rather than $\frac{4}{3}$, which means

* Measured by the Newark Air Service at Newark Airport, Newark, New Jersey.

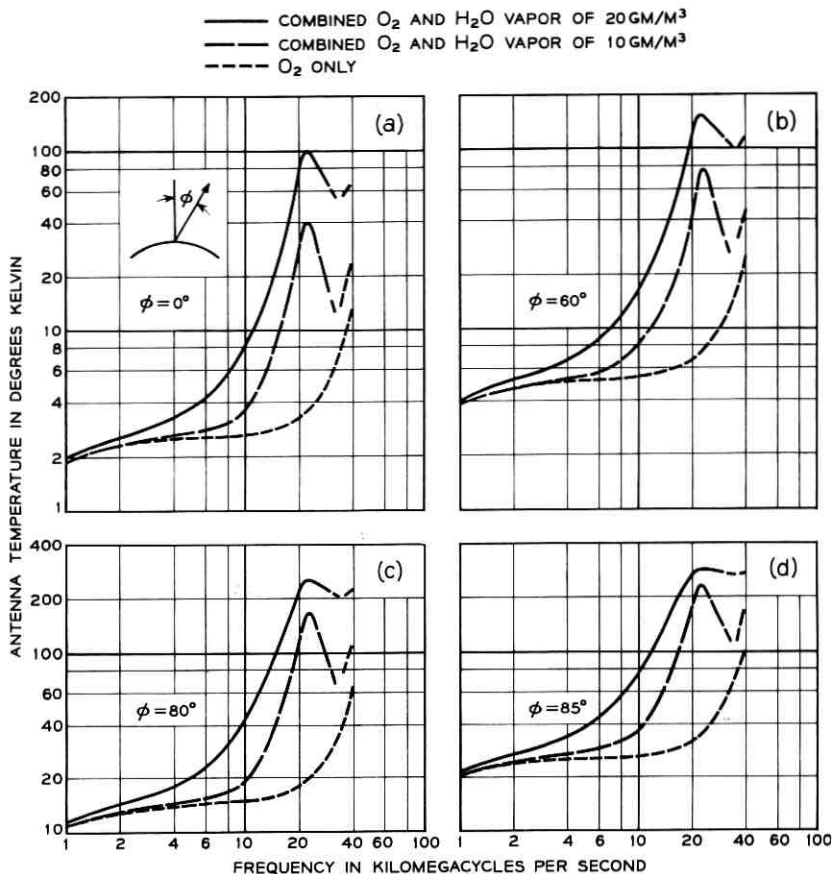


Fig. 2 — Typical sky noise temperatures for various antenna positions. In each case, lower curve is for very dry weather, center curve for average weather, and upper curve for very humid weather.

that for large values of ϕ the antenna beam has a much longer path length through this atmosphere than through the one discussed above. This earth radius and the data of Fig. 1(b) were used in (1) to compute the upper curves of Fig. 2. Because of the rather extreme models chosen for the atmospheres upon which the upper and lower curves are based, they, in the absence of super-refraction and condensation, represent upper and lower bounds on the sky noise one may expect from the atmosphere in the microwave band.

Calculation of the effect of rain on sky noise temperature is difficult, in spite of the fact that the theoretical power absorption coefficient⁴ is

known to agree well with measurements of absorption by rain in the microwave band. Difficulty arises because condensed water in the atmosphere is not horizontally stratified. Moreover, the distribution of condensed water as a function of altitude varies considerably with meteorological conditions. For example, weather radar measurements have shown that rain may originate as high as 45,000 feet.* However, if the distribution of water along the path of the antenna beam were known, neglecting scattering by the raindrops, one would obtain its effect on noise temperature from (1) by substituting for $\alpha_1 + \alpha_2$ the appropriate $\alpha(r)$ for the water distribution. Thus, a measurement of sky noise is an integrated measure of the water content in the antenna beam.

III. APPARATUS

For measurement of sky noise, it is desirable that the antenna and receiver proper have low intrinsic noise, relatively large bandwidth, and good stability. If these conditions are met, accurate absolute values of sky noise can be obtained and rapid changes in the level can be observed. A receiving system⁵ that meets these requirements consists of a traveling-wave maser⁶ and a horn-reflector antenna.

This type of antenna was designed for microwave radio relay systems and is shown in Fig. 3; it consists of a section of a paraboloid fed by a pyramidal horn. The antenna is broadband and is well matched. However, its virtues as a low-noise antenna stem from the fact that the far side and rear lobes are about 30 db below the isotropic level, thus allowing only a small amount of noise from the earth (which is a good absorber at microwaves) to enter the antenna; also, only a short length of transmission line is required between the horn and receiver proper (0.1 db transmission loss at room temperature produces $7\frac{1}{2}^\circ\text{K}$ of noise). At 6 kmc, the beamwidth of the antenna shown in Fig. 3 is about 1.75 degrees and its gain is 40 db. The antenna mount permits movement in the elevation plane only, but in this plane the antenna beam may be moved from horizon to horizon. To permit this movement, a rotary joint is used in the waveguide run between the antenna and the maser. The length of this waveguide run is minimized to reduce waveguide losses and their associated noise contributions. A means of ensuring the rapid exit of water from the antenna is of practical concern, since the aperture is open to the elements; a few holes in the seam formed by the parabolic section and horn satisfy this requirement.

The essential features of the receiver proper are shown in Fig. 4 in

* Private communication through the courtesy of J. S. Marshall, McGill University, Montreal, Canada.

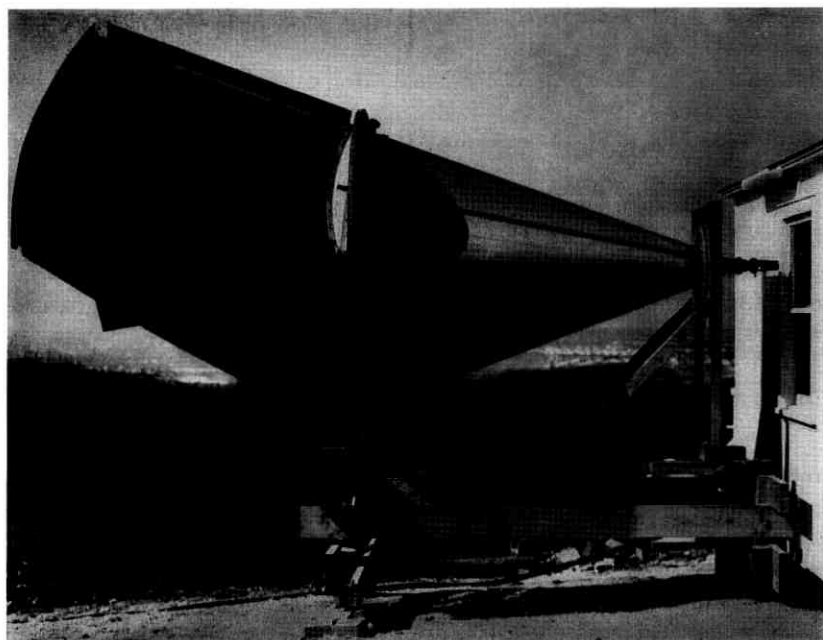


Fig. 3 — Horn-reflector antenna.

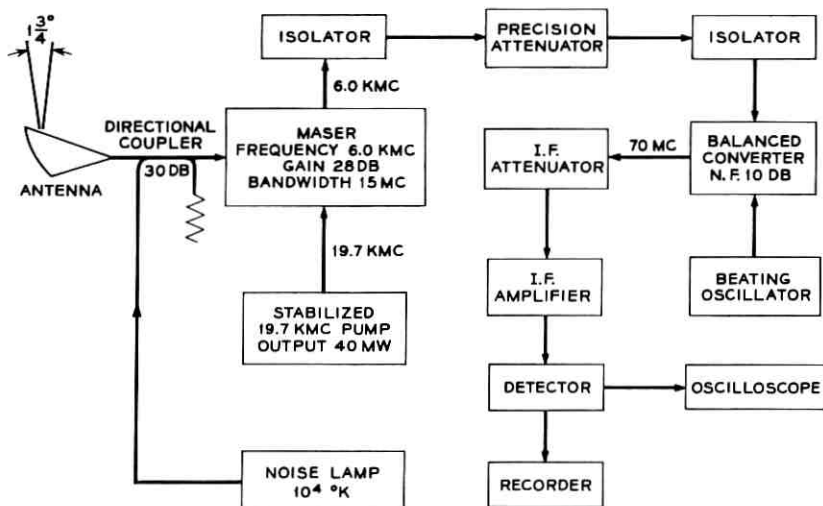


Fig. 4 — Low-noise receiver for 6 kmc.

block diagram form. At 6 kmc, the traveling-wave maser has a gain of 28 db and a bandwidth of 15 mc at a liquid helium temperature of 1.5°K. The magnetic field for the maser is provided by a permanent magnet. The pumping power is obtained from a stabilized 19.7-kmc klystron with an output of approximately 40 mw. Isolators are used between maser and attenuator and between attenuator and balanced converter to prevent the possibility of noise power re-entering the maser and, due to any input mismatch, reappearing as a net increase in input temperature. The remainder of the system utilizes a balanced crystal converter and IF amplification.

IV. MEASUREMENT TECHNIQUES

The over-all input noise temperature of the receiver consists of a constant term, which includes noise due to the antenna, input waveguide loss, maser, and the contribution from the second stage; these amount to 15.5°K, which is the limit of the sensitivity of the receiver proper. An additional term is the sky noise temperature, which varies as a function of the microwave absorption in the volume of the sky subtended by the antenna beam; on a clear day the sky noise is about 3°K, which increases the input noise to 18.5°K.

During periods of relatively stable atmospheric conditions a comparison method employing an argon noise lamp is used to measure the input temperature. The noise lamp is coupled into the input waveguide through a 30-db directional coupler. With the noise lamp switched off, the room temperature load which terminates the noise lamp introduces about 0.3°K; when the noise lamp is fired, the input noise is increased by 10°K. The change in noise level is determined by using a precision microwave attenuator. This is set at zero with the noise lamp off, and the IF gain of the measuring equipment is adjusted to obtain a suitable output level of detected noise; then the noise lamp is fired, and the precision attenuator is adjusted to give the same noise output from the measuring set.

The noise temperature at the input to the converter, T_{ic} , with the noise lamp off, is given by

$$T_{ic} = G_m(T_a + T_m), \quad (2)$$

where

G_m is the gain of the maser,

T_a is the effective antenna noise temperature (waveguide, antenna, and sky), and

T_m is the effective noise temperature of the maser.

In this equation a small noise due to losses between maser and converter has been neglected. When the noise lamp is fired, the noise input to the converter becomes

$$T_{ic} = (T_a + T_m + T_L)G_m A + (1 - A)T_0, \quad (3)$$

where

T_L is the additional noise introduced into the input waveguide by the noise lamp,

T_0 is ambient room temperature, and

A is the reciprocal of the additional attenuation introduced by the precision attenuator.

Let $T_i = T_a + T_m$; then, solving (2) and (3),

$$T_i = \frac{AT_L}{1 - A} + \frac{T_0}{G_m}.$$

The second term of this equation is the noise contribution of the precision attenuator; with maser gains in the order of 30 db, it is less than 0.5°K.

The noise-lamp method of measurement, although quite accurate, is time-consuming and does not lend itself readily to continuous measurement, since atmospheric conditions can and do change rapidly. For a continuous record of the changes in sky noise, the detected output of the intermediate frequency amplifier is applied to a two-channel Sanborn recorder. On one channel the full-scale sensitivity is of the order of 250°K, while on the second channel it is $\pm 6^\circ\text{K}$. It should be noted that a 15-mc band is preserved up to the recorder and that the recorder has a maximum bandwidth of 80 cps. Simultaneously, the output is displayed on an oscilloscope.

Ground-rain rate data obtained from a calibrated tipping-bucket rain gauge located about 50 feet from the antenna are fed to a timing channel on the recorder; each tip of the rain gauge bucket indicates that 0.01 inch of rain has fallen. This type of gauge is fairly accurate, provided the wind is not excessive. The time interval between two consecutive bucket tips is carefully measured, and the rainfall rate is then plotted at the mean value of time of two consecutive tips. This procedure was used for all recorded bucket tips.

V. DISCUSSION OF MEASURED DATA

5.1. *The Water Vapor Effect*

Fig. 5 shows a set of data taken on October 7, 1960, with sky temperature plotted as ordinate and antenna beam position as abscissa. This

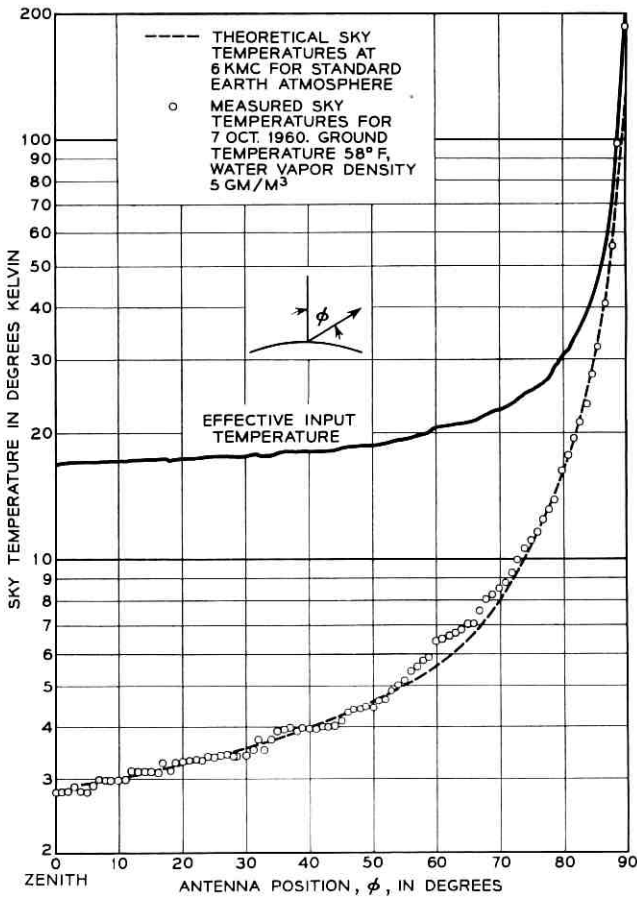


Fig. 5 — Sky noise temperature for a dry clear day.

particular day was dry and clear; based on temperature and humidity measurements made at the receiving site, the absolute water vapor density was 5 grams per cubic meter at the ground. The curve labeled "effective input temperature" shows the values of the first-circuit input noise temperature for various antenna positions; these values are the sum of the sky temperature plus the noise temperature contributed by input waveguide losses, maser noise, and antenna. The actual sky temperatures for the various antenna beam positions are plotted as dots. These data were obtained by averaging groups of measurements taken for each antenna beam position, as the antenna was moved from zenith to horizon in one-degree increments. For comparison, the dashed curve

represents the theoretical sky temperature based on a standard earth's atmosphere, as obtained from the central curves of Fig. 2.

From the relative smoothness of the dotted curve in Fig. 5 and the absence of abrupt changes, one concludes that the back and side lobes of the antenna allow only a small amount of noise radiation from the earth into the system. There is little difference in sky noise between that measured for a water vapor density of 5 grams per cubic meter and the theoretical values based on a water vapor density of 10 grams per cubic meter. A period of two and one-half hours was required to complete this set of measurements. During this time atmospheric conditions could have changed somewhat, even though no change was visibly detectible.

In order to determine the effects of water vapor density, antenna temperatures were measured for various sky conditions at random times during the year 1959-60. Fig. 6 shows the increase in sky temperature above zenith temperature for various water vapor densities. Again, the values for water vapor density were based on measurements made at the receiving site. Measurements with the antenna beamed 60° , 85° , and $87\frac{1}{2}^\circ$ from the zenith are shown. As the antenna beam approaches the horizon, it looks through greater amounts of the earth's atmosphere, so that the water vapor effect is accentuated. The data show more or less what we might expect: the noise temperatures increase with increasing water vapor density. The data shown here are conservative because they are taken relative to the zenith temperature for the day that the measurements were made; in other words, if this were not done, the slope of the curves would be slightly greater. With the antenna beam at 60° from the zenith, the sky noise temperature increases about 1° from winter to summer; at 85° from zenith, it increases 15° from winter to summer. These increases are partly due to changes in the effective radius of the earth, since the refractive index profile also changes with water vapor content. The scatter in the data is partly due to the following: water vapor measurements are only made at the receiving site; the distribution of water vapor with height is unknown, and hence the changes in the effective earth's radius are also unknown.

The dashed curves in Fig. 6 connect three theoretical noise temperatures for water vapor contents 0, 10, and 20 grams per cubic meter, as obtained from Fig. 2. Especially in the case of the 85° curve, agreement with the experimental data is considered to be good.

5.2 *The Rain Effect*

Sky temperature measurements taken on a drizzly day (May 18, 1960) for various zenith angles of the antenna beam are shown in Fig. 7; the

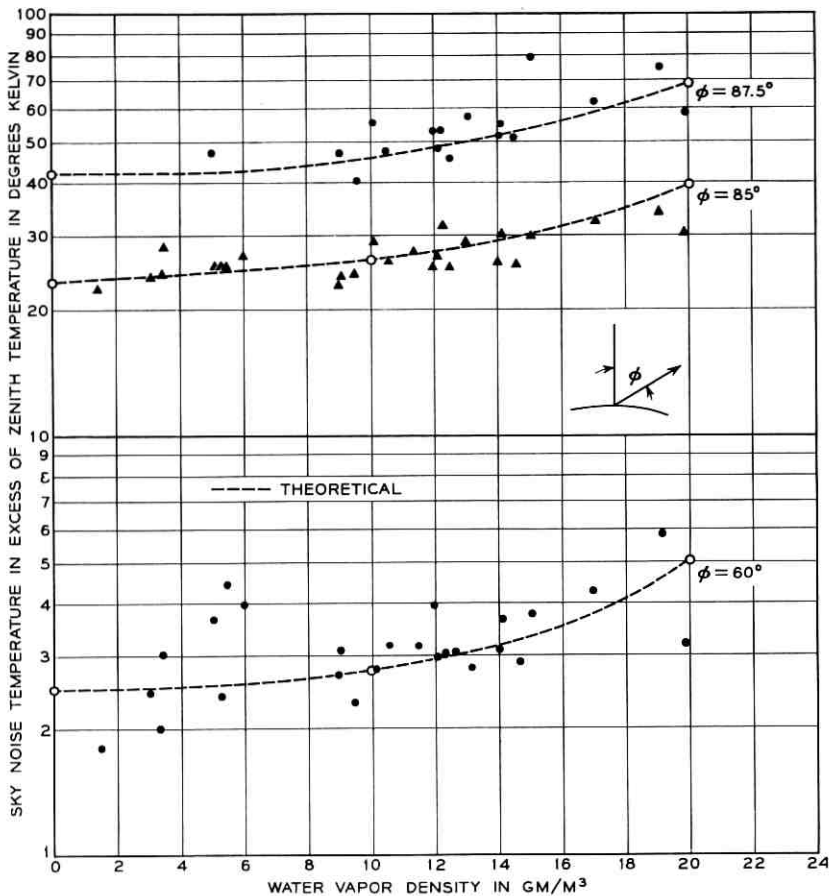


Fig. 6 — The effect of water vapor on sky noise temperature.

ground rain rate of about 6 millimeters per hour is also indicated on the figure.* The increase in sky temperature due to the light rain is readily apparent. The curve titled "18 May 1960 2:45 p.m." in this group is interesting, for here, at beam position of 85° and greater, the noise temperature is essentially the same as it is for the dry day. Hence it appears as though the rain was fairly local at the time the measurements were made. All data discussed so far were obtained using the RF attenuator method.

* It should be emphasized that the indicated rainfall rates are measured at the antenna site. When the beam approaches the horizon, however, it looks through several hundred miles of atmosphere, which may contain rain of various densities. In this case, the measured rain rate at the receiving site is of little significance.

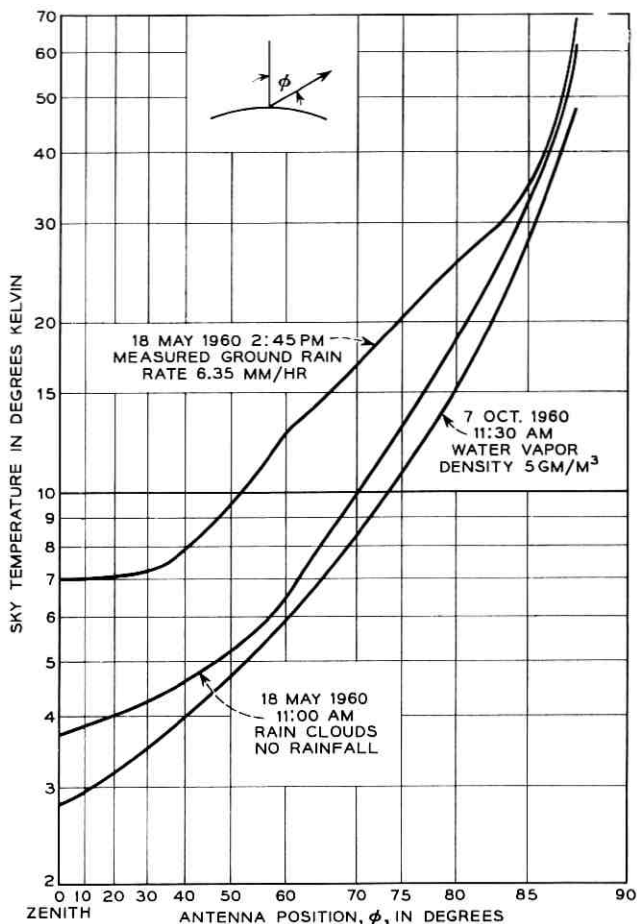


Fig. 7 — Sky noise temperatures for dry and wet days.

Sky temperature data for rapidly changing sky conditions which occur during periods of rainfall, obtained by means of the recorder, are shown in Figs. 8 through 12. These were selected as representative of the changes which usually occur in zenith sky temperature just prior to and on arrival of the rain at the ground. The left ordinate is zenith sky temperature, the right ordinate ground rain rate, with time on the abscissa. Before the arrival of the rain and after its departure from the location, the sky temperature may undergo variations on the order of several degrees. One should note on the measured data that an abrupt increase in sky noise is observed well before any significant rain is measured at

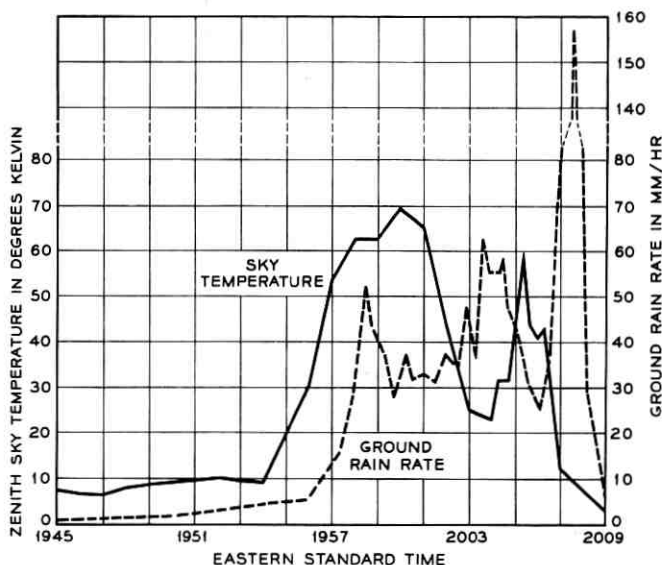


Fig. 8 — Sky noise temperatures observed during a rain storm (June 24, 1960); note that the sky temperature increases before the rain rate.

the ground; hence the sky noise is a precursor of rain at the ground. Most probably, the precursory action is due to a slope in the rain column as the water falls from a fairly localized cloud which is moving with the upper wind.* Thus the antenna beam sees a maximum water density before it is observed at the ground.

Of special interest are the data shown in the last portion of Fig. 9

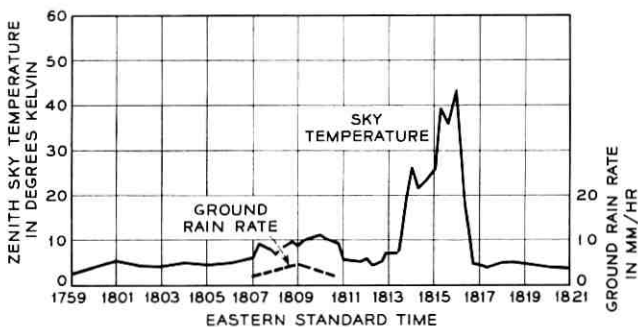


Fig. 9 — Relatively high sky noise with little or no measurable rainfall (June 24, 1960).

* See Fig. 9 of Ref. 7.

Note that, although the sky temperature increased more than 40°K , no rain was recorded. It is apparent here that the antenna beam looked through a cloud containing a large amount of condensed water. Even though no rainfall was recorded at the receiving site, the water from this cloud was no doubt released further along the path of the storm. The data in Fig. 9 are in direct contrast with those in the last portion of Fig. 8, for in the latter case a rain rate in excess of 150 millimeters per hour produced only moderate sky temperatures. A good example of the rapid and large changes that can occur in sky temperature is shown in Fig. 10: in a period of about 40 seconds it increased 66°K . The abrupt break in the sky temperature curve was due to a 20-second power interruption caused by lightning; the remainder of the interval was used in re-checking the calibration of the measuring equipment.

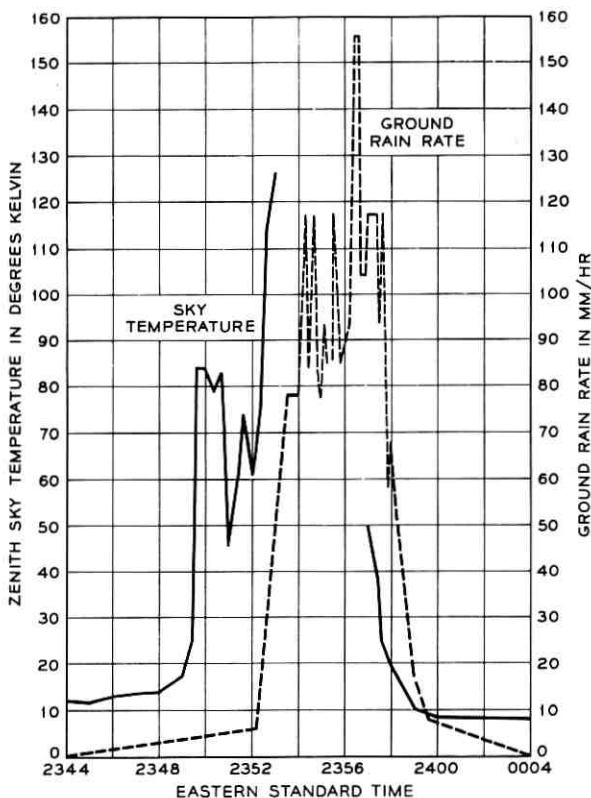


Fig. 10 — Data obtained during a severe thunderstorm (June 17, 1960).

There appears to be no detailed correlation between measured ground rain rate and zenith sky temperature; for example, in Fig. 11, fairly high sky temperatures were recorded when ground rain rates were low. This lack of correlation is further indicated in Fig. 12(a), for during the period shown the noise is fairly high but no measurable amount of rain fell. Fig. 12(b), however, shows that small variations in zenith temperature can occur with correspondingly low rain rates. Since data of this type are a measure of the water content of the atmosphere, they are of some meteorological significance and perhaps may serve as adjuncts to measurements by weather radar.

Since rain noise in space receivers plays the counterpart of signal fading in line-of-sight communications, a percentage-time distribution of sky noise is required for making reliability estimates. Although a relatively small statistical sample has been obtained so far, Fig. 13 shows both distributions of zenith sky noise temperature and ground rain rates. Additional measurements at the zenith and especially at beam positions near the horizon, where the background noise is higher, are needed to improve the statistics and to indicate the limit for a communication system. The median value of zenith sky temperature is 14.7°K or

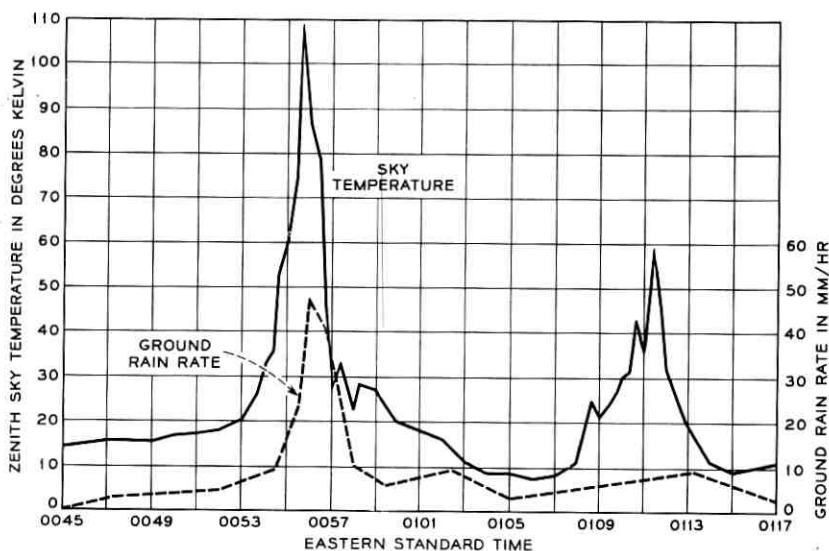


Fig. 11 — Example of lack of correlation between measured ground rain rate and zenith sky noise temperature (data of June 18, 1960). Compare these data with those in Fig. 8.

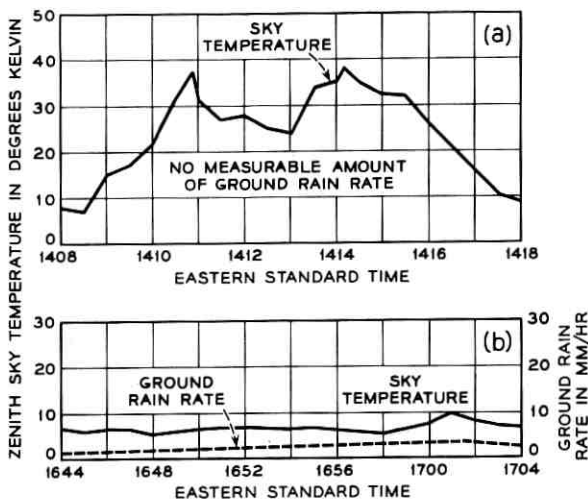


Fig. 12 — (a) Example of lack of correlation between measured ground rain rate and sky temperature (June 3, 1960); (b) small variations in sky noise temperature (July 27, 1960).

7.2 db above 2.8°K , the zenith temperature for an atmosphere containing a water vapor density of 10 grams per cubic meter, the corresponding median rain rate being 5 millimeters per hour. The noise temperature follows a log-normal distribution with a standard deviation of approximately 4 db. It should be cautioned that the relationships between the noise and rain rate distributions do not necessarily apply for any other general location, even though the ground rain rate distribution at that location is known.

It was mentioned in Section II that, if the distribution of rain along the antenna beam were known, one could calculate its noise contribution, since the absorption coefficient, α , is a known function of wavelength and rain density. If the assumption is made that the raindrops are uniformly distributed over a given interval along the beam, then the absorption $\alpha(r)$ is a constant over the interval. Calculations based upon (1), measured 6-kmc sky noise and measured ground rain rate, then can be made for the noise temperature which should obtain at other frequencies in the microwave band. For example, using this procedure,* a

* It should be noted that scattering of noise radiated from the earth into the antenna by the raindrops contributes to the total sky noise. However, calculations show that the scattering contribution is small. For example, again using a rainfall rate of 47.0 millimeters per hour and a measured 6 kmc temperature of 110°K , the contribution by scattering in the far field of the antenna is 3.3°K , the near field contribution 0.6°K , and the total scattered contribution 3.9°K , which is only 3.5 per cent of the total measured value.

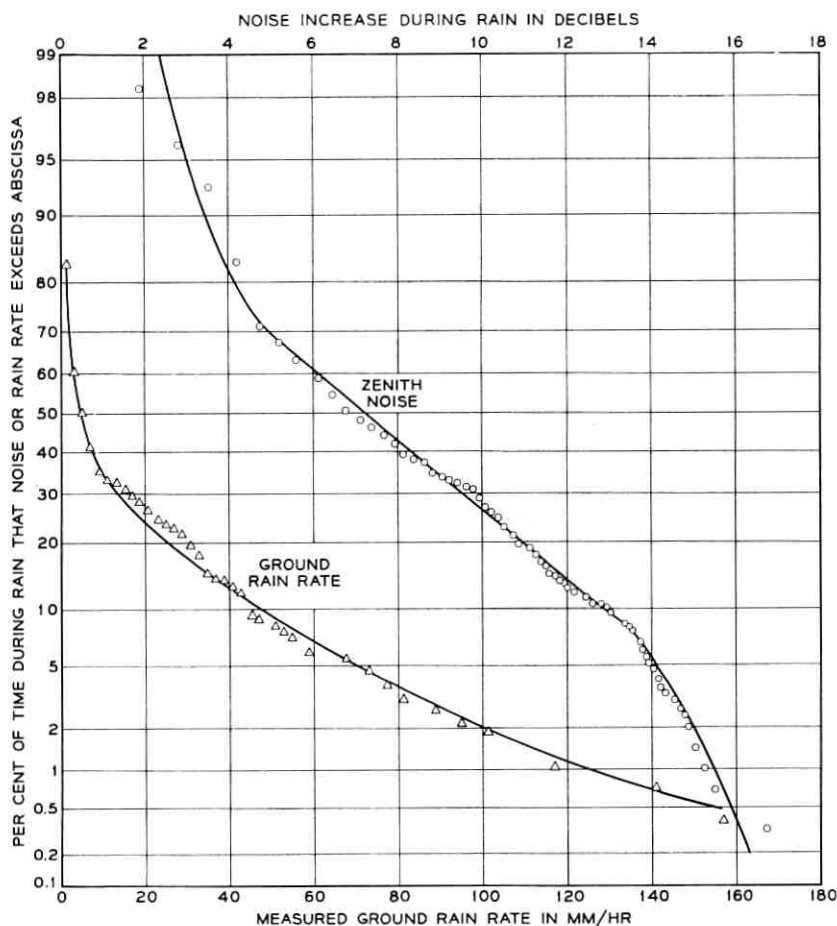


Fig. 13 — Percentage time distributions of zenith sky noise and ground rain rates, for eight rain periods from March 3, 1960 to July 27, 1960.

rainfall rate of 47.0 millimeters per hour that produced a measured temperature of 110°K at 6.0 kmc would produce a corresponding temperature of 34°K at 4.0 kmc, a decrease of 5.0 db.

Snow is another form of condensed water prevalent in the atmosphere. Noise measurements were made during a snowfall on January 27, 1961, but these show only a slight increase over those obtained on a clear day. Since snow is a fairly low-loss dielectric at centimeter wavelengths, this result is not unexpected.

VI. CONCLUSIONS

Based upon measurements of sky noise temperatures made in atmospheres devoid of rain at a frequency of 6 kmc, the following seasonal changes in sky temperature can be expected: at an antenna beam position 60° from the zenith, approximately $1\frac{1}{2}^\circ\text{K}$; at 85° , about 15°K . If the assumption is made that the desirable low-noise frequency band extends from 2 to 8 kmc, again neglecting the effects of rain, theoretical calculations establish the following limits: with an antenna beamed 60° from the zenith and for a frequency of 2 kmc, the seasonal change in sky temperature is 0.6°K ; for a frequency of 8 kmc, approximately 5°K . With an antenna beam 85° from the zenith at a frequency of 2 kmc, this change is 1.0°K and, at 8 kmc, it is 26°K . Data obtained at 6 kmc, with the antenna beam in the favorable zenith position, indicate that rain can increase the sky noise temperature as much as 17 db. In a space communication system, these changes in noise level are expected to play the counterpart of signal fading in line-of-sight communications; therefore suitable margins must be allowed for them.

REFERENCES

1. Pierce, J. R., and Kompfner, R., Transoceanic Communications by Means of Satellites, *Proc. I.R.E.*, **47**, 1959, p. 372; see also *B.S.T.J.*, **40**, July, 1961, Project Echo issue.
2. Hogg, D. C., Effective Antenna Temperature Due to Oxygen and Water Vapor in the Atmosphere, *J. Appl. Phys.*, **30**, 1959, p. 1417.
3. Tolbert, C. W., and Straiton, A. W., Anomalous Atmospheric Water Vapor Absorption of Millimeter Wavelength Radiation, U.R.S.I. meeting, fall 1960, Boulder, Colorado.
4. Gunn, K. L. S., and East, T. W. R., The Microwave Properties of Precipitation Particles, *Quart. J. Roy. Meteor. Soc.*, **80**, 1954, p. 522.
5. DeGrasse, R. W., Hogg, D. C., Ohm, E. A., and Scovil, H. E. D., Ultra-Low-Noise Antenna and Receiver Combination for Satellite or Space Communication, *Proc. Nat. Elect. Conf.*, **15**, 1959, p. 370.
6. DeGrasse, R. W., Schulz-DuBois, E. O., and Scovil, H. E. D., The Three-Level Solid State Traveling-Wave Maser, *B.S.T.J.*, **38**, 1959, p. 305.
7. Gunn, K. L. S., and Marshall, J. S., The Effect of Wind Shear on Falling Precipitation, *J. Meteor.*, **12**, 1955, p. 339

The Calculation of Metallic Hall Constants: Topological Aspects and Applications to Copper

By JOHN R. KLAUDER

(Manuscript received March 30, 1961)

The theoretical prediction of the Hall constant in metals appropriate to high magnetic fields depends, in part, on a repeated computation of a specific topological invariant: Each of several parallel two-dimensional planes in crystal-momentum space must be analyzed to find, on a single sheet of the Fermi surface, the number of nonintersecting closed contours that surround an arbitrary point. This number is related to the effective sign of the charge of the "relevant" electronic carriers in the plane — carriers whose contribution to the Hall constant is based on their orbital area. Applications of the technique discussed in this paper have been made to the Pippard model for copper, and various carrier contours are illustrated that were found in the course of three calculations of the Hall constant in cases where the magnetic field direction lay along a principal crystallographic axis.

I. INTRODUCTION

When a current-carrying metallic sample is placed in a uniform magnetic field, a change in the resistance is generally observed. This phenomenon, known as *magnetoresistance*, is the basis for an important experimental technique for studying the band structure of metals. If a simple model is taken to represent a metal (i.e., an effective mass approximation, a constant relaxation time, etc.), then the expected change in resistance can be theoretically predicted. Even with a more realistic model, it is found that one experimentally observable feature can still be predicted, namely, the Hall constant. The analytical expression for the Hall constant [presented in (4)] is an integral over a family of parallel two-dimensional planes. For each plane the integrand is pro-

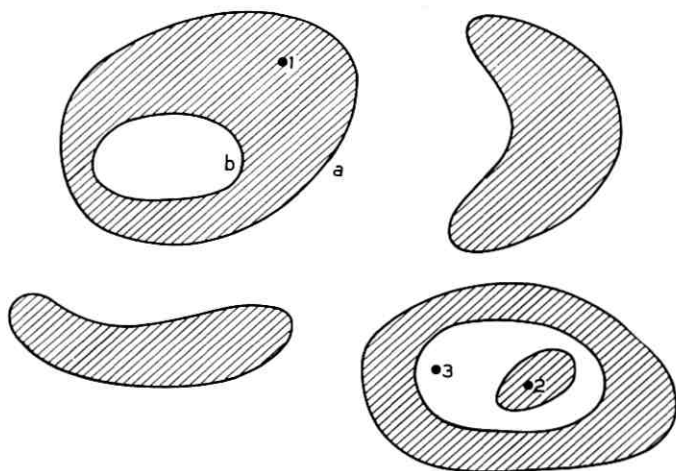


Fig. 1 — Several Fermi energy contours, shown in heavy lines, generated by the intersection of the plane $P_z = \text{constant}$ with a hypothetical constant energy surface.

portional to the area of that part of the plane surrounded by an odd number of zero-contours of a function defined in that plane.

Prediction of this resistance parameter turns out in general to be a somewhat complicated analytical test, and this paper discusses some interesting aspects of a numerical method developed by the author to calculate the Hall constant. In part, the calculation procedure repeatedly requires a *topological* investigation to be carried out on a two-dimensional plane in order to evaluate a particular topological invariant associated with that plane.

Stripped of the particular physical problem which motivates the study, the essence of the topological inquiry is the following: Suppose there is a container with many pieces of string of varying lengths. The ends of each piece of string are tied together so that each piece forms a closed loop. As many loops of string as desired are placed on a table, subject only to the requirement that they do not touch one another; it should be emphasized that loops within loops are permitted (see Fig. 1 for a suitable example). Now, select a point on the table not on a string itself, and ask for the number of loops which surround that point (thus, for example, point 1 in Fig. 1 is surrounded by one loop, loop *a*). This is clearly a topological question, since the number of loops which surround a point is unaffected by an arbitrary displacement of each loop — a displacement which in addition does not permit any loop to pass through the point of investigation.

A simple pictorial model should help to clarify the role of topology in the calculation of the high-field Hall constant. Fig. 2 shows a hypothetical energy surface with cubic symmetry, composed of "spheres" and "cylinders." The intersection of this surface with each of a family of parallel planes defines the set of contours that are pertinent for the Hall constant calculation. For definiteness, consider those planes, two of which are illustrated in Fig. 2, perpendicular to the $[001]$ axis. If the plane, like plane *a* in Fig. 2, cuts across the cylinders, then it is the area of the various cylindrical cross sections that contributes to the Hall constant calculation. However, for plane *b*, the second plane shown in Fig. 2, the contribution to the Hall constant is based on the planar area *outside* the spheres and cylinders, since now the regions in that plane surrounded by an odd number of contours lie outside the cylinders. Other planes are analyzed in a similar fashion.

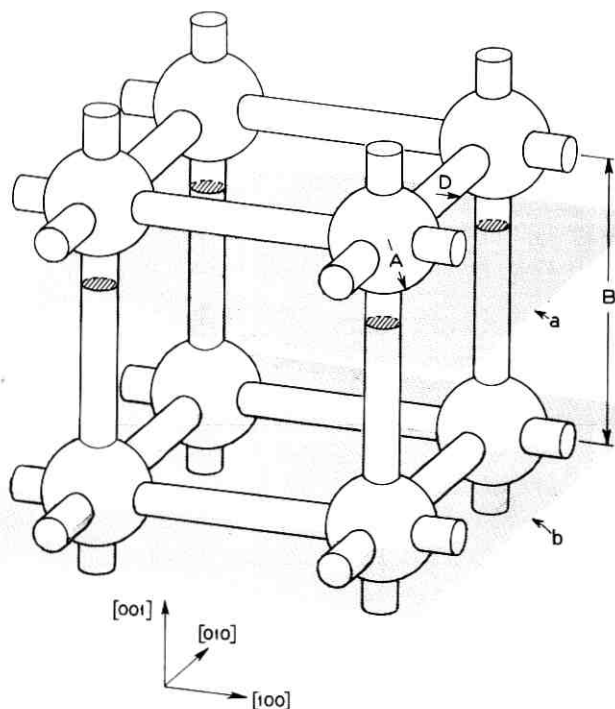


Fig. 2 — A portion of an ideal Fermi surface model of cubic symmetry composed of spheres of radius A and cylinders of radius D . The cube edge is of length B , and represents also the length of the cubic Brillouin zone, which is centered about each sphere. Two (001) planes, *a* and *b*, illustrate several different planar orbits.

In Section II, we define the customary resistance parameters in terms of the general resistance tensor that determines the electric field in terms of an applied current. The analytical expression for the high-field Hall constant is related to the direction of the magnetic field, the Fermi surface of the metal, and the symmetry of the crystal. In Section III those aspects of our calculation procedure that study the topology of nonintersecting planar contours are discussed in detail, and a logical procedure is developed, suitable for study with the aid of a digital computer, by which the "relevant" area may be ascertained in the most general case. In addition, a brief discussion is included of the significance of open contours, i.e., those contours that possess a net direction and never close upon themselves.

Section IV presents some of the numerical results obtained in the calculation of the Hall constant for copper. These results are based on the phenomenological model for copper that was proposed by Pippard.¹ Several figures are shown that depict orbits, both closed and open, computed for various charge carriers in planes transverse to the three axes [001], [011], and [111].

II. BACKGROUND AND ORIGIN OF THE PHYSICAL PROBLEM

In the most general situation, the relation between an applied current density (j_i) and the induced electric field (E_k) is described by a resistivity tensor (ρ_{kl}) according to the equation

$$E_k = \rho_{kl} j_l. \quad (1)$$

Here, and in what follows, all indices run from 1 to 3, and the summation convention is adopted. In the presence of a magnetic field (H_l), the tensor ρ_{kl} becomes a function of H_l , restricted by the Onsager relations so that $\rho_{kl}(-H_m) = \rho_{lk}(H_m)$. Two invariant expressions are commonly taken as descriptors of the resistivity tensor. The first is the resistivity ρ , defined by

$$\rho = \rho_k n_k n_l, \quad (2)$$

a double summation involving the unit vector n_l which is in the direction of the current j_l . The second invariant parameter is the Hall constant R_H , defined by

$$R_H = \frac{\frac{1}{2}(\rho_{kl} \epsilon_{klm} H_m)}{H^2}, \quad (3)$$

where, as usual, $H^2 = H_l H_l$ and ϵ_{klm} is the three-index, completely antisymmetric tensor defined as +1 (or -1) whenever (klm) is an even (or odd) permutation of (123). Observe that the resistivity, (2), is

sensitive only to the symmetric part of the resistance tensor, while the Hall constant, (3), depends on the antisymmetric part of ρ_{ikl} .

The theory developed by Lifshitz, Azbel, and Kaganov² and by Lifshitz and Peschanskii³ treats conduction in a magnetic field in a very general way. The underlying picture of their theory is most easily visualized in crystal-momentum space, the dual space of the ordinary three-dimensional space in which the metallic lattice lies. At very low temperatures the motion of the charge carriers in a metal is confined to a unique energy surface, the Fermi surface; the carrier motion is also confined to lie in one of the planes in momentum space perpendicular to the direction of the magnetic field H_l . The periodicity of the lattice in physical space leads to a periodicity of the reciprocal lattice in momentum space, and the periodicity of the reciprocal lattice is imposed on the energy surfaces imbedded in momentum space.

If a closed segment of the Fermi surface is contained within a Brillouin zone, the fundamental zone of periodicity which we identify with the usual or "proximity" zone, then the orbits on which the charge carriers move are all *closed* contours, when studied for any direction of the magnetic field. If the region immediately within an orbit contains energies *less* than the Fermi energy E_f (a value which defines the Fermi surface), then that orbit is said to be "electron-like." Conversely, if the immediately interior region has energies *greater* than E_f , then the orbit is "hole-like."

When the Fermi surface encounters the boundary of the Brillouin zone, then the periodicity requirement can lead to a connected surface extending throughout the entire momentum space. For instance, the simple cubic model in Fig. 2 represents a connected Fermi surface composed of "spheres," each in a separate Brillouin zone, that are connected by "cylinders." On such an open surface, some orbits — open orbits — can have an infinite extension, never becoming closed. In particular, planar open orbits directed along the cylinders exist for the model illustrated in Fig. 2. Nevertheless, even for open Fermi surfaces, certain directions of the magnetic field (whose direction dictates the family of parallel planes on which the orbits lie) lead to all closed orbits; i.e., each orbit in every plane is closed, at least in some finite number of Brillouin zones. The fundamental property of all closed orbits characterizes the class of magnetic field directions for which the Hall constant R_H can be theoretically predicted. Under these circumstances, the analytical expression,^{2,3}

$$R_H = \frac{h^3}{2ec} \left(\int dp_z \oint p_x dp_y \right)^{-1}, \quad (4)$$

defines R_H , where h is Planck's constant, $e > 0$ is the value of the electronic charge, and c is the speed of light. The integration over p_z represents a summation over the various momentum planes perpendicular to the magnetic field; now, and hereafter, the z direction is taken along the direction of the magnetic field. The domain of integration in (4) can be restricted to a single Brillouin zone, as will be discussed in Section 3.1.1. The contour integral extends around closed orbits in each $p_x p_y$ plane. If an orbit is electron-like, the contour integral in (4) is a positive number proportional to the area of the interior domain. On the other hand, if an orbit is hole-like, the contour integral in (4) is a *negative* number proportional to the area of the interior domain.

After suitable scale factors are introduced it suffices to study a completely *geometrical* factor in the Hall constant, which we call G : Let $p_i \rightarrow P_i$, where the P_i represent appropriate dimensionless variables; then

$$G \equiv \int dP_z \oint P_x dP_y. \quad (5)$$

This integral is based only on (a) the *direction* of the magnetic field, which selects the z direction, (b) the several plane *surfaces*, fixed in crystal-momentum space, which define the Brillouin zone boundary, and (c) the Fermi *surface*, also fixed in crystal-momentum space, on which the orbits are constrained to lie. In many substances the Fermi surface is multisheeted in each Brillouin zone. In this case a " G " may be determined for each sheet independently according to (5); the true " G " is the sum of the various individual ones. For simplicity, we confine our further analysis to a Fermi surface with only one sheet.

III. STUDY OF THE GEOMETRICAL FACTOR IN THE HALL CONSTANT

3.1 Computational Procedure

3.1.1 Domain of Integration and Subdivision of the Calculation

The formal definition previously given^{2,3} for (4) [or (5)] states that the integration domain covers all of momentum space (infinite in extent), but that the redundancy of the total number of Brillouin zones (also infinite) is to be divided out. The periodicity of the energy surface suggests that the numerical result obtained from an integration over this weighted infinite domain should be equal to the result of a momentum integration restricted to only one Brillouin zone. Indeed, if all the closed orbits lay within a single zone (as on a small sphere), then

the integration domain could readily be restricted to only one Brillouin zone.

However, when the Fermi surface is open, it is a common occurrence that all of the different closed orbits do *not* lie within a single zone. For example, for certain directions of the magnetic field, some of the larger orbits may close only after excursions through a number of different zones. It is the possibility that there may be some large orbits, which necessitates the seemingly redundant integration over a correspondingly larger domain of momentum space, so that within each P_z -plane, the true nature of the larger orbits, i.e., whether they are electron- or hole-like, may be ascertained directly.

If an independent determination of the nature of all the orbits — say, a “list” of their electron- or hole-like character — were already in hand, then G could be calculated by an integral over only one Brillouin zone. This limited integration would be valid because the “list” would already specify the true nature of any orbit for which only a *portion* of that orbit lay on a single plane within one Brillouin zone. It is this general scheme that we employ to limit the domain of integration to the customary first Brillouin zone centered about the point $P_x = P_y = P_z = 0$.

Must we examine every orbit in an arbitrary P_z -plane for our “list”? No, fortunately, as the following argument shows. In a fixed momentum plane, simultaneous electron- and hole-like orbits can exist on a single sheet of the Fermi surface only if one orbit type encloses another. Several different orbits are pictorially indicated in Fig. 1, separated by clear and shaded regions which represent, for example, energies above and below the Fermi energy. For the particular P_z -plane shown in Fig. 1, the geometrical factor G would acquire contributions from the entire area within the orbit marked a and also, with the opposite sign, the entire area interior to the orbit marked b . The net result consists in including, with a single sign, only the shaded area between the orbits a and b . This same line of reasoning applies to the remainder of the orbits in this plane, so that just the shaded area within the boundaries of the first Brillouin zone contributes to G . As a result only the “shaded area,” or, as we shall call it hereafter, the “relevant area,” corresponding to a *unique* sign of the energy relative to the Fermi energy, enters into the calculation of the Hall constant for each plane $P_z = \text{constant}$. Thus, to obtain the equivalent of a “list” of the orbits suitable for calculating G , it suffices to study just one point. If this point is surrounded by an odd (even) number of contours, then it is within (outside) the relevant area; all points whose energy E relative to the Fermi energy E_f is of the same (opposite) sign lie in the area relevant in calculating G .

3.1.2 *Certain Orbit Characteristics Determined by a Topological Study*

The relevant area in Fig. 1 is independent of whether the interior points have energies greater or less than the Fermi energy. Indeed, Fig. 1 graphically illustrates the following property: *Every point in the relevant area is enclosed by an odd number of closed orbits.* For example, point 1 is within a shaded region of Fig. 1 and is enclosed by one loop, and so on for points 2 and 3. Furthermore, this *characterization* of the points in the relevant area is invariant under any deformation of the contours such that no contour passes through the point under study. This invariance property is topological in nature, being independent of specific details of the orbit shapes.

In order to accurately determine the number of contours surrounding a given point, a portion of the plane $P_z = \text{constant}$ must be "scanned" — a portion generally much larger than that lying in the first Brillouin zone. Orbits which surround a point can only be found if they are actually encountered in the investigation; however, the restriction to a finite region of investigation is not serious in practice.

In the selection of a logical procedure to find the number of contours enclosing a particular point, simplicity is certainly a desirable feature. One very simple means of counting the contours (modulo 2, which is all that is necessary here) would seem to be the following: Pick a point not on a contour (i.e., $E \neq E_f$) and, starting there, investigate the energy values for a large sequence of closely spaced points all lying in a straight line. If an odd (even) number of crossings of the Fermi energy is found, the original point lies inside (outside) the relevant area. This intuitive procedure is, unfortunately, not always correct. Fig. 3 illustrates a set of contours in a fixed plane and shows three possible paths that could be used to determine the topology using the preceding scheme. Both paths 1 and 2 encounter three contours, and both emanate from shaded regions, i.e., relevant regions, as would be predicted by the above rule. However, path 3 illustrates the difficulty with this technique. In this example, the path ends inside a "sand-trap," and only two contours are encountered. Thus, we would erroneously conclude that the initial point did not lie within the relevant area. Clearly, a more sophisticated procedure of investigation is required that avoids the possibility of ending in a "sand-trap" by skirting around it.

Several alternate schemes of investigation come to mind which have the desired property of skirting around "sand-traps." In each of these an original test point is selected, and a series of points lying in some preselected straight line are tested for the crossing of a Fermi surface contour. If a contour is encountered, the *contour* is then followed, by

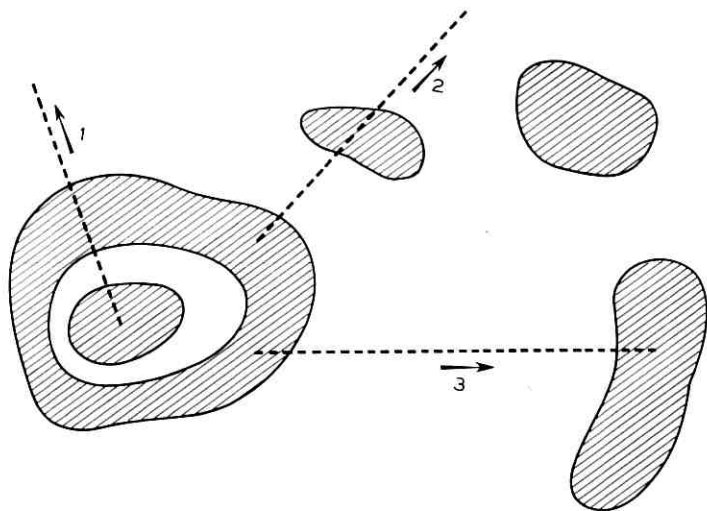


Fig. 3 — Fermi energy contours in a plane of fixed P_z . Three different paths, directed as shown, are tested in an attempt to determine where the shaded or “relevant” area lies.

always turning (with no loss of generality) initially to the left. The schemes differ in how far the contour is followed before the path leaves it.

One suggestive scheme has the path leave the contour as soon as it can proceed, without crossing the contour itself, in a radial direction whose origin is the original test point. A simple study reveals several fallacies in this scheme. Another scheme, of slightly greater complexity, has the path follow the contour until an extension of the original straight line of investigation is first encountered. At this point the path leaves the contour and continues in the original direction searching for a new contour. This scheme is likewise not foolproof in general, and must be abandoned. Among the contours it fails to analyze correctly is one shaped like the letter G if the test point is below the “hook” and the direction to be followed is upwards. Of course, in the general case, the location of the original point and the direction of inquiry are quite arbitrary, relative to the position and orientation of the contours to be studied.

The method of investigation illustrated in Fig. 4, although still more complicated, leads to an accurate count of the orbits surrounding any point. When a contour is encountered in the present scheme, that contour is followed *completely* until the point is reached at which the contour

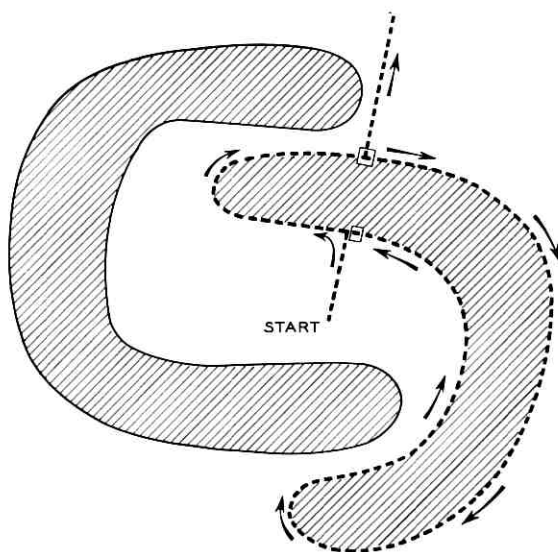


Fig. 4 — A satisfactory means of investigating an arbitrary layout of contours to establish the number of orbits which surround any point. It is applied here to a plane of fixed P_z that contains two distinct contours.

was originally engaged. While the entire contour is being examined, the number of points on the contour are counted that lie along a line which is an extension of the original line of investigation. Whether this number is odd or even determines whether that contour does or does not surround the original test point.* After the contour is completed, the path jumps to the furthest of these points, and then proceeds radially again along the original direction searching for a new contour to study. If a new contour is found, it is studied in completely the same fashion. Fig. 4 schematically illustrates the steps in a complete study of one contour by the present technique: the quest for the contour, the entire tracing of the contour, and the subsequent leaving in the original direction to continue the topological investigation. This investigation procedure is the one employed in the topological study of the various P_z -planes in calculating the geometrical part of the Hall constant by (5). A straightforward numerical integration, not discussed here, was used to determine the appropriate area in each plane that contributes to G .

* The more lengthy procedure of summing the polar angle variations along the contour can and has been used. This method can provide an exact count of the number of contours surrounding the original point, rather than an exact count modulo 2 provided by the method discussed in the text.

3.2 Additional Features of a Calculation Encountered in Practice

It was pointed out in Section II that the Hall constant cannot be theoretically predicted on the basis of (5) for magnetic fields which lead to a sizable, or nonnegligible, number of open orbits. However, this does not mean that open orbits will not be encountered in practical calculations of the Hall constant. On the contrary, open orbits may be present for one of several reasons. For example, all the magnetic field directions which support open orbits may not be known *a priori* for a phenomenological Fermi surface; only by trial and error can all such directions be discovered.

Another reason why open orbits may be encountered is based on the fact that numerical computations on digital machines can never be made with absolute precision. In particular, in the technique employed, a path of investigation is said to follow a contour so long as the value of the energy E satisfies the relation $|E - E_f| < \epsilon$, where ϵ is a small, nonzero number ($\epsilon \approx 10^{-4}E_f$), fixed by the gradient of the energy function and by practical computation requirements. It is entirely possible that two closed orbits which come close to each other cannot (computationally) be distinguished from a single connected orbit if the energy values satisfy $|E - E_f| < \epsilon$ along some path that connects the contours. If a sequence of closely spaced closed orbits exists, an error may be repeatedly or even unpredictably made, so that the path of investigation does not close. While the likelihood of errors being made decreases as ϵ is made smaller, a proviso must be made for encountering either one of two types of "open" contours: (a) valid open orbits for which the path proceeds through momentum space along an open contour possessing a net directivity, or (b) "lost" open orbits, which may or may not indicate the presence of true open contours, for which the path of investigation gets lost and wanders rather aimlessly throughout momentum space. The first type of open orbit, i.e., a valid open orbit, is detected to a high accuracy by placing an upper bound on the radial distance from the test point that the path travels in momentum space; the second type, i.e., a "lost" open orbit, is sensed by an upper bound on the number of individual points of investigation that can be made along any single contour. Whether an open orbit declared "valid" is truly open or just part of an exceedingly extended closed orbit is not too important, since the magnetic field is, after all, finite, and greatly extended closed orbits become indistinguishable from open orbits whenever the orbit circumference c divided by the mean free scattering length l is comparable with the magnetic field strength, suitably normalized.

Such considerations can in fact act as a guide in selecting the upper bounds placed on the path's excursion.

The presence of open contours generated by "tunnelling through an energy barrier" (i.e., along a path such that $|E - E_f| < \epsilon$) is not necessarily an academic case when ϵ is small. There is a small but very important class of magnetic field directions that lead to both closed and open orbits, but to the latter only on a set of P_z -planes of measure zero. The hypothetical connected Fermi surface illustrated in Fig. 2 is one for which sets of open orbits of measure zero exist when the magnetic field is along the principal crystalline axis [001], or equivalently along either the [010] or [100] directions.

Even in the presence of an infinitesimal number of open orbits, the geometrical part of the Hall constant is still theoretically predicted according to (5), since the contribution from those few P_z -planes is negligible. In fact, those magnetic field directions that lead to a negligible number of open orbits are often among the most interesting, for they are axes that possess high symmetry. When a Hall calculation is carried out for such a direction and one of the few open orbits is encountered (in one of the two ways discussed above), that entire P_z -plane is disregarded; no investigation of the size of the relevant area is or even can be made. Since the open orbits are ideally of zero measure, they should be encountered in the computation on at most a few of the large number of parallel planes studied that are perpendicular to the magnetic field. An open orbit of this kind, ideally belonging to a set of measure zero, will be illustrated in the next section.

Finally, there is some additional information that can be found from a knowledge of the extent and of the directions of the open orbits themselves. This additional information pertains to the expected behavior of the other magnetoresistance parameter, the resistivity ρ defined in (2). Generally speaking, it follows from theory^{2,3} that, on the one hand, the resistivity saturates in a high magnetic field for those magnetic field directions which lead to all closed orbits, and, on the other hand, the resistivity continues to increase as H^2 for those magnetic field directions which possess a nontrivial set of open orbits, all of which lead off in the same net direction. Thus, to predict qualitatively the resistivity behavior for various magnetic field directions, it suffices to study the geometrical shape of the Fermi surface in order to see for which directions of the magnetic field open orbits are permitted.

Conversely, the analysis of high-field resistivity data gained from measurements on high-purity metal single crystals is one of the best means available to gain information regarding the topology of the Fermi

surface, and even to provide some quantitative limitations on the shape of the Fermi surface. Therefore, a valuable auxiliary feature of a general procedure to calculate the Hall constant is the ability to shed some light on the expected resistivity behavior of a particular Fermi surface model by displaying the open orbits that are encountered for an arbitrary magnetic field direction. Preliminary discussions of the interpretation of the observed resistivity behavior for copper have been given elsewhere,^{4,5,6,7} and an extensive paper is in preparation in collaboration with J. E. Kunzler that will discuss magnetoresistance measurements and their explanation in some detail.

IV. APPLICATION TO METALS: COPPER

In this section we present some of the results of Hall constant analyses for a particular model of the Fermi surface of copper. Copper, or any of the noble metals, is particularly attractive for a discussion of the present type, since it involves a Fermi surface with only one "sheet"; the essential ingredients needed to discuss more complex metals are present without any unnecessary complications.

Copper has a face-centered cubic space lattice whose cube edge is 3.61 angstroms long. It follows that the reciprocal lattice in crystal-momentum space is body-centered cubic. The Brillouin zone that is characteristic of this lattice is an octahedron, truncated by a cube at the six vertices. If we make use of the cubic symmetry, then two inequivalent planes suffice to define the Brillouin zone boundary: one face perpendicular to the [111] axis — this face forms one of the eight surfaces of the octahedron; and three equivalent faces perpendicular respectively to the [100], [010], and [001] axes — these faces truncate the octahedron.

Because copper has only one valence electron, the region within the first zone that is also contained within the Fermi surface must equal one-half the total volume of the first Brillouin zone. Subject to this volume requirement, the confining surface may, in principle, be arbitrarily deformed. Suppose, as an extreme, the Fermi surface assumes the free-electron spherical shape. It is readily deduced from the volume constraint that this sphere does not intersect the boundary of the Brillouin zone, although it does come close to the octahedral faces. In this case the Fermi surface would consist of a number of disconnected spheres, each lying completely within a separate zone of periodicity. Only electron-like orbits would exist; no hole-like or open orbits would be present. Consequently, the Hall constant would depend on the sum of all the electron-like areas, i.e., it would depend simply on the volume within the Fermi surface. The observed Hall constant would be the same for

any orientation of the magnetic field, and by (5), $G \equiv G_{\text{free}} = \frac{1}{2}$, expressed in units in which the volume of the Brillouin zone is unity. Such a state of affairs would be present if the conduction electrons were free and felt no perturbing influence from the lattice.

In real copper the electrons are, of course, not free, and the Fermi energy surface deviates from a simple sphere. Pippard¹ in 1957 was the first investigator to successfully derive a phenomenological, nonspherical Fermi surface for copper. Speaking picturesquely, the surface he found from anomalous skin effect measurements possesses "arms" or "necks", which extend outward from the Fermi surface, that actually contact the hexagonal Brillouin zone faces associated with the $\langle 111 \rangle$ directions, i.e., the $[111]$ and equivalent directions. Each area of contact is about 10 per cent of the area of one of the hexagonal zone faces. The Fermi surface of copper therefore consists of a sequence of sphere-like "bodies" situated on a body-centered lattice that are completely interconnected by "arms" which lie along $\langle 111 \rangle$ directions.

An "arm" along one of the $\langle 111 \rangle$ directions supports open orbits whose net direction is along the same $\langle 111 \rangle$ direction. However, in copper (and surely in various other metals) the shape of the "bodies" and "arms" is such as to provide support for open orbits along additional directions. The origin of the support for additional open orbits is straightforward and is discussed elsewhere.⁸ Resistivity measurements bear out in detail the presence of the primary open orbits along $\langle 111 \rangle$ directions, as well as additional open orbits that the $\langle 111 \rangle$ "arms" support,⁷ principally in the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions.

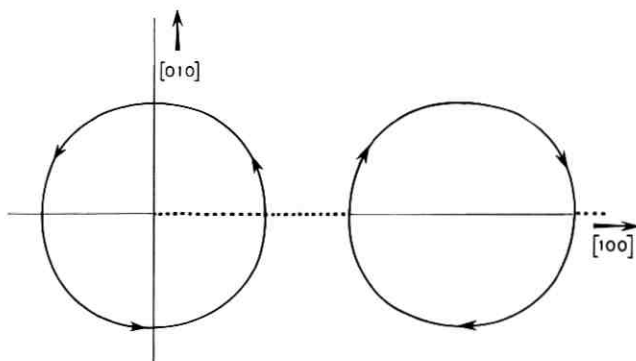


Fig. 5 — Contours on the Pippard Fermi surface in the $P_z = 0$ plane transverse to the $[001]$ axis. The topological path of investigation proceeded along the dotted line and around the contours in the direction of the arrows. This plane was determined to be electron-like since at the initial point $E < E_f$.

The periodic energy surface determined by Pippard has been expressed to an accuracy of 1 per cent by Garcia-Moliner⁹ in terms of a three-parameter Fourier series. His analytical expression has been employed in calculations on IBM 704 and 7090 computers to study the Hall constant for various directions of the magnetic field, based on the procedure developed in this paper. In these calculations, electron-like, hole-like, and open orbits were all encountered. Some of the contours that are present when the magnetic field is along either a $\langle 100 \rangle$, $\langle 110 \rangle$, or $\langle 111 \rangle$ type direction are shown in Figs. 5 through 9. In these figures the path of investigation is explicitly exhibited by which the relevant area was determined according to the procedure indicated in Fig. 4: The dotted line denotes that part of the path, always proceeding to the right, for which contours are being sought. The heavy line or lines represent the part of the path on which, in the direction indicated by the arrows, a contour on the Fermi surface is being followed.

Figs. 5 through 7 represent the behavior in three parallel (001) planes,

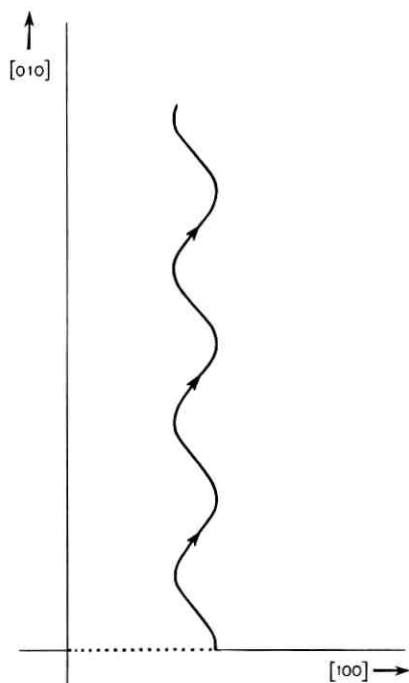


Fig. 6 — An open contour on the Pippard Fermi surface in the $P_z = 0.585$ [if center to (001) face of Brillouin zone $\equiv 1$] plane transverse to the $[001]$ axis.

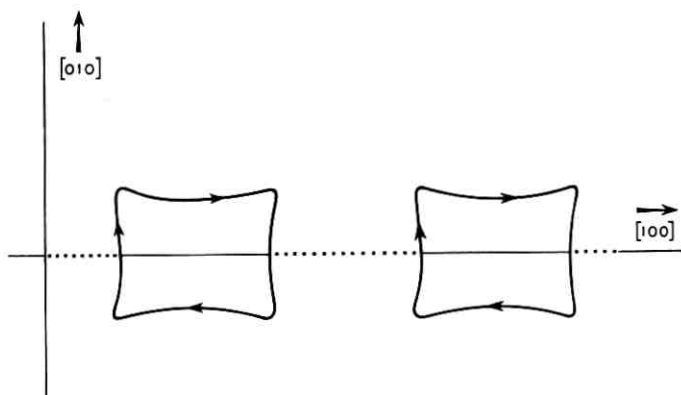


Fig. 7 — Contours on the Pippard Fermi surface in the $P_z = 0.575$ plane transverse to the $[001]$ axis. This plane was determined to be hole-like since at the initial point $E < E_f$.

the first at $P_z = 0$, which shows the contours of electron-like behavior around the “belly” of the Fermi surface. Fig. 6 shows an open orbit, ideally belonging to a set of measure zero, located on the (001) plane at $P_z = 0.585$, where for present purposes the distance to the (001) face of the Brillouin zone boundary is taken as unity. Just a short distance below this plane, at $P_z = 0.575$, the carrier behavior is quite different, being hole-like; the Fermi surface contours for this plane are shown in Fig. 7.

In Fig. 8, hole-like contours are illustrated that lie in the $P_z = 0$ plane when the magnetic field is along the $[011]$ axis. Oscillations in gold and silver of similar so-called “dog-bone” orbits have been observed in de Haas-van Alphen experiments.¹⁰

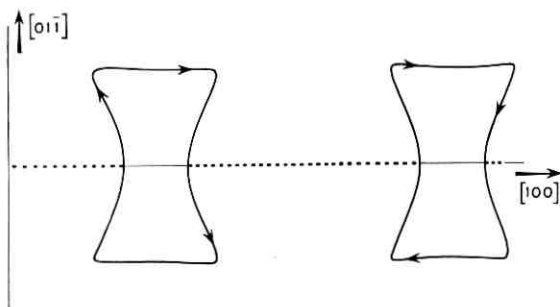


Fig. 8 — Contours (“dog-bone” orbits) on the Pippard Fermi surface in the $P_z = 0$ plane transverse to the $[011]$ axis. This plane was determined to be hole-like since at the initial point $E < E_f$.

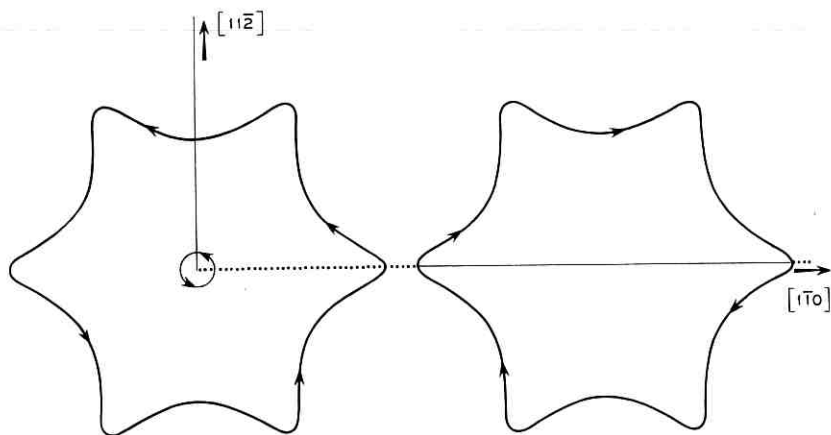


Fig. 9 — Contours on the Pippard Fermi surface in the $P_z = 0.866$ plane transverse to the $[111]$ axis, which is tangent to one of the hexagonal faces of the Brillouin zone. This plane was determined to be hole-like, since at the initial point $E < E_f$.

Fig. 9 shows the contours found when the magnetic field was in the $[111]$ direction and the particular plane of investigation lay on the surface of the Brillouin zone, i.e., $P_z = \frac{1}{2}\sqrt{3} \cong 0.866$, expressed in the same units as above. The small contour represents a cross section of one of the “necks” which connects the sphere-like “bodies.” However, the large hole-like contour, which surrounds the small one, determines the behavior in the plane for purposes of the Hall constant calculation. An investigation of the contours that terminated before the first big contour had been reached would have given an erroneous result, although it would have covered all of the area within the first Brillouin zone.

In the planes transverse to either one of the two types of symmetry directions, $\langle 100 \rangle$ or $\langle 111 \rangle$, closed orbits of both electron- and hole-like character exist in different planes, this being possible because, by rotational symmetry, only a finite (hence negligible) number of intermediate planes can occur that contain open orbits. This is not the case for copper if the magnetic field lies along a $\langle 110 \rangle$ direction, and the analysis of this particular direction requires an individual study, which will be given presently. For the former two symmetry directions the predicted and observed¹¹ values of the geometrical factor G —divided by the free-electron value G_f —are tabulated along with the Hall constant R_H in Table I. The ratio G/G_f measures the effective number of metallic electrons that are contributed per atom. The observed Hall voltage is very sensitive to field orientation in the neighborhood of the $\langle 111 \rangle$ direction,¹¹ and accounts for the increased experimental uncertainty. This extreme

sensitivity with orientation can be readily understood: It is due to the rapid transformation of hole-like regions similar to that illustrated in Fig. 9 into electron-like regions as the magnetic field orientation deviates from a $\langle 111 \rangle$ axis.

When the magnetic field is in, say, the [011] direction the Fermi surface of copper supports not only planes of hole-like or electron-like behavior separated by a finite number of planes with open orbits, but also a small yet nontrivial band of planes that support [100]-directed open orbits. In principle, these open orbits are broken into very large but nevertheless closed orbits if the magnetic field is reoriented very near to the [011] direction with a slight component along the [100] direction. If the angle of deviation from the [011] direction is small, then only an extremely small number of any new open orbits could have been introduced. Consequently, it appears quite permissible to treat the Hall constant in the immediate neighborhood of a $\langle 110 \rangle$ direction as numerically significant *if* the magnetoresistance exhibits a point of saturation in the proximity of the $\langle 110 \rangle$ axis, which indicates that at some point open orbits were actually absent experimentally, *and* if the measured Hall voltage is a smooth function of the magnetic field orientation in that vicinity. The experimental results support both of the requirements in question: The magnetoresistance shows a point of saturation near the $\langle 110 \rangle$ axes,* and the Hall voltage shows a broad flat peak in the neighborhood of the $\langle 110 \rangle$ directions.^{7,11} Thus this Hall measurement is significant, and it was used to obtain the observed Hall constant for the $\langle 110 \rangle$ directions shown in Table I. The theoretically calculated value for Table I was computed by assuming that the magnetic field was along the [011] axis, so as to determine the principal contribution, including that from the hole-like regions. The band of planes with open orbits, which was actually detected, was then reinterpreted as a band of planes having electron-like behavior, the behavior each of these planes would acquire if the magnetic field were tilted ever so slightly in the [100] direction.

The precision of the various Hall constant predictions is estimated at 1 to 2 per cent. Better agreement with the experimental Hall constants can be obtained if a Fermi surface model is used whose "arms" or "necks" are increased in diameter. With fatter "arms," more hole-like planes, such as in Figs. 7 and 8, will be present for fields along $\langle 100 \rangle$ and $\langle 110 \rangle$ axes, and, as a consequence, higher Hall constants would be predicted. It is interesting that the results of de Haas-van Alphen measurements¹⁰

* See Ref. 7, especially Fig. 1b, which shows the saturation behavior near to an [011] axis.

TABLE I—A COMPARISON BETWEEN THE PREDICTED AND OBSERVED HALL CONSTANT FOR THREE INEQUIVALENT FIELD ORIENTATIONS IN COPPER

| Magnetic Field Directions | Predicted* | | Observed† | |
|---------------------------|--------------|---------------|--------------|---------------|
| | G/G_{free} | $R_H\ddagger$ | G/G_{free} | $R_H\ddagger$ |
| <100> | .71 | 10.3 | 0.59 | 12.6 ± .5 |
| <110> | .72 | 10.2 | 0.65 | 11.4 ± .5 |
| <111> | .32 | 22.7 | 0.37 | 20.0 ± 5.0 |

* Based on the Garcia-Moliner⁹ expression of the Pippard Fermi surface of copper.

† Based on the data of Kunzler and Klauder.¹¹

‡ The numbers in this column when multiplied by 10^{-12} represent the Hall constant expressed in volt-cm/abamp gauss.

also suggest that a more realistic copper Fermi surface would have fatter "arms" than the Pippard model.

V. ACKNOWLEDGMENTS

It is a pleasure to thank C. Herring for bringing this problem to my attention and for his interest in this work. Many valuable conversations with J. E. Kunzler are greatly appreciated, and a critical reading of the manuscript by E. I. Blount is acknowledged. The author wishes to thank Miss B. Cetlin for her advice and encouragement.

REFERENCES

1. Pippard, A. B., *Phil. Trans. Roy. Soc.*, **A250**, 1957, p. 325.
2. Lifshitz, I., Azbel, M., and Kaganov, M., *J. Exptl. Theoret. Phys. (U.S.S.R.)*, **31**, 1956, p. 31; *Soviet Phys. J.E.T.P.*, **4**, 1957, p. 41.
3. Lifshitz, I., and Peschanskii, V., *J. Exptl. Theoret. Phys. (U.S.S.R.)*, **35**, 1958, p. 1251; *Soviet Phys. J.E.T.P.*, **8**, 1959, p. 875ff.
4. Kunzler, J. E., and Klauder, J. R., *Bull. Am. Phys. Soc.*, **5**, 1960, p. 150.
5. Alekseevskii, N., and Gaidukov, Y., *J. Exptl. Theoret. Phys. (U.S.S.R.)*, **36**, 1959, p. 447; *Soviet Phys. J.E.T.P.*, **9**, 1959, p. 311.
6. Priestley, M. G., *Phil. Mag.*, **5**, 1960, p. 111.
7. Klauder, J. R., and Kunzler, J. E., *Proceedings of the International Conference on Fermi Surfaces*, John Wiley & Sons, New York, 1961.
8. Klauder, J. R., and Kunzler, J. E., *J. Phys. Chem. Solids*, **18**, 1961 p. 256.
9. Garcia-Moliner, F., *Phil. Mag.*, **3**, 1958, p. 207.
10. Schoenberg, D., *Phil. Mag.*, **5**, 1960, p. 105.
11. Kunzler, J. E., and Klauder, J. R., *Phil. Mag.*, in press.

The Resistance of Organic Materials to Attack by Marine Bacteria at Low Temperatures

By PRISCILLA L. STEINBERG

(Manuscript received April 13, 1961)

The use of various organic materials in submarine cable structures has brought about an interest in the anticipated behavior of such materials in a sea water exposure. Organic materials have been exposed in the laboratory at atmospheric pressure to marine bacteria at both 20°C and 5°C under aerobic conditions and 20°C under anaerobic conditions. The utilization of the test materials by the bacteria has been determined by either measurements of oxygen consumption or hydrogen sulfide production.

None of the polyethylenes, synthetic fibers, or other thermoplastics such as polytetrafluoroethylene, polycarbonate, or polyamide was utilized by either aerobic or anaerobic microorganisms. All the elastomers tested were utilized by aerobic bacteria, but a majority of them were resistant to attack by anaerobic bacteria. Jute fibers served as a carbon source for both aerobic and anaerobic bacteria. In general, the poly(vinyl chloride) materials plasticized with an external plasticizer served as a source of energy for aerobic microorganisms at both temperatures, and six of these were also utilized by anaerobic bacteria. It appears that the added external plasticizer was responsible for the attack, since further tests showed the pure resin and the internally plasticized poly(vinyl chloride) to be resistant to both types of bacteria.

Comparison of the data from 20°C and 5°C aerobic tests indicates greater microbial activity at the warmer temperature. Calculated rates of utilization of the susceptible materials showed that the elastomeric compounds were generally consumed ten times as fast when incubated at 20°C as at 5°C. The poly(vinyl chloride) compounds when compared similarly showed approximately a five-to-one ratio, while jute fibers were utilized twice as fast at 20°C than at 5°C.

I. INTRODUCTION

The interest of the Bell System in the extent of marine bacterial activity at great ocean depths stems from the knowledge that certain organic materials are susceptible to attack by such organisms. For example, studies made by ZoBell and Beckwith¹ on rubber products and by ZoBell² on hydrocarbons have indicated that marine bacteria are capable of utilizing these materials. Therefore, one phase of a biological testing program (see Fig. 1) initiated by the Laboratories in 1954 has been designed to determine the resistance of organic materials that were possibly applicable in ocean cable construction to attack by both aerobic and anaerobic marine bacteria. The first data collected from these laboratory tests have been reported earlier by Snoke,³ and showed that in many instances specific organic materials were capable of being utilized by marine bacteria at 20°C. Attack was predominately by aerobes, little attack having occurred due to sulfate-reducing anaerobes.

The 20°C temperature is representative of certain shallow water environments and is in the range which supports high microbial activity. Ocean bottom temperatures at depths greater than 1000 meters range from 5° to -1.5°C.⁴ The present work, concerned with deep water temperatures, involves the testing of the resistance of the materials to aerobic bacteria at an incubation temperature of 5°C and the comparison of the results with those at 20°C. It was assumed initially that the rate of activity would be somewhat lower at 5°C, and since, in most cases, there had been little attack by anaerobic organisms at 20°C, no 5°C anaerobic tests were conducted.

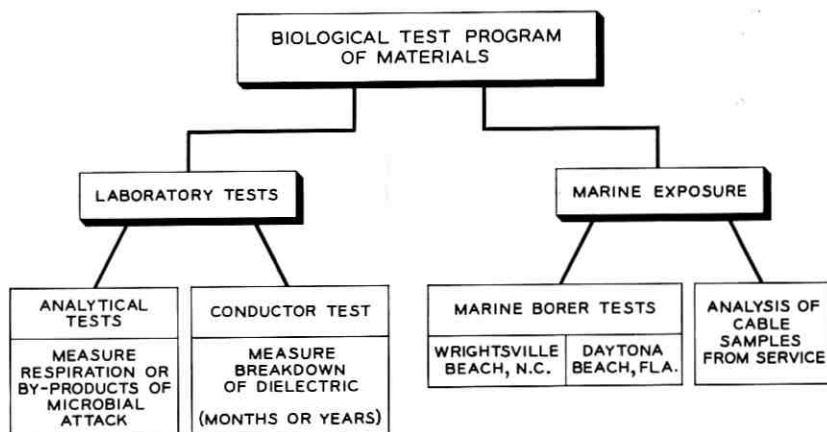


Fig. 1 — Outline of marine biological test program.

II. BACKGROUND

Marine bacteria are found in varying quantities over the entire ocean, with the greatest concentration being in coastal waters and in the region of the sediment-water interface. In spite of their minute size (generally less than one micron in length and one micron in diameter) they are so plentiful that the estimated mass of bacteria in the oceans has been placed at ten million tons.⁵ The concentration at a particular time at any location reflects the existing balance between productive and destructive factors. The numbers of bacteria are predominately due to the amount of organic matter present, since this provides the food source for the bacteria. An example of the influence of the quantity of organic matter on the bacterial population can be seen in the studies made by Hesse on water from the junction of the Atlantic Ocean and the colder Arctic Ocean.⁶ The change in temperature encountered here kills large quantities of marine organisms, thus providing an increase in the organic food supply, and has resulted in a great increase in the number of bacteria.

Other environmental factors may also have some influence on the bacterial population through effects on the supply of organic matter. It has been noted that a direct relationship exists between the bacteria and the diatoms, unicellular marine plants. These plants are capable of synthesizing organic matter from the simple chemical substances produced during the bacterial decomposition of extraneous organic material. Therefore, in areas of abundant diatoms there may also be an abnormally high number of marine bacteria. One area with a characteristically high population of marine bacteria is the mud-water interface. The population increase here is due to the rich food supply caused by constantly sinking marine plant and animal remains. Of this utilizable organic matter, Waksman and Carey found that approximately 60 per cent consumed by the bacteria is oxidized to carbon dioxide and water, and that the remainder is used to build new bacterial cells.⁷

Marine bacteria have been found in bottom deposits at depths of 10,280 meters below sea level, where the hydrostatic pressure is approximately 1000 atmospheres.⁸ Cattell⁹ indicates that single-cell microorganisms can tolerate pressures up to 6000 atmospheres. Several species of bacteria studied by Sanborn¹⁰ were active at a temperature of -5°C , which is slightly lower than typical bottom temperatures in the deep ocean. However, in spite of the ability of marine bacteria to adapt to different conditions, there are several factors that do restrict their number in the ocean. First, the amount of organic matter in the open sea is limited, and of this supply a significant part cannot be utilized by the

bacteria. Second, bacteria are a major part of the food supply consumed by protozoa and other small marine animals. Third, as the rate of bacterial multiplication increases, a toxic substance injurious to their development may begin to form and thus partly inhibit further growth.

Temperature seems to influence the bacterial environment, mainly as a factor controlling the extent of activity, but with little effect on gross population figures. Studies have indicated that bacteria found in the coldest waters multiply more slowly but live longer. ZoBell⁴ has also found that this relationship holds true for sea water stored in glass receptacles at incubation temperatures ranging from 0° to 25°C. Temperature may restrict the number of species of bacteria present in a given location, but the only time it becomes a factor influencing the total bacterial population is prior to the establishment of an equilibrium within the system.

This brief outline of marine bacteria illustrates the complexities facing the experimenter who is interested in trying to simulate their ecology. Since there are so many variables, any laboratory tests will necessarily be limited by the particular set of conditions satisfied during incubation of the test bottles. The tests reported here were conducted under laboratory conditions using aged and filtered sea water, laboratory-incubated enrichment cultures, and small test samples. Due to the great number of samples, all could not be tested at one time. Also, any chemical changes which might occur in the materials due to prolonged exposure were not reflected. The data collected in these tests do provide immediate information on the relative performances of the materials under the test conditions, and can be used as a gauge in judging materials until natural exposure data are available.

The first results from early inspections of samples exposed directly in the ocean were reported by Snoke;³ more recent data will be made available in a future paper.

III. DESCRIPTION OF LABORATORY TESTS

The two types of laboratory tests conducted are designed to test the resistance of the various organic materials to both aerobic and sulfate-reducing bacteria. The latter test, since sulfate-reducing bacteria are ubiquitous, gives an indication of the extent of anaerobic attack. In both cases, the procedure consists in exposing the test materials to bacteria which must utilize the material as a carbon source if they are to live.

Whenever possible, the materials to be tested are obtained in the form of film about 4 mils thick. Each bottle contains enough sample to expose a surface area of 12.9 square centimeters, which, in the case of film, is a 1-inch-square sample. The elastomeric materials are approximately

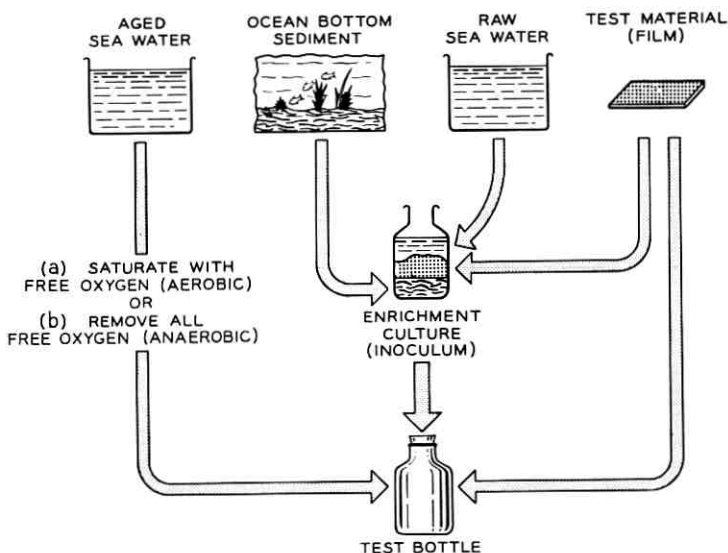


Fig. 2 — Flow chart of laboratory bacteriological test.

75 mils thick and are therefore cut in $\frac{1}{16}$ -inch-square samples. It is difficult to measure the surface area of granular samples and, as a result a selected mass, namely 0.05 gram, is used for all such samples or for jute fiber. The sample size for the other fibers, namely acrylonitrile-vinyl chloride copolymer and 6-6 polyamide (Source Ib)*, was determined by an approximation of the desired 12.9 square centimeter surface area.

3.1 *Aerobic Bacteria*

The biochemical oxygen demand (BOD) type of test measures the amount of dissolved oxygen consumed by aerobic bacteria during an eight-week incubation of the organic materials in aged and inoculated sea water. Since the test samples supply the only source of carbon available to the bacteria, the rate of oxygen consumption, which is a measure of the respiratory process of the living system, becomes a direct measure of the utilization of the material. The flow chart representing the steps involved in the test preparation is shown in Fig. 2.

At least eight weeks prior to running the test, sea water and marine sediment are collected from an area on the Atlantic coast. Some of the material to be tested is added to a portion of the collected sea water and sediment, and the mixture is inoculated from pooled marine en-

* Indicates source of supply.

richment cultures obtained through the courtesy of C. E. ZoBell of Scripps Institute of Oceanography. These latter were prepared using a wide range of organic materials including elastomers, carbohydrates, coal products, and other aromatic compounds. The enrichment cultures made up for each test sample are incubated at 25°C for at least six weeks prior to testing. The purpose of preparing separate cultures is to enable bacteria capable of utilizing the materials to develop preferentially under ideal conditions. In the meantime, the remainder of the raw sea water is filtered and aged in the dark until the utilizable organic content is about one part per million or less.

When these preliminary procedures have been carried out, the aged sea water is placed in a carboy and aerated with oxygen for 16 hours. Prior to filling the test bottles, a combined enrichment culture containing a few drops from each of the incubated cultures of the sample group being tested is added to the aerated sea water in a concentration of 1 milliliter per 10 liters of sea water. This concentration of inoculum provides maximum microbial activity with the least amount of added organic material. A slight positive oxygen pressure is placed on the water in the carboy until after the test bottles are filled, thus assuring oxygen saturation throughout the period preceding incubation.

Ten replicates of each sample are incubated in 60-milliliter glass-stoppered bottles in a constant-temperature water bath. Normally, two replicate samples are sacrificed for measuring the dissolved oxygen of the water following 0, 1, 2, 4, and 8 weeks of incubation. However, when a high rate of attack is anticipated the incubation intervals are shortened in order to obtain significant titration data before all the available oxygen has been utilized by the bacteria. Usually six materials are included in each test run with one set of control bottles. The 10 control bottles included contain inoculated sea water and are prepared exactly the same as the test bottles except for the absence of test material.

The dissolved oxygen content of the test bottles before incubation and following the different incubation periods is measured using a modified Winkler method. The basic Winkler procedure consists of the three steps shown in Fig. 3: (1) oxidation of manganous hydroxide in a highly alkaline solution, (2) acidification of the product in the presence of an iodide, and (3) volumetric titration of the liberated iodine with a standard thiosulfate solution using a starch indicator.

The oxygen consumed is calculated in parts per million from the results of the titrations of the sea water prior to incubation and following the longest period of incubation. The difference between these two values over and above any loss of oxygen in the control samples constitutes the biochemical oxygen demand (BOD) with the test material.

WINKLER OXYGEN DETERMINATION IN BOD TEST

(a) OXIDATION OF MANGANOUS HYDROXIDE IN A HIGHLY ALKALINE SOLUTION



(b) ACIDIFICATION IN THE PRESENCE OF AN IODIDE



(c) FREE IODINE TITRATED WITH A STANDARD THIOSULFATE SOLUTION

STARCH INDICATOR
(BLUE)STARCH INDICATOR
(COLORLESS)

HYDROGEN SULFIDE DETERMINATION IN ANAEROBIC TEST



(STARCH INDICATOR REMAINS BLUE FOLLOWING COMPLETION OF REACTION)

Fig. 3 — Reactions involved in laboratory marine analytical tests.

3.2 *Anaerobic Bacteria*

In the anaerobic test procedure the volume of hydrogen sulfide produced as a by-product during the utilization of the organic materials by anaerobic sulfate-reducing bacteria is measured as an indication of anaerobic attack.

Except for the handling of the aged sea water, the preparations for the anaerobic test are quite similar to those for the BOD test (see Fig. 2). In this test the sea water is boiled vigorously for 10 minutes and siphoned immediately into a carboy. A slight positive nitrogen pressure is placed on the carboy until the water is cooled enough to add the inoculum. Additional nitrogen pressure is kept on the carboy while the bottles are filled in order to prevent free exchange of oxygen with the water. Since the anaerobes function much more slowly than the aerobic bacteria the incubation periods selected are 0, 4, 8, 12, and 16 weeks in the water bath at 20°C.

Following incubation the hydrogen sulfide dissolved in the sea water is titrated with an iodine solution using starch as an indicator (see Fig. 3). Hydrogen sulfide content calculated over and above that found in the inoculated controls is considered a measure of anaerobic attack of the material in test.

IV. MATERIALS TESTED

The individual compounds which have been tested against aerobic and anaerobic organisms are listed in Table I. They included elastomers, poly(vinyl chloride), polyethylene, polypropylene, other thermoplastics, natural and synthetic fibers, and casting compounds.

4.1 *Elastomers*

With the exception of silicone rubber, all rubber samples were compounded at Bell Telephone Laboratories, steam cured, and supplied in sheet form approximately 75 mils thick. Although a few materials were submitted as both jackets and insulation the composition of the two varied considerably in some cases, and the two compounds behaved quite differently in test.

According to ZoBell and Beckwith,¹ even a highly purified nonvulcanized rubber sample can be utilized by marine bacteria. The presence of numerous waxes, oils, soaps, stearic acid, and phenolic compounds used in the compounding of the elastomers tends to increase the susceptibility of the rubber compounds to microbial attack. To date, no work has been done to determine which components in the elastomers were primarily responsible for the extent of attack experienced in the laboratory.

4.2 *Poly(Vinyl Chloride)*

All poly(vinyl chloride) samples submitted in sheet form were approximately 2 to 5 mils thick. All of the known plasticizers in the externally plasticized poly(vinyl chloride) samples contained typical organo-metallic-type stabilizers (such as barium, cadmium, and lead), fatty acid lubricants in low concentrations, and, in some cases, small quantities of inorganic fillers.

The unplasticized, rigid poly(vinyl chloride), although not containing a specific external plasticizer, was compounded with a significant amount of fatty acid lubricant and is therefore compared throughout the text with the externally plasticized compounds. This compound contained approximately eight to ten times as much fatty acid as was used in compounding the two internally plasticized acrylate copolymers designated semiflexible copolymers A and B. The rigid poly(vinyl chloride) should not be confused with poly(vinyl chloride) resin also discussed here, which was a pure poly(vinyl chloride) powder containing no added plasticizer or lubricant. Polyester A, C, and E plasticizers were fatty

TABLE I—MATERIALS TESTED AGAINST AEROBIC AND ANAEROBIC MARINE BACTERIA

| Material | Chemical Department Designation |
|--|---------------------------------|
| Elastomers | |
| GR-S jacket | BTL 54-14 |
| Natural rubber jacket | BTL 54-23 |
| Butyl rubber jacket | BTL 54-19 |
| Neoprene red lead jacket (with clay) | BTL 54-9 |
| Neoprene red lead jacket (without clay) | BTL 54-1 |
| Natural rubber insulation | BTL 54-24 |
| GR-S insulation | BTL 54-22 |
| Silicone rubber | BTL 53-83 |
| GR-A jacket | BTL 54-18 |
| Neoprene (ZnO, MgO jacket) | BTL 54-164 |
| Poly(vinyl chloride) — plasticizer used | |
| Polyester E (BTL 46-55) | P5503115 |
| Polyester A | BTL 24-54 |
| Di-2-ethylhexyl phthalate (DOP), Source 1 — sheet | P5606412 |
| None — rigid | P5503087 and P5402049 |
| Tri-2-ethylhexyl phosphate (TOP) | P5502081 |
| Nitrile rubber/polyester C | P5502074 |
| Nitrile rubber | P5502076 |
| Tricresyl phosphate (TCP) — sheet | P5311580 |
| Di-2-ethylhexyl phthalate (DOP), Source 3 | P5511691 |
| Proprietary compound | P5511690 |
| Di-2-ethylhexyl phthalate (DOP), Source 2 | P5502082 |
| Tricresyl phosphate (TCP) — granular | BTL 529-53 |
| Di-2-ethylhexyl phthalate (DOP), Source 1 — granular | BTL 23-54 |
| None (semiflexible copolymer A) | P5502078 |
| None (semiflexible copolymer B) | P5502077 |
| None (PVC resin) | P5510645 |
| Polyethylene | |
| 0.2 melt index + antioxidant — sheet | P5308390 |
| 0.2 melt index + antioxidant + 5% butyl rubber-sheet | P5602065 and P5403122 |
| 0.2 melt index (Source B) — sheet | P5312587 |
| 2.0 melt index (ASTM D1238) — granular | P5310466 |
| 0.2 melt index (Source A) — granular | P5304156 |
| 0.2 melt index (Source B) — granular | P5312587 |
| 0.2 melt index + antioxidant — granular | P5308390 |
| 0.2 melt index + antioxidant + 5% butyl rubber — granular | P5308396 |
| 0.7 melt index (high density) natural + antioxidant — sheet | P5512792 |
| 0.7 melt index (high density) + carbon black + antioxidant — sheet | P5503133 |
| Polypropylene | |
| Polypropylene + antioxidant A | |
| Polypropylene + antioxidant B | |

TABLE I—CONTINUED

| Material | Chemical Department Designation |
|--|--|
| Other thermoplastics | |
| Cellulose acetate Cellulose triacetate 6-6 polyamide (Source Ia) 6 polyamide (Source II) 6-10 polyamide (Source III) Polycarbonate Polyether Polyethylene terephthalate Polytetrafluoroethylene Trithene film | |
| Fibers | |
| Acrylonitrile-vinyl chloride copolymer 6-6 polyamide (Source Ib) | |
| Casting compounds | |
| Epoxide Styrene-polyester, silica filled | LRM 392, L2 (sheet) LRM 326, L1 (sheet) |

acid compounds with polyester C having been compounded with nitrile rubber for use as a test material.

Two of the compounds containing di-2-ethylhexyl phthalate (DOP) as the external plasticizer are labeled Source 1 and Source 2, and may be differentiated by the Shore A hardness test.* Source 1 had a Shore A figure of 88, indicating that it contained more plasticizer than Source 2, which had a Shore A figure of 62. The concentration of di-2-ethylhexyl phthalate in Source 3 is not known.

4.3 Polyethylene

The 0.2 melt index (Source A) and 2.0 melt index samples were prepared only in granular form. Data for some of the other polyethylenes were tabulated for both sheet and granular forms, but comparisons were only made when both series of tests involved the use of a similar form of material.

4.4 Polypropylene

Both polypropylene samples were submitted in sheet form about 15 mils thick. One sample contained antioxidant A, which included a phenolic compound, and the other contained an unknown antioxidant (B).

* See ASTM D-1706-59 T.

4.5 *Other Thermoplastics*

Most of the other thermoplastics tested were supplied as films between 1 and 5 mils thick. Of the polyamide samples tested, Source Ia is a 6-6 polyamide, Source II is a 6 polyamide, and Source III is a 6-10 polyamide. The cellulose acetate sample contained 54.5 per cent acetic acid, while the cellulose triacetate contained 61.5 per cent acetic acid.

4.6 *Fibers*

The jute used was an untreated roving. Acrylonitrile-vinyl chloride copolymer and 6-6 polyamide (Source Ib) were both supplied as yarn. Since both of these latter materials had been subjected to spinning oils during manufacture, they were exposed in test both as unwashed and washed samples. Petroleum ether was used as a wash medium.

4.7 *Casting Compounds*

The granulated epoxide compound was a casting resin set up through the use of an epoxy hardener. The sheet compound was a silica filled epoxy resin which had been pigmented with titanium oxide and hardened with an amine hardener. The resulting compound was quite brittle and about 13 mils thick.

The granulated styrene-polyester compound was also silica filled. The similar sheet material provided was filled with silica (39.7 parts) and glass fiber (5 parts) and pigmented white with TiO_2 . This compound was approximately 10 mils thick and was very brittle.

V. TEST RESULTS

In discussing the results of this test program the word "attack" is used quite freely when describing the utilization of organic materials by the bacteria. The word "attack" is used here to indicate that a material can serve as a source of food for marine bacteria rather than to indicate that it suffers rapid visible deterioration. In natural exposure this "attacked" material might perform quite satisfactorily for many years; in fact, many similar materials have been used successfully in marine environments over the years. However, the accelerated test data are valuable in screening various materials as to their comparative expected performances. Such data are of value in the selection of materials for submarine applications until long-range tests have been in operation for a considerable length of time. Also, when the natural exposure data do become available they can be correlated with the laboratory figures to provide additional information.

Figs. 4 and 5 show the extent of attack of the elastomers and the externally plasticized poly(vinyl chloride) samples as functions of time. They are based on data collected from the Winkler titrations, which measure the dissolved oxygen content of the sea water after the various incubation intervals. Each individual point on the curve was calculated as follows:

$$\frac{A - B_x}{A} \times 100 - C = \text{per cent of oxygen consumed} \\ \text{during particular incubation interval,}$$

where

A = dissolved oxygen (ppm) present in test bottles prior to incubation (average of two replicate samples),

B_x = dissolved oxygen (ppm) present in test bottles following the incubation period under consideration (average of two replicate samples),

C = per cent oxygen loss in the control or inoculated sea water bottles during the same incubation interval (x).

Whenever the available supply of dissolved oxygen was utilized entirely, no control factor (C) was used, and the result was reported as 100 per cent consumption. In general, the loss of oxygen in the control bottles during eight weeks' incubation at 20°C was between 10 and 15 per cent of the total dissolved oxygen present. In the tests run at an incubation temperature of 5°C, 5 to 10 per cent of the total oxygen was usually lost during the same interval.

In a few instances of low oxygen consumption the two-week titration figure showed more dissolved oxygen remaining than at the end of one week of incubation. Such anomalies may result from the fact that either (a) the control bottles showed more loss of oxygen than usual after two weeks of incubation or less than usual in the bottles titrated following one week, (b) the systems were not exact replicates, or (c) there was a small variation in the initial supply of oxygen. In such instances the higher oxygen consumption value was entered for both test periods to reflect the higher rate of attack.

To permit a ready comparison of the attack of different materials by aerobic bacteria, an "index of oxygen consumption" was calculated for each material. This index is the average daily consumption for the incubation period up to 100 per cent consumption or to the longest test interval (eight weeks), whichever is shorter. Thus the units of the index are properly "per cent oxygen consumed by the bacteria per day." It is

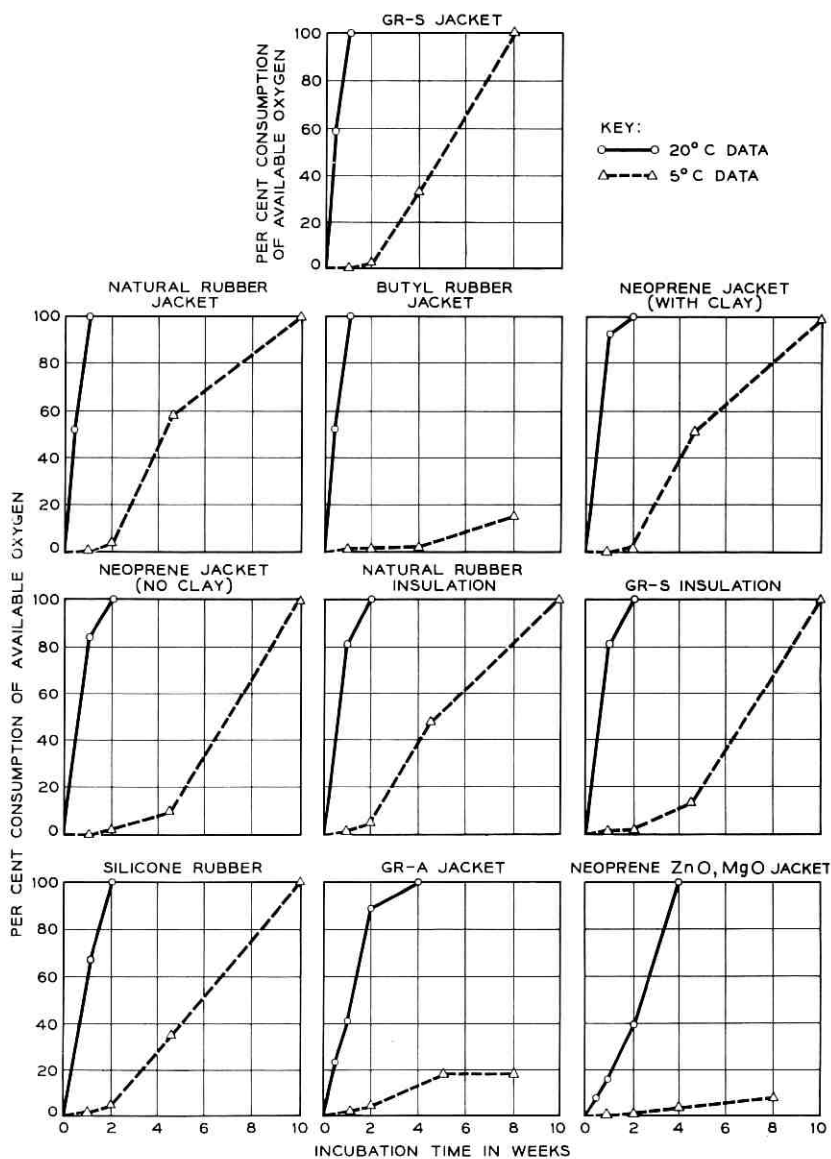


Fig. 4— Individual biochemical oxygen demand (BOD) test results on elastomers.

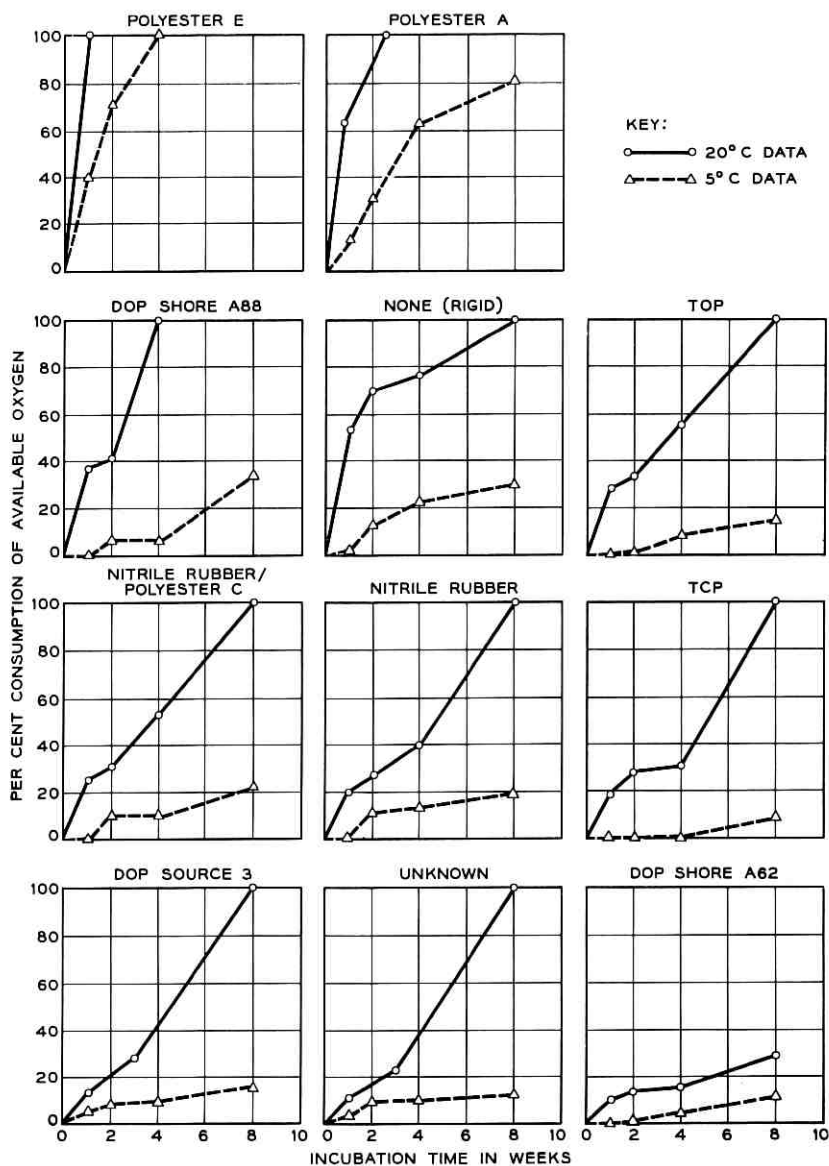


Fig. 5 — Individual biochemical oxygen demand (BOD) test results on poly(vinyl chloride) plastics; samples are designated by plasticizer used.

preferable, however, to disregard the units in dealing with the index, since the consumption rate is seldom linear; hence the index cannot be used to calculate oxygen consumption reliably. Its value is in providing a single figure which qualitatively reflects relative rates of attack while the curves of Figs. 4 and 5 provide more detailed comparisons.

The indices of oxygen consumption figures for all test materials are listed in Table II. In the interpretation of these data, a material was arbitrarily judged a potential food source for the marine bacteria if the index of oxygen consumption exceeds 0.5 per cent per day in the 20°C tests. Such materials are indicated by asterisks. Also, if BOD test results showed a system to be capable of utilizing more than 0.1 per cent of the available oxygen per day in the 5°C tests, the quality of the performance of this material was questioned.

The comparison of the indices calculated for the individual materials at 5° and 20°C has been made by a general grouping of the materials according to their rates of utilization, and may be found in Table III.

A list of materials that were not attacked in aerobic or anaerobic tests may be found in Table IV.

VI. DISCUSSION OF TEST RESULTS

Curves drawn for all of the externally plasticized poly(vinyl chloride) samples and the elastomers (Figs. 4 and 5) suggest that there are three general patterns of oxygen consumption by the marine bacteria. Essentially these are as follows: *Type A* — oxygen consumption rate follows a linear pattern as a function of time; *Type B* — oxygen consumption rate increases as the length of incubation time increases; and *Type C* — oxygen consumption rate decreases as the length of incubation time increases. In all cases involving the BOD testing of elastomeric materials at 20°C the oxygen was consumed in a manner illustrated by the Type A slope, in which the rate of oxygen consumption appears to be a linear function dependent on the length of incubation time. However, in most of these tests the total available oxygen was utilized so quickly that very little titration data were collected, and it was rather difficult to generalize concerning the manner of attack of the materials. In many cases only two points could be considered, since the exact time at which 100 per cent of oxygen had been consumed was not known. When the same materials were incubated at 5° only neoprene (containing zinc oxide and magnesium oxide) jacket was attacked at a rate directly proportional to time. GR-A jacket was utilized more readily at the beginning of the test, and the remainder of the elastomers were utilized more easily as the incubation time increased.

TABLE II — COMPILATION OF DATA FROM 20°C AND 5°C BOD TESTS AND 20°C ANAEROBIC TEST

| Material tested | Form of material tested | Individual 20° BOD test results — per cent available oxygen consumed | | | | Individual 5° BOD test results — per cent available oxygen consumed | | | | Index of oxygen consumption — per cent per day | | Anaerobic test results — total H ₂ S produced | |
|---|-------------------------|--|---------|---------|---------|---|---------|-------------------|-------------------|--|------|--|------|
| | | 1 week | 2 weeks | 4 weeks | 8 weeks | 1 week | 2 weeks | 4 weeks | 8 weeks | 20°C | 5°C | | |
| | | 3 days | 1 week | 1 week | 1 week | 3 days | 1 week | 1 week | 1 week | 0 | 0 | 19.9* | 1.2* |
| GR-S jacket | sheet | 59.8 | 100 | | | 0 | 3.5 | 34.3 | 100 | 17.8* | 1.8* | 3.0* | 0 |
| Natural rubber jacket | sheet | 53.4 | 100 | | | 0 | 3.5 | (32 days) 58.8 | (70 days) 100 | 17.6* | 0.3* | 9.3* | 0 |
| Butyl rubber jacket | sheet | 52.7 | 100 | | | 0 | 2.0 | 2.0 | 15.0 | 13.3* | 1.6* | 0.1 | 0 |
| Neoprene red lead jacket (with clay) | sheet | 93.2 | 100 | | | 0 | 1.0 | (32 days) 51.4 | (70 days) 100† | 12.0* | 1.4* | 0 | 0 |
| Neoprene red lead jacket (without clay) | sheet | 84.1 | 100 | | | 0.3 | 2.0 | (32 days) 10.2 | (70 days) 100 | 11.7* | 1.5* | 1.7* | 0 |
| Natural rubber insulation | sheet | 81.6 | 100 | | | 1.3 | 4.5 | (32 days) 48.3 | (70 days) 100† | 11.6* | 1.4* | 0.4 | 0 |
| GR-S insulation | sheet | 81.5 | 100 | | | 1.3 | 1.3 | (32 days) 13.0 | (70 days) 100 | 9.7* | 1.4* | 0 | 0 |
| Silicone rubber | sheet | 68.1 | 100 | | | 1.6 | 4.9 | (32 days) 35.8 | (70 days) 100† | 6.4* | 0.3* | 0.1 | 0 |
| GR-A jacket | sheet | 41.3 | 89.0 | 100 | | 1.0 | 4.1 | 18.0 | 18.0 | 2.8* | 0.1* | 0.4 | 0 |
| Neoprene jacket (ZnO, MgO) | sheet | 15.5 | 39.0 | 100 | | 0 | 0.1 | 2.8 | 7.4 | 2.8* | 0.1* | 0.4 | 0 |

Elastomers

Poly(vinyl chloride)—designated by plasticizer

| | sheet | 100 (5 days) 62.8 | (17 days) 100 | | | 39.7 | 71.7 | 100 | | 14.3* | 5.1* | 82.3* |
|---|----------|-------------------------|-------------------|------|------|------|------|-------|------|-------|------|-------|
| Polyester E | sheet | | | | | | | | | | | |
| Polyester A | granular | | | | | 13.3 | 31.7 | 61.5† | 81.2 | 10.3* | 2.2* | 47.1* |
| | sheet | 37.3 | 41.3 | 100 | | 0 | 6.2 | 6.2 | 34.0 | 3.0* | 0.6* | 1.5* |
| Di-2-ethylhexyl phthalate (DOP) — Source 1 | sheet | | | | | | | | | | | |
| Unplasticized (rigid) | granular | 52.4 | 69.5 | 76.6 | 100 | 2.4 | 12.6 | 22.2 | 29.6 | 2.7* | 0.5* | 5.0* |
| | sheet | 28.7 | 34.8 | 56.2 | 100 | 0.6 | 1.4 | 8.4 | 14.1 | 2.0* | 0.3* | 0.4 |
| Tri-2-ethylhexyl phosphate (TCP) | | | | | | | | | | | | |
| Nitrile rubber/polyester C | sheet | 26.1 | 32.4 | 54.9 | 100 | 1.2 | 8.8 | 8.8 | 22.2 | 2.0* | 0.4* | 18.0* |
| | sheet | 19.9 | 27.2 | 40.5 | 100† | 0.7 | 11.6 | 13.2 | 18.6 | 1.8* | 0.3* | 10.9* |
| Nitrile rubber | sheet | 18.2 | 28.0 | 30.7 | 100† | 0 | 0 | 0 | 9.5 | 1.8* | 0.2* | 0 |
| | sheet | 13.9 | (3 weeks) 28.8 | | 100 | 6.2 | 8.3 | 9.3 | 15.0 | 1.4* | 0.3* | 0 |
| Tricresyl phosphate (TCP) | granular | | | | | | | | | | | |
| Di-2-ethylhexyl phthalate (DOP) — Source 3 | granular | 10.4 | (3 weeks) 22.2 | | 100 | 2.7 | 9.3 | 9.3 | 12.8 | 1.1* | 0.2* | 0 |
| Unknown | granular | | | | | | | | | | | |
| Di-2-ethylhexyl phthalate (DOP) — Source 2 | sheet | 10.5 | 14.1 | 16.2 | 28.6 | 0 | 2.0 | 4.8 | 11.4 | 0.5* | 0.2* | 0 |
| Tricresyl phosphate (TCP) | granular | 14.4 | (17 days) 45.7 | 64.7 | 100 | | | | | 2.3* | | 0 |
| Di-2-ethylhexyl phthalate (DOP) — Source 1 | granular | 11.1 | (17 days) 44.3 | 52.8 | 100 | | | | | 1.9* | | 0 |
| None—copolymer A | sheet | 2.1 | 4.0 | 5.4 | 5.4 | 0.3 | 0.3 | 0.3 | 0.4 | 0.1 | 0 | 0.4 |
| None—copolymer B | sheet | 0 | 0 | 0 | 0 | 0.8 | 1.2 | 3.4 | 3.4 | 0 | 0.06 | 0 |
| None-poly(vinyl chloride) resin | granular | 1.0 | 1.0 | 1.8 | 1.8 | 0.6 | 1.9 | 1.9 | 4.3 | 0.03 | 0.08 | 0 |

* Denotes material utilized by marine bacteria.

† Denotes figure used to determine Index of Oxygen Consumption in cases of exception to the rule.

TABLE II — CONTINUED

| Material tested | Form of material tested | Individual 20° BOD test results — per cent available oxygen consumed | | | | Individual 5° BOD test results — per cent available oxygen consumed | | | | | Index of oxygen consumption — per cent per day | | Anaerobic test results — total H ₂ S produced |
|--|-------------------------|--|---------|---------|---------|---|---------|---------|---------|------|--|----------|--|
| | | 1 week | 2 weeks | 4 weeks | 8 weeks | 1 week | 2 weeks | 4 weeks | 8 weeks | 20°C | 5°C | 16 weeks | |
| | | | | | | | | | | | | | |
| Polyethylene | | | | | | | | | | | | | |
| 0.2 melt index + antioxidant | sheet | 5.4 | 10.8 | 10.8 | 10.8 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0 | 0 |
| 0.2 melt index + antioxidant + 5% butyl rubber | sheet | 2.4 | 4.4 | 5.5 | 9.9 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0 | 0.7 |
| 0.2 melt index (Source B) | sheet | 1.9 | 6.9 | 7.6 | 7.6 | 1.2 | 2.0 | 2.0 | 3.4 | 0.1 | 0.06 | 0.06 | 0.3 |
| 2.0 melt index — standard molecular weight polyethylene | granular | 0 | 0 | 0 | 0 | | | | | 0 | | | 0 |
| 0.2 melt index (Source A) high molecular weight polyethylene | granular | 0 | 0 | 0 | 6.4 | 0 | 0.8 | 0.8 | 0.8 | 0.1 | 0.01 | 0.01 | 0 |
| 0.2 melt index (Source B) | granular | 0 | 0 | 0 | 2.7 | | | | | 0.05 | | | 0 |
| 0.2 melt index + antioxidant | granular | 0 | 0 | 0 | 0 | | | | | 0 | | | 0 |
| 0.2 melt index + antioxidant + 5% butyl rubber | granular | 0 | 0 | 0 | 0 | | | | | 0 | | | 0 |
| 0.7 melt index (high density + antioxidant), natural | sheet | 0 | 0 | 0 | 0.8 | 0 | 0 | 0 | 0 | 0.01 | 0 | 0.01 | 0 |
| 0.7 melt index (high density), black | sheet | 1.8 | 1.8 | 1.8 | 1.8 | 0 | 0 | 0 | 0 | 0.03 | 0 | 0.03 | 0 |

Polypropylene

| | | | | | | | | | | | | |
|-------------------------------|-------|-----|-----|------|------|-----|-----|-----|-----|------|------|-----|
| Polypropylene + antioxidant A | sheet | 2.2 | 2.3 | 3.3 | 3.3 | 0 | 0 | 0.7 | 2.9 | 0.06 | 0.05 | 0 |
| Polypropylene + antioxidant B | sheet | 3.2 | 5.7 | 11.7 | 18.4 | 0.4 | 1.1 | 2.1 | 3.2 | 0.3 | 0.06 | 1.3 |

Other thermoplastics

| | | | | | | | | | | | | |
|-----------------------------|-------|-----|-----|-----|-----|-----|-----|-----|------|------|------|---|
| Cellulose acetate | film | 2.0 | 2.0 | 4.0 | 6.7 | 0 | 0 | 0.7 | 13.6 | 0.1 | 0.2 | 0 |
| Cellulose triacetate | film | 4.9 | 4.9 | 4.9 | 4.9 | 0 | 0 | 3.3 | 3.3 | 0.09 | 0.06 | 0 |
| 6-6 polyamide (Source Ia) | film | 0 | 1.6 | 4.0 | 8.4 | 0.4 | 1.1 | 1.1 | 2.2 | 0.2 | 0.04 | 0 |
| 6 polyamide (Source II) | film | 4.2 | 4.2 | 4.4 | 4.4 | 0 | 2.2 | 2.2 | 2.2 | 0.08 | 0.04 | 0 |
| 6-10 polyamide (Source III) | film | 1.1 | 1.1 | 1.1 | 1.5 | 0.3 | 0.3 | 0.3 | 1.4 | 0.03 | 0.03 | 0 |
| Polycarbonate | film | 2.7 | 2.7 | 3.2 | 4.3 | 0 | 1.4 | 1.4 | 1.4 | 0.08 | 0.03 | 0 |
| Polyether | film | 0.7 | 0.7 | 2.0 | 2.3 | 0 | 0.4 | 1.4 | 1.4 | 0.04 | 0.03 | 0 |
| Polyethylene terephthalate | film | 0.4 | 1.6 | 1.6 | 1.6 | 0 | 0 | 0 | 0 | 0.03 | 0 | 0 |
| Polytetrafluoroethylene | sheet | 1.5 | 1.5 | 1.5 | 1.5 | 0 | 0.8 | 0.8 | 1.6 | 0.03 | 0.03 | 0 |
| Trithene film | film | 0.3 | 0.3 | 0.8 | 0.8 | 0 | 0 | 0 | 1.8 | 0.01 | 0.03 | 0 |

Fibers

| | | | | | | | | | | | | |
|---|-------|------|-------------------|------|-----|-----|-----|-----|------|------|------|-------|
| Jute, untreated roving | fiber | 18.5 | (17 days) 56.6 | 45.0 | 100 | 1.6 | 7.3 | 9.4 | 54.2 | 2.0* | 1.0* | 51.6* |
| Acrylonitrile-vinyl chloride copolymer (washed) | fiber | 0 | 0 | 0.7 | 0.7 | 0 | 0 | 0 | 0.4 | 0.01 | 0.01 | 0.5 |

* Denotes material utilized by marine bacteria.

TABLE II—CONTINUED

| Material tested | Form of material tested | Individual 20° BOD test results — per cent available oxygen consumed | | | | | Individual 5° BOD test results — per cent available oxygen consumed | | | | | Index of oxygen consumption — per cent per day | | Anaerobic test results — total H ₂ S produced |
|---|-------------------------|--|---------|---------|---------|--------|---|---------|---------|-------|------|--|--|--|
| | | 1 week | 2 weeks | 4 weeks | 8 weeks | 1 week | 2 weeks | 4 weeks | 8 weeks | 20°C. | 5°C. | 16 weeks | | |
| Fibers — Continued | | | | | | | | | | | | | | |
| Acrylonitrile-vinyl chloride copolymer (unwashed) | fiber | | | | | 0 | 0 | 0 | 1.9 | | 0.03 | 0.5 | | |
| 6-6 polyamide (Source Ib) (washed) | fiber | 1.1 | 1.1 | 1.1 | 1.5 | 0 | 0 | 0 | 0 | 0.03 | 0 | 0.2 | | |
| 6-6 polyamide (Source Ib) (unwashed) | fiber | 0 | 0 | 0 | 3.5 | 0 | 0 | 0 | 0 | 0.06 | 0 | 0 | | |
| Casting compounds | | | | | | | | | | | | | | |
| Styrene-polyester, silica filled | sheet | 4.5 | 4.5 | 9.1 | 22.5 | 0 | 0 | 1.6 | 2.0 | 0.4 | 0.04 | 0.6 | | |
| Styrene-polyester, silica filled | granular | 13.2 | 14.3 | 14.3 | 25.0 | | | | | 0.4 | | 0 | | |
| Epoxide casting compound | granular | 3.8 | 5.7 | 5.7 | 6.6 | | | | | 0.1 | | 0 | | |
| Epoxide casting compound | sheet | 0 | 0 | 0.1 | 2.0 | 0 | 0 | 0.2 | 2.5 | 0.04 | 0.04 | 0 | | |

TABLE III—COMPARISON OF 20° AND 5°C BOD TEST RESULTS
BASED ON INDEX OF OXYGEN CONSUMPTION

| Material tested | Per cent oxygen consumed per day | |
|---|----------------------------------|----------|
| | 20°C test | 5°C test |
| Elastomers | | |
| GR-S jacket Natural rubber jacket Neoprene red lead jacket (with clay) Neoprene red lead jacket (without clay) Natural rubber insulation GR-S insulation | 10-20% | 1-2% |
| GR-A jacket Neoprene (ZnO, MgO jacket) | under 10% | under 1% |
| Butyl rubber jacket | 10-20% | under 1% |
| Silicone rubber | under 10% | 1-2% |
| Jute | | |
| Untreated roving | 2% | 1% |
| Poly(vinyl chloride), designated by plasticizer | | |
| Polyester E Polyester A | over 10% | over 2% |
| Di-2-ethylhexyl phthalate (DOP) — Source 1 None (rigid) | 2.5-10% | 0.5-2% |
| Tri-2-ethylhexyl phosphate (TOP) Nitrile rubber/polyester C Nitrile rubber Tricresyl phosphate (TCP) Di-2-ethylhexyl phthalate (DOP) — Source 3 Proprietary compound Di-2-ethylhexyl phthalate (DOP) — Source 2 | 0.5-2.5% | 0.1-0.5% |

Examination of the individual curves containing the BOD data for the externally plasticized poly(vinyl chloride) plastics shows some similarity to the rates of attack found in the rubber samples. At 20°C nine different test materials were utilized essentially in a manner which would fall into the straight-line or Type A oxygen-consumption rate curve. The two exceptions were the rigid unplasticized poly(vinyl chloride), which began to be utilized more slowly as incubation time increased, and the poly(vinyl chloride) plasticized with tricresyl phosphate, which began to be attacked more rapidly as time increased. At the 5°C incubation temperature only the poly(vinyl chloride) plasticized with poly-

TABLE IV—MATERIALS NOT ATTACKED IN EITHER BOD OR ANAEROBIC TESTS

| |
|--|
| Poly(vinyl chloride) |
| Semiflexible copolymer A (no plasticizer added) Semiflexible copolymer B (no plasticizer added) Pure poly(vinyl chloride) resin |
| Polyethylene |
| 2.0 melt index, standard molecular weight 0.2 melt index (Source A), high molecular weight 0.2 melt index (Source B), high molecular weight 0.2 melt index + antioxidant 0.2 melt index + 5% butyl rubber + antioxidant 0.7 melt index (high density) + antioxidant 0.7 melt index (high density) + carbon black + antioxidant |
| Polypropylene |
| Polypropylene + antioxidant A Polypropylene + antioxidant B |
| Other thermoplastics |
| Polyamide (Sources Ia, II, and III) Polyethylene terephthalate Trithene film Polytetrafluoroethylene Polyether Cellulose triacetate Cellulose acetate Polycarbonate |
| Casting compounds |
| Epoxide casting compound Styrene-polyester, silica filled |
| Fibers |
| Acrylonitrile-vinyl chloride copolymer 6-6 polyamide (Source Ib) |

ester E was utilized following the Type A pattern, while five samples — the rigid unplasticized poly(vinyl chloride) and those plasticized with polyester A, nitrile rubber, DOP Source 3, and a proprietary plasticizer — were attacked more slowly with an increase in incubation time. The five other poly(vinyl chloride) plasticizers — DOP Source 1, TOP, nitrile rubber/polyester C, TCP, and DOP Source 2 — were more easily attacked as incubation time increased.

It is highly possible that the slope variations noted in the individual

Type B curves could have resulted from testing under laboratory conditions. For instance, when a material has been attacked quite slowly at the beginning of a test and then more rapidly as the test progresses, it may be due to the fact that an equilibrium had to be established within the confines of the 60-milliliter test bottle before the bacteria could function effectively against the test material. Or perhaps the bacteria were unable to attack until a leachable plasticizer or lubricant or other compounding material had been released into the sea water in a form or amount that would be acceptable to the bacteria. Another factor which could be partially responsible for this type of situation was the fact that the enrichment cultures which were used for the 5°C tests were originally incubated at 25°C; this temperature could have temporarily inactivated bacteria, allowing them to become active only after incubation at 5°C.

With the Type C curve, laboratory conditions could also be held partially responsible for its shape. Possible explanations for a greater rate of attack at the beginning rather than at the end of the test might be that toxic materials were being released into the sea water, thus continually diminishing the rate of attack, or perhaps that the establishment of an equilibrium brought about a dormant system causing a decrease in microbial activity. This suggests that attack of the same materials in a natural marine exposure might follow a different pattern.

In comparing the curves for the same materials under different incubation temperatures it was noted that the only materials which showed similar patterns of attack at both temperatures were the poly(vinyl chloride) samples plasticized with polyester E and TCP, the unplasticized rigid poly(vinyl chloride), and the elastomer neoprene (containing zinc oxide and magnesium oxide) jacket. In all cases the rate of utilization was noticeably higher during the 20°C test.

In the remainder of the testing of rubber compounds the greatest difference between the patterns of oxygen consumption at 20° and 5° was noted during the first two weeks of incubation. The temperature differential appeared to be responsible for the length of incubation time needed to set up an active microbial system. This situation was also encountered in the testing of the four poly(vinyl chloride) samples plasticized with DOP Source 1, TOP, nitrile rubber/polyester C and DOP Source 2 at 5°C, in which the oxygen consumption rate increased as incubation time increased. These same materials followed the Type A pattern of oxygen depletion when incubated at 20°C.

The remaining four externally plasticized poly(vinyl chloride) plastics were attacked essentially in direct proportion to the extent of time in

test when incubated at 20°C, but exhibited a higher rate of attack at the beginning than at the end when tested at 5°C.

One other interesting feature was noted when comparing the curves of the individual materials at the two incubation temperatures. In a few cases, such as with the DOP Source 1 plasticized poly(vinyl chloride) sample, there was a short interval in both curves where the slope had leveled off. In the cited example this interval represented the second week of the 20°C test and the third and fourth week of the test, when the temperature was held at 5°C.

In comparing the indices of oxygen consumption of the individual materials at the two temperatures, the most noticeable feature appears to be that the elastomeric materials, when incubated at 20°C, were utilized essentially ten times faster than the same materials incubated at 5°C. Furthermore, examination of the relationship between BOD tests at 20° and 5°C involving the externally plasticized poly(vinyl chloride) samples shows that the ratio of attack was approximately 5 to 1. The ratio of the rates of attack witnessed in the BOD testing of jute was 2 to 1. No other ratios were noted for less utilized material, since the indices for both 20° and 5° data were too small to allow any accurate comparisons.

This classification system does not necessarily mean that the test data for each material exhibit this type of ratio. For instance, the relationship of attack at the two temperatures for polyester E plasticized poly(vinyl chloride) would be 14.3 to 5.1, for a ratio of 2.8 to 1 instead of 5 to 1, and for natural rubber insulation it would be 7.8 to 1 instead of 10 to 1. One extreme case was the BOD results for neoprene (containing zinc oxide and magnesium oxide) jacket, which gave the material an index of 2.8 per cent per day and a place on the "under 10 per cent" for 20°C data, while the 5°C index of 0.1 per cent per day correctly placed the material on the related "under 1 per cent" list. However, this produced an actual 28.0 to 1 ratio in the rates of attack. This is a broad classification of the performance ratios and, although most of the materials fall essentially within the assumed ratios, it is not infallible.

The materials which failed to fall into this group relationship were the elastomers butyl jacket and silicone rubber. The ratio of the results of the silicone rubber BOD tests were 9.7 to 1.4, thus being fairly close to the range of anticipated behavior. The placement difficulty came from the choice of 10.0 rather than 9.0 or even 9.7 as a breaking point between the two groups, and does not indicate any peculiarity of sample performance. In the case of butyl rubber jacket, the BOD tests have been rerun in an effort to determine its odd behavior when compared to the

remainder of the elastomers. However, in all tests conducted to date butyl jacket has not been readily attacked at 5°C. The reason for this phenomenon is not fully understood.

In studying these figures it may be noted that the elastomers were utilized by aerobic bacteria more readily than was any other group of materials. The total available oxygen was consumed in all of the ten BOD tests run on elastomers at 20°C. The time required for complete consumption ranged from 7 to 28 days with the indices of oxygen consumption varying from 19.9 to 2.8 per cent per day. In the 5°C tests the elastomers were usually attacked but at a slower rate, and butyl jacket, GR-A jacket and neoprene jacket containing zinc oxide and magnesium oxide still showed some resistance after eight weeks of incubation. Only three of the elastomers, namely butyl and natural rubber jackets and natural rubber insulation, served as a source of energy for the anaerobic bacteria.

The remaining general class of materials which provided a food source for the marine bacteria was the externally plasticized poly(vinyl chloride) samples. In all poly(vinyl chloride) tests that showed high oxygen consumption the materials in question had been compounded with fatty-acid-type lubricants or external plasticizers that contained fatty acids. Three of the group which were studied and showed no attack were the pure resin and two internally plasticized semiflexible copolymers. These observations indicate that the plasticizer or lubricant and not the poly(vinyl chloride) itself was utilized as a source of carbon. With incubation at 5°C the rate of attack was much lower, but the samples plasticized with polyesters E and A, fatty acid compounds, were attacked considerably faster at both temperatures than the remaining poly(vinyl chloride) samples. Also, these samples were readily attacked by anaerobes during 16 weeks of incubation.

In all tests conducted jute fiber was readily utilized, but it is rather difficult to compare rates of attack on jute with the various sheet materials due to the great variation in material surface area made available to the bacteria. The performances of cellulose acetate at 5°C and styrene-polyester and polypropylene containing antioxidant B at 20° were close to the arbitrary 0.5 per cent per day consumption index, but cannot be reported as definitely showing attack.

The majority of the anaerobic test data available was reported previously by Snoke³ but are also presented here in Table II in order to include all the laboratory findings in a single document. In general, only materials that were attacked by aerobic organisms have tended to show attack by anaerobic sulfate-reducing bacteria, and usually this was to a

lesser degree and over a longer period of incubation. Since the rate of attack was usually very slow, the anaerobic tests were only run at 20°, where, it was assumed, the higher hydrogen sulfide production would be. There were three cases of doubtful performances in this group, natural rubber insulation, poly(vinyl chloride) DOP Source 1 plasticizer, and the polypropylene sample containing antioxidant B.

VII. SUMMARY

The following general statements are made regarding the relationship between the 5° and 20°C BOD tests run on 57 materials to date:

1. Oxygen was consumed significantly during incubation at both temperatures for all the externally plasticized poly(vinyl chloride) samples and for all the elastomeric materials except neoprene jacket containing zinc oxide and magnesium oxide when incubated at 5°C. Jute was also utilized at both temperatures.

2. In each of the above exceptions the microbial activity, measured by the oxygen consumed, was greater during the 20°C incubation tests. The ratio of the rates of attack at the two temperatures for the elastomers was generally 10 to 1, while it was 5 to 1 for the externally plasticized poly(vinyl chloride) samples and 2 to 1 for the untreated jute roving.

3. No one figure could be used to describe adequately the rate of attack in any test, since the rates appeared to vary during the several intermediate incubation intervals. Sometimes the rate of attack seemed to be in direct proportion to the length of time of incubation, but quite often the rates either decreased or increased with an increase in incubation time. The type of rate curve for the same material at the two temperatures was often dissimilar.

4. One number was selected for each material representing an assumed average oxygen consumption per day. This figure, designated the Index of Oxygen Consumption, is presented as a means of simplifying the presentation of the data obtained from each test.

5. The oxygen consumption with three materials — cellulose acetate when incubated at 5°C and silica-filled styrene-polyester and polypropylene containing antioxidant B at 20°C — although very low, was somewhat higher than comparable control values.

6. The following were not attacked at either temperature:

- a. Polyethylene

- b. Internally plasticized poly(vinyl chloride) copolymers and pure poly(vinyl chloride) resin

c. Other thermoplastics with the possible exception of cellulose acetate at 5°C

d. Epoxide casting compound

e. Synthetic fibers acrylonitrile-vinyl chloride copolymer and 6-6 polyamide (Source Ib).

Although incubated for a longer time, the samples exposed to anaerobic conditions showed less utilization by sulfate-reducing bacteria. Only three elastomers, six externally plasticized poly(vinyl chloride) samples, and jute showed noticeable attack following 16 weeks incubation under anaerobic conditions at 20°C.

VIII. ACKNOWLEDGMENT

The teamwork of the chemical research department in providing the test samples and general assistance during the entire accelerated laboratory program is greatly appreciated. A special word of thanks is due to J. B. DeCoste, who coordinated the efforts related to this program within his department.

REFERENCES

1. ZoBell, C. E., and Beckwith, J., The Deterioration of Rubber Products by Microorganisms, *J. Amer. Water Works Assoc.*, **36**, 1944, p. 439.
2. ZoBell, C. E., Action of Microorganisms on Hydrocarbons, *Bact. Rev.*, **10**, 1946, nos. 1-2.
3. Snoke, L. R., Resistance of Materials to Marine Biological Attack, *B.S.T.J.*, **36**, 1957, p. 1095.
4. ZoBell, C. E., *Marine Microbiology*, Chronica Botanica Co., Waltham, Mass., 1944.
5. ZoBell, C. E., Bacteria of the Marine World, *Sci. Monthly*, **55**, 1942, p. 320.
6. Hesse, E., Bakteriologische Untersuchungen auf einer Fahrt nach Island Spitzbergen und Norwegen im Juli 1913, *Centralbl. für Bakt.*, **72**, 1914, p. 454.
7. Waksman, S. A., and Carey, C. L., Decomposition of Organic Matter in Sea Water by Bacteria. II. Influence of Addition of Organic Substances upon Bacterial Activities, *J. Bact.*, **29**, 1935, p. 545.
8. Morita, R. Y., and ZoBell, C. E., Bacteria in Marine Sediments, Office of Naval Research, Research Reviews, July 1956, p. 21.
9. Cattell, M., The Physiological Effects of Pressure, *Biol. Rev.*, **11**, 1936, p. 441.
10. Sanborn, J. R., Certain Relationships of Marine Bacteria to the Decomposition of Fish, *J. Bact.*, **19**, 1930, p. 375.

The Influence of Crystallization Conditions on Radiation Effects in Polyethylene — I. Crystallization from Dilute Solution and from the Melt

By R. SALOVEY and A. KELLER

(Manuscript received January 16, 1961)

Samples of Marlex 50 crystallized from dilute solution (yielding single crystals) and in the bulk form were irradiated by high-energy electrons and the solubilities compared. The single-crystal specimens remained fully soluble after a dose of 20 megarads, which rendered the bulk two-thirds to three-quarters insoluble. Since the initial free-radical concentration was identical and the chemical changes induced were not widely different in the two kinds of specimens, it is suggested that intramolecular crosslinks, occurring in a larger number in the single crystals, are responsible for this behavior. At high doses the same solubility limit is approached by both specimens, which is ascribed to the possibility of chain cutting becoming increasingly ineffective in the single crystal. In the light of these findings, the limitations of the conventional treatment of radiation effects are discussed and the possible significance of the present results for morphological problems is pointed out.

A number of further effects are reported in connection with melting behavior and spin resonance spectra that are at variance with results in the literature.

I. INTRODUCTION

The effect of ionizing radiation on polyethylene has been extensively studied in the past^{1,2} but no investigation has been reported concerning the irradiation of polyethylene single crystals.³ We have examined the influence of morphology on the irradiation behavior of polyethylene. These studies are purely exploratory and many of the physical tools involved were used only in "key" tests. We did not intend to pursue every effect found in the course of these experiments as each might

merit a specific research program. Indeed, the present research was largely tailored so as to be accommodated by the limited tenure of one of us* at Bell Telephone Laboratories. For this reason, some experiments which appear to be obvious corollaries of the present work could not be performed, and a rather arbitrary line had to be drawn to terminate this exploratory work. Nevertheless, we feel that our findings in their present state convey sufficient information of value to have it published in its present form.

The irradiation behavior of polyethylene is a difficult and controversial subject, which is being extensively and carefully studied in different laboratories.⁴⁻¹⁰ We have attempted to minimize uncertainties by employing a primarily comparative approach to the subject. In trying to establish differences between morphologically different polyethylenes, specimens under comparison received strictly identical treatments. Consequently, the many uncertainties that beset this field do not affect our principal experimental findings, which always refer to differences in identically treated materials.

This first paper will refer to single-crystal polyethylene prepared from dilute solution as compared with the melt-crystallized bulk specimen. We recall that such single crystals consist of uniform lamellae which are about 100 angstroms thick, the molecules approximately normal to the plane of the lamellae in a sharply and regularly folded configuration.³ The bulk is usually considered as an aggregate of amorphous-crystalline fringed micelles. The recognition of single crystals and numerous morphological observations support the idea of identifying similar structures in the bulk. The relation between single crystals and the bulk material is one of the key problems in present structural studies of polymers. It is difficult to ascertain whether the bulk has the same structure as the single crystals or represents a different kind of aggregation solely on the basis of direct structural examinations. Consequently, novel lines of approach are essential. Besides its intrinsic interest from the point of view of irradiation chemistry, we hope that this exploratory work will be relevant to the problem of polymer morphology.

II. EXPERIMENTAL DETAILS AND RESULTS

2.1 *Specimens*

The material studied was the linear polyethylene Marlex 50. Two kinds of preparations were used. In one case, the polymer was dissolved

* A. Keller.

in boiling xylene and crystallized isothermally at lower temperatures. Concentrations ranged from 0.02 to 0.05 per cent and crystallizations were conducted at 70° or 85°C. A suspension of single crystals formed which could readily be verified microscopically. The suspended material was collected by filtering and was vacuum-dried. Low-angle X-ray scattering³ revealed that the fold lengths were 110 and 130 angstroms for crystallization temperatures of 70° and 85°C respectively. The second kind of material consisted of melt-crystallized flakes. Preceding irradiation, specimens were placed in Pyrex tubes, pumped in high vacuum overnight, and sealed.

2.2 Irradiation

Irradiations were conducted within vacuum-sealed tubes in a beam of 1 million electron-volt electrons using a High Voltage Engineering Co. Van de Graaff accelerator. In order to minimize heating, samples were placed on a plate cooled by circulating cold water. The dose rate was determined by measuring the attenuation in current, with a Faraday cage, on interposing a known thickness of polyethylene in the electron beam. Moreover, current absorption by the tube was measured and accounted for. Dose rates of about 5 megarads per minute were ordinarily employed. The dose of 20 megarads was used for most comparative purposes for reasons to become apparent below, but a few experiments with other doses ranging from 2 to 128 megarads will also be quoted. Unless specifically stated, specimens were fused at 150°C while still sealed off after irradiation, which should lead to maximal changes in properties owing to the activation of the free radicals formed.^{6,7}

2.3 Solubility

Solubility determinations represent the principal measurements in this work. They were carried out by Soxhlet extraction with refluxing xylene, a treatment which fully extracts the unirradiated starting material.

Bulk and single-crystal specimens showed very marked differences in their irradiation characteristics. When irradiated by our standard 20-megarad dose the bulk became much less soluble, leaving an insoluble gel fraction of 65–75 per cent. The single-crystal preparation irradiated with the same 20-megarad dose remained fully soluble, leaving no gel in the extraction thimble. Most of the work which followed was aimed at evaluating this remarkable difference.

An investigation of the full dose-solubility characteristics could not

be fully accommodated within the present experimental program. Nevertheless, we can quote a few data obtained with other doses. The radiation-solubility characteristics of the bulk have been known from some earlier work by one of us.* Even if conditions were not rigorously identical with those in the present 20-megarad experiments, they are representative of the over-all trend. Accordingly, the gel point is roughly 2 megarads; 10 megarads already gives a gel content of 60 per cent, rising to 70 per cent at 20 megarads and to 80 per cent at 40 megarads. In the case of the single crystals, a gel content of 26.5 per cent was obtained with a dose of 48 megarads. A dose of 85 megarads yielded 76 per cent, and one of 128 megarads, 86 per cent gel. The last two irradiations were carried out under somewhat altered operational conditions of the generator, which necessitated leaving the specimens in air. Even if not strictly comparable with the "standard" experiment, these figures certainly bear out the fact that at higher doses a limiting value comparable with that of the bulk is reached.

To sum up the over-all trend: the bulk shows the first signs of insolubility at less than one-tenth of the dose needed to produce the first gel in the single crystals. It becomes two-thirds to three-quarters insoluble at a dose where the solubility of the single crystals is still unaffected, but at sufficiently high doses both specimens converge to approximately the same limiting solubility.

When a single-crystal preparation was fused prior to irradiation, the 20-megarad dose resulted in 71.4 per cent gel; consequently, the sample behaved as the bulk material did. This means that the change in solubility is due to the difference in the morphology alone and not to some change introduced in the course of specimen preparation. No obvious change in the X-ray pattern could be observed on irradiation within the dose range studied in either type of specimen. Further, the irradiated and still soluble single crystal (irradiated to our standard 20-megarad dose), dissolved thereafter, showed no significant difference in its subsequent crystallization behavior from that of similarly treated unirradiated material.

2.4 *Viscosity Measurement*

The intrinsic viscosities of two single-crystal preparations, one before and one after irradiation with 20 megarads, were compared in dilute xylene solution at 100°C. The results were 1.42 deciliters per gram for

* R. Salovey.

the unirradiated and 1.38 for the irradiated specimens, the difference considered to be within experimental error.

2.5 *Melting Behavior*

It was found that the melting behavior was largely dominated by rate effects unless special precautions were taken to heat the samples slowly. We were not aware of this in the first experiments, which consisted of placing small fragments of irradiated and unirradiated polyethylene side by side on a coverglass lying on a temperature gradient bar. The coverglass was moved gradually along the bar towards higher temperatures. Since the unirradiated single-crystal preparation goes through a series of pronounced premelting and recrystallization effects, specimens were first fused once so that the initial morphology was identical in all of them. The irradiated bulk material retained its shape even after fusion, so that the melting point could not be readily deduced from the consistency alone. The irradiated single-crystal specimen, however, was found to melt 4 to 5°C lower (at 130 to 132°C) than the corresponding unirradiated samples (135 to 136°C) in this test. It is worth noting that the consistency above the melting point proved to be uniquely related to the solubility. Irradiated specimens which did not flow when heated to 140 to 150°C invariably proved to be largely insoluble. Thus, the solubility results could always be predicted from the appearance of the sample when fused in the course of our standard treatment after irradiation.

More accurate melting-point measurements were determined by a precision dilatometric technique set up by C. J. Aloisio and S. Matsuoka of Bell Telephone Laboratories. Heating rates were very slow, the entire procedure taking about seven days. Just below the melting point, the temperature was raised in very small increments, allowing the materials to recrystallize. The unirradiated single crystals showed pronounced recrystallization effects, i.e., remelting and recrystallization to higher densities, in contrast to the smaller effects evidenced by the irradiated single crystals (which had already been fused following irradiation). The melting point itself was the same for all four irradiated and unirradiated specimens: 135.5 to 136°C. Accordingly, the previously observed melting-point depression in the irradiated single crystals was due to the relatively rapid rate of heating. We did not observe the melting-point depression of 4 to 5°C reported in the literature for bulk specimens irradiated with comparable doses.⁸ After fusion, the dilatometer was cooled to 125°C, and the samples recrystallized at this temperature.

Both irradiated samples attained the previously observed specific volumes, in contrast to the unirradiated samples, which reached higher densities when cooled at 125°C from the melt.

2.6 *Electron Spin Resonance*

Electron spin resonance measurements were made with a Varian EPR spectrometer in order to assess the free-radical concentration resulting from irradiation. We compared a single-crystal preparation and a bulk specimen irradiated with 20 megarads. Irradiation was conducted at low temperatures by placing the specimen tubes on a block of solid carbon dioxide wet with acetone. In this case the samples were not fused immediately afterwards, as was done in the other experiments, but the tubes were dropped into liquid nitrogen. The specimens were subsequently warmed for successively increasing lengths of time, and spin resonance measurements were made (with the sample in an unirradiated part of the tube).

Measurements at liquid nitrogen temperatures revealed broad absorption peaks in both specimens. These were identical in position, shape, and intensity for single-crystal and bulk specimens. On exposure to room temperature the strength of the absorption decayed, while a multiple-line pattern emerged from the band. The decay was much faster for the single crystals, where the peak decreased by a factor of eight after two hours and by a further 40 per cent after a total time of 16 hours. The corresponding decrease in the bulk was only by a factor of two, taking place during the first two hours, and resulting in a seven-fold difference between the two specimens. On heating to 74°C the absorption in the bulk decreased by 60 per cent after half an hour, and a further decrease to one-third of this value occurred after 20 minutes at 120°C. The single-crystal spectrum remained essentially unchanged by this treatment. Thus, the two peaks became roughly comparable in the two specimens. The peaks remained unaltered in both specimens on fusion at 152°C for half an hour, but a further decrease by a factor of two occurred in the single crystals when they were kept fused overnight. The spectra persisted even when the tubes were opened.

The initial decay could also be seen visually. Both specimens were brown after irradiation at low temperatures, but the color faded rapidly on warming to room temperature.

We do not attempt any specific assignment of the band, particularly as in some detail our findings differ from those reported in literature.^{7,9} For the present purpose it suffices to conclude that the free radicals

responsible for the absorption were present in comparable concentration in both kinds of irradiated specimens, but showed different decay patterns in the two.

2.7 *Infrared Spectra*

We examined infrared absorption of single crystals and bulk specimens and corresponding samples irradiated with 20 megarads. There were only slight differences between the original bulk and single-crystal specimens. On irradiation the vinyl peaks (terminal unsaturation) at 910 and 990 cm^{-1} decreased in both specimens but to a lesser extent in the single crystals, where it decreased by a factor of two, in contrast to the bulk, where the decrease was by a factor of seven (as referred to the weak methyl band at 1370 cm^{-1}). There was a pronounced increase of the transvinylene band (internal unsaturation) at 965 cm^{-1} , which was practically identical in the two samples (in the single crystals the increase was about 20 per cent larger).

2.8 *Hydrogen Evolution*

A bulk and a single-crystal specimen irradiated with 20 megarads were analyzed for free hydrogen with a mass spectrometer. Single crystals gave off 4.19×10^{-4} mole hydrogen per mole of CH_2 and the bulk gave off 8.17×10^{-4} . After this analysis the specimens were fused and analyzed again for free hydrogen which might have been originally occluded. Only a slight increase was found, the final figures being 4.40×10^{-4} mole per mole of CH_2 in the single crystals and 8.28×10^{-4} mole per mole of CH_2 in the bulk. The single-crystal preparation gave off noticeable amounts of larger fragments roughly comparable in volume to the hydrogen evolved. Thus, in this measurement the single crystals gave off about one-half the amount of hydrogen as did the bulk.

III. DISCUSSION

It was initially apparent from the solubility measurements that the samples crystallized from the melt and from dilute solution responded differently to irradiation. Here we shall discuss possible causes of this striking difference. The decrease in solubility on irradiation is usually associated with the formation of crosslinks leading to an infinite network. In fact, solubility is often used as a measure of crosslinking on the basis of Flory's theory of network formation.^{1,2,11} Accordingly, the first signs of insolubility are observed when an infinite network first forms. The

dose where this occurs is referred to as the *gel point* because insoluble material appears in the gel consistency. With increasing doses the relative amount of gel increases and rapidly attains a high value. For a given crosslink concentration, the gel point is also a function of the molecular weight; shorter molecules lead to enhanced solubility. Moreover, the onset of gelation is somewhat retarded by the incidence of scission, which accompanies crosslinking in the irradiation of polyethylene.^{1,2}

Our observations first raise the question whether or not the radiation has an intrinsically much smaller effect on the single crystals than on the bulk. For this, as for most of the later arguments, we concentrate on the specimens which received 20 megarads, since in this case the differences were the most clear cut; the single crystals remained soluble, while about 70 per cent of the bulk became insoluble. Certain references in the literature state that irradiation should have a much smaller effect on the crystalline than on the amorphous portions.^{6,7} Accordingly, single crystals should be less affected. However, L. Mandelkern* has found that the effects of irradiation on the solubility of crystalline and amorphous polymer depend on the radiation dose. In any case, the crystallinity content of the bulk specimen is not sufficiently different from that of the single crystals (it is about 90 per cent crystalline) to account for our effect, quite apart from the ambiguities encountered in definitions of the amount of crystallinity. However, our spin resonance results indicate that the primary effect of the radiation is comparable in bulk and single-crystal specimens, the concentration of the alkyl radicals being about the same (also indicated by the apparent discoloration). It is true that, from the one comparison pursued, the hydrogen evolution in the single-crystal specimens is only half of that found in the bulk for a dose of 20 megarads. If this means that the radiation is only half as effective in introducing chemical changes in the single crystals, the single crystals should behave like the bulk irradiated to 10 megarads. But a dose of 10 megarads still renders the bulk material about 60 per cent insoluble. Consequently, a factor of two in the hydrogen evolution could not account for our observations. Thus, the possibility that the effect of the radiation on the single crystals is sufficiently less to account for the observed increase in solubility can be ruled out.

Another alternative is that the scission-to-crosslinking ratio is higher in the single crystal, while the total effect of the radiation (scission plus crosslinking) remains comparable.

Infrared results indicate that the terminal vinyl concentration is

* Private communication.

higher in the irradiated single crystal than in the corresponding bulk. If scission occurs by disproportionation leading to vinyl groups, this, at first sight, may seem to support the suggestion of increased scission. However, the vinyl concentration is invariably found to go down on irradiation in spite of the assumed scission by disproportionation.⁵ So does it also in both of our specimens, but less in the single crystals. Since the basic reason for the decrease in the vinyl concentration is obscure, differences in the relative amount of decrease cannot be safely invoked as evidence for differences in scission in the present argument. End-linking has been suggested to explain the terminal vinyl decay.¹² This would increase the size of the molecule and facilitate gelation, which would be in line with the more rapid lowering of the solubility in the bulk. However, whether this speculative factor could account for the large differences in the two specimens remains questionable.

The possibility of scission leading to methyl end groups through hydrogen uptake is not usually considered. We could not observe any increase in methyl concentration from the infrared spectra, but owing to the relative weakness of the band this is not conclusive. As far as one can surmise, such a mechanism could consume part of the hydrogen evolved and explain why less hydrogen forms on irradiation of single crystals. However, this hydrogen deficiency could result from a recombination of the radical and hydrogen, which might occur more readily in the better-ordered single-crystal system. Thus, while the infrared evidence raises some interesting possibilities, it remains inconclusive as regards scission. Neither is there evidence for an increase in scission from other properties, such as viscosity or melting point. In fact, there is no evidence for scission at all other than the existence of a limiting solubility at high doses. Since this limiting solubility is comparable for both samples, it is unlikely that the amount of scission should be much different at the smaller doses.

The theory of network formation on which all arguments in this field are based assumes that crosslinks are randomly distributed between different molecules, intramolecular crosslinks being ignored. If in two specimens the total number of crosslinks were the same, but in one of them part of the crosslinks were *intramolecular*, the latter one would evidence enhanced solubility. In other words, part of the crosslinks would not contribute to the formation of an infinite network, hence they would not contribute to the decrease of solubility. The existence of some intramolecular crosslinks ("ringlinks") has been suggested¹³ to account for the lack of material balance on irradiating bulk specimens. Since single crystals consist of folded chains, intramolecular crosslinking would be

expected to be appreciable. In fact, it was with this expectation that the present investigation was undertaken.

Besides the enhanced probability of intramolecular crosslinking in a folded structure, the probability of multiple intermolecular linking between a finite number of molecules in close proximity is also enhanced, in view of the extensive contacts between ribbons of folded molecules. Such crosslinks will not contribute to the formation of an infinite network, hence will be ineffective in lowering the solubility. The presence of these crosslinks should, however, show up in the viscosity. The lack of a significant change in our viscosity figures might be due to a counterbalancing of two effects: a reduction due to internal crosslinks making the molecule more compact and the increase of the molecular weight.

Attractive and probable as the explanation based on intramolecular crosslinking may sound, no final confirmation could be brought to bear on the issue within the range of the present work. The salient point is that there is no direct chemical method for the assessment of crosslinks. Usually the solubility criterion is used, and this is clearly restricted to intermolecular crosslinks. In fact, our experiments demonstrate the limitations of this criterion. The situation is identical with changes in the elastic modulus, which are also used as a measure of crosslinking,^{1,2} since this is also based on the formation of infinite networks through intermolecular crosslinks. The simultaneous appearance of insolubility in the irradiated material and that of rubbery consistency on fusion are due to this reason. This leaves us with the possibility of studying internal configuration in greater detail, which would require extensive and careful work on solution properties beyond the scope of the present investigation.

The fact that the limiting gel content is comparable for both samples in spite of the significant initial difference needs some comments. It is customary to characterize the scission/crosslinking ratio (p/q) by a plot of $s + \sqrt{s}$ versus $1/r$, where s is the soluble fraction and r is the dose: then the intercept along the ordinate gives p/q .⁴ Since this intercept will be determined by solubility figures at high doses, different curves tending to the same limiting value will give the same p/q . In such cases the same p/q is assigned to the corresponding samples, and the differences at lower doses are attributed to differences in the type of initial molecular weight distributions.⁴ (The molecular weight distributions will tend to become identical — random — in the course of scission.) In our case, however, the molecular weight distribution is identical in the two samples. The fact that after fusion the single-crystal preparation showed the same radiation characteristics as the bulk proves that no molecular weight or distribution change sufficient to modify the radiation behavior could

have occurred during the preparation of the single crystals. This means that the practically identical p/q values obtained from the limiting solubility cannot characterize both samples; p/q must vary with the dose at least in one, if not both, specimens.

The Charlesby-Pinner⁴ treatment is based on the assumption that both crosslinking and scission are proportional to the dose received. We now suggest that this assumption could still be valid for the chemical changes induced, but that the effectiveness of these in altering the solubility might vary. We have seen that the initial solubility of the single-crystal specimen could be accounted for by assuming that part of the crosslinking is ineffective because of intramolecular (or multiple intermolecular) crosslinks leading to a large effective p/q . With large doses, the effective p/q decreases to that corresponding to the bulk. Plausibly, this could be due to a decrease in the effectiveness of scission while the true p/q may remain unaltered. If scission occurred within a loop formed by an intramolecular crosslink (or multiple intermolecular link) the number of molecules would not increase; hence the cutting would be ineffective as regards lowering the solubility. The chance for scission to become ineffective in this way increases with crosslinking, hence with the dose. In the light of this picture, it appears even less probable that scission should be unusually high at low doses in the single-crystal preparation, which strengthens our contention that intramolecular crosslinking rather than unusually high scission should be responsible for the high solubility in the 20-megarad region.

If the explanation based on intramolecular crosslinking proves to be true it would imply that a given molecular segment within the crystal in the bulk is in an environment of segments predominantly belonging to different molecules, in contrast to single crystals, where it would be partly in the vicinity of other parts of the same molecule. This would suggest that chain folding would be less frequent in the bulk, if present at all. Thus, it appears that this newly found radiation effect might contribute substantially to the present controversy as to whether or not folded molecules found in single crystals are also representative of the bulk. However, it will be shown in the following paper¹⁴ that such conclusions would be premature. It will be demonstrated that the effect of the morphology intervenes in a more subtle way and that one cannot exclude folded structures on the present evidence.

IV. ACKNOWLEDGMENT

The diversity of experiments reported in this work resulted from the encouragement and assistance of various members of Bell Telephone

Laboratories to whom we are indebted. Particularly, we thank C. J. Aloisio and S. Matsuoaka for dilatometry, S. H. Glarum for electron spin resonance measurements, J. P. Luongo for infrared studies, E. E. Francois for mass spectrometry, and W. P. Slichter for his many kindnesses. In addition, we are indebted to Professor F. C. Frank of the University of Bristol for helpful discussions.

REFERENCES

1. Charlesby, A., *Atomic Radiation and Polymers*, Pergamon Press, New York and London, 1960, ch. 13.
2. Bovey, F. A., *The Effects of Ionizing Radiation on Natural and Synthetic High Polymers*, Interscience Publishers, New York, 1958, p. 97.
3. Keller, A., *Phil. Mag.*, **2**, 1957, p. 1171; *Makromol. Chem.*, **34**, 1959, p. 1.
4. Charlesby, A., and Pinner, S. H., *Proc. Roy. Soc.*, **A249**, 1959, p. 367.
5. Lawton, E. J., Balwit, J. S., and Powell, R. S., *J. Polymer Sci.*, **32**, 1958, p. 257.
6. Lawton, E. J., Powell, R. S., and Balwit, J. S., *J. Polymer Sci.*, **32**, 1958, p. 277.
7. Loy, B. R., *J. Polymer Sci.*, **44**, 1960, p. 341.
8. Mandelkern, L., Roberts, D. E., Halpin, J. C., and Price, F. P., *J. Am. Chem. Soc.*, **82**, 1960, p. 46.
9. Lawton, E. J., Balwit, J. S., and Powell, R. S., *J. Chem. Phys.*, **33**, 1960, p. 395.
10. Williams, T. F., and Dole, M., *J. Am. Chem. Soc.*, **81**, 1959, p. 2919.
11. Flory, P. J., *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, N. Y., 1953, p. 353.
12. Dole, M., Milner, D. C., and Williams, T. F., *J. Am. Chem. Soc.*, **80**, 1958, p. 1580.
13. Dole, M., Milner, D. C., and Williams, T. F., *J. Am. Chem. Soc.*, **79**, 1957, p. 4809.
14. Salovey, R., and Keller, A., this issue, p. 1409.

The Influence of Crystallization Conditions on Radiation Effects in Polyethylene —

II. Crystallization from Concentrated Solutions

By R. SALOVEY and A. KELLER

(Manuscript received January 24, 1961)

Samples of Marlex 50 were crystallized from solution in the concentration range of 0.1 to 2 per cent and were irradiated with a dose of 20-megarad electrons. If crystallized at 70°C the materials were largely soluble, but if crystallization was conducted at 85°C the samples became two-thirds insoluble at from 0.4 to 0.6 per cent concentration upwards. Morphological examination of samples crystallized from 1 per cent solution revealed that the mode of crystallization was lamellar in all cases, but in the case of 70°C crystallization the lamellae were largely separated while in the 85°C case they were grown together in a compacted mass. After the latter was broken up by ultrasonic treatment this sample stayed fully soluble after 20-megarad electron irradiation. This finding indicates that the crosslinks are formed predominantly at the lamellar interfaces tying the lamellae together when these are in close contact. This might imply that the radiation is most effective at the sites where the chains fold. The present exploratory work shows the importance of the mode of packing of the lamellae, and indicates that this might be an important factor in distinguishing between the different morphologies observed in general.

I. INTRODUCTION

It was shown in the preceding paper¹ that polyethylenes crystallized from the melt and from dilute solution were affected differently by electron irradiation, the former becoming insoluble at much lower doses than the latter. In addition to their inherent interest for radiation studies, these findings should also be of value for morphological investigations, as the morphological state of the specimen may be reflected in the effect of the radiation on solubility. As was concluded in the preceding paper,

the different behavior of the bulk and dilute solution crystallized specimens indicates different morphologies. In particular, the higher solubility of the irradiated solution crystallized specimens is readily explicable in terms of intramolecular crosslinks forming within the folded chains of such single-crystal preparations. Accordingly, the lower proportion or even absence of such intramolecular linkages in the bulk would indicate that the molecules are not folded in such specimens. This appears to support the view that we are dealing with two distinct states of matter in the two kinds of specimens, a point which is nowadays widely debated.^{2,3}

According to recent theoretical work by Lauritzen and Hoffman,⁴ crystallization from dilute solution should occur via chain folding and that from the melt through the traditional bundle-like aggregation, a transition occurring between the two as the concentration of the solution is increased. Although these authors have modified their view since and consider the possibility of a folded crystallization in all cases,* we accepted their first suggestion as a working hypothesis at the time, particularly in view of the large differences found in the two specimens in our radiation work. In order to explore the possible transition between the two kinds of behavior we examined polymers crystallized from more concentrated solutions. This work again is of a purely exploratory nature subject to the same limitations mentioned in the introduction of the preceding paper.

II. EFFECT OF CONCENTRATION AND TEMPERATURE OF CRYSTALLIZATION ON THE SOLUBILITY

2.1 *Methods*

Our material was again the linear polyethylene Marlex 50. Specimens were crystallized at either 70° or 85°C from xylene solutions of varying concentrations. Crystallization at 70°C was very rapid for the concentrated solutions, turbidity setting in nearly immediately on lowering the temperature of the solution to 70°C. All samples gave discrete low-angle reflections corresponding to spacings of 110 and 130 angstroms in the samples crystallized at 70 and 85°C respectively. Irradiation experiments were carried out under the same circumstances as in the preceding paper.¹ Only one dose size (that of 20 megarads) was used, since it was with this dose that the most pronounced differences between melt and dilute solution crystallized specimens were found. Solubility changes were

* Private communication.

TABLE I—SOLUBILITY OF POLYETHYLENE (CRYSTALLIZED AT VARIOUS CONCENTRATIONS AND TEMPERATURES) AFTER IRRADIATION TO 20 MEGARADS

| Concentration | Gel fraction | |
|---------------|----------------------|----------------------|
| | Crystallized at 70°C | Crystallized at 85°C |
| 0.1% | — | 1.2% |
| 0.3 | 7.6% | 1.4 |
| 0.4 | — | 4.7 |
| 0.6 | — | 53.6 |
| 0.8 | — | 57.5 |
| 1.0 | 2.4 | 64.4 |
| 2.0* | 7.7 | — |

* There was an undissolved residue left. The actual concentration must have been lower than 2 per cent.

determined by using the same Soxhlet extraction method described previously.¹ All irradiations reported here were carried out under strictly identical conditions, samples under comparison having been irradiated simultaneously.

2.2 Solubility Results

The results are listed in Table I. As can be seen, the solubility remained practically unaffected in the whole accessible concentration range for specimens crystallized at 70°C, giving a very small gel content even for the most concentrated solution (the relatively high figure at 0.3 per cent is undoubtedly anomalous, but this does not affect the general conclusion). Samples crystallized at 85°C, however, showed a remarkable transition. For this reason the 85° series was closely examined. Here, the gel content was insignificant up to a concentration of 0.4 per cent, the material remaining essentially fully soluble. Beyond this stage however, the gel content rose suddenly to a value comparable with bulk crystallized samples irradiated to the same dose.

According to these results, all the specimens crystallized at 70°C behaved in a similar manner independent of concentration. Since the single-crystal nature of specimens formed at low concentrations is established, there is no reason to expect any departure from this at the higher concentrations on the basis of the radiation results (although the persistence of single crystals to such high concentrations may appear surprising). The transition from full solubility to two-thirds insolubility with increasing concentration in the case of 85°C crystallization may appear to be consistent with the expected transition from single crystal

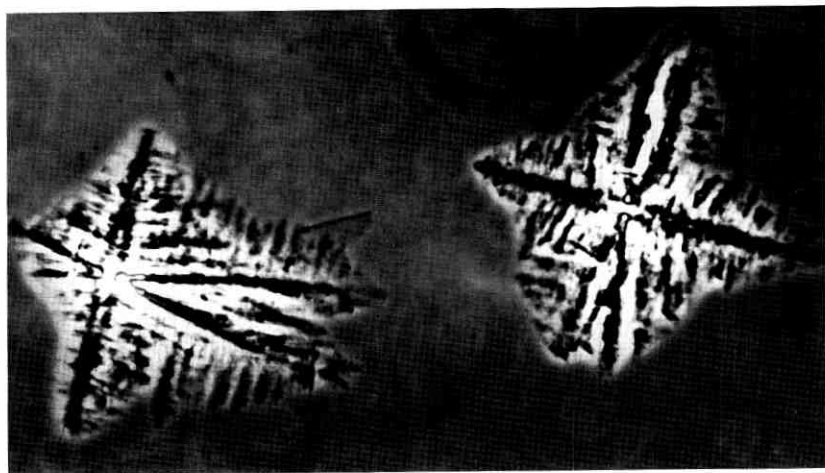


Fig. 1 — Typical polyethylene dendrites formed at 70°C from a 1 per cent xylene solution, photographed in a mixture of xylene and acetone. (Phase-contrast microscopy; magnification of original photograph 600X.)

to fringed micelle; the dependence on the temperature of crystallization, however, would remain unexplained.*

2.3 Microscopic Examinations

Specimens crystallized from 1 per cent solution at 70° and 85°C were examined by phase-contrast microscopy. Figs. 1 and 2 were taken of a suspension formed at 70°C. Fig. 1 clearly represents dendritic single crystals; Fig. 2 needs some further comments. In another investigation by one of us in collaboration with Mitsuhashi⁶ it was established that the consecutive terraces of a polyethylene crystal need not be grown together but can be spatially separated, the whole crystal resembling more a loose helicoid (tighter at the center than at the edges) than a

* One could invoke some of the implications of the kinetic approach to crystallization.⁴ Accordingly, in concentrated solutions nucleation via inter- and intramolecular processes might compete. At high temperatures bundle-like crystallization would be favored in accordance with the considerations of Lauritzen and Hoffman. At low temperatures (high supercoolings) the nucleation rate might be so high that several intramolecular nuclei might form within the same molecule (a possibility predicted by Frank and Tosi⁶) even before a bundle from different molecules could develop by the aggregation of different chains. This would result in further crystal growth occurring by an aggregation of tiny chain folded crystals. Frank and Tosi estimate a crystallization temperature of 70°C in xylene where such a process could set in. Thus, as it stands this could be in line with our observations. Further observations which followed, however, indicate that other factors are at play which make this proposed explanation questionable.



Fig. 2 — Multilayer dendrites in the same preparation as in Fig. 1. (Phase-contrast microscopy; magnification of original photograph 600X.)

solid spiral pyramid. This could be directly confirmed by observing the crystals turning in the liquid, when the individual layers could be seen to be separated and diverging outwards from the center. The more complicated structure in Fig. 2 is essentially such a multilayer crystal in an unusual projection. When such crystals are dried down, this three-dimensional structure collapses, leading to obscured areas in the interior of the crystal. It follows that such confused central areas usually observed in dried-down crystals need not be indicative of a departure from the chain folded lamellar structure. The crystals in Figs. 1 and 2 were in suspension when photographed to avoid this obscuring effect. We conclude that the specimen crystallized at 70°C has the usual lamellar single crystal habit throughout.

Fig. 3 shows a typical example of aggregates formed at 85°C. As seen, they are compact and nearly circular (although reminiscent of spherulites, they do not show a clear Maltese cross extinction pattern between crossed polaroids). In view of the marked difference in solubility after



Fig. 3 — Typical compact crystal formations of polyethylene formed at 85°C from 1 per cent xylene solution; the crystals are in their mother liquor. (Phase-contrast microscopy; magnification of original photograph 1000 \times .)

irradiation, we anticipated a difference in morphology between preparations formed at 70°C and at 85°C. However, the specimen crystallized at 85°C also proved to be lamellar, the distinct fibrous structure corresponding to lamellae seen edge-on. The observation of loose lamellae protruding beyond the edges of the more compact mass also suggests that the aggregates in Fig. 3 consist of closely packed lamellae. Thus the morphology may not be basically dissimilar in the two specimens. Both would consist essentially of chain-folded lamellae; in the 70°C crystallization the lamellae within a crystal are separated by the liquid in the sense described above, while in the 85°C crystallization they are compacted in a solid mass. It would follow, therefore, that the difference in the radiation behavior may be connected with *the mode of packing of the lamellae*. We attempted to separate these lamellae in samples crystallized at 85°C to see whether the high solubility of irradiated specimens formed at 70°C could be obtained.

2.4 *Disintegration Experiments*

The disintegration of aggregates formed at 85°C (Fig. 3) was effected by means of ultrasonic radiation. Suspensions in xylene were exposed to radiation from an ultrasonic generator operating at 40 kc and 200 watts. The first trials showed that disintegration of the aggregates occurred and samples were then subjected to treatments of increasing duration.*

2.5 *Microscopic Examination*

Increasing extents of disintegration are indicated in Fig. 4. It will be seen that the compact masses are broken up into smaller units consisting of fewer fibrous entities (lamellae seen edge-on). The full morphological implications will not be discussed here. However, we suggest that the layers in Fig. 4 are probably multilayers themselves (not the basic 100-angstrom ones), and point out the frequent occurrence of pyramidal fragments consistent with other observations in this field.^{7,8,9,10} Besides the prominent structure in the photographs, a finer background structure is also visible which increases with the duration of the ultrasonic treatment. This is partly a lamellar structure of the basic 100-angstrom thickness and partly a ribbon-like fibrillar structure which may have its origin in broken-up lamellar units.

2.6 *Solubility of Irradiated Disintegrated Specimens*

Suspensions of polyethylene prepared by crystallizing a 1 per cent solution at 85°C and treated ultrasonically for 4 and 8 hours were filtered and prepared for irradiation with a 20-megarad dose in the usual manner. The irradiation of four samples was carried out simultaneously: the two just mentioned, a third sample which was a portion of the above preparation that received the eight-hour ultrasonic treatment but was fused subsequently, and a fourth sample that was crystallized under the same conditions but was not treated ultrasonically. When the samples were heated to 150°C after irradiation, while still *in vacuo* as part of the usual procedure,¹ marked differences could be observed in the melting behavior. The untreated and the ultrasonically treated but subsequently fused sample both retained their shape and showed no outward signs of melting, while the two remaining ultrasonically treated samples flowed

* The geometry of the set-up available for these tests permitted only a very inefficient utilization of energy. Consequently the times to be quoted are unusually long in comparison with results since obtained with other generators by one of us (A. Keller). Nevertheless, they give a comparative measure of the treatments received by the different samples.

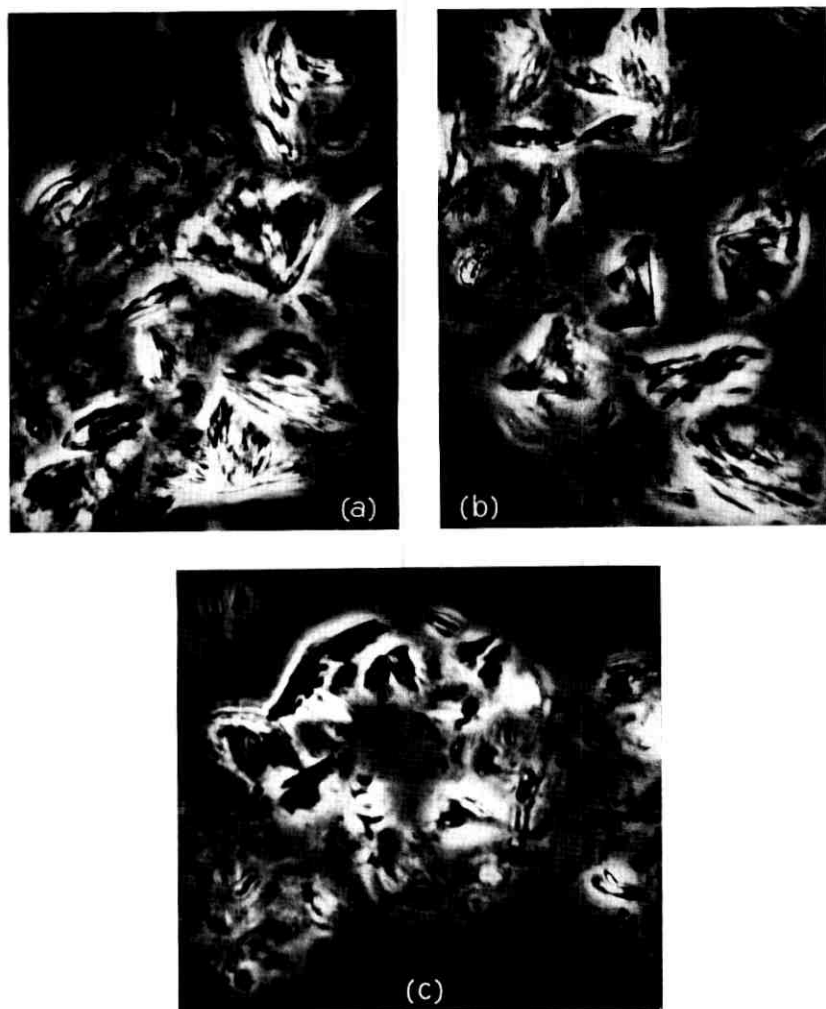


Fig. 4 — The same preparation as in Fig. 3, treated ultrasonically for (a) one and one-half hours, (b) four hours, and (c) eight hours. (Phase-contrast microscopy; magnification of original photograph 1060X).

as unirradiated polyethylene does. As described in the preceding paper, the absence of clear melting was indicative of intermolecular crosslinks and was invariably associated with appreciable insolubility. In fact, both ultrasonically treated and unfused specimens were fully soluble, the untreated sample giving a gel fraction of 60.9 per cent, and the sample

which received an ultrasonic treatment of eight hours and was subsequently fused giving 27.8 per cent gel.

III. DISCUSSION OF RESULTS

The complete solubility of the two ultrasonically treated specimens indicated that the behavior characteristic of single-crystal preparations had been restored. If the differences between specimens showing complete solubility, on the one hand, and large insolubility, on the other, after a 20-megarad dose were due to inherently different molecular textures, this situation could not be altered by ultrasonic treatment. Clearly, bundle-like structures could not be changed into folded ones in this way. However, morphological observations suggest that both kinds of samples consist essentially of lamellae which are packed differently. Accordingly, if the lamellae are closely packed in the sample crystallized at 85°C crosslinking is more likely to occur between them. Such interlamellar crosslinks would be necessarily intermolecular and would lead to a decrease in solubility in agreement with our experimental results. This raises a further problem. Only a small fraction of the molecule (2 to 4 per cent, the part in the folds) is in the lamellar interfaces. If crosslinking could occur with equal probability anywhere along the chain, there would only be an insignificant chance for it to occur at the interface linking two lamellae. On this basis, interlamellar links could not have the significant effect our results imply. It appears therefore that the folds may be "sensitive points" where crosslinking may occur with greater frequency than elsewhere.

This situation is complicated by the fact that the gel fraction of the sample which was treated ultrasonically for eight hours and subsequently fused did not regain the value of 60.9 per cent given under exactly the same conditions by the untreated control (but only yielded about half of this quantity). Probably an appreciable amount of chain-cutting was caused by the ultrasonic radiation. Then the increase of the solubility of the two unfused ultrasonically treated specimens cannot be interpreted in terms of physical changes alone. A significant part of it must be due to the degradation of the molecule. The amount of chain-cutting is expected to depend on the time of ultrasonic treatment; the sample treated for four hours was still fully soluble on irradiation and the minimum treatment required might well be still shorter. This means that our treatment and consequent degradation were unnecessarily drastic. Therefore, perhaps a more complete reversibility of the gel content on fusion would still be compatible with full solubility of the specimen prior to

fusion. Fuller clarification of this point could not be accommodated in the present program. The present results have established the fact that the mode of aggregation of the lamellae can have a profound effect on the irradiation behavior. In view of scission during the ultrasonic treatment, it cannot be stated with certainty that the *full* difference in irradiation behavior between the two specimens under question can be accounted for by the mode of aggregation of the lamellae.

We have tacitly implied that the contact between layers is of the van der Waals type only. It is likely that the better compacting of the units in the specimen crystallized at 85°C is due to a few molecules running through two or more layers before folding back, or that some folds might interpenetrate. This situation is more likely to arise with concentrated systems. If this were the case, these joining molecules would have to be cut for the lamellae to be separated, which would mean that some of the molecular degradation may be necessarily linked with the physical separation of the layers.

We have seen that the packing of the layers is strongly influenced by the temperature of crystallization, layers grown at higher temperatures being more closely stuck or linked together. At this stage, we can offer no explanation for this behavior.

IV. SOME GENERAL CONSIDERATIONS

In Section III, discussion was restricted to the interpretation of differences between two kinds of solution crystallized specimens, one crystallized at 70°C and the other at 85°C from 1 per cent xylene solution. The question arises as to whether the explanation cited could be extended to account for the differences between solution-crystallized specimens (which retained their full solubility after a 20-megarad dose) and bulk specimens (which became two-thirds to three-quarters insoluble after the same dose). In the light of present findings it can certainly be stated that this difference need not be associated with a basic difference in molecular texture, i.e., either with the presence of folded molecules or of bundles. In fact, in the case of specimens crystallized from 1 per cent solution at 85°C, no basic departure from single-crystal texture was found other than more compact packing of lamellae, perhaps arising from some attachment of these units. It may not be unreasonable, therefore, to consider the possibility of the same effect being responsible for bulk behavior. This would imply that the bulk might consist of chain folded lamellae. Indeed, we would expect closer packing and stronger connectedness of such lamellae in the bulk. Such a picture of the bulk

would not contradict the results of the present radiation studies. On the contrary, it would be supported by analogy with the behavior of specimens crystallized from concentrated solutions as well as by existing evidence of lamellar structures in the bulk.^{11,12} In view of the complexity of the subject, it is not intended to imply that the present evidence is in any way conclusive. (For other views see, for example, Chiang and Flory.³)

In conclusion, the present exploratory work has drawn attention to limitations in the usual treatments of the effects of irradiation which consider random intermolecular crosslinking and random scission only. Moreover, it contributes to current morphological studies by indicating the importance of packing of crystal lamellae. Finally, radiation effects are shown to be sensitive indicators of the morphological state of the polymer, and thus should take their place among other structural tools.

V. ACKNOWLEDGMENTS

We sincerely thank Prof. F. C. Frank of the University of Bristol and D. W. McCall, H. D. Keith, and F. J. Padden of Bell Telephone Laboratories for helpful discussions.

REFERENCES

1. Salovey, R., and Keller, A., *this issue*, p. 1397.
2. Keller, A., *Makromol. Chem.*, **34**, 1959, p. 1.
3. Chiang, R., and Flory, P. J., *Am. Chem. Soc. meeting*, New York, September 1930.
4. Lauritzen, J. I., and Hoffman, J. D., *J. Res. N.B.S.*, **64A**, 1960, p. 73.
5. Frank, F. C., and Tosi, M. P., to be published.
6. Mitsuhashi, S., and Keller, A., *Polymer*, **2**, 1961, p. 109.
7. Bassett, D. C., Frank, F. C., and Keller, A., *Nature*, **184**, 1959, p. 810.
8. Bassett, D. C., and Keller, A., *Phil. Mag.*, **6**, 1961, p. 345.
9. Reneker, D. H., and Geil, P. H., *J. Appl. Phys.*, **31**, 1960, p. 1916.
10. Niegisch, W. D., and Swan, P. R., *J. Appl. Phys.*, **31**, 1960, p. 1906.
11. Fischer, E. W., *Z. Naturforsch.*, **12a**, 1957, p. 753.
12. Sella, C., and Trillat, J. J., *C. R. Acad. Sci.*, **248**, 1959, p. 410.

Compression, Filtering, and Signal-to-Noise Ratio in a Pulse-Modulated System

By E. A. MARCATILI

(Manuscript received April 19, 1961)

Some aspects of the transmission of Gaussian pulses in a frequency-division multiplex system have been calculated. It is assumed that each channel transmitter includes an amplifier, which may be nonlinear, with input and output filters, and that each channel receiver includes a linear amplifier with a thermal noise source at its input plus input and output filters. The influence that (a) transmitting amplifier compression, (b) maximally flat channel-dropping filters of first or second order, and (c) distribution of filtering in the system have on the signal-to-noise ratio, time crosstalk, and frequency crosstalk is calculated.

I. INTRODUCTION

A proposed long distance waveguide system¹ will transmit a large number of wideband PCM channels via a single waveguide by means of a frequency multiplexing arrangement. In previous papers^{2,3} an idealized frequency multiplex system was analyzed. These papers considered a system in which each channel signal, consisting of on-off carrier pulses, arrives at a detector together with three unwanted signals: (a) time crosstalk, due to leading and trailing edges of neighboring pulses; (b) frequency crosstalk, which is the interference from neighboring channels; and (c) thermal noise* generated essentially in the first amplifier of the receiver. Since these unwanted signals cause errors in the reading of the pulses, the filtering characteristics are closely related to the probability of errors. It has been shown that for any given probability of error it is possible to design the filters in such a way that a very desirable result is achieved—namely, the simultaneous minimization of time spacing between successive pulses, of frequency separation between neighboring channels, and of signal-to-noise ratio.

* Throughout the paper, when we talk about "noise" we mean thermal noise.

These conclusions were derived by analyzing a linear system consisting of a transmitting filter and a receiving filter. Such a scheme does not give quantitative answers to three practical questions:

i. Because RF power is expensive it is desirable to drive the transmitting amplifiers as hard as possible, and consequently they must operate in the nonlinear portion of their characteristic. What is the influence of nonlinearity on the system's performance?

ii. Filters between transmitting and receiving amplifiers reduce the signal without changing the noise. Consequently, in order to increase signal-to-noise ratio it would be desirable to have no filter between those amplifiers. But then the frequency crosstalk would be prohibitively high. How much filtering before and after each amplifier makes a reasonable compromise?

iii. All of the immediately available channel-dropping filters that connect transmitters and receivers with the transmission media^{4,5,6} are approximately Butterworth filters of first or second order. Pulses belonging to one channel will lose part of their power in passing through neighboring channel-dropping filters. How is the frequency spacing of channels affected by this loss?

The object of this paper is to answer those questions. More specifically, we want to find out the influences that compression, distribution of filtering with respect to the amplifiers, and the use of the available dropping filters have on time crosstalk, frequency crosstalk, and signal-to-noise ratio. The results extend those already derived previously.^{2,3}

In Section II the system to be analyzed is described and the fundamental formulas are introduced. The influences of neighboring channel-dropping filters and nonlinearity of amplifiers are described in Sections III and IV. Signal-to-noise ratio is evaluated in Section V, and some design examples are given in Section VI. Section VII contains a résumé of results and conclusions.

II. SYSTEM DESCRIPTION AND ANALYSIS

The system to be analyzed, Fig. 1, is ideally phase-equalized. The effect of imperfect equalization can be derived⁷ as a perturbation of the results obtained in this paper.

Table I shows that the transfer characteristic of each transmitter, excluding the amplifier is Gaussian and that that of each receiver is given by the sum of two slightly displaced Gaussian functions, which approximates the characteristic of a maximally flat Butterworth filter of third order and is easy to handle mathematically. Quantitative reasons for the selection of these filters can be found in a previous paper.³

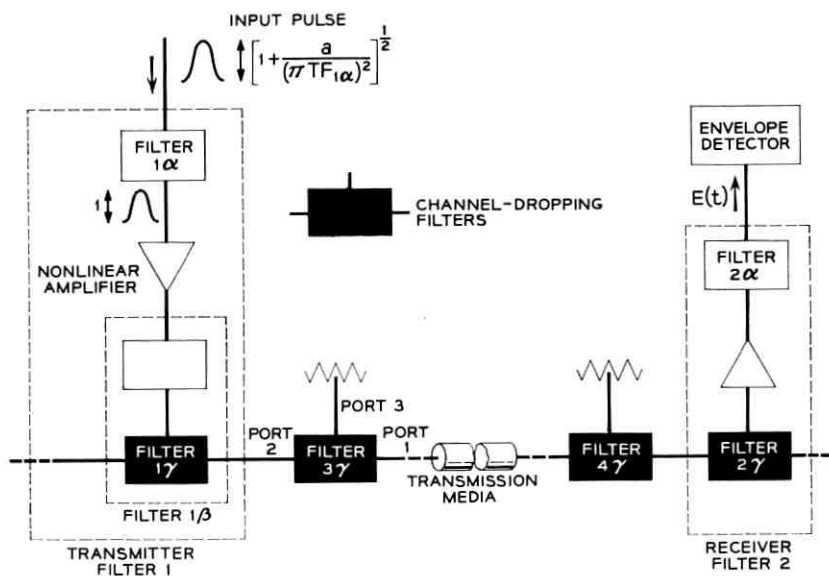


Fig. 1 — Block diagram of system (see definitions of filters in Table I).

Qualitative reasons are that these filters make a good compromise for low time crosstalk and frequency crosstalk. In effect, in order to minimize the tails of the output pulses and consequently the time crosstalk, it would be best to have both filters Gaussian. On the other hand, in order to minimize frequency crosstalk, it would be best to have both filters maximally flat (Butterworth) of high order. A reasonable compromise is achieved by making one of them Gaussian and the other maximally flat of high order. A transfer characteristic relatively easy to achieve without too much midband attenuation is that of a maximally flat filter of third order, which we adopt.

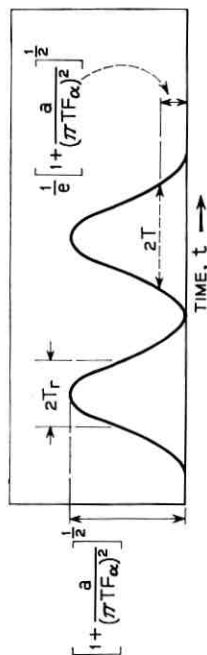
If the transmitter and receiver are centered at the same frequency, the transmission through them measures the insertion loss through a channel. If they are centered at different frequencies, the transmission through them measures the frequency crosstalk.

Each transmitter includes three filters and a nonlinear amplifier. The filter 1α preceding the amplifier and the combination of the two following it, 1β , are Gaussian. Each receiver consists of a linear amplifier between two filters 2α and 2γ whose combined transfer characteristic is approximately maximally flat of third order. This scheme allows one to evaluate the effects of the amount of filtering preceding and following each amplifier.

TABLE I—DEFINITIONS OF SYMBOLS

| Filters in Fig. 1 | Transfer Characteristics | 3 db Bandwidth | Center Frequency | Notes |
|--|--------------------------------|----------------|------------------|---|
| 1 Transmitter, excluding nonlinear amplifier | Gaussian | $2F_1$ | f_1 | $a = 0.346$ |
| 1 α Filter before nonlinear amplifier | Gaussian | $2F_\alpha$ | f_1 | $F_1^{-2} = F_\alpha^{-2} + F_\beta^{-2}$ |
| 1 β Filter after nonlinear amplifier | Gaussian | $2F_\beta$ | f_1 | $F_1^{-2} = F_\alpha^{-2} + F_\beta^{-2}$ |
| 1 γ Channel-dropping filter at transmitting end | Maximally flat (q th order) | $2F_{1\gamma}$ | f_1 | |
| 2 Receiver | ~Maximally flat (3rd order) | $2F_2$ | f_2 | $b = 1.3$ |
| 2 γ Channel-dropping filter at receiving end | Maximally flat (r th order) | $2F_{2\gamma}$ | f_2 | $n = 1.61$ |
| 2 α Noise filter | Maximally flat (q th order) | $2F_{1\gamma}$ | f_3 | |
| 3 γ Neighboring channel-dropping filter at transmitting end | Maximally flat (q th order) | $2F_{1\gamma}$ | f_3 | |
| 4 γ Neighboring channel-dropping filter at receiving end | Maximally flat (r th order) | $2F_{2\gamma}$ | f_4 | |

Note — In Fig. 1, if $f_1 = f_2$, the output is transmission through channel; if $f_1 \neq f_2$, the output is crosstalk.



Envelope input (on-off Gaussian pulses).

The channel-dropping filters 1γ , 2γ , 3γ , and 4γ connecting the transmission medium and transmitters to receivers are maximally flat of first or second order. No channel-dropping filters have been indicated in Fig. 1 to the left of the transmitter nor to the right of the receiver, because such filters will not influence the transfer characteristic between transmitter and receiver.

A Gaussian pulse modulating a carrier is the input to the system. The amplitude has been selected in such a way that the amplitude of the pulse reaching the nonlinear transmitter's amplifier is unity. The envelope of the output pulse reaching the detector has been calculated under the following assumptions:

- i. Of all the neighboring channels, only those with center frequencies immediately adjacent to the carrier influence the output pulse.
- ii. The transfer characteristics of the channel-dropping filters of the neighboring channels can be approximated by the first three terms of their power series expansions around the carrier frequency.

The envelope $E(t)$ of the pulse reaching the detector has been calculated in (79) of Appendix A:

$$E(t) = \left| \sum_{N=0}^{\infty} \left[I_N - \frac{i}{2\pi\varphi_1} \frac{\partial I_N}{\partial t} + \frac{1}{(2\pi\varphi_2)^2} \frac{\partial^2 I_N}{\partial t^2} \right] \right|. \quad (1)$$

Before detailing the exact meaning of each letter we will give a rough functional interpretation of this expression: φ_1 and φ_2 have to do with the presence of neighboring channel-dropping filters; if these filters had no influence φ_1 and φ_2 would be infinity and the envelope $E(t)$ would be reduced to

$$E(t) = \left| \sum_{N=0}^{\infty} I_N \right|. \quad (2)$$

The infinite number of terms is due to the nonlinearity of the transmitting amplifier, since in general the output is given by a summation of powers of the input. Those powers are an odd number $2N + 1$, and, if the term with $N = 0$ is the only one different from zero, the amplifier is linear; if only the terms with $N = 0$ and $N = 1$ are different from zero, the amplifier has compression of third order; if only the terms with $N = 0$, $N = 1$, and $N = 2$ are different from zero, the amplifier has compression of third and fifth order; and so forth.

Now we define the symbols in (1):

$$I_N = A_N \frac{T_N}{T} \sqrt{\frac{B_N}{(1+S)(1+R)}} e^{(n\mu A/2)^2 B_N - 4\rho^2 b(1-b\mu^2 A^2 B_N) - B_N(t/T)^2} \cdot \cosh \left[2n\rho(1 - b\mu^2 A^2 B_N) + in\mu A B_N \frac{t}{T} \right], \quad (3)$$

$$A_N = 2^{-2N} \binom{2N+1}{N} a_{2N+1}, \quad (4)$$

$$T_N = T \sqrt{\frac{1 + \frac{a}{(\pi T F \alpha)^2}}{2N+1}}, \quad (5)$$

$$B_N = \frac{1}{1 + A^2(a + b\mu^2) - 2N \left(\frac{T_N}{T}\right)^2}, \quad (6)$$

$$\varphi_1 = \frac{f_1 - f_3}{\frac{qS}{1+S} \pm \frac{rR}{1+R}}, \quad (7)$$

$$\varphi_2 = \frac{f_1 - f_4}{\left[\frac{q(2q+1)S}{2(1+S)} + \frac{r(2r+1)R}{2(1+R)} \mp \frac{qrSR}{(1+S)(1+R)} - \frac{3}{2} \frac{(qS)^2}{(1+S)} - \frac{3}{2} \frac{(rR)^2}{(1+R)} \right]^{1/2}}, \quad (8)$$

$$R = \left(\frac{F_{2\gamma}}{f_1 - f_4} \right)^{2\gamma}, \quad (9)$$

$$S = \left(\frac{F_{1\gamma}}{f_1 - f_3} \right)^{2\gamma}; \quad (10)$$

$a = 0.346$ is a parameter normalizing the bandwidth of a Gaussian filter,

$b = 1.3$ and $n = 1.61$ are parameters that normalize the bandwidth of an approximately maximally flat filter of third order,

f_1 is the center frequency of the transmitter and the carrier frequency,

f_2 is the center frequency of the receiver,

f_3 and f_4 are the center frequencies of neighboring channels,

$2F_1$ is the transmitting bandwidth measured at half power assuming a linear amplifier,

$2F_2$ is the receiving bandwidth measured at half power assuming a linear amplifier,

$2F_{1\alpha}$ is the half-power bandwidth of the Gaussian filter preceding the transmitting amplifier,

$2F_{1\beta}$ is the half-power bandwidth of the Gaussian filter following the transmitting amplifier,

$2F_{1\gamma}$ is the half-power bandwidth of the transmitting Butterworth channel-dropping filter,

$2F_{2\gamma}$ is the half-power bandwidth of the receiving Butterworth channel-dropping filter,

q is the order of the Butterworth channel-dropping filter at the transmitting end,

r is the order of the Butterworth channel-dropping filter at the receiving end,

$2T$ is the width of the input Gaussian pulse measured 8.686 db down,

$$\mu = \frac{F_1}{F_2} \quad (11)$$

is the ratio of transmitting to receiving bandwidths,

$$\rho = \frac{|f_1 - f_2|}{2F_2} \quad (12)$$

is the ratio of channel spacing to receiving bandwidth,

and

$$A = \frac{1}{\pi T F_1} \quad (13)$$

a_{2N+1} are the odd coefficients in the power series

$$v_0 = \sum_0^{\infty} a_\nu v_i^\nu \quad (14)$$

that relates the input and output amplitudes v_i and v_0 of the nonlinear transmitting amplifier.

Where double signs are indicated, the upper one is to be used if $f_3 = 2f_1 - f_4$ (neighboring channels at opposite side of carrier f_1), and the lower one if $f_3 = f_4$ (neighboring channels at the same side of carrier f_1).

If the transmitting amplifier is linear ($a_\nu = 0$ for $\nu \neq 1$) and there is no interference from neighboring channels ($f_3 = f_4 = \infty$), expression (1) measures the transmission through a channel (if $f_1 = f_2$) or the frequency crosstalk (if $f_1 \neq f_2$), as calculated previously.³

III. THE INFLUENCE OF NEIGHBORING CHANNEL-DROPPING FILTERS ON SIGNAL-TO-NOISE RATIO AND CROSSTALK

Assume:

(a) Linear transmitting amplifier. Then

$$A_N = 0 \quad \text{for } N \neq 0, \quad (15)$$

and only the first term of the summation (1) remains.

(b) Transmitter and receiver centered at the same frequency,

$$f_1 = f_2. \quad (16)$$

Call $E_1(t)$ the output envelope (1) when conditions (15) and (16) are satisfied, and normalize this output to the value $E_0(0)$ that measures the peak of the output ($t = 0$) that would occur if the channels were infinitely spaced, ($f_3 = f_4 = \infty$). Then

$$\frac{E_1(t)}{E_0(0)} = \left\{ \frac{\left[\left[1 + \frac{1}{2\pi\varphi_2} \frac{\partial^2}{\partial t^2} \right] e^{-(t/T)^2 B_0} \cos \frac{\mu A B_0 n t}{T} \right]^2}{(1+S)(1+R)} + \frac{\left[\frac{1}{2\pi\varphi_1} \frac{\partial}{\partial t} e^{-(t/T)^2 B_0} \cos \frac{\mu A B_0 n t}{T} \right]^2}{(1+S)(1+R)} \right\}^{1/2}. \quad (17)$$

For $t = 0$,

$$\frac{E_1(0)}{E_0(0)} = \frac{1 - \frac{2B_0}{(2\pi\varphi_2 T)^2} [1 - \frac{1}{2} (\mu A n)^2 B_0]}{\sqrt{(1+S)(1+R)}}. \quad (18)$$

The insertion loss in db due to the proximity of neighboring channel dropping filters is measured by

$$20 \log \frac{E_0(0)}{E_1(0)}.$$

This function has been plotted for maximally flat channel dropping filters of first order ($q = r = 1$) in Fig. 2 and for maximally flat filters of second order ($q = r = 2$) in Fig. 3. The abscissas

$$\rho = \frac{|f_1 - f_3|}{2F_2} = \frac{|f_1 - f_4|}{2F_2}$$

are proportional to channel spacing. The rows correspond to different ratios $F_{2\gamma}/F_2$ between the bandwidth of the channel-dropping receiving filter and the bandwidth of the receiver. The different columns correspond to different ratios $F_{1\gamma}/F_1$ between the bandwidth of the channel dropping transmitting filter and bandwidth of the transmitter. The solid lines correspond to the ratio between the transmitting and the receiving bandwidths $F_1/F_2 = 1.3$, the dashed lines to the ratio $F_1/F_2 = 1$; the upper solid and dashed lines in each set of curves correspond to $\pi T F_1 = 0.5$ and the lower solid and dashed lines correspond to $\pi T F_1 = 1$. We have chosen not to include as another variable the selection of signs in the expression for φ_2 , (8), because its effect is small. The two possible

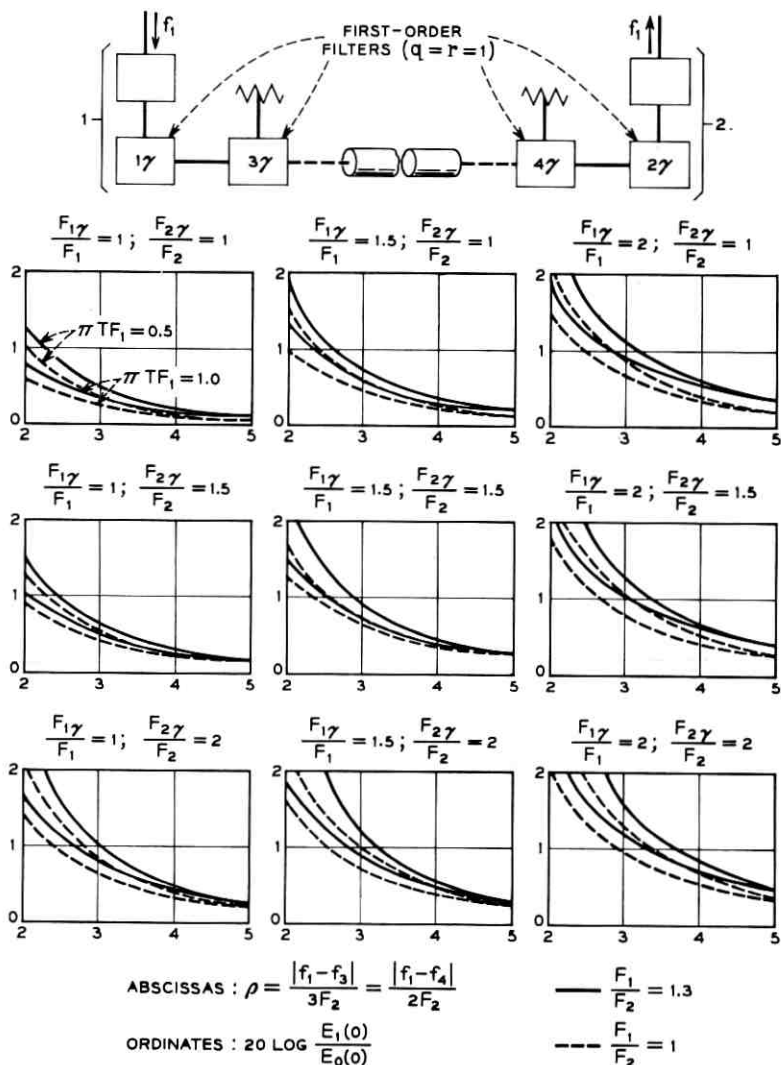


Fig. 2 — Insertion loss due to first-order neighboring channels.

choices of signs correspond to frequencies f_3 and f_4 of the neighboring channels on the same or different sides of f_1 . The first case introduces less insertion loss, but for the ranges of values selected in Figs. 2 and 3, and for small abscissas, the difference in ordinates for the two cases is smaller than 0.5 db. As a compromise, Figs. 2 and 3 have been calculated by averaging the results.

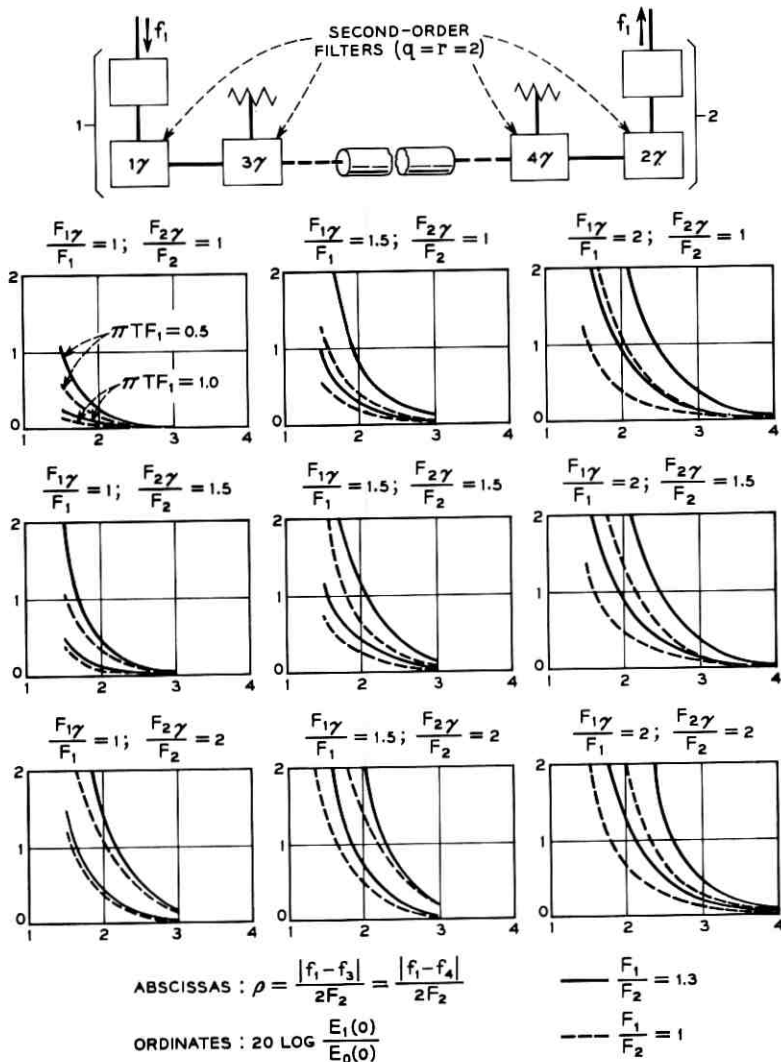


Fig. 3 — Loss due to second-order neighboring channels.

Let us concentrate on Fig. 2, ($q = r = 1$; that is, channel-dropping filters of first order). What parameters yield a system with low insertion loss due to neighboring channels? For a given abscissa (normalized channel spacing ρ), the smallest ordinate occurs in the lowest line in the upper left set of curves. This line is characterized by

$$F_{1\gamma} = F_{2\gamma} = F_1 = F_2; \quad \pi T F_1 = 1.$$

As could be expected, the system should have

- (a) long input pulses,
- (b) wide and equal bandwidths in transmitters and receivers,
- (c) narrow and equal channel-dropping filters at transmitting and receiving ends.

In order to compare Figs. 2 and 3 we compute an example:

For

$$\frac{F_1}{F} = 1, \quad \frac{F_{1\gamma}}{F_1} = \frac{F_{2\gamma}}{F_2} = 1.5, \quad \pi T F_1 = 1, \quad \text{and} \quad \rho = \frac{(f_1 - f_3)}{2F_2} = 2,$$

the insertion loss due to neighboring channels is 1.25 db if $q = r = 1$ in Fig. 2 (channel-dropping filter of first order) and 0.25 db if $q = r = 2$ in Fig. 3 (channel-dropping filter of second order). Second-order filters make the insertion loss due to neighboring channels 1 db better than that of first-order filters. But this does not necessarily mean that the total insertion loss will be 1 db better in a practical case. There are two reasons for this statement:

i. A second-order maximally flat filter, has more cavities and consequently more heat loss than does one of first order. We don't know how much the difference is, but assuming it is 0.5 db, the increase of heat loss in the two channel-dropping filters is 1 db, and consequently the advantage cited previously is only apparent. This point becomes more important for smaller tolerable insertion loss due to neighboring channels (large ρ).

ii. Consider the transmitter's total Gaussian transfer characteristic

$$y_1 = e^{-a[|f| - f_1]/F_1]^2} \quad (19)$$

and the transfer characteristic of its q th order channel-dropping filter

$$y_{1\gamma} = \frac{1}{\sqrt{1 + \left(\frac{|f| - f_1}{F_{1\gamma}}\right)^{2q}}}. \quad (20)$$

The transfer characteristic of the rest of the filter should be $y_1/y_{1\gamma}$. There are values of the frequency f for which this function is larger than unity, and therefore it is impossible to realize a passive filter with the desired transfer characteristic. The difficulty is solved by allowing the transfer characteristic of the transmitting filter to be not (19) but

$$\tilde{y}_1 = \frac{e^{-a[|f| - f_1]/F_1]^2}}{K}, \quad (21)$$

where K is a constant greater than unity and equal to the maximum

value that $y_1/y_{1\gamma}$ can achieve; K expressed in db is an added insertion loss.

The previous reasoning applies also to the receiver, but, since we are finally going to be interested in signal-to-noise ratios, the value of K applied to the receiver drops out.

In order to have an idea of the values that K may achieve in Gaussian transmitters we have calculated some examples:

| | $\frac{F_{1\gamma}}{F_1}$ | $20 \log K$ |
|--|--|-------------|
| First-order maximally flat channel-dropping filter ($q = r = 1$): | $\left. \begin{array}{l} 1 \\ 1.5 \\ 2 \end{array} \right\}$ | 0.3 db |
| Second-order maximally flat channel-dropping filter ($q = r = 2$): | $\left. \begin{array}{l} 1 \\ 1.5 \\ 2 \end{array} \right\}$ | < 0.1 db |

In general, the broader the channel dropping filter and the flatter the transfer characteristic of the transmitter or receiver, the smaller the added insertion loss $20 \log K$. This added insertion loss can be neglected except when $F_{1\gamma} = F_1$ and $q = r = 1$.

Now we turn to the increase of time and frequency crosstalk that is introduced by the presence of neighboring channel dropping filters. It is shown in Appendix B that, for systems in which the increase of insertion loss due to neighboring channel-dropping filters is not very large, (around 1 db), the increase of both time and frequency crosstalk is negligible.

IV. INFLUENCE OF THE NONLINEARITY OF THE TRANSMITTING AMPLIFIER ON SIGNAL-TO-NOISE RATIO AND CROSSTALK

Let us assume that, in Fig. 1, (a) the transmitting amplifier is driven so hard that the input signal is not linearly amplified, and (b) the presence of neighboring channel dropping filters can be neglected; this means that $f_3 = f_4 = \infty$ and, from (7) and (8),

$$\varphi_1 = \varphi_2 = \infty. \quad (22)$$

Call $E_2(t)$ the output envelope (1) when (22) is satisfied and normalize this output to the value $E_0(0)$ that measures the peak of the output ($t = 0$) that would occur if there were no compression ($A_N = 0$ for $N \neq$

0), and with the transmitter and receiver centered at the same frequency, $f_1 = f_2$. Then,

$$\frac{E_2(t)}{E_0(0)} = \left| \sum_{N=0}^{\infty} \frac{A_N}{A_0} \frac{T_N}{T_0} \sqrt{\frac{B_N}{B_0}} e^{(n\mu A/2)^2 (B_N - B_0) - 4\rho^2 b(1 - b\mu^2 A^2 B_N) - B_N(t/T)^2} \cdot \cosh \left[2n\rho(1 - b\mu^2 A^2 B_N) + in\mu A B_N \frac{t}{T} \right] \right|. \quad (23)$$

As explained in Appendix A, this expression has a simple physical interpretation. The Gaussian pulse of unit amplitude and width

$$2T \sqrt{1 + \frac{a}{(\pi T F_a)^2}}$$

entering the nonlinear amplifier Fig. 1 produces an output that is a summation of Gaussian pulses each characterized by the integer N . The larger N is, the smaller are the amplitude and width of the corresponding pulse. Because of the linearity of the circuit following the amplifier, the envelope of the normalized output of the system (23) is the envelope of the sum of the transients produced by the Gaussian pulses.

For

$$t = \rho = 0, \quad (24)$$

(23) measures the normalized maximum output intensity through a channel ($\rho = 0$) which, expressed in db, we call output compression:

$$20 \log \frac{E_3(0)}{E_0(0)} = 20 \log_{10} \left[1 + \sum_{N=1}^{\infty} \frac{A_N}{A_0} \frac{T_N}{T_0} \sqrt{\frac{B_N}{B_0}} e^{(n\mu A/2)^2 (B_N - B_0)} \right]. \quad (25)$$

Naturally, if there were no compression (linear amplifier), A_N would be zero for all N and the output compression would be zero db.

The amplifier compression characterizes the nonlinear amplifier. It measures in db the ratio between the amplitude of a sine wave input and the amplitude of the output without harmonics. It is derived from (25), (5), (6), and (13), making $T = \infty$:

$$20 \log \frac{E_4(0)}{E_0(0)} = 20 \log \left(1 + \sum_1^{\infty} \frac{A_N}{A_0} \right). \quad (26)$$

Assuming the summations in (25) and (26) to be small compared to unity, it is possible to expand the logarithms in series and to retain the first term; then the ratio between output compression and amplifier compression is

$$\frac{\log \frac{E_3(0)}{E_0(0)}}{\log \frac{E_4(0)}{E_0(0)}} = \frac{\sum_1^{\infty} \frac{A_N}{A_0} \sqrt{\frac{B_N}{(2N+1)B_0}} e^{(n\mu A/2)^2 (B_N - B_0)}}{\sum_1^{\infty} \frac{A_N}{A_0}}. \quad (27)$$

Furthermore, if the amplifier has only compression of order $2N + 1$, the previous expression becomes

$$\frac{\log \frac{E_3(0)}{E_0(0)}}{\log \frac{E_4(0)}{E_0(0)}} = \sqrt{\frac{B_N}{(2N+1)B_0}} e^{(n\mu A/2)^2 (B_N - B_0)}. \quad (28)$$

This function has been plotted in Fig. 4. In abscissas we carry both $(F_1/F_{1\beta})^2$, the square of the ratio between the total transmitting bandwidth and the bandwidth of the filter after the amplifier, and $F_{1\alpha}/F_{1\beta}$, the ratio between bandwidth of the filters preceding and following the

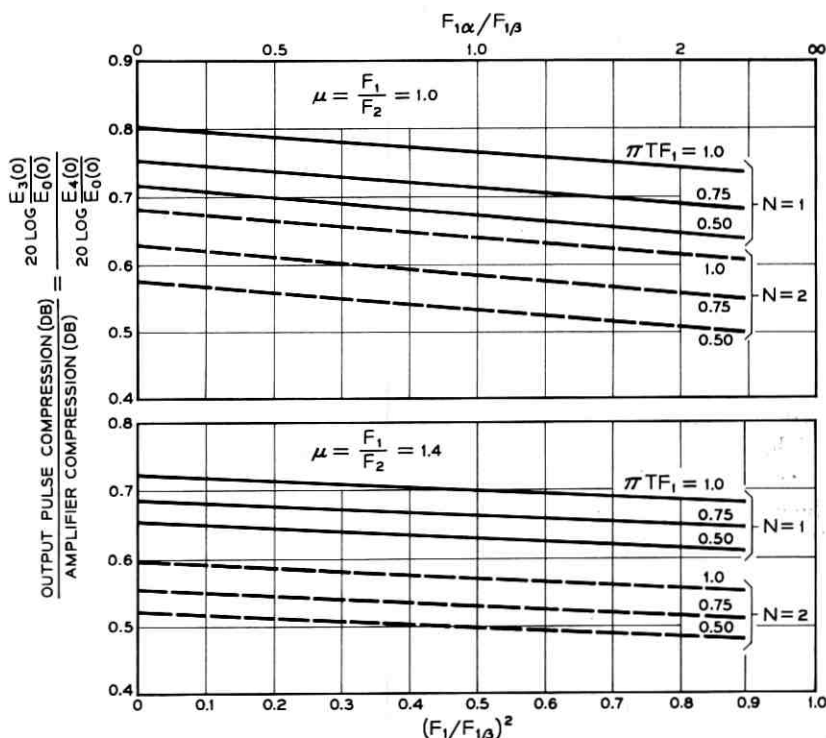


Fig. 4 — Compression of output pulse due to nonlinearity of amplifier.

amplifier in the transmitting end. The two functions are related by the following expression:

$$\frac{F_{1\alpha}}{F_{1\beta}} = \left[\left(\frac{F_{1\alpha}}{F_1} \right)^2 - 1 \right]^{\frac{1}{2}}.$$

The upper and lower sets of curves correspond to different ratios between transmitting and receiving bandwidths $\mu = F_1/F_2$. The solid lines correspond to amplifier compression of third order ($2N + 1 = 3$) and the dashed lines to amplifier compression of fifth order ($2N + 1 = 5$). The parameter πTF_1 is proportional to input pulse width $2T$, and transmitting bandwidth $2F_1$.

Let us use these curves for an example. Suppose that the amplifier compression is -2 db. This means that the amplifier output is 2 db below what it would be if the amplifier were linear. Let us assume further that the amplifier's compression is of third order ($N = 1$) and that

$$\mu = \frac{F_1}{F_2} = 1; \quad \frac{F_{1\alpha}}{F_{1\beta}} = 1; \quad \pi TF_1 = 1.$$

For these parameters the upper curve in Fig. 4 yields

$$\frac{\text{output compression}}{\text{amplifier compression}} = 0.763,$$

and, since the amplifier compression is -2 db, the system output compression is -1.53 db. This means that the amplifier compresses more than the system of which it is a part.

The output compression can be reduced by decreasing πTF_1 and by increasing F_1/F_2 , $F_{1\alpha}/F_{1\beta}$, and N . In other words, for given input pulse width T , the output compression is reduced by placing as much filtering as possible after the amplifier and by using an amplifier that works like a limiter (large N). These conclusions were to be anticipated if we recall that the compression of order $2N + 1$ introduces a spurious Gaussian pulse in the output of the amplifier. Its effect on the output can be reduced not only by filtering as much as possible but also by making that spurious pulse very narrow (large N). Using the example elaborated previously and making $N = 2$, the output compression is reduced from -1.53 db to -1.4 db.

Because of the compression of the top of the output pulse, time and frequency crosstalks are increased by

$$20 \log \frac{E_3(0)}{E_0(0)} \text{ db},$$

and the signal-to-noise ratio is reduced by the same amount with respect to the values that would be obtained if the amplifier were linear.

The crosstalks are different from the linear case, not only because of the compression of the top of the output pulse but also because the tails of the pulses are changed, thus modifying the time crosstalk (extra time crosstalk), and because the narrow spurious pulses have a wide spectrum that change the frequency crosstalk (extra frequency crosstalk). It is shown in Appendix C that in computing time crosstalk it is slightly conservative not to take into account the extra time crosstalk, and in computing frequency crosstalk it is conservative to ignore both the output compression and the extra frequency crosstalk.

V. SIGNAL-TO-NOISE RATIO

We want to calculate the maximum signal-to-noise ratio in the system shown in Fig. 1. Assume (a) the transmitter and receiver tuned at the same frequency $f_1 = f_2$, (b) the nonlinearity of the transmitting amplifier of order $2N + 1$, and (c) small insertion losses (about 2 db) due to compression and proximity of neighboring channel-dropping filters.

Then, except for a constant measuring the insertion loss of the transmission media, the signal-to-noise ratio in db is derived from (1) and (100) of Appendix D:

$$20 \log \frac{E(0)}{\sigma} = 20 \log \left[\frac{E(0)}{\sigma} \right]_0 + 20 \log \frac{E_1(0)}{E_0(0)} + 20 \log \frac{E_3(0)}{E_0(0)}, \quad (29)$$

in which the first term,

$$20 \log \left[\frac{E(0)}{\sigma} \right]_0 = 20 \log \left(\frac{2}{\pi} \right)^{\frac{1}{2}} \frac{\sqrt{\left[1 + \frac{a}{(\pi T F_{1a})^2} \right] B_0 e^{(n\mu A/2)^2 B_0}}}{\left\{ W_0 F_2 \left[1 + \frac{1}{(-2)^r} \left(\frac{F_2}{F_{2r}} \right)^{2r} \frac{\partial^r}{\partial b^r} \right] \frac{1 + e^{n^2/2b}}{\sqrt{b}} \right\}^{\frac{1}{2}}}} \quad (30)$$

measures in db the reference signal-to-noise ratio in the system — that is, the one that would be obtained if there were no compression and no neighboring channels. The other terms are already-known corrections. The second term measures the insertion loss due to neighboring channels (18) and is plotted in Figs. 2 and 3; the third term measures the insertion loss due to compression (25), and can be derived from Fig. 4.

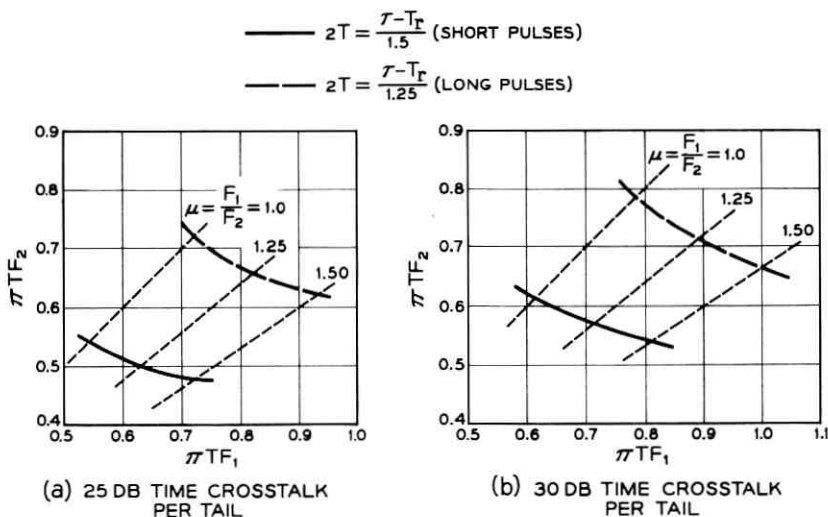


Fig. 5 — Relation between tolerable crosstalk filtering and pulses.

Before plotting the first term we are going to establish a functional relationship among several parameters. This relationship answers the following question: Given the tolerable time crosstalk at the detector of a system in which the input is a train of Gaussian pulses $2T$ wide and τ seconds apart and the sampling time is $2T_r$, what are the possible combinations of bandwidths $2F_1$ and $2F_2$, of the Gaussian transmitter and maximally flat receiver respectively, that satisfy those demands? The answers have been given in Figs. 4 and 7 of Ref. 3. We reproduce them in a convenient way in Figs. 5(a) and 5(b). They correspond to 25 and 30 db of time crosstalk per tail. The parameter for the solid curves is

$$2T = \frac{\tau - T_r}{1.5} \quad (\text{short pulses})$$

and for the dashed curves it is

$$2T = \frac{\tau - T_r}{1.25} \quad (\text{long pulses}).$$

Straight lines of constant ratios $\mu = F_1/F_2$ have been added for convenience. The curves have been limited to the region between the straight lines $\mu = 1$ and $\mu = 1.5$ because, as will be seen later, that is the region where the signal-to-noise ratios pass through desirable maxima.

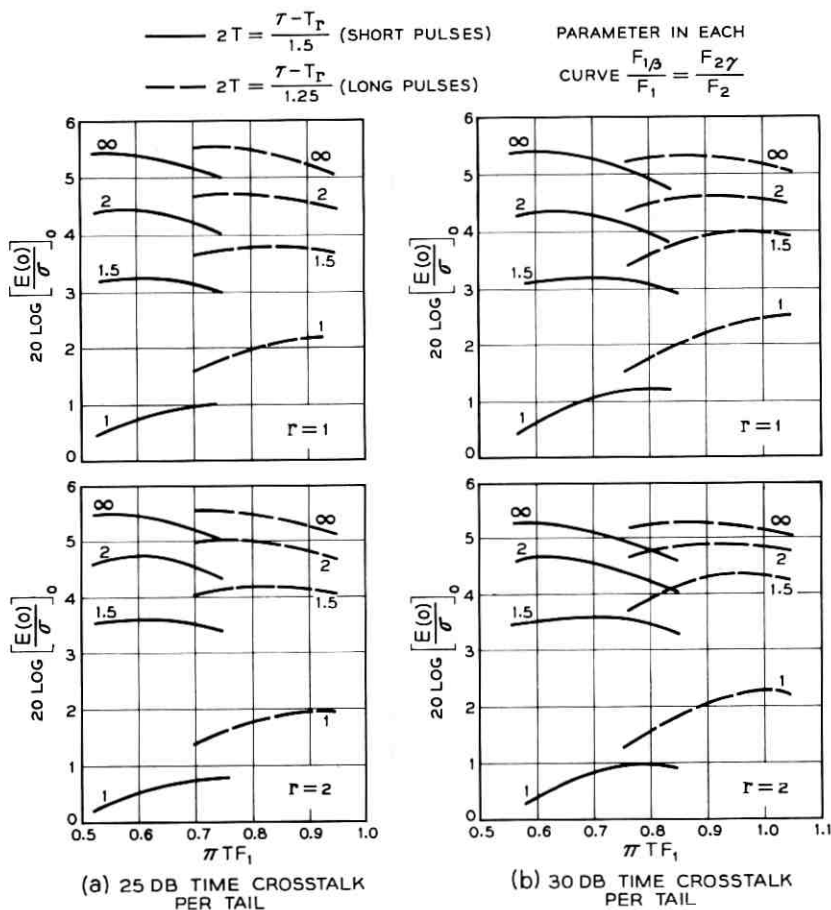


Fig. 6 — Reference signal-to-noise ratio in db.

The values of F_1 , F_2 , and T from the curves in Fig. 5 have been substituted in (30), and this reference signal-to-noise ratio expressed in db has been plotted with arbitrary reference in Fig. 6. The first and second rows of curves correspond to receiving channel-dropping filters of order $r = 1$ and $r = 2$ respectively. The first and second columns of curves correspond to 25 and 30 db of time crosstalk per tail. The solid lines correspond to shorter input pulses than do the dashed ones, and the parameters on each curve are different ratios between the bandwidths $2F_{1\beta}$ and $2F_{2\gamma}$ of the filters between amplifiers and transmission mediums and the bandwidths $2F_1$ and $2F_2$ of transmitter and receiver.

Let us examine how sensitive the signal-to-noise ratio is to each parameter.

Any curve in the first row ($r = 1$) and the corresponding one in the second ($r = 2$) differ only by a few tenths of a db, and therefore the signal-to-noise ratio is fairly insensitive to the order r of the receiving channel dropping filter.

Corresponding curves in the first and second columns (different time crosstalk) have similar ordinates. This means that crosstalk can be reduced at the expense of bandwidth without changing signal-to-noise level. Signal-to-noise level is improved by changing from short input pulses (solid lines) to long input pulses (dashed lines), but the advantage is paid for in channel bandwidth.

Finally, we find that signal-to-noise ratio is very sensitive to the parameter

$$\frac{F_{1\beta}}{F_1} = \frac{F_{2\gamma}}{F_2},$$

which is essentially proportional to the bandwidth of the filter connecting transmitting and receiving amplifiers. The signal-to-noise ratio grows by widening this bandwidth, but it cannot be widened indefinitely because it increases the insertion loss due to neighboring channel-dropping filters, Figs. 2 and 3. Nevertheless, curves corresponding to

$$\frac{F_{1\beta}}{F_1} = \frac{F_{2\gamma}}{F_2} = \infty$$

have been included in Fig. 6 to show the maximum achievable signal-to-noise ratio.

VI. DESIGN EXAMPLES

Given

- $1/\tau = 160$ mc, pulse repetition rate;
- $T = T_r$, equal sampling time and input pulse width;
- 25 db time crosstalk per tail;
- 2 db, third-order amplifier compression;
- $q = r = 1$, channel-dropping filters of first order;

calculate $2T, 2F_1, 2F_2, 2F_{1\alpha}, 2F_{1\beta}, 2F_{1\gamma}, 2F_{2\gamma}, |f_1 - f_3|$. In order to have low insertion loss due to neighboring channels and large signal-to-noise ratio we select

$$\mu = \frac{F_1}{F_2} = 1, \quad (31)$$

$$F_{1\gamma} = F_{1\beta} = F_{2\gamma}. \quad (32)$$

From (a), (b), (31), and Fig. 5(a), we obtain for short input pulses

$$\pi T F_1 = 0.54, \quad (33)$$

$$2T = 3.125 \text{ millimicroseconds}, \quad (34)$$

$$2F_1 = 2F_2 = 220 \text{ mc.} \quad (35)$$

From (c), (e), (33), and Fig. 6, the reference signal-to-noise ratio results:

$$20 \log \left[\frac{E(0)}{\sigma} \right]_0 = \begin{cases} 4.4 \text{ db} & \text{for } \frac{F_{1\beta}}{F_1} = \frac{F_{1\gamma}}{F_1} = \frac{F_{2\gamma}}{F_2} = 2 \\ 3.2 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 1.5 \end{cases} \quad (36)$$

From (d), (31), (33), and Fig. 4, we calculate the insertion loss due to compression and consequent increase in time crosstalk per tail:

$$20 \log \frac{E_0(0)}{E_3(0)} = \begin{cases} 0.7 \times 2 = 1.4 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 2 \\ 0.68 \times 2 = 1.36 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 1.5 \end{cases} \quad (37)$$

From (e), (31), (32), and (33), the insertion loss due to neighboring channels is given approximately in the upper dotted curves in the center and lower right set of curves in Fig. 2. The abscissa which determines the frequency occupancy $f_1 - f_3$, should be selected as small as possible. We will elaborate two examples:

Example 1.

We select

$$\rho = \left| \frac{f_1 - f_3}{2F_2} \right| = 2. \quad (38)$$

The insertion loss due to neighboring channels is

$$20 \log \frac{E_0(0)}{E_1(0)} = \begin{cases} \sim 4 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 2 \\ 1.7 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 1.5 \end{cases} \quad (39)$$

The total signal-to-noise ratio (29) derived from (36), (37), and (39) is

$$20 \log \frac{E(0)}{\sigma} = \begin{cases} 4.4 - 1.4 - 4 = -1 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 2 \\ 3.2 - 1.36 - 1.7 = 0.14 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 1.5 \end{cases} \quad (40)$$

Because of the better signal-to-noise ratio we adopt

$$\frac{F_{1\gamma}}{F_1} = 1.5 \quad (41)$$

and deduce from (32), (34), (35), (38), and (41):

$$\begin{aligned} 2T &= 3.125 \text{ millimicroseconds,} \\ 2F_1 &= 2F_2 = 220 \text{ mc,} \\ 2F_1 &= 2F_{1\beta} = 2F_{2\gamma} = 330 \text{ mc,} \\ 2F_{1\alpha} &= \frac{2F_1}{\sqrt{1 - \left(\frac{F_1}{F_{1\gamma}}\right)^2}} = 295 \text{ mc,} \end{aligned} \quad (42)$$

$$|f_1 - f_3| = 440 \text{ mc.}$$

Example 2.

We select

$$\rho = \frac{|f_1 - f_3|}{2F_2} = 3. \quad (43)$$

The insertion loss due to neighboring channels is (Fig. 2),

$$20 \log \frac{E_0(0)}{E_1(0)} = \begin{cases} 1.3 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 2 \\ 0.7 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 1.5 \end{cases} \quad (44)$$

Then from (29), (36), (37), and (44), the total signal-to-noise ratio is

$$20 \log \frac{E(0)}{\sigma} = \begin{cases} 4.4 - 1.4 - 1.3 = 1.7 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 2 \\ 3.2 - 1.36 - 0.7 = 1.14 \text{ db} & \text{for } \frac{F_{1\gamma}}{F_1} = 1.5 \end{cases} \quad (45)$$

Selecting the best signal-to-noise ratio, we adopt

$$\frac{F_{1\gamma}}{F_1} = 2 \quad (46)$$

and deduce from (32), (34), (35), (43), and (46):

$$\begin{aligned} 2T &= 3.125 \text{ millimicroseconds,} \\ 2F_1 &= 2F_2 = 220 \text{ mc,} \\ 2F_{1\gamma} &= 2F_{1\beta} = 2F_{2\gamma} = 440 \text{ mc,} \\ 2F_{1\alpha} &= \frac{2F_1}{\sqrt{1 - \left(\frac{F_1}{F_{1\gamma}}\right)^2}} = 254 \text{ mc,} \end{aligned} \quad (47)$$

$$|f_1 - f_3| = 660 \text{ mc.}$$

In the first example, the signal-to-noise ratio (40) and the channel spacing (42) are 0.14 db and 440 mc respectively. In the second example these values, (45) and (47), are 1.7 db and 660 mc. The first example provides better frequency occupancy and the second provides greater spacing between successive repeaters.

VII. RÉSUMÉ OF RESULTS AND CONCLUSIONS

The influence of (a) neighboring channel-dropping filters, (b) non-linearity of transmitting amplifiers, and (c) filtering preceding and following the transmitting and receiving amplifiers on the signal-to-noise ratio, time crosstalk and frequency crosstalk in a frequency-division multiplex PCM system has been calculated.

Neighboring channel-dropping filters reduce the reference* signal-to-noise ratio by increasing the signal insertion loss, Figs. 2 and 3. If these filters are of first order, the channel spacing is determined by the tolerable insertion loss they introduce and not by the frequency crosstalk, as calculated in Ref. 3. If these filters are of order higher than one, the channel spacing may be determined by the tolerable insertion loss or by the tolerable frequency crosstalk, (Figs. 5 and 8 of Ref. 3), depending on which demands more channel spacing. Neighboring channel-dropping filters have no practical bearing on time crosstalk.

Nonlinearity of the transmitting amplifiers decreases the reference signal-to-thermal-noise ratio and increases time crosstalk by the same amount, Fig. 4, but it is conservative to ignore its effect on frequency crosstalk.

* Signal-to-noise ratio in the absence of neighboring channels and with linear amplifiers.

The tolerance to compression is increased by concentrating as much filtering as possible between the nonlinear transmitting amplifier and the detector.

Even for compression as large as 2 db the increase in frequency and time crosstalk is smaller than the amplifier compression. This means that the system can be readily designed to take advantage of the greater power available from an amplifier operating in the nonlinear region.

Finally, reducing the filtering between transmitting and receiving amplifiers increases the reference signal-to-noise ratio, Fig. 6, but it also increases the signal insertion loss due to the presence of neighboring channel-dropping filters, Figs. 2 and 3.

APPENDIX A

Transmission Through Nonlinear Amplifier and Filters

Given a system consisting of five filters and a nonlinear amplifier, Fig. 1, we want to calculate the envelope of the response to a certain input.

The transfer characteristics of the cascaded filters are:

$$y_{1\alpha} = e^{-a(|f|-f_1)/F_{1\alpha}}^2, \quad (48)$$

$$y_{1\beta} = e^{-a(|f|-f_1)/F_{1\beta}}^2, \quad (49)$$

$$y_2 = e^{-b(|f|-f_2)/F_2} \cosh n \frac{|f| - f_2}{F_2}, \quad (50)$$

$$y_{3\gamma} = \frac{1}{\sqrt{1 + \left(\frac{F_{1\gamma}}{|f| - f_3}\right)^{2q}}}, \quad (51)$$

$$y_{4\gamma} = \frac{1}{\sqrt{1 + \left(\frac{F_{2\gamma}}{|f| - f_4}\right)^{2r}}}, \quad (52)$$

where

$$a = 0.346,$$

$$b = 1.3,$$

$$n = 1.61,$$

and f_1, f_2, f_3, f_4 are the center frequencies of the different filters; $2F_{1\alpha}, 2F_{1\beta}, 2F_2, 2F_{1\gamma}, 2F_{2\gamma}$ are their half-power bandwidths; and $y_{1\alpha}$ and $y_{1\beta}$ are Gaussian transfer characteristics and y_2 is the transfer characteristic of an approximately third-order maximally flat filter.³

The characteristic $y_{3\gamma}$ is the transfer function between two ports of a three-port channel-dropping filter^{4,5,6} proposed for use in a long distance waveguide communication system.¹ This filter 3γ (see Fig. 1) has port 1 matched at all frequencies, and the transfer function between ports 1 and 3 is that of a phase-equalized maximally flat filter of order q , bandwidth $2F_{1\gamma}$, and center frequency f_3 . Because of conservation of energy it follows that the phase equalized transfer function between ports 1 and 2 is (51). $y_{4\gamma}$ is similar to $y_{3\gamma}$.

Between the first two filters 1α and 1β in Fig. 1 there is a nonlinear amplifier. Its infinitely wide band output voltage $v_0(t)$ as a function of the input $v_i(t)$ is

$$v_0(t) = \sum_0^{\infty} a_v [v_i(t)]^v. \quad (53)$$

The input to the system,

$$i(t) = \sqrt{1 + \frac{a}{(\pi T F_\alpha)^2}} e^{-(t/T)^2} \cos 2\pi f_1 t, \quad (54)$$

is a Gaussian pulse $2T$ seconds wide, at 8.686 db down, that modulates a carrier f_1 . The amplitude

$$\sqrt{1 + \frac{a}{(\pi T F_\alpha)^2}}$$

has been selected to make the maximum amplitude of the signal $v_i(t)$ reaching the amplifier unity.

In order to evaluate the system's output we start calculating the amplifier's output $v_0(t)$. This is achieved in (53) by expressing $v_i(t)$ in terms of the input $i(t)$ to the system

$$v_0(t) = \sum_0^{\infty} a_v \left[\int_{-\infty}^{\infty} g(f) y_{1\alpha} e^{i2\pi f t} df \right]^v, \quad (55)$$

in which

$$g(f) = \sqrt{\pi} \frac{T}{2} \sqrt{1 + \frac{a}{(\pi T F_\alpha)^2}} (e^{-[\pi T (f-f_1)]^2} + e^{-[\pi T (f+f_1)]^2}) \quad (56)$$

is the Fourier transform of $i(t)$ of (54).

Substituting (48) and (56) in (55), performing the integration, and assuming

$$e^{-(\pi T F_1)^2} \ll 1, \quad (57)$$

one derives the explicit value of the amplifier's output:

$$v_0(t) = \sum_0^{\infty} a_\nu \cos^\nu 2\pi f_1 t \exp \left[- \frac{\nu}{1 + \frac{a}{(\pi T F_\alpha)^2}} \left(\frac{t}{T} \right)^2 \right]. \quad (58)$$

Since

$$\cos^\nu 2\pi f_1 t = \frac{1}{2^\nu} \sum_{r=0}^{\nu} \binom{\nu}{r} \cos (\nu - 2r) 2\pi f_1 t, \quad (59)$$

where

$$\binom{\nu}{r} = \frac{\nu!}{(\nu - r)! r!}, \quad (60)$$

(58) becomes

$$v_0(t) = \sum_{\nu=0}^{\infty} \sum_{r=0}^{\nu} \frac{a_\nu}{2^\nu} \binom{\nu}{r} \exp \left[- \frac{\nu}{1 + \frac{a}{(\pi T F_\alpha)^2}} \left(\frac{t}{T} \right)^2 \right] \cos (\nu - 2r) 2\pi f_1 t. \quad (61)$$

A further simplification: assume that only frequencies in the neighborhood of f_1 are allowed to pass through the filters following the amplifier. Physically, this means that all harmonics of f_1 are filtered out; mathematically, this is translated by keeping in (61) only the terms in which

$$\nu - 2r = \pm 1. \quad (62)$$

Consequently,

$$v_0(t) = \cos 2\pi f_1 t \sum_0^{\infty} A_N e^{-(t/T_N)^2}, \quad (63)$$

where

$$A_N = 2^{-2N} \binom{2N+1}{N} a_{2N+1}, \quad (64)$$

$$T_N = T \sqrt{\frac{1 + \frac{a}{(\pi T F_\alpha)^2}}{2N+1}}. \quad (65)$$

The output from the nonlinear amplifier $v_0(t)$ is, then, a summation of Gaussian pulses each of amplitude A_N and width $2T_N$ measured 8.686 db down.

Using $v_0(t)$ as input to the second filter 1β in Fig. 1, we calculate the Fourier transform of each term with the help of (56) and deduce the system's output:

$$e(t) = \frac{\sqrt{\pi}}{2} \sum_0^{\infty} A_N T_N \int_{-\infty}^{\infty} y_{1\beta} y_2 y_{3\gamma} y_{4\gamma} e^{-[\pi T_N (|f| - f_1)]^2 + i2\pi f t} df. \quad (66)$$

We perform a change of variable,

$$f = \varphi + f_1, \quad (67)$$

and assume

$$e^{-(\pi T_N f_1)} \ll 1. \quad (68)$$

Then the envelope of the output (66) becomes

$$E(t) = \sqrt{\pi} \left| \sum_0^{\infty} A_N T_N \int_{-\infty}^{\infty} e^{-(\pi T_N \varphi)^2 + i2\pi \varphi t} y_{1\beta} y_2 y_{3\gamma} y_{4\gamma} d\varphi \right|, \quad (69)$$

where $y_{1\beta}$, y_2 , $y_{3\gamma}$, $y_{4\gamma}$ are deduced from (49), (50), (51), (52), and (67) to be

$$y_{1\beta} = e^{-a(\varphi/F_{1\beta})^2}, \quad (70)$$

$$y_2 = e^{-b(\varphi/F_2)^2} \cosh n \frac{\varphi + f_1 - f_2}{F_2}, \quad (71)$$

$$y_{3\gamma} = \frac{1}{\sqrt{1 + \left(\frac{F_{1\gamma}}{\varphi + f_1 - f_3}\right)^{2q}}}, \quad (72)$$

$$y_{4\gamma} = \frac{1}{\sqrt{1 + \left(\frac{F_{2\gamma}}{\varphi + f_1 - f_4}\right)^{2r}}}. \quad (73)$$

In order to solve the integral in (69), we simplify part of the integrand by noticing that, on account of the exponential functions, most of the contribution to the result comes from small values of φ , and consequently $y_{3\gamma}$, $y_{4\gamma}$ can be expanded in power series. We keep the first three terms

$$y_{3\gamma} y_{4\gamma} = \frac{1 + \frac{\varphi}{\varphi_1} - \left(\frac{\varphi}{\varphi_2}\right)^2}{\sqrt{(1+S)(1+R)}}, \quad (74)$$

where

$$S = \left(\frac{F_{1\gamma}}{f_1 - f_3}\right)^{2q}, \quad (75)$$

$$R = \left(\frac{F_{2\gamma}}{f_1 - f_4}\right)^{2r}, \quad (76)$$

$$\varphi_1 = \frac{f_1 - f_3}{\frac{qS}{1+S} \mp \frac{rR}{1+R}}, \quad (77)$$

$$\varphi_2 = \frac{f_1 - f_4}{\left[\frac{q(2q+1)S}{2(1+S)} + \frac{r(2r+1)R}{2(1+R)} \pm \frac{qrSR}{(1+S)(1+R)} - \frac{3}{2} \left(\frac{qS}{1+S} \right)^2 - \frac{3}{2} \left(\frac{rR}{1+R} \right)^2 \right]^{\frac{1}{2}}}. \quad (78)$$

Where double signs are indicated, the upper one is to be used if $f_3 = 2f_1 - f_4$ (neighboring channels at opposite sides of carrier f_1), and the lower one if $f_3 = f_4$ (neighboring channels at the same side of carrier f_1).

Substituting (70), (71), and (74) in (69) and integrating, one obtains the output envelope

$$E(t) = \left| \sum_{N=0}^{\infty} \left[I_N - \frac{i}{2\pi\varphi_1} \frac{\partial I_N}{\partial t} + \frac{1}{(2\pi\varphi_2)^2} \frac{\partial^2 I_N}{\partial t^2} \right] \right|, \quad (79)$$

in which

$$I_N = A_N \frac{T_N}{T} \sqrt{\frac{B_N}{(1+S)(1+R)}} e^{(n\mu A^2/2)B_N - 4\rho^2 b(1-b\mu^2 A^2 B_N) - (t/T)^2 B_N} \cdot \cosh \left[2n\rho(1 - b\mu^2 A^2 B_N) + in\mu A B_N \frac{t}{T} \right], \quad (80)$$

$$B_N = \frac{1}{1 + A^2(a + b\mu^2) - 2N \left(\frac{T_N}{T} \right)^2}, \quad (81)$$

$$\mu = \frac{F_1}{F_2}, \quad (82)$$

$$\rho = \frac{|f_1 - f_2|}{2F_2}. \quad (83)$$

APPENDIX B

Influence of Channel-Dropping Filters on Crosstalk

We want to compare the frequency and time crosstalks of the system in Fig. 1, assuming a linear amplifier with the already known³ frequency and time crosstalk of the same system obtained assuming not only linear amplifier but also no neighboring channel-dropping filters.

The maximum frequency crosstalk between neighboring channels is

deduced from (69), making

$$A_N = 0 \text{ for } N \neq 0 \text{ (linear amplifier),}$$

$f_1 \neq f_2$ (transmitter and receiver centered at neighboring frequencies),

$f_3, f_4 \neq f_1$ (f_3 and f_4 are the center frequencies of neighboring channel-dropping filters),

$$t = 0 \text{ (in order to maximize the integrand).}$$

Then

$$E_1(0) = \sqrt{\pi} A_0 T_0 \int_{-\infty}^{\infty} e^{-(\pi T_0 \varphi)^2} y_{1\beta} y_{2\gamma} y_{3\gamma} y_{4\gamma} d\varphi. \quad (84)$$

We eliminate the effect of neighboring channel-dropping filters by assuming f_3 and f_4 to be infinitely large frequencies. Then $y_{3\gamma}$ and $y_{4\gamma}$ given in (72) and (73) become unity and

$$E_0(0) = \sqrt{\pi} A_0 T_0 \int_{-\infty}^{\infty} e^{-(\pi T_0 \varphi)^2} y_{1\beta} y_2 d\varphi. \quad (85)$$

This result has been calculated elsewhere³ and consequently is known. The reader can check that $E_1(0)$ is slightly smaller than $E_0(0)$, because the product $y_{3\gamma} y_{4\gamma}$ in (84) is slightly smaller than one for values of φ for which the rest of the always-positive integrand contributes substantially to the result. Therefore, it is slightly conservative to ignore the effect of the neighboring channel-dropping filters on frequency crosstalk.

In order to evaluate the influence of neighboring channel-dropping filters on time crosstalk, we must compare the maximum values of two transients during the sampling time. The first transient (17) is that of a pulse through a system with neighboring channel-dropping filters; the second transient is that of the same pulse through the same system except for the removal of the neighboring channel-dropping filters. The last transient is obtained from (17) making $f_3 = f_4 = \infty$,

$$\frac{E_0(t)}{E_0(0)} = e^{-(t/T)^2 B} \cos \frac{\mu A B n t}{T}, \quad (86)$$

and the time crosstalk that is derived from this expression is known.³

If the pulses must be closely spaced, the maximum time crosstalk must be fixed by the sloping part of the transients, and consequently the broadening of the pulse due to the presence of the neighboring channel-dropping filters indirectly measures the time crosstalk. For small values of t , (17) and (24) yield, after power series expansion,

$$\frac{E_1(t)}{E_0(0)} \cong 1 - (1 + \frac{1}{2} n^2 \mu^2 A^2 B) \left[1 - \frac{2B(1 + \frac{1}{2} n^2 \mu^2 A^2 B)}{(2\pi\varphi_1 T)^2} \right] B \frac{t^2}{T^2}, \quad (87)$$

$$\frac{E_0(t)}{E_0(0)} \cong 1 - (1 + \frac{1}{2}n^2\mu^2A^2B)B \frac{t^2}{T^2}. \quad (88)$$

Since expressions (87) and (88) have the same functional dependence, the transient of a system with neighboring channel-dropping filters is equivalent to that of a system without neighboring channel-dropping filters that operates with a wider input pulse

$$2T_1 = 2T \left[1 + B \frac{1 + \frac{1}{2}n^2\mu^2A^2B}{(2\pi\varphi_1T)^2} \right]. \quad (89)$$

It is possible to prevent any widening, $T_1 = T$, by making $\varphi_1 = \infty$. This is achieved in (6) by selecting the center frequencies f_3 and f_4 of the neighboring channels at opposite sides of the center frequency f_1 of the channel in which the time crosstalk is being studied. But even if φ_1 is finite, the relative pulse width increase $(T_1/T) - 1$ is negligibly small for typical cases. For example, if one selects the parameters of a system with 1 db insertion loss due to neighboring channel-dropping filters (center set of curves, Fig. 2):

$$\frac{F_{1\alpha}}{F_1} = \frac{F_{2\alpha}}{F_2} = 1.5,$$

$$\mu = \frac{F_1}{F_2} = 1,$$

$$q = r = 1,$$

$$\pi TF_1 = 0.5,$$

$$\frac{|f_1 - f_3|}{2F_2} = \frac{|f_1 - f_4|}{2F_2} = 2.6,$$

it follows from (89) that

$$\frac{T_1}{T} - 1 = 0.025.$$

APPENDIX C

Influence of Nonlinearity of the Transmitting Amplifier on Frequency and Time Crosstalk

The nonlinearity of the transmitting amplifier changes crosstalk by compressing the top of the output pulse and because the spurious pulses generated by the amplifier change the tail of the main output pulse

(extra time crosstalk) and part of their spectra is fed into neighboring channels (extra frequency crosstalk).

The time crosstalk at the instant $t = \zeta T$ is by definition the ratio between the envelope of the channel output at that instant and the maximum output expressed in db. From (23), calling $E_3(t)$ the value acquired by $E_2(t)$ if $\rho = 0$ ($f_1 = f_2$), it is

$$20 \log \frac{E_3(\zeta T)}{E_3(0)} = 20 \log e^{-B_0 \zeta^2} \left| \cos n\mu A B_0 \zeta \right| \\ - 20 \log \left[1 + \sum_1^{\infty} \frac{A_N}{A_0} \frac{T_N}{T_0} \sqrt{\frac{B_N}{B_0}} e^{(n\mu A/2)^2 (B_N - B_0)} \right] \quad (90) \\ + 20 \log \left[1 + \sum_1^{\infty} \frac{A_N}{A_0} \frac{T_N}{T_0} \sqrt{\frac{B_N}{B_0}} e^{[(n\mu A/2)^2 - \zeta^2] (B_N - B_0)} \right. \\ \left. \cdot \frac{\cos n\mu A B_N \zeta}{\cos n\mu A B_0 \zeta} \right].$$

The first term is the main contribution to time crosstalk and has been computed elsewhere.³ The second term is the contribution due to output compression, $20 \log [E_3(0)/E_0(0)]$, (25), and the third term is what we called extra time crosstalk in the text.

Assume:

- (a) only one term of nonlinearity (only compression of third, fifth, etc. order),
- (b) small output compression (2 db at most),
- (c) such a pessimistic combination of parameters that

$$\cos n\mu A B_0 \zeta = \cos n\mu A B_N \zeta = 1. \quad (91)$$

Then (90) becomes

$$20 \log \frac{E_3(\zeta T)}{E_3(0)} \cong 20 \log e^{-B_0 \zeta^2} - (1 - e^{-\zeta^2 (B_N - B_0)}) 20 \log \frac{E_3(0)}{E_0(0)}. \quad (92)$$

For the range of values of the different parameters given in Fig. 4 and $\zeta > 2$,

$$e^{-\zeta^2 (B_N - B_0)} \ll 1, \quad (93)$$

and consequently it is a slightly conservative assumption to neglect the extra time crosstalk due to compression.

Now we turn to frequency crosstalk. The maximum frequency crosstalk is by definition the ratio between $E_3(0)$, the maximum of the envelope output in a neighboring channel, and $E_2(0)$, the maximum output

in the channel. $E_2(0)$ is calculated from (23), making $t = 0$, and $E_3(0)$ is also derived from (23), making $t = \rho = 0$. The ratio expressed in db is

$$20 \log \frac{E_2(0)}{E_3(0)} = 20 \log e^{-4\rho^2 b(1-b\mu^2 A^2 B_0)} \cosh 2n\rho(1 - b\mu^2 A^2 B_0) \\ - 20 \log \left[1 + \sum_{N=1}^{\infty} \frac{A_N}{A_0} \frac{T_N}{T_0} \sqrt{\frac{B_N}{B_0}} e^{(n\mu A/2)^2 (B_N - B_0)} \right] \\ + 20 \log \left[1 + \sum_{1}^{\infty} \frac{A_N}{A_0} \frac{T_N}{T_0} \sqrt{\frac{B_N}{B_0}} e^{(\mu A/2)^2 [n^2 + (4\rho b)^2] (B_N - B_0)} \right. \\ \left. \cdot \frac{\cosh 2n\rho(1 - b\mu^2 A^2 B_N)}{\cosh 2n\rho(1 - b\mu^2 A^2 B_0)} \right]. \quad (94)$$

The first term is the main contribution to frequency crosstalk and it has been computed elsewhere.³ The second term is the contribution due to output compression $20 \log [E_3(0)/E_0(0)]$, (25), and the third term is the extra frequency crosstalk.

Assuming that the amplifier has only one term of nonlinearity and that the output compression is small, the second term of (94) is simplified to

$$20 \log \frac{E_3(0)}{E_0(0)} \cong 8.686 \frac{A_N}{A_0} \frac{T_N}{T_0} \sqrt{\frac{B_N}{B_0}} e^{(n\mu A/2)^2 (B_N - B_0)}, \quad (95)$$

and consequently (94) can be rewritten

$$20 \log \frac{E_2(0)}{E_3(0)} = 20 \log e^{-4\rho^2 b(1-b\mu^2 A^2 B_0)} \cosh 2n\rho(1 - b\mu^2 A^2 B_0) \\ - 20 \log \frac{E_3(0)}{E_0(0)} + 20 \log \left[1 + 2.3 \log \frac{E_3(0)}{E_0(0)} e^{(2\mu A \rho b)^2 (B_N - B_0)} \right. \\ \left. \cdot \frac{\cosh 2n\rho(1 - b\mu^2 A^2 B_N)}{\cosh 2n\rho(1 - b\mu^2 A^2 B_0)} \right]. \quad (96)$$

If the channels must be crowded, ρ is small, but in that case the extra frequency crosstalk can be expanded in series and the second and third terms tend to cancel each other. In the limit, for $\rho = 0$, $20 \log [E_2(0)/E_3(0)] = 0$, as it should be, since by definition $E_2(0)$ and $E_3(0)$ coalesce to the same value.

On the other hand, for large values of ρ the extra frequency crosstalk (third term) becomes more important than the frequency crosstalk due to compression (second term), but then the total frequency crosstalk $20 \log [E_2(0)/E_3(0)]$ is negligibly small.

Consider, for instance, the following system, in which the parameters $\pi TF_1 = 0.5$, $F_1/F_2 = 1.4$, $F_{1\beta} = F_{2\alpha} = \infty$, $N = 2$, and 2 db of amplifier compression have been chosen to exaggerate the influence of nonlinearity on frequency crosstalk. Substituting these values in (96),

$$20 \log \frac{E_2(0)}{E_3(0)} = \begin{cases} -29.1 + 1.04 - 2.9 = -30.96 \text{ (in db)} & \text{for } \rho = 2, \\ -67.2 + 1.04 - 37.1 = -103.26 \text{ (in db)} & \text{for } \rho = 3. \end{cases} \quad (97)$$

The first term corresponds to frequency crosstalk for a linear amplifier, the second term is the frequency crosstalk correction due to output compression, and the third term is the extra frequency crosstalk. For small ρ ($\rho = 2$) it is conservative to ignore the second and third terms. For large ρ ($\rho = 3$) the total frequency crosstalk is negligibly small. Therefore, for all values of ρ and for all practical purposes the total frequency crosstalk can be calculated as if the amplifier were linear.

APPENDIX D

Noise Level at the Detector.

The distribution of the filtering before and after the first amplifier of the receiver, Fig. 1, is important because it determines the thermal noise power reaching the detector.

If the noise from the amplifier is white and the power density is W_0 watts per cycle, then the mean noise power received by the detector is

$$\sigma^2 = W_0 \int_{-\infty}^{\infty} |y_{2\alpha}|^2 df, \quad (98)$$

where

$$y_{2\alpha} = \frac{y_2}{y_{2\gamma}} = e^{-b(|f|-f_2)/F_2)^2} \cosh n \frac{|f| - f_2}{F_2} \sqrt{1 + \left(\frac{|f| - f_2}{F_{2\gamma}}\right)^{2r}} \quad (99)$$

is the transfer characteristic of the filter connecting the amplifier and the detector. This characteristic has been found as the ratio between y_2 , the transfer characteristic of the receiver (approximately third-order maximally flat), and $y_{2\gamma}$, the transfer characteristic of the receiving channel-dropping filter (maximally flat of order r).

Performing the integration (98), the average noise power received by the detector is

$$\sigma^2 = \sqrt{\frac{\pi}{2}} W_0 \left[1 + \frac{1}{(-2)^r} \left(\frac{F_2}{F_{2\gamma}} \right)^{2r} \frac{\partial^r}{\partial b^r} \right] \left(\frac{e^{n^2/2b} + 1}{\sqrt{b}} \right) F_2. \quad (100)$$

REFERENCES

1. Miller, S. E., Waveguide as a Communication Medium, *B.S.T.J.*, **33**, 1954, pp. 1209-1265.
2. Marcatili, E. A., Errors in Detection of RF Pulses Embedded in Time Crosstalk, Frequency Crosstalk, and Noise, *B.S.T.J.*, **40**, 1961, pp. 921-950.
3. Marcatili, E. A., Time and Frequency Crosstalk in Pulse-Modulated Systems, *B.S.T.J.*, **40**, 1961, pp. 951-970.
4. Marcatili, E. A., Channel-Dropping Filter in the Millimeter Region Using Circular Electric Modes, *I.R.E. Trans.*, **MTT-9**, 1961, pp. 176-182.
5. Marcatili, E. A., Mode Conversion Filters, *B.S.T.J.*, **40**, 1961, pp. 149-184.
6. Marcatili, E. A., A Circular Electric Hybrid Junction and Some Channel-Dropping Filters, *B.S.T.J.*, **40**, 1961, pp. 185-196.
7. Wheeler, H. A., The Interpretation of Amplitude and Phase Distortion in Terms of Paired Echoes, *Proc. I.R.E.*, **27**, 1939, pp. 359-386.

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. Member American Mathematical Society, Association for Symbolic Logic, Institute of Mathematical Statistics, Society for Industrial and Applied Mathematics, Mind Association, Phi Beta Kappa.

D. C. HOGG, B.Sc., 1949, University of Western Ontario; M.Sc., 1950, and Ph.D., 1953, McGill University; Bell Telephone Laboratories, 1953—. His work has included studies of artificial dielectrics for microwaves, diffraction of microwaves, and over-the-horizon and millimeter wave propagation. He was concerned with studies of sky noise and analysis of performance characteristics of microwave antennas most recently. Senior member I.R.E.; member Commission 2, U.R.S.I.; Sigma Xi.

A. KELLER, Ph.D., University of Bristol (England). Mr. Keller was a visiting scholar at Bell Telephone Laboratories during the summer of 1960. He is a specialist in morphology of high polymers. He is now a member of the physics staff at the University of Bristol.

ARCHIE P. KING, B.S., 1927, California Institute of Technology; Seismological Laboratory, Carnegie Institute of Washington, 1927–30; Bell Telephone Laboratories, 1930—. He has been engaged in ultra-high-frequency radio research and has concentrated on waveguide transmission and waveguide transducers, and components for low-loss circular electric wave transmission. Senior member I.R.E.

JOHN R. KLAUDER, B.S., 1953, University of California; M.S., 1956, Stevens Institute of Technology; M.A., 1957, and Ph.D., 1959, Princeton University; Bell Telephone Laboratories, 1953—. He was a member of the Communications Development Training Program from 1953 to

1956 and from 1956 to 1959 he held a C.D.T. Fellowship at Princeton, where he studied quantum field theory. Since returning to Bell Laboratories he has been engaged in theoretical studies of properties of solids. Member American Physical Society, Phi Beta Kappa, Sigma Xi, Tau Beta Pi.

J. A. LEWIS, B.S., 1944, Worcester Polytechnic Institute; Sc.M., 1948, and Ph.D., 1950, Brown University; Corning Glass Works, 1950-51; Bell Telephone Laboratories, 1951—. Mr. Lewis has been engaged in mathematical research in theoretical mechanics, piezoelectric crystal vibrations, heat transfer, and stress analysis. Member American Mathematical Society, Society for Industrial and Applied Mathematics.

G. D. MANDEVILLE, Monmouth Junior College; Rutgers University; Western Electric Co., 1939-49; Bell Telephone Laboratories, 1949—. With Western Electric, Mr. Mandeville was concerned with radar development and shop testing equipment. He headed the shop test equipment prove-in section for three years. With Bell Laboratories he has been engaged in various areas of measurement associated with guided-wave research.

E. A. MARCATILI, Aeronautical Engineer, 1947, and E.E., 1948, University of Cordoba (Argentina); Research staff, University of Cordoba, 1947-54; Bell Telephone Laboratories, 1954—. He has been engaged in theory and design of filters in multimode waveguides. More recently he has concentrated on waveguide systems research. Member I.R.E., Physical Association of Argentina.

DOREN MITCHELL, B.S., 1925, Princeton University; American Telephone & Telegraph Co., 1925-34; Bell Telephone Laboratories, 1934—. Mr. Mitchell's early work was concerned with field studies of transmission on long telephone and radio circuits, including work on various types of voice operated devices. During World War II he worked on military projects, including transmission systems and the problem of laying wire from airplanes. He also founded the Somerset Mechanics School to provide vocational training to residents of that county. Since the war, he has worked on radio systems and data transmission systems. Now working on satellite communication systems as Head, Satellite Systems Studies Department. Fellow of the IRE, licensed professional engineer and member of the Armed Forces Communications and Electronics Association and the American Association for the Advancement of Science.

RONALD SALOVEY, B.S., 1954, Brooklyn College; A.M., 1957, and Ph.D., 1959, Harvard University; Bell Telephone Laboratories, 1958—. Mr. Salovey has been engaged in research in polymer chemistry. Member American Chemical Society, Phi Beta Kappa, Sigma Xi.

R. A. SEMPLAK, B.S., 1961, Monmouth College; Bell Telephone Laboratories, 1955—. Mr. Semplak has been engaged in beyond-the-horizon propagation and sky temperature measurements with a maser. He also was concerned with the Project Echo satellite communications experiment.

PRISCILLA L. STEINBERG, B.A., 1954, Mount Holyoke College; Bell Telephone Laboratories, 1954—. Mrs. Steinberg has been engaged in laboratory studies of the resistance of organic materials to marine bacteria.

KENNETH G. VAN WYNEN, E.E., 1925, Cornell University; M.E.E., 1932, Polytechnic Institute of Brooklyn; American Telephone and Telegraph Co., 1925-34; Bell Telephone Laboratories, 1934—. He has been engaged in studies relating to improved telephone transmission and to air traffic control problems, and recently has been concerned with the development of a personal signaling system. His current assignment is responsibility for visitors programs at Murray Hill. Senior member I.R.E.; member American Association for the Advancement of Science.

