The Bell System Technical Journal

Vol. XXVII *January, 1948* *No. 1*

An Experimental Multichannel Pulse Code Modulation System of Toll Quality

By L. A. MEACHAM and E. PETERSON

Pulse Code Modulation offers attractive possibilities for multiplex telephony via such media as the microwave radio relay. The various problems involved in its use have been explored in terms of a 96-channel system designed to meet the transmission requirements commonly imposed upon commercial toll circuits. Twenty-four of the 96 channels have been fully equipped in an experimental model of the system. Coding and decoding devices are described, along with other circuit details. The coder is based upon a new electron beam tube, and is characterized by speed and simplicity as well as accuracy of coding. These qualities are matched in the decoder, which employs pulse excitation of a simple reactive network.

I. Introduction

IN THE development of systems for transmitting telephonic speech, much effort has been directed toward minimizing the effects of noise picked up in the transmission medium. The system described in this paper represents one method which has been successful in eliminating completely such effects under appropriate and practical conditions. In this method, known as Pulse Code Modulation[1,2,3] (PCM), telephone waves are represented by sequences of on-off constant-amplitude pulses.

Perfect reception of such pulses demands simply recognition of whether any pulse exists or not. Recognition can be carried out effectively in the presence of noise and interference amounting to a substantial fraction of the pulse amplitude. In contrast, telephonic waves carry information by subtle amplitude variations in the course of time. High quality telephone reception accordingly demands a much lower ratio of noise and interference—lower by as much as 50 decibels.

The magnitude of this figure exhibits one good reason for exploring the possibilities of PCM. Another potent reason, which is of particular importance in multi-link transmission, is that, with pulses involving just two standard values, regeneration can be used at repeater points and at the receiver to wipe out transmission impairments. The regeneration process consists of the pro-

[1] A. H. Reeves, U. S. Patent 2,272,070.

[2] "Telephony by Pulse Code Modulation", W. M. Goodall, *Bell System Technical Journal*, July, 1947.

[3] "Pulse Code Modulation", H. S. Black and J. O. Edson; presented June 11, 1947 at the Montreal Summer Meeting of the American Institute of Electrical Engineers, and to be published in the *A. I. E. E. Transactions*.

duction of a properly formed standard pulse, free of noise, to correspond with each received pulse, even though the latter may be considerably misformed. The sole proviso here is that before regeneration the level of noise and distortion in each link be kept below the comparatively large threshold value at which a mark cannot be distinguished from a space. If this holds good throughout the transmission path then literally the received pulses can be made as good as new. In contrast it is impossible fully to repair or to regenerate signals not involving standard values of amplitude and of time. With such signals distortion and noise in each span contribute to the total which therefore increases with the system length.

To sum up, conversion of speech to a code of pulses and spaces permits telephony to assume certain new and desirable properties; ability to work with small signal-to-noise ratios, and ability to regenerate any number of times with no degradation of quality. These advantages do not accrue without certain penalties. Conversion of speech waves to pulse form and back to speech involves a certain degree of apparatus complexity at the terminals. This complexity is not decreased by the need to handle pulses at high speeds, of the order of a million per second. Here radar and television circuit techniques are helpful. Another characteristic is that a greater band width is occupied in the transmission medium. This arises through the operation of two factors, of which one is the use of double sideband in pulse transmission (as against single sideband in carrier telephony), and the second involves the number of pulses used in the code. The relatively wide band required can best be accommodated in the microwave region and it happens that the properties of on-off pulse transmission can be used there to particular advantage.

The PCM system to be described was set up to evaluate experimentally the problems involved in providing multichannel facilities of toll system quality. It was designed to accommodate 96 channels. For experimental studies of such things as crosstalk and methods of multiplexing channels, a fraction of the total number of channels is sufficient and only 24 of the 96 were built. These are arranged as two groups of 12 channels each. The channels of a group are assembled on a time division basis. Assembly of the groups is carried out on a frequency division basis, each group amplitude-modulating its own carrier. In a planned alternative arrangement the group pulses may be narrowed and interlaced to put all 96 channels in time division on a single carrier, but this alternative will not be explored here.

The assignment of 12 channels per group fits in well with the present arrangement of carrier telephone circuits used in the Bell System plant, such as Types *J*, *K*, and *L*.[4] Use of time division for a group of this size involves pulses with

[4] "A Twelve-Channel Carrier Telephone System for Open-Wire Lines," B. W. Kendall and H. A. Affel, *Bell System Technical Journal*, January, 1939. "Coaxial Systems in the United States," M. E. Strieby, *Signals*, January–February, 1947.

repetition rates up to 672 kilocycles, and pulse durations as short as a quarter microsecond. These pulses, obtainable from more or less standard types of vacuum tube circuits, can be distributed from point to point in the equipment without too much difficulty. Amplitude modulators and demodulators at two neighboring carriers—65 and 66.5 megacycles—then serve to bring the PCM signals into the intermediate frequency range for transmission to and from the microwave equipment.

The speech quality of the overall system in respect to such factors as band width, volume range, noise, distortion, and crosstalk more than meets the requirements generally imposed upon such systems.

Figure 1 is a front view photograph of the experimental apparatus setup with covers removed from one bay to show typical construction. The two end bays contain intermediate frequency modulators and demodulators required for the two groups. In addition voice frequency terminating sets and jacks are mounted here, together with testing equipment. The center bay and the one to the right of it are identical; each includes apparatus for handling a group of twelve message channels. Transmitting equipment is mounted in the upper half, and receiving equipment in the lower half of each bay. The remaining bay, second from the left, holds all the timing equipment needed to furnish control pulses for operating eight of the message bays, a total of 96 channels. Included are circuits for synchronizing the receiver. Individual regulated power supplies are mounted near their loads on the several bays.

Figure 2 is a rear view of the same equipment. Cables in the four horizontal ducts shown carry control pulses from the timing bay to the 12-channel group bays. These ducts are large enough in cross-section to handle all the cables required for a complete 96-channel terminal.

II. Functional Problems Involved

The broad problems brought together in building this system may be considered under the four classes following:

1. The pulse code modulation problem; to convert signal waves to pulse patterns.
2. The multiplex problem; to aggregate channels into groups and groups into a supergroup.
3. The transmission problem; to fit the system into the minimum required band width, and to remove the effects of transmission impairments.
4. The pulse code demodulation problem; to convert pulse patterns back to the original signal waves.

These are to be discussed from a functional standpoint to provide background for discussion of the specific equipment.

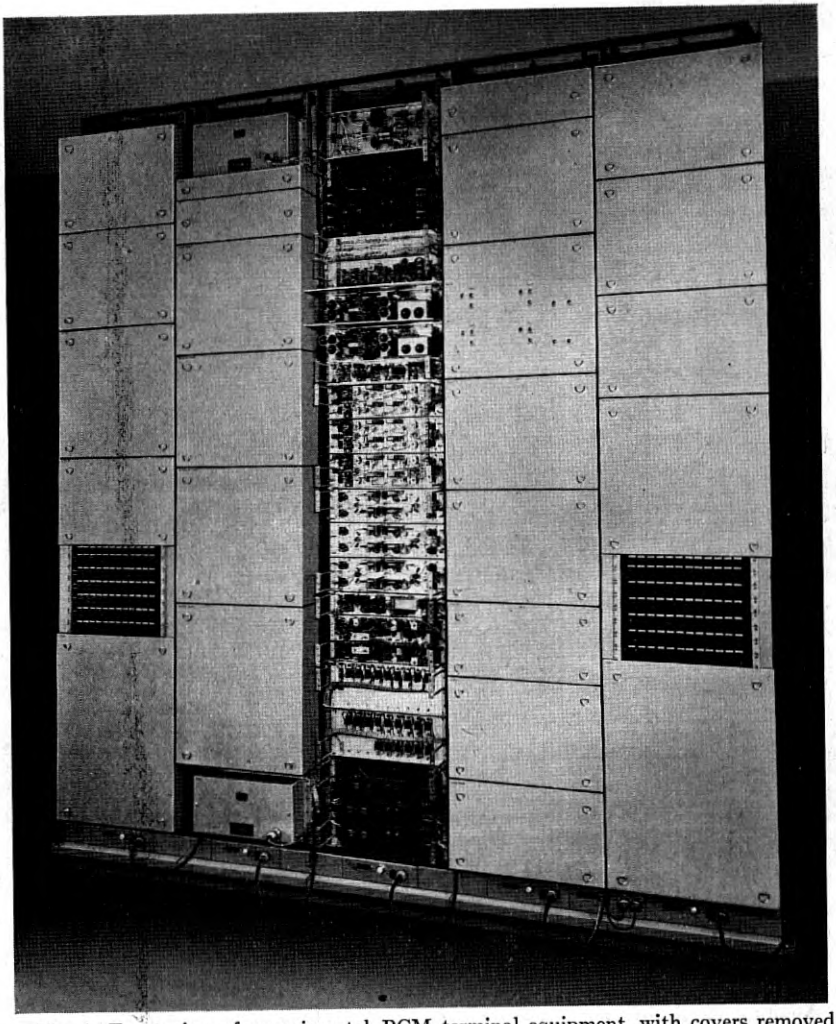


Fig. 1. Front view of experimental PCM terminal equipment, with covers removed from a 12-channel group bay.

Pulse Code Modulation

Sampling. A basic premise of pulse modulation systems is that the information content of a wave can be conveyed by samples taken at sufficiently frequent, equally spaced time intervals. The interval should be no greater than half the period of the highest frequency speech component to be reproduced or, otherwise expressed, the sampling rate should be not less than twice the frequency of the highest speech component present. This provision insures

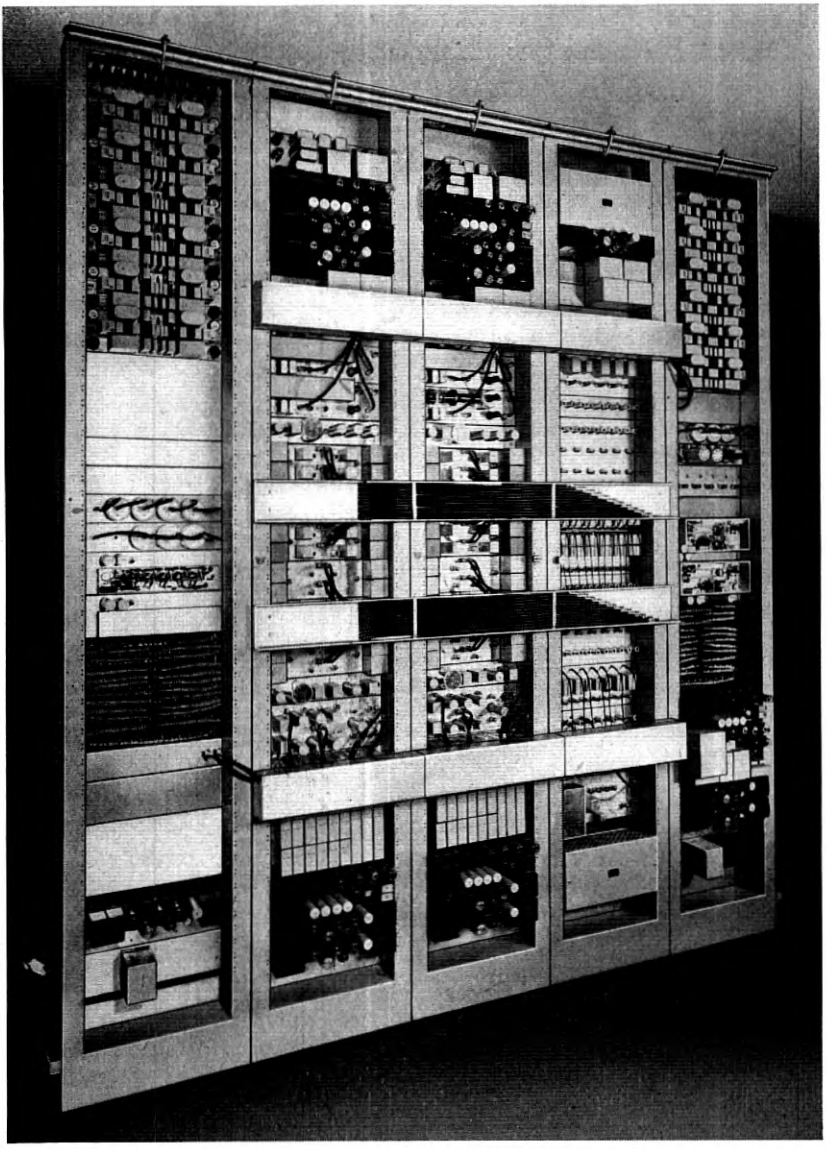


Fig. 2. Rear view. Cables in the horizontal ducts carry timing pulses leftward from the synchronizing bay.

that sidebands produced by sampling do not overlap to introduce distortion, as illustrated in Fig. 3. For a speech band extending up to 3400 cycles, a reasonable sampling rate is eight kilocycles.

Samples may be intermittently-transmitted portions of the signal wave, of appreciable duration, or they may be essentially instantaneous amplitudes.

To afford time for coding, the instantaneous samples may be maintained at constant value for an appropriate interval—a process here referred to as "holding."

Quantization. The fundamental operation of PCM is the conversion of a signal sample into a code combination of on-off pulses. In any practical system a continuous range of signal values cannot be reproduced since only a finite number of combinations can be made available. Each combination stands for a specific value, of course, so that we wind up by representing a continuous range of amplitudes by a finite number of discrete steps. This process is spoken of as quantization, a quantum being the difference between two adjacent discrete values. Graphically this means that a straight line representing the relation between input and output samples in a linear continuous

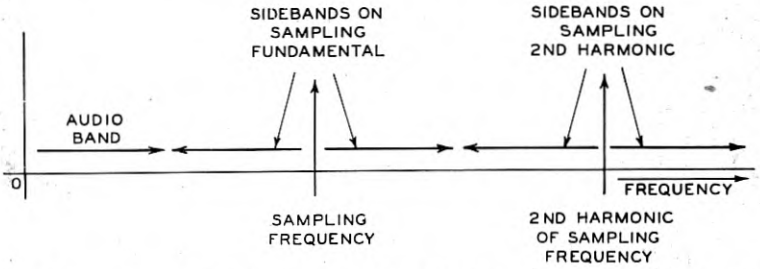


Fig. 3. Spectrum of a sampled audio band, illustrating separation of components when the sampling frequency is at least twice the top audio frequency.

system is here replaced by a flight of steps as in Fig. 4a. The midpoints of the treads fall on the straight line, and the height of the step is the quantum.

Manifestly the greatest error inherent in quantization amounts to half a step. Hence the quality of reproduction may be measured by the size of that interval, which depends upon the total number of steps in the amplitude range covered. With n pulses assigned to represent an amplitude range, the maximum number of discrete steps is 2^n, and the size of each step is proportional to 2^{-n} times the amplitude range.

This error shows up as a noiselike form of distortion, affecting background noise in the absence of speech, and accompanying speech as well. The distortion actually consists of a multiplicity of harmonics and high order modulation products between signal components and the sampling frequency scattered fairly evenly over the audio spectrum. If the audio signal is a simple sine wave, these many products may be identified individually; but for speech or other complex signals they merge into an essentially flat band of noise that sounds much like thermal noise. Since the level of this distortion is fixed by the quantum size, an adequate number of steps must be provided for the lowest

amplitude sounds it is necessary to transmit. Considering that consonant levels may be of the order of 30 decibels below vowels, and that weak talkers may be of the order of 30 decibels down from loud talkers, it is clear that amplitudes as far below maximum as 60 decibels require at least a few steps.

Ordinarily this would involve a large number of pulses for transmission, with a consequent increased complexity of terminal apparatus, and an in-

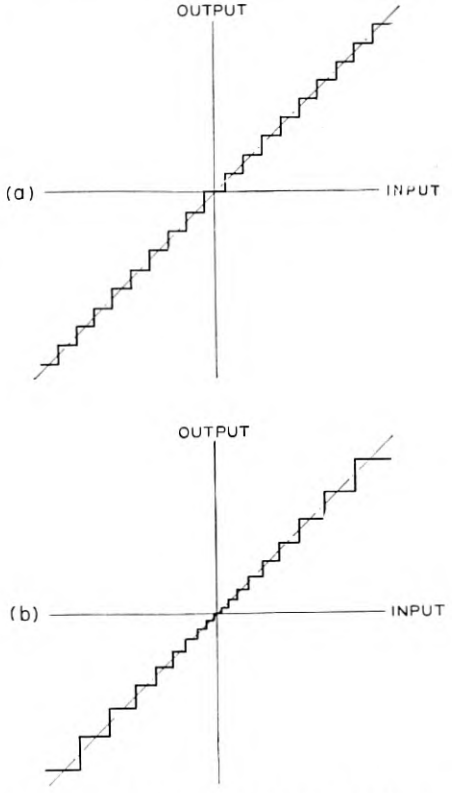


Fig. 4. Relation between input and quantized output, with quantization uniform in (a) and tapered in (b).

creased frequency band per channel. A more practical solution is to employ tapered steps rather than uniform ones. In this way a given number of steps can be assigned in greater proportion to the low amplitudes than to the highs, as shown in Fig. 4b. There results a degree of step subdivision sufficient to care adequately for the low-amplitude sounds including background noise. A small penalty is paid at the upper end of the amplitude scale because of the proportionately smaller number of steps available there, bringing in higher

quantizing noise in this range. How large the effect may be is determined by the degree of taper. To apportion the noise at different levels, the distribution of steps over the amplitude range can be varied.

Preliminary studies of quantization, involving listening tests and noise measurements for various numbers of digits and various kinds of taper, led to the choice of a seven-digit code (128 steps) for the present system. The taper employed reduces the smallest steps 26 decibels below the average size and increases the largest ones about 6 decibels.

Coding. The coder is required to set up a pulse code combination for each quantized signal value. A great many codes are conceivable, but in practice a simple one in which the pulses correspond to digits of the binary number system allows greatest simplicity at the receiver.

While coders may take a wide variety of forms, they can be arranged in three categories according to the way in which they evaluate speech amplitudes. In the first category an amplitude is measured by counting out, with a binary counter for example, the number of units contained in it one by one until the residue amounts to less than a unit. In the second, the amplitude is measured by comparison with one digit value after another, proceeding from the most significant digit to the least, and subtracting the digit value in question each time that value is found to be smaller than the amplitude (or its residue from the previous subtraction). In the third, amplitude is measured in toto by comparison with a set of scaled values. Modulators disclosed by Reeves[1] and by Black and Edson[3] are of the first category. That described by Goodall[2] belongs to the second, and the one described in the present paper is in the third category. Generally speaking the number of operations and the time required for coding decrease in going from the first to the third. Rapid coding is obviously desirable since it allows more channels to be handled in time division by common equipment.

Multiplex

Channels may be multiplexed by arranging them in time sequence, or by arranging them along the frequency scale. These methods are known as time division and as frequency division, respectively. The first is accomplished by gating, or switching, at precisely fixed times. One way of doing this impresses a more or less rectangular pulse on one of the grids of a gate tube, so that a signal wave may be transmitted during the gating pulse. The second method here refers to the use of amplitude modulators, each supplied with an appropriate carrier, which translate the signals to their assigned positions on the frequency scale. To avoid crosstalk in a time-division system, operations in group equipment common to a number of channels must proceed without memory of the amplitudes of preceding channels, requiring build up and decay times to be held within limits. This implies a sufficiently wide pass band

together with phase linearity. For the same purpose in a frequency-division system, filters are used to select and to combine channels. To avoid crosstalk the filters must be sufficiently selective, and amplitude non-linearity must be held within limits.

In the present system, the pulse code is delivered by the coder as on-off pulses in time sequence. It is therefore natural to organize the pulses of the different channels so that they appear in sequence, thus forming a time-division multiplex. Most types of coder require an appreciable length of time, after delivering the pulses of one channel, to prepare for coding the next. This preparation time may be afforded conveniently, without introducing gaps in the pulse train between assignments of consecutive channels, by providing two coders, which take turns at the channels in each time-division group.

As the number of channels increases, evidently the time interval which can be assigned to each channel must be reduced since all of them must be fitted into one period of an 8-kilocycle wave. Similarly the allowable duration of a code or digit pulse becomes shorter as the number of time-division channels in a group is increased. Then too, pulses tend to become more difficult to generate and transmit as their duration decreases. For these reasons it is desirable, and eventually it becomes necessary, to restrict the number of channels included within a time-division group. Frequency division may then be used to aggregate several time-division groups.

For our purposes groups of 12 channels are multiplexed by time division. With seven digits per channel, each group has a capacity of 672,000 pulses per second. To combine eight of the groups for a 96-channel system we again have a choice between frequency division and time division. The equipment of Fig. 1 is laid out to accommodate either procedure. In the first case each group is assigned a carrier for amplitude modulation, as used in actual tests to be described. For the second case the pulse durations would be cut down by a factor of eight, and the pulses from the different groups interlaced. Here the supergroup would have a capacity of 5,376,000 pulses per second.

In carrying out coding operations, and in multiplexing on a time-division basis, various control pulses are required which differ in repetition frequency, in time of occurrence, and in duration. These may be generated from a stable base frequency oscillator through the use of harmonics or, alternatively, of sub-harmonics. In a time-division system which requires a variety of flat-topped waves for switching operations, the use of sub-harmonics fits naturally. Frequency step-down circuits of the multivibrator type produce waves either approximating the desired forms directly, or requiring only simple circuits for reshaping. In contrast harmonic generation requires filters for component wave selection, more and more elaborate in structure as the order of the wanted harmonic goes up. Then, after selection, the harmonic has to be amplified and limited or otherwise shaped for switching purposes. Generally speaking

we need less apparatus and less space if we use multivibrator step-downs. Another advantage is that noise in the base frequency supply produces proportionately less phase jitter with sub-harmonics. While multivibrators do not have high inherent stability, they are capable of great precision when suitably controlled. For these reasons frequency step-downs are used to generate all the timing waves of the system.

Timing and gating operations require accurate time alignment of pulses, which may be accomplished by suitably delaying one set with respect to the other. For this purpose use is made of delay networks or cables, or delay multivibrators. Pulse durations may be varied through the use of interference effects between given pulses and their delayed replicas. Additional timing and gating wave forms are required for regeneration and for assembly in time-division multiplex. All such control pulses are economically generated at a single common point rather than by a number of local generators, and can then be supplied to the message equipment by common power amplifiers via shielded cable.

Transmission

In this section we are to consider the general factors entering into satisfactory reception of on-off pulses including such limitations as those on band width and noise. These are to be viewed while keeping in mind the procedures available for pulse regeneration.

If we start with a rectangular pulse like one of those generated at the transmitter, the corresponding frequency spectrum exhibits lobes extending indefinitely on the frequency scale, with progressively decreasing amplitudes as indicated in Fig. 5a. When such a pulse is passed through a linear phase filter[5] which discriminates against frequencies beyond the first lobe, the resulting pulse is practically of the sinusoidal form shown in Fig. 5b. Reducing the high-frequency response to the extent prescribed rounds the corners of the transmitted pulse, its duration at half value remaining equal to that of the original input pulse. In practice, transmission characteristics depart from phase linearity and low frequency cutoffs exist. Both effects introduce irregularities into the pulse response. While actual pulses therefore differ slightly in detail from the idealized picture given above, that picture will be retained for simplicity in discussion.

The band width needed for good pulse transmission can be minimized by making pulses as wide as possible. But since a specified number of pulses have to be put into a given time interval (84 pulses in $\frac{1}{8000}$ sec.) we must limit our broadening at a value where the presence or absence of a pulse may be clearly

[5] A suitable filter is one with a Gaussian characteristic, the loss in decibels varying as the square of frequency and having a value of 1 neper (8.68 decibels) at a frequency equal to 1/T.

determined in the presence of noise and interference. This spaces the sinusoidal pulses so that they overlap at their half-value points as illustrated in Fig. 5c. There it will be observed that no matter how many pulses occur in succession, the maximum amplitude of the pulse train is no different from that of a single pulse. The spectra of all such pulse trains have a common envelope, shown dashed in Fig. 5c.

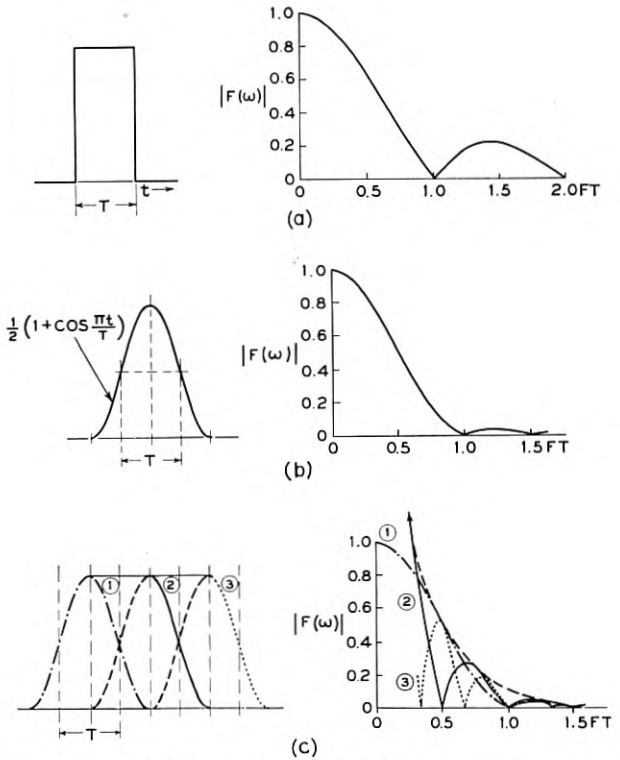


Fig. 5. Pulse forms and their associated amplitude spectra for: a, rectangular pulse; b, single lobe of a sinusoidally varying pulse; c, succession of pulses, each like that in (b).

Further, the band needed to transmit any sequence of such pulses is the same as that needed for a single pulse. This can readily be shown by using the familiar relation between transient build-up time and band width. Moreover, this conclusion is consistent with statistical analysis of all possible pulse combinations, which yields a spectrum of the form of Fig. 5b.

The relation between pulse duration and band width here described gives close to the optimum ratio of signal to noise and interference for the system considered. Narrowing the band would increase build-up and decay times, leading to reduced pulse amplitude and to increased pulse overlap. Thus,

although the narrower band would admit less extraneous noise, margins over noise and interference would be reduced. Widening the band, on the other hand, would allow pulses to build up and decay faster, but would not increase the pulse height. Thus the same signal would result, but the widened band would pass increased noise, and again the margins would be reduced. The optimum band width represents the most useful compromise in efforts to reduce noise and interference and to increase signal.

The filter characteristic we have been discussing is that of the entire link taking in all selectivity inserted between the practically rectangular pulses originally generated at the transmitter and the pulses delivered to the PCM receiver. Filters at both transmitting and receiving terminals of the link are included, the greater part of the overall selectivity being located at the receiver. With about 1.5 microsecond available per pulse at half amplitude, filters spaced 1.5 megacycle apart accommodate the double sideband and keep the interference between groups within tolerable limits.

To establish the presence of pulses we can set up an amplitude threshold equal to half the normal pulse amplitude, and test to see if that threshold is exceeded at a time near the center of the assigned pulse position. Selecting the threshold at half amplitude provides equal margin against the possibility of noise and interference bringing the full pulse amplitude, or mark, below threshold and bringing the nominal space above threshold. Testing at the pulse position midpoint maximizes this margin.

The amplitude threshold is set up by slicing a thin section horizontally out of the pulse at its half-amplitude level by means of an amplitude discriminator. Evidently, this procedure restores the flat top of the pulse. To complete regeneration by restoring the pulse epoch to a standard value the sliced pulses are gated at the midpoints of their proper intervals with narrow pulses supplied by the timing equipment. By these two pulse regenerating processes—slicing and gating—noise and interference are made impotent to produce errors until they attain a substantial fraction of the pulse amplitude. With the effects of noise and distortion thus eliminated, the only noise inherent in the system is that of quantization. In long systems having many repeater points, regeneration has to be practiced at spans short enough to permit cleaning out noise and distortion before it piles up above threshold. Thus in this system transmission impairments have to be considered only for the span between regeneration points; with their effects limited below threshold they are not carried over from one span to the next.

Where PCM groups are multiplexed by frequency division, amplitude nonlinearity of the system must be kept within limits. Otherwise intermodulation products may fall within transmission bands, adding to interference. Overlapping of neighboring filter bands also must be kept within bounds to reduce direct crosstalk between the pulse trains. With pulses arranged exclusively

in time-division multiplex, however, amplitude non-linearity of itself is not a factor. In this case the limitation comes on the departure from a suitable attenuation characteristic and from constant delay with respect to frequency. Distortion from these sources may increase the pulse duration so that excessive overlap of adjacent pulses will leave less margin available for interference and noise.

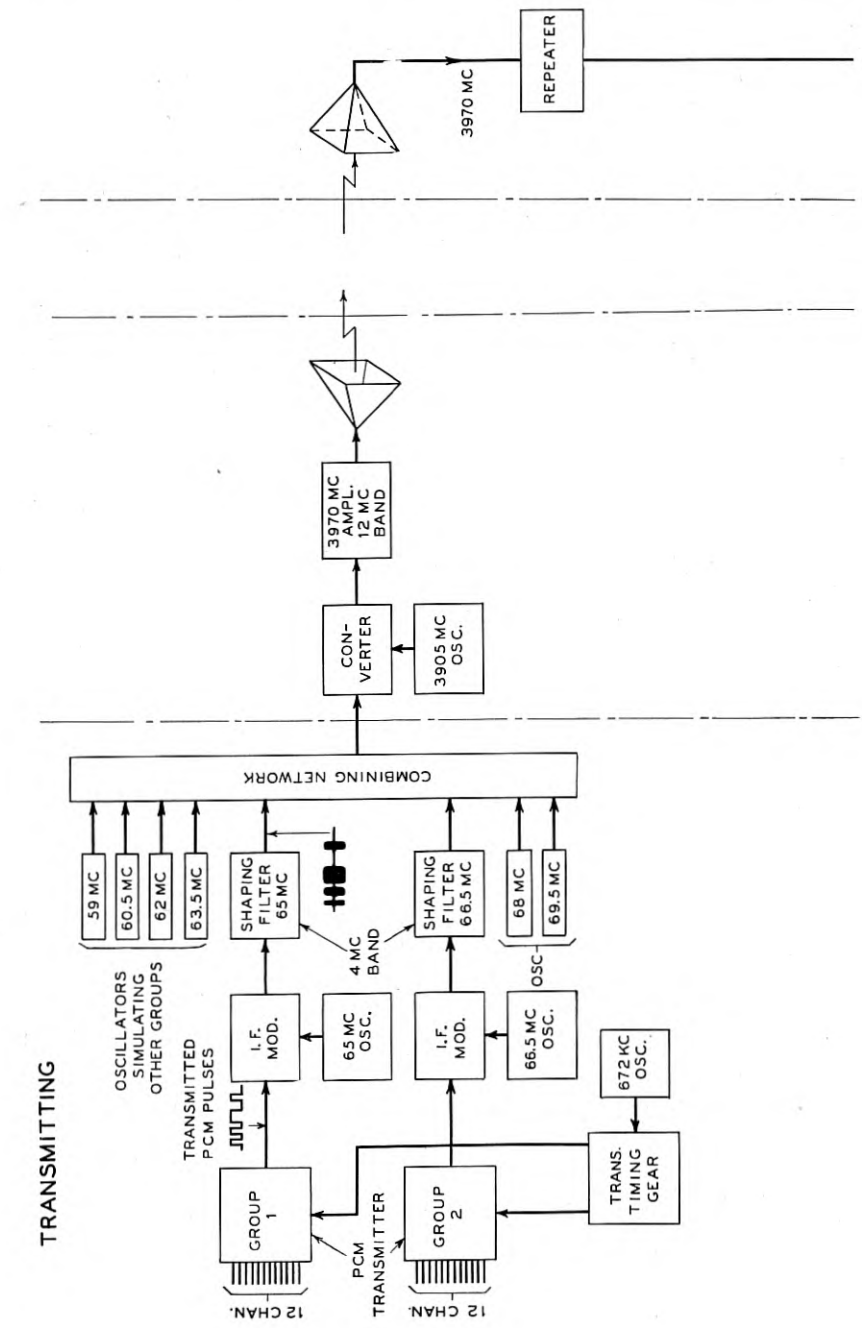
Pulse Code Demodulation

At the receiver, the regenerated code pulses are operated upon to recreate as closely as possible the original signal by operations complementary to those at the transmitter. Unfortunately, however, no process at the receiver can undo the effects of quantization which remain as noise, so that the quanta must obviously be made small enough from the beginning to permit satisfactory speech quality.

In each time-division group alternately working decoders may be used, in the same way as the two coders at the transmitter. Routing is effected by suitably timed gates. Conversion of a pulse code to amplitude may be accomplished by causing each pulse of a code combination to contribute an amplitude corresponding to the binary digit it represents, and then summing the contributions. When tapered steps are employed, due consideration must be given to overall linearity, discussed subsequently. The resulting output is a pulse-amplitude-modulated signal, which is then distributed to the channels of the group by an electronic commutator. Reconstruction of the signal from these distributed pulses is accomplished by simple filtering, which serves to remove components extraneous to speech introduced by the sampling procedure and tied up with the sampling rate.

Production of the necessary timing pulses at the receiving end proceeds in much the same manner as at the transmitter except for one thing. That is, instead of being initiated by a local oscillator, the receiver timing must be linked to the input pulses so that they may be properly routed. This involves the problem of synchronizing or, to borrow a term from television, framing. Use of this term is based upon the similarity of the sequence of PCM digits within a single sampling period to the complete ordered array of television picture elements. Preferably framing should be done with a minimum of time interval and of band width.

One method of synchronizing pulse systems employs a marker pulse which serves to initiate a timing sequence for each frame at the receiver. Here the marker pulse must differ sufficiently from the other pulses to permit its rapid and certain identification. This is ordinarily done by making the marker several times as long as any message pulse. In PCM, however, where digit pulses are run together in many code combinations, the marker would have to be very long to be clearly distinguishable, thereby cutting into channel ca-


TRANSMITTING
12 CHAN.
GROUP 1
PCM TRANSMITTER
12 CHAN.
GROUP 2
TRANS. TIMING GEAR
672 KC OSC.
TRANSMITTED PCM PULSES
OSCILLATORS SIMULATING OTHER GROUPS
I.F. MOD.
65 MC OSC.
I.F. MOD.
66.5 MC OSC.
OSC
4 MC BAND
SHAPING FILTER 65 MC
SHAPING FILTER 66.5 MC
59 MC
60.5 MC
62 MC
63.5 MC
68 MC
69.5 MC
COMBINING NETWORK
CON-VERTER
3905 MC OSC.
3970 MC AMPL. 12 MC BAND
3970 MC
REPEATER

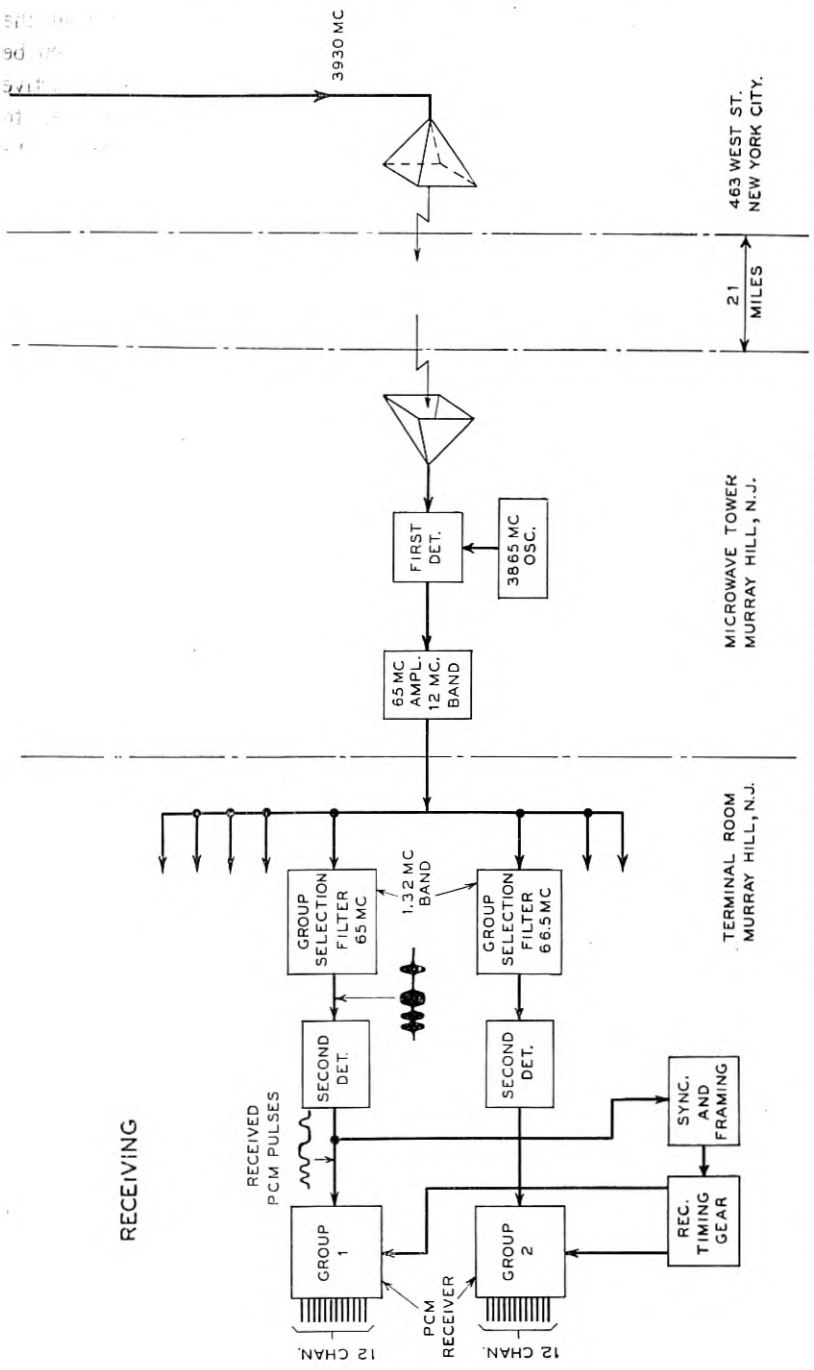


Fig. 6. Block diagram of the experimental PCM system.

pacity. This inefficiency can be avoided by deriving the pulsing rate from the pulse train through the use of a narrow band filter. Framing may then be effected by using in each frame just one digit pulse, which is given a distinctive repetition rate. With this method, a certain amount of time is required to establish synchronism when the system is started up. In the system to be described the framing time is less than one-tenth second—a tolerable value.

III. Experimental System

In the light of the foregoing discussion, the block diagram of the experimental system shown in Fig. 6 is believed to be largely self-explanatory. It will be noted that for microwave transmission the modulated intermediate-frequency signals are simply translated in frequency to the 4000-megacycle band. The shaping filters and group selection filters shown have approximately Gaussian characteristics, in accord with transmission considerations noted earlier. The band widths shown for these filters apply between points one neper down from the midband loss. Band widths given for the amplifiers, on the other hand, refer to their regions of essentially flat response.

Overall measurements are facilitated and the amount of experimental apparatus minimized by looping the radio path through a non-regenerative microwave repeater 21 miles away in New York. Both ends of the 24 channels are thus made available at Murray Hill, New Jersey, the location of the apparatus pictured in Fig. 1. With this arrangement, and using conventional 4-wire voice frequency terminating sets, 24 people are able to engage in 12 simultaneous conversations through the system.

PCM Transmitter. The transmitting equipment for an individual 12-channel group is shown schematically in Fig. 7. Each audio input[6] is passed through a 3400-cycle low-pass filter and through a limiter which chops off the positive and negative peaks of any signal exceeding a prescribed maximum amplitude. This limit is chosen to penalize the loudest talkers to the degree customary in toll system practice. The inputs then enter a "collector" circuit, which assembles samples of the channels in time division multiplex on a common lead. Although it functions electronically under control of pulses from the timing bay, the circuit so resembles a mechanical commutator that this analogy has been used in the schematic. The period of rotation of the "contact arm" is 125 microseconds (8 kilocycles), and a conducting path is formed to the common multiplex lead from each channel circuit in turn for $\frac{1}{12}$ of this period, or $10\frac{5}{12}$ microseconds. It should be

[6] In telephone terminology, these 4-wire inputs are normally at the −13 decibel level point; i.e., 13 decibels below the transmitting level at the toll test board. Strap connections are provided, as in the Western Electric A-2 channel bank, for adapting the system to inputs 3 decibels smaller. Similarly the final 4-wire outputs are delivered at the +4 (or +7) decibel point. The normal gain through a link is thus 17 decibels but can be set as much as 6 decibels greater.

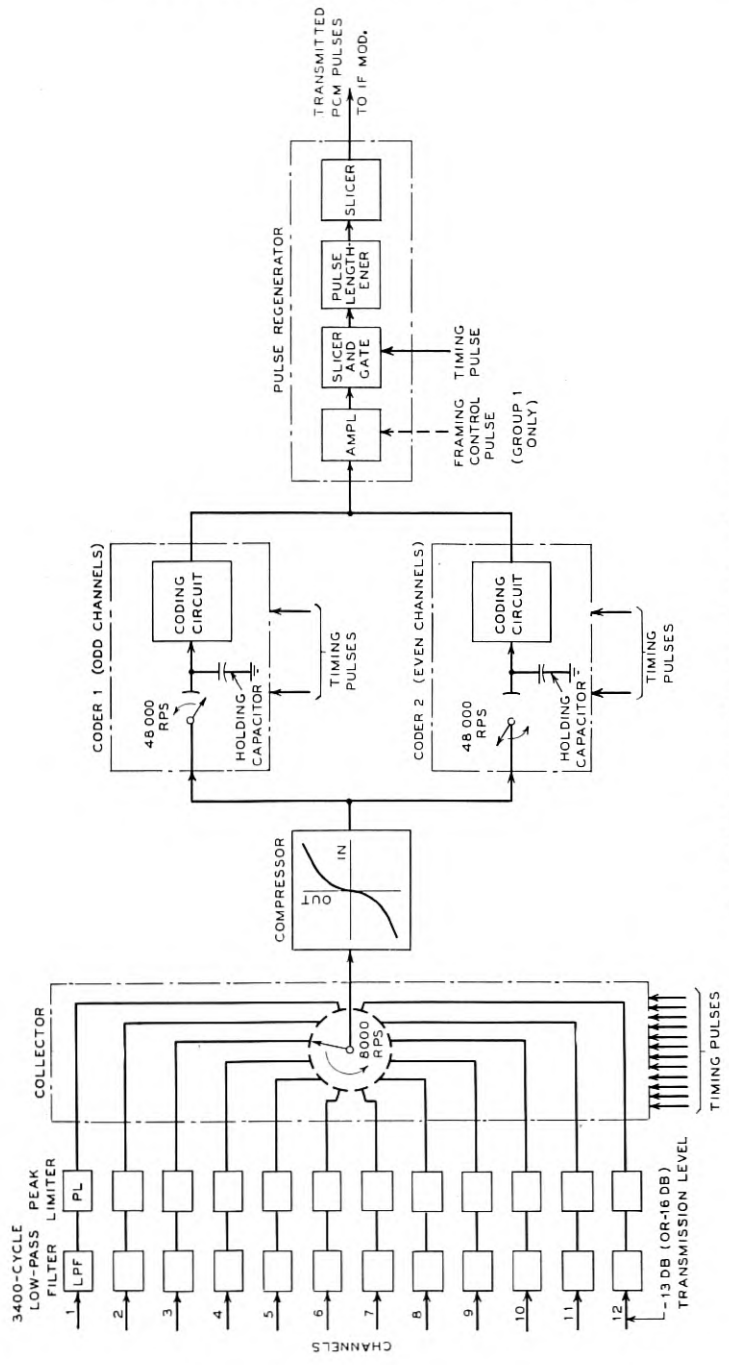


Fig. 7. PCM transmitter for a 12-channel group.

noted that the channel samples thus collected are not "held" at this stage; i.e., each sample does not remain constant in potential during its assigned interval, but rather changes to follow the wave form of the audio signal in the corresponding channel.

The multiplex signal is supplied to an instantaneous compressor, which employs silicon rectifier elements to give an input-output characteristic of the general form indicated within its block (Fig. 7). To understand the purpose of this device, we must recall the discussion of quantizing noise given earlier. There it was found desirable to provide a tapered distribution of step heights in the staircase-like quantizing characteristic, thus devoting a considerable number of small steps to the treatment of background noise and low-level signals. Although coders have been devised which inherently deal with signal amplitudes in this graded manner, it has been found more practicable in the present system to apply amplitude compression to the samples before coding, and to divide this compressed amplitude range into uniform steps in the coder. The result is a tapered step distribution with respect to the original uncompressed scale of amplitudes, details of the distribution being determined by the shape of the compression characteristic.

It may be well to note here that this method can be used in reverse at the receiver, with the decoding performed on an equal-step basis, and the resultant samples passed through a complementary instantaneous expandor. If the compression and expansion are truly complementary, the overall characteristic relating amplitudes of input and output samples will be linear except for the tapered array of quantizing steps (Fig. 4b).

Incidentally, no added band width in the transmission path of this system is required to accommodate the instantaneous compandor action.

After compression, the multiplex signal is delivered at low impedance to the inputs of two coders. In Fig. 7 the switch analogy is called upon again to illustrate the routing of alternate samples to the "odd" and "even" coders, and concurrently the storage of these samples on "holding capacitors" to keep them unchanged during the coding operation. Here the switches rotate at 48,000 revolutions per second, each one closing six times in a complete 8-kilocycle frame. The contact segment is drawn as a short arc to indicate a brief closure, actually lasting about 5 microseconds and occurring while the switch of the collector is in contact with a single segment. When the circuit is thus completed from a particular channel to the holding capacitor, the voltage on the latter very rapidly assumes and then follows, for the remainder of the 5-microsecond interval, the potential of the compressed version of the channel signal. When the circuit opens, the latest state of charge, which is essentially an "instantaneous" sample, is left on the capacitor, and is thus held for about 16 microseconds—until the next closure. These sampling operations occur alternately in the two coders.

By a process to be considered later, each coder produces a 7-digit PCM code representation of its set of samples. The two coders deliver their code groups alternately on a common output lead during the final $10\frac{5}{12}$ microseconds of each 16-microsecond holding interval mentioned above. Individual pulses last about 1.5 microsecond, although as delivered from the coders they are somewhat irregular in timing and waveform. It will be noted that the interval allotted to the code group from each channel is just

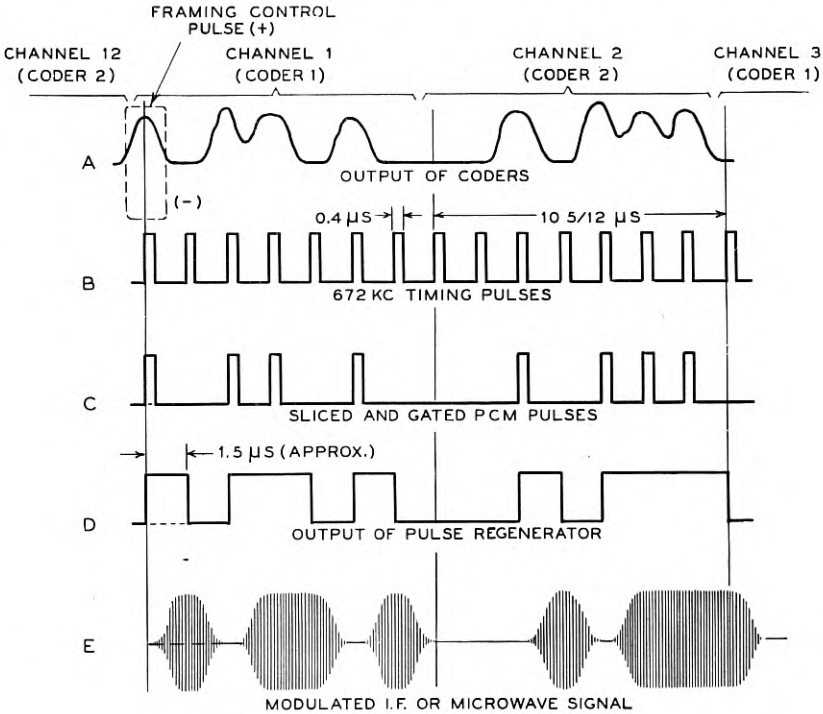


Fig. 8. Waveforms of the pulse regenerator.

$\frac{1}{12}$ of the 125-microsecond frame period and that a continuous train of code groups is thereby produced.

The common output circuit of the coders goes to a pulse regenerator which standardizes the pulses in height by slicing, and in time by gating, as illustrated in Fig. 8. The peaks of the coder output pulses (line A) are lined up in time with the gate control pulses (line B) supplied from the timing gear. The latter have a constant repetition frequency of 672 kilocycles and a uniform pulse length of 0.4 microsecond. Accordingly, the sliced and gated PCM pulses (line C) are also 0.4 microsecond long, and require lengthening to fill their allotted 1.5-microsecond intervals. This is accomplished by a

circuit in the pulse regenerator which first doubles the length of each pulse by adding thereto its own delayed reflection obtained from a short-circuited delay cable, and then slightly less than doubles it again by a similar process using a longer cable. A final slicing, to eliminate amplitude irregularities acquired in the lengthening process, yields square pulses as shown in line *D*, with adjacent pulses merged into a single longer pulse. This is the final output signal delivered to the intermediate-frequency modulator. Passage of the modulator output through a shaping filter results in rounded pulses (line *E*) suitable for transmission over the radio relay path.

In this regenerating apparatus, provision is also made for introducing a "framing control pulse," supplied from the timing bay and normally applied only to Group 1, although any other group may be used if desired. This pulse is about 1.5 microsecond long, and occurs once in each 8-kilocycle frame, but has opposite polarity in successive frames. It is timed to synchronize with the first digit of the Channel 1 code and is large enough in amplitude to override the pulse or space put out by the coder in that position. Hence in the final PCM output from Group 1, pulse 1 of Channel 1 is alternately present and absent regardless of the audio signal. This arrangement, used in automatic "framing" of the receiving timing gear as described in the following section, thus borrows the least significant digit from one channel of the system, leaving that channel usable, but with 6-digit instead of 7-digit quality. A 4-kilocycle tone of very low amplitude which it introduces in that single channel is made inaudible by the low-pass filter in the final audio output.

Synchronization. The connection of a transmitting channel to its proper receiving circuit in the time-division part of the system requires the two terminals to be synchronized: timing operations at the receiver must follow closely those at the transmitter. In a broad and general way this timing matter amounts to getting a local clock to keep the same time as a distant standard clock. Here the criterion of good timekeeping might be thought fussy by some standards; we cannot work with a discrepancy as long as a microsecond for the very good reason that incorrect routing of pulses would then result, associated with intolerably large decoding errors. Three provisions are made to take care of this situation. First, the framing is automatically monitored at all times. Second, if the system is out of frame—as it may be after transmission has been temporarily interrupted—the monitor circuit hunts for and establishes synchronism. Third, whenever the system is not properly synchronized and framed, all message circuits are cut off to avoid resulting noise and crosstalk.

For the purpose of this description we can use a mechanical analogy once more and picture all the transmitting channels of a time-division group arranged in order around a circle (Fig. 9). This time, however, we let each

small division of the circle constitute a commutator bar which is activated by the information of a particular code digit. Thus as the brush is stepped around at a uniform rate it puts digit information on the line in proper sequence. In each complete revolution there appears a set of on-off pulses constituting one frame. With twelve channels in the group, and seven digits to each channel, one revolution of the brush arm covers the frame of 84 digit positions. The revolution period is 125 microseconds. Roughly one and a half microsecond is allotted to each digit, and the brush steps 672,000 times a second. The driving force for stepping is supplied by a

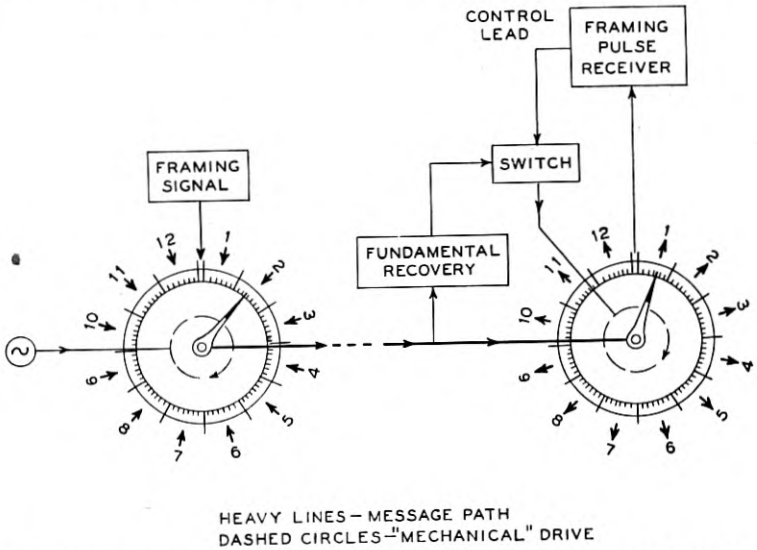


Fig. 9. Representation of the method used for framing, in relation to a single 12-channel time-division group.

stable oscillator, its frequency determined by a quartz crystal. One of the digit positions is taken away from its message duties and assigned for framing purposes, as described in the last section.

Our clock now is a one-handed affair with the brush arm for the hand. Its scale has the customary twelve main divisions, but each main division is divided into seven parts, one for each digit. Hour and minute markings refer to channel and digit, respectively.

At the receiving end we imagine a similar clock structure to be provided. Input pulses go directly to the clock hand which is stepped around to distribute pulses to the twelve message channels, plus a framing channel which takes up but a single digit space. Here the clock has no independent driving source, since it must follow the transmitter. Accordingly, the basic 672-

kilocycle frequency is recovered from the input pulse train by means which include filtering in a narrow-band crystal filter. The filter band is narrow enough to select the base frequency with negligible modulation and noise. That is, sidebands produced by pulse keying are attenuated to negligible proportions, as is the background noise. It is this output which serves to step the brush arm around the frame at precisely the transmitter frequency.

With the transmitter frequency recovered, both hands are stepped around their dials at the same rate. While this is a necessary condition for running the receiver, it is not a sufficient one since most likely the system will not be framed. In fact the odds are 83 to 1 against the two clocks indicating the same time if connection to the receiver is initiated at random times. If we had to deal with ordinary clocks, both in view, the resetting procedure could be accomplished by moving one clock-hand to agree with the other at one fell swoop. But in the PCM case resetting has to be done more discretely since only one dial position per frame is viewed in the framing receiver. Accordingly, an orderly procedure is set up for locating the framing pulse which consists in examining digit positions one by one until the framing pulse is reached. After any one position is viewed long enough to establish the absence of the framing pulse, the receiving clock is set back one digit position and the next position viewed.

This resetting or framing procedure is governed by the framing receiver through its control of a switch which connects the recovered base frequency to the driving mechanism of the clock. If the channels are correctly routed, so that it is the framing pulse which is being viewed by the framing receiver, the switch is left closed, and the 672-kilocycle wave steps the clockhand around the dial without interruption. But if the system is not correctly framed the framing receiver does not get its distinctive pulses. In this case the switch is opened every little while for the duration of a single pulse interval, stopping the local receiver clock during that interval while the transmitter clock advances one digit position. In effect the receiving clockhand is set back precisely one digit interval with respect to the standard. Thus the next digit pulse is brought into the framing receiver. If again the monitored input turns out to be other than the framing pulse, the stopping process is brought into play once more; this process is repeated until the system is framed.

As pointed out in the last section, the framing pulse is alternately absent and present in successive frames, corresponding to a 4-kilocycle rate. This is readily distinguishable from any of the message pulses, which in practice are found to have little energy content at this frequency. The framing receiver accordingly includes a resonant circuit tuned to four kilocycles. In the hunting process, eight frame periods are allowed between successive interruptions of the clock drive, to give the resonant circuit sufficient time

to build up above a threshold when the system is first framed. With this time interval thus fixed, each clock position is maintained for about one millisecond. Hence the time required to frame the receiver varies between one and eighty-three milliseconds, depending upon the epoch at which the system connection is established.

PCM Receiver. The received PCM signals of a 12-channel group are filtered from the frequency multiplex while in the intermediate-frequency state and then detected to the original signal band, as indicated in Fig. 6. They consist of rounded pulses, nominally sinusoidal in shape, but more or less distorted by transmission defects and accompanied by noise and interference.

These signals are supplied as input to the PCM receiver shown in Fig. 10. They are first sliced in amplitude, the slice being taken at approximately half the average or noise-free pulse height. Code groups of seven pulses are then routed alternately to two decoders, which handle even and odd channels respectively. The routing function is represented in the drawing by a two-segment commutator (A) rotating at 48,000 revolutions per second. Before entering the actual decoding circuit, the pulses are again sliced to secure very great uniformity, and are gated with 0.4-microsecond, 672-kilocycle pulses from the receiver timing equipment. Immediately after the arrival of the seventh digit of each code group of these standardized pulses, the decoding circuit produces a voltage on its low-impedance output lead proportional to the quantized amplitude represented by the code. As in the case of coding, details of the decoding process will be reserved for a later section of this paper.

The decoded amplitudes are available only momentarily; therefore it is desirable to sample each one at the proper time and store the result as a charge on a holding capacitor. This sampling process is represented by switch B in one decoder and B' in the other. Here the switch closures last only two microseconds, and values are held for about 19 microseconds.

The next step is to assemble the six samples from odd channels held successively by one capacitor and the six from even channels held by the other into a single time-division multiplex. Switch C performs this operation, rotating at 48,000 revolutions per second, and making contact alternately with the output circuits of the odd and even decoders.

This 12-channel multiplex signal is passed through an instantaneous expandor, the purpose of which has been noted in an earlier section. To simplify the problem of making the input-output characteristic of this circuit accurately complementary to that of the compressor, the two devices are designed to use identical silicon units. In the expandor, however, the nonlinear device is employed in the feedback path of a broadband amplifier rather than in the direct transmission path, thus giving the inverse character-

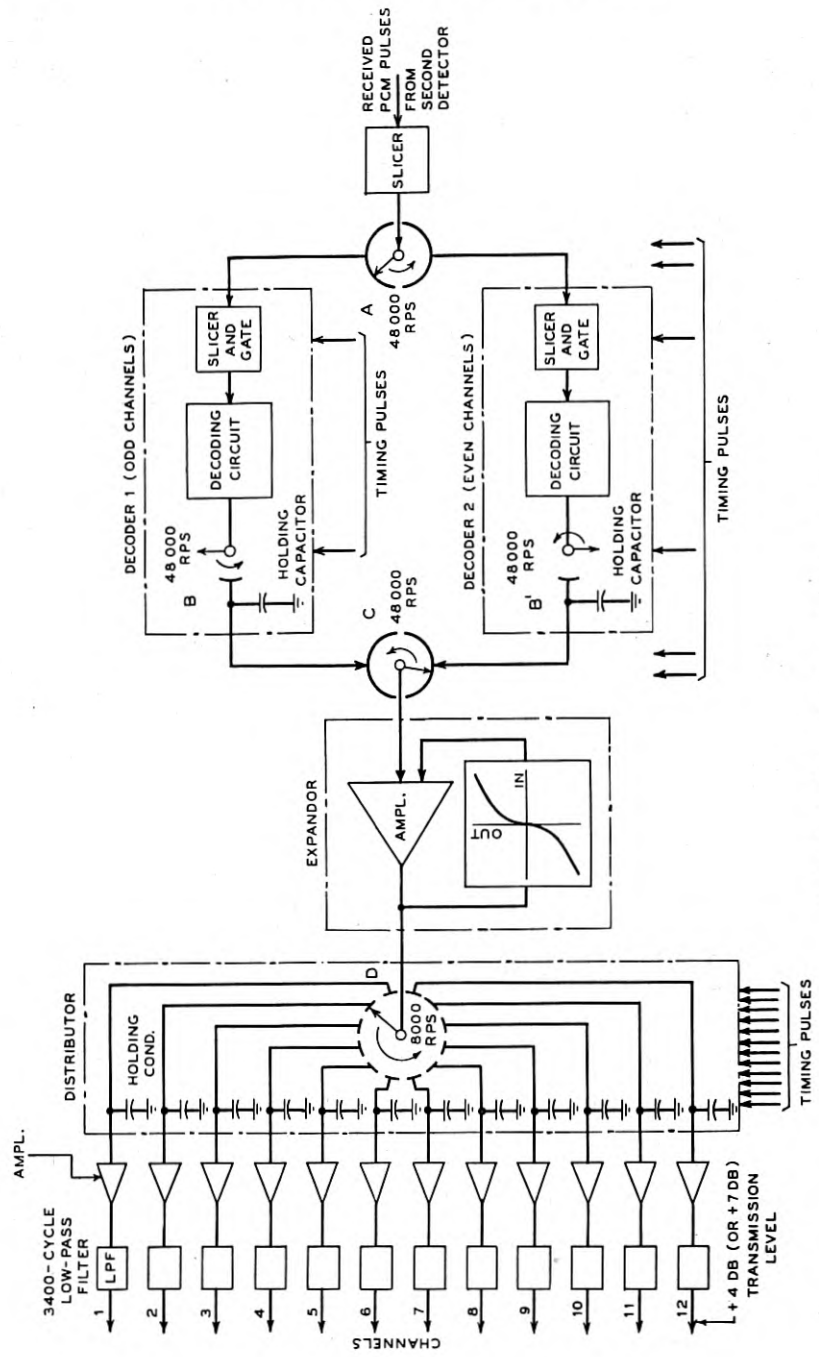


Fig. 10. PCM receiver for a 12-channel group.

istic. To allow any compressor to be used with any expandor, all the silicon elements are matched to a chosen standard unit, using selection and resistance "padding" methods. By such means, and by use of sufficient loop gain in the feedback amplifier, very satisfactory overall linearity of the system has been attained.

At the output of the expandor the waveform of the multiplex signal is essentially the same as that at the input of the compressor in the transmitting terminal. The samples are distributed to their respective channel destinations by a distributor D, resembling the collector described earlier. The rotation rate is 8000 revolutions per second, but the duration of contact on any one segment is only five microseconds instead of the possible full twelfth of the 125-microsecond frame period. This effective narrowing of the contact segments is done to allow the closure to occur well within the interval in which the circuit is completed by switch C from the output of a particular decoder. Each of the 12 segments of the distributor is provided with a holding capacitor, which stores its allotted samples for the full 125-microsecond frame period. The potential on any one of these capacitors thus changes at 125-microsecond intervals from one quantized sample amplitude to the next derived from the same original speech wave.

This potential is of sufficient magnitude to require use of only a simple single-stage triode amplifier for the output of each channel. Lengthening the samples by holding, as described, helps to make this possible by causing the amplifier to deliver useful power continuously instead of on a fractional time basis.

The only disadvantage of using lengthened pulses arises from an effect, very similar to the "aperture effect" encountered in sound movies, which introduces a curving slope across the audio gain-frequency characteristic of the system. In the present case the gain drops about three decibels as the frequency goes from the lowest to the highest value of interest. This slope can be corrected by a simple equalizing network, as shown in Fig. 11. In the present system the equalization is incorporated in the low-pass filter at the input of each audio channel. This is preferable to equalizing at the output, where power is at a premium.

The outputs from the channel amplifiers are passed through 3400-cycle low-pass filters, identical with the input filters except for omission of the equalization, and are delivered to standard voice-frequency circuits at the same levels[6] as are provided by a type J, K, or L carrier system.

IV. Component Circuits

Many of the circuit techniques used in the experimental system are conventional, others are more or less unfamiliar, and still others are believed to be novel. In the following some of the more important building blocks are

described. These include sampling circuits, the instantaneous compressor, slicers, and the PCM coder and decoder.

Sampling Circuits. The function of opening and closing a circuit at prescribed instants, represented by rotating switches in the block schematics, is actually performed[7] by one or the other of the devices shown in Fig. 12, employing diodes and triodes, respectively.

The diode type (Circuit a) is normally biased "open" by rectified charges stored on the two large capacitors. While in this condition it presents an extremely high series resistance between the low-impedance input and the load. But when a flat-topped pulse is impressed upon the pulse transformer, the aforesaid high impedance changes to a low value (of the order of 100

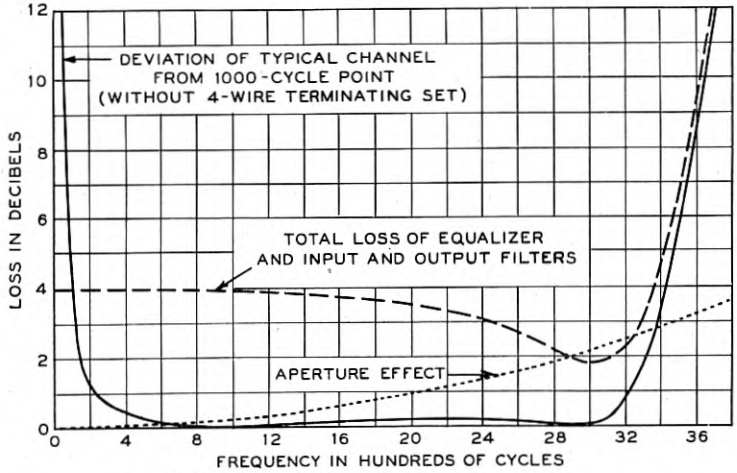


Fig. 11. Equalization for the aperture effect.

ohms), for the duration of the pulse. The upward-sloping tops of the pulses sketched in Fig. 12 represent predistortion to allow for transmission through the pulse transformer, for the purpose of obtaining rectangular pulses at the tube itself.

In the other variety (Circuit b) the plate-cathode paths of the two triodes are connected directly in parallel, conducting in opposite directions. The grids are both arranged to be biased below cut-off during the "open" condition of the sampler by grid rectification of pulses, and to be driven strongly positive by a pulse when a low-resistance conducting path is required between source and load.

These two types have their respective advantages. For example, the low capacitance to ground which the diode type affords across its load makes it

[7] A single exception is the case of switch A of Fig. 10. In this case two gated slicer circuits are used, as described subsequently.

preferable for use in the collector, where twelve samplers are multipled to a common load. On the other hand the triode type affords a d-c. path (without the series blocking capacitors that are required by circuit a) between source and load. In cases where a holding capacitor is to be charged to a succession of sample amplitudes which must be kept mutually independent to avoid crosstalk, the d-c. path is very desirable for it avoids "memory" effects associated with passage of the charging currents through the blocking capacitors. A further useful property of the triode circuit is its ability to

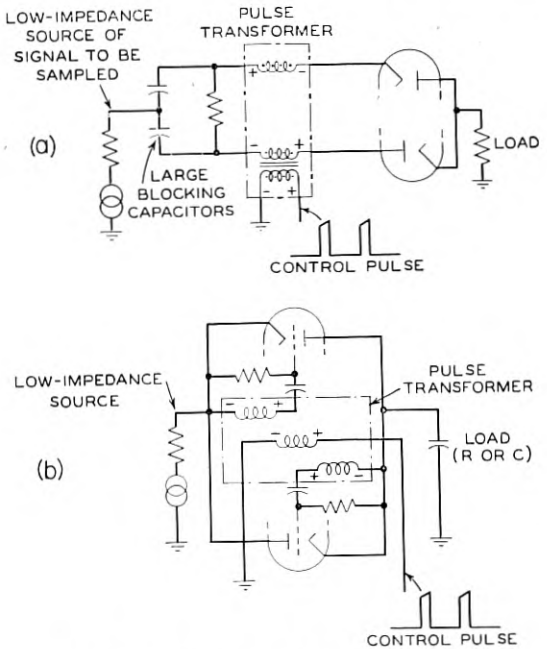


Fig. 12. Diode and triode sampling circuits.

sample signals of greater amplitude than that of the control pulses. With the diode circuit, the signal amplitude must be kept less than half the pulse height.

Instantaneous Compressor. The simple configuration of the non-linear circuit used in the compressor and in the feedback path of the expandor is shown in Fig. 13. Two selected silicon rectifiers are connected in parallel, but poled oppositely, with a small padding resistor (R_1 or R_2) in series with each. A parallel padding resistor (R_3) of large value is also provided. The direct-current resistance of this combination varies from about 6000 ohms at zero signal to about 190 ohms at the peak of a full-load signal. Input is applied as a current through the relatively high resistor R_4, and the voltage

drop across the varistor unit constitutes the compressed output. Although the development of this compressor has been a problem of many interesting aspects, it must suffice here to point out that copper-oxide elements are unsuitable at the speeds involved because of excessive capacitance, that silicon is superior to at least the presently available types of germanium

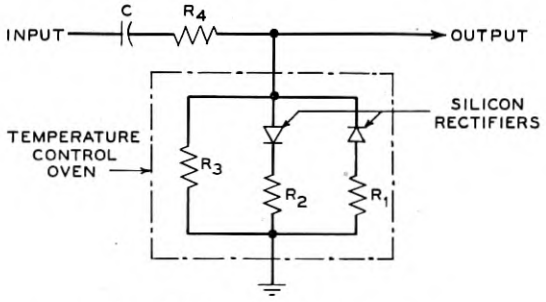


Fig. 13. Instantaneous compressor.

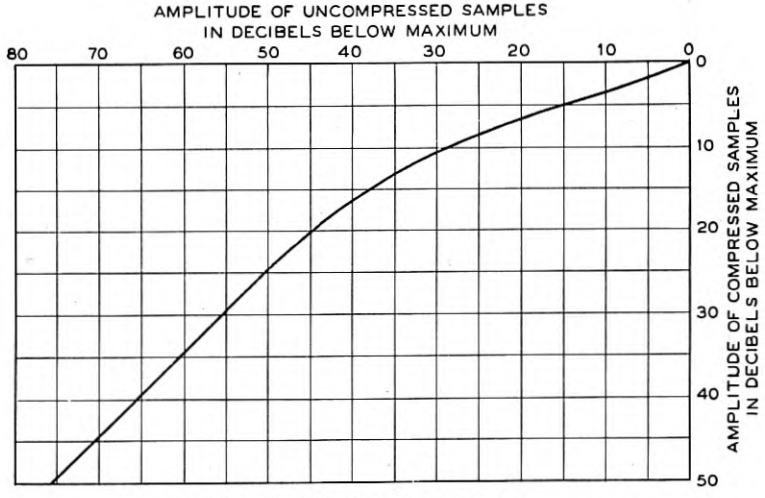


Fig. 14. Characteristic of the instantaneous compressor.

varistor in regard to freedom from certain memory or hysteresis effects believed to be thermal in origin, and that a satisfactory yield of matched silicon elements appears to be obtainable by the selection and resistance padding methods employed. Temperature control is used for the sake of constancy of the compression characteristic. Aging has been found negligible over a period of many months.

The actual compression characteristic afforded by these units is plotted

in Fig. 14 in terms of compressed signal amplitude vs. uncompressed signal amplitude, both in decibels.

Slicers. The slicer circuit shown in Fig. 15 resembles a conventional single-trip multivibrator in configuration, but functions somewhat differently because of the choice of parameters. In particular, capacitor C is made large enough so that the potential drop across it does not vary appreciably during normal operation, and plate resistor R_1 is given a small value such that the gain around the feedback path of the circuit is approximately unity when both triodes are in their active regions. Germanium varistors VR_1 and VR_2 maintain desired bias conditions regardless of the number of pulses or spaces in the input signal. Unlike the single-trip multivibrator, which when tripped remains so until the charge on C has had time to relax,

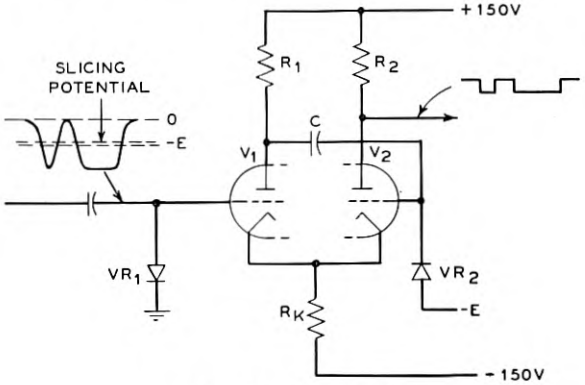


FIG. 15. Slicer circuit.

this circuit trips whenever the input signal falls through a narrow potential range near the value E, and trips back again when the input rises through the same potential range. A square wave of constant amplitude representing an accurate thin slice of the input is thus made available across resistor R_2. The circuit is capable of high speed, slicing pulses as narrow as 0.1 microsecond.

Time gating may readily be included among the functions of this circuit, through the addition of a triode having its plate and cathode connected directly to the plate and cathode, respectively, of tube V_1. The grid of this added triode is held normally at ground potential, and is pulsed in the negative direction by approximately rectangular gating pulses, having an amplitude of about $2E$. The total current passing through the common cathode resistor R_K is essentially constant, and if either or both of the paralleled tubes are conducting at a given instant this current is carried by one or shared by both of them. Hence the tripping action, involving transfer of

the cathode current to V_2, cannot occur as long as the grid of either V_1 or the added tube is positive with respect to the critical slicing potential. Tripping does occur whenever this limitation is removed, and thus the desired gating and slicing functions are performed concurrently.

In the experimental system this circuit is used to regenerate the PCM pulses at the common output of the two coders, and again at the input of each decoder to slice the received pulses and to sort out code groups of odd and even channels.

Coders. The method of binary coding used in this system was originally suggested by F. B. Llewellyn. It employs a novel electron beam tube. This device, pictured in Fig. 16, carries out simultaneously the two functions of quantizing and of coding. The tube is about $10\frac{1}{2}$ inches long and $2\frac{1}{4}$ inches in diameter. It has the 128 combinations of the 7-digit code laid out permanently as holes in an "aperture plate," and translates sample amplitudes from the form of beam deflections directly into PCM pulse symbols. Figure 17 shows one of these tubes in its socket on the rear of a coder panel, and above it a rectangular permalloy shield which covers the coding tube of the other coder serving the same 12-channel group.

As shown schematically in Fig. 18 the coder includes, in addition to the coding tube, a sampling and holding circuit which sorts out the odd (or even) channels from the input multiplex signal, push-pull amplifiers for vertical and horizontal beam deflections, and simple arrangements for blanking, focusing and centering. Within the tube are shown, from left to right, a conventional electron gun, vertical and horizontal deflection plates, a rectangular "collector" for secondary electrons, a "quantizing grid," the "aperture plate" and finally a "pulse plate." Figure 19 shows the target end of the tube, as viewed from a point near the gun. "Digit holes" in the aperture plate, laid out in accordance with the desired binary code, may be seen behind stretched parallel wires of the quantizing grid. One may count 64 narrow holes separated by equally narrow bars of metal in the left-hand vertical column, 32 holes in the next column, 16 in the next, and so on for seven columns. There are 129 grid wires uniformly spaced and accurately aligned so as to mask the upper and lower edges of every one of these holes when viewed from the geometric "point of origin" of the beam.

Stored audio samples from the sampling and holding circuit provide potential for the vertical deflection, with zero at the center, positive amplitudes in the upper half, and negative amplitudes in the lower half of the target area. A sawtooth sweep provides the horizontal deflection. The beam is blanked while deflection potentials are being changed to move it upward or downward from one sample amplitude to the next. When first restored, the beam strikes in the left-hand unperforated region of the aperture plate, and is then swept linearly across from left to right. Electrons which pass through the

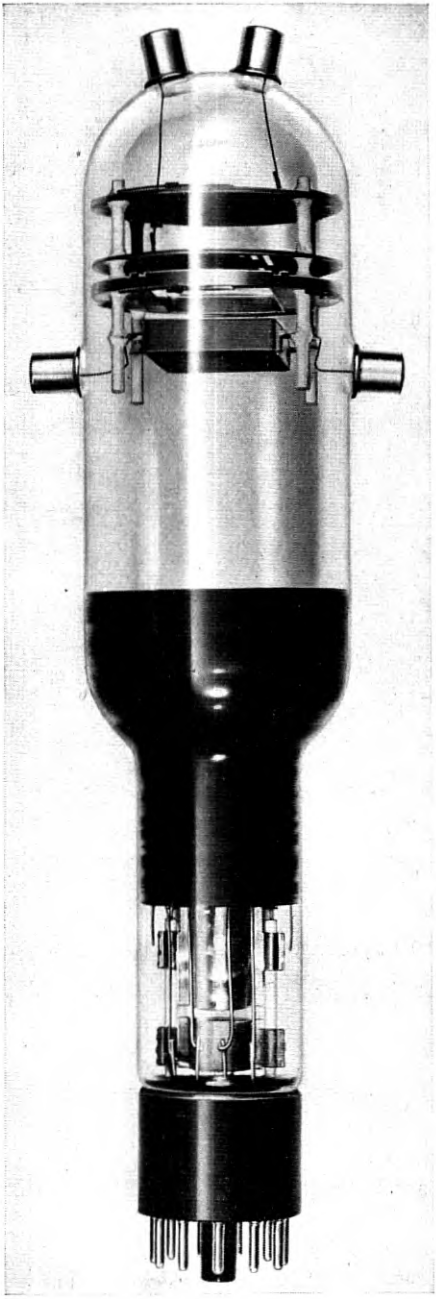


Fig. 16. Coding tube. This new electron device transforms speech samples into pulse codes.

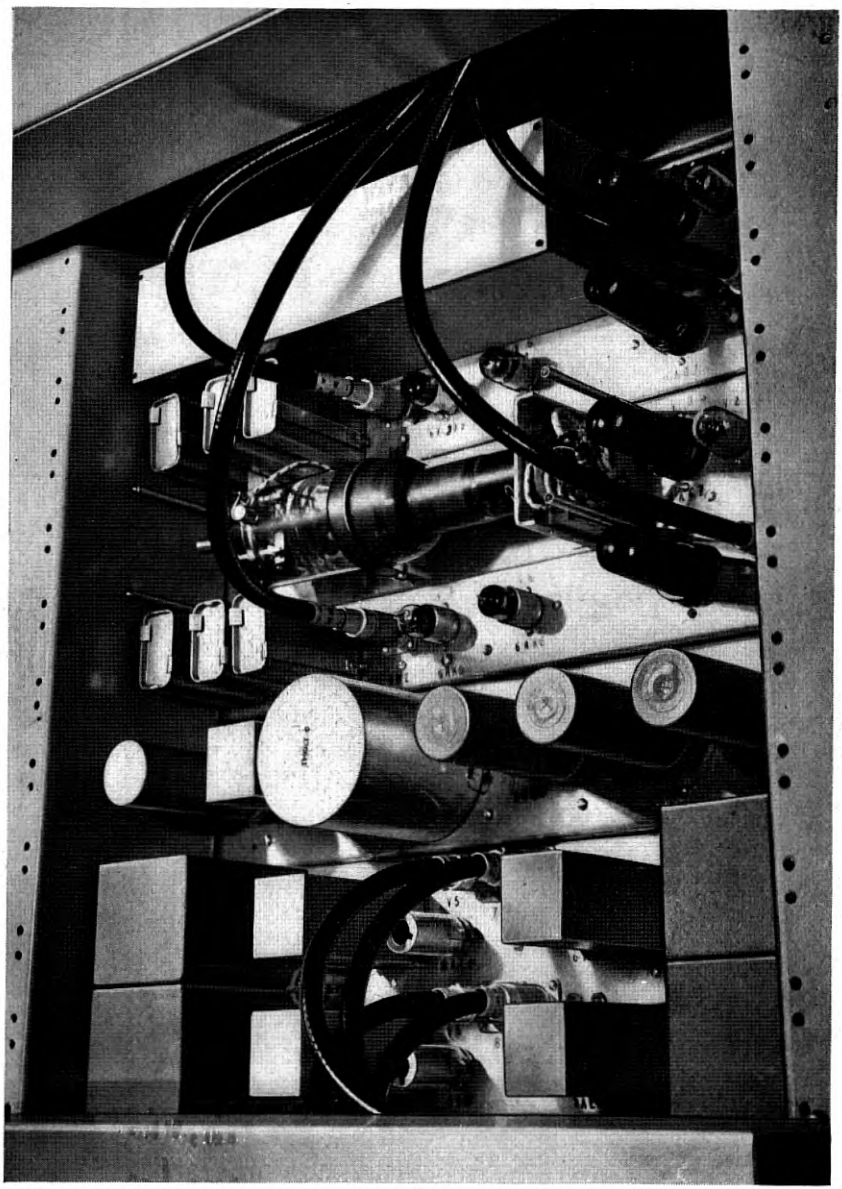


Fig. 17. Rear view of part of the collector (bottom panel), compressor, and two coders.

digit holes during the sweep are caught by the pulse plate, forming pulses which are amplified, gated and lengthened in the pulse regenerator to constitute the desired PCM signals. Retrace of the sweep occurs while the

beam is blanked, and is simultaneous with the application of the succeeding audio sample.

The wires of the quantizing grid are used to guide the beam so that it can illuminate only the particular row of apertures which correspond to the initial vertical deflection. Without this feature, erroneous codes would be produced when the beam straddled the edges of apertures or crossed from one amplitude level into another, as a result of electrode misalignment or possible slight drift in potential of an applied sample. The guiding action (basically proposed by W. A. Marrison and applied to the PCM coder at the suggestion of G. Hecht) is obtained by means of feedback from the quantizing grid to the vertical deflection amplifier.

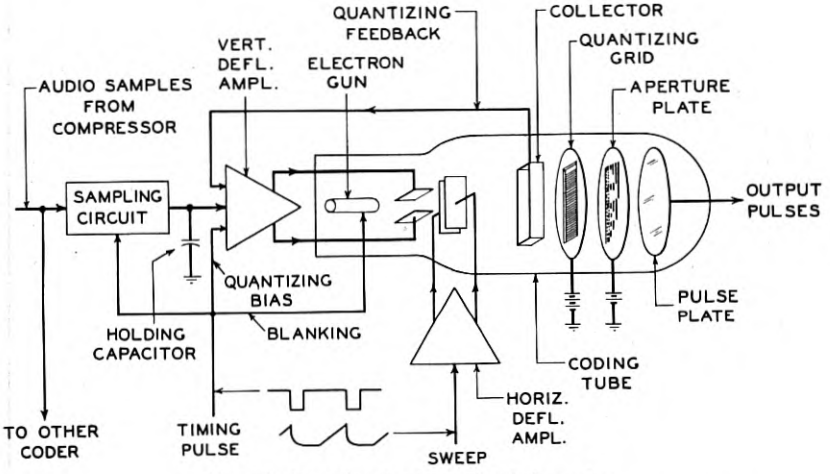


Fig. 18. Functional schematic of the coder.

The feedback signal is actually a current taken from the positively biased collector, which draws to it secondary electrons from the grid wires. The portion of the beam current striking the grid varies as a cyclic function of the vertical deflection. It follows that for some spot positions the value $\mu\beta$ in the feedback loop is positive, for others negative; hence with proper amplifier design there is a stable and an unstable region associated with each wire.

The spot can come to rest (vertically) only within one of the stable regions. In order to locate it consistently near the center of such a region, and thus gain equal margins against "hopping" upward or downward across a wire, a "quantizing bias" is introduced into the vertical deflection amplifier, along with the feedback and the signal samples. This bias is a current of opposite polarity from the unidirectional feedback current, and of magnitude equal

to the average between the two values of feedback current which exist when the beam falls (1) directly on a grid wire and (2) midway between two wires. One may regard this bias as pressing the beam upward against a wire, resisted by downward pressure associated with the feedback current. The latter

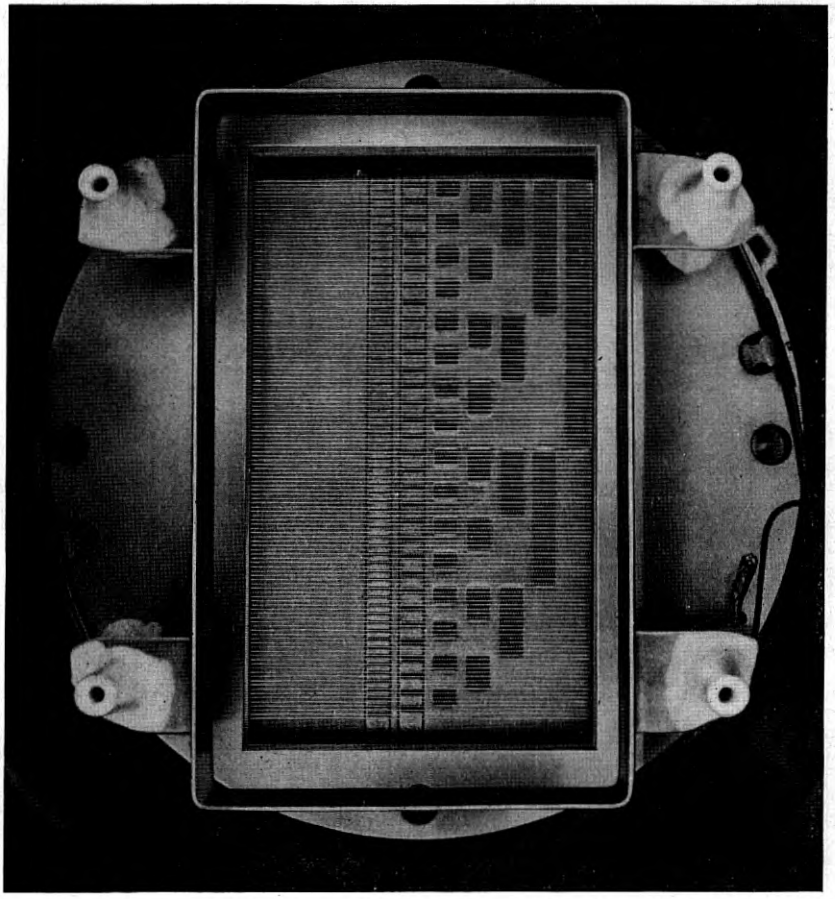


Fig. 19. Interior of the coder tube, viewed from the gun end.

increases as the beam approaches the wire, and equilibrium is reached when the beam is about half way between positions (1) and (2) mentioned above. The feedback current is actually intermittent, turned off and on by the blanking pulse, and careful analysis shows it necessary to make the bias intermittent also, with its wave fronts synchronous with those of the feedback current. This is readily accomplished by deriving the bias from the blanking signal itself.

With this arrangement the beam, suddenly turned on, moves either upward or downward from its initial "unquantized" position to the nearest position of stable equilibrium. Quantization is completed in less than a microsecond. Thereafter, as the beam is swept horizontally across the target area, it remains pressed upward against the lower surface of its guiding wire. Quantization is thus maintained until the end of the sweep, when blanking occurs. The margins against hopping across wires, while quanti-

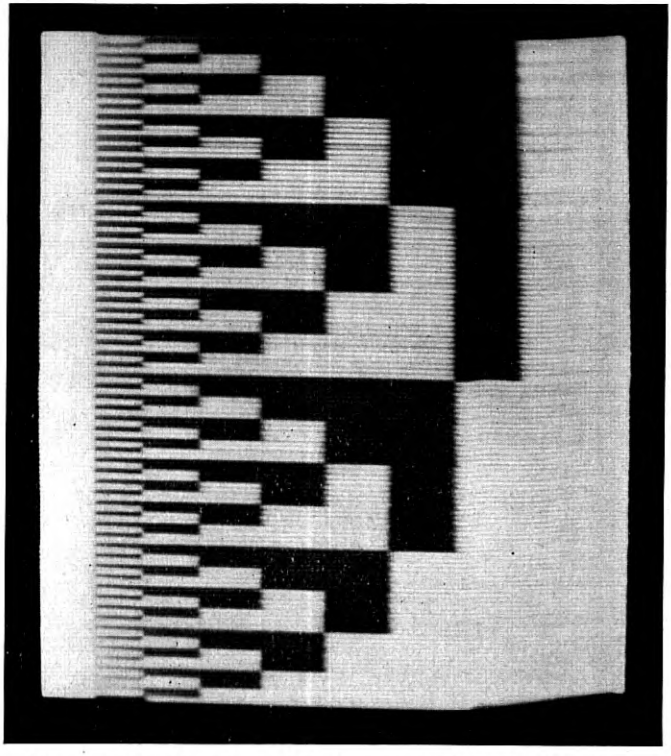


Fig. 20. "Television picture" of the aperture plate of a coder tube.

zation is in effect, are related to the amount of loop gain in the feedback path. In the present system feedback measuring about 20 decibels is provided, and will counteract changes which, without feedback, would move the beam two grid wires in either direction from the initial position.

This coding process, based on the electron beam coding tube and making use of a two-dimensional permanent layout of the code, is more straightforward than the various coding processes which depend upon counting or sequential comparison. Accordingly it leads to higher speeds and greater circuit simplicity.

Figure 20 illustrates the coding accuracy obtained. This photograph of

the screen of a test oscilloscope may be thought of as a television picture of the aperture plate of the coding-tube. To produce the pattern an audio-frequency sawtooth wave of full-load amplitude was applied to the sampling and holding circuit at the input of a coder. The resulting sample amplitudes, quantized by the coding tube and falling into all the possible 128 steps of the quantizing characteristic with uniform regularity, were used to energize the vertical deflection of the test oscilloscope. An ordinary synchro-

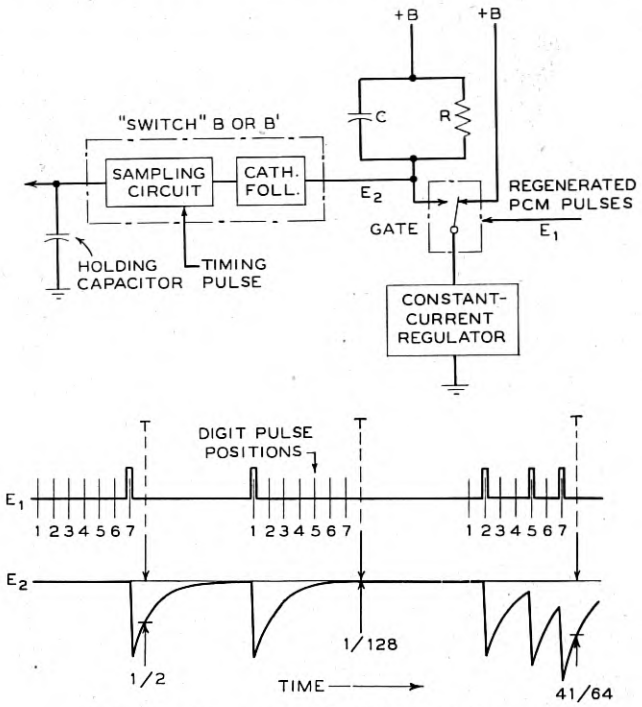


Fig. 21. Shannon decoding circuit and waveforms.

nized sweep provided the horizontal deflection. The square PCM pulses, delivered by the coder via the pulse regenerator, were applied to the intensity control.

Thus the code pulses corresponding to each quantized amplitude were made to appear as a row of blanks in a horizontal trace at the proper relative height. This pattern is very useful in studying coder performance.

Decoders. The decoding method is an impressively simple one originally proposed by C. E. Shannon. In its basic form it employs a pulsed resistance-capacitor circuit as illustrated in Fig. 21. Upon arrival of each pulse of the code, an identical increment of charge is placed upon the capacitor of the

device. The time constant $t = RC$ is such that, during any single pulse interval, whatever charge is on this capacitor decays precisely 50% in amplitude. It follows that the charge remaining at some chosen instant after the arrival of a complete code group consists of contributions of all its pulses, weighted in a binary manner. That is, if we define the contribution of a pulse in the final digit position as $\frac{1}{2}$, then contributions of $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, $\frac{1}{32}$, $\frac{1}{64}$ and $\frac{1}{128}$, respectively, are made by pulses in successively earlier positions. Any value from 0 to $\frac{127}{128}$, in steps of $\frac{1}{128}$, may thus be produced. Of course the digit holes in the aperture plate of the coder are laid out to make this straight-forward scheme workable. Since samples of low-level audio signals are coded near the center of the aperture plate, the corresponding decoded values lie in the neighborhood of $\frac{1}{2}$.

The basic Shannon decoder, then, comprises the resistance-capacitor circuit, means for supplying it with precisely controlled units of charge at precisely determined times, and a sampling and holding circuit (represented by switch B or B' in Fig. 10) to measure and store the decoded potential which is fleetingly present across the capacitor at a regularly recurring instant T, following the final pulse position. The scheme employed to inject the identical charges involves a regulated source of current and a gate to admit this current to the resistance-capacitor circuit under control of the regenerated PCM pulses. Two successive slicing operations and careful gating, as described earlier, make these pulses more than adequately uniform.

The wave form sketches of Fig. 21 show three typical decoding cycles. In the first, a single pulse in digit position 7 produces a decoded amplitude of $\frac{1}{2}$; in the second, a pulse in position 1 gives $\frac{1}{128}$; and in the third, pulses at 2, 5 and 7 provide a decoded value of $\frac{41}{64}$. It may readily be verified that the provision of an idle channel interval between operations (following from the alternate use of two decoders) allows the residue of one decoding operation to decay to a negligible value (never larger than $\frac{1}{128}$ of a single step height or "quantum") by the time of the next consecutive sampling. Experimentally, interchannel crosstalk from this source is virtually nonexistent.

In the foregoing it has been emphasized that precise timing is required for this basic Shannon decoder. Although the necessary accuracy was actually obtained without great difficulty in early tests, a modification has also been introduced which eases the requirements to a very marked extent. This scheme, devised by A. J. Rack, employs a damped resonant circuit in conjunction with the resistance-capacitance elements in a manner illustrated by Fig. 22. The natural frequency of the resonant circuit L, C_2, R_2 is made equal to the PCM pulse rate, and the time constant of the damped oscillation ($t=2R_2C_2$) is matched to that of the circuit R_1, C_1. The same charging

pulses pass through both sections of the circuit; hence by proper choice of C_1 and C_2 the amplitudes of the damped sine wave and the exponential may be proportioned so that the rate of change of their combined potential becomes zero at successive points one pulse period apart. In fact it has been found possible to make both the first and the second time derivatives of potential equal to zero simultaneously at such points.

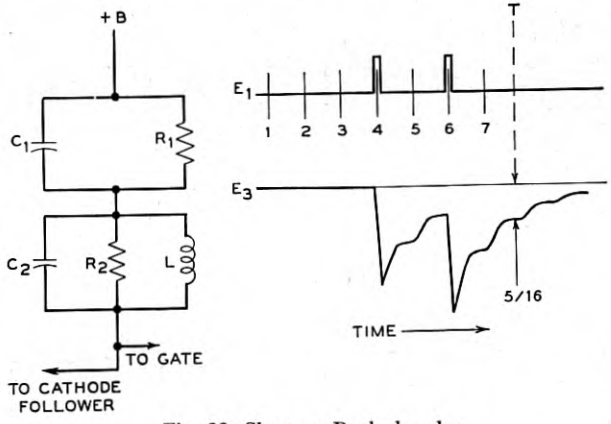


Fig. 22. Shannon-Rack decoder.

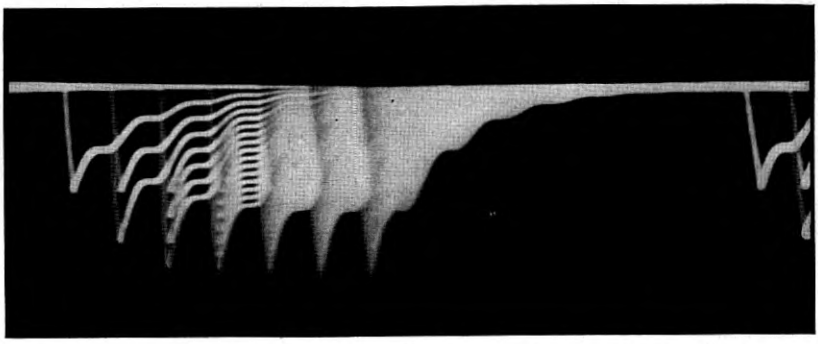


Fig. 23. Output of Shannon-Rack decoder for a signal giving 100% modulation.

By this modification, not only the times of application of the charges but the time of sampling is made much less critical. Of course this presumes that the sampling circuit is designed to complete its operation near the center of a level region. In Fig. 22, the voltage transient due to a typical pair of pulses is sketched, and Fig. 23 shows an oscilloscope screen, on which the waveforms delivered by the cathode follower of one of the Shannon-Rack decoders of the system are superimposed for a full set of 128 possible code combinations in sequence.

V. Performance

In general the behavior of the system has shown promise for toll plant application. Among other things the stability realized in the adjustments of the coder and decoder, and the apparent absence of aging or other drift in the compressor and expandor have been gratifying. A daily check of the focus and centering in the coders and of the time constants in the decoders appears adequate to keep them in optimum adjustment. The synchronizing

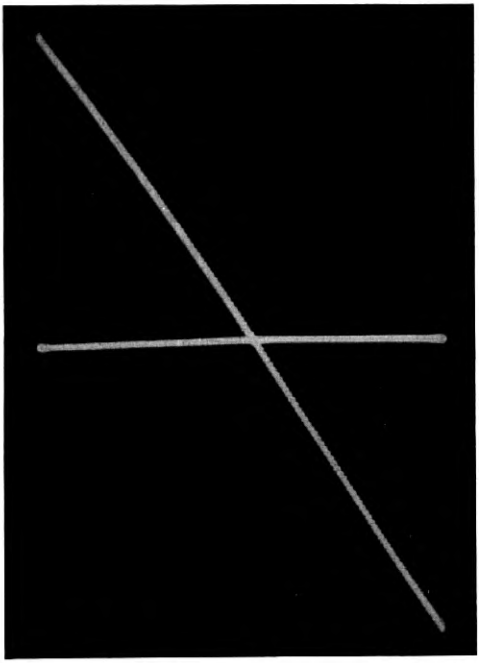


Fig. 24. Output of decoder (vertical) vs. input to coder (horizontal).

gear has been found equally easy to maintain. On the other side of the ledger, occasional breakdowns have pointed up the need for alarm and automatic replacement facilities in any version of the system which might be developed for commercial service.

A few measured characteristics are given in this section in addition to the compression and audio gain-frequency characteristics already shown (Figs. 14 and 11, respectively).

Input-Output Characteristics. The diagonal trace in the oscilloscope pattern of Fig. 24 shows the relationship between the input of a coder (horizontal deflection) and the output of the corresponding decoder (vertical). For this pattern a full-load audio signal was applied to the input of the odd

coder only (without passing through the compressor), and the output was taken directly from the common output of the two decoders. Thus the six odd channels took turns transmitting the signal while the six even channels produced the horizontal center-line. Uniform quantizing steps may be seen along most of the trace, but are obscured near the ends by defocusing of the test oscilloscope.

A similar pattern, obtained with the compressor and expandor included in the transmission path, is shown in Fig. 25. Here tapered steps may be discerned, as well as a small amount of non-linearity due to residual imper-

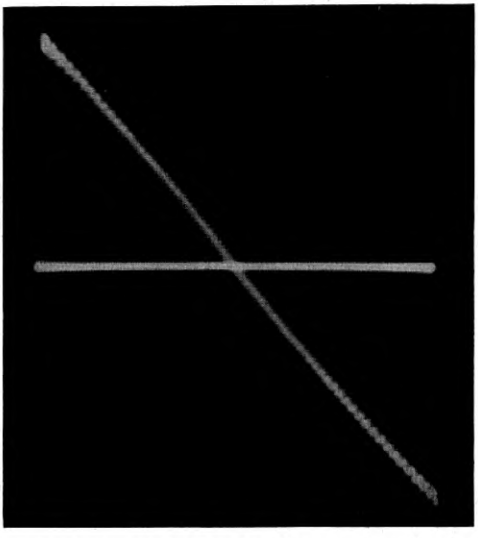


Fig. 25. Output of expandor (vertical) vs. input to compressor (horizontal).

fections in the companding. The effect of the channel peak choppers is not included in this pattern.

Two measured overall input-output characteristics appear in Fig. 26, for the case of a typical single channel and for five channels patched in tandem through 17-decibel pads on a 4-wire basis. The latter simulates a possible extreme case of a long circuit in which for some reason it is desired to decode to audio at four junction points as well as at the final terminal. It should not be confused with a series of spans between which the PCM pulses are amplified or regenerated without decoding. In the latter case, of course, the overall audio characteristics are independent of the number of spans.

Audio Noise. Quantizing was found to be the only significant source of noise in the received audio signals. Noise levels measured in the absence of speech are shown in Fig. 27. The measurements are given for various num-

bers of channels from one to ten, connected in tandem as described in the preceding paragraph. Two scales of ordinates are shown in this figure. On

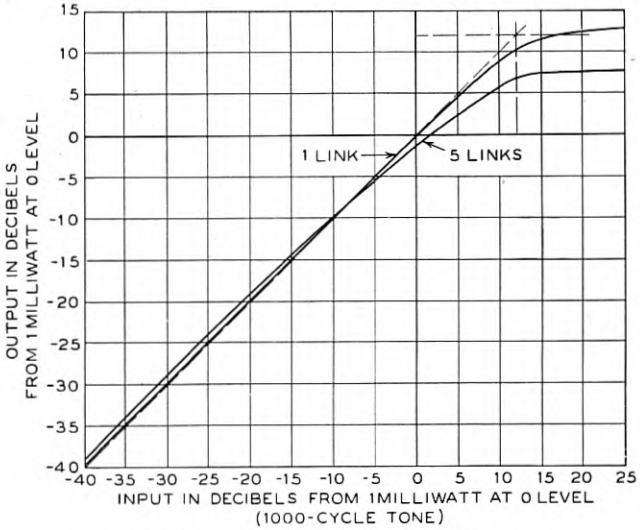


Fig. 26. Input-output characteristics of PCM channels.

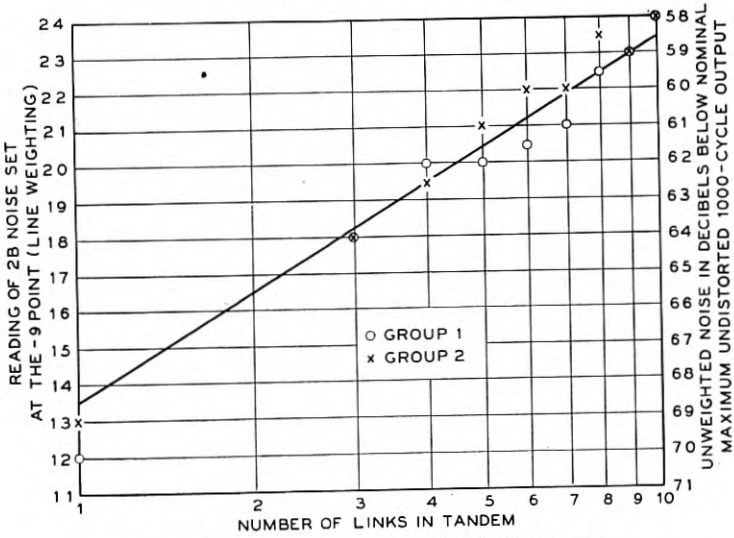


Fig. 27. Noise measurements on idle PCM channels.

the left is a reference scale of weighted noise employed in the Western Electric 2B Noise Set; on the right, a scale which relates the corresponding unweighted root-mean-square noise to the nominal maximum undistorted

single-frequency output of the system (taken to be 9 decibels above a milliwatt at zero level). This output corresponds to an input just reaching the peak limiters. Thus in a single link the idle circuit noise is down about 68 decibels from the full-load sine wave. For five links in tandem, the readings are at least 8 decibels below the accepted limit (29 decibels on the left-hand scale) for noise at the end of a 4000-mile circuit. Quantizing noise is found to increase approximately three decibels for each doubling of the number of links, as is generally the case with other forms of random noise.

For a single channel, or a small number of channels in tandem, the idle circuit noise varies considerably with the vertical centering of the coding tube. This may be understood by noting that if the zero-signal operating point effectively rests near the center of a "tread" on the quantizing staircase, (Fig. 4) a small amount of power hum or other disturbance may simply move it back and forth on the same tread, in which case quantizing noise is entirely absent. On the other hand, if the operating point is near a "riser" the small disturbances may cause it to joggle from one tread to another, producing noise. The measurements given in Fig. 27 were obtained with a very quiet input circuit and with the centering adjusted for maximum noise. The "joggling" was produced largely by residual power hum.

A-B tests to compare PCM transmission over a single link with direct transmission over a noise-free circuit of the same audio band were carried out using a wide range of talker volumes. In such tests only a few experienced observers were able to pick the PCM path consistently. When the PCM circuit included five tandem links most observers could tell the difference, but all judged the quality to be more than satisfactory for toll service.

Crosstalk. It has been pointed out earlier in this paper that interchannel crosstalk in a PCM system can occur only in the terminal equipment. Considerable care was exercised, particularly in the design of the time-division parts of the system, to hold individual sources of crosstalk to 70 decibels or better. As a result, measurements using a single-frequency test tone and a current analyzer have shown the overall crosstalk from any one channel to any other to be down 66 decibels in the worst cases.

Very severe tests have also been made in which a loud talker was connected to ten channels of a group simultaneously and crosstalk into either of the two remaining channels (one odd, one even) was listened to, and measured with the 2-B Noise Set. In such tests unintelligible crosstalk could be detected, which seemed to consist of changes in the quality of the quantizing noise occurring at a syllabic rate. The 2-B readings averaged about a decibel above the quantizing noise of a single idle channel with occasional peaks reaching the 17-decibel point on the reference scale.

In tests involving 24 talkers in 12 simultaneous conversations, crosstalk was practically undetectable.

Radio Interference and Noise. To obtain experimental confirmation of the expected tolerance to high interference levels in the radio path, the output of an oscillator, tunable through the band near 65 megacycles, was superimposed upon the received intermediate-frequency signal at the input to the group selection filters. With this controllable interference tuned near the center of the Group 1 filter, and its amplitude set 6 decibels below the peak amplitude of the (noise-free) pulses, errors were so plentiful that the demodulator did not remain synchronized. Proper framing was restored when the amplitude difference was increased to 7 decibels, but enough decoding errors remained to give intolerable audio noise. At 8 decibels only an occasional crackle of noise was observed, and at 9 decibels reception was perfectly normal. Similar tests of Group 2 gave the same results except, of course, that synchronization was not affected. The fact that noise-free transmission was not maintained quite up to the ideal 6-decibel point is due principally to the width (0.4 microsecond) of the gate which is applied to the rounded PCM pulses entering the receiving equipment. If necessary the ideal could undoubtedly be approached more closely by reducing this width, thus admitting only the extreme peaks and troughs of the signal.

The effects of actual fluctuation noise were studied by sending the PCM pulses over the radio path at reduced level. The boundary between good and bad transmission was not so sharp as with the continuous-wave interference, as should be expected because of the random nature of the noise. Flawless reception occurred when the root-mean-square signal at the peak of a pulse was greater by 18 decibels than the root-mean-square noise, both measurements being made at the output of a group selection filter.

VI. Acknowledgment

The system described is the result of the co-ordinated efforts of many people. In particular, the writers wish to express their appreciation to Mr. R. W. Sears for his development of the coding tube, and to acknowledge the contributions of certain others who have been actively concerned with various phases of the project. These include Messrs. R. L. Carbrey, A. E. Johanson, J. M. Manley, G. W. Pentico, A. J. Rack, P. A. Reiling, L. R. Wrathall, and the mechanics and wiremen who have kept to their usual high standards throughout countless circuit changes.

The project was under the overall direction of Mr. C. B. Feldman.

Electron Beam Deflection Tube for Pulse Code Modulation

By R. W. SEARS

Introduction

PULSE code transmission systems[1] in which successive signal amplitude samples are transmitted by pulse code groups require special modulators. The essential operational requirements of a pulse code modulator are: (1) to quantize or measure the signal amplitude sample to the nearest step in the discrete amplitude scale transmitted by the pulse code system, and (2) to generate the group of on-off pulses identifying the step.

Several methods have been proposed[2, 3, 4] in which quantization and pulse formation were performed with circuits employing conventional electron tubes. The circuits involved sequential and comparison operations and were not easily adapted to a multi-channel time division system because of limitations in coding speeds and the complexity of the equipment. An electron beam deflection tube has been developed which, together with associated beam positioning and sweep circuits, performs the modulation rapidly, making possible the sequential modulation of a number of channels in time division multiplex.

The electronic principles, design and characteristics of the experimental tube are described in the present paper.

Conversion from Signal Input to Pulse Code Output

An electrical input voltage may be converted into an output code pulse group with the electron beam deflection tube shown in Fig. 1a. An aperture or code masking plate is arranged perpendicular to the axis of the electron gun at the focal point. The coordinates of the aperture plate are aligned with the deflection axes of the X and Y deflector plate pairs. The electron beam strikes the output plate when it passes through an opening in the aperture plate.

An input voltage of appropriate value applied to the Y deflector plates will deflect the beam to point "a" of the aperture plate as indicated in Fig. 1a. A linear sweep voltage applied to the X deflection plates, while the

[1] An Experimental Multi-Channel Pulse Code Modulation System of Toll Quality, L. A. Meacham and E. Peterson, this issue.

[2] A. H. Reeves, *U. S. Patent* #2,272,070, Feb. 3, 1942.

[3] H. S. Black and J. O. Edson, paper presented June 11, 1947 at A. I. E. E. meeting; Montreal, Canada.

[4] W. M. Goodall, *Bell System Technical Journal*, July 1947.

input voltage on the Y deflection plates is held constant, causes the electron beam to sweep across the aperture plate along the dashed line *a-b*. A time sequence of output pulses is produced at the output plate when the beam passes through the apertures of the code plate along the path *a-b*.

A series of output pulses or a "pulse group" is characterized by the presence of pulses at time positions corresponding to the several vertical columns of apertures. The code plate shown in Fig. 1 is laid out in accordance with

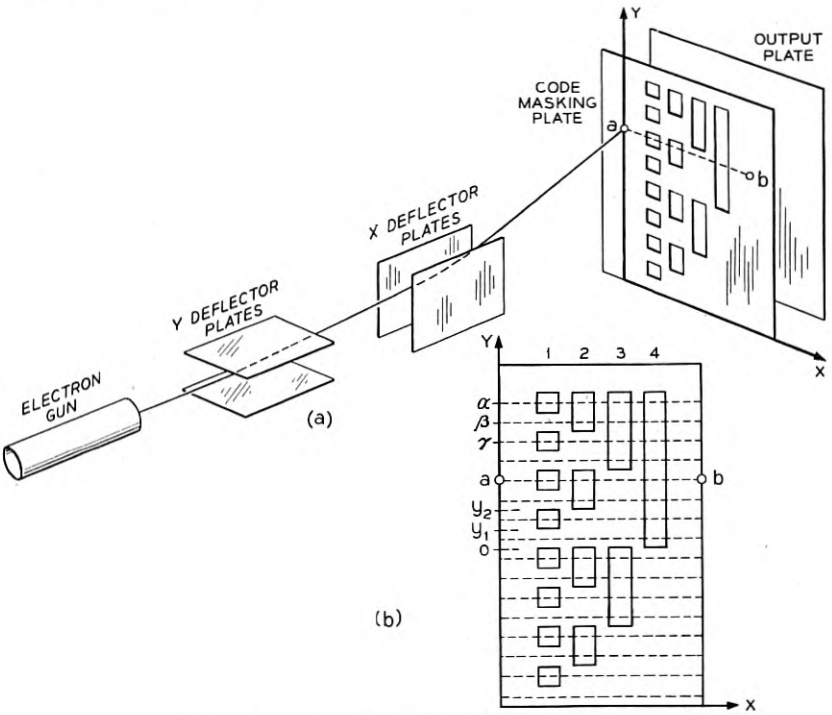


Fig. 1—Electron beam deflection tube for coding.

the binary number system pattern[5] which was chosen for the present pulse code work because of the simplicity of decoding. The four-digit code plate shown in the tube of Fig. 1 provides for coding only 16 amplitude values and was used to facilitate the illustration. The tube developed for the ex-

[5] The four vertical columns of the four-digit code plate in Fig. 1b provide four positions in time for output pulses. Vertical columns marked 1, 2, 3 and 4 correspond to the first, second, third and fourth digits, respectively, of the pulse code transmitted. Pulses located in time in accordance with this notation are given "weights" of 1, 2, 4 and 8, respectively. Sixteen pulse group combinations are indicated by the 16 dashed horizontal lines in Fig. 1b. The pulse group defined by beam sweep α corresponds to a total "weight" of 15; beam positions β and γ correspond to total "weights" of 14 and 13, respectively, and so on, with the bottom horizontal dashed line in the figure corresponding to zero.

perimental pulse code system uses a seven-digit code plate which provides for coding 128 amplitude values.

The vertical distances between the centers of successive code sweep positions (horizontal dashed lines in Fig. 1b) are made equal so that the codes are spaced by equal input voltage steps. A continuous range of input signal amplitudes will result in a continuous range of horizontal sweep positions. With an infinitely small electron beam, input signal amplitudes in the range from o to y_1 will produce a single output pulse group and input signal sample amplitudes from y_1 to y_2 will produce another output pulse group. This process of dividing the total input amplitude range into finite steps and arranging that input voltages falling within each step produce one and only one output pulse group is called quantization.

The tube of Fig. 1 will only quantize effectively if the electron beam is infinitely small and the sweep and aperture plate axes are aligned exactly. With a finite beam size, there will be sweep positions for which the beam straddles and sweeps out a combination of two adjacent codes.

Precise quantization and a uniform pulse output are required. The problems of quantizing, alignment and uniform pulse output have been solved by the use of a wire grid, called the quantizing grid, located in front of the aperture plate.

Quantization of Beam Position by Feedback

The quantizing grid consists of a horizontal array of grid wires aligned parallel to the code sweep or X axis of the aperture plate. The grid spacings and alignment are such that a wire lies between each adjacent pair of code groups as viewed from the various incident angles of the deflected electron beam. The quantizing grid, by means of an electrical feedback path to the signal deflection plates, divides the input signal range into a number of equal steps and positions the electron beam to the proper level for the code corresponding to the voltage step within which the signal amplitude sample falls. The quantizing grid wires also constrain the electron beam during the formation of the output code pulses, so that it must sweep out the code initially selected. In general, wires or shaped electrodes of any sort located where the electron beam can impinge thereon and connected in feedback relation to the deflection system constrains the electron beam to move in patterns prescribed by these electrodes and are thus called beam guides.

The coding tube with quantizing grid and feedback circuit is shown schematically in Fig. 2. The electrode line-up, reading from left to right in the figure, consists of an electron gun, deflection system, secondary electron collector, quantizing grid, aperture plate and output plate.

For the present purpose, a consideration of the collector and output plate electrodes is omitted and it is assumed that the grid does not emit secondary

electrons when bombarded. Electrons which strike the grid produce a current in the grid circuit while electrons that miss the grid have no effect in the grid circuit. The electron beam current intercepted by the grid will be dependent on the y deflection or position of the electron beam. The current is a maximum when the beam is centered on a grid wire and a minimum when it is centered between two grid wires and varies with beam position as indicated by the curve in the lower portion of Fig. 3. The curve is constructed for the case in which the beam diameter is slightly greater than the space between two grid wires. The current to the grid never becomes zero for any beam position. It may be thought of as having a "d-c. component" B. Amplifier 2 in Fig. 2 introduces a bias which cancels this "d-c. component" so that the feedback voltage is symmetrical about zero.

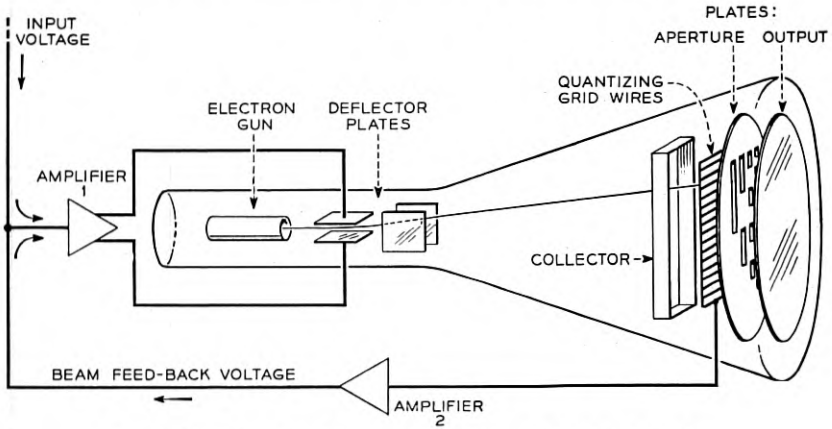


Fig. 2—Coding tube with quantizing grid and circuit schematic.

The grid current is amplified and the voltage developed is applied in feedback relation to the Y deflection plates as shown in Fig. 2. The case in which a positive feedback potential deflects the beam in the same direction as that for a positive signal voltage will be considered.

The beam deflecting voltage is equal to the sum of signal and feedback voltages and the beam position is a linear function of the deflection voltage so we may write

$$-e + Dy = e_f \tag{1}$$

where e is the input signal voltage, D the deflection constant, y the beam deflection or position and e_f the feedback voltage. The feedback voltage e_f is a periodic function of beam position y. Equation 1 therefore defines equilibrium beam positions for input signal voltage e. Equilibrium beam positions in accordance with Equation 1 are determined graphically in the

top portion of Fig. 3. The feedback voltage representing the right-hand side of Equation 1 is plotted as a function of beam position. The electron beam will have several possible positions of equilibrium at points ($p1$, $p2$, $p3$, $p4$, $p5$, $p6$ and $p7$) where the deflection line D erected from $-e$, representing the left-hand portion of Equation 1, crosses the feedback curve. These are the only beam positions for which the deflection potential (signal plus feedback) attains correct values for corresponding beam positions.

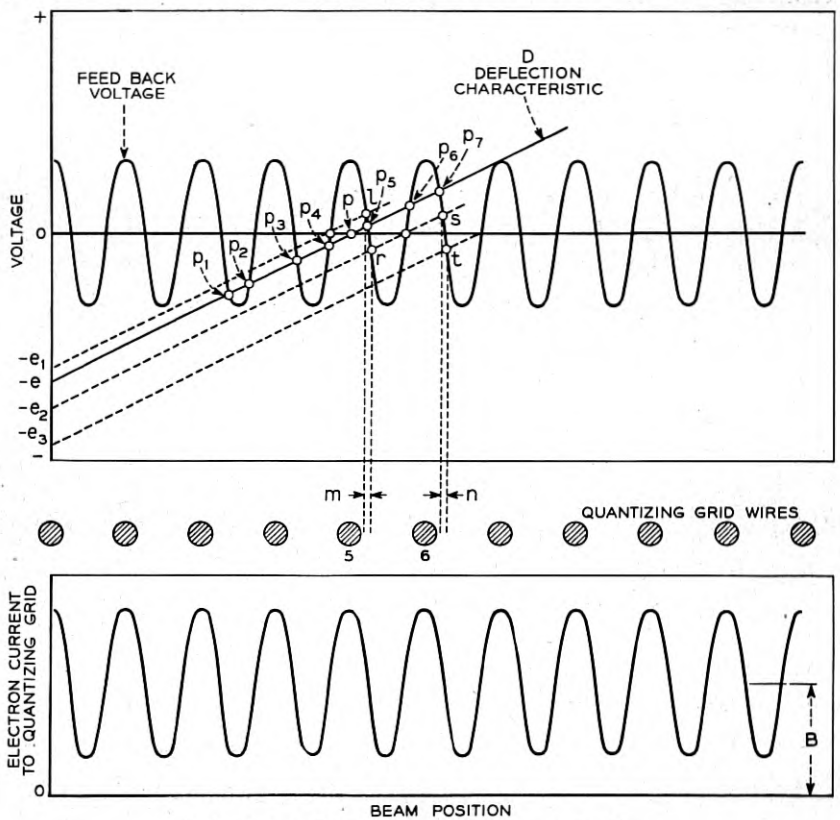


Fig. 3—Graphical representation of quantization.

However, only positions $p1$, $p3$, $p5$ and $p7$ will be in a true state of equilibrium. This can be seen as follows: Consider the beam at point $p3$; if the beam is perturbed toward the right, the feedback voltage tending to deflect the beam to the left increases. The opposite action ensues if the beam is perturbed in the left-hand direction. Point $p3$ is a true equilibrium point. On the other hand, consider the beam at position $p4$. If the beam is perturbed to the right, the feedback voltage tending to deflect it to the right increases and the beam continues to move until it reaches the equilibrium

position $p5$. If the beam is perturbed to the left from position $p4$ it will continue to move to the left until it reaches position $p3$.

The number of possible equilibrium beam positions for an input signal sample depends on the maximum values of the feedback voltage and the slope of the deflection characteristic. It is necessary that only one equilibrium beam position be available for a particular small range of signal voltage. This can be achieved as follows: if a signal voltage e is established with the feedback circuit inoperative, the beam will be at a position p, Fig. 3. When the feedback circuit is activated with the signal voltage held at e, the beam will move from p to $p5$. With this procedure, signals in the range from e_1 to e_2 will result in equilibrium beam positions between points l and r on the curve. Thus, for signal voltages within the range from e_1 to e_2 the beam will fall in the small spacial interval m, whereas beam positions for input signals in the same range without feedback would vary from grid wire 5 to grid wire 6. Likewise, signal voltages between e_2 and e_3 will cause the beam to assume positions between s and t in the spacial interval n. The electron beam may be thought of as "leaning" on one side of a grid wire for a finite signal voltage range and on the same side of the next grid wire for an adjacent signal voltage range.

If the feedback voltage is of the opposite polarity to that assumed above, the quantizing action proceeds in the same manner except that the quantized beam positions lie at the left of the wires. The beam may be thought of as "leaning" on the opposite side of the grid wire.

The proper quantizing action is obtained by establishing and holding the signal voltage with the feedback circuit inoperative and then activating this circuit. The feedback circuit may be deactivated and activated by either (a) blanking and deblanking the electron beam, or (b) defocusing and focusing the electron beam by applying the proper voltage change to the beam control or focusing electrodes of the gun, respectively.

Since the grid wires are parallel to the horizontal rows of aperture holes, the feedback action constrains the beam to sweep out the code group initially selected even though the sweep axis is tilted slightly with respect to the grid wires and aperture plate. The maximum swing of the feedback voltage at the deflection plates should be about three or four times the value of the voltage required to deflect the beam from one code group to the next in order to provide ample protection against the beam jumping from one code group to the next code group during the sweep.

The Experimental Coding Tube

The seven-digit experimental tube developed for pulse code transmission system trials utilizing the electrode lineup shown schematically in Fig. 2 is pictured in Fig. 4. The electron gun assembly and the target plate assembly

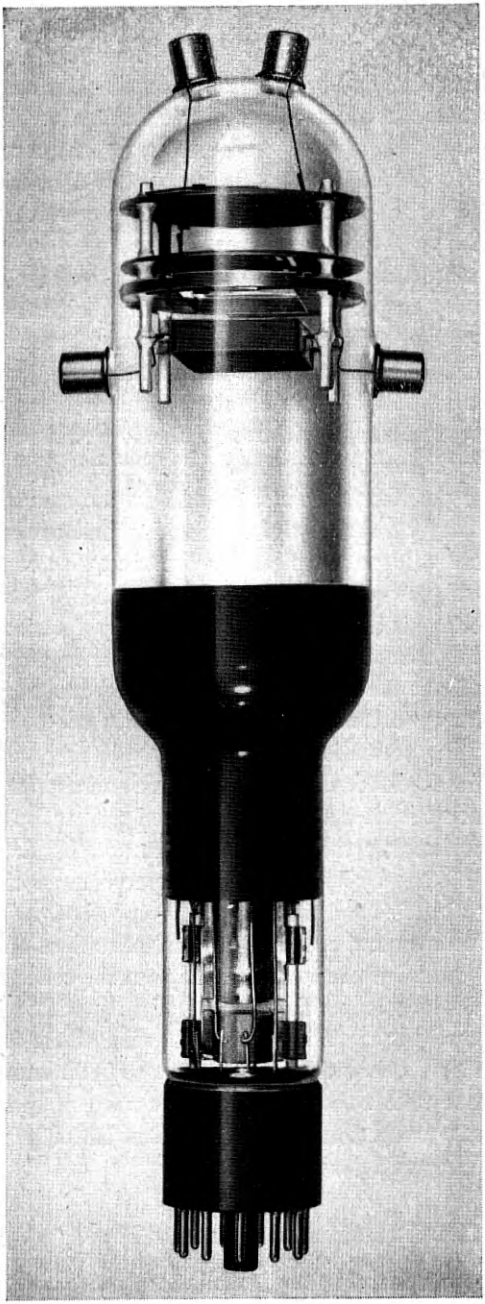


Fig. 4—Experimental seven-digit coding tube.

are sealed in at opposite ends of the tube envelope. The over-all length of the tube is $11\frac{1}{4}''$ with a maximum bulb diameter of $2\frac{1}{4}''$.

The electron gun operates at a final anode potential of 1000 volts with a beam current of approximately 10 microamperes. A potential of about 100 volts applied to the signal input deflection plates deflects the beam from the center to the top of the aperture plate. This corresponds to a maximum deflection angle of $10\frac{1}{2}°$.

The four electrodes of the target assembly, secondary collector, quantizing grid, aperture plate and output plate are shown in the photograph of Fig. 5 from bottom to top respectively. The secondary collector is a simple rectangular shaped electrode. The quantizing grid consists of a circular frame with a parallel array of grid wires stretched across a rectangular opening. The aperture plate is a thin disc with apertures arranged in a binary pattern which provides for a seven-digit code. The output plate is a thin circular disc. Both aperture and output plates are coated with a carbon layer to suppress secondary electron emission from their surfaces.

The parts of the target assembly are held in accurate alignment in a jig and cemented and held in position on four ceramic rods. The entire assembly is held rigidly in the tube envelope by means of spacers attached to the quantizing grid and output plate.

The target assembly is aligned with the electron gun and deflection plate axes by means of lineup tools in the glass lathe at the time the final seal is made at the center of the glass envelope. It has been possible to hold the alignment of the deflection axes with the aperture plate to within slightly less than 1° with this construction.

The construction of the quantizing grid may be seen more clearly in the photograph of Fig. 6. The grid frame has raised portions on two sides of the rectangular opening. These are milled with a series of grooves for each grid wire. The grid laterals are affixed in the grooves by brazing and are thus accurately spaced with respect to each other and to assembly lineup holes which can be seen spaced around the edge of the grid frame. The wires are held taut by means of a flat spring which is welded to the grid frame and supplies tension to stretch the lateral wires. The grid wires are 4.0 mils in diameter, processed to have a secondary emission coefficient of about 3. The laterals are spaced 11.6 mils between centers.

The openings in the aperture plate are made by a punching operation. The area of the aperture plate covered by the first and second digit columns is milled to a thickness of 5 mils in order to facilitate accurate punching of the smallest apertures. The apertures in the first digit column (bottom horizontal row in Fig. 5) are rectangular .012″ x .062″. The seven-digit code pattern provides 128 different output pulse groups. The wide openings

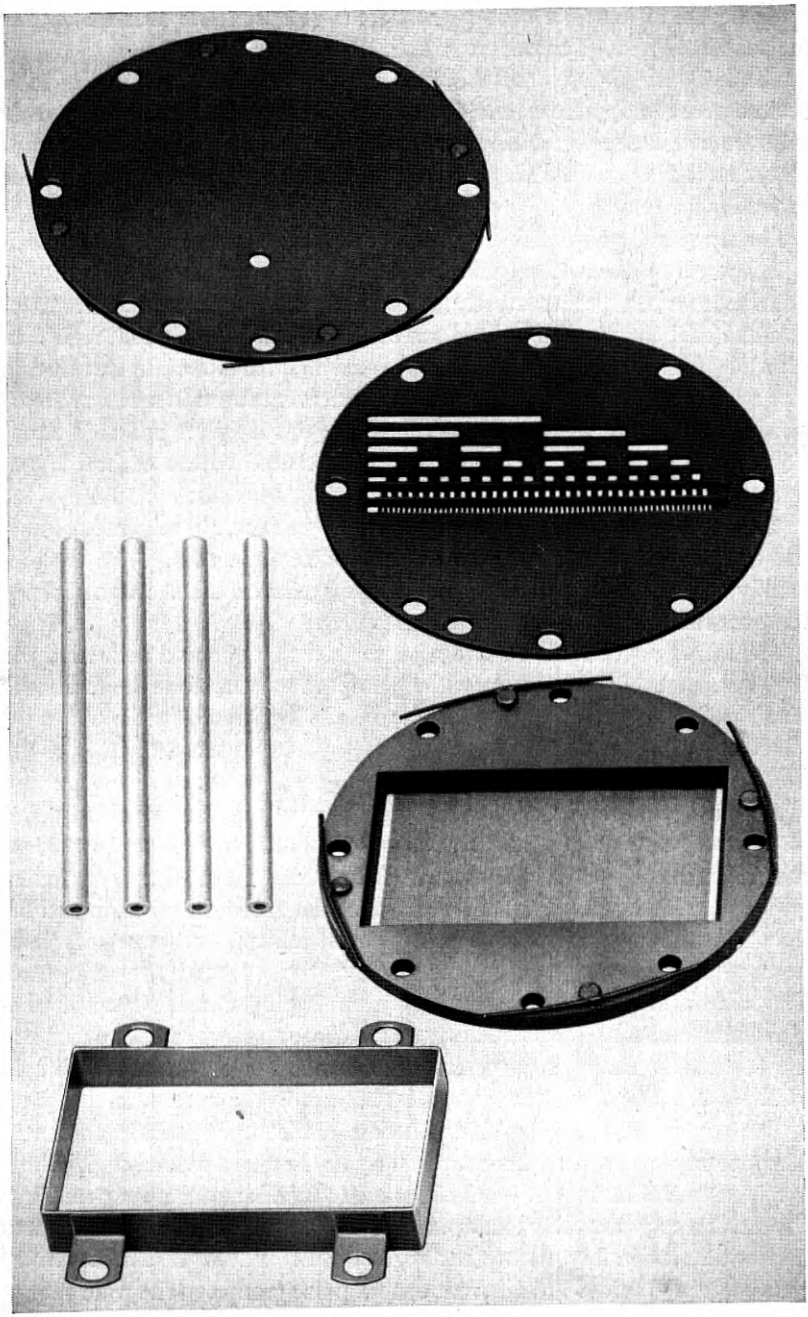


Fig. 5—Target electrodes.

at the left hand side of the aperture plate in Fig. 5 provide a peak amplitude range for which the output pulse group consists of all seven pulses. This in effect provides a peak limiting action.

Leads from the four electrodes of the target assembly are brought out directly to terminal caps on the side and end of the tube envelope to decrease the interelectrode capacitances and to facilitate direct connection to external circuits.

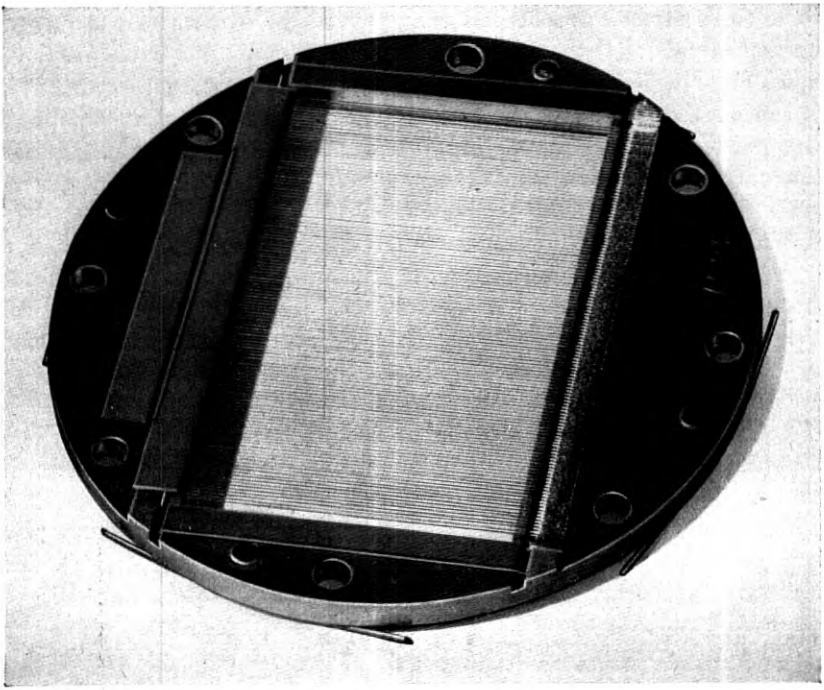


Fig. 6—Quantizing grid.

Operational and Design Considerations

The quantizing action depends on the periodic variation of electron beam current intercepted by the wires of the quantizing grid with beam position. Secondary electron emission from the grid and other electrodes was assumed to be negligible with the feedback circuit connected directly to the quantizing grid as shown in Fig. 2. The uniform suppression of secondary electrons is difficult to achieve even though the grid is coated and processed for a low secondary ratio. It is also difficult to prevent secondary electrons produced at the aperture plate from being collected by the grid.

A preferred method of operation utilizes the secondary electrons produced

at the grid by the impinging primary beam for the quantizing action rather than the directly intercepted electron current as heretofore assumed. The secondary electron collector located in front of the quantizing grid is maintained at a positive potential and collects most of the secondaries from the grid. There is, of course, a correspondence between the secondary electron current and the fraction of the beam current intercepted by the grid wires. The quantizing circuit is made by connecting the feedback path to the secondary collector and the quantizing action proceeds as described heretofore. This method has the following advantages over the direct primary current method: (1) the collector current as a function of beam position is much more regular; (2) the swing between maximum and minimum current is considerably larger because of the secondary emission multiplication at the grid surface; and (3) the capacitance between collector and ground is lower than the capacitance between the closely spaced grid and aperture plate.

With secondary electron current feedback, the aperture plate is operated at a positive potential relative to the grid to suppress secondary electrons from the aperture plate. The proportion of the secondary emission from the grid collected by the aperture plate is small compared with that collected by the secondary collector. High velocity secondaries originating at the aperture plate are, however, able to penetrate the retarding field and strike the grid. These energetic secondaries produce low-velocity secondaries at the grid which flow to the secondary collector electrode. This alters the character of the secondary collector current in accordance with the spacial pattern of the apertures in the code plate. The surface of the aperture plate is carbonized to reduce the emission of high-velocity secondaries. The spacial variation of the quantizing current is reduced to less than 10% of the total quantizing current swing with this treatment.

The periodic nature of the quantizing current with beam position must be as uniform as possible as regards both the "a-c" and "d-c" components. The maximum swing should also be as large as possible to permit the use of the lowest possible impedance in the feedback path to obtain a wideband characteristic for fast coding.

The factors which determine the maximum current swing are electron beam current, secondary emission coefficient of the grid wires, electron beam spot size, grid wire diameter and spacing. The quantized beam falls approximately halfway between the center of a grid wire and the midpoint between wires. The beam must be small enough that its edge does not extend into the region beyond the wire on which it "leans" by an appreciable amount. This is obtained with the electron beam focused to a spot slightly smaller than that for maximum quantizing current amplitude.

Optimum performance has been obtained with the electron beam focused to a radius[6] of about 5 mils for the grid spacings previously specified.

The principal irregularities in the periodic quantizing current are caused by variations in secondary emission coefficient of the grid and aperture plate and deflection defocusing of the electron beam. The secondary emission from the grid surface can be made sufficiently uniform by careful processing. This is illustrated in Fig. 7 which is a trace of an oscilloscope

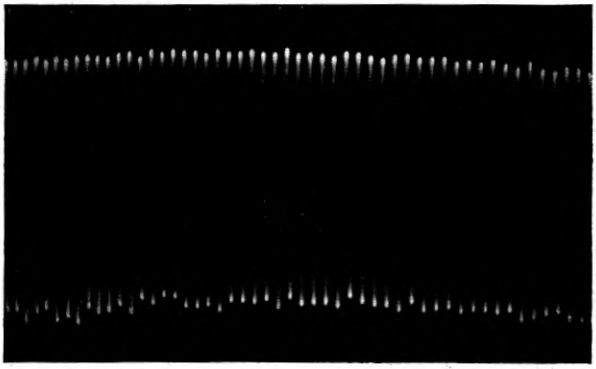


Fig. 7—Variation of quantizing current with beam deflection.

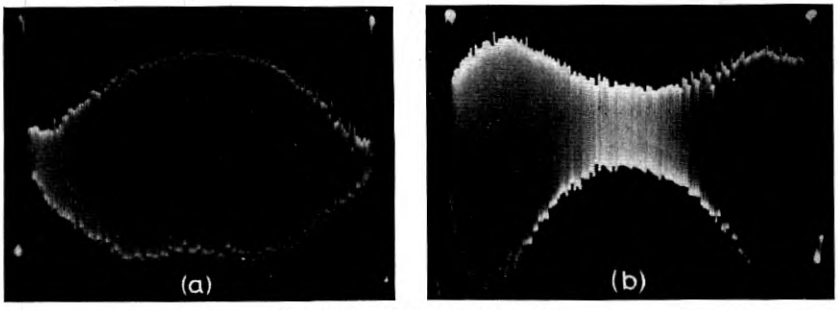


Fig. 8—Deflection defocussing.

presentation of the secondary collector current as the beam is deflected by a linear sweep at right angles to the grid laterals. The trace was limited to cover a sufficiently small number of wires to show details of its shape. Figure 8a illustrates the effect of deflection defocusing. The sweep was expanded to cover all of the grid wires and the electron beam focused for maximum amplitude in the center region. The curve is so compressed that

[6] Distance from the center of the beam to a point at which the electron current density is 5% of its value at the center.

the individual oscillations of the current between maximum and minimum values can hardly be resolved. The envelope of the curve indicates the extent to which the swing is reduced at the two ends by the increase in electron beam spot size which results from deflection defocusing. In Fig. 8b, the focusing voltage has been changed by 12 volts and it can be seen that the beam is in focus at the maximum deflection angles and out of focus in the center region. With an intermediate or compromise focus voltage, the tube will quantize satisfactorily over the entire code range. Best results have been obtained with the tube in the experimental pulse code system by the use of a simple circuit which changes the focus voltage in a linear manner with the rectified or absolute value of the input deflecting signal thus compensating for the deflection defocusing of the tube.

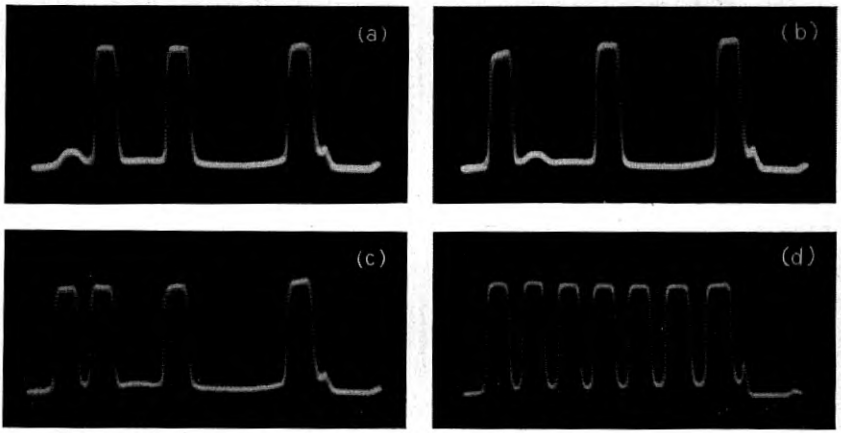


Fig. 9—Typical pulse code outputs.

The output plate is usually operated at a positive potential relative to the aperture plate with attendant suppression of secondary emission from the former. Output pulses of opposite polarity may be obtained by operating the output plate at a negative potential and processing this plate to have a secondary emission ratio greater than unity.

The several groups of output pulses shown in Fig. 9 are illustrative of the tube output. Successive digit pulse positions occur from left to right in the figures. The small kink in the base line at the first and second digit pulse positions in codes a and b respectively are small error pulses caused by a small portion of the edge of the beam overlapping a grid wire when the beam is quantized. Error pulses of this magnitude are readily eliminated by slicing when the tube is used in the pulse code transmission system.

Coding tubes have operated satisfactorily for long periods of time in the experimental multi-channel pulse code system and have required minor adjustments of potentials no more than once a day. The tube operates in this equipment with a code sweep time of 10 microseconds.

Acknowledgment

The writer wishes to acknowledge the original suggestions of Dr. F. B. Llewellyn and Mr. G. Hecht and the help of his colleagues. The contributions and wholehearted cooperation of Messrs. C. B. Feldman, L. A. Meacham and their associates who concurrently developed the pulse code system and circuits associated with the coding tube, have been invaluable. Mr. A. Salecker deserves a great deal of credit for the detailed mechanical design of the tube and his work in producing tube models.

The development was under the over-all direction of Mr. J. R. Wilson and Dr. S. B. Ingram.

Metallic Delay Lenses

By WINSTON E. KOCK

A metallic lens antenna is described in which the focussing action is obtained by a reduction of the phase velocity of radio waves passing through the lens rather than by increasing it as in the original metal plate lens. The lens shape accordingly corresponds to that of a glass optical lens, being thick at the center and thin at the edges. The reduced velocity or "delay" is caused by the presence of conducting elements whose length in the direction of the electric vector of the impressed field is small compared to the wavelength; these act as small dipoles similar to the molecular dipoles set up in non-polar dielectrics by an impressed field. The lens possesses the relatively broad band characteristics of a solid dielectric lens, and since the conducting element can be made quite light, the weight advantage of the metal lens is retained. Various types of lenses are described and a theoretical discussion of the expected dielectric constants is given. An antenna design which is especially suitable for microwave repeater application is described in some detail.

Introduction

THE metal lens antennas described by the writer elsewhere[1] comprised rows of conducting plates which acted as wave guides; a focussing effect was achieved by virtue of the higher phase velocity of electromagnetic waves passing between the plates. Higher phase velocity connotes an effective index of refraction less than unity, and a converging lens therefore assumes a concave shape. The relation between the index of refraction n, the plate spacing, a, and the wavelength λ

$$n = \sqrt{1 - (\lambda/2a)^2}, \tag{1}$$

indicates that the refractive index varies with wave length. As a consequence, such lenses exhibit "chromatic abberation"; i.e. the band of frequencies over which they will satisfactorily operate is limited. Although some of these lenses may have ample bandwidth (15% to 20%) for most microwave applications, others, having large apertures in wavelengths, may have objectionable bandwidth limitations. For example, the lens of Fig. 1, having an aperture diameter of 96 wavelengths, has a useful bandwidth of only 5%.

As a means for overcoming these band limitations, the metallic lenses of this paper were developed. They are light in weight and possess an index of refraction which can be made sensibly constant over any desired band of microwave frequencies. They avoid the weight disadvantages of glass or plastic lenses, and retain the tolerance and shielding advantages of the lens

[1] W. E. Kock, *Bell Laboratories Record*, May 1946, p. 193; *Proc. I. R. E.*, Nov. 1946, p. 828.

over the reflector antenna. Because electromagnetic waves passing through them are slowed down or "delayed" (as in the glass lenses of optics), they are called delay lenses; and since the elements in the lenses which produce the delay are purely metallic, they are called metallic delay lenses.

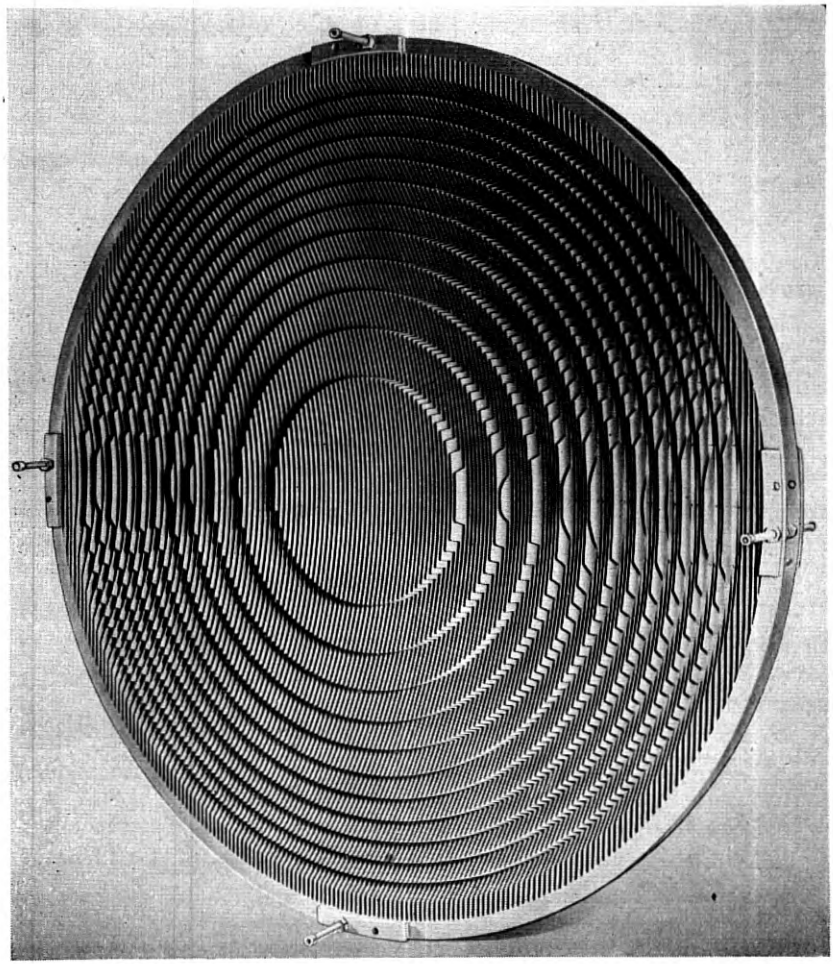


Fig. 1—Waveguide metallic lens having an aperture of 96 wavelengths and a useful bandwidth of 5%.

PART I—EXPERIMENTAL

Fundamental Principles

The artificial dielectric material which constitutes the delay lens was arrived at by reproducing, on a much larger scale, those processes occurring

in the molecules of a true dielectric which produce the observed delay of electromagnetic waves in such dielectrics. This involved arranging metallic elements in a three-dimensional array or lattice structure to simulate the crystalline lattices of the dielectric material. Such an array responds to radio waves just as a molecular lattice responds to light waves; the free electrons in the metal elements flow back and forth under the action of the alternating electric field, causing the elements to become oscillating dipoles similar to the oscillating molecular dipoles of the dielectric. In both cases, the relation between the effective dielectric constant ϵ of the medium, the density of the elements N (number per unit volume) and the dipole strength (polarizability α of each element) is approximately given by

$$\epsilon = \epsilon_0 + N\alpha \tag{2}$$

where ϵ_0 is the dielectric constant of free space.

There are two requirements which are imposed on the lattice structure. First, the spacing of the elements must be somewhat less than one wavelength of the shortest radio wave length to be transmitted, otherwise diffraction effects will occur as in ordinary dielectrics when the wavelength is shorter than the lattice spacing (X-ray diffraction by crystalline substances). Secondly, the size of the elements must be small relative to the minimum wavelength so that resonance effects are avoided. The first resonance occurs when the element size is approximately one half wavelength, and for frequencies in the vicinity of this resonance frequency the polarizability α of the element is not independent of frequency. If the element size is made equal to or less than one quarter wavelength at the smallest operating wavelength, it is found that α and hence ϵ in equation 2 is substantially constant for all longer wavelengths.

Since lenses of this type will effect an equal amount of wave delay at all wavelengths which are long compared to the size and spacing of the objects, they can be designed to operate over any desired wavelength band. For large operating bandwidths, the stepping process[2] is to be avoided, since the step design is correct only at one particular wavelength. Such unstepped lenses are thicker, but the diffraction at the steps is eliminated and a somewhat higher gain and superior pattern compared to a stepped lens is achieved. By tilting the lens a small amount, energy reflected from it is prevented from entering the feed line and a good impedance match between the antenna and the line can be maintained over a large band of frequencies.

Another way of looking at the wave delay produced by lattices of small conductors is to consider them as capacitative elements which "load" free

[2] The lens of Fig. 1 has 12 steps.

space, just as parallel capacitors on a transmission line act as loading elements to reduce the wave velocity. Consider a charged parallel plate air condenser with its electric lines of force perpendicular to the plates. Its capacity can be increased either by the insertion of dielectric material or by the insertion of insulated conducting objects between the plates if the objects have some length in the direction of the electric lines of force. This is because such objects will cause a rearrangement of the lines of force (with a consequent increase in their number) similar to that produced by the shift, due to an applied field, of the oppositely charged particles comprising the molecules of the dielectric material. The conducting elements in the lens may thus be considered either as portions of individual condensers, or as objects which, under the action of the applied field, act as dipoles and produce a dielectric polarization, similar to that formed by the rearrangement of the charged particles comprising a non-polar dielectric.[3] Either viewpoint leads to the delay mechanism observed in the focussing action of the artificial dielectric lenses to be described.

Experimental Models

We turn now to experimental exemplifications of lenses built in accordance with the principles outlined above.

(a) *Sphere Array*

One of the simpler shapes of conducting elements to be tried was the sphere. Figure 2 is a sketch of an array of conducting spheres arranged approximately in the shape of a convex lens. The spheres are mounted on insulated supporting rods; the microwave feed horn and receiver are shown at the right. The focal length is f, the radius of the lens "aperture" is y, the maximum thickness is x and not only the spacings s_1 and s_2 but also the size of the spheres are small compared to the wavelength. Rays A and B are of equal electrical length because ray A is slowed down or delayed in passing through the lens. Figure 3 is a photograph of the lens of Fig. 2; it also portrays a similar sphere array lens made of steel ball bearings supported by sheets of polystyrene foam[4] which have holes drilled in them to accept the spheres. In both cases the balls are arranged in a symmetrical lattice. It will be shown below that the polarizability α of a conducting

[3] Polar dielectrics have arrangements of charged particles which are electric dipoles even before an external electric field is applied; the field tends to align these and the amount of polarization (and hence the dielectric constant) that they exhibit depends upon temperature, since collisions tend to destroy the alignment. Non-polar (or hetero-polar) molecules have no dipole moment until an electric field is applied; the polarization of such materials (and of the artificial dielectrics we are discussing) is accordingly independent of temperature. See, for example, Debye, "Polar Molecules", Chap. III.

[4] Styrofoam (Dow). Because of its low density (1 to 2 pounds per cu. ft.), its contribution to the wave delay is negligible ($\epsilon_r = 1.02$).

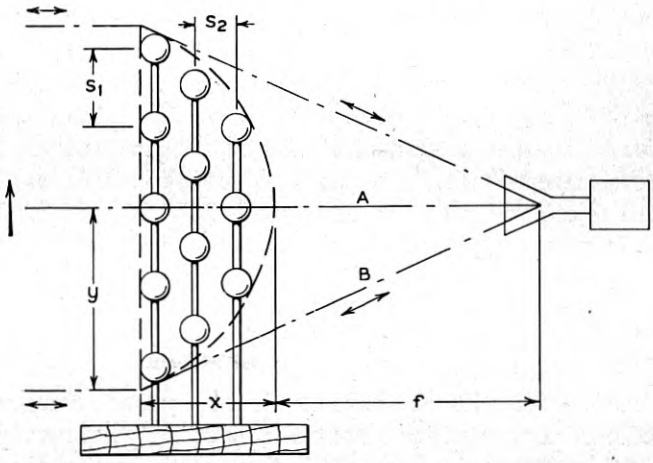


Fig. 2—Lattice of conducting spheres arranged to form a convex lens.

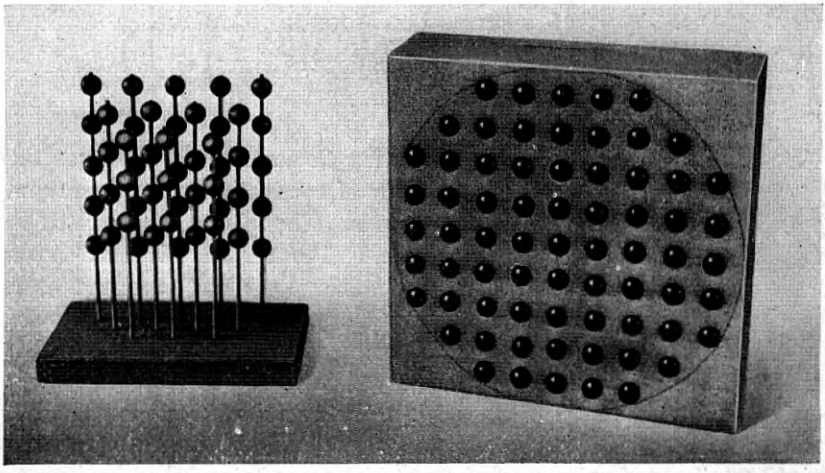


Fig. 3—Left: The sphere lens of Fig. 2. Right: Sphere lattice supported by foam sheets.

sphere of radius a is $4\pi\epsilon_0 a^3$, so that, for static fields, the relative dielectric constant is, from (2),

$$\epsilon_r = \epsilon/\epsilon_0 = 1 + 4\pi N a^3, \tag{3}$$

where N = number of spheres per unit volume.

For most dielectrics, the index of refraction is simply the square root of the relative dielectric constant. However, in the case of the sphere lens at microwaves eddy currents on the surface of the spheres prevent the magnetic

lines of forces from penetrating them and it will be seen later that this effect causes the expected value of the index of refraction to be smaller than that determined by the usual equation

$$n^2 = \epsilon_r. \tag{4}$$

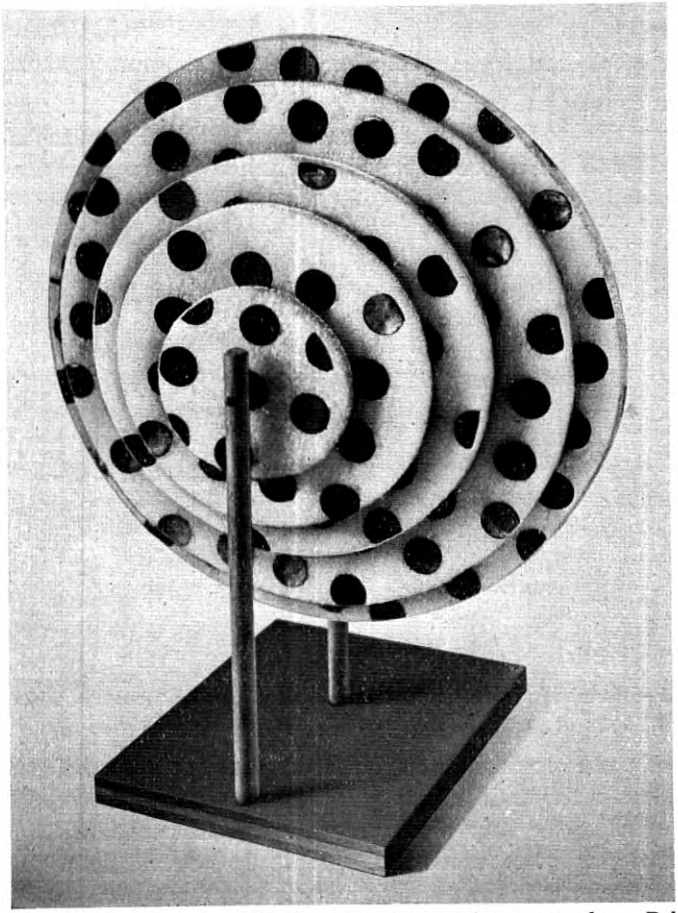


Fig. 4—Lattice of conducting disks arranged to form a plano-convex lens. Polystyrene foam sheets support the disks.

To avoid this effect the elements should be shaped so as not to alter the magnetic lines of force. This can be done by making their dimension in the direction of propagation of the impressed waves negligibly small.

(b) *Disk Array*

One accordingly arrives at the lens design shown in Fig. 4 in which the spheres are replaced by copper foil disks lying in planes parallel to the

impressed E and H vectors. The supporting sheets are again polystyrene foam. The foil disks have negligible thickness so that the magnetic lines are unaffected as shown in Fig. 5. Equation (4) is therevore valid even at radio frequencies and the index of refraction of this material is obtainable from (4) and (2) with α equal to $\frac{16\,\epsilon_0\,a^3}{3}$, as shown in the last section.

(c) *Strip Array*

Both the sphere and disk type lenses have the advantage that they will perform equally well on horizontally or vertically polarized waves. If the lens is required to focus only one type of wave polarization the disks can be replaced by thin, flat, conducting strips extending in length in the direction

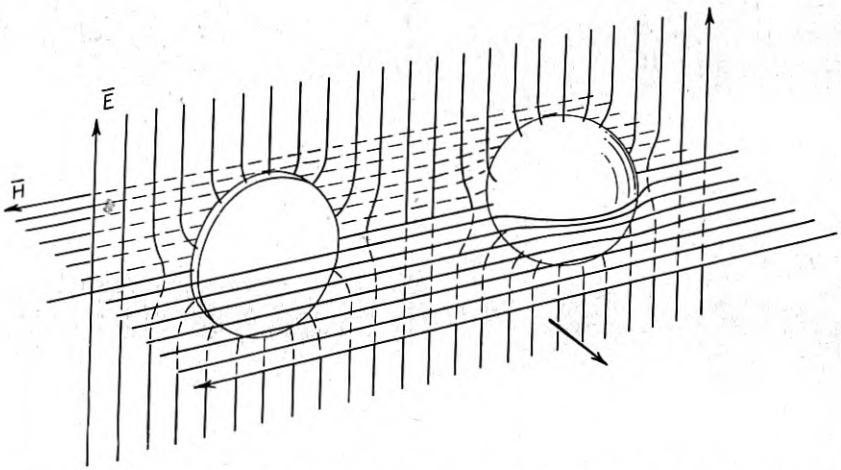


Fig. 5—Disturbance of the magnetic field is avoided by the use of disks instead of spheres.

of the magnetic vector of the applied field. A simple method of constructing such a lens is shown in Fig. 6. Slabs of dielectric foam are slotted to a depth equal to the strip width and each slab is marked with the profile contour necessary to produce the final plano-convex lens. The metal strips are then cut to the length indicated on the profile and inserted in the slots. With the strips in place the unit slabs are stacked on top of one another and held in a mounting frame to form the complete lens. Figure 7 shows one slab of a 10-foot strip type lens being assembled and Fig. 8 shows a six-foot square lens half assembled. Figure 9 shows directional patterns of a 3-foot diameter lens of this type made up of $\frac{3}{4}$ inch x .005 inch strips spaced $\frac{3}{8}''$ in the slabs and the slabs $1\frac{1}{2}''$ thick. The three patterns were taken over a 12% band of frequencies with the lens purposely illuminated by

a low directivity feed at the focal point. This produced an almost uniform illumination across the aperture at all three wavelengths so that the side lobes were not too well suppressed. However, the deep minima in all three directional patterns indicate the relative absence of curvature of the emerging phase front; this signifies that the strip dielectric has a negligible variation of refractive index over the indicated wavelength band.

(d) *Sprayed Sheet Lenses*

The disk lens or strip lens can be constructed in the manner indicated in Fig. 10, which shows two lenses made by spraying conducting paint on thin dielectric sheets through masks. This results in round dot or square dot

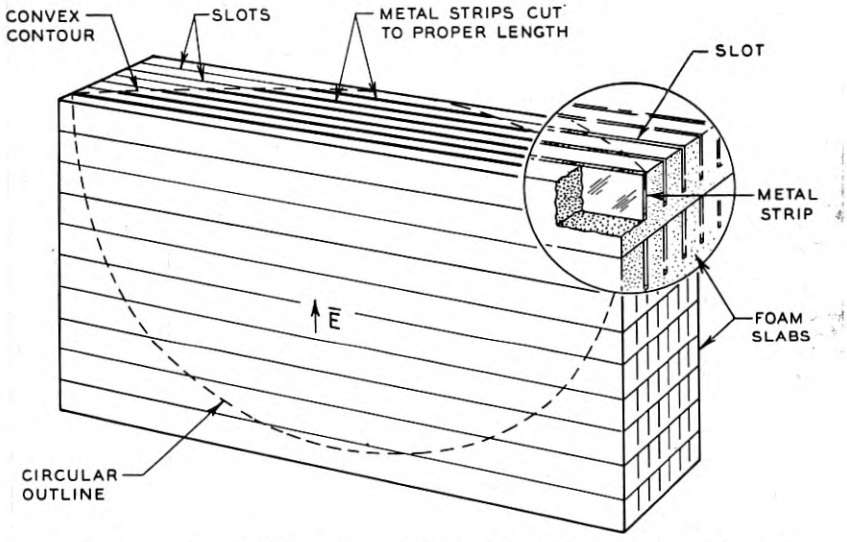


Fig. 6—Construction of a delay lens employing metallic strips as the delay elements.

patterns on the sheets and the size circle used on each sheet determines the three-dimensional contour when the sheets are stacked up to form the lens. Those in the photograph are spaced by wooden spacers as shown in the sketch of Fig. 11; for large lenses it would probably be preferable to cement the sheets to polystyrene foam spacers, thereby making a solid foam lens. Because of the small size of the elements, the lens on the left in Fig. 10, was effective at wavelengths as short as 1.25 cm, in addition to longer wavelengths (3, 7 and 10 cm). The lens of Fig. 12 was made by metal spraying metallic tin directly on circular foam slabs through masks, and the foam disks then cemented together.

Strip lenses can also be constructed in this way. The lens in Fig. 13,

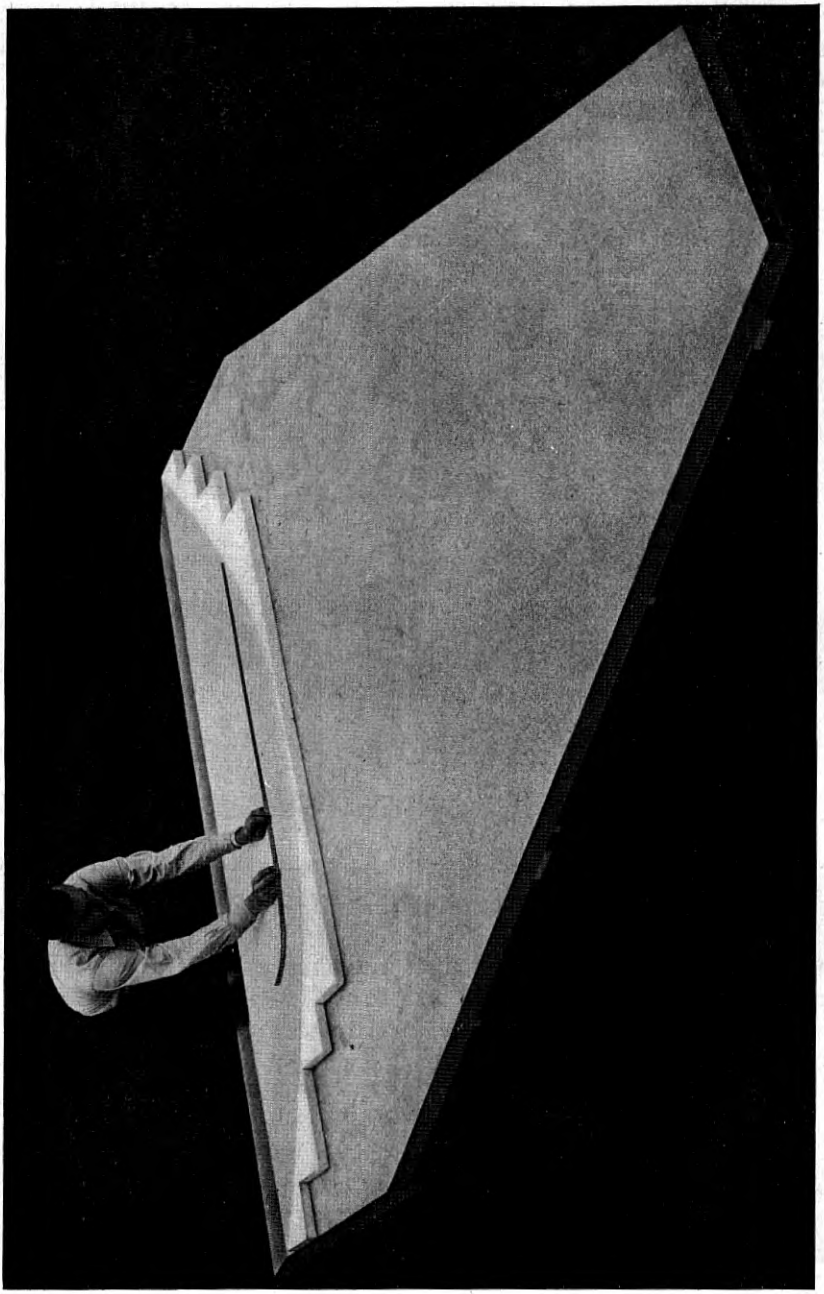


Fig. 7—Inserting metal strips into a profile plate of a 10-foot strip lens.

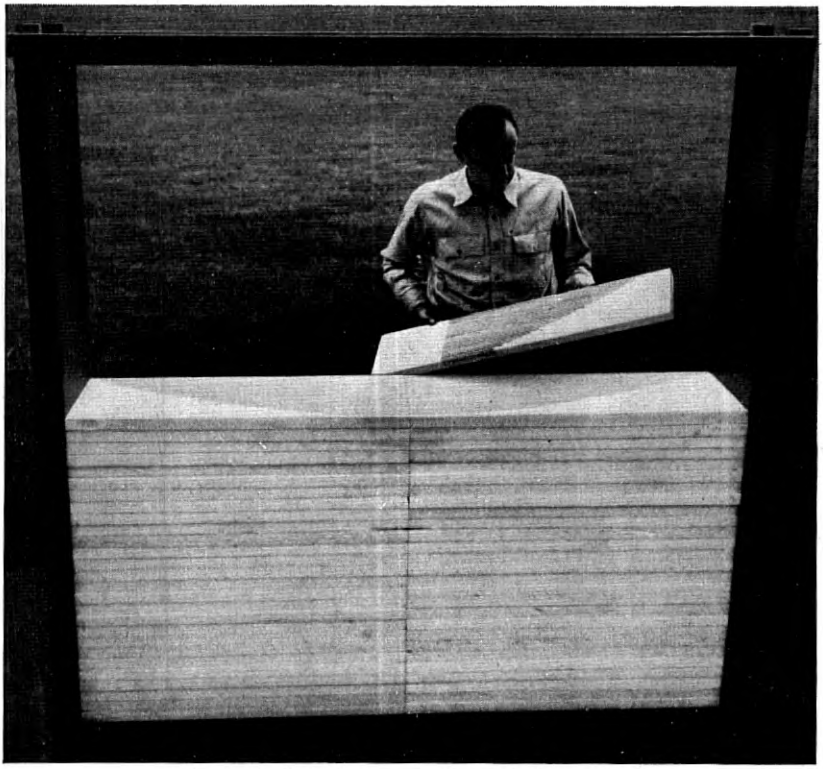


Fig. 8—A six-foot square strip lens half assembled.

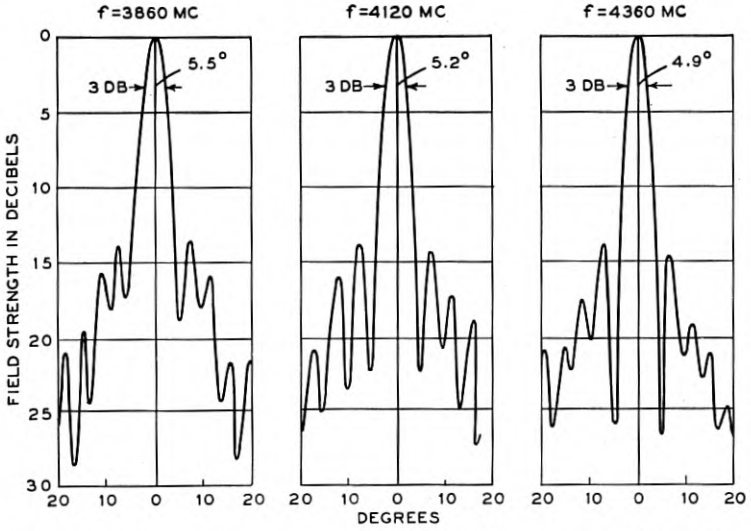


Fig. 9—Directional patterns of a 3-foot diameter lens at several frequencies.

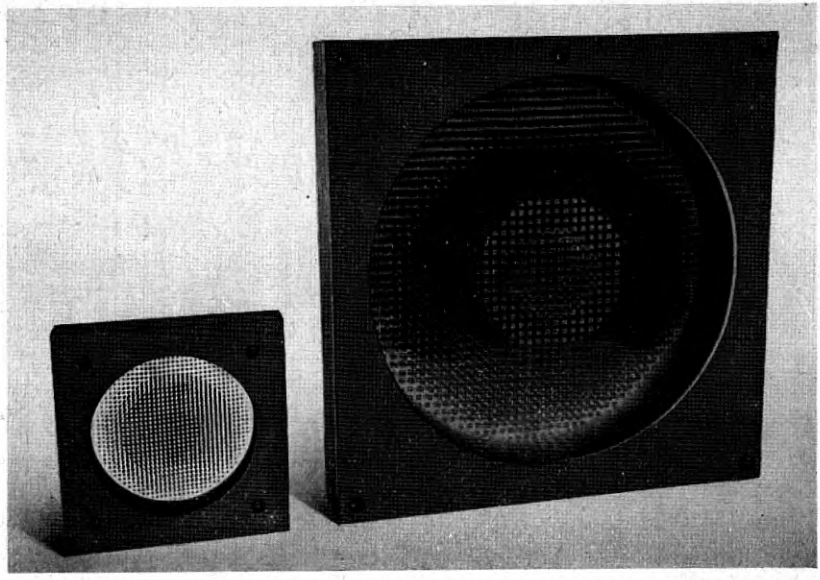


Fig. 10—Lenses constructed by spraying conducting paint on cellophane sheets.

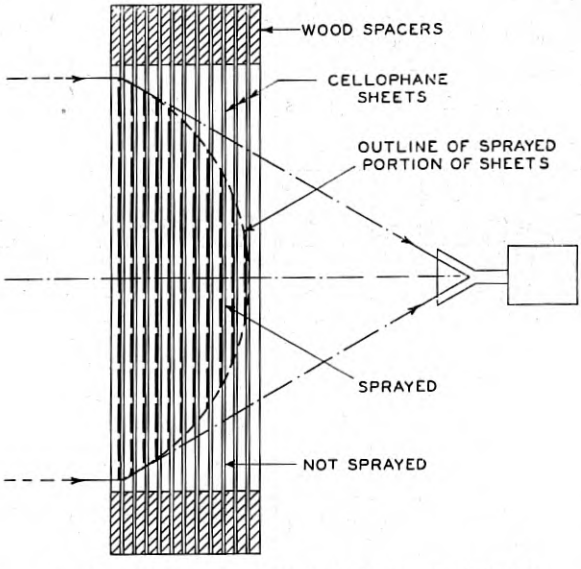


Fig. 11—Construction details of the lenses of Fig. 10.

was made by affixing copper foil strips to cellophane sheets and stacking the sheets with no spacing other than the sheets themselves between them.

This extremely close spacing and the staggered arrangement of strips correspond to a heavy capacitative loading and introduced so much delay per unit length that the measured effective dielectric constant of this lens proved to be 225 (i.e. the index of refraction was 15). Because of such a high dielectric constant the reflection losses at the surface of this lens are high,

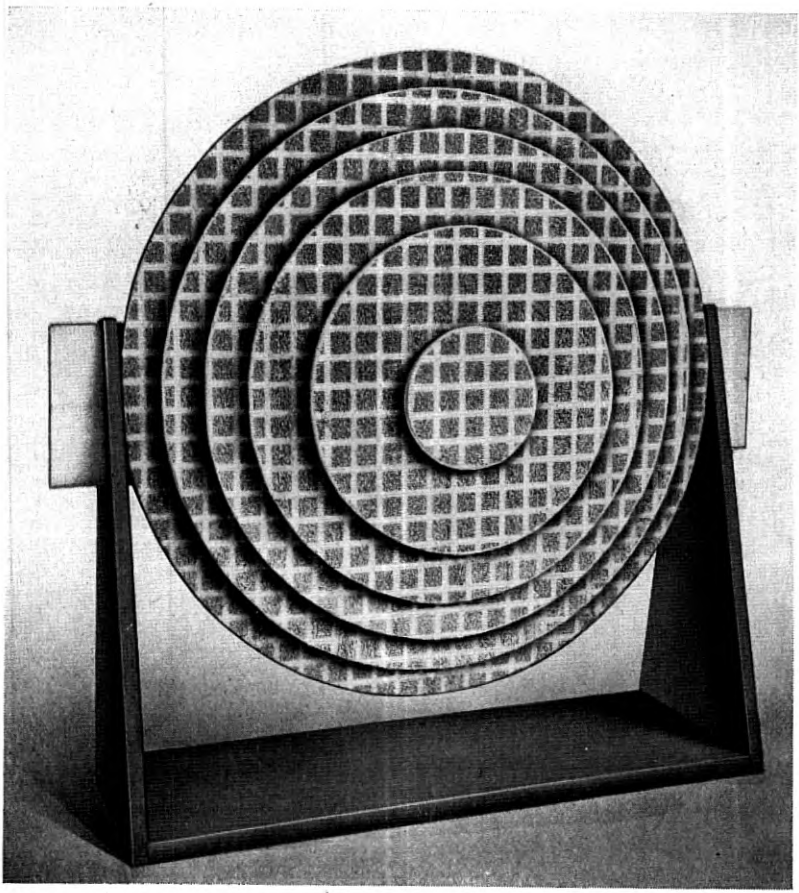


Fig. 12—Lens formed by metal-spraying tin directly onto polystyrene foam sheets.

so that, for efficient operation, tapered or quarter wave matching sections on each surface would be necessary.

(e) *Frequency Sensitive Delay Lenses*

When the conducting elements approach a half wavelength in their length parallel to the electric vector, resonance effects occur and the artificial dielectric behaves like an ordinary dielectric near its region of anomalous

dispersion.[5] The index of refraction of an artificial dielectric using $\frac{3}{4}''$ metallic elements would increase rapidly as the wavelength approached $1\frac{1}{2}''$ until, at $\lambda = 1\frac{1}{2}''$ the dielectric would be opaque. At still lower wavelengths, the material would appear to have an index of refraction less than one.

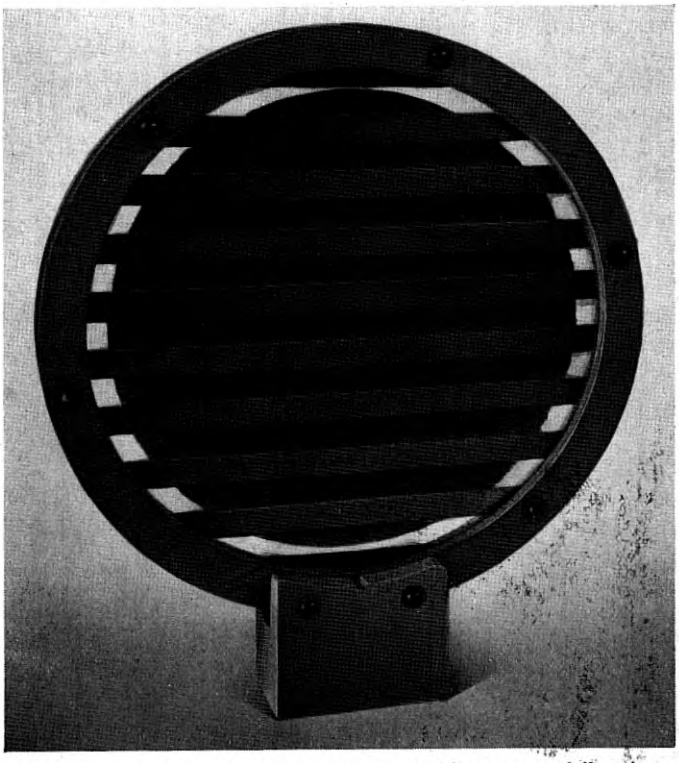


Fig. 13—Closely spaced strip construction comprising copper foil strips affixed to cellophane sheets. Juxtaposition of the sheets yielded an effective index of refraction of 15.

Elements, such as rods, which are $\lambda/2$ long, have a very broad resonance band, and the region of anomalous dispersion in a dielectric utilizing such

[5] Anomalous dispersion occurs in optical substances when the frequency of the incident light wave approximates one of the vibrational resonance frequencies of the molecule. On the long wavelength side of this resonance region the index of refraction is greater than one and increases rapidly as the resonance wavelength is approached. Dispersion, which is the change of index of refraction with frequency, is therefore very high, but it is the "normal" type of dispersion. At resonance, the absorption of the wave is high and the substance becomes opaque. At still shorter wavelengths, the resonance phenomenon acts to make the index of refraction less than its long wavelength value, often less than one, and the index again varies rapidly with wavelength, but because this is an "abnormal" type of dispersion, it is called the region of anomalous dispersion.

rods is very large, i.e., the dispersion is not very great. If it is desired to have a highly dispersive material, this band can be considerably reduced by the process of tilting the rods so that they are more nearly perpendicular to the electric vector. They thereby become "loosely coupled" to the incident wave and acquire a higher Q. Some unsymmetrical arrangement such as that shown in Fig. 14 is necessary to insure that the radiation damping of the elements is actually reduced, since a uniform tilt of all the elements would allow the array to radiate, unhindered, a wave polarized parallel to the elements. Measurements of the index of refraction of a dielectric made up of successive arrays of rods arranged as in Fig. 14 are

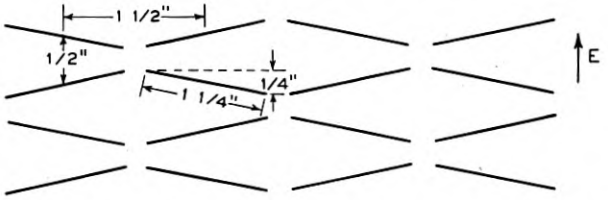


Fig. 14—Array of metal rods to produce a narrow dispersion band.

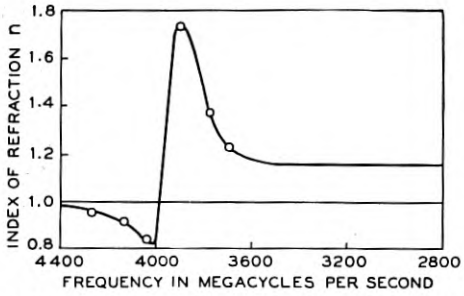


Fig. 15—Measured index of refraction of a metallic dielectric comprising the tilted rods of Fig. 14.

given in Fig. 15, and one observes a marked similarity to the behavior of the index of refraction of dielectrics in the region of their anomalous dispersion. Because of the rapid change of n with wavelength, such material may be useful as a means of separating narrow band radio channels by the use of prisms or by lenses having several feed horns located at the optimum focal points for the various frequency bands involved.

Microwave Repeater Antenna

For radio relay applications, there are three electrical characteristics which antennas should possess. The first is high gain (effective area), as this will reduce the path loss and accordingly the requirements on trans-

mitter power. The second is good directional qualities so as to minimize interference from outside sources and also interferences between adjacent antennas. The third is a good impedance match so that reflections between the antenna and the repeater equipment will not distort the transmitted signal. These characteristics should preferably be attainable without the imposition of severe mechanical or constructional requirements.

The shielded lens type of antenna (lens in the aperture of a horn) lends itself well to repeater work because of its moderate tolerance requirements, its good directional properties associated with the shielding, and the desirable impedance characteristic obtainable by tilting the lens. The delay lens offers the additional advantage of broad band performance with the consequent possibility of operating on several bands widely separated in wavelength. In this section, construction details and performance of a 6-foot square aperture strip type delay lens will be discussed.

(a) *Design of the Artificial Dielectric*

The operating frequency band envisioned for this antenna was 3700 to 4200 megacycles (λ = 7.15 to 8.1 cm), and to keep the element length sufficiently well removed from the half-wave resonant length a value of $\frac{3}{4}''$ for the strip width was chosen. The index of refraction of solid polystyrene is approximately 1.61 and this introduces a reflection loss (mismatch loss) at each surface of 0.225 decibels. To reduce this loss to 0.18 db. the artificial dielectric was designed to have an n of 1.5 as this still did not impart too great a thickness to the lens. A combination of strip spacings which yields an n of 1.5 was found to be $\frac{3}{8}''$ in the horizontal direction and $1\frac{5}{16}''$ center to center spacing in the vertical direction as shown in Fig. 16. The construction method of Fig. 6 was used which involved inserting .002″ copper strip into slots cut in foam sheets.

(b) *Lens Design*

Several lens shapes were possible: (1) bi-convex, (2) plano-convex with the flat side toward the feed, and (3) plano-convex with the curved side toward the feed. For a given thickness and therefore weight of lens, the third possibility produces the shortest over-all structure of lens plus horn feed, and it was accordingly selected. The curved side is then a hyperboloid of revolution as shown in Fig. 17 and for the chosen focal length of 60″, the profile, as calculated from the equation shown for n = 1.5, reaches a maximum thickness of 16″. To eliminate the reflection from the lens into the feed, a lens tilt could have been employed. It was found that a quarter wave offset of one half of the lens relative to the other half could accomplish this same purpose, because reflected rays from one half of the lens then undergo a one half wavelength longer path in returning to the feed and the reflections from the two halves cancel. As this process, however, is completely effective only at one frequency, the final lens design employed both a

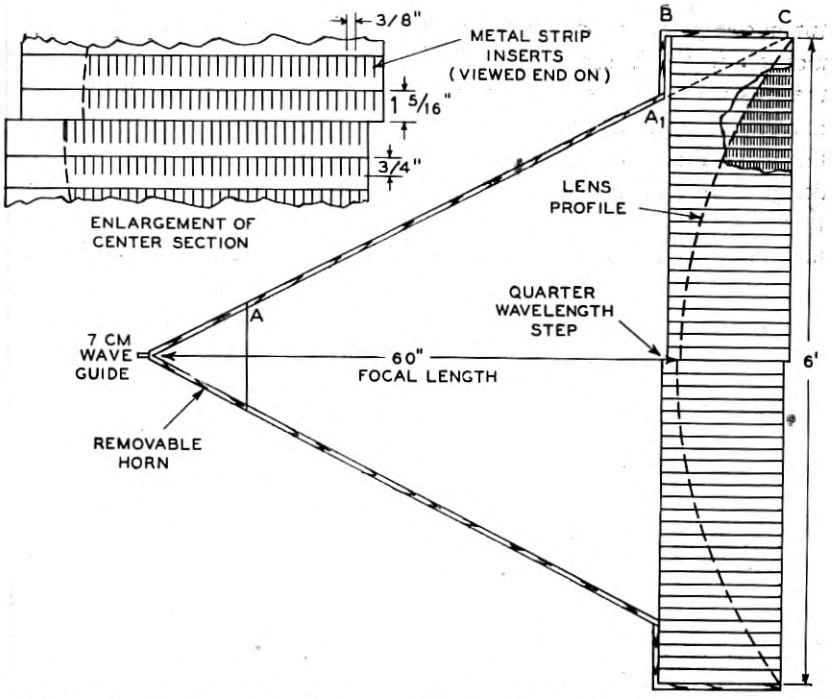


Fig. 16—Construction of a six-foot square shielded delay lens for repeater application. The wooden horn has a metal interior surface.

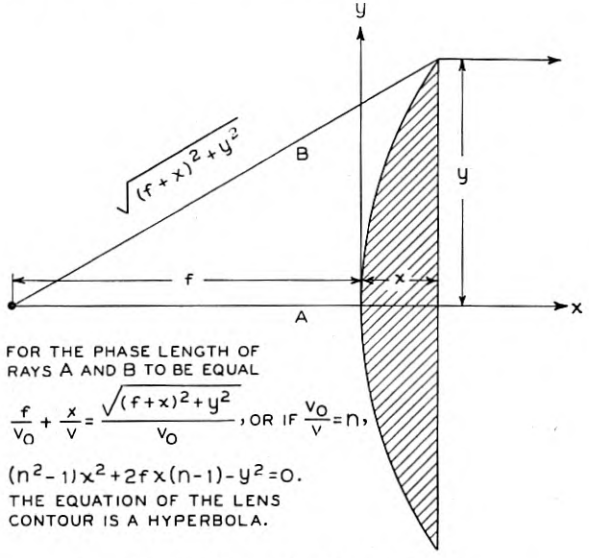


Fig. 17—Profile equation for a delay lens.

tilt in one plane and a quarter wave offset in the other plane. To utilize most efficiently the space afforded the antenna on the top of the relay tower, the lens aperture was made square. Since an unstepped lens has, by its nature, a circular aperture, the four corners of the horn aperture must be filled in with lens material as sketched in Fig. 18. The step height is designed for midband wavelength and in the present case follows the equation of Fig. 17 with the focal length reduced from the value used for the main lens section by $K\lambda/(n-1)$ where K is an integer. That integer K is

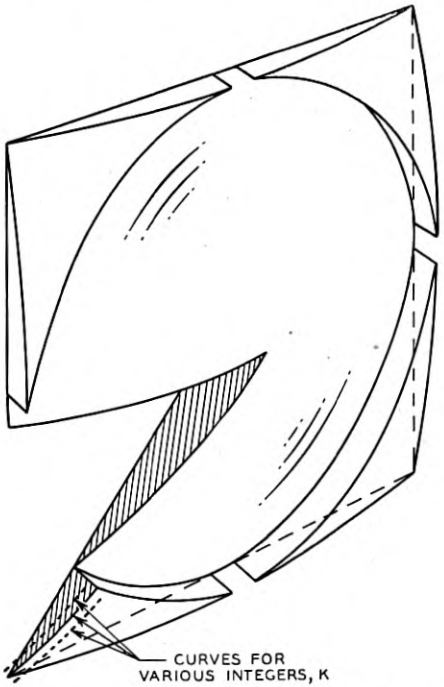


Fig. 18—Filling in the four corners of a square horn aperture with lens material.

selected which brings the step profile nearest the horn corners as indicated in Fig. 18.

(c) *Horn Shield*

Lenses can be energized by means of a small feed horn placed at the focal point, but to obtain the best directional properties a full metallic shield extending from the wave guide feed up to the sides of the lens should be employed. The use of foam slabs in the construction of the delay lens prevents this shield from extending completely up to the lens as indicated by the extension of A-A_1 (shown dotted) in Fig. 16. Optical formulas

indicate, however, that for such large apertures, the amount of diffraction (spreading of the waves outside of the pyramid formed by the dotted lines) will be small, and it is only necessary that the sides AA_1, A_1B and BC all be conducting to insure good back lobe suppression and other desired properties associated with a horn shield.

(d) *Performance*

The gain of this antenna over an isotropic radiator is plotted in Fig. 19. The top curve is the theoretical gain of a uniform current sheet of the same area (6′ x 6′), the lower curve the gain of a 6′ x 6′ area having 60% effective area. The points, which fall approximately on the lower curve, are the experimentally measured gain values of this antenna at the frequencies

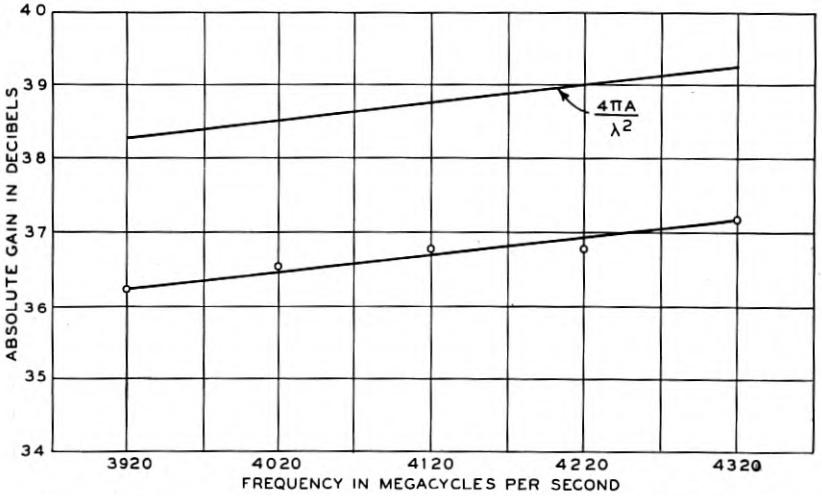


Fig. 19—Measured gain characteristics of the six-foot square shielded lens of Fig. 16. The lower line indicates 60% effective area and the circles are experimental points.

indicated. The constant percentage effective area indicates that the index of refraction of the lens material remains quite constant over the indicated frequency band. In contrast to this, the 10′ x 10′ metal plate lenses[1] ($n < 1$) exhibit, at the band edges, a falling off of $1\frac{1}{2}$ decibels from midband gain for a 10% wavelength band.

The magnetic plane pattern of the lens when fed by a 6-inch square feed horn is shown in Fig. 20. The remarkable symmetry of the minor lobes in Fig. 20 and also in Fig. 9 shows that the phase fronts of the waves radiated by these antennas are very accurately flat. This result emphasizes the tolerance advantages of the lens over the reflector. It is believed that the

[1] Loc. cit.

further improvement in pattern of this lens over previous wave guide type metal lenses is attributable to the absence of steps, as they tend to introduce diffraction effects. The symmetry of the pattern also indicates a high degree

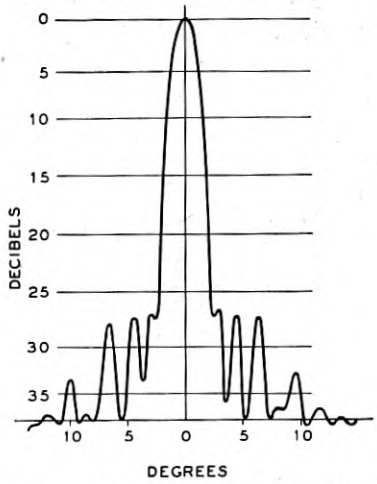


Fig. 20—Directional pattern of the lens of Figs. 8 and 16 when fed with a six-inch square electromagnetic horn.

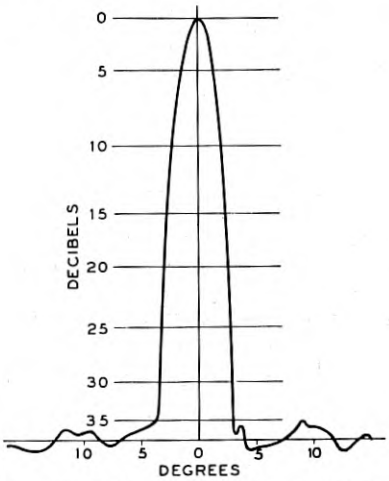


Fig. 21—Directional pattern of the same lens when enclosed with a full horn shield.

of homogeneity of the dielectric (the lens is 16 inches thick); this is a property not always shared by ordinary dielectrics such as polystyrene.

When a full horn shield is used, the illumination across the aperture has a very strong taper (somewhat stronger than a cosine taper because of the wide flare angle of the horn). This results in a very effective suppression of the close-in side lobes as shown in Fig. 21. In the vertical plane, the side

lobes are not very well suppressed because the illumination is only slightly tapered. However, for repeater work, lobes in the vertical plane are not as objectionable as lobes in the horizontal plane because interfering signals generally originate from other stations lying in the horizon plane.

The impedance match of the shielded lens antenna is affected by the discontinuity of the wave guide at the expanding horn throat and by whatever energy is reflected back into the feed line from the lens. By tuning means, the throat mismatch can be held to less than 0.2 db SWR (1.02 V.S.W.R.) over the 500 mc band and by a combination of lens tilt and quarter-wave step in the lens the SWR due to the energy reflected from the lens was also held to less than 0.2 db over the band.

PART II—THEORETICAL CONSIDERATIONS

We turn now to a consideration of the electromagnetic theory underlying the operation of artificial dielectrics. If the polarizability of the individual conducting element employed is known, equation (2) will permit a calculation of the effective dielectric constant. Before proceeding to this, however, a brief review of dipole moment, dipole potential, and dielectric polarization will be given (MKS units).

Dipole Moment, Potential and Polarization

Two charges, $+q$ and $-q$, displaced a small distance from one another, constitute an electric dipole. If the vector joining them is called $\boldsymbol{ds}$, the dipole moment is defined as

$$\boldsymbol{m} = q\, \boldsymbol{ds}. \tag{5}$$

The potential V, at any point, due to a point-charge q is defined as

$$V = q/4\pi\epsilon r, \tag{6}$$

where r is the distance from the point in question to the charge q. The potential V due to a dipole of moment $\boldsymbol{m}$ is

$$V = \frac{|m|}{4\pi\epsilon r^2} \cos\theta, \tag{7}$$

where θ is the angle between $\boldsymbol{r}$ and $\boldsymbol{m}$.

A conducting object becomes polarized when placed in an electric field. Its dipole moment $\boldsymbol{m}$ depends upon the field strength $\boldsymbol{E}$ and upon its own polarizability α:

$$\boldsymbol{m} = \alpha \boldsymbol{E}. \tag{8}$$

If there are N elements per unit volume, the polarization $\boldsymbol{P}$ of the artificial dielectric is

$$\boldsymbol{P} = N\alpha \boldsymbol{E}. \tag{9}$$

But $\boldsymbol{P}$ is related to the displacement vector $\boldsymbol{D}$ and the dielectric constant as follows:

$$\boldsymbol{D} = \epsilon \boldsymbol{E} = \epsilon_0 \boldsymbol{E} + \boldsymbol{P}, \tag{10}$$

so that $\epsilon = \epsilon_0 + N\alpha$ which is equation (1). A knowledge of α thus permits a determination of ϵ. We obtain α from (8) by first finding the dipole moment $\boldsymbol{m}$ of the particular shape of element when immersed in a uniform field $\boldsymbol{E}$.

CALCULATION OF DIELECTRIC CONSTANTS OF ARTIFICIAL DIELECTRICS[6]

(1) *Conducting Sphere*

Consider a perfectly conducting sphere immersed in an originally uniform field of potential

$$V = -Ey = -Er \cos \theta. \tag{11}$$

The free charges on the sphere are displaced by the applied field and it thereby becomes a dipole whose moment $\boldsymbol{m}$ we wish to determine. The external potential field is the sum of the applied potential and the dipole potential, and from (7) and (11) we have

$$V_{out} = -Er \cos \theta + \frac{m \cos \theta}{4\pi\epsilon_0 r^2}. \tag{12}$$

The internal field is zero because the sphere is conducting. At a boundary between two dielectrics, there is the requirement[7]

$$V_{\text{outside}} = V_{\text{inside}}. \tag{13}$$

Equation (13) gives, at $r = a$ (the radius of the sphere),

$$-Ea \cos \theta + \frac{m \cos \theta}{4\pi\theta_0 a^2} = 0, \tag{14}$$

or

$$m = 4\pi\epsilon_0 E a^3, \tag{15}$$

the dipole moment of the sphere. From (8) we see that the polarizability of the conducting sphere is accordingly $4\pi\epsilon_0 a^3$, from which equation (3) follows.

(2) *Magnetic Effects of a conducting sphere array*

The above calculations on a conducting sphere assume an electrostatic field. At microwaves, the rapidly varying fields induce eddy currents on the surface of the sphere which prevent the magnetic lines of force from penetrating the sphere. The magnetic lines are perturbed as shown in

[6] The author is indebted to Dr. S. A. Schelkunoff for the polarizability formulas given in this memorandum.

[7] Smythe, "*Static and Dynamic Electricity*", McGraw-Hill, 1939, p. 19.

Fig. 5. Now a conducting sphere in an alternating magnetic field is equivalent to an oscillating magnetic dipole,[8] and we should observe an effective permeability for a sphere array.[9] The magnetic dipole field is, however, opposed to the inducing field, in other words, the dipole moment is negative. Smythe[7] shows that the magnetic polarizability of a conducting sphere of radius a in a high frequency field is

$$\alpha_m = -2\pi\mu_0 a^3. \tag{16}$$

The effective relative permeability of an array of N spheres per unit volume is therefore

$$\mu_r = 1 - 2\pi a^3 N. \tag{17}$$

The index of refraction was given in equation (4) as the square root of the dielectric constant. This is strictly true only when the permeability of the dielectric is unity. We have seen above, however, that the sphere array at microwaves possesses an effective permeability given by (17), and therefore (4) is not valid. The correct expression for n is

$$n = \sqrt{\mu_r \epsilon_r}, \tag{18}$$

and the effective refractive index of a sphere array is accordingly

$$n = \sqrt{(1 - 2\pi N a^3)(1 + 4\pi N a^3)}, \tag{19}$$

which is smaller than that given by (3) and (4). The disk and strip arrays, besides being lighter, avoid the diminishing effect on the refractive index caused by the perturbing of the magnetic lines.

(3) *Conducting Circular disk*

The determination of the dipole moment of a disk involves the use of ellipsoidal coordinates and will not be carried through here. Smythe[7] gives an expression for the torque on a flat disk of radius a in a uniform field in Gaussian units. In M.K.S. units, his formula becomes

$$T = 4\pi\epsilon_0 \frac{2a^3 E^2 \sin 2\theta}{3\pi} = 4\pi\epsilon_0 \frac{2a^3 E^2 2 \sin\theta\cos\theta}{3\pi}$$

$$= \left[\frac{16a^3\epsilon_0}{3}(E\sin\theta)\right][E\cos\theta]. \tag{20}$$

The first bracket represents the dipole moment, the second the field, and the product gives the torque. When the plane of the disks is parallel to E $\sin\theta$ is one, and from $\boldsymbol{m} = \alpha\boldsymbol{E}$, we have

$$\alpha = \frac{16\epsilon_0 a^3}{3}. \tag{21}$$

[8] T. S. E. Thomas, *Wireless World*, Dec. 1946, p. 322.
[9] L. Lewin, *Jour. I. E. E.*, Part III, Jan. 1947, p. 65.
[7] Loc. Cit. Eq. 14, p. 397.
[7] Ibid, eq. 7, p. 163.

so that, from (2),

$$\epsilon_r = 1 + \tfrac{16}{3} N a^3 \tag{22}$$

(4) *Strips*

The calculation of the dipole moment of a thin conducting strip as used in the strip lens of Fig. 15 involves two dimensional elliptic coordinates and will also be omitted here. Again, the torque is given in Smythe[7] for an elliptic dielectric cylinder, from which we obtain,

$$\alpha = \frac{\pi \epsilon_0 s^2}{4}, \tag{23}$$

where s is the strip width, so that

$$\epsilon_r = 1 + \frac{\pi}{4} s^2 n, \tag{24}$$

where n is the number of strips per sq. unit area looking end on at the strips.

(5) *Validity of the polarizability equations*

Equation (2), which expresses the dielectric constant to be expected from an array of N elements each having a polarizability of α, was derived (by (8), (9) and (10)) by assuming that the field acting on an element, and tending to polarize it, was the impressed field $\boldsymbol{E}$ alone. This is a satisfactory assumption when the separation between the objects is so large that the elements themselves do not distort the field acting on the neighboring elements. Such is not the case when the value of ϵ_r exceeds 1.5 or thereabouts. For the usually desired values of ϵ_r of 2 or 3 it is thus seen that the above formulas such as equations (22) and (24) will yield only qualitative results and that the exact spacings of the elements to produce a desired refractive index will have to be determined by experimental methods.

For lattices having 3-dimensional symmetry, an improvement over equation (2), which takes into account not only the impressed field $\boldsymbol{E}$ but also the field due to the surrounding elements, is the so-called Clausius-Mosotti equation:

$$\frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} = \frac{N\alpha}{3\epsilon_0} \tag{25}$$

This, along with a similar expression for the permeability [replacing (17)], would permit a fairly accurate determination from (18) of n for the conducting sphere array.

[7] Ibid, eq. 6, p. 97.

Resonance Effects

It was stated earlier that the size of the elements should be small relative to a half wavelength in order for the refractive index to be independent of frequency. A qualitative idea of this criterion can be obtained by an elementary analysis of forced oscillations of dipoles. It is known that a dielectric medium which is composed of elements that resonate under the action of an alternating electric field, such as atoms having bound electrons, will exhibit a dielectric constant which varies with frequency:[10]

$$\epsilon_r = 1 + \frac{k}{f_0^2 - f^2}, \tag{26}$$

where f_0 is the frequency of resonance of the element, f the frequency of the incident radiation and k a proportionality constant. Thus when f is small relative to f_0, ϵ_r is practically independent of f.

As a means of estimating the change in refractive index with frequency of metal delay lenses, we consider a specific example: Let $n = 1.50$ when the elements are $\lambda/4$ in length, i.e. when $f^2 = \frac{1}{4}f_0^2$, then the last term of (26) equals 1.25. Decreasing f by 20% reduces n from 1.50 to 1.46. Thus the change in n from midband to the edges of a $\pm$ 10% band is about .02. From this, at 7 cm, the phase front, even for a lens 30″ thick, should remain plane to within $\pm\frac{1}{10}$ wavelength over this 20% band of wavelengths. If the members had been made $\frac{1}{8}$ wavelength long, the variation in n from the design frequency all the way down to D.C. would have amounted to only 1.2%.

Summary

A *metallic dielectric* is constructed by arraying conducting elements in a three-dimensional lattice structure. For electromagnetic waves whose wavelength is long compared to the size and spacing of the elements, this structure displays an effective dielectric constant and index of refraction which is sensibly constant over wide frequency bands. Lenses can be designed according to these principles which will focus microwaves and longer radio waves as a glass lens focusses light waves. Such lenses have the advantage of broad-band performance over the earlier waveguide type metal lenses and they retain the advantages of light weight over dielectric lenses. As microwave antennas, they are superior to parabolic dish reflectors from the standpoint of warping and twisting tolerance, profile tolerance, directional properties and impedance match. By eliminating the steps in the lens, the directive patterns are made cleaner and an increase in absolute

[10] See for example, Joos, *Theoretical Physics*, Blackie & Son, Book 4, Chapter 4.

gain results. Because of the broad-band properties of the artificial dielectric, this improved gain can be maintained over a very wide band of frequencies. The lenses can be built to focus waves of any polarization and, if desired, the dielectric can be designed to exhibit strong dispersion. Theoretical calculations of the expected dielectric constant are in fairly good agreement with experiment for values less than 1.5; for higher values an accurate determination of the true value must be obtained experimentally.

ACKNOWLEDGEMENT

The author wishes to express his appreciation to Dr. S. A. Schelkunoff for his advice and consultation, and to the members of the Holmdel Radio Research Laboratory, especially to Mr. William Legg, for assistance and cooperation during the course of this work.

A Non-reflecting Branching Filter for Microwaves

By W. D. LEWIS and L. C. TILLOTSON

Microwave branching filters are required as integral parts of multi-channel microwave radio relay systems. These filters must have characteristics which are difficult to attain if one attempts to extend familiar lower frequency techniques to the microwave region. A novel network configuration, through which currently anticipated requirements can be met without excessive difficulty, is described in this paper.

In this configuration individual constant resistance channel dropping units are formed of appropriate assemblies of two hybrid circuits, two band reflection filters and two quarter wavelengths of line. An assembly of N channel dropping units in cascade then forms an N channel constant resistance branching network.

The mechanical and electrical characteristics of a practical five channel branching filter of this type are described. As a result of experience with this prototype filter it can be stated with some safety that these requirements can be fulfilled with a network of this type. Experimentally observed impedance, insertion loss and phase characteristics were fully satisfactory. In addition the circuit appears to be flexible enough both electrically and mechanically to fulfill the various types of systems needs which may be encountered at branch points or when channels must be added or interchanged.

Introduction

PRESENT plans for point-to-point communication by means of microwave radio relay systems call for the operation of several radio channels between each pair of repeaters. A proposed frequency plan for the 4000 mc common carrier band (3700 to 4200 mc) specifies channels 20 mc wide spaced 40 mc center to center. A possible arrangment for a five-channel radio repeater station is illustrated in Fig. 1. This arrangement is calculated to utilize the available frequency space in an efficient and technically sound manner.

If this channel disposition is to be achieved without a costly increase in the number of antennas and the size of the supporting towers, radio frequency branching networks must be provided which connect the individual transmitting or receiving circuits at each repeater point to a common antenna (Fig. 1). If this connection is to be made without excessive loss of power these branching devices must have adequate adjacent channel rejection, low ohmic loss, and good impedance match in the channel bands. An excellent impedance match is especially desirable if circuit disturbances resulting from echoes in the long waveguide lines which lead from the filter assemblies to the antennas are to be minimized (Fig. 1).

Since the type of microwave radio repeater now planned[1] obtains most of its gain at intermediate frequencies, the IF amplifiers will reject all spurious

[1] See H. T. Friis, "Microwave Repeater Research", to appear in the April 1948 issue of *B. S. T. J.*

signals entering the receiver except those in the vicinity of the receiver image bands. Because of this, suppression requirements on the branching filters, except possibly in the vicinity of receiver image bands, are not severe.

Problems connected with the design of suitable microwave branching filters naturally differ considerably from previous filter problems. The discrimination and band utilization requirements are readily met, but the

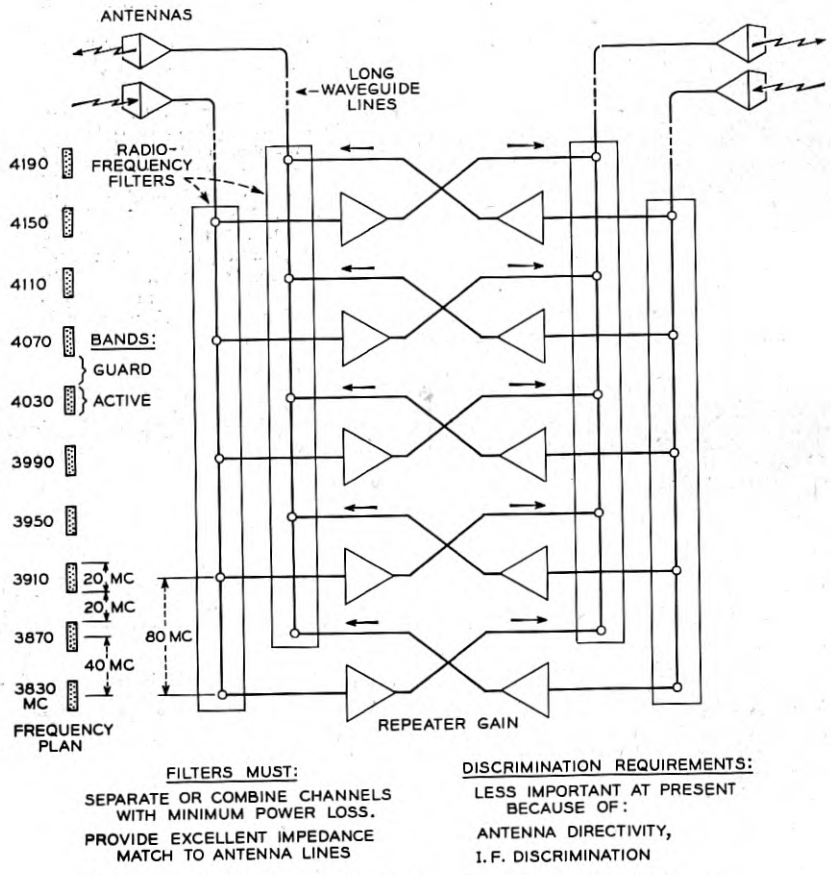


Fig. 1—Possible five-channel radio repeater station schematic diagram.

impedance control required is not easy to obtain by familiar filter techniques. Not more than about 5% variation in impedance or about 0.5 db standing wave ratio in the channels could be tolerated. At lower frequencies channel passing networks which can be connected in series or parallel to form a channel branching filter can be designed on the basis of lumped circuit theory and built of coils, condensers and resistances, but in the microwave region

simple parallel or series connections and simple lumped circuit elements do not exist. Although in the interests of flexibility it would be desirable to add or substitute individual channels without affecting other channels in a branching filter, the convenient possibility of doing so at a high impedance level on vacuum tube grids is not yet available in the microwave region.

A satisfactory two-channel waveguide branching filter has been constructed following partially 'classical' methods. This filter, designed for the New York-Boston experimental radio relay system, is composed of two channel-passing cavity filters each connected to a common input line through one arm of an E plane Y junction, the waveguide analogue of a series connection (Fig. 2). This solution, although relatively simple where only two channels are required, becomes quite complex when more than two are involved, since in every channel the sum of the interactions of all the inactive filters on

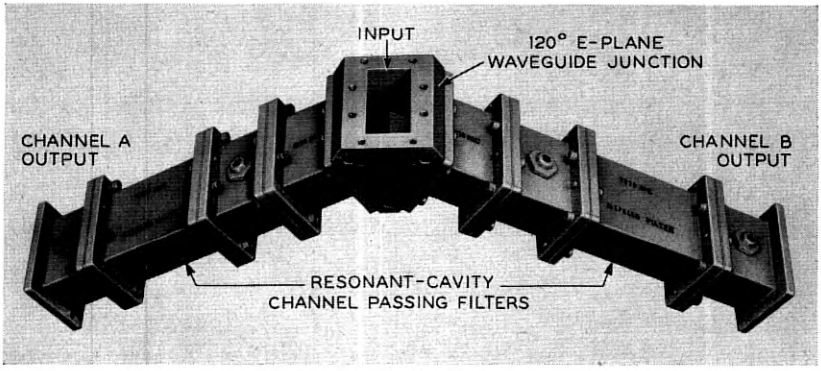


Fig. 2—Branching filter for New York–Boston experimental radio relay system.

transmission through the active filter must be zero. It is evidently not easy to satisfy this condition, particularly since in doing so one must take accurate account of the change with frequency of the effective phase shift of all waveguide connecting lines. And even if such a solution were found it would be valid for only one set of channels, so that the problem must be solved all over again for every change in channel arrangement.

As a result of these difficulties and after a few attempts to overcome them, it became apparent that a more flexible method of microwave filter construction should be found. Constant resistance filters, which provide discrimination by absorbing or diverting the unwanted incident waves rather than by reflecting them, can be useful in any frequency range. In the microwave region, where the shortest connecting pipe may be many wavelengths long, constant resistance devices are particularly helpful. Accordingly a constant resistance channel-dropping network was devised which

could extract one channel from the line, while permitting others to pass through it without disturbance. Several of these networks were then placed in cascade to make up the required filter.

The Hybrid Channel-Dropping Unit

After several possibilities were considered the constant resistance channel-dropping circuit illustrated schematically in Fig. 3 was selected.[2] This circuit is made up of two hybrid junctions, two identical channel reflection filters tuned to the dropped channel, and two quarter wavelength sections of line.

In order to understand the operation of this circuit, the properties of a hybrid circuit, Fig. 4(a), must first be understood. This circuit, which has

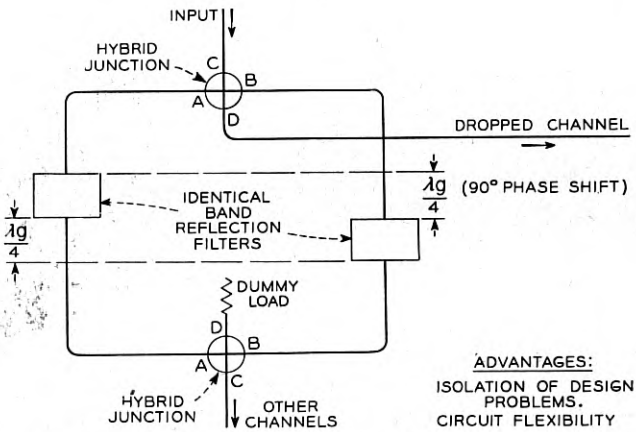


Fig. 3—Possible constant impedance channel dropping filter.

been embodied at voice frequencies as the hybrid coil and at microwaves as the hybrid junction (E-H plane T junction)[3] can be represented schematically as in Fig. 4(b). Let us connect four transmission lines, A, B, C, D, each terminated in its characteristic impedance to the four pairs of terminals of the hybrid circuit. Then if each of these lines is matched to the pair of terminals to which it is connected, the following characteristics will result. Power in transmission line C flowing towards the junction will divide equally into lines A and B, flow away from the junction and be absorbed in loads A and B. None of this power will appear in line D or be reflected back into

[2] Lumped element networks with properties similar to those of this circuit have been devised by Vos and Laurent, U. S. Patent 1,920,041, and Bobis, U. S. Patent 2,044,047. A. G. Fox of these laboratories has independently devised similar microwave circuits.

[3] W. A. Tyrrell, Hybrid Circuits for Microwaves, Proc. I. R. E., Vol. 35, pp. 1294–1306, November 1947.

line C. Similarly, power in line D flowing towards the junction, will appear equally in A and B but not in C or back in D. If these characteristics hold, then by the principle of reciprocity similar characteristics must hold for lines A and B. If proper planes of reference are chosen this behavior can be described in a slightly different, but equivalent, manner. If waves in both A and B flow towards the junction the vector sum of the voltages of these times a constant (0.707) appears in C and the vector difference times the same constant appears in D, but nothing is reflected back into A or B. An equivalent statement can be made if the waves start in C and D.

With the properties of the hybrid in mind, and if it is assumed that the hybrids, the identical reflection filters and the quarter wave lines are perfect and free of ohmic loss, the operation of the circuit of Fig. 3 is easy to understand, and is as follows: A wave entering from arm C of the input hybrid is divided equally into the two arms A and B. None of the power in this wave

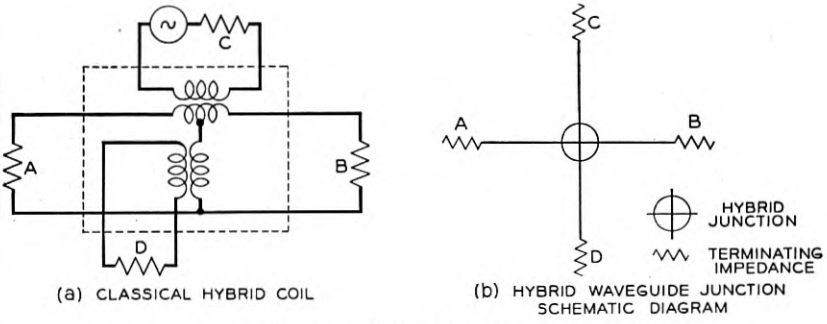


Fig. 4—a. Classical hybrid coil.
b. Hybrid junction schematic diagram.

is reflected back into the arm C or appears initially in arm D. The two equal components of the wave now travel along the lines which are connected to the arms A and B of the input hybrid. If the frequency lies outside the band of the reflection filters the waves travel through these filters and appear in phase in the arms A and B of the output hybrid. The vector sum of these two waves appears in the arm C of the output hybrid and has an amplitude equal to that of the original input wave. Consequently all energy in the input line incident on this network, except that lying in the band of the reflection filters, will pass through it to the output line.

If now the frequency of the input wave lies within the band of the reflection filters, the two equal components of the wave traveling away from the input hybrid will be reflected at the filters and will travel back towards the input hybrid. One of these components must, however, travel twice through an extra quarter wavelength of line, and will therefore be reversed in phase; with

respect to the other component when it reappears at the input hybrid. The two components will consequently combine in the fourth or difference arm of the input hybrid to form a wave equal to the original input wave.

The circuit of Fig. 3 is only one of a general class of hybrid filter circuits. From one viewpoint these circuits resemble spectroscopes, and, from another, ordinary lumped circuits. We could, for example, replace the band reflection circuits of Fig. 3 with any two identical four-terminal networks. This circuit would still retain its constant resistance character. One of its branches would contain all energy reflected by the four terminal networks; the other would contain all energy passed by them.

In particular if the two identical filters are of the channel passing type, input waves within this channel will be transmitted by both filters and will combine at the output hybrid and appear in arm C. All of the other channels will be reflected by the filters and thus will appear in arm D of the input hybrid, provided that the assumed 90° phase shift holds over a band which includes all of the channels.

The particular configuration of Fig. 3 was chosen to minimize the effect of practical limitations. The dropped channel width (20 mc) is only a small fraction of its midband frequency (about 4000 mc). Consequently when band reflection filters are used in Fig. 3, the change with frequency of the nominally quarter wave sections of guide does not seriously affect the performance of the filter and the impedance match of arm D of the hybrids needs to be good only in the vicinity of the dropped channel.

The circuit shown in Fig. 3 is a constant resistance network which drops the channel corresponding to the reflection filter. Several in sequence as indicated in Fig. 5 constitute a constant resistance channel branching filter. In the sections to follow we will give an account of the physical configuration and electrical performance of a branching network designed according to this pattern to meet the radio frequency requirements of a typical practical radio relay system containing five 20 mc radio frequency channels, spaced 80 mc center to center.

The Waveguide Hybrid

In choosing a hybrid configuration which could be successfully used in the network of Fig. 5, both electrical and mechanical requirements were considered. Since frequency f_n passes through $2n-1$ hybrids it is evident that if acceptable overall performance is to be obtained, the balance and impedance characteristics of each hybrid must be excellent. A broad-band balance can be obtained with relative ease by attaching the 'driven' arms (A and B on Fig. 5) symmetrically. Fortunately the strict impedance requirement applies to only one of the two 'driving' arms (C and D), the other being required to transmit only a single channel.

Because of the reentrant nature of the circuit of Fig. 5 it is evidently desirable, if not essential, to employ a hybrid with a convenient mechanical layout. If, for example, a familiar E-H plane junction type of hybrid were used in combination with waveguide filters built in a straight piece of waveguide, twenty-four E or H plane right angle waveguide bends would be required in the construction of each six-channel branching network. To avoid this extra complication and expense a hybrid configuration with 'driven' output arms parallel to the well matched 'driving' input arm was sought.

The hybrid configuration settled upon was the one shown in Fig. 6. Here the electrical and geometrical requirements discussed above are met simultaneously by connecting the driving arm C to the symmetrically located driven arms A and B through a smoothly tapered E plane Y junction. The taper is approximately one half wavelength long in the center of the 3700–4200

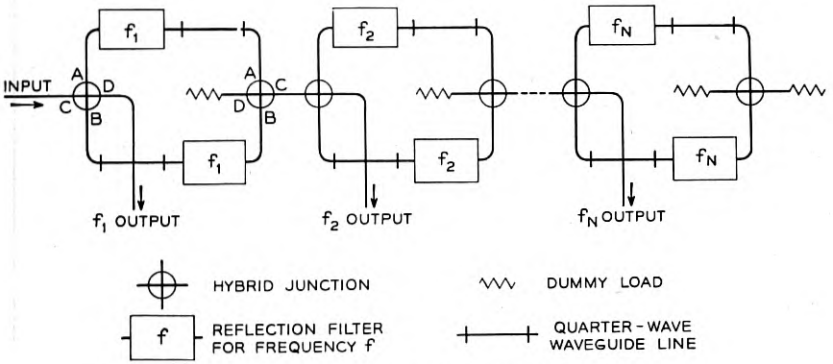


Fig. 5—N Channel branching filter.

mc band. The driving arm D is connected to arms A and B through a coaxial line. This line is coupled to waveguide D by means of a conventional probe. The center conductor traverses the Y junction space in such a way that it is normal to the electric vectors of guides A, B, and C, and is effectively coupled to A and B but not to C by means of a probe P fastened through it normally (Fig. 6).

The Band Reflection Filters

The ideal reflection filter for the circuit of Figs. 3 and 5 would reflect perfectly within a certain band and pass perfectly outside of this band. However, the ratio of bandwidth to band spacing in a given branching filter (20 mc to 80 mc) is such that a sufficiently good approximation to this ideal can be obtained in theory if each reflection filter employs only three resonances, Fig. 7(a). These could be effectively series resonant circuits placed at quarter

wave intervals along the guide, properly distributed in impedance level and all tuned close to the center frequency of the channel to be extracted, Fig 7(b). The practical question was to find how to obtain these resonances in an easily constructed and adequately adjustable form.

Early experiments indicated that a probe inserted in the broad side of the guide far enough so that its end formed an appreciable capacitance with the opposite side could be made to resonate in a series resonant fashion. Impedance levels available through this means of coupling were, however, far lower than required. Accordingly an alternate method in which the probe is

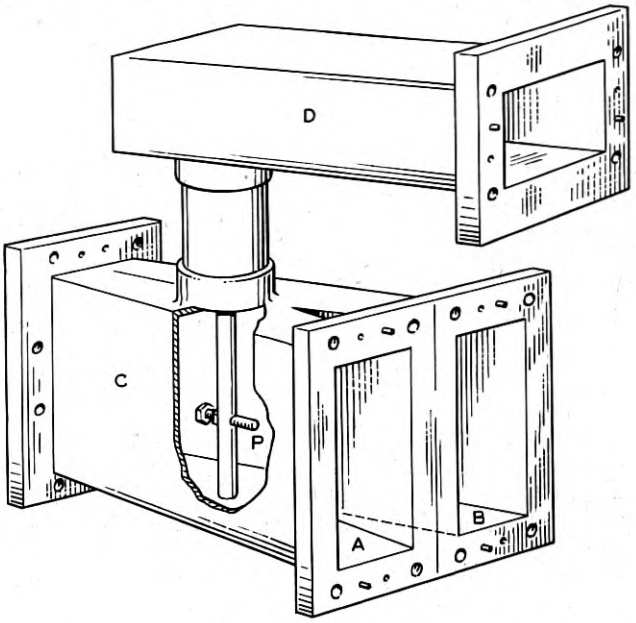


Fig. 6—Hybrid junction.

inserted in the narrow side of the guide was selected (Fig. 8). The probe is made long enough to approach the opposite narrow wall of the guide and is tipped by a capacitative disk. The capacity of the disk and consequently the resonant frequency of the circuit is adjusted by means of a screw in the wall just opposite the disk. Since this rod is inserted perpendicular to the narrow guide wall it is normally uncoupled to the principal mode in the guide. An adjustable coupling is achieved by inserting a screw in the broad side of the guide just above the probe. Insertion of this screw disturbs the symmetry of the field and couples the rod to the guide. Increasing the screw insertion increases the coupling and consequently varies the impedance level of the

equivalent series resonant circuit continuously from infinity down to any value within the range required.

It should be pointed out that impedance and frequency adjustments are not in practice completely independent, but do at least permit the realization of any required value. Furthermore, mutual coupling between the probes interferes with the exact realization of the circuits of Fig. 7. This coupling does not interfere, however, with the realization of satisfactory filter characteristics.

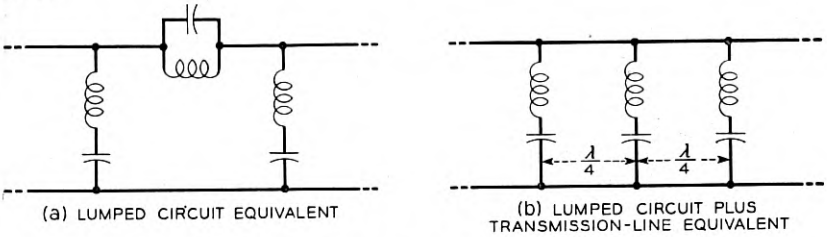


Fig. 7—Band reflection filter.

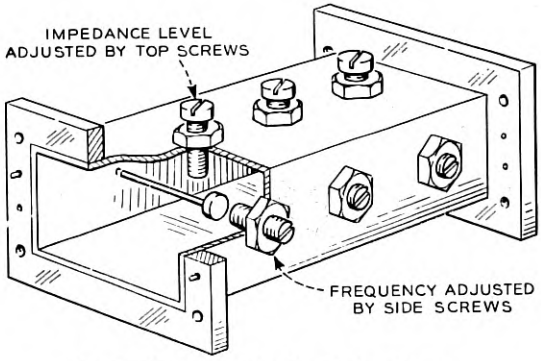


Fig. 8—Waveguide band reflection filter.

When the reflection filter of Fig. 7 was embodied in waveguide form through use of the series resonant circuits just described, the configuration illustrated in Fig. 8 resulted. It was found that, with the exception of a slight change in disk to guide wall spacing, a single configuration could be tuned to any of the desired channels.

Electrical Performance

Individual components as described above were constructed, adjusted and assembled to form a five-channel branching network (Fig. 9). The electrical performance of this network was trimmed systematically, then was measured

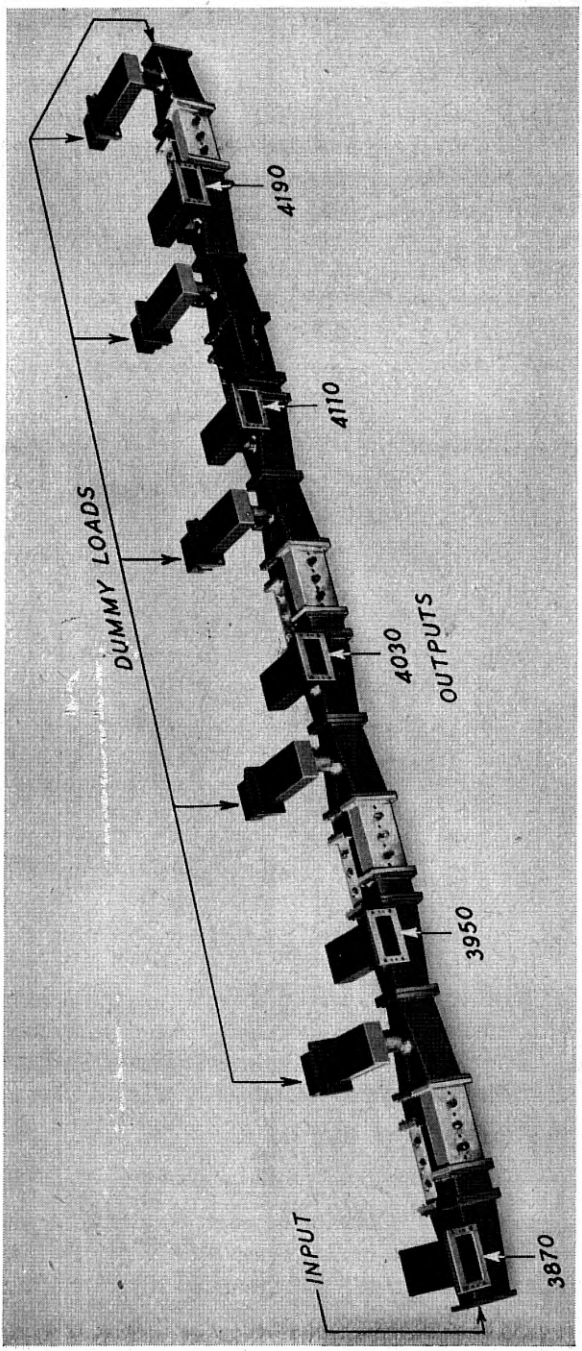


Fig. 9—Five-channel branching filter.

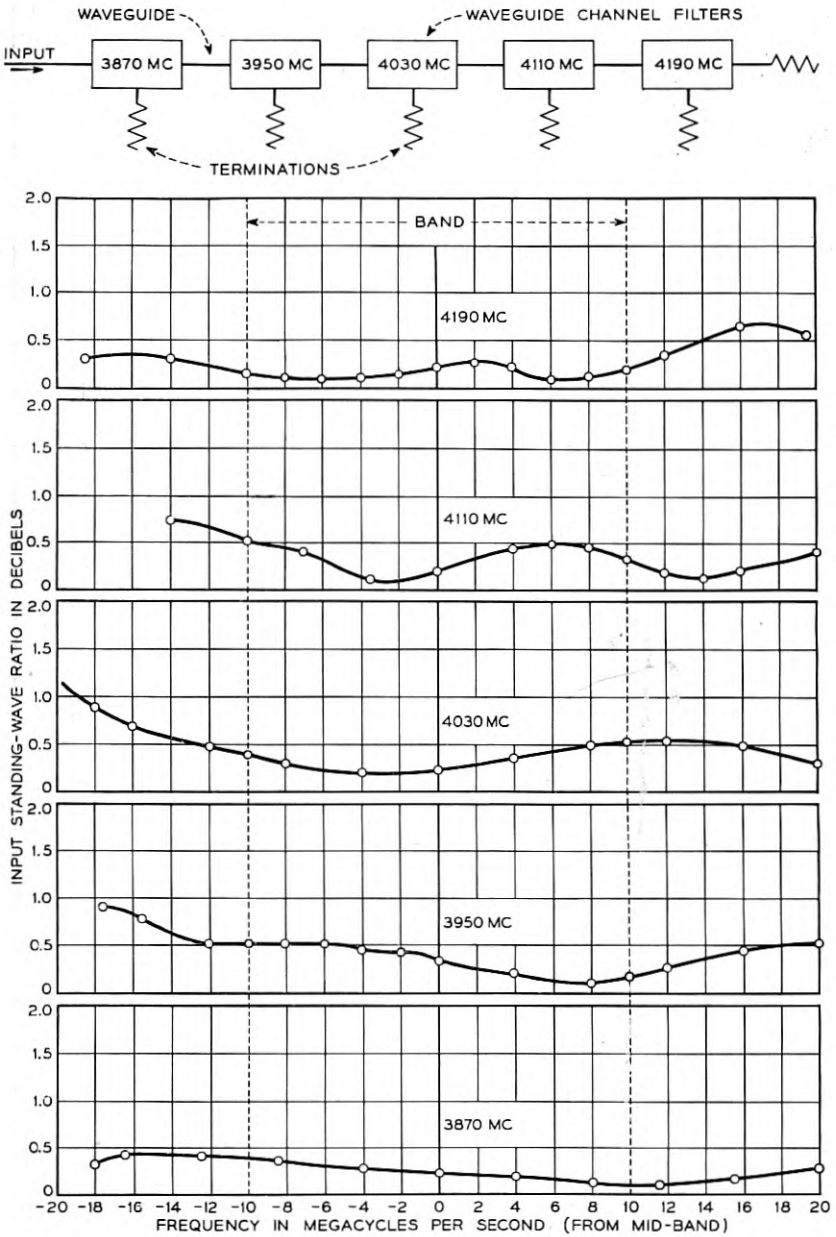


Fig. 10—Input standing wave ratio of hybrid branching filter.

point by point with a double detection measuring set. With all outputs terminated the standing wave ratio observed at the input line was under 0.6 db in

all channels. This quantity is plotted in Fig. 10. The insertion loss measured between the input line and the various output lines varied from about 0.5 db in the lowest frequency channel closest to the input to about 1.0 db in the highest frequency channel farthest from the input. This loss is plotted in Fig. 11. Since 0.5 db was approximately the loss observed in each individual

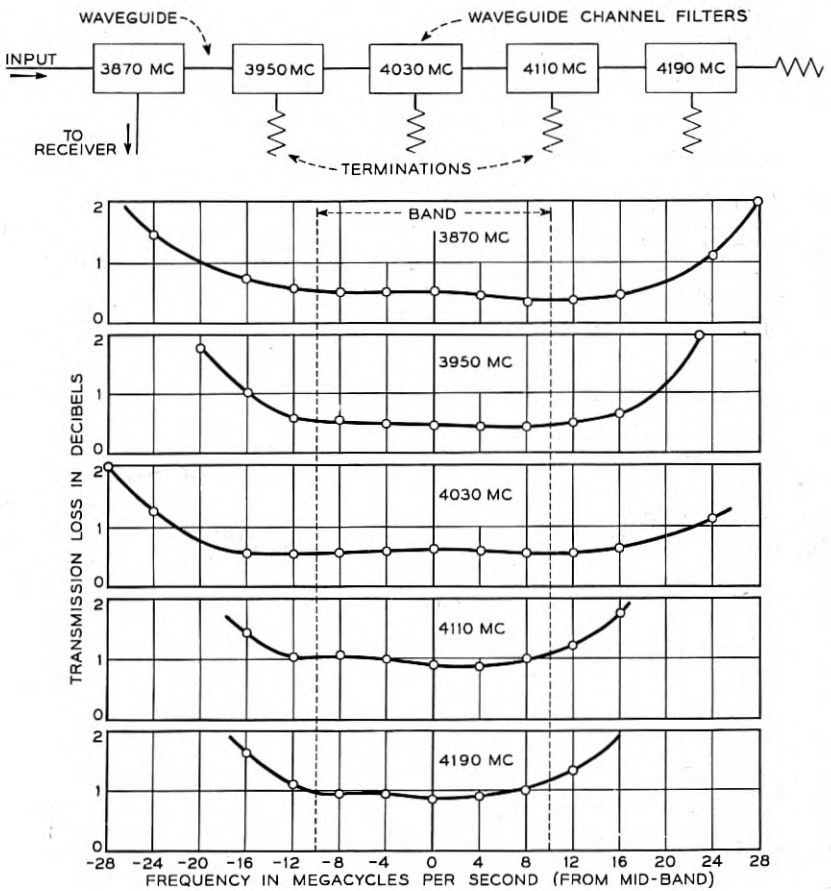


Fig. 11—Transmission loss of hybrid branching filter.

channel dropping unit, it appears logical to assume that the progressive increase in loss is due to passage of the higher frequencies through the lower frequency circuit components.

The measured discrimination against other channels and image responses was 20 db or more. This modest amount of selectivity is sufficient since the primary function of the present filter is branching, and hence the amount of

discrimination needed is only that required to prevent a significant energy loss in the channel bands. If crosstalk considerations indicate that more r-f discrimination is required, this can be obtained by placing auxiliary filters in the branch arms. This can be done without complication since the branching filter is a constant resistance device.

The delay distortion was examined but was found to be less than the errors of measurement, i.e. 2 millimicroseconds.

ACKNOWLEDGEMENTS

The authors, in carrying out the work described in this paper, have been helped by discussions with many people. Particular credit is due to J. R. Pierce for a variety of stimulating suggestions, and to their colleagues at the Holmdel Radio Laboratory for much fundamental work on microwave transmission.

A Note on a Parallel-Tuned Transformer Design

By V. C. RIDEOUT

An analysis of the parallel-tuned transformer used in radio-frequency amplifiers has been made for a slightly over-coupled case. The resulting design formulas are simple and practical.

Two cases are discussed: (a) the so-called matched transformer, with resistance loading on each side; (b) a transformer with loading on one side only, which has the same pass-band and phase characteristics as the matched transformer, but gives 3 db more gain when used as an interstage.

A special arrangement of (a) where the matched transformer design is used with one resistor removed, giving a transformer with a considerably double-humped pass-band characteristic and about 6 db more gain, is also discussed.

Introduction

THE parallel-tuned transformer has been used in radio-frequency amplifiers for many years.[1] In an excellent paper,[2] Christopher based design formulas on the principles of the broad band filter. In other publications, simple circuit analyses have been used and design formulas have been based on the assumption of a small ratio of bandwidth to mid-frequency which often fails to be adequate when the wide bands required for modern television and multiplexing services are encountered.

A transitionally flat transformer design may be obtained by setting the first three derivatives of the absolute value of the transfer impedance, with respect to frequency, equal to zero. The resulting design formulas are somewhat unwieldy.

The transformer design to be described here is based upon two simple circuit conditions[3] applied to the fundamental case of a parallel-tuned transformer with resistance loading on each side:

I. Both sides of the transformer are tuned to the same frequency.

II. The transmission loss is zero at the tune frequency. (This condition is responsible for the term "matched transformer" used to describe this case).

The resulting transformer has a slightly double-humped characteristic with less than 0.005 db dip for a coupling coefficient of 0.5. Because of the slight overcoupling this transformer design gives a little more gain and bandwidth than the critically coupled case. Its main advantage lies in the fact that simpler design formulas can be used.

[1] H. T. Friis and A. G. Jensen, High Frequency Amplifiers, *B. S. T. J.*, Vol. III, April 1924, pp. 181–205.

[2] A. J. Christopher, Transformer Coupling Circuits for High-Frequency Amplifiers, *B. S. T. J.*, Vol. XI, Oct. 1932, pp. 608–621.

[3] This analysis is based on an old unpublished report by H. T. Friis.

Derivation of Design Formulas

(a) *Matched, or Symmetrically Loaded Transformer.*

In the circuit of Fig. 1:

$$\left.\begin{aligned} E_1 &= I_1(R_1 + 1/j\omega C_1) + I_2(-1/i\omega C_1) \\ 0 &= I_1(-1/i\omega C_1) + I_2(j\omega L_1 + 1/j\omega C_1) + I_3(-j\omega M) \\ 0 &= I_2(-j\omega M) + I_3(j\omega L_2 + 1/j\omega C_2) + I_4(-1/j\omega C_2) \\ 0 &= I_3(-1/j\omega C_2) + I_4(R_2 + 1/j\omega C_2) \end{aligned}\right\} \quad (1)$$

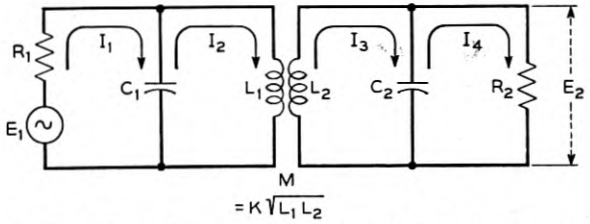


Fig. 1.—Parallel tuned, matched, or symmetrically loaded transformer.

Eliminating I_1, I_2 and I_3 among these equations gives:

$$\begin{aligned} E_1/I_4 = {} & \frac{L_1R_2 + L_2R_1}{M} + \omega^2\left[\left(M - \frac{L_1L_2}{M}\right)(C_1R_1 + C_2R_2)\right] \\ & + j\left(\omega^3\left[\left(M - \frac{L_1L_2}{M}\right)C_1C_2R_1R_2\right] + \omega\left[\frac{R_1R_2}{M}(L_1C_1 + L_2C_2)\right.\right. \\ & \left.\left. - \left(M - \frac{L_1L_2}{M}\right)\right] - \frac{1}{\omega}\cdot\frac{R_1R_2}{M}\right) \end{aligned} \quad (2)$$

If the coefficient of mutual coupling is k, and the first and second inductance and capacitance are resonant at ω_1 and ω_2 radians per second, respectively, we can substitute in (2) the expressions,

$$M = k\sqrt{L_1L_2},\ L_1 = 1/\omega_1^2C_1,\ L_2 = 1/\omega_2^2C_2,\ I_4 = E_2/R_2 \quad (3)$$

This gives the general expression,

$$\begin{aligned} E_1/E_2 = {} & \frac{R_1\sqrt{C_1C_2}}{k}\left\{\left[\frac{\omega_1}{\omega_2}\cdot\frac{1}{R_2C_2} + \frac{\omega_2}{\omega_1}\cdot\frac{1}{R_1C_1}\right]\right. \\ & - \omega^2\left[\frac{1-k^2}{\omega_1\omega_2}\left(\frac{1}{R_2C_2} + \frac{1}{R_1C_1}\right)\right] \\ & + j\left(\frac{-\omega^3}{\omega_1\omega_2}(1-k^2) + \omega\left[\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} + \frac{1-k^2}{\omega_1C_1R_1\omega_2C_2R_2}\right]\right. \\ & \left.\left. - \frac{\omega_1\omega_2}{\omega}\right)\right\} \end{aligned} \quad (4)$$

The application of circuit condition I, ($\omega_1 = \omega_2 = \omega_0$) to (4) gives,

$$E_1/E_2 = \frac{R_1\omega_0\sqrt{C_1C_2}}{k}\left\{\left(\frac{1}{\omega_0C_1R_1} + \frac{1}{\omega_0C_2R_2}\right)\right.$$
$$- \left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\left(\frac{1}{\omega_0C_1R_1} + \frac{1}{\omega_0C_2R_2}\right) \tag{5}$$
$$\left. - j\left[\left(\frac{\omega}{\omega_0}\right)^3(1-k^2) - \frac{\omega}{\omega_0}\left(2 + \frac{1-k^2}{\omega_0^2C_1C_2R_1R_2}\right) + \frac{\omega_0}{\omega}\right]\right\}$$

Circuit condition II (zero transmission loss at $\omega = \omega_0$) is satisfied if $|E_1/E_2| = 2\sqrt{R_1/R_2}$. Substituting this condition in (5) and solving for k gives:

$$k^2 = \frac{\omega_0^2C_1C_2R_1R_2 + 1 \pm j(\omega_0C_1R_1 - \omega_0C_2R_2)}{(\omega_0C_1R_1)^2 + (\omega_0C_2R_2)^2 + (\omega_0^2C_1C_2R_1R_2)^2 + 1} \tag{6}$$

Since k must be real,

$$\omega_0C_1R_1 = \omega_0C_2R_2 \tag{7}$$

and from (6) and (7)

$$\omega_0C_1R_1 = \omega_0C_2R_2 = \sqrt{1-k^2}/k \tag{8}$$

From (5) and (8) the transmission formula for this transformer is,

$$\frac{E_1}{E_2} = \frac{\sqrt{1-k^2}}{k^2}\sqrt{\frac{C_2}{C_1}}\left[\frac{2k}{\sqrt{1-k^2}}\left(1 - \left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\right)\right.$$
$$\left. - j\left(\left(\frac{\omega}{\omega_0}\right)^3(1-k^2) - \frac{\omega}{\omega_0}(2+k^2) + \frac{\omega_0}{\omega}\right)\right] \tag{9}$$

or, $\frac{E_1}{E_2} = \left|\frac{E_1}{E_2}\right| e^{j\phi}$, where,

$$\left|\frac{E_1}{E_2}\right| = \frac{\sqrt{1-k^2}}{k^2}\left(\frac{C_2}{C_1}\right)^{\frac{1}{2}}\left\{\frac{4k^2}{1-k^2}\left[1 - \left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\right]^2\right.$$
$$\left. + \left[\left(\frac{\omega}{\omega_0}\right)^3(1-k^2) - \frac{\omega}{\omega_0}(2+k^2) + \frac{\omega_0}{\omega}\right]^2\right\}^{\frac{1}{2}} \tag{10}$$

$$\phi = \arctan\frac{\sqrt{1-k^2}}{2k} \cdot \frac{(\omega/\omega_0)^4(1-k^2) - (\omega/\omega_0)^2(2+k^2) + 1}{(\omega/\omega_0)^3(1-k^2) - \omega/\omega_0} \tag{11}$$

The transmission loss, $\mathfrak{L}$, defined as the ratio of the maximum available power from the generator to the output power may be obtained from (10),

$$\mathfrak{L} = \frac{R_2}{4R_1}\left|\frac{E_1}{E_2}\right|^2 = \left\{\frac{1}{k^2}\left[1 - \left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\right]^2 + \frac{1-k^2}{4k^4}\left[\left(\frac{\omega}{\omega_0}\right)^3(1-k^2) - \frac{\omega}{\omega_0}(2+k^2) + \frac{\omega_0}{\omega}\right]^2\right\} \tag{12}$$

or the loss in decibels equals 10 log $\mathfrak{L}$.

The transmission delay, δ, may be obtained from (11).

$$\delta = \frac{d\phi}{d\omega} = \frac{\left(\frac{\omega}{\omega_0}\right)^4(1-k^2)^2 - \left(\frac{\omega}{\omega_0}\right)^2(1-k^2)^2 + 4k^2 - 1 + \left(\frac{\omega_0}{\omega}\right)^2}{\mathfrak{L}} \cdot \frac{\sqrt{1-k^2}}{2k^3\omega_0} \tag{13}$$

seconds.

The input impedance seen from the generator terminals is,

$$Z_{in} = R_1\left\{\frac{\left(\frac{\omega}{\omega_0}\right)k\sqrt{1-k^2} - j\left[1 - \left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\right]}{\left[\frac{\sqrt{1-k^2}}{k}\left(\frac{\omega_0}{\omega}\right)\right]\left[-1 + 2\left(\frac{\omega}{\omega_0}\right)^2 - \left(\frac{\omega}{\omega_0}\right)^4(1-k^2)\right] - j\left[1 - \left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\right]}\right\} \tag{14}$$

The formulas (12), (13) and (14) give the important characteristics of this transformer. Table I gives expressions for these characteristics in the general case, and for $k = 0.5$, for several picked frequencies. Of these frequencies ω_0 and $\omega_0/\sqrt{1-k^2}$ are the frequences of zero transmission loss, $\omega_0/\sqrt[4]{1-k^2}$ is the geometric midband, and $\omega_0/\sqrt{1+k}$ and $\omega_0/\sqrt{1-k}$ are the cut-off frequencies of an infinite filter made of identical transformers of this type.

The input impedance, transmission loss, and the transmission delay are plotted against ω/ω_0 for a matched transformer with $k = 0.5$ in Fig. 2. From these curves, and from Table I the following important characteristics of this transformer will be noted:

(1) The pass-band is approximately symmetrical with frequency, rather than with the logarithm of frequency, as in circuits which have a low-pass analog.

TABLE I

Relative Frequency ω/ω_0	Transmission Loss (db)		Delay $\times \omega_0$ (radians)		Input Impedance	Input Standing Wave Ratio in db. $k = 0.5$
	General Case	$k = 0.5$	General Case	$k = 0.5$		
1	0	0	$\frac{2\sqrt{1-k^2}}{k}$	3.464	R_1	0
$1/\sqrt{1-k^2}$	0	0	$\frac{2\sqrt{1-k^2}}{k}$	3.464	R_1	0
$1/\sqrt[4]{1-k^2}$	$\frac{0.136k^4}{2-k^2}$ approx.	0.0048	$\frac{3+\sqrt{1-k^2}}{4+k^4/8(2-k^2)} \cdot \frac{2\sqrt{1-k^2}}{k}$ approx.	3.344	$R_1\left(\frac{\frac{k\sqrt[4]{1-k^2}}{1-\sqrt{1-k^2}}-j}{\frac{2-\sqrt{1-k^2}}{k}-j}\right)$	0.585
$1/\sqrt{1+k}$	$10 \log \frac{5-k}{4}$	0.52	$\frac{6-k}{5-k} \cdot \frac{2\sqrt{1-k^2}}{k}$	4.234	$R_1(1+j\sqrt{1-k})$	6.02
$1/\sqrt{1-k}$	$10 \log \frac{5+k}{4}$	1.39	$\frac{6+k}{5+k} \cdot \frac{2\sqrt{1-k^2}}{k}$	4.094	$R_1(1-j\sqrt{1+k})$	10.06

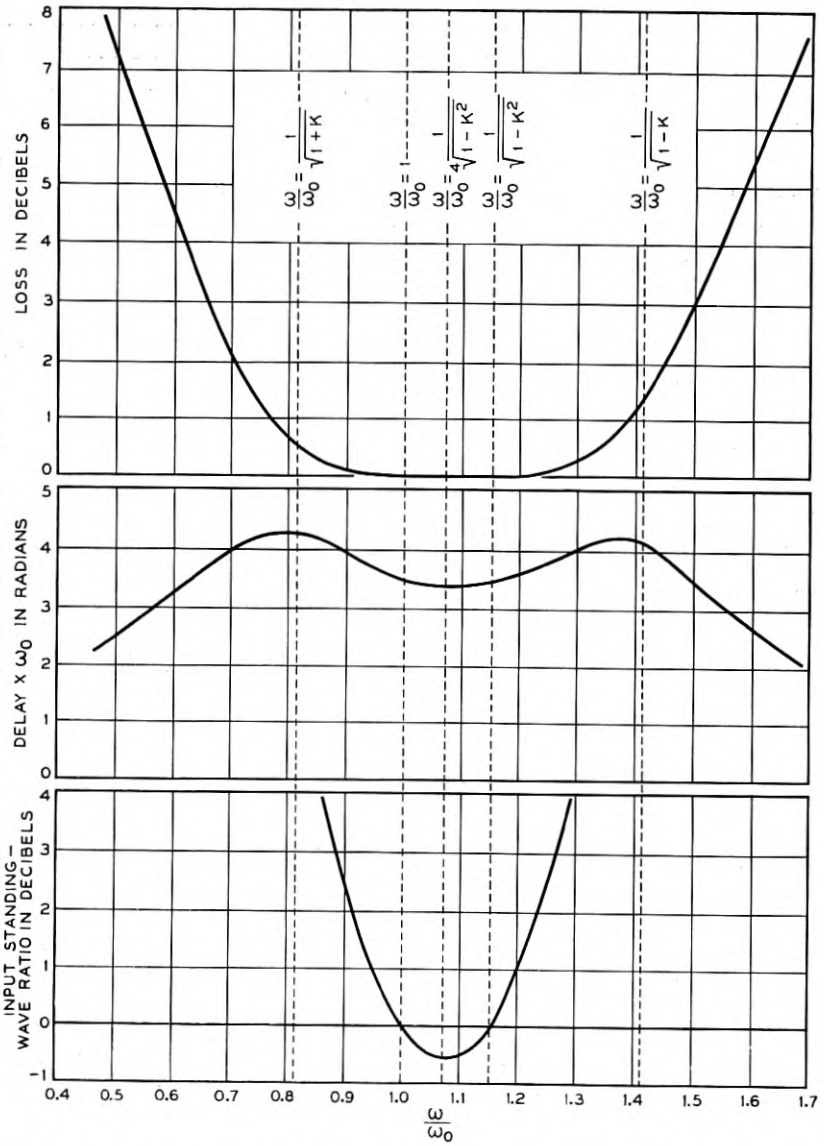


Fig. 2.—Loss, delay, and input standing-wave ratio versus relative frequency for the matched transformer for k = 0.5. Loss and delay curves also apply to the case of the mis-matched transformer for k = 0.632.

(2) The delay curve is also quite symmetrical. Circuits which can be derived from low-pass forms have more delay on the low-frequency side, where loss increases faster. The symmetry of the delay curve makes phase compensation easier.

(3) The input standing-wave ratio passes through zero on either side of mid-band, and although it has a 0.6 db dip at mid-band for the case of $k = 0.5$, it is under one db over a larger frequency range than in the transitionally coupled case.

(4) The bandwidth between points one db down on the loss curve is very nearly equal to the bandwidth between the cut-off frequencies. This bandwidth is simply related to the tune frequency f_0 and the geometric midband frequency f_m as shown below.

$$\left.\begin{aligned} \Delta f &= f_0/\sqrt{1-k} - f_0/\sqrt{1+k} \\ &\cong f_0 k/\sqrt{1-k^2} \\ &\cong f_m k/\sqrt[4]{1-k^2} \end{aligned}\right\} \tag{15}$$

Thus, given required values of f_m, Δf, C_1 (or R_1), and C_2 (or R_2), the matched transformer can be designed by the use of formulas (3), (8) and (15). These formulas are summarized at the end of this paper.

A transformer of different phase and loss characteristics results if the loading resistance is removed from one side of a matched transformer. Equation (5), if $R_2 = \infty$, becomes,

$$\frac{E_1}{E_2''} = \frac{\sqrt{1-k^2}}{k^2}\sqrt{\frac{C_2}{C_1}}\left[\frac{k}{\sqrt{1-k^2}}\left(1-\left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\right) - j\left(\left(\frac{\omega}{\omega_0}\right)^3(1-k^2) - 2\frac{\omega}{\omega_0} + \frac{\omega_0}{\omega}\right)\right] \tag{16}$$

At $\omega = \omega_0$,

$$|E_1/E_2''|^2 = C_2/C_1 \tag{17}$$

while from (10) the corresponding ratio for the matched transformer is,

$$|E_1/E_2|^2 = 4C_2/C_1 \tag{18}$$

Thus this mis-matched transformer has 6 db more gain at the tune frequency than the corresponding matched transformer, whether used as an interstage, an output or an input transformer.

The transmission loss referred to the loss at $\omega = \omega_0$ is,

$$\mathfrak{L}'' = \frac{C_1}{C_2}\left|\frac{E_1}{E_2''}\right|^2 = \frac{1}{k^2}\left[1-\left(\frac{\omega}{\omega_0}\right)^2(1-k^2)\right]^2 + \frac{1-k^2}{k^4}\left[\left(\frac{\omega}{\omega_0}\right)^3(1-k^2) - 2\frac{\omega}{\omega_0} + \frac{\omega_0}{\omega}\right]^2 \tag{19}$$

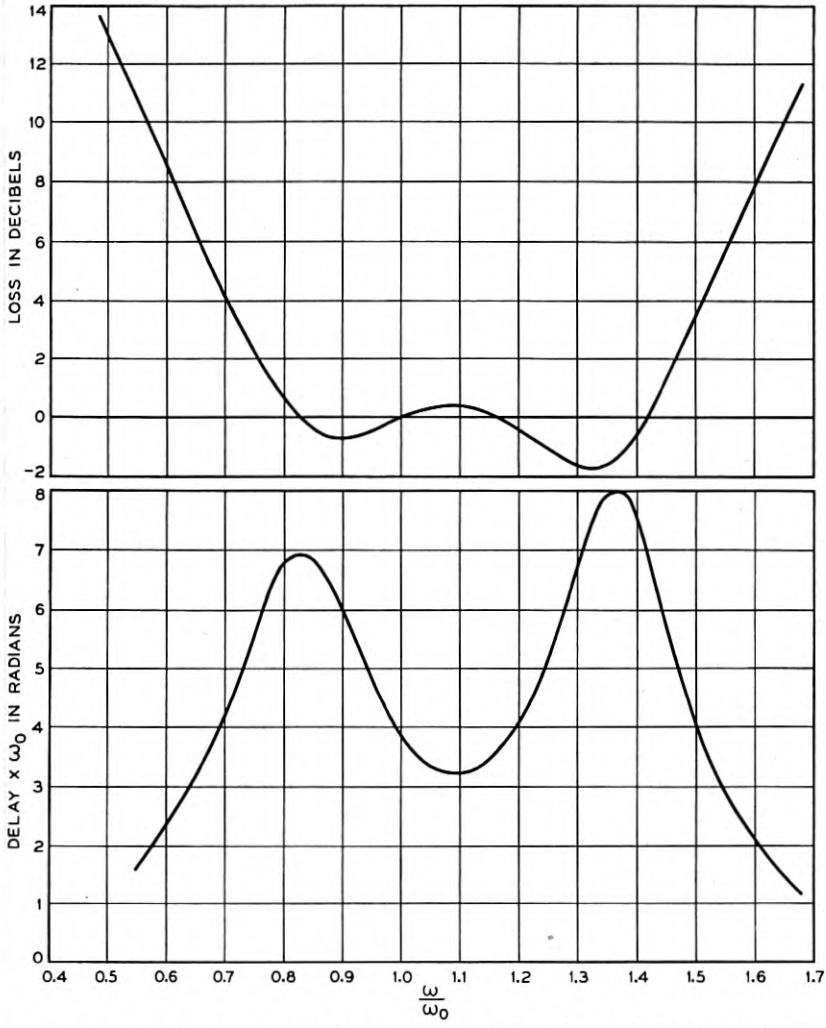


Fig. 3.—Loss and delay versus relative frequency for the mis-matched transformer obtained by making R_2 in Fig. 1 infinite, for k = 0.5.

The transmission delay δ' in this case is,

$$\delta' = \frac{\sqrt{1-k^2}}{\omega_0 k^3} \cdot \frac{\left(\frac{\omega}{\omega_0}\right)^4 (1-k^2)^2 - \left(\frac{\omega}{\omega_0}\right)^2 (1-k^2) + (3k^2-1) + \left(\frac{\omega_0}{\omega}\right)^2}{\mathfrak{D}''} \text{ seconds.} \tag{20}$$

The loss is equal at four points,

$$\omega = \omega_0, \quad \omega = \omega_0/\sqrt{1 - k^2}, \quad \omega = \omega_0/\sqrt{1 - k}, \quad \omega = \omega_0/\sqrt{1 + k}.$$

The curves of Fig. 3 show the loss and delay for this mis-matched transformer for the case where $k = 0.5$. The double-humped band-pass characteristic of the mis-matched transformer can be compensated for in an amplifier by de-tuning slightly to equalize the humps and using a single-tuned circuit in another interstage in which,

$$Q = \omega_m CR \cong \sqrt{1/k^2 - 3/4 - 3k^2/32} \tag{21}$$

The above combination of a mis-matched transformer and a single-tuned transformer, in successive interstages, will give approximately 6 db more gain than if two matched transformers were used. The actual gain obtained will depend upon the ratio of the input and output capacitances. This gain advantage carries with it the penalty of increased sensitivity to tube capac-

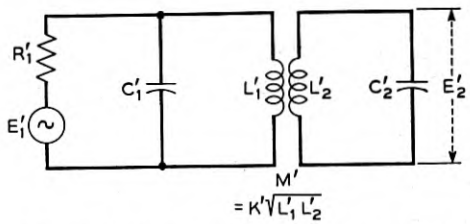


Fig. 4.—Mismatched or unsymmetrically loaded transformer.

itance variations. However, the mis-matched transformer may be used to advantage in the first interstage of an amplifier to minimize the effect of the second tube on the signal-to-noise ratio. It may also be used in the last interstage to offset any lack of power-handling capacity in the second-last tube.

(b) *Mismatched or Unsymmetrically Loaded Transformer*

A transformer with only one loading resistor, as in Fig. 4, can always be designed to have the same bandpass and phase characteristics as the matched transformer[4] shown in Fig. 1. No power transfer will occur, but the ratio of output voltage to the square root of the available input power can be determined, and set equal to this ratio for the matched transformer, except

[4] This important principle has been described by S. Darlington in a paper: Synthesis of Reactance 4-Poles, *Journal of Mathematics and Physics*, Vol. XVIII, Sept. 1939, pp. 257–353. Independently, this principle was demonstrated to the author by W. J. Albersheim in an unpublished memorandum dealing with the transitionally flat transformer.

for a constant factor. For the matched case this ratio is, from equation (9),

$$\sqrt{\frac{E_1^2/4R_1}{E_2^2}} = \left| \frac{1}{k}\sqrt{\frac{C_2}{C_1R_1}} - \left(\frac{\omega}{\omega_0}\right)^2 \frac{1-k^2}{k}\sqrt{\frac{C_2}{C_1R_1}} - j\left(\frac{\omega}{\omega_0}\right)^3 \frac{(1-k^2)\sqrt{1-k^2}}{2k^2}\sqrt{\frac{C_2}{C_1R_1}} + j\left(\frac{\omega}{\omega_0}\right)\frac{(2+k^2)\sqrt{1-k^2}}{2k^2}\sqrt{\frac{C_2}{C_1R_1}} - j\left(\frac{\omega_0}{\omega}\right)\frac{\sqrt{1-k^2}}{2k^2}\sqrt{\frac{C_2}{C_1R_1}} \right| \quad (22)$$

For the mismatched transformer the corresponding ratio can be obtained from equation (4) by letting R_2 approach infinity, giving,

$$\sqrt{\frac{E_1'^2/4R_1'}{E_2'^2}} = \left| \frac{1}{2k'}\frac{\omega_2'}{\omega_1}\sqrt{\frac{C_2'}{C_1'R_1'}} - \frac{\omega^2}{\omega_1'\omega_2'} \cdot \frac{1-k'^2}{2k'}\sqrt{\frac{C_2'}{C_1'R_1'}} - j\frac{\omega^3}{\omega_1'\omega_2'}\frac{1-k'^2}{2k'}\sqrt{R_1'C_1'C_2'} + j\omega\left(\frac{\omega_2'}{\omega_1'}+\frac{\omega_1'}{\omega_2'}\right)\frac{\sqrt{R_1'C_1'C_2'}}{2k'} - j\frac{\omega_1'\omega_2'}{\omega}\frac{\sqrt{R_1'C_1'C_2'}}{2k'} \right| \quad (23)$$

The two transformers are to have the same phase characteristics, and the same band-pass characteristics except for a constant factor N where,

$$N = \sqrt{\frac{E_1^2/4R_1}{E_2^2}} \div \sqrt{\frac{E_1'^2/4R_1'}{E_2'^2}} \quad (24)$$

The terms involving voltages in (22) and (23) can be eliminated by combining with (24). Equating the coefficients of like powers of ω gives (25), below.

$$\left.\begin{aligned}
\frac{1}{k}\sqrt{\frac{C_2}{C_1R_1}} &= \frac{1}{2k'}\sqrt{\frac{C_2'}{R_1'C_1'}}\,N\frac{\omega_2'}{\omega_1'} \\
\frac{1-k^2}{k}\sqrt{\frac{C_2}{C_1R_1}}\frac{1}{\omega_0^2} &= \frac{1-k'^2}{2k'}\sqrt{\frac{C_2'}{C_1'R_1'}}\left(\frac{N}{\omega_1'\omega_2'}\right) \\
\frac{1-k^2}{k}\sqrt{R_1C_1C_2}\cdot\frac{1}{\omega_0^2} &= \frac{1-k'^2}{k'}\sqrt{R_1'C_1'C_2'}\left(\frac{N}{\omega_1'\omega_2'}\right) \\
\frac{2+k^2}{k}\sqrt{R_1C_1C_2} &= \frac{\sqrt{R_1'C_1'C_2'}}{k'}\left(\frac{\omega_2'}{\omega_1'}+\frac{\omega_1'}{\omega_2'}\right)N \\
\frac{\sqrt{R_1C_1C_2}}{k}\omega_0^2 &= \frac{\sqrt{R_1'C_1'C_2'}}{k'}(\omega_1'\omega_2'N)
\end{aligned}\right\} \quad (25)$$

The last three equations were simplified by the use of equation (8).

Equations (25) yield the five relations below:

$$\left.\begin{aligned} R_1'C_1' &= R_1C_1/2 \\ \omega_1 &= \omega_0 \\ \omega_2 &= \omega_0/\sqrt{1+k^2} \\ k' &= \frac{\sqrt{2}\,k}{\sqrt{1+k^2}} \\ N &= 2\sqrt{\frac{C_2}{C_2'}} \end{aligned}\right\} \quad (26)$$

Summary of Design Formulas

A transformer must usually meet certain requirements as to bandwidth and mid-frequency. Loss curves for various values of k, plotted against relative frequency may be used. To a very close approximation the geometric mean frequency $f_m = \omega_m/2\pi$, may be used for the mid-band frequency, and Δf, the bandwidth between cut-off points, may be used for the bandwidth between one db points. Then,

$$\omega_m = \omega_0/\sqrt[4]{1-k^2} \quad (27)$$

And, from (15),

$$k \cong \frac{\Delta f/f_0}{\sqrt{1+(\Delta f/f_0)^2}} \quad (28)$$

The other necessary relations for the matched transformer and for the mismatched transformer with the same transmission characteristics, as derived from equations (3), (8) and (26) are given below.

(a) Matched Transformer.

$$\left.\begin{aligned} \frac{1}{\sqrt{L_1C_1}} &= \omega_0 \\ \frac{1}{\sqrt{L_2C_2}} &= \omega_0 \\ \frac{M}{\sqrt{L_1L_2}} &= k \\ R_1C_1 = R_2C_2 &= \frac{\sqrt{1-k^2}}{\omega_0 k} \end{aligned}\right\} \quad (29)$$

(b) Mismatched Transformer:

$$\left.\begin{aligned} \frac{1}{\sqrt{L_1' C_1'}} &= \omega_0 \\ \frac{1}{\sqrt{L_2' C_2'}} &= \frac{\omega_0}{\sqrt{1+k^2}} \\ \frac{M'}{\sqrt{L_1' L_2'}} &= k' = \frac{\sqrt{2}\,k}{\sqrt{1+k^2}} \\ R_1' C_1' &= \frac{\sqrt{1-k^2}}{2\omega_0 k} \end{aligned}\right\} \quad (30)$$

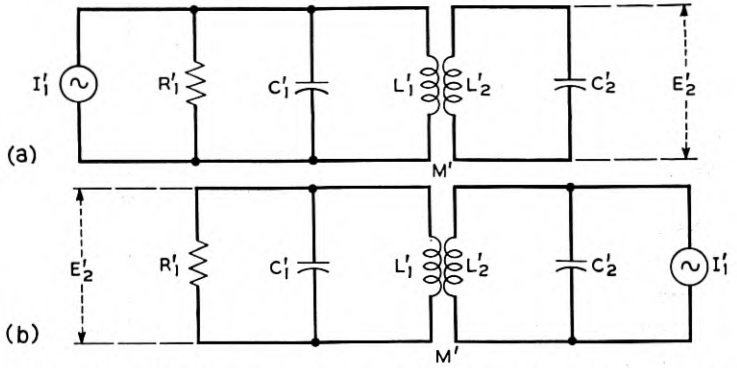


Fig. 5.—Mismatched transformer (a) fed by constant current generator, (b) with input and output reversed.

The ratio of the output voltages in the two cases is given by:

$$\left|\frac{E_2'}{E_2}\right|^2 = N^2 \left(\frac{E_1'^{\,2}/4R_1'}{E_1^{\,2}/4R_1}\right) = 4\,\frac{R_1 C_2}{R_1' C_2'}\left|\frac{E_1'}{E_1}\right|^2 \quad (31)$$

The relative gains of the two transformers, as given by equation (31), will be examined for three important cases:

(1) *Input Transformer*—In this case $R_1 = R_1'$, and $C_2 = C_2'$, so that from (31), the mismatched transformer gives four times, or 6 db more gain.

(2) *Output Transformer*—The voltage source in Fig. 4 may be replaced by the constant current source $I_1' = E_1'/R_1'$, by Norton's Theorem, as shown in Fig. 5(a). By the Reciprocity Theorem we may exchange the positions of the current source and the output voltage as shown in Fig. 5(b), without changing the value of the latter. This may also be done for the

matched transformer. The resultant circuit is that for a pentode output transformer in each case, where the plate resistance of the tube is so high compared to circuit impedance that we can consider it to be a constant current generator of $g_m E_g$ amperes, where g_m is the grid-plate transconductance and E_g the grid signal voltage. Thus in this case the mis-matched transformer also gives 6 db more gain than the matched transformer.

(3) *Interstage Transformer*—If the voltage generators of Figs. 1 and 4 are replaced by their equivalent current generators $g_m E_g = E_1/R_1$ and $g'_m E'_g = E'_1/R'_1$, then (31) becomes,

$$\left|\frac{E'_2}{E_2}\right|^2 = 4\,\frac{R'_1 C_2}{R_1 C'_2}\left|\frac{g'_m E'_g}{g_m E_g}\right|^2 \tag{32}$$

The first relation in (26) may be substituted in (32) giving,

$$\left|\frac{E'_2}{E_2}\right|^2 = 2\left|\frac{C_1 C_2 g'_m E'_g}{C'_1 C'_2 g_m E_g}\right|^2 \tag{33}$$

Thus, in the interstage case, this mismatched transformer has only 3 db more gain than the matched transformer.

Here, as in the case of the first mismatched transformer discussed in this paper, there is an increased sensitivity to capacity variations. In practice it appears that this is not always serious in wide-band intermediate-frequency amplifiers. Some improvement in noise figure and power handling may be obtained, as before, without encountering difficulties with double-peaked pass-band characteristics.

The design formulas given in this paper have been used successfully in the design of experimental intermediate frequency amplifiers in the 65 mc region having band widths of the order of 10 to 20 mc.

Statistical Properties of a Sine Wave Plus Random Noise

By S. O. RICE

Introduction

IN SOME technical problems we are concerned with a current which consists of a sinusoidal component plus a random noise component. A number of statistical properties of such a current are given here. The present paper may be regarded as an extension of Section 3.10 of an earlier paper,[1] "Mathematical Analysis of Random Noise", where some of the simpler properties of a sine wave plus random noise are discussed.

The current in which we are interested may be written as

$$\begin{aligned} I &= Q\cos qt + I_N \\ &= R\cos(qt + \theta) \end{aligned} \tag{3.4}$$

where Q and q are constants, t is time, and I_N is a random (in the sense of Section 2.8 of Reference A) noise current. When the second expression involving the envelope R and the phase angle θ is used, the power spectrum of I_N is assumed to be confined to a relatively narrow band in the neighborhood of the sine wave frequency $f_q = q/(2\pi)$. This makes R and θ relatively slowly (usually) varying functions of time.

In Section 1, the probability density and cumulative distribution of I are discussed. In Section 2, the upward "crossings" of I (i.e., the expected number of times, per second, I increases through a given value I_1), are examined.

The probability density and the cumulative distribution of R are given in Section 3.10 of Reference A. The crossings of R are examined in Section 4 of the present paper.

The statistical properties of θ', the time derivative of the phase angle θ, are of interest because the instantaneous frequency of I may be defined to be $f_q + \theta'/(2\pi)$. The probability density of θ' is investigated in Section 5 and the crossings of θ' in Section 6. θ' is a function of time which behaves somewhat like a noise current and may accordingly be considered to consist of an infinite number of sinusoidal components. The problem of determining the "power spectrum" $W(f)$ of θ', i.e., the distribution of the mean square value of the components as a function of frequency, is attacked in

[1] *B.S.T.J.* 23 (1944), 282–332 and 24 (1945), 46–156. This paper will be called "Reference A".

Sections 7 and 8. The correlation function of θ' is expressed in terms of exponential integrals in Section 7, the power spectrum of I_N being assumed symmetrical and centered on f_q. In Section 8, values of $W(f)$ are obtained for the special case in which the power spectrum of I_N is centered on f_q and is of the normal-law type.

It is believed that some of the material presented here may find a use in the study of the effect of noise in frequency modulation systems. For example, the curves in Section 8 yield information regarding the noise power spectrum in the output of a primitive type of system. Also, the procedure employed to obtain the expression (5.7) for $\bar{\theta}'$ may be used to show that if

$$Q\cos[(A/\omega_0)\cos\omega_0 t + qt] + I_N = R\cos(qt + \theta)$$

the sinusoidal component of $d\theta/dt$ is[2]

$$-A\,(1 - e^{-\rho})\sin\omega_0 t$$

where ρ is the ratio $Q^2/(2\overline{I_N^2})$. This illustrates the "crowding effect" of the noise. The statistical analysis associated with R and θ of equations (3.4) (when the sine wave is absent) is similar to that used in the examination of the current returned to the sending end of a transmission line by reflections from many small irregularities distributed along the line. This suggests another application of the results.

ACKNOWLEDGMENT

I am indebted to a number of my associates for helpful discussions on the questions studied here. In particular, I wish to thank Mr. H. E. Curtis for his suggestions regarding this subject. As in Reference A, all of the computations for the curves and tables have been performed by Miss M. Darville. This work has been quite heavy and I gratefully acknowledge her contribution to this paper.

1. PROBABILITY DISTRIBUTION OF A SINE WAVE PLUS RANDOM NOISE

A current consisting of a sine wave plus random noise may be represented as

$$I = Q\cos qt + I_N \tag{1.1}$$

where Q and q are constants, t is the time, and I_N is a random noise current. The frequency, in cycles per second, of the sine wave is $f_q = q/(2\pi)$. In all

[2] The first person to obtain this result was, I believe, W. R. Young who gave it in an unpublished memorandum written early in 1945. He took the output of a frequency modulation limiter and discriminator to be proportional to either the signal frequency or to the instantaneous frequency (assumed to be distributed uniformly over the input band) of I_N according to whether Q is greater or less than the envelope of I_N. His memorandum also contains results which agree well with several obtained in this paper.

our work we denote the power spectrum of I_N by $w(f)$ and its correlation function by $\psi(\tau)$. The mean square value of I_N is denoted by ψ_0.

The study of the probability distribution of I is essentially a study of the integral[3]

$$p(I) = \frac{1}{\pi\sqrt{\psi_0}} \int_0^\pi \varphi\left[\frac{I - Q\cos\theta}{\sqrt{\psi_0}}\right] d\theta \tag{1.2}$$

where

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{1.3}$$

and $p(I)$ is the probability density of I, i.e. $p(I)dI$ is the probability that a value of current selected at random will lie in the interval $I, I + dI$. Another expression for $p(I)$ is given by equation (3.10–6) of Reference A, namely

$$p(I) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-izI - \psi_0 z^2/2} J_0(Qz)\, dz \tag{1.4}$$

where $J_0(Qz)$ denotes the Bessel function of order zero.

The substitutions

$$y = \frac{I}{\sqrt{\psi_0}}; \qquad a = \frac{Q}{\sqrt{\psi_0}} \tag{1.5}$$

enable us to write (1.2) as

$$p_1(y) = \sqrt{\psi_0}\, p(I) = \frac{1}{\pi} \int_0^\pi \varphi(y - a\cos\theta)\, d\theta, \tag{1.6}$$

where $p_1(y)$ denotes the probability density of y. This is the expression actually studied. Curves showing $p_1(y)$ and the cumulative distribution function

$$\begin{aligned} \int_{-\infty}^{I} p(I_1) dI_1 &= \int_{-\infty}^{y} p_1(y_1) dy_1 \\ &= \frac{1}{\pi} \int_0^\pi \varphi_{-1}(y - a\cos\theta)\, d\theta, \end{aligned} \tag{1.7}$$

where

$$\varphi_{-1}(x) = \int_{-\infty}^{x} \varphi(x_1) dx_1 = \tfrac{1}{2} + \tfrac{1}{2}\, erf\,(x/\sqrt{2}) \tag{1.8}$$

[3] W. R. Bennett, "Response of a Linear Rectifier to Signal and Noise," *Jour. Acous. Soc. Amer.* Vol. 15 (1944), 164–172, and *B.S.T.J.* Vol. 23 (1944), 97–113.

are shown in Figs. 1 and 2. The curves for $a = 10$ and $a = \sqrt{10}$ were computed by Simpson's rule from (1.6) and (1.7), and the curves for $a = 1$ were computed from the series (1.10) given below. Since both $\varphi(x)$ and

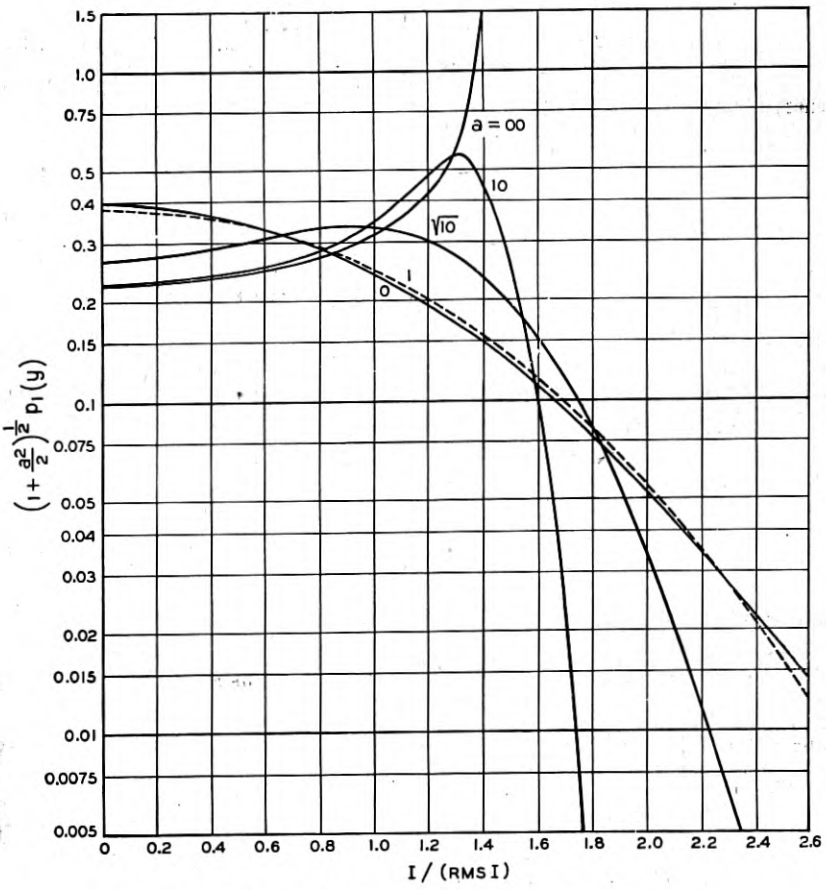


Fig. 1—Probability density of sine wave plus noise.

$I = Q\cos qt + I_N$, $a = Q/\sqrt{\psi_0}$, $y = I/\sqrt{\psi_0}$, $\sqrt{\psi_0}$ = rms I_N
$p_1(y)\, dI/\sqrt{\psi_0}$ = probability total current lies between I and $I + dI$
$y(1 + a^2/2)^{-1/2} = I/(\text{rms } I)$. Curves are symmetrical about $y = 0$.

$\varphi_{-1}(x)$ are tabulated[4] functions the integrals in (1.6) and (1.7) are well suited to numerical evaluation.

[4] $\varphi(x)$ is given directly and $\varphi_{-1}(x)$ may be readily obtained from W.P.A., "Tables of Probability Functions," Vol. II, New York (1942). The functions $\varphi^{(m)}(y)$ are tabulated in Table V of "Probability and its Engineering Uses" by T. C. Fry (D. Van Nostrand Co., 1928).

The form assumed by $p_1(y)$ as the parameter a becomes large is examined in the latter portion (from equation (1.12) onwards) of the section.

Series which converge for all values of a but which are especially suited for calculation when $a \leq 1$ may be obtained by inserting the Taylor's series (in powers of x) for $\varphi(y + x)$ and $\varphi_{-1}(y + x)$, $x = -a \cos \theta$, in (1.6) and (1.7) and integrating termwise. When we introduce the notation[4]

$$\varphi^{(m)}(y) = \frac{d^m}{dy^m}\varphi(y) = \frac{1}{\sqrt{2\pi}}\frac{d^m}{dy^m}e^{-y^2/2} \tag{1.9}$$

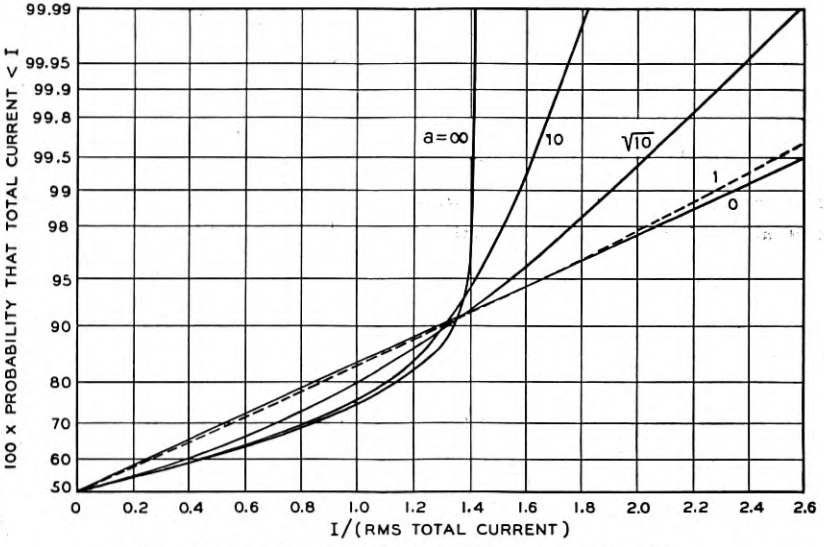


Fig. 2—Cumulative distribution of sine wave plus noise.
Ordinate = $100 \int_{-\infty}^{v} p_1(y_1)\, dy_1$. See Fig. 1 for notation.

we obtain

$$\begin{aligned} p_1(y) &= \sum_{n=0}^{\infty} \frac{1}{n!n!}\left(\frac{a}{2}\right)^{2n} \varphi^{(2n)}(y) \\ \int_{-\infty}^{v} p_1(y_1)dy_1 &= \varphi_{-1}(y) + \sum_{n=1}^{\infty} \frac{1}{n!n!}\left(\frac{a}{2}\right)^{2n} \varphi^{(2n-1)}(y) \end{aligned} \tag{1.10}$$

The second equation of (1.10) may be shown to be valid by breaking the interval $(-\infty, y)$ into $(-\infty, 0)$ and $(0, y)$. In the first part,

$$\int_{-\infty}^{0} p_1(y_1)\, dy_1 = \varphi_{-1}(0)$$

since both sides have the value 1/2. In the second, term by term integration is valid since the series integrated are uniformly convergent as may be seen from the inequality

$$|\varphi^{(n)}(y)| \leq \left(\frac{n!}{2\pi}\right)^{1/2}\left(\frac{2}{\pi n}\right)^{1/4}[1 + 0(n^{-1}) + 0(y^2 n^{-1})], \qquad (1.11)$$

in which we suppose that y remains finite as $n \to \infty$. This may be obtained by using the known behavior of Hermite polynomials of large order.[5]

When $Q \gg rms\ I_n$ so that a is very large the distribution approaches that of a sine wave, namely

$$p_1(y) \sim \begin{cases} 0, & |y| > a \\ (a^2 - y^2)^{-1/2}/\pi, & |y| < a \end{cases}$$
$$\int_{-\infty}^{y} p_1(y_1)dy_1 \sim \frac{1}{2} + \frac{1}{\pi}\arcsin\frac{y}{a}, \qquad |y| < a \qquad (1.12)$$

In order to study the manner in which the limiting expressions (1.12) are approached it is convenient to make the change of variable

$$x = y - a\cos\theta, \qquad d\theta = [a^2 - (y - x)^2]^{-1/2}\,dx$$
$$z = x - y + a$$

in (1.6). We obtain

$$p_1(y) = \frac{1}{\pi}\int_{y-a}^{y+a} \varphi(x)\,[a^2 - (y - x)^2]^{-1/2}\,dx$$
$$= \frac{1}{\pi}\int_0^{2a} \varphi(z + y - a)[z(2a - z)]^{-1/2}\,dz. \qquad (1.13)$$

An asymptotic (as a becomes large) expression for $p_1(y)$ suitable for the middle portion of the distribution where $a - |y| \gg 1$ may be obtained from the first integral in (1.13). Since the principal contributions to the value of the integral come from the region around $x = 0$ we are led to expand the radical in powers of x and integrate termwise. Legendre polynomials enter naturally since they are sometimes defined as the coefficients in such an expansion. Replacing the limits of integration $y + a$ and $y - a$ by $+\infty$ and $-\infty$, respectively and integrating termwise gives

$$p_1(y) \sim \frac{(a^2 - y^2)^{-1/2}}{\pi}\left[1 + \sum_{n=1}^{\infty}(-)^n\,\frac{1.3.5.\ (2n - 1)}{(a^2 - y^2)^n}\,P_{2n}(it^{1/2})\right]$$
$$= \frac{(a^2 - y^2)^{-1/2}}{\pi}\left[1 + \frac{3t + 1}{2(a^2 - y^2)} + \frac{3(35t^2 + 30t + 3)}{8(a^2 - y^2)^2} + \cdots\right] \qquad (1.14)$$

[5] A suitable asymptotic formula is given in Orthogonal Polynomials, by G. Szegö, *Am. Math. Soc. Colloquium, Pub.*, Vol. 23, (1939), p. 195.

where $t = y^2/(a^2 - y^2)$ and $P_{2n}(\ )$ denotes the Legendre polynomial of order $2n$. We have written this as an asymptotic expansion because it obviously is one when y, and hence t, is zero in which case

$$P_{2n}(0) = (-)^n \frac{1.3.5 \cdots (2n-1)}{2.4. \cdots 2n}$$

When y is near a or is greater than a, a suitable asymptotic expansion may be obtained from the second integral in (1.13) by expanding $(2a - z)^{-1/2}$ in powers of $z/(2a)$ and integrating termwise. The upper limit of integration, $2a$, may be replaced by ∞ since $\varphi(z + y - a)$ may be assumed to be negligibly small when z exceeds 2a. We thus obtain

$$\begin{aligned} p_1(y) &\sim \frac{1}{\pi} \sum_{n=0}^{\infty} \frac{(\tfrac{1}{2})_n}{n!} \left(\frac{1}{2a}\right)^{n+1/2} \int_0^{\infty} \varphi(z + y - a) z^{n-1/2}\, dz \\ &= \frac{\varphi(y-a)}{\pi} \sum_{n=0}^{\infty} \frac{(\tfrac{1}{2})_n}{n!} \left(\frac{1}{2a}\right)^{n+1/2} \int_0^{\infty} e^{-z(y-a)-(z^2/2)} z^{n-1/2}\, dz \end{aligned} \tag{1.15}$$

where we have used the notation

$$(\alpha)_0 = 1, \qquad (\alpha)_n = \alpha(\alpha + 1) \cdots (\alpha + n - 1).$$

The integrals occurring in (1.15) are related to the parabolic cylinder function[6] $D_m(x)$. Their properties may be obtained from the known properties of these functions or may be obtained by working directly with the integrals.

Suppose now that a is very large so that only the leading term in the series (1.15) for $p_1(y)$ need be retained.

Then

$$p_1(y) \sim a^{-1/2} F(y - a) \tag{1.16}$$

where

$$F(s) = \pi^{-1} 2^{-1/2} \int_0^{\infty} \varphi(z + s) z^{-1/2}\, dz \tag{1.17}$$

By writing out $\varphi(z + s)$, expanding $exp\ (-zs)$ in a power series, and integrating termwise we see that

$$\begin{aligned} F(s) &= \frac{\varphi(s) 2^{-5/4}}{\pi} \sum_{\ell=0}^{\infty} \frac{\Gamma\left(\frac{\ell}{2} + \frac{1}{4}\right)}{\ell!} (-s\sqrt{2})^{\ell} \\ &= (2\pi)^{-1} s^{1/2} \varphi(s/\sqrt{2}) K_{\frac{1}{4}}(s^2/4) \end{aligned} \tag{1.18}$$

where K denotes a modified Bessel function.[7] The relation (1.18) may also

[6] Whittaker and Watson, "Modern Analysis," 4th ed. (1927), 347–351.

[7] A table of $K_{\frac{1}{4}}(x)$ is given by H. Carsten and N. McKerrow, Phil. Mag. S7, Vol. 35 (1944), 812–818.

be obtained from pair 923.1 of "Fourier Integrals for Practical Applications," by G. A. Campbell and R. M. Foster.[8]

A curve showing $F(y-a)$ plotted as a function of $y-a$ is given in Fig. 3. It was obtained from the relation

$$F(s) = 2^{1/4}\pi^{-3/2}\chi(-s/\sqrt{2})$$

where

$$\chi(x) = \int_0^\infty e^{-(x-w^2)^2}\,dw$$

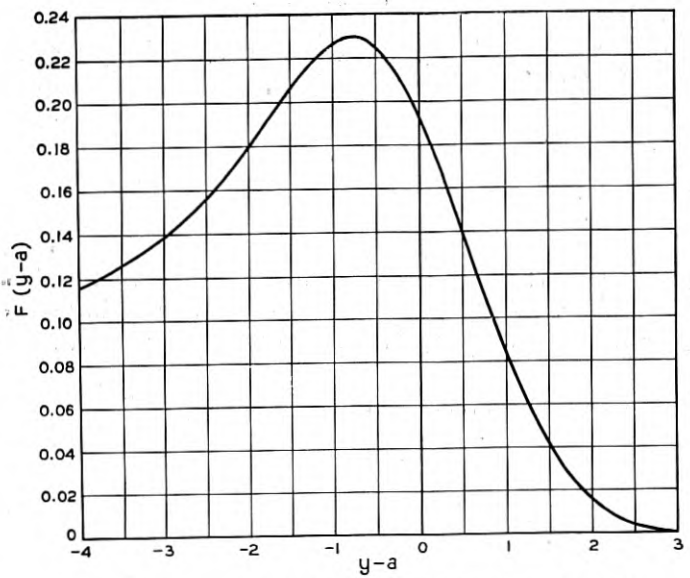


Fig. 3—Probability density of sine wave plus noise.
When rms $I_N << Q$ and I is near Q, $p_1(y) \sim a^{-1/2}F(y-a)$, $y-a = (I-Q)/(\text{rms } I_N)$.
See Fig. 1 for notation.

This function has been tabulated by Hartree and Johnston.[9]

The probability that I exceeds Q, or that y exceeds a, is, integrating the second of expressions (1.13),

$$\int_a^\infty p_1(y)\,dy = \frac{1}{\pi}\int_0^{2a}\frac{dz}{\sqrt{z(2a-z)}}\int_z^\infty \varphi(x)\,dx.$$

An asymptotic expansion may be obtained by expanding $(2a-z)^{-1/2}$ as in the derivation of (1.15) but we shall be content with the leading term.

[8] *Bell Telephone System Monograph* B-584.
[9] *Manchester Lit. and Phil. Soc. Memoirs*, v. 83, 183–188, Aug., 1939.

Using

$$\int_0^\infty z^{-1/2}\,dz \int_z^\infty \varphi(x)\,dz = \int_0^\infty \varphi(x)\,dx \int_0^x z^{-1/2}\,dz = 2^{1/4}\pi^{-1/2}\Gamma(\tfrac{3}{4})$$

we obtain

$$\int_a^\infty p_1(y)\,dy \sim 2^{-1/4}\pi^{-3/2}\Gamma(\tfrac{3}{4})a^{-1/2} = 0.185\cdots a^{-1/2} \tag{1.19}$$

For use in computations we list the following values

$$\Gamma(\tfrac{1}{4}) = 3.62561, \qquad \Gamma(\tfrac{3}{4}) = 1.22542, \qquad \Gamma(\tfrac{5}{4}) = 0.90640$$

2. Expected Number of Crossings of I per Second

In this section, we shall study two questions. First, what is the probability $P(I_1, t_1)dt$ of I increasing through the value I_1 (i.e. of I passing through the value I_1 with positive slope) during the infinitesimal interval $t_1, t_1 + dt$? Second, what is the expected number $N(I_1)$ of times per second I increases through the value I_1. When I_1 is zero, 2N(0) is the expected number of zeros per second, and when I_1 is large $N(I_1)$ is approximately equal to the expected number of maxima lying above the value I_1 in an interval one second long.

We start on the first question by considering the random function

$$z = F(a_1, a_2, \cdots a_N; t)$$

where the a's are random variables. The probability that the random curve obtained by plotting z as a function of t increases through the value $z = z_1$ in the interval $t_1, t_1 + dt$ is[10]

$$dt \int_0^\infty \eta p(z_1, \eta; t_1)\,d\eta \tag{2.1}$$

where $p(\xi, \eta; t_1)$ denotes the probability density of the random variables

$$\xi = F(a_1, a_2, \cdots, a_N; t_1)$$

$$\eta = \left[\frac{\partial F}{\partial t}\right]_{t=t_1}.$$

In our case z becomes the current I defined by equation (1.1). The method used to obtain equation (3.3–9) of Reference A may also be used to show that the quantity $p(I_1, \eta, t_1)$ (which now appears in (2.1)) is given by

$$p(I_1, \eta, t_1) = \frac{\pi N_0}{-\psi_0''}\varphi(y - a\cos qt_1)\varphi(x + b\sin qt_1) \tag{2.2}$$

[10] This result is a straightforward generalization of expression (3.3-5) in Section 3.3 of Reference A where references to related results by M. Kac are given. A formula equivalent to (2.1) has also been given by Mr. H. Bondi in an unpublished paper written in 1944. He applies his formula to the problem studied in Section 4.

where $\varphi(\ )$ denotes the normal law function defined by equation (1.3) and

$$-\psi_0'' = 4\pi^2 \int_0^\infty w(f) f^2\, df, \qquad N_0 = \frac{1}{\pi}\sqrt{\frac{-\psi_0''}{\psi_0}}, \qquad y = \frac{I_1}{\sqrt{\psi_0}},$$
$$a = \frac{Q}{\sqrt{\psi_0}}, \qquad x = \frac{\eta}{\sqrt{-\psi_0''}}, \qquad b = \frac{Qq}{\sqrt{-\psi_0''}} = \frac{2af_q}{N_0}. \tag{2.3}$$

Equation (3.3–11) of Reference A shows that N_0 is the expected number of zeros per second which I_N would have if it were to flow alone.

Let $P(I_1, t_1)dt$ be the probability that I will increase through the value I_1 during the interval $t_1, t_1 + dt$. Then (2.1) and (2.2) give

$$\begin{aligned} P(I_1\ t_1) &= \int_0^\infty \eta p(I_1, \eta, t_1)\, d\eta \\ &= \pi N_0 \varphi(y - a\cos qt_1) \int_0^\infty x\varphi(x + b\sin qt_1)\, dx. \end{aligned} \tag{2.4}$$

The integral in (2.4) is of the form

$$\begin{aligned} \int_0^\infty x\varphi(x + v)\, dx &= \varphi(v) - v\int_v^\infty \varphi(x)\, dx \\ &= -\frac{v}{2} + \varphi(v) + v\int_0^v \varphi(x)\, dx \\ &= -v + \varphi(v) + v\varphi_{-1}(v) \\ &= -\frac{v}{2} + (2\pi)^{-1/2}\, {}_1F_1\left(-\frac{1}{2}; \frac{1}{2}; -\frac{v^2}{2}\right) \end{aligned} \tag{2.5}$$

where v replaces $b\sin qt_1$ and ${}_1F_1$ denotes a confluent hypergeometric function.

The distribution of the crossings at various portions of the cycle (of the sine wave) may be obtained by giving special values to qt_1 in (2.4).

The expected number of times I increases through the value I_1 in one second is

$$\begin{aligned} N(I_1) &= \operatorname*{Limit}_{T\to\infty} \frac{1}{T}\int_0^T P(I_1, t_1)\, dt_1 \\ &= N_0 \int_0^\pi \varphi(y - a\cos\theta)\left[\varphi(b\sin\theta) + b\sin\theta \int_0^{b\sin\theta} \varphi(x)\, dx\right] d\theta \end{aligned} \tag{2.6}$$

where we have used (2.4) and the second equation of (2.5). The integrand in (2.6) is composed of tabulated functions and is of a form suited to numerical integration. Expanding $\varphi(y - a\cos\theta)$ in (2.6) as in the derivation

of (1.10), replacing the quantity within the brackets by the series shown in the last equation of (2.5) ,and integrating termwise leads to

$$N(I_1) = N_0(\pi/2)^{1/2} \sum_{n=0}^{\infty} \frac{\varphi^{(2n)}(y)}{n!n!} \left(\frac{a}{2}\right)^{2n} {}_1F_1\left(-\tfrac{1}{2}; n+1; -\frac{b^2}{2}\right) \tag{2.7}$$

The series (2.7) converges for all values of a, y, and b. This follows from the inequality (1.11) which may be applied to $\varphi^{(2n)}(y)$, and from the fact that the ${}_1F_1$ is less than $exp\ (b^2/2)$ as may be seen by comparing corresponding terms in their expansions.

The expected number of zeros, per second, of I is $2N(0)$ where we have set I_1, and hence y, equal to zero. In this case the integral in (2.6) may be simplified somewhat and we obtain

$$2N(0) = N_0\left[e^{-\alpha} I_0(\beta) + \frac{b^2}{2\alpha} Ie\left(\frac{\beta}{\alpha}, \alpha\right)\right] \tag{2.8}$$

where $I_0(\beta)$ is the Bessel function of order zero and imaginary argument and

$$\alpha = \frac{a^2 + b^2}{4}, \qquad \beta = \frac{a^2 - b^2}{4}$$

$$Ie(k, x) = \int_0^x e^{-u} I_0(ku)\, du.$$

The integral $Ie(k, x)$ is tabulated in Appendix I.

I have been unable to obtain a simple derivation of (2.8). It was originally obtained from the following integral

$$N(I_1) = \frac{N_0}{2} \int_{-\pi}^{\pi} d\theta\, \varphi(y - a\cos\theta) \int_0^{\infty} x\varphi(x + b\sin\theta)\, dx \tag{2.9}$$

which may be derived from the second equation of (2.4) and the first of (2.6). Setting I_1 and y equal to zero and writing out the φ's gives

$$2N(0) = \frac{N_0}{2\pi} \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dx\ x \exp\left[-\tfrac{1}{2}(x^2 + 2bx\sin\theta + a^2\cos^2\theta + b^2\sin^2\theta)\right].$$

Equation (2.8) was obtained by applying the method of Appendix III to this expression.

3. Definitions and Simple Properties of R and Θ

The remaining portion of this paper is concerned with the envelope R and the corresponding phase angle θ. These quantities are introduced and some of their simpler properties discussed in this section.

Suppose that the frequency band associated with I_N is relatively narrow

and contains the sine wave frequency f_q. The noise current may be resolved into two components, one "in phase" and the other "in quadrature" with $Q \cos qt$. Using the representation (2.8–6) of reference A and proceeding as in Section 3.7 of that paper:

$$I_N = \sum_{n=1}^{M} c_n \cos (\omega_n t - \varphi_n) \tag{3.1}$$

$$= \sum_{n=1}^{M} c_n \cos [(\omega_n - q)t - \varphi_n + qt]$$

$$= I_c \cos qt - I_s \sin qt \tag{3.2}$$

where

$$I_c = \sum_{n=1}^{M} c_n \cos [(\omega_n - q)t - \varphi_n]$$

$$I_s = \sum_{n=1}^{M} c_n \sin [(\omega_n - q)t - \varphi_n] \tag{3.3}$$

$$\omega_n = 2\pi f_n, \qquad f_n = n\Delta f, \qquad c_n^2 = 2w(f_n)\Delta f$$

$w(f)$ denotes the power spectrum of I_N and the φ_n's are random variables distributed uniformly over the interval $(0, 2\pi)$.

The total current I may be written as

$$I = Q \cos qt + I_N$$

$$= (Q + I_c) \cos qt - I_s \sin qt$$

$$= R \cos \theta \cos qt - R \sin \theta \sin qt \tag{3.4}$$

$$= R \cos (qt + \theta)$$

where we have introduced the envelope function R and the phase angle θ by means of

$$R \cos \theta = Q + I_c$$

$$R \sin \theta = I_s \tag{3.5}$$

Since I_c and I_s are functions of t whose variations are relatively slow in comparison with those of $\cos qt$, the same is true of R and (usually) θ.

A graphical illustration of equations (3.4) and (3.5) which is often used is shown in Fig. 4.

In accordance with the usual convention used in alternating current theory, the vector OQ is supposed to be rotating about the origin O with angular velocity q. If I_N happened to have the frequency $q/2\pi$, its vector representation QT would be fixed relative to OQ. In general, however, the

length and inclination of QT will change due to the random fluctuations of I_N. Thus the point T will wander around on the plane of the figure. If rms I_N is much less than Q, T will be close to the point Q most of the time. In this case

$$R = [(Q + I_c)^2 + I_s^2]^{1/2} \sim Q + I_c$$

$$\theta = \tan^{-1} \frac{I_s}{Q + I_c} \sim \frac{I_s}{Q} \tag{3.6}$$

$$\frac{d\theta}{dt} \sim \frac{d}{dt}\frac{I_s}{Q} = \frac{I_s'}{Q}$$

and a number of statistical properties of R and θ may be obtained from the corresponding properties of noise alone when we note that I_c, I_s, and I_s' behave like noise currents whose power spectra are concentrated in the lower portion of the frequency spectrum.

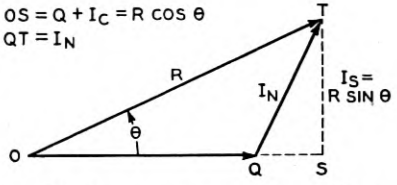


Fig. 4—Graphical representation of $I = Q\cos qt + I_N$.

By squaring both sides of equations (3.1) and (3.3) and then averaging with respect to t and the φ_n's we may show that I_c, I_s, and I_N all have the same rms value, namely $\psi_0^{1/2}$.

It may be seen from (3.3) that the power spectra of I_c and I_s are both given by

$$w(f_q + f) + w(f_q - f) \tag{3.7}$$

where it is assumed that $0 \leq f \ll f_q$. Likewise the power spectrum of the time derivative I_s' of I_s is

$$4\pi^2 f^2[w(f_q + f) + w(f_q - f)] \tag{3.8}$$

This follows from the representation of I_s' obtained by differentiating the expression (3.3) for I_s with respect to t, the procedure being the same as in the derivation of equation (7.2) in Section 7. The power spectra shown in Table 1 were computed from equations (3.7) and (3.8).

The correlation function for I_c, and hence also for I_s, is, from equations (A2–1) and (A2–3) of Appendix II,

$$\overline{I_c(t)I_c(t + \tau)} = g = \int_0^\infty w(f) \cos 2\pi(f - f_q)\tau \, df$$

where the bar denotes an average with respect to t and g is a function of τ. From (A2–3) the correlation function for I_s' is $-g''$ where the double prime on g denotes the second derivative with respect to τ.

Attention is sometimes fixed upon the variation in distance between successive zeros of I. The time between two successive zeros of I at, say, t_0 and t_1 is the time taken for $qt + \theta$, as appearing in $R\cos(qt + \theta)$, to increase by π. This assumes that the envelope R does not vanish in the interval. For the moment we write θ as $\theta(t)$ in order to indicate its dependence on the time t. Then t_0 and t_1 must satisfy the relation

$$qt_1 + \theta(t_1) - qt_0 - \theta(t_0) = \pi \tag{3.9}$$

Since $\theta(t)$ is a relatively slowly varying function we write

$$\theta(t_1) - \theta(t_0) = (t_1 - t_0)\theta'(t_0) + (t_1 - t_0)^2\theta''(t_0)/2 + \cdots$$

TABLE 1
POWER SPECTRA OF I_N, I_c, I_s, AND I_s'

I_N	I_c and I_s	I_s'
$w(f) = w_0 = \psi_0/\beta$ for $f_q - \beta/2 < f < f_q + \beta/2$ $w(f) = 0$ elsewhere f_q = mid-band frequency	$2w_0$ for $0 < f < \beta/2$ 0 elsewhere	$8\pi^2 f^2 w_0$ for $0 < f < \beta/2$ 0 elsewhere
$w(f) = w_0 = \psi_0/\beta$ for $f_q - \beta < f < f_q$ $w(f) = 0$ elsewhere f_q = top frequency	w_0 for $0 < f < \beta$ 0 elsewhere	$4\pi^2 f^2 w_0$ for $0 < f < \beta$ 0 elsewhere
$w(f) = \dfrac{\psi_0}{\sigma\sqrt{2\pi}} e^{-(f-f_q)^2/(2\sigma^2)}$	$\dfrac{2\psi_0}{\sigma\sqrt{2\pi}} e^{-f^2/(2\sigma^2)}$	$\dfrac{8\pi^2 f^2 \psi_0}{\sigma\sqrt{2\pi}} e^{-f^2/(2\sigma^2)}$

where the primes denote differentiation with respect to t. When this is placed in (3.9) and terms of order $(t_1 - t_0)^2$ neglected, we obtain

$$\frac{1}{2(t_1 - t_0)} = \frac{q}{2\pi} + \frac{1}{2\pi}\theta'(t_0) \tag{3.10}$$

which relates the interval between successive zeros to θ'.

The expression on the right hand side of (3.10) may be defined as the instantaneous frequency:

$$\text{Instantaneous frequency} = f_q + \frac{1}{2\pi}\frac{d\theta}{dt} \tag{3.11}$$

This definition is suggested when $\cos 2\pi ft$ is compared with $\cos(qt + \theta)$ and also by (3.10) when we note that the period of the instantaneous fre-

quency is approximately equal to twice the distance between two successive zeros which is $2(t_1 - t_0)$.

The probability density of R is[11]

$$\frac{R}{\psi_0} \exp\left[-\frac{R^2 + Q^2}{2\psi_0}\right] I_0(RQ/\psi_0) \tag{3.12}$$

where $I_0\,(RQ/\psi_0)$ denotes the Bessel function of order zero with imaginary argument. In Section 3.10 of Reference A, it is shown that the average value of R^n is*

$$\overline{R^n} = (2\psi_0)^{n/2}\Gamma\left(\frac{n}{2} + 1\right) {}_1F_1\left(-\frac{n}{2};\quad 1;\quad -\rho\right), \tag{3.13}$$

where $\rho = Q^2/(2\psi_0)$, of which special cases are

$$\begin{aligned} \bar{R} &= e^{-\rho/2}(\pi\psi_0/2)^{1/2}\,[(1 + \rho)I_0(\rho/2) + \rho I_1(\rho/2)] \\ \overline{R^2} &= Q^2 + 2\psi_0 \end{aligned} \tag{3.14}$$

Curves showing the distribution of R are also given there.

4. Expected Number of Crossings of R per Second

Here we shall obtain expressions for the expected number N_R of times, per second, the envelope passes through the value R with positive slope. When R is large, N_R is approximately equal to the expected number of maxima of the envelope per second exceeding R and when R is small N_R is approximately equal to the expected number of minima less than R. For the special case in which the noise band is symmetrical and is centered on the sine wave frequency f_q N_R is given by the relatively simple expression (4.8).

The probability that the envelope passes through the value R during the interval $t, t + dt$ with positive slope is, from (2.1),

$$dt \int_0^\infty R' p(R, R', t)\, dR' \tag{4.1}$$

where $p(R, R', t)$ denotes the probability density of R and its time derivative R', t being regarded as a parameter.

An expression for $p(R, R', t)$ may be obtained from the probability density of I_c, I_s, I_c', I_s'. From our representation of a noise current and the central limit theorem it may be shown (as is done for similar cases in Part III of Reference A) that the probability distribution of these four variables is

[11] In equation (60-A) of an unpublished appendix to his paper appearing in the *B.S.T.J.* Vol. 12 (1933), 35–75, Ray S. Hoyt gives an integral, obtained by integrating (3.12) with respect to R, for the cumulative distribution of R.

*The correlation function for the envelope of a signal plus noise, together with associated probability densities of the envelope and phase, is given by D. Middleton in a paper appearing soon in the Quart. Jl. of Appl. Math.

normal in four dimensions. If the variables be taken in the order given above the moment matrix is, from equations (A2–2) of Appendix II,

$$M = \begin{bmatrix} b_0 & 0 & 0 & b_1 \\ 0 & b_0 & -b_1 & 0 \\ 0 & -b_1 & b_2 & 0 \\ b_1 & 0 & 0 & b_2 \end{bmatrix} \tag{4.2}$$

where the b's are defined by the integrals in equations (A2–1). The inverse matrix is

$$M^{-1} = \frac{1}{B}\begin{bmatrix} b_2 & 0 & 0 & -b_1 \\ 0 & b_2 & b_1 & 0 \\ 0 & b_1 & b_0 & 0 \\ -b_1 & 0 & 0 & b_0 \end{bmatrix}, \quad B = b_0 b_2 - b_1^2 \tag{4.3}$$

which may be readily verified by matrix multiplication, and the determinant $| M |$ is B^2. The normal distribution may be written down at once when use is made of the formulas given in Section 2.9 of Reference A. The substitutions

$$\begin{aligned} I_c &= R\cos\theta - Q, & I_c' &= R'\cos\theta - R\sin\theta\,\theta' \\ I_s &= R\sin\theta, & I_s' &= R'\sin\theta + R\cos\theta\,\theta' \end{aligned} \tag{4.4}$$

$$dI_c dI_s dI_c' dI_s' = R^2 dR dR' d\theta d\theta'$$

enable us to write

$$\begin{aligned} b_2(I_c^2 + I_s^2) + b_0(I_c'^2 + I_s'^2) & \\ -2b_1(I_c I_s' - I_s I_c') &= b_2(R^2 - 2QR\cos\theta + Q^2) \\ &\quad + b_0(R'^2 + R^2\theta'^2) \\ &\quad - 2b_1 R^2\theta' + 2b_1 Q(R'\sin\theta + R\theta'\cos\theta). \end{aligned}$$

Consequently the probability density of R, R', θ, θ' is

$$\begin{aligned} p(R, R', \theta, \theta') = \frac{R^2}{4\pi^2 B} \exp\Big\{ -\frac{1}{2B}\,[b_2(R^2 - 2QR\cos\theta + Q^2) \\ + b_0(R'^2 + R^2\theta'^2) - 2b_1 R^2\theta' + 2b_1 Q(R'\sin\theta + R\theta'\cos\theta)] \Big\} \end{aligned} \tag{4.5}$$

In this expression R ranges from 0 to ∞, θ from $-\pi$ to π, and R' and θ' from $-\infty$ to $+\infty$. The probability density for R and R' is obtained by

integrating (4.5) with respect to θ and θ' over their respective ranges. The integration with respect to θ' may be performed at once giving

$$p(R, R', t) = \frac{R(2\pi)^{-3/2}}{\sqrt{Bb_0}} \int_{-\pi}^{\pi} d\theta \exp\left\{-\frac{1}{2Bb_0}\left[B(R^2 - 2RQ\cos\theta + Q^2) + (b_0R' + b_1Q\sin\theta)^2\right]\right\} \tag{4.6}$$

From (4.1) and (4.6) it follows that the expected number N_R of times per second the envelope passes through R with positive slope is

$$N_R = \frac{R(2\pi)^{-3/2}}{\sqrt{Bb_0}} \int_{-\pi}^{\pi} d\theta \int_0^{\infty} R'dR' \exp\left\{-\frac{1}{2Bb_0}\left[B(R^2 - 2RQ\cos\theta + Q^2) + (b_0R' + b_1Q\sin\theta)^2\right]\right\} \tag{4.7}$$

When the power spectrum $w(f)$ of the noise current I_N is symmetrical about the sine wave frequency f_q, b_1 is zero and B is equal to b_0b_2. In this case the integrations in (4.7) may be performed. We obtain

$$N_R = \left(\frac{b_2}{2\pi}\right)^{1/2} \frac{R}{b_0} I_0\left(\frac{RQ}{b_0}\right) \exp\left(-\frac{R^2 + Q^2}{2b_0}\right) = \left(\frac{b_2}{2\pi}\right)^{1/2} \times \left[\begin{array}{l}\text{Probability density of} \\ \text{envelope at the value } R\end{array}\right] \tag{4.8}$$

where the second line is obtained from expression (3.12). As will be seen from its definition (A2–1), b_0 is equal to the mean square value ψ_0 of I_N (and also of I_c and I_s).

Introducing the notation

$$\begin{aligned} v &= Rb_0^{-1/2} = R/\text{rms } I_N \\ a &= Ab_0^{-1/2} = Q/\text{rms } I_N, \end{aligned} \tag{4.9}$$

which is the same as that of equations (3.10–15) of Reference A except that there P denotes the amplitude of the sine wave and plays the same role as Q does here, enables us to write (4.8) as

$$N_R = \left[\frac{b_2}{2\pi b_0}\right]^{1/2} vI_0(av)e^{-(v^2+a^2)/2} = \left[\frac{b_2}{2\pi b_0}\right]^{1/2} p(v). \tag{4.10}$$

The function $p(v)$ is plotted as a function of v for various values of a in Fig. 6, Section 3.10, of Reference A.

It is interesting to note that

$$(b_2/b_0)^{1/2}/\pi = \text{Expected number of zeros per second of } I_c \text{ (or of } I_s) \tag{4.11}$$

This relation, which is true even if the noise band is not symmetrical about f_q, follows from equation (3.3–11) of Reference A.

When $Q \gg$ rms I_N and f_q is not at the center of the noise band it is easier to obtain the asymptotic form of N_R from the approximation (3.6),

$$R \sim Q + I_c,$$

instead of the double integral (4.7). When $Q >>$ rms I_N and R is in the neighborhood of Q (as it is most of the time in this case), N_R is approximately equal to the expected number of times I_c increases through the value $I_{c1} = R - Q$ in one second. Thus, regarding I_c as a random noise current we have from expression (3.3–14) of Reference A

$N_R \sim e^{-I_{c1}^2/(2b_0)} \times$ [1/2 the expected number of zeros per second of I_c] and when we use equation (4.11) we obtain

$$N_R \sim \frac{1}{2\pi}(b_2/b_0)^{1/2} e^{-(R-Q)^2/(2b_0)} = \frac{1}{2\pi}(b_2/b_0)^{1/2} e^{-(v-a)^2/2} \tag{4.12}$$

TABLE 2

$w(f) = w_0 = b_0/\beta$ OVER A BAND OF WIDTH β

	b_2	N_R
1. Band extends from $f_q - \beta/2$ to $f_q + \beta/2$	$\pi^2\beta^2 b_0/3$	$(\pi/6)^{1/2}\beta p(v) = 0.724\,\beta p(v)$
2. Same as 1 and in addition $Q = 0$	"	$(\pi/6)^{1/2}\beta v e^{-v^2/2}$
3. Same as 1 and in addition $Q >>$ rms I_N	"	$\sim \frac{\beta}{2\sqrt{3}} e^{-(v-a)^2/2}$
4. Band extends from f_q to $f_q + \beta$ and $Q >>$ rms I_N	$4\pi^2\beta^2 b_0/3$	$\sim \frac{\beta}{\sqrt{3}} e^{-(v-a)^2/2}$

Table 2 lists the forms assumed by (4.10) and (4.12) when the power spectrum $w(f)$ of the noise current I_N is constant over a frequency band of width β. The quantity b_0 in the expressions for b_2 represents the mean square value of I_N.

In the general case where the band of noise is not centered on f_q and where R is not large enough to make (4.12) valid we are obliged to return to the double integral (4.7). Some simplification is possible, but not as much as could be desired. Introducing the notation

$$\alpha = RQ/b_0, \qquad \gamma = b_1 Q(Bb_0)^{-1/2}$$

$$x = (b_0 R' + b_1 Q \sin\theta)(Bb_0)^{-1/2}$$

enables us to write (4.7) as

$$N_R = (2\pi)^{-3/2}(R/b_0)(B/b_0)^{1/2}e^{-(R^2+Q^2)/(2b_0)} \int_{-\pi}^{\pi} d\theta \int_{\gamma \sin\theta}^{\infty} (x - \gamma \sin\theta)e^{\alpha\cos\theta - x^2/2}\, dx \tag{4.13}$$

Part of the integrand may be integrated with respect to x and the remaining portion integrated by parts with respect to θ. The double integral in the second line of (4.13) then becomes

$$\begin{aligned}
&\int_{-\pi}^{\pi} e^{\alpha\cos\theta-(\gamma\sin\theta)^2/2}\, d\theta + \int_{-\pi}^{\pi}\left[\int_{\gamma\sin\theta}^{\infty} e^{-x^2/2}\, dx\right] d[\gamma\alpha^{-1}e^{\alpha\cos\theta}] \\
&\quad = \int_{-\pi}^{\pi} (1 + \gamma^2\alpha^{-1}\cos\theta)e^{\alpha\cos\theta-(\gamma\sin\theta)^2/2}\, d\theta \\
&\quad = \gamma\alpha^{-1}e^{-(\gamma^2+\alpha^2\gamma^{-2})/2} \int_{-\pi}^{\pi} (\gamma\cos\theta + \alpha/\gamma)e^{(\gamma\cos\theta+\alpha/\gamma)^2/2}\, d\theta \\
&\quad = 2\pi \sum_{n=0}^{\infty} \frac{(\frac{1}{2})_n}{n!}\left(-\frac{\gamma^2}{\alpha}\right)^n [I_n(\alpha) + \gamma^2\alpha^{-1}I_{n+1}(\alpha)].
\end{aligned} \tag{4.14}$$

The series is obtained by expanding $\exp[-(\gamma \sin\theta)^2/2]$ in the second equation in powers of $\sin\theta$ and integrating termwise.

5. Probability Density of $\frac{d\theta}{dt}$

As was pointed out in Section 3 the time derivative θ' of the phase angle θ associated with the envelope is closely related to the instantaneous frequency. The probability density $p(\theta')$ of θ' may be expressed in terms of modified Bessel functions as shown by equation (5.4). Curves for $p(\theta')$ are given when the sine wave frequency f_q lies at the middle of a symmetrical band of noise. Although the expressions for $p(\theta')$ are rather complicated, those for the averages $\bar{\theta}'$ and $|\overline{\theta'}|$ given by equations (5.7) and (5.16) are relatively simple.

The probability density $p(\theta')$ may be obtained by integrating the expression (4.5) for $p(R, R', \theta, \theta')$ with respect to R, R', θ. The integration with respect to R', the limits being $-\infty$ and $+\infty$, gives the probability density for R, θ, θ':

$$p(R,\theta,\theta') = \frac{R^2}{4\pi^2}\left(\frac{2\pi}{b_0 B}\right)^{1/2} \exp[-aR^2 + 2bR\cos\theta + c\sin^2\theta - b_2Q^2/(2B)] \tag{5.1}$$

where

$$B = b_0 b_2 - b_1^2 \qquad b = Q(b_2 - b_1\theta')/(2B)$$
$$a = (b_2 - 2b_1\theta' + b_0\theta'^2)/(2B) \qquad c = Q^2 b_1^2/(2Bb_0) = b_1^2\rho/B \tag{5.2}$$
$$\rho = Q^2/(2b_0) \qquad \gamma = b^2/a$$

and b_0, b_1, b_2 are given in Appendix II.

Integrating with respect to R gives the probability density for θ, θ'. Expanding $\exp(2bR\cos\theta)$ in powers of R and integrating termwise,

$$p(\theta, \theta') = \frac{1}{16\pi a}\left(\frac{2}{ab_0B}\right)^{1/2} e^{c\sin^2\theta - b_2b_0\rho/B} \sum_{n=0}^{\infty} \frac{n+1}{\Gamma\left(\frac{n}{2}+1\right)}\left(\frac{b\cos\theta}{a^{1/2}}\right)^n \tag{5.3}$$

When we integrate θ from $-\pi$ to π to obtain $p(\theta')$ the terms for which n is odd disappear and we have to deal with the series, writing γ for b^2/a,

$$\sum_{m=0}^{\infty} \frac{2m+1}{m!}(\gamma\cos^2\theta)^m = (2\gamma\cos^2\theta + 1)\exp(\gamma\cos^2\theta)$$

Thus, the probability density of θ' is

$$p(\theta') = \frac{1}{16\pi a}\left(\frac{2}{ab_0B}\right)^{1/2}\int_{-\pi}^{\pi}(2\gamma\cos^2\theta + 1)\, e^{c\sin^2\theta + \gamma\cos^2\theta - b_2b_0\rho/B}\, d\theta$$
$$= \frac{1}{8a}\left(\frac{2}{ab_0B}\right)^{1/2}\left[(\gamma+1)I_0\left(\frac{\gamma - c}{2}\right) + \gamma I_1\left(\frac{\gamma - c}{2}\right)\right] \tag{5.4}$$
$$\exp\left[\frac{c+\gamma}{2} - \frac{b_2b_0\rho}{B}\right]$$

From (5.2)

$$\gamma - c = \rho\frac{b_2 - 2b_1\theta'}{b_2 - 2b_1\theta' + b_0\theta'^2}$$
$$\frac{c+\gamma}{2} - \frac{b_2b_0\rho}{B} = -\frac{\rho}{2}\frac{b_2 - 2b_1\theta' + 2b_0\theta'^2}{b_2 - 2b_1\theta' + b_0\theta'^2} \tag{5.5}$$

It will be noted that for large values of $|\theta'|$ the probability density of θ' varies as $|\theta'|^{-3}$. Although this makes the mean square value of θ' infinite, the average values $\overline{\theta'}$ and $\overline{|\theta'|}$ of θ' and $|\theta'|$ still exist. In order to obtain $\overline{\theta'}$ it is convenient to return to (4.5) and write

$$\overline{\theta'} = \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dR \int_{-\infty}^{+\infty} dR' \int_{-\infty}^{\infty} d\theta'\, \theta' p(R, R', \theta, \theta') \tag{5.6}$$

The integration with respect to θ' may be performed by setting $R\theta'$ equal to x and using

$$\int_{-\infty}^{+\infty} x\, e^{-\alpha x^2+2\beta x}\, dx = (\beta/\alpha)(\pi/\alpha)^{1/2}\, e^{\beta^2/\alpha}$$

The integral in R' reduces to a similar integral except that the factor x in the integrand is absent. Performing these two integrations and using the definition of B leads to

$$\bar{\theta}' = \frac{1}{2\pi}\frac{b_1}{b_0^2}\int_{-\pi}^{\pi} d\theta \int_0^{\infty} dR\,(R - Q\cos\theta) \cdot \exp\left[-\frac{1}{2b_0}(R^2 - 2QR\cos\theta + Q^2)\right]$$

We may integrate at once with respect to R. When this is done $\cos\theta$ disappears and the integration with respect to θ becomes easy. Thus

$$\bar{\theta}' = (b_1/b_0)\exp[-Q^2/(2b_0)] = (b_1/b_0)e^{-\rho} \tag{5.7}$$

When the noise power spectrum is equal to w_0 in the band extending from $f_0 - \beta/2$ to $f_0 + \beta/2$ and is zero outside the band, $b_1 = 2\pi(f_0 - f_q)b_0$. Hence, from (3.11),

$$\begin{aligned}\text{ave. instantaneous frequency} &= f_q + \bar{\theta}'/(2\pi)\\ &= f_0 + (f_q - f_0)(1 - e^{-\rho})\end{aligned} \tag{5.8}$$

In the remainder of this section we assume the power spectrum of the noise current to be symmetrical about the sine wave frequency f_q. In this case b_1 and c are zero, B is equal to b_0b_2 and (5.4) becomes

$$\begin{aligned}p(\theta') &= \tfrac{1}{2}(b_0/b_2)^{1/2}(1 + z^2)^{-3/2}e^{-\rho+y/2}\\ &\quad [(y+1)I_0(y/2) + yI_1(y/2)]\\ &= \tfrac{1}{2}(b_0/b_2)^{1/2}(1+z^2)^{-3/2}e^{-\rho}\,{}_1F_1\left(\frac{3}{2};\,1;\,y\right)\end{aligned} \tag{5.9}$$

where ${}_1F_1$ denotes a confluent hypergeometric function[12] and

$$z^2 = b_0\theta'^2/b_2, \qquad y = (\gamma)_{b_1=0} = \rho/(1+z^2) \tag{5.10}$$

When the noise power spectrum is constant in the band extending from $f_q - \beta/2$ to $f_q + \beta/2$ (see Table 2, Section 4)

$$(b_2/b_0)^{1/2} = 3^{-1/2}\beta\pi, \qquad z = 3^{1/2}\theta'/(\beta\pi) \tag{5.11}$$

[12] The relation used above follows from equation (66) (with misprint corrected) of W. R. Bennett's paper cited in connection with equation (1.2).

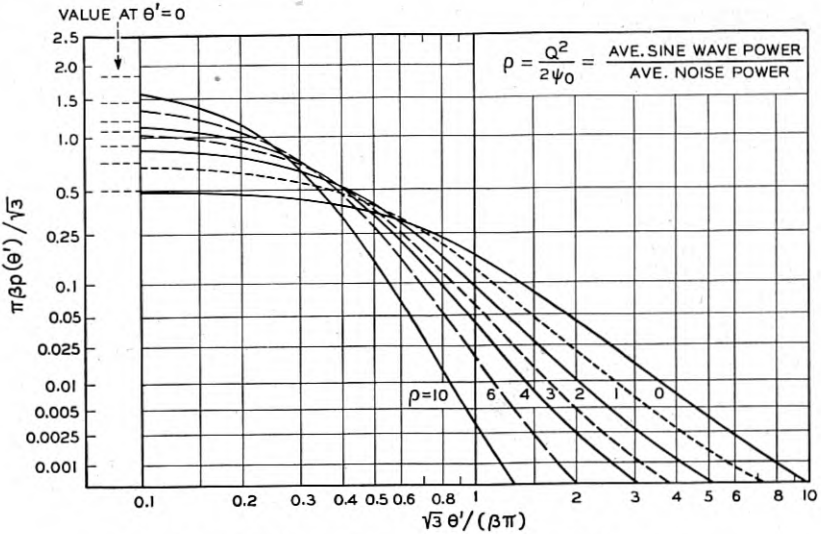


Fig. 5—Probability density of time derivative of phase angle.

$p(\theta')\,d\theta'$ = probability that the value of $d\theta/dt$ at an instant selected at random lies between θ' and $\theta' + d\theta'$. The power spectrum of I_N is constant in band of width β centered on f_q and is zero outside this band.

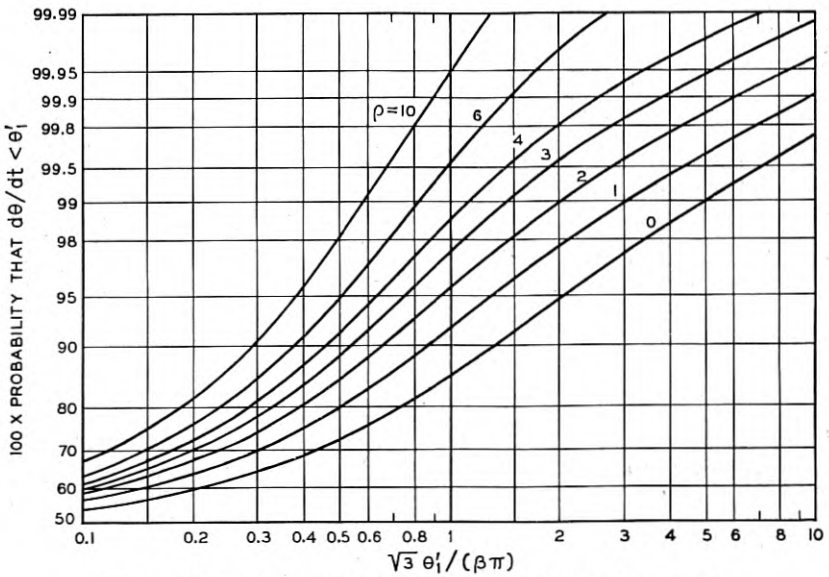


Fig. 6—Cumulative distribution of time derivative of phase angle.

Notation explained in Fig. 5.

The probability density $p(\theta')$ of θ' and its cumulative distribution, obtained by numberical integration, are shown in Figs. 5 and 6.

The probability that θ' exceeds a given θ_1' is equal to the probability that z exceeds z_1, where z_1 denotes $(b_0/b_2)^{1/2}\theta_1'$, and both probabilities are equal to

$$\frac{e^{-\rho}}{2}\int_{z_1}^{\infty}(1+z^2)^{-3/2}\,{}_1F_1\left[\frac{3}{2};\,1;\,\rho(1+z^2)^{-1}\right]dz \tag{5.12}$$

The probability that $\theta' > \theta_1'$ becomes $e^{-\rho}/(4z_1^2)$ as $\theta_1' \to \infty$.

When $Q \gg$ rms I_N the leading term in the asymptotic expansion of the ${}_1F_1$ in (5.9) gives

$$p(\theta') \sim \frac{1}{\sigma\sqrt{2\pi}}e^{-\theta'^2/(2\sigma^2)}, \qquad \sigma^2 = b_2/Q^2 \tag{5.13}$$

when it is assumed that $z^2 \ll 1$. This expression holds only for the central portion of the curve for $p(\theta')$. Far out on the curve, $p(\theta')$ still varies as θ'^{-3}. Equation (5.13) may be obtained directly by using the approximation (3.6) that θ' is nearly equal to I_s'/Q and noticing that b_2 is the mean square value of I_s'.

If the sine wave is absent, ρ is zero and

$$p(\theta') = \tfrac{1}{2}(b_0/b_2)^{1/2}(1+z^2)^{-3/2} \tag{5.14}$$

which is consistent with the results given between equations (3.4–10) and (3.4–11) of Reference A. In this case (5.12) becomes

$$\frac{1}{2} - \frac{z_1}{2}(1+z_1^2)^{-1/2} \tag{5.15}$$

Although the standard deviation of θ' is infinite an idea of the spread of the distribution may be obtained from the average value of $|\theta'|$. Setting b_1 equal to zero in (4.5) in order to obtain the case in which the noise band is symmetrical about the sine-wave frequency leads to

$$\overline{|\theta'|} = \frac{2}{4\pi^2 b_0 b_2}\int_0^{\infty}dR\int_{-\pi}^{\pi}d\theta\int_{-\infty}^{+\infty}dR'\int_0^{\infty}d\theta'\,\theta'\,R^2$$
$$\exp\tfrac{1}{2}\left[-(R^2 - 2QR\cos\theta + Q^2)/b_0 - (R'^2 + R^2\theta'^2)/b_2\right]$$

The integrals in R', θ' cause no difficulty and the integral in θ is proportional to the Bessel function $I_0(QR/b_0)$. When the resulting integral in R is evaluated[13] we obtain

$$\overline{|\theta'|} = (b_2/b_0)^{1/2}e^{-\rho/2}I_0(\rho/2) \tag{5.16}$$

where $\rho = Q^2/(2b_0)$.

[13] See, for example, G. N. Watson, "Theory of Bessel Functions," Cambridge (1944), p. 394, equation (5).

When ρ is zero equation (5.16) agrees with a result given in Section 3.4 of reference A, namely, for an ideal band pass filter

$$\frac{\text{ave}\,|\,\tau - \tau_1\,|}{\tau_1} = \frac{f_b - f_a}{\sqrt{3}(f_b + f_a)}$$

where τ is the interval between two successive zeros and τ_1 is its average value. τ is equal to $t_1 - t_0$ of our equation (3.10) from which it follows that

$$(\tau - \tau_1)/\tau_1 \approx -\theta'/q \tag{5.17}$$

6. Expected Number of Crossings of θ and $\frac{d\theta}{dt}$ per Second

After a brief study of the expected number of times per second the phase angle θ increases through 0 and through π (where it is assumed that $-\pi < \theta \leq \pi$) expressions are obtained for the expected number $N_{\theta'}$ of times per second the time derivative of θ increases through the value θ'.

The point T shown in Fig. 4 of Section 3 wanders around, as time goes by, in the plane of the figure. How many times may we expect it to cross some preassigned section of the line OQ in one second? To answer this problem we note that, from expression (2.1), the probability that θ increases through zero during the interval t, $t + dt$ with the envelope lying between R and $R + dR$ is

$$dt\, dR \int_0^\infty \theta' p(R, 0, \theta')\, d\theta' \tag{6.1}$$

where the probability density in the integrand is obtained by setting θ equal to zero in equation (5.1). The expected number of such crossings per second is

$$(2\pi)^{-3/2}\,(b_0 B)^{-1/2} R^2\, dR e^{-b_2(R-Q)^2/(2B)} \int_0^\infty d\theta'\, \theta' \exp\,[-b_0 R^2 \theta'^2/(2B) + b_1 R(R - Q)\theta'/B] \tag{6.2}$$

which may be evaluated in terms of error functions or the function $\varphi_{-1}(x)$ defined by equation (1.8). For the special case in which the power spectrum of the noise current I_N is symmetrical about the sine wave frequency, b_1 is zero and (6.2) yields

$$(2\pi)^{-3/2} b_0^{-1} b_2^{1/2} e^{-(R-Q)^2/(2b_0)}\, dR \tag{6.3}$$

From equation (6.1) onwards we have tacitly assumed that the range of θ is given by $-\pi < \theta \leq \pi$ because setting θ equal to any multiple of 2π in our equations leads to the same result as setting θ equal to zero. This is due to θ occurring only in cos θ and sin θ. When θ increases through the value π,

as it does when the point T crosses, in the downward direction, the extension of the line OQ lying to the left of the point O in Fig. 4, we imagine the value of θ to change discontinuously to the value $-\pi$.

The expected number of times per second θ increases through π may be obtained from (6.2) and, in the symmetrical case, from (6.3) by changing the sign of Q since this produces the same effect as changing θ from 0 to π in $p(R, \theta, \theta')$.

The expected number of crossings per second when R lies between two assigned values may be obtained by integrating the above equations. For example, the number of times per second θ increases through zero with R between Q and R_1 is, from (6.3) for the symmetrical case,

$$(4\pi)^{-1}(b_2/b_0)^{1/2} \operatorname{erf} [(2b_0)^{-1/2} \mid R_1 - Q \mid] \tag{6.4}$$

where we have used the absolute value sign to indicate that R_1 may be either less than or greater than Q and

$$\operatorname{erf} x = 2\pi^{-1/2} \int_0^x e^{-t^2}\, dt \tag{6.5}$$

Expressions for b_0 and b_2 are given by equations (A2–1) of Appendix II. The mean square value of I_N is b_0, and when the power spectrum of I_N is constant over a band of width β, $b_2 = \pi^2\beta^2 b_0/3$.

In much the same way it may be shown that the expected number of times per second θ increases through π with R between 0 and R_1 is

$$(4\pi)^{-1}(b_2/b_0)^{1/2} \{\operatorname{erf} [(2b_0)^{-1/2}(R_1 + Q)] - \operatorname{erf} [(2b_0)^{-1/2}Q]\} \tag{6.6}$$

A check on these equations may be obtained by noting that the expected number of zeros per second of I_s, given by equation (4.11), is equal to twice the number of times θ increases through zero plus twice the number of times θ increases through π. Setting R_1 equal to zero in (6.4), to infinity in both (6.4) and (6.6), and adding the three quantities obtained gives half of (4.11), as it should.

Now we shall consider the crossings of θ'. The equations in the first part of the analysis are quite similar to those encountered in Section 3.8 of Reference A where the maxima of R, for noise alone, are discussed. We start by introducing the variables $x_1, x_2, \cdots x_6$ where

$$x_1 = I_c = R\cos\theta - Q, \qquad x_4 = I_s = R\sin\theta \tag{6.7}$$

and the remaining x's are defined in terms of the derivatives of I_c and I_s and are given by the equations just below (3.8–4) of Reference A.

Here we shall consider the noise band to be symmetrical about the sine

wave frequency f_q so that b_1 and b_3 are zero. Then from equations (3.8–3) and (3.8–4) of Reference A the probability density of $x_1, x_2, \cdots x_6$ is

$$\frac{1}{8\pi^3 b_2 D} \exp\left(-\frac{1}{2D}[b_4(x_1^2 + x_4^2) + 2b_2(x_1x_3 + x_4x_6) + (D/b_2)(x_2^2 + x_5^2) + b_0(x_3^2 + x_6^2)]\right) \tag{6.8}$$

where $D = b_0b_4 - b_2^2$ and the b_n's are given by equations (A2–1). Replacing the x's by their expressions in terms of R and θ, similar to those just above equation (3.8–5) of Reference A, shows that the probability density for R, R', R'', θ, θ', θ'' is

$$\begin{aligned} p(R, R', R'', \theta, \theta', \theta'') = \frac{R^3}{8\pi^3 b_2 D} \exp\Big(-\frac{1}{2D}[b_4(R^2 - 2RQ\cos\theta + Q^2) \\ + (D/b_2)(R'^2 + R^2\theta'^2) + 2b_2(RR'' - R^2\theta'^2) \\ + b_0(R''^2 - 2RR''\theta'^2 + 4R'^2\theta'^2 + 4RR'\theta'\theta'' + R^2\theta'^4 + R^2\theta''^2) \\ - 2b_2Q(R''\cos\theta - R\theta'^2\cos\theta - 2R'\theta'\sin\theta - R\theta''\sin\theta)]\Big) \end{aligned} \tag{6.9}$$

It must be remembered that (6.9) applies only to the symmetrical case in which b_1 and b_3 are zero.

Integrating R' and R'' in (6.9) from $-\infty$ to ∞ gives the probability density of R, θ, θ', θ''. The integration with respect to R'' is simplified by changing to the variable $R'' - R\theta'^2$. The result is

$$\begin{aligned} p(R, \theta, \theta', \theta'') = R^3(2\pi)^{-2}(b_0b_2D)^{-1/2}(1 + u)^{-1/2} \\ \exp\left(-\frac{1}{2b_0}\left[R^2 - 2RQ\cos\theta + Q^2 + b_0R^2\theta'^2/b_2 + \frac{(Qb_2\sin\theta + b_0R\theta'')^2}{(1 + u)D}\right]\right) \end{aligned} \tag{6.10}$$

where $u = 4b_2b_0\theta'^2/D$. The expected number of times per second the time derivative of θ increases through the value θ' is

$$\begin{aligned} N_{\theta'} &= \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dR \int_0^{\infty} d\theta'' \theta'' p(R, \theta, \theta', \theta'') \\ &= \pi^{-2}(b_2\delta/b_0)^{1/2} \int_{-\pi}^{\pi} d\theta \int_0^{\infty} r dr \int_0^{\infty} x dx \\ &\quad \exp[-\gamma r^2 + 2r\alpha\cos\theta - \alpha^2 - \delta(x + \alpha\sin\theta)^2] \end{aligned} \tag{6.11}$$

where we have set

$$r = R(2b_0)^{-1/2} \qquad x = rb_0\theta''/b_2$$
$$\alpha = Q(2b_0)^{-1/2} = \rho^{1/2} \qquad \gamma = 1 + b_0\theta'^2/b_2 = 1 + z^2 \tag{6.12}$$
$$\delta = \frac{b_2^2}{(1+u)D}$$

r being regarded as a constant when the variable of integration is changed from θ'' to x.

The double integral in θ and x occurring in (6.11) is of the same form as (A3–1) of Appendix III and hence may be transformed into (A3–3). Here $a = r\alpha$, $c = -\delta\alpha^2$, $c + b^2 = 0$. The diameter of the path of integration C may be chosen so large that the order of integration may be interchanged and the integration with respect to r performed. The result is again an integral of the form (A3–3) in which $a^2 = 0$. When this is reduced to (A3–6) it becomes

$$N_{\theta'} = e^{-\rho}(2\pi\gamma)^{-1}b_2^{1/2}(b_0\delta)^{-1/2}\,[e^{-\delta\rho/2}I_0(\delta\rho/2) + (1+\gamma\delta)(1+\gamma\delta/2)^{-1}e^{\rho/\gamma}Ie\,\{\gamma\delta(2+\gamma\delta)^{-1},\ \rho/\gamma + \delta\rho/2\}] \tag{6.13}$$

where we have used $Ie(-k, x) = Ie(k, x)$ which follows from the definition (A1–1) given in Appendix I.

When there is no sine wave present, ρ is zero and (6.13) becomes

$$N_{\theta'} = \frac{1}{2\pi\gamma}\left(\frac{b_2}{b_0\delta}\right)^{1/2} = \frac{\sqrt{\dfrac{b_4}{b_2} - \dfrac{b_2}{b_0} + 4\theta'^2}}{2\pi\left(1 + \dfrac{b_0}{b_2}\theta'^2\right)} \tag{6.14}$$

This gives a partial check on some of the above analysis since (6.14) may be obtained immediately by setting α equal to zero in (6.11). Another check may be obtained by letting $\rho \to \infty$ and using $Ie(k, \infty) = (1 - k^2)^{-1/2}$. (6.13) becomes

$$N_{\theta'} \sim (2\pi)^{-1}(b_4/b_2)^{1/2}e^{-\rho z^2} \tag{6.15}$$

which agrees with the result obtained from $\theta' \sim I_s'/Q$.

For the case in which the power spectrum $w(f)$ of the noise is equal to the constant value w_0 over the frequency band extending from $f_q - \beta/2$ to $f_q + \beta/2$,

$$b_0 = \beta w_0, \quad b_2 = \pi^2\beta^3 w_0/3 = \pi^2\beta^2 b_0/3, \quad b_4 = \pi^4\beta^5 w_0/5 = \pi^4\beta^4 b_0/5 \tag{6.16}$$

These lead to

$$z = (b_0/b_2)^{1/2}\theta' = 3^{1/2}\theta'/(\pi\beta) \qquad D/b_2^2 = b_4 b_0/b_2^2 - 1 = 9/5 - 1 = 4/5$$

$$u = 4\, b_2^2 z^2/D = 5z^2 \qquad \delta = \frac{5}{4(1 + 5z^2)} \tag{6.17}$$

$$\gamma = 1 + z^2$$

and the coefficient in (6.13) may be simplified by means of

$$\frac{1}{2\pi\gamma}\left(\frac{b_2}{b_0\delta}\right)^{1/2} = \frac{\beta}{1 + z^2}\left(\frac{1 + 5z^2}{15}\right)^{1/2} \tag{6.18}$$

From (6.14) we see that (6.18) is equal to $N_{\theta'}$ when noise alone is present (and is of constant strength in the band of width β). The curves of $N_{\theta'}/\beta$ versus z shown in Fig. 7 were obtained by setting (6.17) and (6.18) in (6.13). $N_{\theta'}/\beta$ approaches $e^{-\rho}/(z\sqrt{3})$ as $z \to \infty$.

When the wandering point T of Fig. 4 passes close to the point O, θ changes rapidly by approximately π and produces a pulse in θ'. In discussions of frequency modulation θ' is sometimes regarded as a noise voltage which is applied to a low pass filter. Although the closer T comes to O the higher the pulse, the area under the pulse will be of the order of π and the response of the low pass filter may be calculated approximately.

That the pulses in θ' arise in the manner assumed above may be checked as follows. We choose a point relatively far out on the curve for $\rho = 5$ in Fig. 7, say $z = \sqrt{3}\theta'/(\beta\pi) = 1.6$ or $\theta' = 2.9\beta$. The number of pulses per second having peaks higher than 2.9β is roughly $N_{\theta'} = .009\beta$, and half of these have peaks greater than $\theta' = 3.8\beta$ which is obtained from Fig. 7 for $N_{\theta'} = .0045\beta$. From Fig. 6 we see that θ' exceeds 2.9β about .0018 of the time. Thus the average width at the height 2.9β of the class of pulses whose peaks exceed this value is $.0018/(.009\beta) = .2/\beta$ seconds. This figure is to be checked by the width obtained from the assumption that the typical pulse arises when T moves along a straight line with speed v and passes within a distance b of O. We take $\tan\theta = vt/b = \alpha t$ so that $\theta' = \alpha/(1 + \alpha^2 t^2)$. From this expression for θ' it follows that a pulse of peak height 3.8β (the median height) has a width of $.3/\beta$ seconds at $\theta' = 2.9\beta$. This agreement seems to be fairly good in view of the roughness of our work. A similar comparison may be made for $\rho = 0$ by using the limiting forms of (5.15) and (6.18). Here it is possible to compute the average width instead of estimating it from the median peak value. Exact agreement is obtained, both methods leading to an average width of $\pi/(4\theta')$ seconds at height θ'.

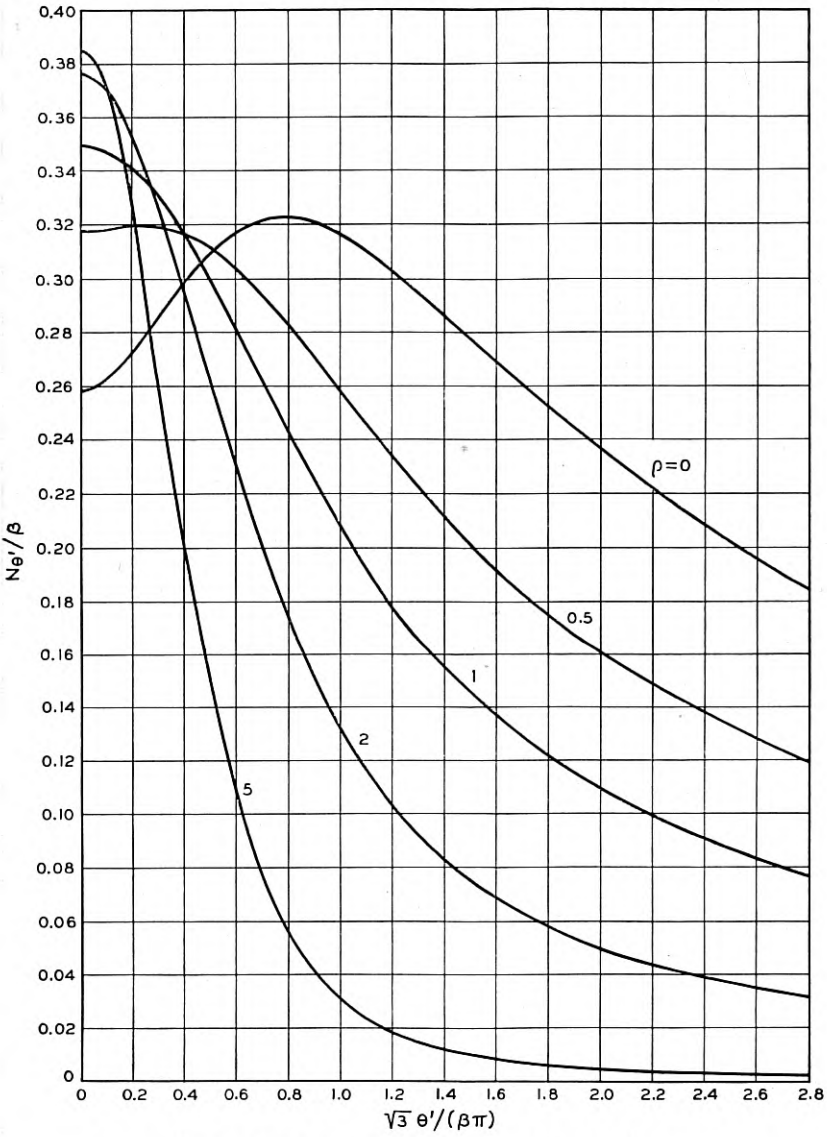


Fig. 7—Crossings of time derivative of phase angle.

$N_{\theta'}$ = expected number of times per second $d\theta/dt$ increases through the value θ'. ρ, β, and the power spectrum of I_N are the same as in Fig. 5.

7. Correlation Function for $\frac{d\theta}{dt}$

In this section we shall compute the correlation function $\Omega(\tau)$ of $\theta'(t)$. We are primarily interested in $\Omega(\tau)$ because it is, according to a fundamental result due to Wiener, the Fourier transform[14] of the power spectrum $W(f)$ of $\theta'(t)$.

We shall first consider the case in which the sine wave power is very large compared with the noise power so that, from (3.6), θ is approximately I_s/Q and θ' approximately I_s'/Q. Then using (A2–3) and (A2–1)

$$\begin{aligned}\Omega(\tau) &= \overline{\theta'(t)\theta'(t+\tau)} \approx Q^{-2}\overline{I_s'(t)I_s'(t+\tau)} \\ &= -g''Q^{-2} = 4\pi^2Q^{-2}\int_0^\infty w(f)(f-f_q)^2\cos 2\pi(f-f_q)\tau\, df\end{aligned} \tag{7.1}$$

When $w(f)$ is effectively zero outside a relatively narrow band in the neighborhood of f_q, as it is in the cases with which we shall deal, (7.1) leads to the relation (divide the interval $(0, \infty)$ into $(0, f_q)$ and (f_q, ∞), introduce new variables of integration $f_1 = f_q - f, f_2 = f - f_q$ in the respective intervals, replace the upper limit f_q of the first integral by ∞, combine the integrals, and compare with (2.1–6) of Reference A)

Power spectrum of $\theta'(t) = W(f)$

$$\approx 4\pi^2f^2Q^{-2}[w(f_q+f) + w(f_q-f)] \tag{7.2}$$

This form is closely related to results customarily used in frequency modulation studies. It should be remembered that in (7.2) it is assumed that $0 < f \ll f_q$ and rms $I_N \ll Q$.

Additional terms in the approximation for $\Omega(\tau)$ may be obtained by expanding

$$\theta = \arctan\frac{I_s}{Q+I_c}$$

in descending powers of Q, multiplying two such series (one for time t and the other for time $t+\tau$) together, and averaging over t. If I_{c1}, I_{s1} and I_{c2}, I_{s2} denote the values of I_c, I_s at times t and $t+\tau$ respectively, the average values of the products of the I's may be obtained by expanding the characteristic function (obtainable from equation (7.5) given below by setting $z_5 = z_6 = z_7 = z_8 = 0$) of the four random variables $I_{c1}, I_{s1}, I_{c2}, I_{s2}$. This method is explained in Section 4.10 of Reference A. When $w(f)$ is symmetrical about f_q it is found that

[14] The form which we shall use is given by equation (2.1–5) of Reference A.

$$\overline{\theta_1\theta_2} = \frac{g}{Q^2} + \frac{g^2}{Q^4} + \frac{8g^3}{3Q^6} + \cdots$$

$$\Omega(\tau) = \overline{\theta_1'\theta_2'} = -\frac{d^2}{d\tau^2}\overline{\theta_1\theta_2} \tag{7.3}$$

$$= -\frac{g''}{Q^2} - \frac{2}{Q^4}(gg'' + g'^2) - \frac{8}{Q^6}(g^2g'' + 2g'^2g) + \cdots$$

From the exact expression for $\Omega(\tau)$ obtained below it is seen that the last equation in (7.3) is really asymptotic in character and the series does not converge. We infer that this is also true for the first equation of (7.3).

We shall now obtain the exact expression for the correlation function $\Omega(\tau)$ of $\theta'(t)$ when f_q is at the center of a symmetrical band of noise. At first sight it would appear that the easiest procedure is to calculate the correlation function for $\theta(t)$ and then obtain $\Omega(\tau)$ by differentiating twice. However, difficulties present themselves in getting θ outside the range $-\pi, \pi$ since θ enters the expressions only as the argument of trigonometrical functions. Because I could not see any way to overcome this difficulty I was forced to deal with θ' directly. Unfortunately this increases the complexity since now the distribution of the time derivatives of I_c and I_s also must be considered.

We have

$$\tan\theta = \frac{I_s}{Q + I_c}, \qquad \sec^2\theta = 1 + \left(\frac{I_s}{Q + I_c}\right)^2$$

$$\theta' = \frac{(Q + I_c)I_s' - I_sI_c'}{\sec^2\theta(Q + I_c)^2} = \frac{(Q + I_c)I_s' - I_sI_c'}{(Q + I_c)^2 + I_s^2}$$

and the value of $\overline{\theta'(t)\theta'(t+\tau)}$ is the eight-fold integral

$$\Omega(\tau) = \int_{-\infty}^{+\infty} dI_{c1} \cdots \int_{-\infty}^{+\infty} dI_{s2}'\, p(I_{c1}, \cdots, I_{s2}') \frac{(Q + I_{c1})I_{s1}' - I_{s1}I_{c1}'}{(Q + I_{c1})^2 + I_{s1}^2} \times \frac{(Q + I_{c2})I_{s2}' - I_{s2}I_{c2}'}{(Q + I_{c2})^2 + I_{s2}^2} \tag{7.4}$$

where $p(I_{c1}, \cdots, I_{s2}')$ is an eight-dimensional normal probability density. As before, the subscripts 1 and 2 refer to times t and $t + \tau$, respectively. The most direct way of evaluating the integral (7.4) is to insert the expression for $p(I_{c1}, \cdots, I_{s2}')$ and then proceed with the integration. Indeed, this method was used the first time the integral (7.4) was evaluated. Later it was found that the algebra could be simplified by representing $p(I_{c1}, \cdots, I_{s2}')$ as the Fourier transform of its characteristic function. The second procedure will be followed here.

The characteristic function for $I_{c1}, I_{c2}, I_{s1}, I_{s2}, I'_{c1}, I'_{c2}, I'_{s1}, I'_{s2}$ is, from (A2–2) and (A2–3) of Appendix II and Section 2.9 of Reference A,

$$\begin{aligned}\text{ave. } \exp i[z_1 I_{c1} + z_2 I_{c2} + z_3 I_{s1} + z_4 I_{s2} + z_5 I'_{c1} + z_6 I'_{c2} + z_7 I'_{s1} + z_8 I'_{s2}] \\ = \exp\left(-\tfrac{1}{2}\right)[b_0(z_1^2 + z_2^2 + z_3^2 + z_4^2) + b_2(z_5^2 + z_6^2 + z_7^2 + z_8^2) \\ +2b_1(z_1 z_7 + z_2 z_8 - z_3 z_5 - z_4 z_6) \\ +2g(z_1 z_2 + z_3 z_4) + 2g'(z_1 z_6 - z_2 z_5 + z_3 z_8 - z_4 z_7) \\ -2g''(z_5 z_6 + z_7 z_8) + 2h(z_1 z_4 - z_2 z_3) \\ +2h'(z_1 z_8 + z_2 z_7 - z_3 z_6 - z_4 z_5) - 2h''(z_5 z_8 - z_6 z_7)].\end{aligned} \tag{7.5}$$

Since we have included b_1, h, h', h'' this holds when f_q is not necessarily at the center of the noise band. However, henceforth we return to our assumption that f_q is placed at the center of a symmetrical noise band and take b_1, h, h', h'' to be zero.

The probability density of $I_{c1}, \cdots I'_{s2}$ which is to be placed in (7.4) is the eight-fold integral

$$\begin{aligned}p(I_{c1}, \cdots I'_{s2}) = (2\pi)^{-8} \int_{-\infty}^{+\infty} dz_1 \cdots \int_{-\infty}^{+\infty} dz_8 \\ \exp[-iz_1 I_{c1} - \cdots - iz_8 I'_{s2}] \times [\text{ch.f.}]\end{aligned} \tag{7.6}$$

where "ch.f." denotes the characteristic function obtained by setting b_1, h, h', h'' equal to zero on the right hand side of (7.5).

The integral (7.4) for $\Omega(\tau)$ may be written as

$$\Omega(\tau) = J_1 - J_2 - J_3 + J_4 \tag{7.7}$$

where J_1 is the 16-fold integral

$$\begin{aligned}J_1 = \int_{-\infty}^{+\infty} dI_{c1} \cdots \int_{-\infty}^{+\infty} dI'_{s2} (2\pi)^{-8} \int_{-\infty}^{+\infty} dz_1 \cdots \int_{-\infty}^{+\infty} dz_8 \\ \exp[-iz_1 I_{c1} - \cdots - iz_8 I'_{s2}] \\ \frac{(Q + I_{c1})(Q + I_{c2}) I'_{s1} I'_{s2}}{[(Q + I_{c1})^2 + I_{s1}^2][(Q + I_{c2})^2 + I_{s2}^2]} \times [\text{ch.f.}]\end{aligned} \tag{7.8}$$

and J_2, J_3, J_4 are obtained from J_1 by replacing the product $(Q + I_{c1})(Q + I_{c2}) I'_{s1} I'_{s2}$ by $(Q + I_{c1}) I'_{s1} I_{s2} I'_{c2}$, $I_{s1} I'_{c1} (Q + I_{c2}) I'_{s2}$, $I_{s1} I'_{c1} I_{s2} I'_{c2}$ respectively.

The integration with respect to I_{c1} and I_{s1} in (7.8) may be performed at once. We replace $Q + I_{c1}$ and I_{s1} by x and y, respectively, and use

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \, \frac{x}{x^2 + y^2} e^{-izx - i\zeta y} = \frac{-2\pi i z}{z^2 + \zeta^2}. \tag{7.9}$$

The integration with respect to I_{c2} and I_{s2} may be performed in a similar manner. In this way we obtain a 12-fold integral.

The integrations with respect to the I''s may be performed by using

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} dI \int_{-\infty}^{+\infty} e^{-izI} f(z)\, dz = f(0)$$
$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} I\, dI \int_{-\infty}^{+\infty} e^{-izI} f(z)\, dz = -i\left[\frac{df(z)}{dz}\right]_{z=0} \tag{7.10}$$

The result is the four-fold integral

$$J_1 = (2\pi)^{-2}\int_{-\infty}^{+\infty} dz_1 \cdots \int_{-\infty}^{+\infty} dz_4 \frac{z_1 z_2(g'' - g'^2 z_3 z_4)}{(z_1^2 + z_3^2)(z_2^2 + z_4^2)} \tag{7.11}$$
$$\exp\left[-(b_0/2)(z_1^2 + z_2^2 + z_3^2 + z_4^2) - g(z_1 z_2 + z_3 z_4) + iQ(z_1 + z_2)\right].$$

In the same way J_2, J_3, J_4 may be reduced to the integrals obtained from (7.11) by replacing $z_1 z_2(g'' - g'^2 z_3 z_4)$ by $-g'^2 z_1^2 z_4^2$, $-g'^2 z_2^2 z_3^2$ and $z_3 z_4(g'' - -g'^2 z_1 z_2)$, respectively. When the J's are combined in accordance with (7.7) we obtain an integral which may be obtained from (7.11) by replacing $z_1 z_2(g'' - g'^2 z_3 z_4)$ by

$$g''(z_1 z_2 + z_3 z_4) + g'^2(z_1 z_4 - z_2 z_3)^2 \tag{7.12}$$

The terms $z_1^2 + z_3^2$ and $z_2^2 + z_4^2$ in the denominator may be represented as infinite integrals. Interchanging the order of integration and expressing (7.12) in terms of partial derivatives of an exponential function leads to the six-fold integral

$$\Omega(\tau) = (4\pi)^{-2}\int_0^\infty du \int_0^\infty dv \left[-g''\frac{\partial}{\partial g} + g'^2\frac{\partial^2}{\partial \alpha^2}\right]_{\alpha=0} \int_{-\infty}^{+\infty} dz_1 \cdots \int_{-\infty}^{+\infty} dz_4$$
$$\exp\left[-(b_0 + u)(z_1^2 + z_3^2)/2 - (b_0 + v)(z_2^2 + z_4^2)/2\right. \tag{7.13}$$
$$\left. - g(z_1 z_2 + z_3 z_4) - \alpha(z_1 z_4 - z_2 z_3) + iQ(z_1 + z_2)\right]$$

where the subscript $\alpha = 0$ indicates that α is to be set equal to zero after the differentiations are performed.

When the four-fold integral in the z's is evaluated (7.13) becomes

$$\Omega(\tau) = \int_0^\infty du \int_0^\infty dv \left[-g''\frac{\partial}{\partial g} + g'^2\frac{\partial^2}{\partial \alpha^2}\right]_{\alpha=0}$$
$$\frac{1}{4D}\exp\left[-Q^2(2b_0 - 2g + u + v)/(2D)\right] \tag{7.14}$$
$$= \int_0^\infty du \int_0^\infty dv[(g'^2 - gg'')(2 - 2F + Q^2/g) - g'^2 Q^2/g]e^{-F}/(4D_0^2)$$

where

$$D = (b_0 + u)(b_0 + v) - g^2 - \alpha^2, \qquad F = Q^2(2b_0 - 2g + u + v)/(2D_0)$$

and D_0 denotes the value of D obtained by setting $\alpha = 0$. When differentiating with respect to α it is helpful to note that

$$\frac{\partial^2 f(D)}{\partial \alpha^2} = f''(D)\left(\frac{\partial D}{\partial \alpha}\right)^2 + f'(D)\,\frac{\partial^2 D}{\partial \alpha^2}$$

and that only $f'(D) = df/dD$ need be obtained since $\partial D/\partial \alpha$ vanishes when $\alpha = 0$.

In order to reduce the double integral to a single integral we make the change of variables

$$\begin{gathered} r = Q^2(b_0 + u - g)/(2D_0) \equiv \frac{Q^2(b_0 + u - g)}{2[(b_0 + u)(b_0 + v) - g^2]} \\ s = Q^2(b_0 + v - g)/(2D_0), \qquad F = r + s \\ \partial(r, s)/\partial(u, v) = -rs/D_0, \qquad 4srD_0 = Q^2[Q^2 - 2g(r + s)] \end{gathered} \tag{7.15}$$

The limits of integration for r and s are obtained by noting that the points $(0, 0)$, $(\infty, 0)$, (∞, ∞), $(0, \infty)$ in the (u, v) plane go into $(Q^2/(2b_0 + 2g), Q^2/(2b_0 + 2g))$, $(Q^2/(2b_0), 0)$, $(0, 0)$ $(0, Q^2/(2b_0))$, respectively, in the (r, s) plane. It may be verified that the region of integration in the (r, s) plane is the interior of the quadrilateral obtained by joining the above points by straight lines. Equation (7.14) may now be written as

$$\begin{aligned} \Omega(\tau) &= \iint \left\{ \frac{(g'^2 - gg'')(2 - 2r - 2s + Q^2/g) - g'^2 Q^2/g}{Q^2[Q^2 - 2g(r + s)]} \right\} e^{-r-s}\, dr\, ds \\ &= \frac{g'^2 - gg''}{2g^2}\, y_1 - \frac{g'^2}{2g^2}\, y_2 \end{aligned} \tag{7.16}$$

where y_1 and y_2 are the dimensionless quantities

$$\begin{aligned} y_1 &= \iint \frac{2g^2(2 - 2r - 2s + Q^2/g)}{Q^2[Q^2 - 2g(r + s)]}\, e^{-r-s}\, dr\, ds \\ y_2 &= \iint \frac{2ge^{-r-s}\, dr\, ds}{Q^2 - 2g(r + s)} \end{aligned}$$

It is seen that

$$y_1 = 2gQ^{-2}\left\{ y_2 + \iint e^{-r-s}\, dr\, ds \right\}. \tag{7.17}$$

Since the integrands are functions of $r + s$ alone we are led to apply the transformation

$$\iint_A f(r + s)\, dr\, ds = \int_0^{\alpha} uf(u)\, du + \int_{\alpha}^{2\beta} \frac{\alpha(2\beta - u)}{2\beta - \alpha} f(u)\, du \tag{7.18}$$

where A is the area enclosed by the quadrilateral whose vertices are at the points (r, s) given by $(0, 0)$, $(0, \alpha)$, (β, β), $(\alpha, 0)$ and it is assumed that β and α are positive. u is a new variable and is not the one introduced in (7.13).

Setting $\alpha = Q^2/(2b_0)$ and $\beta = Q^2/(2b_0 + 2g)$, using (7.18), and introducing the notation

$$\begin{aligned} \rho &= Q^2/(2b_0), & k &= g/b_0 \\ \xi &= Q^2/(2g) = \rho/k, & \lambda &= \frac{Q^2}{b_0 + g} = \frac{2\rho}{1 + k} \end{aligned} \tag{7.19}$$

permits us to write

$$\begin{aligned} \iint_A e^{-r-s}\, dr\, ds &= \int_0^\rho u e^{-u}\, du + \int_\rho^\lambda \frac{\rho(\lambda - u)}{\lambda - \rho} e^{-u}\, du \\ &= 1 - \frac{\lambda e^{-\rho}}{\lambda - \rho} + \frac{\rho e^{-\lambda}}{\lambda - \rho} \end{aligned} \tag{7.20}$$

and (7.17) yields

$$y_2 = \frac{\rho}{k} y_1 - 1 + \frac{2}{1 - k} e^{-\rho} - \frac{1 + k}{1 - k} e^{-2\rho/(1+k)} \tag{7.21}$$

where we have expressed λ in terms of ρ and k.

The double integral defining y_2 may be treated in the same way as (7.20):

$$y_2 = \iint_A \frac{e^{-r-s}\, dr\, ds}{\xi - r - s} = \int_0^\rho \frac{u e^{-u}\, du}{\xi - u} + \int_\rho^\lambda \frac{\rho(\lambda - u) e^{-u}\, du}{(\lambda - \rho)(\xi - u)}$$

Writing $u = \xi - (\xi - u)$ and $\lambda - u = \lambda - \xi + (\xi - u)$ in the two numerators leads to

$$\begin{aligned} y_2 = \xi \int_0^\rho \frac{e^{-u}\, du}{\xi - u} - \int_0^\rho e^{-u}\, du \\ - \xi \int_\rho^\lambda \frac{e^{-u}\, du}{\xi - u} + \frac{\rho}{\lambda - \rho} \int_\rho^\lambda e^{-u}\, du \end{aligned} \tag{7.22}$$

where we have used $\rho(\lambda - \xi)/(\lambda - \rho) = -\xi$ to simplify the coefficient of the third integral. When the second and fourth integrals are evaluated, their contribution to y_2 is found to be equal to the terms independent of y_1 on the right of (7.21). Hence, comparison of equations (7.21) and (7.22) shows that

$$y_1 = \int_0^\rho \frac{e^{-u}\, du}{\xi - u} - \int_\rho^\lambda \frac{e^{-u}\, du}{\xi - u} \tag{7.23}$$

The integrals in (7.23) may be evaluated in terms of the exponential integral $Ei(x)$ defined by, for x real,

$$Ei(x) = \int_{-\infty}^{x} e^t\, dt/t = C + \tfrac{1}{2}\log_e x^2 + \sum_{n=1}^{\infty} \frac{x^n}{n!n}$$
$$\sim e^x \sum_{n=0} n!/x^{n+1}$$

where $C = .577\cdots$ is Euler's constant and Cauchy's principal value of the integral is to be taken when $x > 0$. We set $t = \xi - u$ and obtain

$$y_1 = e^{-\rho/k}\left\{Ei[\rho/k] - 2Ei[\rho(1-k)/k] + Ei\left[\frac{\rho(1-k)}{k(1+k)}\right]\right\}$$

where we have again expressed ξ and λ in terms of ρ and k.

A power series for y_1 which converges when $-1/3 \leq k < 1$ may be obtained by expanding the denominators of the integrands in (7.23) in powers of u/ξ and integrating termwise:

$$\begin{aligned} y_1 = {} & \xi^{-1}[1 - 2e^{-\rho} + e^{-\lambda}] \\ & + 1!\xi^{-2}[1 - 2(1 + \rho/1!)e^{-\rho} + (1 + \lambda/1!)e^{-\lambda}] \\ & + 2!\xi^{-3}[1 - 2(1 + \rho/1! + \rho^2/2!)e^{-\rho} + (1 + \lambda/1! + \lambda^2/2!)e^{-\lambda}] \\ & + \cdots \end{aligned} \tag{7.24}$$

The following special values may be obtained from the equation given above. When $\rho = 0$

$$\begin{aligned} y_1 &= -\log_e(1 - k^2) \\ y_2 &= 0 \end{aligned} \tag{7.25}$$

This result may also be obtained by evaluating the integral obtained when we set $Q = 0$, $z_1 = r_1 \cos\theta_1$, $z_3 = r_1 \sin\theta_1$, $z_2 = r_2 \cos\theta_2$, $z_4 = r_2 \sin\theta_2$ in (7.11) and (7.12).

Near $k = 1$,

$$\begin{aligned} y_1 &\approx e^{-\rho}[Ei(\rho) - C - \log_e \rho(1 - k^2)] \\ y_2 &\approx \rho y_1 - 1 + (1 + \rho)e^{-\rho} \end{aligned} \tag{7.26}$$

Near $k = 0$,

$$y_1 \approx k(1 - e^{-\rho})^2/\rho, \qquad y_2 \approx y_1 \tag{7.27}$$

except when $\rho = 0$ in which case y_1 is approximately k^2.

When ρ is large

$$\begin{aligned} y_1 &\sim \frac{k}{\rho} + \frac{1!k^2}{\rho^2} + \frac{2!k^3}{\rho^3} + \frac{3!k^4}{\rho^4} + \cdots \\ y_2 &\sim -1 + \frac{\rho}{k}y_1 \sim \frac{1!k}{\rho} + \frac{2!k^2}{\rho^2} + \cdots \end{aligned} \tag{7.28}$$

except near $k = 1$ where both y_1 and y_2 have logarithmic infinities. The asymptotic expansion (7.3) for $\Omega(\tau)$, which was obtained by the first method

of this section, may be checked by inserting (7.28) in the expression (7.16) for $\Omega(\tau)$ in terms of y_1 and y_2.

Values of y_1 and y_2 tabulated as functions of k for various values of ρ are given in Table 3. Negative values of k have not been considered since they

TABLE 3

VALUES OF y_1 AND y_2 USED IN COMPUTATION OF CORRELATION FUNCTION OF $d\theta/dt$

$$\Omega(\tau) = \overline{\theta'(t)\theta'(t+\tau)} = [g'^2(y_1 - y_2) - gg''y_1]/(2g^2)$$

$$g \equiv g(\tau) = \int_0^\infty w(f) \cos 2\pi(f - f_0)\tau \, df, \quad k = g(\tau)/g(0)$$

k	Values of y_1					Values of y_2			
	ρ					ρ			
	0	.5	1	2	5	.5	1	2	5
0	0	0	0	0	0	0	0	0	0
.1	.01005	.03526	.04224	.03854	.02000	.03171	.04147	.03936	.02051
.2	.04082	.08043	.09003	.07979	.04105	.06550	.08654	.08275	.04283
.3	.09431	.1379	.1452	.1246	.06292	.1022	.1363	.1315	.06702
.4	.1744	.2110	.2102	.1740	.08586	.1432	.1926	.1870	.09384
.5	.2877	.3056	.2886	.2296	.1101	.1914	.2579	.2515	.1238
.6	.4463	.4278	.3860	.2942	.1358	.2481	.3368	.3289	.1576
.7	.6733	.5953	.5129	.3721	.1636	.3220	.4379	.4269	.1975
.8	1.0216	.8416	.6914	.4729	.1941	.4275	.5803	.5602	.2461
.84	1.2228	.9798	.7888	.5242	.2075	.4866	.6593	.6318	.2693
.88	1.4890	1.1590	.9127	.5866	.2219	.5641	.7619	.7226	.2964
.90	1.6607	1.2742	.9898	.6241	.2296	.6138	.8260	.7752	.3114
.92	1.8734	1.4144	1.0834	.6686	.2378	.6753	.9058	.8486	.3294
.94	2.1507	1.5948	1.2024	.7217	.2466	.7550	1.0093	.9333	.3498
.96	2.5459	1.8486	1.3668	.7939	.2566	.8711	1.1558	1.0546	.3752
.97	2.8285	2.0251	1.4815	.8414	.2623	.9474	1.2605	1.1366	.3849
.98	3.2289	2.2762	1.6405	.9073	.2690	1.0704	1.4081	1.2548	.4119
.99	3.9170	2.7080	1.9066	1.0127	.2778	1.2773	1.6610	1.4505	.4429
.995	4.6072	3.1341	2.1721	1.1125	.2846	1.4838	1.9175	1.6416	.4705
.997	5.1175	3.4445	2.3622	1.1866	.2889	1.6367	2.1048	1.7859	.4893

are not required for the case in which I_N has a normal law power spectrum, the case discussed in the next section.

8. POWER SPECTRUM OF $\frac{d\theta}{dt}$ WHEN I_N HAS NORMAL LAW POWER SPECTRUM

The problem of computing the power spectrum $W(f)$ of $\theta'(t)$ appears to be a difficult one.* In order to obtain an answer without an excessive amount of work we have had to do two things which are rather restrictive. First, we confine our attention to the case in which the power spectrum $w(f)$ of

*Since the above was written the general f. m. problem has been studied by D. Middleton. He generalizes our (7.11) and (7.12), introduces polar coordinates, expands the integrand in powers of g, and integrates termwise. $W(f)$ then follows somewhat as in a.m. theory.

I_N is of the normal law type (our method could be applied to other types but g' and g'' would be more complicated functions of τ and Table 3 would have to be extended to negative values of k, if they should occur). Second, we resort to numerical integration to obtain a portion of $W(f)$. Because of the second item our results are either tabulated or are given as curves, shown in Figs. 8 and 9, except when $Q = 0$ (noise only) in which case the power spectrum of θ' is given by the series (8.7).

The power spectrum of I_N is assumed to be

$$w(f) = \frac{\psi_0}{\sigma\sqrt{2\pi}} e^{-(f-f_q)^2/(2\sigma^2)} \tag{8.1}$$

The mean square value of I_N is equal to that of a noise current whose power spectrum has the constant value of $\psi_0/(\sigma\sqrt{2\pi})$ over a band of width $f_b - f_a = \sigma\sqrt{2\pi} = \sigma 2.507$. The value of $w(f)$ is one quarter of its mid-band value at the points $f - f_q = \pm\sigma\sqrt{2\log_e 4} = \pm\sigma 1.665$ (the 6 *db* points) and the distance between these points is 3.330σ. Integration of (8.1) shows that the mean square value of I_N is ψ_0 in accordance with our customary notation. The mid-band value of $w(f)$ is $\psi_0/(\sigma\sqrt{2\pi})$.

Assuming $f_q \gg \sigma$ and evaluating the integrals (A2–1) of Appendix II defining b_0 and g gives

$$\begin{gathered} b_0 = \psi_0, \qquad g = \psi_0 e^{-2(\pi\sigma\tau)^2} = \psi_0 e^{-u^2/2} \\ g'/g = -uu' = -2\pi\sigma u, \qquad g''/g = -(2\pi\sigma)^2(1 - u^2) \\ \frac{g''g - g'^2}{g^2} = -(2\pi\sigma)^2, \qquad k = g/b_0 = e^{-u^2/2} \end{gathered} \tag{8.2}$$

where we have set

$$u = 2\pi\sigma\tau, \qquad u' = 2\pi\sigma \tag{8.3}$$

and the primes on g and u denote differentiation with respect to τ. The correlation function is accordingly, from (7.16).

$$\Omega(\tau) = 2\pi^2\sigma^2(y_1 - u^2 y_2) \tag{8.4}$$

If $\theta'(t)$ be regarded as a noise current its power spectrum is

$$W(f) = 4\int_0^\infty \Omega(\tau)\cos 2\pi f\tau \, d\tau \tag{8.5}$$

When noise alone is present, ρ is zero and (7.25) yields

$$\Omega(\tau) = -2\pi^2\sigma^2 \log_e (1 - k^2) = -2\pi^2\sigma^2 \log_e (1 - e^{-u^2}) \tag{8.6}$$

In this case the power spectrum is, from (8.3), (8.5), and (8.6),

$$\begin{aligned} W_N(f) &= -4\pi\sigma \int_0^\infty \cos(uf/\sigma) \log (1 - e^{-u^2})\, du \\ &= 2\sigma\pi^{3/2} \sum_{n=1}^{\infty} n^{-3/2} e^{-f^2/(4n\sigma^2)}, \end{aligned} \tag{8.7}$$

the series being obtained by expanding the logarithm and integrating termwise. When this equation was used for computation it was found convenient to apply the Euler summation formula to sum the terms in the series beyond the (N − 1)st. Writing b for $f^2/(4\sigma^2)$, the series in (8.7) becomes

$$\begin{aligned} &1^{-3/2}e^{-b/1} + 2^{-3/2}e^{-b/2} + \cdots (N-1)^{-3/2}e^{-b/(N-1)} \\ &+ (\pi/b)^{1/2} \operatorname{erf} [(b/N)^{1/2}] + N^{-3/2}e^{-b/N}\left[\frac{1}{2} - \frac{1}{12N}\left(-\frac{3}{2} + \frac{b}{N}\right)\right. \\ &\left. + \frac{1}{720N^3}\left(-\frac{105}{8} + \frac{105}{4}\frac{b}{N} - \frac{21}{2}\frac{b^2}{N^2} + \frac{b^3}{N^3}\right) + \cdots\right] \end{aligned} \tag{8.8}$$

When b is zero the sum[15] of the series is $2.61237 \cdots$. The values for $\rho = 0$ in Table 4 were computed by taking $N = 12$ in (8.8). As $b \to \infty$ the dominant term in (8.8) is seen to be the one containing erf (choose N so that $b = N^{3/2}$). Hence as $f \to \infty$

$$W_N(f) \sim 4\pi^2\sigma^2/f. \tag{8.9}$$

When both noise and the sine wave are present it is convenient to split the power spectrum into three parts. The first part, $W_1(f)$, is proportional to $W_N(f)$, the power spectrum with noise alone. The second part $W_2(f)$ is proportional to the form $W(f)$ assumes when rms $I_N \ll Q$ and the third part $W_3(f)$ is of the nature of a correction term. This procedure is suggested when we subtract the leading terms in the expressions (7.26) and (7.27) (corresponding to $k = 1$ and $k = 0$, respectively) from y_1. Likewise we subtract the leading term in y_2, (7.27), at $k = 0$ but do not bother to do so at the end $k = 1$ because u^2y_2 approaches zero there. We therefore write

$$\begin{aligned} y_1 - u^2y_2 &= [y_1 + e^{-\rho}\log (1 - k^2) - k(1 - e^{-\rho})^2/\rho - u^2y_2 \\ &+ u^2k(1 - e^{-\rho})^2/\rho] - e^{-\rho} \log (1 - k^2) + (1 - u^2)k(1 - e^{-\rho})^2/\rho \\ &= Z(u) - e^{-\rho} \log (1 - k^2) - \frac{g''(2\pi\sigma)^{-2}}{b_0\rho} (1 - e^{-\rho})^2 \end{aligned} \tag{8.10}$$

[15] "Theory and Application of Infinite Series," Knopp, (1928), page 561.

where $Z(u)$ denotes the function enclosed by the brackets in the first equation and the expressions for g''/g and k in (8.2) have been used in the replacement of $(1 - u^2)k$.

TABLE 4

VALUES OF $W_3(f)/(4\pi^2\sigma)$

$\frac{6f}{\sigma\pi}$	$\rho = 0$	0.5	1.0	2.0	5.0
0	0	−.03517	−.03891	−.02444	−.001948
1	0	−.03003	−.03196	−.01830	−.001814
2	0	−.01717	−.01486	−.003304	.004052
3	0	−.002436	.004014	.01252	.008225
4	0	.008757	.01730	.02244	.01027
6	0	.01478	.02157	.02167	.007665
8	0	.01018	.01366	.01237	.003505
10	0	.005768	.007378	.006201	.001437
12	0	.004027	.004463	.003552	.0006439

VALUES OF $W(f)/(4\pi^2\sigma)$

0	.7369	.4118	.2322	.07529	.003017
1	.7098	.4294	.2672	.1134	.02342
2	.6439	.4516	.3231	.1784	.05828
3	.5542	.4225	.3225	.1947	.06852
4	.4623	.3496	.2654	.1580	.01590
6	.3195	.2178	.1508	.07554	.01540
8	.2390	.1553	.1019	.04506	.005325
10	.1908	.1215	.07768	.03206	.002726
12	.1595	.1003	.06306	.02511	.001719

Inserting (8.10) in the expression (8.4) for $\Omega(\tau)$ and taking the Fourier transform (8.5) leads to

$$\begin{aligned} W(f) &= W_1(f) + W_2(f) + W_3(f) \\ W_1(f) &= e^{-\rho} W_N(f) \\ W_2(f) &= -\frac{(1 - e^{-\rho})^2}{2b_0\rho} \int_0^\infty g'' \cos 2\pi f\tau \, d\tau \\ &= \frac{(1 - e^{-\rho})^2}{\rho} (2\pi f)^2 \frac{e^{-f^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}} \\ W_3(f) &= 4\pi\sigma \int_0^\infty Z(u) \cos (uf/\sigma) \, du \end{aligned} \tag{8.11}$$

In these equations $W_N(f)$ is obtained from (8.7), and $W_2(f)$ by two-fold integration by parts to reduce g'' to g then evaluating the integral obtained

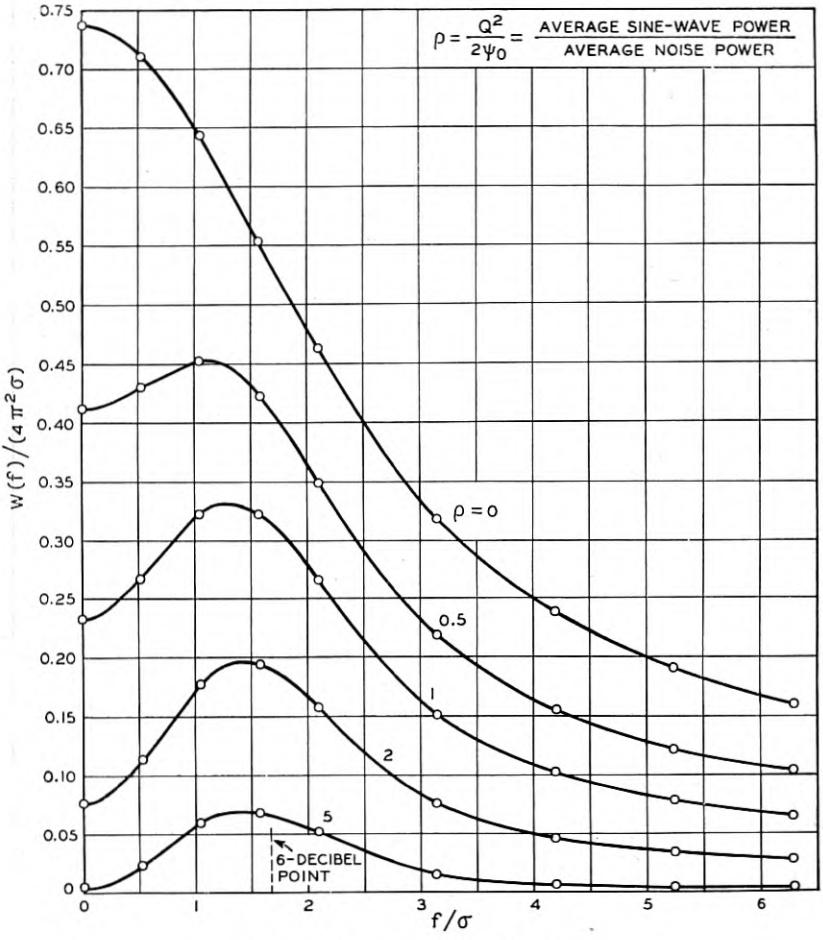


Fig. 8—Power spectrum of $d\theta/dt$.

Power spectrum of I_N is assumed to be

$$\psi_0(\sigma\sqrt{2\pi})^{-1} \exp\,[-(f - f_q)^2/(2\sigma^2)].$$

In this expression f is a frequency near f_q. The f in $W(f)$ and in the abscissa is a much lower frequency. $W(f)$ = power spectrum of $\theta' = d\theta/dt$, θ' being regarded as a random noise current. Dimensions of $W(f)\,df$ same as $(d\theta/dt)^2$ or (radians)^2/sec.^{2}.

by substituting the expression (8.2) for g. That $W(f)$ approaches $W_2(f)$ as $\rho \longrightarrow \infty$ follows when expression (8.11) for $W_2(f)$ is compared with the limiting form (8.13) given below.

Instead of dealing with $W(f)$ it is more convenient to deal with $(4\pi^2\sigma)^{-1}$ $W(f)$ which is the sum of the three components

$$(4\pi^2\sigma)^{-1}W_1(f) = \frac{e^{-\rho}}{2\sqrt{\pi}} \sum_{n=1}^{\infty} n^{-3/2} e^{-f^2/(4n\sigma^2)}$$

$$(4\pi^2\sigma)^{-1}W_2(f) = \frac{(1 - e^{-\rho})^2}{\rho\sqrt{2\pi}} \left(\frac{f}{\sigma}\right)^2 e^{-f^2/(2\sigma^2)} \tag{8.12}$$

$$(4\pi^2\sigma)^{-1}W_3(f) = \frac{1}{\pi} \int_0^{\infty} Z(u) \cos(uf/\sigma)\, du$$

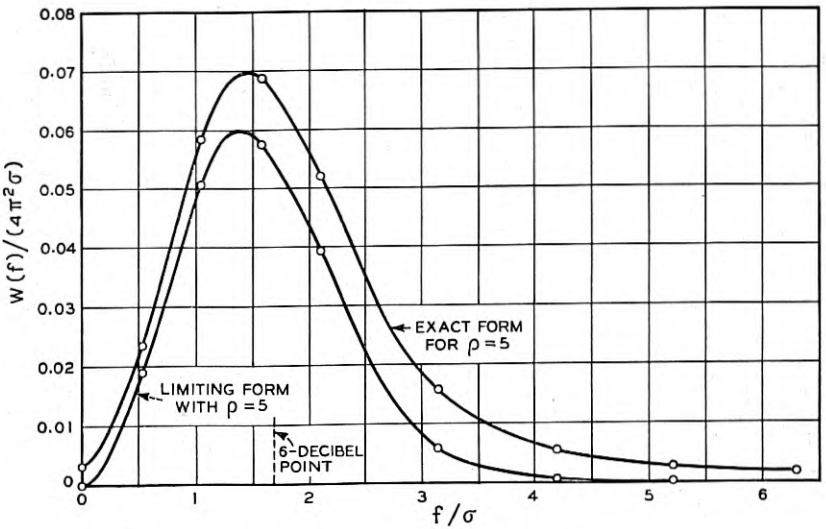


Fig. 9—Approach of $W(f)$ to limiting form.

As $\rho \to \infty$, $W(f) \to 4\pi^2\sigma\ (\rho\sqrt{2\pi})^{-1}\ (f/\sigma)^2 \exp[-f^2/(2\sigma^2)]$.

The integral involving $Z(u)$ has been computed by Simpson's rule, y_1 and y_2 being obtained from Table 3, with the results shown in the first section of Table 4. The value of $W_2(f)$ may be computed directly, and $W_1(f)$ may be obtained from $W_N(f)$. The values of these two functions together with those of W_3 (f) enable us to compute the values of $(4\pi^2\sigma)^{-1}W(f)$ given in Table 4 and plotted in Fig. 8.

Since, as is shown by (8.9), $W_N(f)$ varies as $1/f$ for large values of f, the areas under the curves of Fig. 8 become infinite. This agrees with the fact that the mean square value of θ' is infinite.

The values of $(4\pi^2\sigma)^{-1}W(0)$ for ρ equal to 0, .5, 1, 2, and 5 are .7369, 4118, 2322, .07529, and .003017 respectively. When these values are plotted on

semi-log paper they tend to lie on a straight line whose slope suggests that $W(0)$ decreases as $e^{-\rho}$ when ρ becomes large.

The limiting form assumed by $W(f)$ as $\rho \to \infty$ is given by equation (7.2). When the normal law expression (8.1) assumed in this section for the power spectrum of I_N is put in (7.2) we find that

$$W(f) \to \frac{4\pi^2\sigma}{\rho\sqrt{2\pi}} \left(\frac{f}{\sigma}\right)^2 e^{-f^2/(2\sigma^2)} \tag{8.13}$$

Fig. 9 shows that for $\rho = 5$ the limiting form (8.13) agrees quite well with the exact form computed above.

Both (7.2) and (8.13) show that, for small values of f, the power spectrum of θ' varies as f^2 when $\rho >> 1$. This is in accord with Crosby's* result that the voltage spectrum of the random noise in the output of a frequency modulation receiver is triangular when the carrier to noise ratio is large. When this ratio becomes small he finds that the spectrum becomes rectangular. Fig. 8 shows this effect in that the areas under the curves between the ordinates at $f = 0$ and $f = \lambda\sigma$ (where λ is some number, generally less than unity, depending on the ratio of the widths of the i.f. and audio bands) become rectangles, approximately, as ρ decreases.

APPENDIX I

The Integral $Ie\ (k, x)$

The integral[16]

$$Ie(k,x) = \int_0^x e^{-u} I_0(ku)\, du, \tag{A1-1}$$

where $I_0(ku)$ denotes the Bessel function of imaginary argument and order zero, occurs in Sections 2 and 6. The following special cases are of interest.

$$\begin{aligned} Ie(0,\ x) &= 1 - e^{-x} \\ Ie(1,\ x) &= xe^{-x}[I_0(x) + I_1(x)] \\ Ie(k,\ \infty) &= \frac{1}{\sqrt{1-k^2}} \end{aligned} \tag{A1-2}$$

The second of these relations is due to Bennett.[17]

* M. G. Crosby, "Frequency Modulation Noise Characteristics," Proc. I. R. E. Vol. 25 (1937), 472–514. See also J. R. Carson and T. C. Fry, "Variable Electric Circuit Theory with Application to the Theory of Frequency Modulation," B.S.T.J. Vol. 16 (1937), 513–540.

[16] The notation was chosen to agree with that used by Bateman and Archibald (Guide to Tables of Bessel Functions appearing in "Math. Tables and Aids to Comp.", Vol. 1 (1944) pp. 205–308) to discuss integrals used by Schwarz (page 248).

[17] It is given in equation (62) of the reference cited in connection with our equation (1.2) in Section 1.

The values in the table given below were computed by Simpson's rule for numerical integration. The work was checked at several points by using

$$Ie(k, x) = \sum_{n=0}^{\infty} (k/2)^{2n} \frac{(2n)!}{n!n!} A_n$$

where

$$A_n = 1 - \left[1 + x + \frac{x^2}{2!} \cdots + \frac{x^{2n}}{(2n)!}\right] e^{-x}$$

When x is so large that $Ie(k, x)$ is nearly equal to $Ie(k, \infty)$ we have

$$Ie(k, x) \sim (1 - k^2)^{-1/2} - [2k(1 - k)]^{-1/2}(2/\sqrt{\pi}) \int_{t_1}^{\infty} e^{-t^2}\, dt$$

where $t_1 = \sqrt{x(1 - k)}$. However, this was not found to be especially useful in checking the values given in the table.

TABLE OF $Ie(k, x) = \int_0^x e^{-u} I_0(ku)\, du$

x	k						
	0	.2	.4	.6	.8	.9	1.0
0	0	0	0	0	0		0
.2	.1813	.1813	.1814	.1815	.1816		.1818
.4	.3297	.3298	.3303	.3311	.3322		.3337
.6	.4512	.4517	.4530	.4554	.4586		.4629
.8	.5507	.5516	.5545	.5593	.5661		.5749
1.0	.6321	.6337	.6386	.6468	.6584		.6736
.2	.6988	.7012	.7086	.7209	.7386		.7620
.4	.7534	.7567	.7669	.7841	.8089		.8422
.6	.7981	.8025	.8157	.8383	.8712		.9157
.8	.8347	.8401	.8566	.8850	.9267		.9839
2.0	.8647	.8712	.8910	.9255	.9766		1.0476
.2	.8892	.8968	.9201	.9607	1.0217		1.1075
.4	.9093	.9179	.9446	.9916	1.0627		1.1642
.6	.9257	.9354	.9655	1.0186	1.1001		1.2183
.8	.9392	.9499	.9831	1.0424	1.1345		1.2699
3.0	.9502	.9618	.9982	1.0635	1.1661		1.3195
.2	.9592	.9718	1.0110	1.0822	1.1953		1.3672
.4	.9666	.9800	1.0220	1.0988	1.2223		1.4132
.6	.9727	.9868	1.0314	1.1136	1.2475		1.4578
.8	.9776	.9925	1.0394	1.1268	1.2708		1.5010
4.0	.9817	.9971	1.0463	1.1386	1.2926		1.5430
.2	.9830	1.0010	1.0522	1.1492	1.3130		1.5839
.4	.9877	1.0043	1.0574	1.1587	1.3320		1.6237
.6	.9899	1.0070	1.0619	1.1672	1.3499		1.6625
.8	.9918	1.0092	1.0657	1.1749	1.3666		1.7005
5.0	.9933	1.0111	1.0690	1.1818	1.3823		1.7376
5.4	.9955	1.0140	1.0743	1.1937	1.4110		1.8095

TABLE—*Continued*

x	k						
	0	.2	.4	.6	.8	.9	1.0
5.8	.9970	1.0160	1.0783	1.2034	1.4364		1.8786
6.2	.9980	1.0174	1.0814	1.2114	1.4590		1.9452
6.6	.9986	1.0183	1.0837	1.2180	1.4792		2.0097
7.0	.9991	1.0190	1.0854	1.2234	1.4972		2.0722
7.4	.9994	1.0195	1.0867	1.2278	1.5134		2.1328
7.8	.9996	1.0198	1.0876	1.2375	1.5279		2.1917
8.2	.9997	1.0201	1.0885	1.2346	1.5409		2.2491
8.6	.9998	1.0202	1.0891	1.2371	1.5526		2.3050
9.0	.9999	1.0203	1.0896	1.2393	1.5631		2.3597
10.0	1.0000	1.0205	1.0902	1.2431	1.5852	1.9207	2.4910
11.0	1.0000	1.0206	1.0907	1.2456	1.6024	1.9668	2.6157
12.0	1.0000	1.0206	1.0909	1.2471	1.6158	2.0066	2.7347
13.0	1.0000	1.0206	1.0910	1.2482	1.6263	2.0411	2.8487
14.0	1.0000	1.0206	1.0910	1.2488	1.6346	2.0711	2.9584
15.0	1.0000	1.0206	1.0911	1.2492	1.6412	2.0973	3.0641
∞	1.0000	1.0206	1.0911	1.2500	1.6667	2.2942	∞

x	k			
	.86	.90	.96	1.0
15.0	1.8773	2.0973	2.5810	3.0641
16.0	1.8899	2.1201	2.6371	3.1663
17.0	1.9006	2.1403	2.6894	3.2653
18.0	1.9095	2.1579	2.7381	3.3614
19.0	1.9171	2.1737	2.7837	3.4548
20.0	1.9235	2.1870	2.8263	3.5457
∞	1.9597	2.2942	3.5714	∞

APPENDIX II

SECOND MOMENTS ASSOCIATED WITH I_c AND I_s

The in-phase and quadrature components of the noise current I_N

$$I_c(t) = \sum_{n=1}^{M} c_n \cos [(\omega_n - q)t - \varphi_n]$$
$$I_s(t) = \sum_{n=1}^{M} c_n \sin [(\omega_n - q)t - \varphi_n] \tag{3.3}$$

are closely related to the envelope R and phase angle θ of the total current, this relationship being being shown by the equations (3.4) and (3.5). $I_c(t)$ and $I_s(t)$ and their time derivatives may be regarded as random variables. In much of our work we have to deal with the probability distribution of these random variables. By virtue of the representation (3.3) and the central limit theorem[18] this distribution is normal in the several variables. The coefficients in the quadratic form occurring in the exponent are deter-

[18] Section 2.10 of Reference A.

mined by the second moments of the variables.[19] Here we state these moments. Some of the moments have already been given in Sections 3.7 and 3.8 of Reference A. For the sake of completeness we shall also give them here. The new results given below are derived in much the same way as those given in Reference A.

Let

$$\begin{aligned} b_n &= (2\pi)^n \int_0^\infty w(f)(f - f_q)^n \, df \\ b_0 &= \int_0^\infty w(f) \, df = \psi_0 \\ g &= \int_0^\infty w(f) \cos 2\pi(f - f_q)\tau \, df \\ h &= \int_0^\infty w(f) \sin 2\pi(f - f_q)\tau \, df \end{aligned} \tag{A2-1}$$

and let g', g'', h', h'' denote the first and second derivatives of g and h with respect to τ. For example,

$$g' = -2\pi \int_0^\infty w(f)(f - f_q) \sin 2\pi(f - f_q)\tau \, df$$

Incidentally, in many of our cases $w(f)$ is assumed to be symmetrical about f_q. This introduces considerable simplification because b_1, b_3, b_5, $\cdots$, h, h', h'', reduce to zero.

The following table gives values of b_n's and g for two cases of frequent occurrence

	Ideal band pass filter centered on f_q	Normal law filter centered on f_q, $f_q \gg \sigma$
$w(f)$	w_0 for $f_a < f < f_b$ and zero elsewhere	$\frac{\psi_0}{\sigma\sqrt{2\pi}} e^{-(f-f_q)^2/2\sigma^2}$
b_0	$w_0(f_b - f_a)$	ψ_0
b_2	$\pi^2 w_0(f_b - f_a)^3/3$	$4\pi^2\sigma^2\psi_0$
b_4	$\pi^4 w_0(f_b - f_a)^5/5$	$48\pi^4\sigma^4\psi_0$
g	$(\pi\tau)^{-1} w_0 \sin \pi(f_b - f_a)\tau$	$\psi_0 e^{-2(\pi\sigma\tau)^2}$

If we write I_c, I_c', I_c'' for $I_c(t)$, $I_c'(t)$, $I_c''(t)$, where the primes denote differ-

[19] Section 2.9 of Reference A.

entiation with respect to t, and do the same for $I_s(t)$ and its derivatives we have, from Section 3.8 of Reference A,

$$\begin{gathered}
\overline{I_c^2} = \overline{I_s^2} = b_0, \qquad \overline{I_c I_s} = 0 \\
\overline{I_c I_s'} = -\overline{I_c' I_s} = b_1, \qquad \overline{I_c I_c'} = \overline{I_s I_s'} = 0 \\
\overline{I_c'^2} = \overline{I_s'^2} = -\overline{I_c I_c''} = -\overline{I_s I_s''} = b_2, \qquad \overline{I_c' I_s'} = \overline{I_c I_s''} = \overline{I_s I_c''} = 0 \\
\overline{I_c' I_s''} = -\overline{I_s' I_c''} = b_3, \qquad \overline{I_c' I_c''} = \overline{I_s' I_s''} = 0 \\
\overline{I_c''^2} = \overline{I_s''^2} = b_4, \qquad \overline{I_c'' I_s''} = 0
\end{gathered} \tag{A2-2}$$

When we deal with moments in which the arguments of the two variables are separated by an interval τ as in (see the last of equations (3.7–11) of Reference A)

$$\overline{I_c(t) I_s(t+\tau)} = h,$$

it is convenient to denote the argument t by the subscript 1 and the argument $t + \tau$ by 2. Then our example becomes

$$\overline{I_{c1} I_{s2}} = h$$

We shall need the following moments of this type.

$$\begin{gathered}
\overline{I_{c1} I_{c2}} = \overline{I_{s1} I_{s2}} = g, \qquad \overline{I_{c1} I_{s2}} = -\overline{I_{c2} I_{s1}} = h \\
\overline{I_{c1} I_{c2}'} = \overline{I_{s1} I_{s2}'} = -\overline{I_{c1}' I_{c2}} = -\overline{I_{s1}' I_{s2}} = g' \\
\overline{I_{c1} I_s'} = \overline{I_{c2} I_{s1}'} = -\overline{I_{c1}' I_{s2}} = -\overline{I_{c2}' I_{s1}} = h' \\
\overline{I_{c1}' I_{c2}'} = \overline{I_{s1}' I_{s2}'} = -g'', \qquad \overline{I_{c1}' I_{s2}'} = -\overline{I_{c2}' I_{s1}'} = -h''
\end{gathered} \tag{A2-3}$$

It should be remembered that in these equations the primes on the I's denote differentiation with respect to t while the primes on g and h denote differentiation with respect to τ.

APPENDIX III

Evaluation of a Multiple Integral

Several multiple integrals encountered during the preparation of this paper were initially evaluated by the following procedure. The integral was first converted into a multiple series by expanding a portion of the integrand and integrating termwise. It was found possible to sum these series when one of the factorials in the denominator was represented as a contour integral. This reduced the multiple integral to a contour integral and sometimes the latter could be evaluated.

We shall illustrate this procedure by examining the integral

$$I = \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dx\, x \exp\left[-x^2 + 2a\cos\theta + 2bx\sin\theta + c\sin^2\theta\right] \tag{A3-1}$$

Expanding that part of the exponential which contains the trigonometrical terms and integrating termwise gives

$$I = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{\ell=0}^{\infty} \frac{a^{2n} b^{2m} c^{\ell} \pi \Gamma(\ell + m + \tfrac{1}{2})}{n!\,\ell!\,(\ell + m + n)!\,\Gamma(m + \tfrac{1}{2})}$$

where we have used

$$2^{2n}\,\Gamma(n + \tfrac{1}{2})n! = \sqrt{\pi}(2n)!$$

We next make the substitution

$$\frac{1}{(\ell + m + n)!} = \frac{1}{2\pi i}\int_C \frac{e^t\, dt}{t^{\ell+m+n+1}} \tag{A3-2}$$

where the path of integration C is a circle chosen large enough to ensure the convergence of the series obtained when the order of summation and integration is changed. The summations may now be performed:

$$\begin{aligned} I &= \frac{1}{2i}\int_C dt\, e^{t+a^2/t} \sum_{m=0}^{\infty} b^{2m} t^{-m-1}(1 - ct^{-1})^{-m-1/2} \\ &= \frac{1}{2i}\int_C \frac{t^{-1/2}(t - c)^{1/2}}{t - c - b^2}\, e^{t+a^2/t}\, dt \end{aligned} \tag{A3-3}$$

C encloses the pole at $c + b^2$ and the branch point at c as well as the origin.

When a^2 is zero the integral may be reduced still further. Let c be complex and b such that the point $c + b^2$ does not lie on the line joining 0 to c. Deform C until it consists of an isolated loop about $c + b^2$ and a loop about 0 and c, the latter consisting of small circles about 0 and c joined by two straight portions running along the line joining 0 to c. The contributions of the small circles about 0 and c vanish in the limit. Along the portion starting at 0 and running to c, $\arg(t - c) = -\pi + \arg c$, and along the portion starting at c and running to 0, $\arg(t - c) = \pi + \arg c$. On both portions $\arg t = \arg c$. Bearing this in mind and setting $t = c\sin^2\theta$ on the two portions gives

$$I_{a=0} = \pi b(c + b^2)^{-1/2} e^{c+b^2} + 2c\int_0^{\pi/2} \frac{\cos^2\theta e^{c\sin^2\theta}}{b^2 + c\cos^2\theta}\, d\theta \tag{A3-4}$$

The integral may be expressed in terms of the function

$$Ie(k, x) = \int_0^x e^{-u} I_0(ku)\, du$$

by noting that

$$\int_0^\pi \frac{e^{-\alpha-\beta\cos v}}{\alpha+\beta\cos v}\,dv = \int_0^\pi dv\left[\frac{1}{\alpha+\beta\cos v} - \int_0^1 e^{-t(\alpha+\beta\cos v)}\,dt\right]$$

$$= \pi(\alpha^2-\beta^2)^{-1/2} - \pi\int_0^1 e^{-\alpha t} I_0(\beta t)\,dt \qquad \text{(A3-5)}$$

$$= \pi(\alpha^2-\beta^2)^{-1/2} - (\pi/\alpha)Ie(\beta/\alpha,\alpha)$$

Thus

$$I_{a=0} = \pi e^{c/2} I_0(c/2) + (\pi b^2/\alpha)e^{b^2+c} Ie\left(\frac{c}{2\alpha},\alpha\right) \qquad \text{(A3-6)}$$

where

$$\alpha = b^2 + c/2 \qquad \text{(A3-7)}$$

Noise in Resistances and Electron Streams

By J. R. PIERCE

TECHNICALLY correct results in a field are achieved initially in diverse and often confusing and complicated ways. Sometimes, such results are later brought together to give them a more unified form and a sounder basis; such critical summary and exposition is of great value. In quite another way, a worker who uses results established in a field will discover many plausible reasons for believing the results, and he will find eventually that an air of inevitability and "understanding" pervades the subject. Such "understanding" is not to be confused with the process of rigorous proof carried out step by step, but it can help in organizing and making use of a body of related material.

The field of "noise", especially as it affects electron devices and communications in general, is one particularly troublesome to engineers. The sound work on the subject has commonly involved mathematics and especially statistical ideas unfamiliar to many who must deal with the practical problems of noise. In early papers on noise, a great deal of heat was generated in acrimonious controversy between two schools, one of which assigned a uniform noise spectrum to certain noise sources, while the other held this to be inadmissible and got identical answers by more recondite means. Happily, a recent paper by S. O. Rice[1] clearly presents both approaches. Rice's paper further provides a fine broad summary of noise problems together with considerable original material. It does not extend far into the field of electronics.[2]

The reader who has sufficient time could achieve a profound "understanding" of the circuit aspects of noise by reading Rice's paper. The understanding would involve familiarity with much mathematics useful in itself. To many engineers, however, this might prove a lengthy and painful process.

The writer proposes to present here a series of plausible arguments for believing certain facts about noise. Both simple circuit considerations and "electronic" effects (as, space charge reduction of noise) are included. The arguments presented are not intended to be original and it is not claimed that they are rigorous; they do seem to be easily understood, and to help in remembering and in using some important practical material. Starting points of the arguments, or "postulates", have been chosen on the basis of familiarity, not simplicity. No effort is made to point out all of the hidden assumptions in the arguments, but a few important ones are indicated.

An initial warning should be made that quantum effects treated in Nyquist's original paper on Johnson noise, but afterwards much neglected, are entirely disregarded here.

I. Johnson Noise[3]

In 1926, in an investigation of amplifiers with exceedingly high grid resistances, J. B. Johnson discovered that a resistance acts as a noise generator having an open-circuit voltage with a mean square value

$$\overline{v^2} = 4kTRB. \tag{1}$$

Here and subsequently, lower case letters v and i will be used in referring to noise voltages and currents. In (1), $\overline{v^2}$ is the mean square value of noise voltage components of frequency lying in a small bandwidth B (sometimes

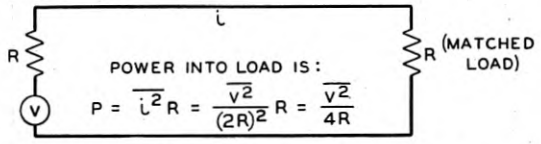


Fig. 1—Relations between noise power, noise voltage and noise current can be derived by assuming the noise source to be a voltage in series with a resistance.

called df or Δf), k is Boltzman's constant, and R is resistance. We easily see from Fig. 1 that the maximum noise power which can be made to flow from the resistance into a load (that which will flow into a matched load) is

$$P = \frac{\overline{v^2}}{4R} = kTB. \tag{2}$$

This "available noise power" is a convenient alternative formulation.

If an impedance has a reactive as well as a resistive component, the open circuit noise is given by (1) where R is the resistive component; if an admittance has a conductance G the noise may be represented as an impressed current (that which flows when the admittance is short circuited) of magnitude

$$\overline{i^2} = 4kTGB. \tag{3}$$

We see from (1) that if two resistances are connected in series, the total *squared* noise voltage is the sum of the *squares* of the noise voltages produced by the resistances separately, and from (3) we see that the noise currents of conductances connected in shunt also add by summing squares. This rule of addition holds for adding the noise of all independent sources. Of course, if noise from the same noise source reaches a point by different paths, the

voltage or current components near any frequency should be added directly with due regard for phase.

Johnson noise is related to many physically similar phenomena such as Brownian motion and the random fluctuations in position observed in the coils of very sensitive galvanometers.

The simplest derivation of (1), (2) or (3) is that given by Nyquist[4] in a companion paper to Johnson's. Consider a long lossless transmission line of length L terminated at each end in resistances equal to its characteristic impedance. Imagine line and terminations in thermal equilibrium at a temperature T, as shown in Fig. 2. If electrical energy flows from the resistance at 1 to that at 2, then equal energy must then flow from 2 to 1, as any net gain or loss of energy would violate the second law of thermodynamics.

Now, suppose that we suddenly close the switches at 1 and 2, short circuiting the ends of the line. The line now becomes a resonator, having resonant

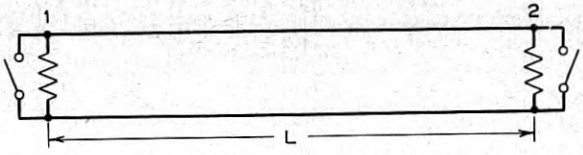


Fig. 2—Two resistances terminating a transmission line act as generators of thermal noise power traveling along the line.

frequencies such that the line is n half wavelengths long. The resonant frequencies will be

$$f = n(c/2L). \tag{4}$$

Here n is an integer and c is the velocity of light. The frequencies are separated by frequency intervals

$$\Delta f = (c/2L). \tag{5}$$

The energy which originally flowed right to left and left to right between the resistances is now reflected at the ends. It may be expressed as the thermal energy associated with the resonant modes of the line. According to statistical mechanics, there is an energy kT associated with each resonant mode. The energy per unit bandwidth is obtained by dividing this by the frequency interval between modes, given by (5) and is

$$w = kT/\Delta f = kT/(c/2L). \tag{6}$$

Since it takes a wave a time L/c to pass completely through the line, this energy w represents the energy per unit bandwidth which flowed into the

line from both resistances over a period L/c. If p is the power per unit bandwidth from one resistance, then

$$\begin{aligned} 2p(L/c) &= w = kT/(c/2L) \\ p &= kT. \end{aligned} \tag{7}$$

Or, we may say that the power flow from a resistance into a matched load (the available power) is, for a bandwidth B

$$P = kTB. \tag{8}$$

Sometimes it may be desired to know the mean squared fluctuation voltage integrated over all frequencies. Carrying out such an integration for the voltage between a pair of terminals connected by a complicated network would seem to be a difficult procedure. However, if the pair of terminals is shunted by a capacitance, the integrated fluctuation voltage can be obtained by direct application of the principles of statistical mechanics.

In a lumped network composed of capacitive, inductive and resistive elements* each capacitance and each inductance constitutes a degree of freedom; that is, the electrical state of the network can be specified completely by specifying the voltage across each capacitance and the current in each inductance**. According to statistical mechanics, the average stored energy per degree of freedom is $kT/2$. The stored energy in a capacitance is $Cv^2/2$. Thus, the mean squared noise voltage of all frequencies across a capacitance C must be

$$\overline{v^2} = kT/C. \tag{9}$$

Similarly, the mean squared noise current of all frequencies flowing in an inductance L is

$$\overline{i^2} = kT/L. \tag{10}$$

We have conveniently thought of Johnson noise as generated in the resistances in a network. We need not change this concept and say that the voltage and current of (9) and (10) are generated in the capacitance or inductance any more than we would say that the thermal velocities of molecules are generated by the molecules' mass. Relations (9) and (10) merely represent necessary consequences of the laws of statistical mechanics as, indeed, does (1).

It is of some interest to illustrate the use of (9) and its connection with (1)

* Strictly, such a lumped network is an unrealizable ideal. There are no pure capacitances, inductances, or resistances. The conditions under which actual condensers, coils and resistors can be represented satisfactorily by these idealizations must be judged by measurement or calculation or by past experience or intuition.

** In enumerating the degrees of freedom, capacitances in series or shunt are lumped together as one element; the same holds true for inductances.

by a very simple example. Consider a resonant circuit consisting of a capacitance C, an inductance L and a resistance R_0, all in parallel. The resistive component of the impedance across this circuit is

$$R = \frac{R_0}{1 + Q^2\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)^2} \tag{11}$$

$$Q = R_0\omega_0 C = R_0/\omega_0 L \tag{12}$$

$$\omega_0 = 1/\sqrt{LC}. \tag{13}$$

Here ω_0 is the resonant frequency of the circuit and Q has its usual meaning.

From (1) we see that as R_0, the resistance at resonance ($\omega = \omega_0$) is made higher, the noise voltage for frequencies near resonance increases. However, if we regard ω_0 and C in (12) as fixed, we see that as R_0 is increased the Q of the circuit is increased, the frequency range over which R is high is decreased, and R actually becomes lower far from resonance. (9) tells us that the mean square noise voltage integrated over all frequencies remains constant as R_0 is changed.

It is found that for a high Q circuit, the noise is much like a carrier of frequency ω_0 modulated by low-frequency noise. If we let the radian frequency of this "noise modulation" be $(\omega - \omega_0)$, then the mean square amplitude of the noise modulation varies with frequency about as R given by (11) varies with $(\omega - \omega_0)$.

II. Schottky Noise or Shot Noise

In 1918 Schottky[5] described the "Schrot-Effekt": the noise in vacuum tubes due to the corpuscular nature of the electron convection current. This is commonly known as "shot noise." The magnitude of this noise is usually derived by means quite different from those used here.

Johnson noise is necessarily associated with any electrical resistance, whatever its nature. Now, consider a close spaced planar diode shown in Fig. 3 consisting of two opposed emitting cathodes, each emitting a current I_0. Suppose the whole diode is held at the same temperature. There are no batteries or other sources of power aside from thermal energy; the only electrical energy flow must then be Johnson noise, ascribable to the resistance of the diode.

Assume that the cathodes both have the same uniform work function. Then when the diode is short circuited, each electron emitted from cathode 1 will reach cathode 2, and each electron emitted from cathode 2 will reach cathode 1.* If cathode 2 were made negative, all the electrons from 2 would

* It is here assumed that I_0 is small enough so that depression of potential due to space charge is avoided.

continue to reach 1, but some of the low-velocity electrons leaving 1 would be turned back from 2.

It is well known[6] that if a Maxwellian velocity distribution is assumed for the electrons leaving 1, the electrons which can overcome the retarding field and reach 2 are found to constitute a current

$$I = I_0 \epsilon^{eV/kT}. \tag{14}$$

Here I_0 is the total current carried by electrons leaving 1 and V is the voltage of 2 with respect to 1, which has been assumed to be negative.

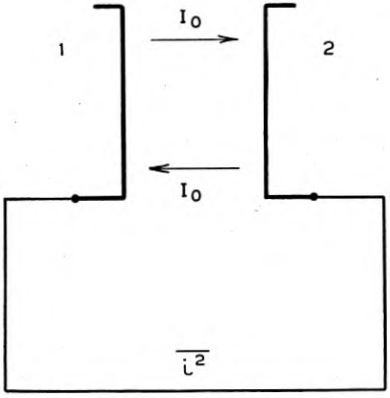


Fig. 3—An electronic resistance formed by two opposed cathodes at the same temperature acts as a generator of thermal noise.

By differentiating (14) we can obtain the diode conductance G at $V = 0$, and we find

$$G = \frac{e}{kT} I_0. \tag{15}$$

From (3) when the diode is short circuited and the voltage is zero we have a mean square noise current

$$\overline{i^2} = 4kTGB = \frac{eI_0}{kT}(4kTB) = 4eI_0B. \tag{16}$$

This noise is the sum of the noise due to two independent noise sources (the noise in the two currents I_0). That due to either current I_0 is*

$$\overline{i^2} = 2eI_0B. \tag{17}$$

* In this section, we are concerned with short transit angles only and no distinction need be made between the current induced in the circuit, i, and the electron convection current.

This is the expression for shot noise in a randomly emitted current, as in temperature limited emission or in photo electric emission.

III. Noise Other Than Shot Noise: Electron Multipliers

Let us consider a class of systems in which the average output current is proportional to the average input current, in which an electron of charge, e entering produces an output charge, ne instantaneously, and in which the probability that any electron will produce n electrons is p_n.

If the input current is I_0, the average output current is

$$I = \bar{n} I_0 \tag{18}$$

$$\bar{n} = \sum_n n p_n . \tag{19}$$

It is easy to persuade ourselves that any frequency component of current, noise or signal, will produce an output current $\bar{n}$ times as great; this happens to be true, and we will use the fact.

Let us consider our device when it has randomly emitted electrons as an input. At the output we will see appear groups of 1, 2, 3 etc. electrons, each group caused by the entrance of a single electron. If I_0 is the total input current, the output current consisting of groups of n electrons is

$$I_n = n I_0 p_n . \tag{20}$$

Each group carries a charge ne. We may now use (17) to write the noise in the part of the current carried by groups of n electrons, replacing the electronic charge, e, by the group charge, ne

$$\overline{i_n^2} = 2(ne)(nI_0 p_n)B. \tag{21}$$

As there is no correlation between entering electrons, the total mean square output noise current will be the sum of the noise components carried by groups consisting of different numbers n of electrons. Summing (19) with respect to n we obtain

$$\overline{i_t^2} = 2eI_0 B\overline{n^2} \tag{22}$$

$$\overline{n^2} = \sum_n n^2 p_n . \tag{23}$$

Now, the input has been taken as having shot noise. A part of the noise output is to be attributed to this input shot noise amplified by the device; that is, it will be $\bar{n}^2$ times the input shot noise.

$$\overline{i_s^2} = \bar{n}^2 2eI_0 B. \tag{24}$$

The part of the noise output current due to the fact that an electron does

not produce $\bar{n}$ electrons, but may produce 0, 1, 2$\cdots$ etc. electrons, must be the difference between (22) and (24), or

$$\overline{i_1^2} = 2eI_0B(\overline{n^2} - \bar{n}^2). \tag{25}$$

The quantity in parentheses is the mean square deviation in n.*

If the input current has any noise components $\overline{i_n^2}$, then the total noise output component will be

$$\overline{i^2} = \overline{i_1^2} + \bar{n}^2\overline{i_n^2}. \tag{26}$$

By applying (26) successively to stage after stage the noise output of a multistage electron multiplier can be evaluated (if one knows $(\overline{n^2} - \bar{n}^2)$).[7]

Wonder is sometimes expressed that current can be noisier than shot noise, in which the time of electron arrival is purely random. Obviously, we can have more than shot noise only if there is something non-random about the time of electron arrival, and the argument above discloses just what this is; it is the arrival of electrons in bunches.** We can easily see how erratic even large currents would be if electrons were bound together in groups having a total group charge of a coulomb, all the electrons in a group arriving simultaneously. Reverting to our shot noise formulas, we may illustrate this by assuming a perfect multiplier with a shot noise input, in which each input electron produces exactly N output electrons. Arguing from the shot noise equation (17) and replacing e by Ne we should expect an output noise current

$$\overline{i^2} = 2(Ne)I_1B \tag{27}$$

where I_1 is the output current; we get exactly the same result by assuming the input noise current squared amplified by N^2

$$\begin{aligned}\overline{i^2} &= (2eI_0B)N^2 \\ &= 2(Ne)\,(NI_0)B \\ &= 2(Ne)\,I_1B\,.\end{aligned} \tag{28}$$

* The mean square deviation is the sum with respect to n of the square of the deviation from the mean value of n, $\bar{n}$.

$$\Sigma\,(n - \bar{n})^2 p_n = \Sigma\, n^2 p_n - 2\bar{n}\,\Sigma\, np_n + \bar{n}^2\,\Sigma\, p_n.$$

The summation in the first term is $\overline{n^2}$, that in the second term is $\bar{n}$ and that in the third term is unity. Hence

$$\Sigma\,(n - \bar{n})^2 p_n = (\overline{n^2} - \bar{n}^2).$$

** Anything, (such as transit time difference for electrons within a bunch) which tends to break up the bunches will reduce the noise—and the signal as well. Such noise reduction involves a return to a more nearly random flow.

Conversely, we are led to wonder whether a current less bunched than that produced by random emission might not have less noise. The most smoothly distributed current we can imagine is that of f_0 electrons per second emitted at evenly spaced intervals. Obviously, such a current will have a spectrum consisting of frequencies nf_0, integral multiples of f_0. Thus for $f < f_0$, there will be no "noise" and similarly for $f_0 < f < 2f_0$, $2f_0 < f < 3f_0$, etc.

For a current of 10 ma, $f_0 = 6.3 \times 10^{16}$; thus, even for small currents an evenly spaced emission would have no a-c components in the radio-frequency range; this is a comforting thought in considering space-charge reduction of noise, which is discussed in section 5. However, purely to satisfy our curiosity we may pursue the matter a little further. If we assume that each electron constitutes an instantaneous pulse of current, a simple harmonic analysis shows that the a-c current component of frequency $n f_0$ will have a mean square value

$$\overline{i_n^2} = 2eI_0 f_0. \tag{29}$$

Thus, in each interval f_0 wide centered about a frequency $n f_0$ there will be a mean squared a-c current equal to that which would be associated with the same band for random emission with the same current. By making the emission regular we have not reduced the mean square "noise" current in a broad frequency range; we have merely changed its frequency distribution from a uniform distribution to a distribution of sharp, high peaks.

IV. Partition Noise

Consider a tetrode, shown in Fig. 4, with a cathode current I_c, a screen current I_s, and a plate current I_p.

The grid current is taken as zero. Suppose that the screen is very fine, so that every electron leaving the cathode has the same chance of striking the screen, regardless of its point of departure. We may now regard the function of the screen as that of a peculiarly simple electron multiplier, for which n can be zero (electron striking screen) or 1 (electron passing screen).

The probability of an electron passing the screen is I_p/I_c. Accordingly, from (19) and (23),

$$\bar{n} = I_p/I_c \tag{30}$$

$$\overline{n^2} = I_p/I_c. \tag{31}$$

Suppose we write the noise in the cathode current as

$$\overline{i^2} = \Gamma^2 2eI_c B \tag{32}$$

Here Γ^2, a factor less than unity, is introduced to account for the "space charge noise reduction" in space charge limited flow.

Now, by applying (25) and (26) we obtain for the noise in the plate current

$$\begin{aligned} \overline{i_p^2} &= 2eI_c B(I_p/I_c - (I_p/I_c)^2) + \Gamma^2 2eI_c B(I_p/I_c)^2 \\ \overline{i_p^2} &= 2eI_p B(1 - (1 - \Gamma^2)(I_p/I_c)). \end{aligned} \tag{33}$$

It is to be noted that if $\Gamma^2 = 1$, that is, if the cathode current is random, the noise in the plate current is purely shot noise. The screen cannot make the plate current noisier than shot noise since it does not act to produce bunches of electrons.

The noise in the screen current can be obtained by substituting I_s for I_p

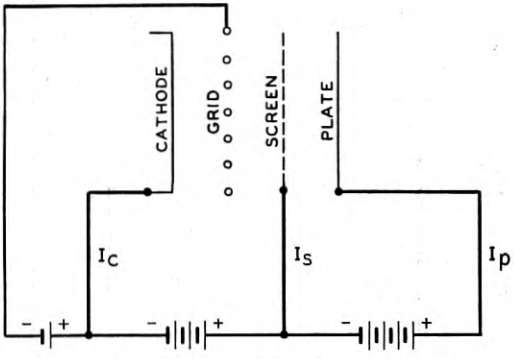


Fig. 4—Electrons randomly hitting or missing the screen grid make a tetrode noisier than a triode.

in (33). There is a correlation between the screen and plate noise currents; the total noise in the screen current plus the plate current must, of course, be

$$\overline{i_s + i_p^2} = \overline{i_c^2} = \Gamma^2 2eI_0 B \tag{34}$$

and not the sum of $\overline{i_p^2}$ and $\overline{i_s^2}$.

Partition noise has been discussed by Thompson, North and Harris.[8]

V. Space Charge Reduction of Noise

In this section an approximate derivation of noise in a space charge limited diode will be presented. The derivation leads to an expression valid for many practical tubes and illustrates the nature of the noise in space charge limited flow.

Consider a parallel plane diode of unit area and spacing x, with an applied voltage V_0, as shown in Fig. 5. When the voltage is applied, the electron convection current in the diode rises to value I_0. Neglecting thermal velocities of electron emission, this current is such that the electronic

"space charge" associated with it causes the voltage gradient at the cathode surface to be zero. A greater current would mean a negative gradient at the cathode and hence no emission; a smaller current would mean a positive gradient at the cathode and unlimited emission. On this basis Child's law is derived, which gives the current per unit area I_0 in amperes in terms of the voltage V_0 and the spacing in centimeters x as

$$I_0 = (2.33)\ 10^{-6} V_0^{3/2}/x^2 . \tag{35}$$

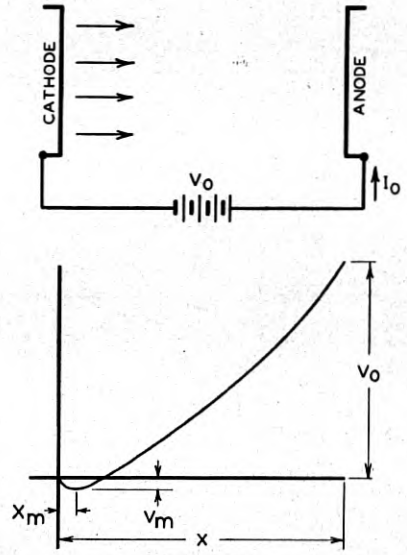


Fig. 5—Part of the electrons leaving the cathode of a diode are turned back before reaching the potential minimum; others proceed to the anode. Ordinarily the greater amount of noise is associated with the space between the potential minimum and the anode.

From (35) we can obtain a useful relation for the conductance G

$$G = \partial I_0/\partial V_0 = (3/2)(I_0/V_0). \tag{36}$$

The resistance R is

$$R = 1/G = (2/3)(V_0/I_0). \tag{37}$$

In actual diodes, the electrons are emitted from the cathode with a thermal velocity distribution; a potential minimum of some negative voltage V_m is formed at some distance x_m from the cathode surface. If the magnitude of the emitted electron current is I_e and the actual current passing the potential minimum is I_0, then because of the Maxwellian velocity distribution we have

$$\begin{aligned} I_0 &= I_e \epsilon^{eV_m/kT} \\ &= I_e \epsilon^{11,600 V_m/T}. \end{aligned} \tag{38}$$

Ordinarily, the magnitude of V_m is very small compared with V_0; I_0 is very small compared with I_e and x_m is very small compared with x.

Suppose V_m were held constant, say, by putting a conducting plane of potential V_m at x_m. Then, the electrons which pass this plane are quite independent of the low energy electrons which are turned back, and hence in the current passing x_m there will be pure shot noise.

$$\overline{i^2} = 2eI_0B. \tag{39}$$

Now suppose we change V_m. The change in I_0 will be, from (38),

$$dI_0 = dV_m/R_m \tag{40}$$

$$R_m = (e\bar{I}_0/kT)^{-1}. \tag{41}$$

If we use a constant current instead of a constant voltage d-c supply, then V_m must fluctuate in such a way as to cause a current equal and opposite to (39), or, there must be a fluctuating voltage $\overline{v_m^2}$ such that

$$\begin{aligned} \overline{v_m^2} &= 2eI_0BR_m^2 \\ &= (1/2)4kTR_mB. \end{aligned} \tag{42}$$

Suppose we consider the noise fluctuation of the anode voltage of a space charge limited diode supplied from a constant-current source. If there were no fluctuations in the voltage drop between the potential minimum at x_m and the anode at x, (42) would give the noise voltage fluctuation of such an "open circuited" diode. Actually, much larger fluctuation voltages are observed, and we must conclude that they arise in the space between the potential minimum and the anode. As the current is constant in this region (by definition—we have assumed a constant-current supply) we are forced to conclude that such fluctuations are due to a variation of mean electron speed in this region. The field at x_m is necessarily zero. If, with a constant current, electrons travel more rapidly between x_m and the anode, there is less electronic charge everywhere in this region, the rate of change of field with distance, and hence, the field, are everywhere smaller, and the voltage between x_m and the anode at x will be smaller.

It is somewhat involved to treat the problem of multi-velocity flow exactly; this has been done by Rack[9] and others[8,10,11]; however, Rack has shown that an approximate treatment yields very nearly the correct result over a fairly wide range of conditions. In this approximation, the stream of electrons with many velocities and a fluctuating mean velocity is replaced by a stream in which all electrons have the same velocity, and this has a mean square fluctuation equal to that of the multi-velocity stream.

Let us now measure x from the potential minimum. Suppose we consider an electron which passed the potential minimum ($x = 0$) at $t = 0$.

The field at the potential minimum is zero. The charge which has flowed in behind the electron at the time t is $-t I_0$. Hence, from Gauss's theorem the potential gradient is

$$\partial V/\partial x = I_0 t/\epsilon \tag{43}$$

where ϵ is the dielectric constant of vacuum. We have for the acceleration

$$\ddot{x} = \frac{e}{m}\frac{I_0 t}{\epsilon}. \tag{44}$$

If at the time $t = 0$ (at the potential minimum), $x = 0$, $\dot{x} = \dot{x}_0$

$$\dot{x} = \frac{e}{m}\frac{I_0}{2\epsilon} t^2 + \dot{x}_0 \tag{45}$$

$$x = \frac{e}{m}\frac{I_0}{6\epsilon} t^3 + \dot{x}_0 t. \tag{46}$$

Now the voltage V between the potential minimum and any point x must be such that

$$\dot{x}^2 - \dot{x}_0^2 = 2\frac{e}{m} V_0 \tag{47}$$

$$V_0 = \frac{1}{2\dfrac{e}{m}}\left(\frac{e}{m}\frac{I_0}{2\epsilon}\right)^2 t^4 + \frac{I_0}{2\epsilon} t^2 \dot{x}_0. \tag{48}$$

At any fixed point x, if we vary $\dot{x}_0$ by a small amount $d\dot{x}_0$, we find by differentiating (46)

$$\frac{dt}{d\dot{x}_0} = -\frac{t}{\left(\dfrac{e}{m}\dfrac{I_0}{2\epsilon} t^2 + \dot{x}_0\right)}. \tag{49}$$

From (48)

$$dV_0 = \frac{I_0 t}{\epsilon}\left(\frac{e}{m}\frac{I_0}{2\epsilon} t^2 + \dot{x}_0\right) dt + \frac{I_0}{2\epsilon} t^2\, d\dot{x}_0. \tag{50}$$

Using (49)

$$dV_0 = -\frac{I_0}{2\epsilon} t^2\, d\dot{x}_0. \tag{51}$$

It now remains to evaluate t. For most cases, the thermal velocities at the potential minimum are so small compared with the velocities in most of the region between the minimum and the anode that we can take the value of t for $\dot{x}_0 = 0$. Then, from (45) and (47)

$$t^2 = \left(\frac{e}{m}\frac{I_0}{2\epsilon}\right)^{-1}\left(2\frac{e}{m} V_0\right)^{1/2}. \tag{52}$$

From (51) and (52)

$$dV_0 = -2^{1/2}\left(\frac{e}{m}\right)^{-1/2} V_0^{1/2}\, d\dot{x}_0. \tag{53}$$

Now, if $\overline{(d\dot{x}_0)^2}$ is the mean square fluctuation in velocity, the mean square fluctuation in voltage will be

$$\overline{v^2} = 2(e/m)^{-1} V_0\, \overline{(d\dot{x}_0)^2}. \tag{54}$$

The assumptions leading to (54) are those leading to Child's law, and thus we can use (37) in connection with (54), giving

$$\overline{v^2} = 3(e/m)^{-1} I_0 R\, \overline{(d\dot{x}_0)^2}. \tag{55}$$

It now remains to evaluate $\overline{(d\dot{x}_0)^2}$, the mean square fluctuation in the velocity of the electrons passing the potential minimum; to do this, we return to (25). Suppose N is the number of input electrons per second. The output current can then be written

$$I_1 = \bar{n}Ne \tag{56}$$

and we can call the fluctuation in it

$$\overline{i^2} = \overline{(\delta\bar{n}Ne)^2}. \tag{57}$$

Equation (25) applies for no fluctuation in I_0 and hence for no fluctuation in N; e is a constant, and thus we may write (25) as

$$\overline{(\delta\bar{n})^2} = \frac{2B}{N}\,(\overline{n^2} - \bar{n}^2). \tag{58}$$

We may generalize this to say that each electron has a probability p of producing some effect of magnitude n and the fluctuation in the magnitude of the effect is $\overline{(\delta\bar{n})^2}$. Before, we said that an electron had a probability p of producing n secondaries. Now we will say instead that an electron has an uncorrelated probability p of having a velocity u, and obtain for the mean fluctuation in the velocity, $\overline{(d\dot{x}_0)^2}$

$$\overline{(d\dot{x}_0)^2} = \frac{2B}{N}\,(\overline{u^2} - \bar{u}^2). \tag{59}$$

In a Maxwellian distribution, the number of electrons passing a plane perpendicular to the direction of motion per second having velocities lying in the range du at u is

$$dn = Au\epsilon^{-(mu^2/2kT_c)}\, du. \tag{60}$$

Here T_c is cathode temperature. We see $\bar{u}$ and $\overline{u^2}$ are

$$\bar{u} = \frac{\int_0^\infty u^2 \epsilon^{-(mu^2/2kT_c)}\, du}{\int_0^\infty u\epsilon^{-(mu^2/2kT_c)}\, du} = \sqrt{\frac{\pi}{2}}\sqrt{\frac{kT_c}{m}} \tag{61}$$

$$\overline{u^2} = \frac{\int_0^\infty u^3 \epsilon^{-(mu^2/2kT_c)}\, du}{\int_0^\infty u\epsilon^{-(mu^2/2kT_c)}\, du} = 2\frac{kT_c}{m}. \tag{62}$$

Accordingly

$$\overline{u^2} - \bar{u}^2 = \tfrac{1}{2}(4 - \pi)\frac{kT_c}{m} \tag{63}$$

Combining (63) with (59) we obtain

$$\overline{(d\dot{x}_0)^2} = \frac{B}{N}(4 - \pi)\frac{kT_c}{m}. \tag{64}$$

Combining (62) with (53) and remembering that $I_0 = Ne$ we find the mean square open circuit noise voltage to be

$$\begin{aligned} \overline{v^2} &= 3(4 - \pi)\, kT_cRB \\ &= (.644)\, 4kT_cRB. \end{aligned} \tag{65}$$

This is the chief contribution to noise in a space charge limited diode.

Usually R is substantially equal to the plate resistance of the diode (it does not include effects on the cathode side of the potential minimum). Hereafter R will be treated as the total plate resistance of the diode.

VI. Noise in Triodes and Pentodes

Consider the triode shown in Fig. 6. Here we have a cathode, a grid, and a plate. The input admittance of the tube is represented in the diagram by the grid-cathode capacitance C_1 and the grid-plate capacitance C_2. The resistance R_n is a fictitious noise resistance which will be evaluated later. It is assumed to act between the input admittance of the tube and the controlling action of the grid; no current can flow in R_n because the grid as indicated in the diagram is presumed to present an open circuit.

We will regard the cathode-grid region of the triode as an "equivalent diode." The anode voltage of the diode is taken as

$$V_0 = (V_g + V_p/\mu). \tag{66}$$

Here V_g is the grid voltage and V_p the plate voltage of the triode. If the plate voltage is held constant and μ is taken as constant

$$dV_0 = dV_g. \tag{67}$$

Hence, under these conditions

$$\partial I_0/\partial V_0 = G = \partial I_0/\partial V_g. \tag{68}$$

Here G is the conductance of the equivalent diode, the reciprocal of R which appears in (65), and is also the transconductance of the triode.

As we wish to calculate the noise with no a-c grid or plate voltage, and as these through (64) specify the plate voltage V_0 of the equivalent diode,

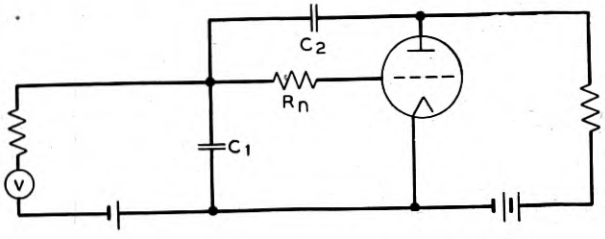


Fig. 6—Low-frequency noise in a triode can be ascribed to a fictitious noise resistance R_n, acting into an open circuit to cause voltage fluctuations on the grid.

the equivalent diode may be regarded as short-circuited. Hence, the noise current will be

$$\begin{aligned} \overline{i^2} &= \overline{v^2}/R^2 \\ &= (.644)\, 4kT_a GB. \end{aligned} \tag{69}$$

If we express this as shot noise reduced by a factor Γ^2 we obtain

$$\begin{aligned} \overline{i^2} &= 2eI_0 \Gamma^2 B \\ \Gamma^2 &= (.644) \frac{2kT_c G}{eI_0}. \end{aligned} \tag{70}$$

Often, the noise expressed by (69) is ascribed as a fictitious noise resistance R_n, at room Temperature T, connected between the grid-cathode capacitance and the "controlling action" of the grid as shown in Fig. 6. This fictitious resistance looks into a complete open circuit; hence, it has a noise voltage

$$\overline{v^2} = 4kTR_nB \tag{71}$$

and produces a noise plate current (for zero load resistance)

$$\overline{i^2} = 4kTR_nBG^2. \tag{72}$$

Comparing (69) with (72) we find

$$R_n = (.644/G)\,(T_c/T_0). \tag{73}$$

Here T_0 is a reference temperature, usually taken as 290° K. The effect of load impedance on signal from this fictitious resistance is treated by purely circuit means.

In pentodes there is noise according to (69) and in addition there is partition noise according to (33). By taking Γ^2 from (70) and equating the noise current given by (33) to (72) the fictitious "noise resistance" of a pentode can be evaluated in terms of g, I_p/I_c and T_c/T.

References

1. S. O. Rice, "Mathematical Analysis of Random Noise," *B. S. T. J.*, XXIII, pp. 282–332, July 1944, and XXIV, pp. 46–156, January, 1945.
2. D. A. Bell has given a valuable summary and discussion of noise and especially of noise in vacuum tubes. "Fluctuations of Electric Current," *Journal of the I. E. E.*, Vol. 93, pt. 3, pp. 37–44, January 1946.
3. J. B. Johnson, "Thermal Agitation of Electricity in Conductors," *Phys. Rev.*, Vol. 32, No. 1, pp. 97–109, July (1928).
4. H. Nyquist, "Thermal Agitation of Electric Charge in Conductors," *Phys. Rev.*, Vol. 32, No. 1, pp. 110–113, July (1928).
5. W. Schottky, "Spontaneous Current Fluctuations in Various Conductors," *Ann. Physik*, 57, pp. 541–567, (1918).
6. Irving Langmuir and Karl T. Compton, "Electrical Discharge in Gases," Part II, *Rev. Mod. Phys.*, Vol. 3, #2, p. 226 (April 1931).
7. W. Shockley and J. R. Pierce, "A Theory of Noise for Electron Multipliers," *Proc. I. R. E.* 26, pp. 321–332, March 1938.
8. B. J. Thompson, D. O. North and W. A. Harris, "Fluctuations in Space-Charge-Limited Currents at Moderately High Frequencies"
   Part I—*R. C. A. Review*, Jan. 1940
   Part II— " Apr. and July 1940
   Part III— " Oct. 1940
   Part IV— " Jan. 1941
   Part V— " Apr. and July 1941.
9. A. J. Rack, "Effect of Space Charge and Transit Time on the Shot Noise in Diodes," *B. S. T. J.*, Vol. 17, No. 4, p. 592, October (1938).
10. W. Schottky, and E. Spenke, "The Space-Charge Reduction of Shot Effect," *Wissenschaftliche Veroffentlichungen aus den Siemens-Werken*, Vol. 16, No. 2, pp. 1–41, July (1937).
11. E. Spenke, "The Space-Charge-Reduction of Shot Effect," *Wiss. Veroff. aus den Siemens-Werken*, Vol. 16, No. 2, p. 19, July (1937).

Abstracts of Technical Articles by Bell System Authors

Gross-linkage of Linear Polyesters by Free Radicals.[1] W. O. BAKER. Reactions fundamental to the use of the new low pressure laminating or casting resins have been studied. The striking property of these plastics, which are usually based on some polyester and a vinyl monomer, is their rapid and easy curing, leading to unique ease of fabrication. This curing, the formation of a permanent three-dimensional polymer network, or gel, is achieved by reaction with a source of free radicals, such as from an organic peroxide. These agents cause polymerization of the vinyl monomer, as was previously understood, but they also seem to incorporate the polyester into the network, even if the polyester contains little or no unsaturation.

Investigation of a series of simple polyesters, the polyundecanoates, showed that free radicals, such as come from the decomposition of benzoyl peroxide, could cross-link or gel the linear, saturated, chains. Apparently the hydrogen atoms in methylene groups next to polar groups like the carbonyl, i.e., the α-hydrogens, are removed by the free radicals. The resulting chain radical attacks an adjacent chain, and a cross-link is formed. The effects of cross-links thus produced on solubility, dilute solution viscosity, melt viscosity and, finally, stress relaxation of the cured solid were examined. Probably the similar activity of α-hydrogen atoms is important in the chemical aging or weathering of plastics and rubbers. It is likewise significant for the vulcanization of many synthetic rubbers.

Rubberlike Products from Linear Polyesters.[2] B. S. BIGGS, R. H. ERICKSON and C. S. FULLER. The polymers which result from the condensation of dibasic acids with propylene glycol are viscous gums which can be vulcanized to rubberlike products. In the unpigmented condition these rubbers are quite weak, but when reinforced with suitable pigments their strength and elongation compare favorably with other synthetic rubbers. Because polyesters of known structure and molecular weight can be easily synthesized, these polymers are useful for the study of the relations between structure and properties in rubberlike materials in general. Factors affecting tensile strength, oil resistance, brittle temperature, and stability are discussed.

Pulse Code Modulation.[3] H. S. BLACK and J. O. EDSON. A radically new modulation technique for multichannel telephony has been developed which involves the conversion of speech waves into coded pulses. This new tech-

[1] *Jour. Amer. Chemical Soc.*, May 1947.
[2] *Indus. & Engg. Chem.*, September 1947.
[3] *Telephony*, August 30, 1947.

nique is called Pulse Code Modulation or simply PCM. An eight-channel system embodying these principles was developed and produced in portable form for field operation. Other work carried on simultaneously by W. M. Goodall (see B. S. T. J., July 1947) resulted in the development of an experimental system using a different method of coding.

In carrying out this new type of modulation, the speech wave applied to each channel is, in effect, transmitted sample by sample, and each sample is represented by a multi-unit code employing on-or-off pulses, hence the term PCM.

This method appears to have exceptional possibilities from the standpoint of freedom from interference. Its full significance in connection with future radio and wire transmission may take some time to reveal.

Stereoscopic Drawings of Crystal Structures.[4] W. L. BOND. A method is presented for getting stereoscopic pairs of atomic structure views given the coordinates of the atoms and cell constants.

Properties of Liquids at High Sound Pressure.[5] H. B. BRIGGS, J. B. JOHNSON and W. P. MASON. When sound of high amplitude is transmitted into a liquid by means of a mechanical driving device, the ultimate limitation to the power that can be transferred is cavitation or breakdown of the liquid under high internal stresses. A study of cavitation has resulted in establishing the following results. Under steady-state conditions, light liquids filled with air cavitate when the negative acoustic pressure reaches the atmospheric pressure. When liquids are degassed, their natural cohesive pressure becomes effective and they will withstand a negative acoustic pressure. It is found that the total negative pressure required to cause cavitation is equal to the sum of the cohesive pressure—tensile strength—and the ambient pressure. Viscous liquids have a higher cohesive pressure and a proportionality has been established between the logarithm of the viscosity and the cohesive pressure. The amount of power that a liquid can withstand increases markedly as the pulse length is shortened.

An explanation of these phenomena is attempted on the basis of Eyring's theory of viscosity, plasticity and diffusion. On this theory natural holes exist in the liquid into which molecules can jump, leaving holes behind them. A jump occurs when the molecule has accumulated enough heat energy to surmount an activation potential barrier of energy value E_0. Cavitation appears to be the result of coalescing of the natural holes in the negative pressure phase of the cycle. Since a molecule has to jump from a hole in order that this can coalesce with another hole, the cavitation pressure is proportional to the activation energy which in turn is proportional to the log-

[4] *The American Mineralogist*, July–August 1947.
[5] *Jour. Acous. Soc. Amer.*, July 1947.

arithm of the viscosity. The increased power-transmitting capacity for short pulse lengths is a result of the finite time taken for the small holes to grow in size to a large enough hole to cause rupture of the liquid.

Modulation in Communication.[6] F. A. COWAN. The fundamentals involved in introducing signals into one medium and transmitting them through another are simplified in this review, so that the relationships between the many varieties of modulations attempted or in contemporary use are formed into a cohesive whole.

Air-borne Magnetometers.[7] E. P. FELCH,* W. J. MEANS,* T. SLONCZEWSKI,* L. G. PARRATT, L. H. RUMBAUGH and A. J. TICKNER.* Developed under the impetus of the submarine menace of World War II, the air-borne magnetometer has found many peacetime uses. Navy airplanes equipped with magnetometers for exploration of Antarctica were used in the recent United States Navy expedition. An expedition now is studying the Aleutian Alaskan volcanos and the Aleutian submarine trench. From there it will proceed to Hawaii and Bikini.

The Generation of Centimeter Waves.[8] H. D. HAGSTRUM. The electronic devices used most extensively, recently, for the generation of centimeter waves are discussed. The physical form, operating capabilities, and the basic physical principles of operation of the triode, velocity-variation, and magnetron oscillators are presented. An attempt is made to show how these oscillators are related to one another. For a variety of reasons, particular emphasis is placed on the magnetron oscillator.

Selective Demodulation.[9] DONALD B. HARRIS. A method of demodulation is proposed in which the output current of the demodulator is a linear function of the input voltage, while at the same time provision is made for producing the necessary product terms which will result in demodulation. Demodulation is brought about by integrating the product of the instantaneous value of the modulated wave by the instantaneous value of a wave having the same frequency and phase as the carrier. Where this method of demodulation is used it is proposed that two carriers in quadrature on the same frequency may be employed, reducing the bandwidth to that required for single-sideband transmission.

It is suggested that the required linear demodulation characteristics may be obtained through the use of "electron-coupled" demodulators. Theoretical considerations indicate that, when demodulation of this type is employed, selectivity ahead of the demodulator may be dispensed with, the

[6] *Elec. Engg.*, September 1947.
[7] *Elec. Engg.*, July 1947.
* *Of the Bell System.*
[8] *Proc. I. R. E.*, June 1947.
[9] *Proc. I. R. E.*, June 1947.

signal-to-noise ratio is improved, greater economy of spectrum space is obtained, the number of tubes required is materially reduced through the use of a common intermediate-frequency amplifier for a number of channels, and any impairment due to the instability of the carrier or oscillator frequency is reduced.

As an example of the possible application of the principles outlined, a hypothetical eight-channel transmission system is described.

The Physical Significance of Birkhoff's Gravitational Equations.[10] HERBERT E. IVES. Birkhoff's gravitational equations are put in terms of *dt* in place of the local time *ds* used by him. The transformed equations show that Lorentzian mass has been used, and to the Newtonian attractive force is added a force normal to the direction of motion, v^2/c^2 times the component of the gravitational force normal to the motion.

Attenuation and Scattering of High-Frequency Sound Waves in Metals and Glasses.[11] W. P. MASON and H. J. MCSKIMIN. By using a pulse method, attenuation and velocity measurements have been made for aluminum and glass rods in the frequency range from 2 to 15 megacycles. The sound pulses are generated by crystals waxed to the surface of the rod. This wax joint limits the band width of the transmitted pulse and measurements are made using long pulses which approach steady state conditions. The reflected pulses show evidence of several normal modes which can be minimized by using specially shaped electrodes. Longitudinal waves show delayed pulses of smaller magnitude that are caused by the longitudinal wave breaking up into reflected longitudinal and shear waves at the boundary. This effect is small if the diameter of the rod is 20 wave-lengths or more.

The measured losses for aluminum rods show a component proportional to the frequency and another component proportional to the fourth power of the frequency. The first component is the hysteresis loss found for most solid materials. The component proportional to the fourth power of the frequency is caused by Rayleigh scattering losses which are the result of differences in the elastic constants between adjacent grains caused by changes in orientation. Calculated scattering losses agree quite well with the measured values. The fourth-power scattering law holds quite well until the grain size is equal to one-third of a wave-length. For higher frequencies the scattering loss increases more nearly with the square of the frequency. Glasses and fused quartz have a loss directly proportional to the frequency, showing that any irregularities must be of very small size.

The Growth of Auditory Sensation.[12] W. A. MUNSON. The integration of sensation with respect to time was studied experimentally by means of tones

[10] *Phys. Rev.*, August 1, 1947.
[11] *Jour. Acous. Soc. Amer.*, May 1947.
[12] *Jour. Acous. Soc. Amer.*, July 1947.

of short duration. Loudness tests were made on sounds persisting from 0.005 to 0.2 second and covering a wide range of levels. The observed increase in magnitude of a sensation as the duration time is increased is attributed to the integration characteristic of the central nervous system, and an equivalent electrical circuit is derived. The circuit analogy is then used in the computation of loudness as a function of the duration of the stimulus.

The Physics of Electronic Semiconductors.[13] G. L. PEARSON. The band theory of solids is capable of explaining such fundamental properties of electronic semiconductors as the dependency of specific resistance on impurity content, the negative temperature coefficient of resistance, the sign of the Hall and thermoelectric effects, and the direction of rectification. Measurements of the specific resistance and the Hall constant enable the calculation of density, mobility, and mean free path of the electric carriers as a function of temperature and impurity.

Automatic Frequency Control of Microwave Oscillators.[14] VINCENT C. RIDEOUT. A method for the automatic frequency control of any type of tunable microwave oscillator is described. In this method a servomechanism is used which includes a wave-guide discriminator circuit, a mercury-contact relay, a 60-cycle amplifier, and a small two-phase induction motor.

Tests made on a preliminary model of a circuit of this type used with a 4000-megacycle oscillator showed that a stability of one part in 50,000 was obtainable. The manner in which such a control system may be used in a microwave repeater is described.

Proposed Method of Rating Microphones and Loudspeakers for Systems Use.[15] FRANK F. ROMANOW and MELVILLE S. HAWLEY. Proposed, is a method of rating microphones and loudspeakers whereby the over-all performance of a sound system may be determined by adding together the microphone and loudspeaker ratings and the gain of the interconnecting network. This sum gives the performance quite accurately for most systems. However, in some combinations of elements correction terms must be added. The formulas for these correction terms are derived.

The proposed microphone and loudspeaker ratings have the additional usefulness of being in a form which permits the comparison of instruments of different impedances.

Sulfur Linkage in Vulcanized Rubber.[16] MILTON L. SELKER and A. R. KEMP. The reaction of 2-methyl-2-butene with sulfur at 141.6°C. was studied. Reaction time and concentration paralleled those common in

[13] *Elec. Engg.*, July 1947.
[14] *Proc. I. R. E.*, August 1947.
[15] *Proc. I. R. E., Waves and Electrons Section*, September 1947.
[16] *Indus. & Engg. Chem.*, July 1947.

rubber-sulfur vulcanization. The results offer further insight into the vulcanization problem. The products of the reaction are liquids of the polysulfide type R-S_x-R, where x varies from 2 to 6 and R is an alkyl or alkenyl group and two solids ($C_5H_6S_3$ and a higher homolog). The polysulfides appear to be somewhat richer in hydrogen than is expected from reaction of two C_5H_{10} molecules, whereas the solids are hydrogen-poor. The structure of an acid anhydride in the sulfur system showing thione-thiol tautomerism is proposed for $C_5H_6S_2$, which is therefore 2,5-dithione-3-methyltetrahydrothiophene. The color changes with reaction time, from yellow to red to black, parallel those of rubber-sulfur vulcanizates. As in rubber-sulfur vulcanization the sulfur reaction rate is directly proportional to time, although the absolute rate is twice that in the polymer system. Starting with equal mole quantities of olefin and sulfur, there is a considerable amount of unreacted olefin in the system when all of the sulfur has reacted. The shorter the reaction time, the higher the value of x in the polysulfide R-S_x-R and the larger the percentage of residues R that are saturated.

On Hearing in Water vs. Hearing in Air.[17] L. J. SIVIAN. The paper deals with the ability of a submerged listener to hear sounds generated in the air above him, compared with their audibility when his head projects above the water. In a theoretical discussion it is shown that at 1000 c.p.s. a loss of the order of 45–55 db might be expected in the in-water audibility relative to the in-air value. This involves a number of assumptions, e.g., that there is no appreciable noise created by the listener's propulsion, and that the effect of hydrostatic pressure unbalance on the eardrum is negligible. A few measurements made at 1000 c.p.s. and 3000 c.p.s. yielded values which are not at variance with the theoretical analysis.

Cathode Phase Inverter Design.[18] C. W. VADERSEN. Part I of this paper covers the general analysis of the cathode coupled phase inverter and develops formulae that enable the designer to compute circuit elements with good accuracy. The theory developed shows that degeneration exists only in the driven side of the amplifier and is limited to 6 decibels. Balance is discussed in terms of the tube parameters and external resistances, its being shown that considerable stability is attainable. A form of the inverter in which the power output stage utilizes a transformer with an unbalanced plate winding is presented. This is shown to give a true power balance in a manner analogous to the unbalanced plate resistor form of the voltage amplifying inverter.

Part II presents a graph of the general design equations and illustrates its use with several working examples.

[17] *Jour. Acous. Soc. Amer.*, May 1947.
[18] *Audio Engineering*, June and July 1947.

Contributors to this Issue

WINSTON E. KOCK, B.E., University of Cincinnati, 1932; M.S., 1933; Ph.D., University of Berlin, 1934. Institute for Advanced Study, Princeton, New Jersey, 1935–36. Director of Electronic Research, Baldwin Piano Company, Cincinnati, Ohio, 1936–42. Bell Telephone Laboratories, Research Department, 1942–. Dr. Kock was engaged in radar antenna work in the Radio Research Department during the war. He is now engaged in microwave and acoustic research.

W. D. LEWIS, A.B. in Communication Engineering, Harvard College, 1935; Rhodes Scholar, Wadham College, Oxford; B.A. in Mathematics, Oxford, 1938; Ph.D. in Physics, Harvard, 1941. Bell Telephone Laboratories, Inc., 1941–. Dr. Lewis was engaged in radar antenna work in the Radio Research Department during the war; he is now engaged in microwave repeater systems research.

L. A. MEACHAM, B.S. in Electrical Engineering, University of Washington, 1929; Certificate of Research, Cambridge University, England, 1930. Bell Telephone Laboratories, 1930–. From 1930 to 1941 Mr. Meacham's work dealt with crystal oscillators, multivibrators, phase shifters, and other devices used in precision standards of frequency. During the war he developed range measuring devices for radar, and has since been concerned with applications of pulse techniques to multiplex telephony.

E. PETERSON, Cornell University, 1911–14; Brooklyn Polytechnic, E.E. 1917; Columbia University, A.M. 1923; Ph.D. 1926. Electrical Testing Laboratories, 1915–17; Signal Corps, U. S. Army, 1917–19. Western Electric Company, Engineering Department, 1919–25; Bell Telephone Laboratories, 1925–. Lecturer in Electrical Engineering, Columbia, 1934–. As circuit research engineer, Dr. Peterson's work has been largely in theoretical studies of non-linear circuits and circuit elements.

J. R. PIERCE, B.S. in Electrical Engineering, California Institute of Technology, 1933; Ph.D., 1936. Bell Telephone Laboratories, 1936–. Engaged in study of vacuum tubes.

S. O. RICE, B.S. in Electrical Engineering, Oregon State College, 1929; California Institute of Technology, 1929–30, 1934–35. Bell Telephone

Laboratories, 1930–. Mr. Rice has been concerned with various theoretical investigations relating to telephone transmission theory.

V. C. RIDEOUT, B.Sc. in Engineering Physics, University of Alberta, 1938; M.S. in Electrical Engineering, California Institute of Technology, 1940. Bell Telephone Laboratories, Inc., 1939–1946. Department of Electrical Engineering, University of Wisconsin, 1946–. During the war Mr. Rideout worked in the Research Department on various components for radar systems; after the war he worked on frequency control systems and on intermediate frequency power amplifiers for microwave repeater systems.

R. W. SEARS, A.B. Ohio Wesleyan University, 1928; M.S., Ohio State University, 1929. Columbia University, 1930–1935. Bell Telephone Laboratories, 1929–. Mr. Sears has been engaged in research work on thermionics and semiconductors. Since 1939 he has been primarily concerned with the development of electron tubes.

L. C. TILLOTSON, B.S. in E.E., University of Idaho, 1938; M.S. in E.E., University of Missouri, 1941. Instructor in Electrical Engineering, University of Missouri, 1940–41. Bell Telephone Laboratories, Inc., 1941–. During the war Mr. Tillotson was engaged in the design and development of wave filters and other transmission networks. In 1946 he was transferred to the Radio Research Department and since that time has been concerned with microwave repeater systems research.